p-ISSN 1607-3274
e-ISSN 2313-688X

Радіоелектроніка Інформатика Управління

Радиоэлектроника Информатика Управление

Radio Electronics Computer Science Control

2015/2

ISSN 1607-3274
9 771607 327005 53>

Запорізький національний технічний університет

Радіоелектроніка, інформатика, управління

Науковий журнал

Виходить чотири рази на рік
№ 2(33) 2015
Заснований у січні 1999 року.
Засновник і видавець – Запорізький національний технічний університет.
ISSN 1607-3274 (друкований), ISSN 2313-688X (електронний).
Запоріжжя
ЗНТУ
2015

Запорожский национальный технический университет

Радиоэлектроника, информатика, управление

Научный журнал

Выходит четыре раза в год
№ 2(33) 2015
Основан в январе 1999 года.
Основатель и издатель – Запорожский национальный технический университет.
ISSN 1607-3274 (печатный), ISSN 2313-688X (электронный).
Запорожье
ЗНТУ
2015

Zaporizhzhya National Technical University

Radio Electronics, Computer Science, Control

The scientific journal

Published four times per year
№ 2(33) 2015
Founded in January 1999.
Founder and publisher – Zaporizhzhya National Technical University.
ISSN 1607-3274 (print), ISSN 2313-688X (on-line).
Zaporizhzhya
ZNTU
2015

Науковий журнал «Радіоелектроніка, інформатика, управління» (скорочена назва – РІУ) видається Запорізьким національним технічним університетом (ЗНТУ) з 1999 р. періодичністю чотири номери на рік.

Зареєстрований Державним комітетом інформаційної політики, телебачення та радіомовлення 29.01.2003 р. Свідоцтво про державну реєстрацію друкованого засобу масової інформації серія КВ №6904.

ISSN 1607-3274 (друкований), **ISSN** 2313-688X (електронний).

Постановою президії ВАК України № 1-05/4 від 26.05.2010 р. журнал **включений до переліку наукових фахових видань України**, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата технічних наук та фізико-математичних наук (радіофізика).

В журналі безкоштовно публікуються наукові статті англійською, російською та українською мовами.

Правила оформлення статей подано на сайті: http://ric.zntu.edu.ua/information/authors.

Журнал забезпечує **безкоштовний відкритий он-лайн доступ** до повнотекстових публікацій.

Журнал дозволяє авторам мати авторські права і зберігати права на видання без обмежень. Журнал дозволяє користувачам читати, завантажувати, копіювати, поширювати, друкувати, шукати або посилатися на повні тексти своїх статей. Журнал дозволяє повторне використання його вмісту у відповідності з CC ліцензією CC-BY.

Опублікованими статтям присвоюється унікальний ідентифікатор цифрового об'єкта DOI.

Журнал реферується та індексується у провідних міжнародних та національних реферативних журналах і наукометричних базах даних, а також розміщується у цифрових архівах та бібліотеках з безкоштовним доступом у режимі on-line (у т. ч. DOAJ, DOI, CrossRef, EBSCO, eLibrary.ru / РИНЦ, Google Scholar, Index Copernicus, INSPEC, ISSN, Ulrich's Periodicals Directory, WorldCat, ВІНІТІ, Джерело), повний перелік яких подано на сайті: http://ric.zntu.edu.ua/about/editorialPolicies#custom-0.

Журнал розповсюджується за Каталогом періодичних видань України (передплатний індекс – 22914).

Тематика журналу містить: радіофізику, мікро-, нано- і радіоелектроніку, апаратне і програмне забезпечення комп'ютерної техніки, комп'ютерні мережі і телекомунікації, теорію алгоритмів і програмування, оптимізацію і дослідження операцій, міжмашинну і людино-машинну взаємодію, математичне і комп'ютерне моделювання, обробку даних і сигналів, управління в технічних системах, штучний інтелект, включаючи системи, засновані на знаннях, і експертні системи, інтелектуальний аналіз даних, розпізнавання образів, штучні нейронні і нейро-нечіткі мережі, нечітку логіку, колективний інтелект і мультиагентні системи, гібридні системи.

Усі статті, пропоновані до публікації, одержують **об'єктивний розгляд**, що оцінюється за суттю без урахування раси, статі, віросповідання, етнічного походження, громадянства або політичної філософії автора(ів).

Усі статті проходять двоступінчасте закрите (анонімне для автора) **резензування** штатними редакторами і незалежними рецензентами – провідними вченими за профілем журналу.

Рекомендовано до видання вченою радою ЗНТУ, протокол № 12 від 03.06.2015 р.

Журнал зверстаний редакційно-видавничим відділом ЗНТУ.

Веб-сайт журналу: http://ric.zntu.edu.ua.

Адреса редакції: Редакція журналу «РІУ», Запорізький національний технічний університет, вул. Жуковського, 64, м. Запоріжжя, 69063, Україна.

Тел: (061) 769-82-96 – редакційно-видавничий відділ

Факс: (061) 764-46-62

E-mail: rvv@zntu.edu.ua

Научный журнал «Радиоэлектроника, информатика, управление» (сокращенное название – РИУ) издается Запорожским национальным техническим университетом (ЗНТУ) с 1999 г. периодичностью четыре номера в год.

Зарегистрирован Государственным комитетом информационной политики, телевидения и радиовещания 29.01.2003 г. (Свидетельство о государственной регистрации печатного средства массовой информации серия КВ №6904).

ISSN 1607-3274 (печатный), **ISSN** 2313-688X (электронный).

Постановлением ВАК Украины №1-05/4 от 26.05.2010 г. журнал **внесен в перечень научных специализированных изданий Украины**, в которых могут публиковаться результаты диссертационных работ на получение научных степеней доктора и кандидата технических наук и физико-математических наук (радиофизика).

В журнале бесплатно публикуются научные статьи на английском, русском и украинском языках.

Правила оформления статей представлены на сайте: http://ric.zntu.edu.ua/information/authors.

Журнал обеспечивает **бесплатный открытый он-лайн доступ** к полнотекстовым публикациям. Журнал разрешает авторам иметь авторские права и сохранять права на издание без ограничений. Журнал разрешает пользователям читать, загружать, копировать, распространять, печатать, искать или ссылаться на полные тексты своих статей. Журнал разрешает повторное использование его содержания в соответствии с CC лицензией CC-BY.

Опубликованным статьям присваивается уникальный идентификатор цифрового объекта DOI.

Журнал реферируется и индексируется в ведущих международных и национальных реферативных журналах и наукометрических базах данных, а также размещается в цифровых архивах и библиотеках с бесплатным доступом on-line (в т .ч. DOAJ, DOI, CrossRef, EBSCO, eLibrary.ru / РИНЦ, Google Scholar, Index Copernicus, INSPEC, ISSN, Ulrich's Periodicals Directory, WorldCat, ВИНИТИ, Джэрэло), полный перечень которых представлен на сайте: http://ric.zntu.edu.ua/about/editorialPolicies#custom-0.

Журнал распространяется по Каталогу периодических изданий Украины (подписной индекс – 22914).

Тематика журнала включает: радиофизику, микро-, нано- и радиоэлектронику, аппаратное и программное обеспечение компьютерной техники, компьютерные сети и телекоммуникации, теорию алгоритмов и программирования, оптимизацию и исследование операций, межмашинное и человеко-машинное взаимодействие, математическое и компьютерное моделирование, обработку данных и сигналов, управление в технических системах, искусственный интеллект, включая системы, основанные на знаниях, и экспертные системы, интеллектуальный анализ данных, распознавание образов, искусственные нейронные и нейро-нечеткие сети, нечеткую логику, коллективный интеллект и мультиагентные системы, гибридные системы.

Все статьи, предлагаемые к публикации, получают **объективное рассмотрение**, которое оценивается по существу без учета расы, пола, вероисповедания, этнического происхождения, гражданства или политической философии автора(ов).

Все статьи проходят двухступенчатое закрытое (анонимное для автора) **рецензирование** штатными редакторами и независимыми рецензентами – ведущими учеными по профилю журнала.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Рекомендовано к изданию ученым советом ЗНТУ, протокол № 12 от 03.06.2015 г.

Журнал сверстан редакционно-издательским отделом ЗНТУ.

Веб-сайт журнала: http://ric.zntu.edu.ua.

Адрес редакции: Редакция журнала «РИУ», Запорожский национальный технический университет, ул. Жуковского, 64, г. Запорожье, 69063, Украина.

Тел.: +38-061-769-82-96 – редакционно-издательский отдел

Факс: (061) 764-46-62

E-mail: rvv@zntu.edu.ua

The scientific journal «Radio Electronics, Computer Science, Control» is published by the Zaporizhzhya National Technical University (ZNTU). since 1999 with periodicity four numbers per year.

The journal is registered by the State Committee for information policy, television and radio broadcasting of Ukraine in 29.01.2003. The journal has a State Registration Certificate of printed mass media (series КВ №6904).

ISSN 1607-3274 (print), **ISSN** 2313-688X (on-line).

By the Decree of the Presidium of All-Ukrainian Certification Commission № 1-05/4 dated 05.26.2010 the journal is **included to the list of the scientific specialized periodicals of Ukraine** where the results of dissertations for D.Sc. and Ph.D in Technical Sciences, Physics and Mathematics may be published.

The journal publishes scientific articles in English, Russian, and Ukrainian free of charge.

The **article formatting rules** are presented on the site: http://ric.zntu.edu.ua/information/authors.

The journal provides policy of **on-line open (free of charge) access** for full-text publications. The journal allow the authors to hold the copyright without restrictions and to retain publishing rights without restrictions. The journal allow readers to read, download, copy, distribute, print, search, or link to the full texts of its articles. The journal allow reuse and remixing of its content, in accordance with a CC license CC-BY.

Published articles have a unique digital object identifier (DOI).

The journal is abstracted and indexed in leading international and national abstractig journals and scientometric databases, and also placed to the digital archives and libraries with a free on-line access (including DOAJ, DOI, CrossRef, EBSCO, eLibrary.ru / РИНЦ, Google Scholar, Index Copernicus, INSPEC, ISSN, Ulrich's Periodicals Directory, WorldCat, VINITI (All-Russian Institute of scientific and technical information), Djerelo), full list of which is presented on the site: http://ric.zntu.edu.ua/about/editorialPolicies#custom-0.

The journal is distributed:by the Catalogue of Ukrainian periodicals (the catalog number is 22914).

The journal scope: radio physics, micro-, nano- and radio electronics, computer hardware and software, computer networks and telecommunications, algorithm and programming theory, optimization and operations research, machine-machine and man-machine interfacing, mathematical modeling and computer simulation, data and signal processing, control in technical systems, artificial intelligence, including knowledge-based and expert systems, data mining, pattern recognition, artificial neural and neuro-fuzzy networks, fuzzy logics, swarm intelligence and multiagent systems, hybrid systems.

All articles proposed for publication receive an **objective review** that evaluates substantially without regard to race, sex, religion, ethnic origin, nationality, or political philosophy of the author(s).

All articles undergo a two-stage **blind peer review** by the editorial staff and independent reviewers – the leading scientists on the profile of the journal.

EDITORIAL BOARD

Recommended for publication by the Academic Council of ZNTU, protocol № 12 dated 03.06.2015.

The journal is imposed by the editorial-publishing department of ZNTU.

The journal web-site is http://ric.zntu.edu.ua.

The address of the editorial office: Editorial office of the journal «Radio Electronics, Computer Science, Control», Zaporizhzhia National Technical University, Zhukovskiy street, 64, Zaporizhzhya, 69063, Ukraine.

Tel.: +38-061-769-82-96 – the editorial-publishing department.

Fax: +38-061-764-46-62

E-mail: rvv@zntu.edu.ua

ЗМІСТ

CONTENTS

p-ISSN 1607-3274. Радіоелектроніка, інформатика, управління. 2015. № 2
e-ISSN 2313-688X. Radio Electronics, Computer Science, Control. 2015. № 2

РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

РАДІОЕЛЕКТРОНИКА ТА ТЕЛЕКОММУНИКАЦИИ

RADIO ELECTRONICS AND TELECOMMUNICATIONS

УДК 621.396.96

Савочкин Д. А.[1], Гимпилевич Ю. Б.[2]

[1]*Аспирант кафедры радиотехники и телекоммуникаций Севастопольского национального технического университета, Севастополь, Украина*

[2]*Д-р техн. наук, профессор, заведующий кафедрой радиотехники и телекоммуникаций Севастопольского национального технического университета, Севастополь, Украина*

ОПТИМИЗАЦИЯ РАЗМЕЩЕНИЯ АНТЕНН ДЛЯ СИСТЕМ ПРОСТРАНСТВЕННОЙ ДВУМЕРНОЙ RFID-ЛОКАЛИЗАЦИИ

Разработана методика поиска оптимального размещения антенн для систем пространственной локализации на основе технологии радиочастотной идентификации (RFID). Методика основана на минимизации средней ошибки при реализации комбинированного метода локализации, включающего в себя алгоритм трилатерации, алгоритм отпечатков и алгоритм пересечений. В ходе проведенного вычислительного эксперимента установлено, что среди ряда рассмотренных паттернов оптимальным для квадратной зоны локализации является вариант размещения антенн в виде квадратной сетки при ориентации зон действия антенн по направлению к центру зоны локализации.

Ключевые слова: RFID, локализация, оптимизация, комбинированный метод, размещение антенн.

НОМЕНКЛАТУРА

HDOP – horizontal dilution of precision;

RFID – radio frequency identification;

a – вектор размещения антенн;

A – множество векторов a;

b – отношение длин большой и малой осей эллипсов зон действия антенн;

$c(a, l)$ – коэффициент покрытия точки l зонами действия антенн при размещении антенн согласно вектору a;

$C(a)$ – коэффициент покрытия зоны локализации зонами действия антенн;

$\overline{C(a)}$ – нормированный коэффициент покрытия зоны локализации зонами действия антенн;

$C_{\max}$ – граница нормирования коэффициента покрытия зоны локализации;

$c_1(a)$ – отношение площади той части зоны локализации, что покрыта зонами действия хотя бы одной антенны, к общей площади зоны локализации;

$c_1'(a)$ – отношение площади той части центральной области зоны локализации, что покрыта зонами действия хотя бы одной антенны, к общей площади центральной области зоны локализации;

$c_3(a)$ – отношение площади той части зоны локализации, что покрыта зонами действия трех или более антенн, к общей площади зоны локализации при размещении антенн согласно вектору a;

d – расстояние между RFID-меткой и антенной;

$e(a, l)$ – ошибка локализации для точки l, обеспечиваемая алгоритмом пересечений при размещении антенн согласно вектору a;

$E(a)$ – средняя ошибка локализации, обеспечиваемая алгоритмом пересечений;

$\overline{E(a)}$ – нормированная средняя ошибка локализации, обеспечиваемая алгоритмом пересечений;

$E_{\max}$ – граница нормирования средней ошибки алгоритма пересечений;

$h(a, l)$ – величина HDOP для точки l при размещении антенн согласно вектору a;

$H(a)$ – усредненная величина HDOP;

$\overline{H(a)}$ – нормированная усредненная величина HDOP;

$H_{\max}$ – граница нормирования величины HDOP;

$k(\varphi)$ – поправочный эллиптический коэффициент для азимута φ на точку размещения метки относительно точки размещения соответствующей антенны;

l – некоторая точка в пространстве зоны локализации;

$\hat{l}(a,l)$ – оценка местоположения метки, формируемая алгоритмом пересечений при ее расположении в точке l;

DOI 10.15588/1607-3274-2015-2-1

$\hat{l}(a,q,l)$ – оценка местоположения метки, формируемая алгоритмом пересечений при ее расположении в точке l в случае принятия ответов только антеннами множества q и размещения антенн согласно вектору a;

L – множество точек зоны локализации;

$|L|$ – площадь зоны локализации;

$L_3(a)$ – множество точек той части зоны локализации, которая покрыта зонами действия как минимум трех различных антенн;

$|L_3(a)|$ – площадь той части зоны локализации, которая покрыта зонами действия как минимум трех различных антенн при размещении антенн согласно вектору a;

N – общее число антенн RFID-системы;

$p(d)$ – вероятность получения антенной ответа от метки, расположенной на расстоянии d;

$p(d, \varphi)$ – вероятность получения антенной ответа от метки, расположенной на расстоянии d с азимутом φ;

$\overline{p}(d,\varphi)$ – вероятность неполучения антенной ответа от метки, расположенной на расстоянии d с азимутом φ;

q – некоторое множество номеров антенн;

q_{all} – множество номеров всех антенн системы;

Q – множество всех возможных множеств номеров антенн, принявших сигналы от метки;

$r(l_1, l_2)$ – функция расстояния между точками l_1 и l_2;

$R_{\max}$ – предельное расстояние между антенной и меткой, после которого вероятность получения ответа от метки антенной равна нулю;

$R_{\min}$ – предельное расстояние между антенной и меткой, до которого вероятность получения ответа от метки антенной равна единице;

$S(a)$ – целевая функция оптимизации размещения антенн при реализации комбинированного метода пространственной локализации;

$v(a, q, l)$ – вероятность получения ответов от метки расположенной в точке l только антеннами множества q и размещения антенн согласно вектору a;

$w(a, l)$ – вероятность размещения метки в точке l;

$w_{\text{отп}}(a)$ – весовой коэффициент алгоритма отпечатков;

$w_{\text{пер}}(a)$ – весовой коэффициент алгоритма пересечений;

$w_{\text{трил}}(a)$ – весовой коэффициент алгоритма трилатерации;

z – величина сдвига антенн относительно края зоны локализации;

$\alpha\ (a, l)$ – геометрическая матрица, используемая при вычислении величины HDOP;

$\beta\ (a, l)$ – ковариационная матрица, используемая при вычислении величины HDOP;

$\varphi\ (l, a_i)$ – азимут на точку l относительно точки размещения антенны a_i;

ψ – угол поворота эллипсов зон действия антенн.

ВВЕДЕНИЕ

Системы пространственной локализации объектов внутри закрытых помещений находят широкое применение в различных сферах человеческой деятельности.

Одной из перспективных технологий, позволяющих реализовать процедуру пространственной локализации, является технология радиочастотной идентификации RFID. В системах RFID-локализации используются специальные RFID-метки, размещаемые на объектах. Местоположение объектов определяется в ходе выполнения алгоритма локализации, который анализирует измерительную информацию, получаемую от меток с помощью нескольких антенн RFID-системы. Важной задачей при этом является повышение точности локализации объектов. Одним из факторов, оказывающих влияние на точность локализации, является размещение антенн в пространстве.

Целью настоящей работы является разработка методики оптимизации размещения антенн RFID-системы для комбинированного метода двумерной пространственной локализации, включающего в себя несколько различных алгоритмов локализации.

1 ПОСТАНОВКА ЗАДАЧИ

В работе мы рассматриваем случай двумерной локализации, когда интересующие нас объекты оснащены RFID-метками и находятся на полу помещения (или на небольшой высоте над уровнем пола), а антенны системы установлены на потолке. Положим, что система содержит определенное количество практически одинаковых антенн, которые должны быть оптимальным образом размещены (заданы их координаты и углы поворота). В качестве критерия оптимальности мы принимаем критерий минимума средней ошибки локализации при реализации комбинированного метода локализации.

Исходными данными для разрабатываемой нами методики являются размер зоны локализации (области в которой могут размещаться объекты локализации с RFID-метками), число антенн RFID-системы, их параметры, а также множество анализируемых векторов размещения антенн. При этом под параметрами антенн мы понимаем в первую очередь форму и размер зон действия антенн – областей, при расположении в которых RFID-меток, их сигналы успешно принимаются антеннами. В качестве выходных данных наша методика должна выдавать значения целевой функции для каждого анализируемого вектора размещения антенн, тем самым позволяя выбрать наилучший вариант.

2 ОБЗОР ЛИТЕРАТУРЫ

В литературе имеется ряд работ, посвященных теме оптимального размещения антенн в пространстве. При этом используется множество различных критериев оптимальности. Так, в [1] представлена методика поиска наилучшего размещения точек доступа в беспроводных сетях. Авторы добиваются максимизации покрытия рабочей области зонами действия антенн, минимизации интерференции между точками доступа, максимизации пропускной способности, а также минимизации числа точек доступа. В [2] помимо максимизации покрытия рабочей области и минимизации общего числа точек доступа авторы добиваются минимизации средних потерь в распространении сигналов. В [3] рассмотрен вопрос оптимального размещения антенн в классической RFID-системе (не обеспечивающей локализацию объектов). В данной работе авторы добиваются минимизации площади областей, в которых пересекаются зоны действия нескольких антенн, минимизации числа «беспо-

лезных» антенн (антенн, внутри областей действия которых не расположены RFID-метки), максимизации числа меток, находящихся внутри зон действия антенн, а также минимизации числа меток, находящихся внутри областей пересечения зон действия антенн.

Известны также работы, посвященные оптимизации размещения антенн при решении задачи локализации объектов в пространстве. В [4] проводится максимизация усредненного отношения величины разрешающей способности к степени изменчивости сигнала во времени. Недостатком данного критерия является необходимость наличия тренировочных (калибровочных) данных. В [5] выполняется минимизация размеров всех образуемых в рабочей области зон неопределенности. При этом под зоной неопределенности понимается такая область, которая образуется путем пересечения двух или более колец с центрами в точках размещения антенн и радиусами, соответствующими минимальным и максимальным дальностям действия антенн. Недостатком этого критерия можно считать его относительно высокую вычислительную сложность. В работе [6] предлагается для максимизации точности локализации добиваться того, чтобы каждая точка зоны локализации была покрыта зонами действия как минимум трех антенн, и при этом минимизировать общее число антенн системы. В [7] авторы стремятся минимизировать среднюю ошибку локализации, используя вероятностную модель распределения измерительной информации в виде уровней сигнала. Недостатком представленного подхода является требование наличия предварительно полученных данных, позволяющих выбрать параметры вероятностной модели.

3 МАТЕРИАЛЫ И МЕТОДЫ

В настоящей работе проводится разработка методики оптимизации размещения антенн RFID-системы для комбинированного метода пространственной локализации. Комбинированный метод основан на объединении нескольких базовых алгоритмов локализации, что позволяет увеличить результирующую точность [8]. В простейшем случае результирующей оценкой местоположения RFID-метки в ходе реализации комбинированного метода считается среднее арифметическое оценок, сформированных базовыми алгоритмами. В этом разделе приводится описание трех базовых алгоритмов: алгоритма трилатерации, алгоритма отпечатков и алгоритма пересечений. Эти алгоритмы входят в состав комбинированного метода локализации, рассматриваемого в дальнейшей работе.

В ходе реализации алгоритма трилатерации (дальномерного алгоритма) для локализуемой метки строятся окружности положения по результатам измерений [8]. Далее находится точка пересечения этих окружностей, которая принимается оценкой положения метки. Радиусы окружностей положения определяются значениями измерительной информации, полученной от метки с помощью антенн (например, уровнями мощности ответных сигналов от метки). Этот алгоритм иллюстрируется рис. 1, на котором изображен случай трилатерации метки с помощью трех антенн (a_1, a_2, a_3 – позиции антенн, d_1, d_2, d_3 – расстояния между антеннами и меткой, крест – результирующая оценка положения метки).

Алгоритм отпечатков требует проведения предварительного этапа сбора измерительной информации в различных точках зоны локализации от тренировочных RFID-меток. Собранная при этом измерительная информация используется как источник эталонных данных на рабочем этапе локализации. При реализации алгоритма для расчета оценки местоположения некоторой метки ее вектор измерительной информации (отпечаток) сравнивается с каждым вектором из сформированной на предварительном этапе базы отпечатков. В ходе сравнения в базе находятся векторы в максимальной степени подобные входному вектору. После этого результирующая оценка местоположения метки рассчитается как среднее тех позиций, для которых были ранее получены найденные в базе векторы. Наиболее известным частным случаем алгоритма отпечатков является алгоритм k ближайших соседей (k nearest neighbors) [8].

Алгоритм пересечений основан на использовании измерительной информации в виде фактов наличия ответов от меток. При его реализации оценка местоположения некоторой RFID-метки рассчитывается путем нахождения области пересечения зон действия антенн, принявших сигналы от метки, с последующим исключением области объединения зон действия антенн, не принявших сигналы от метки [8]. С целью иллюстрации принципа работы алгоритма на рис. 2 изображена схема размещения четырех антенн ($a_1 ... a_4$) и их зон действия (круги). При таком расположении антенн и размере зон действия формируется тринадцать областей возможного местоположения меток (зон позиционирования), центры масс которых могут быть приняты в виде результирующих оценок алгоритма.

Для создания методики оптимизации размещения антенн RFID-системы при использовании комбинированного метода пространственной локализации вначале необходимо сформировать факторы, характеризующие оп-

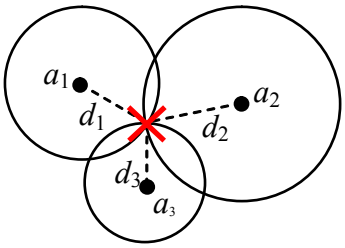


Рисунок 1 – Пример реализации алгоритма трилатерации

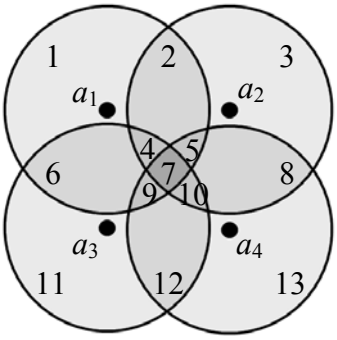


Рисунок 2 – Схема размещения четырех антенн и их зон действия

тимальность каждого из базовых алгоритмов, входящих в состав метода. В данном разделе мы поочередно проводим вывод таких факторов для алгоритмов трилатерации, отпечатков и пересечений, после чего формируем критерий оптимальности комбинированного метода.

В качестве фактора, характеризующего оптимальность размещения антенн при реализации алгоритма трилатерации, мы предлагаем использовать коэффициент горизонтального снижения точности (HDOP). Понятие HDOP широко применяется в теории систем спутниковой навигации для количественной оценки взаимного расположения спутников относительно приемного устройства. Усредненное значение HDOP $H(a)$ можно вычислить путем взятия интеграла от величины HDOP по каждой из точек интересующей нас части зоны локализации:

$$H(a) = \int_{l \in L_3(a)} h(a,l) w(a,l) \mathrm{d}l,$$

где $a = \{a_1, \ldots a_N\}$ – каждый элемент вектора содержит позицию и угол поворота антенны; $w(a, l)$ – вероятность размещения RFID-метки в точке l (позволяет выделить наиболее важные части зоны локализации [7]). В дальнейшей работе мы приняли плотность вероятности размещения меток равномерной: $w(a, l) = 1/|L_3(a)|$.

Мы проводим вычисление усредненного значения HDOP лишь по той части зоны локализации, которая покрыта зонами действия как минимум трех антенн, поскольку классический вариант алгоритма трилатерации для случая двумерной локализации требует получения измерительной информации от метки как минимум тремя различными антеннами. Мы полагаем, что в случае, когда ответы получены двумя или менее антеннами, результирующая оценка комбинированного метода формируется на основе оценок алгоритма отпечатков и алгоритма пересечений, а алгоритм трилатерации не используется.

Известно, что при вычислении значений HDOP для двумерного пространства первоначально формируется геометрическая матрица [9]

$$\alpha(a,l) = \begin{bmatrix} \dfrac{a_{1x} - l_x}{d_1} & \dfrac{a_{1y} - l_y}{d_1} \\ \dfrac{a_{2x} - l_x}{d_2} & \dfrac{a_{2y} - l_y}{d_2} \\ \dfrac{a_{3x} - l_x}{d_3} & \dfrac{a_{3y} - l_y}{d_3} \end{bmatrix}, \quad (2)$$

где $a_1 = (a_{1x}, a_{1y})$, $a_2 = (a_{2x}, a_{2y})$, $a_3 = (a_{3x}, a_{3y})$ – позиции трех ближайших к точке $l = (l_x, l_y)$ антенн; d_1, d_2, d_3 – расстояния от точки l до антенн a_1, a_2, a_3, соответственно.

После этого вычисляется ковариационная матрица

$$\beta(a,l) = \left[\alpha(a,l)^T \alpha(a,l)\right]^{-1}, \quad (3)$$

где $(\cdot)^T$ – операция транспонирования; $(\cdot)^{-1}$ – операция нахождения обратной матрицы.

Окончательно величина HDOP в точке l вычисляется как

$$h(a,l) = \sqrt{\beta(a,l)_{1,1} + \beta(a,l)_{2,2}}. \quad (4)$$

В качестве фактора, характеризующего оптимальность размещения антенн при реализации алгоритма отпечатков, будем использовать коэффициент $C(a)$ покрытия зоны локализации зонами действия антенн. Такой коэффициент может определяться по формуле

$$C(a) = \int_{l \in L} c(a,l) w(l) \mathrm{d}l, \quad (5)$$

где $c(a, l)$ – определяется как число зон действия, внутри которых находится точка l; $w(l) = 1/|L|$ – вероятность размещения RFID-метки в точке l.

В качестве фактора, характеризующего оптимальность размещения антенн при реализации алгоритма пересечений, примем среднюю ошибку $E(a)$ пространственной локализации, обеспечиваемую алгоритмом. Ошибку $E(a)$ можно записать в виде взвешенного интеграла ошибки локализации по каждой из точек зоны локализации:

$$E(a) = \int_{l \in L} e(a,l) w(l) \mathrm{d}l. \quad (6)$$

Ошибка $e(a, l)$ локализации для точки l определяется как расстояние между точкой l и оценкой $\hat{l}(a,l)$ местоположения RFID-метки, формируемой алгоритмом пересечений при расположении метки в точке l

$$e(a,l) = r\left[l, \hat{l}(a,l)\right] = \sqrt{\left[l_x - \hat{l}(a,l)_x\right]^2 + \left[l_y - \hat{l}(a,l)_y\right]^2}. \quad (7)$$

В идеализированном случае при реализации алгоритма пересечений оценка $\hat{l}(a,l)$ положения метки формируется в центре масс той единственной зоны позиционирования, которая включает в себя точку l. Однако из-за различного рода шумов могут возникать ситуации отсутствия ответа от метки при ее расположении внутри зоны действия антенны или наличия ответа при расположении вне зоны действия. Это приводит к возможному смещению оценки местоположения. Учесть подобный эффект можно, реализуя вероятностное вычисление ошибки $e(a, l)$. Для этого нами предлагается учитывать вероятность $p(d)$ получения антенной ответа от метки, расположенной на расстоянии d. Можно предположить, что до некоторого предела $R_{\min}$ такая вероятность близка к единице, а далее она монотонно убывает и достигает нуля на расстоянии $R_{\max}$:

$$p(d) = \begin{cases} 1, \text{ если } d < R_{\min}; \\ 1 - \dfrac{d - R_{\min}}{R_{\max} - R_{\min}}, \text{ если } R_{\min} \le d \le R_{\max}; \\ 0 \text{ если } d > R_{\max}. \end{cases} \quad (8)$$

Выражение (8) справедливо для антенн с изотропной в плоскости локализации диаграммой направленности, формирующих круговые зоны действия. Однако в RFID-системах часто применяются слабонаправленные антенны, при использовании которых форма зон действия принимает эллиптический характер. С целью учета это-

го мы предлагаем рассчитывать значения $R_{\min}$ и $R_{\max}$ с использованием поправочных коэффициентов:

$$\begin{cases} R_{\min}(\varphi) = R_{\min} k(\varphi); \\ R_{\max}(\varphi) = R_{\max} k(\varphi). \end{cases} \quad (9)$$

Эллиптический коэффициент $k(\varphi)$ может быть рассчитан по следующей формуле (выводится из общей формулы эллипса в полярных координатах):

$$k(\varphi) = \frac{b\sqrt{2}}{\sqrt{\left(1-b^2\right)\cos(2\varphi - 2\psi) + b^2 + 1}}. \quad (10)$$

Таким образом, выражение (8) может быть переписано для расчета вероятности $p(d, \varphi)$ получения антенной ответа от метки, расположенной на расстоянии d с азимутом φ:

$$p(d,\varphi) = \begin{cases} 1, \text{ если } d < R_{\min}(\varphi); \\ 1 - \dfrac{d - R_{\min}(\varphi)}{R_{\max}(\varphi) - R_{\min}(\varphi)}, \text{ если } R_{\min}(\varphi) \le d \le R_{\max}(\varphi); \\ 0 \text{ если } d > R_{\max}(\varphi). \end{cases} \quad (11)$$

При этом вероятность неполучения антенной ответа от метки будем определять как

$$\overline{p}(d,\varphi) = 1 - p(d,\varphi). \quad (12)$$

С учетом всего вышесказанного, ошибку $e(a, l)$ локализации для точки l следует определять как среднее взвешенное расстояний между точкой l и всеми возможными оценками местоположения метки при ее расположении в точке l. Весовыми коэффициентами усреднения при этом должны служить вероятности формирования конкретных оценок:

$$e(a,l) = \sum_{q \in Q} r\left[l, \hat{l}(a,q,l)\right] v(a,q,l). \quad (13)$$

Вероятность $\nu(a, q, l)$ будем определять по формуле

$$v(a,q,l) = \prod_{i \in q} p\left[r(l,a_i), \varphi(l,a_i)\right] \prod_{j \in q_{all} \setminus q} \overline{p}\left[r(l,a_j), \varphi(l,a_j)\right], \quad (14)$$

где a_i, a_j – позиции i-й и j-й антенн, соответственно $q_{all} = \{1, ..., N\}$.

Сформировав факторы, характеризующие оптимальность размещения антенн, для всех базовых алгоритмов локализации, мы можем вывести критерий оптимальности для комбинированного метода локализации. При этом целевая функция $S(a)$ оптимизации размещения антенн комбинированного метода должна вычисляться на основе факторов (1), (5) и (6). Критерием оптимальности будем считать максимум целевой функции

$$S(a) \to \max_{a \in A}. \quad (15)$$

Определять целевую функцию $S(a)$ будем в следующем виде:

$$S(a) = c_1'(a)^2 c_1(a)^3 \left[w_{\text{трил}}(a)\overline{H(a)} + w_{\text{отп}}(a)\overline{C(a)} + w_{\text{пер}}(a)\overline{E(a)}\right], \quad (16)$$

где в качестве центральной области понимался квадрат площадью 80% от площади всей зоны локализации, размещенный в ее центре.

В выражении (16) множители $c_1'(a)$ и $c_1(a)$ возводятся во вторую и третью степени, соответственно, поскольку мы полагаем, что важно иметь полное покрытие всей зоны локализации (и особенно важно – ее центральной части) зонами действия хотя бы одной из имеющихся антенн. Возведение величин $c_1'(a)$ и $c_1(a)$ в степень позволяет существенно снизить значение целевой функции для тех вариантов размещения антенн, при которых не вся зона локализации покрыта зонами действия антенн. Этим достигается исключение таких вариантов в ходе процедуры оптимизации, даже если они обеспечивают высокую точность локализации в других частях зоны локализации.

Весовые коэффициенты $w_{\text{трил}}(a)$, $w_{\text{отп}}(a)$, $w_{\text{пер}}(a)$ позволяют учесть степень точности алгоритмов трилатерации, отпечатков и пересечений, соответственно. В настоящей работе мы задали их следующим образом:

$$\begin{cases} w_{\text{трил}}(a) = \dfrac{1}{3} c_3(a); \\ w_{\text{отп}}(a) = w_{\text{пер}}(a) = \dfrac{1}{3}\left[1 + \dfrac{1 - c_3(a)}{2}\right]. \end{cases} \quad (17)$$

Такой вариант задания весовых коэффициентов используется по причине того, что предполагается, что алгоритм трилатерации функционирует только в той части зоны локализации, которая покрыта зонами действия хотя бы трех антенн.

Нормированная величина HDOP $\overline{H(a)}$ определяется как

$$\overline{H(a)} = \begin{cases} 1 - \dfrac{H(a) - 1}{H_{\max} - 1}, \text{ если } H(a) \le H_{\max}; \\ 0, \text{ если } H(a) > H_{\max}. \end{cases} \quad (18)$$

В настоящей работе мы приняли $H_{\max}$ = 2,5 (субъективная граница между «хорошими» и «средними» значениями).

Нормированный коэффициент покрытия зоны локализации зонами действия антенн $\overline{C(a)}$ определяется как

$$\overline{C(a)} = \begin{cases} 1 - \dfrac{C_{\max} - C(a)}{C_{\max}}, \text{ если } C(a) \le C_{\max}; \\ 1, \text{ если } C(a) > C_{\max}. \end{cases} \quad (19)$$

В настоящей работе мы эмпирически выбрали $C_{\max}$ = 4, поскольку ответов от четырех антенн обычно достаточно для высокоточной локализации при использовании метода отпечатков.

Нормированная средняя ошибка $\overline{E(a)}$ алгоритма пересечений определяется как

$$\overline{E(a)} = \begin{cases} 1 - \dfrac{E(a)}{E_{\max}}, \text{ если } E(a) \le E_{\max}; \\ 0, \text{ если } E(a) > E_{\max}. \end{cases} \quad (20)$$

В данной работе для рассматриваемого далее размера зоны локализации и зон действия антенн мы эмпирически приняли E_{max} = 0,6 м.

4 ЭКСПЕРИМЕНТЫ

С целью поиска наилучшего размещения антенн RFID-системы мы провели вычисление величины целевой функции $S(a)$ согласно выражению (16) для различных вариантов размещения 16 антенн. При этом мы рассматривали несколько простых паттернов размещения, представленных на рис. 3. Каждый вариант размещения антенн для каждого из паттернов также характеризовался некоторой величиной сдвига z.

В ходе исследований зона локализации была принята квадратной с размером 5 м × 5 м, а зоны действия антенн задавались в виде эллипсов с размерами малой оси R_{min} = 1,6 м, R_{max} = 1,9 м (при отношении длины большой оси к длине малой, равном b = 1,5625). Рассматривалось три варианта поворота антенн (и, соответственно, поворота их зон действия у) относительно вертикальной оси: 1) 45°; 2) 90°; 3) поворот каждой антенны отдельно по направлению к центру зоны локализации. При этом для каждого паттерна размещения антенн мы рассчитывали величину целевой функции $S(a)$ при изменении значения z от 0 м до 1,5 м с шагом в 5 см.

5 РЕЗУЛЬТАТЫ

Полученные нами результаты представлены на рис. 4 в виде зависимостей величины $S(a)$ от значения z. Наилучшие для каждого паттерна варианты (один для случая поворота всех антенн в одном направлении и один для случая поворота всех антенн к центру зоны локализации) представлены на рис. 5 в виде схем.

6 ОБСУЖДЕНИЕ

Рассмотрим результирующие зависимости величины целевой функции $S(a)$ от сдвига z, представленные на рис. 4. Видно, что согласно разработанному критерию наилучшим является вариант размещения антенн с использованием паттерна сетки при величине сдвига z = 85 см, шаге сетки 110 см и повороте антенн по направлению к центру зоны локализации (рис. 5б). Полученное при этом значение целевой функции равно $S(a)$ = 0,504. Важно отметить, что для всех вариантов рассмотренных паттернов (кроме паттерна треугольников) поворот антенн по направлению к центру зоны локализации позволил добиться увеличения величины целевой функции.

Наихудшие результаты, как и следовало ожидать, в основном были получены при использовании паттернов квадрата и окружности без размещения дополнительной антенны в центре, поскольку в этих случаях центральная часть зоны локализации остается не покрытой зонами действия антенн (рис. 5д–з). Также неудовлетворительные результаты получены для паттерна окружности с размещением дополнительной антенны в центре (рис. 5л). Это происходит из-за того, что, несмотря на внесение центральной антенны, при использовании такого паттерна для любых рассмотренных z в зоне локализации всегда остаются области, которые не покрыты зонами действия антенн (как в центре, так и по краям).

Проведенный в работе экспериментальный анализ не позволяет заявить о том, что полученное наилучшее размещение антенн является оптимальным не только среди рассмотренных, но и среди всех возможных размещений. Однако мы считаем, что разработанный критерий оптимальности адекватен и позволяет добиться минимальной средней ошибки локализации при использовании рассмотренного в работе комбинированного метода локализации.

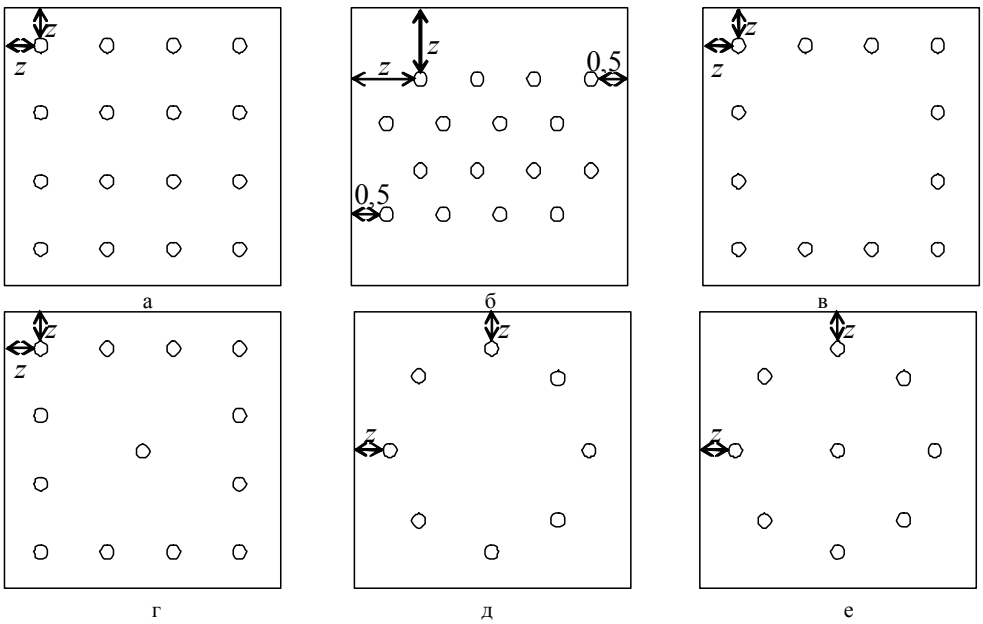


Рисунок 3 – Паттерны размещения антенн: а – сетка, б – треугольники, в – квадрат, г – квадрат с антенной в центре, д – окружность, е – окружность с антенной в центре

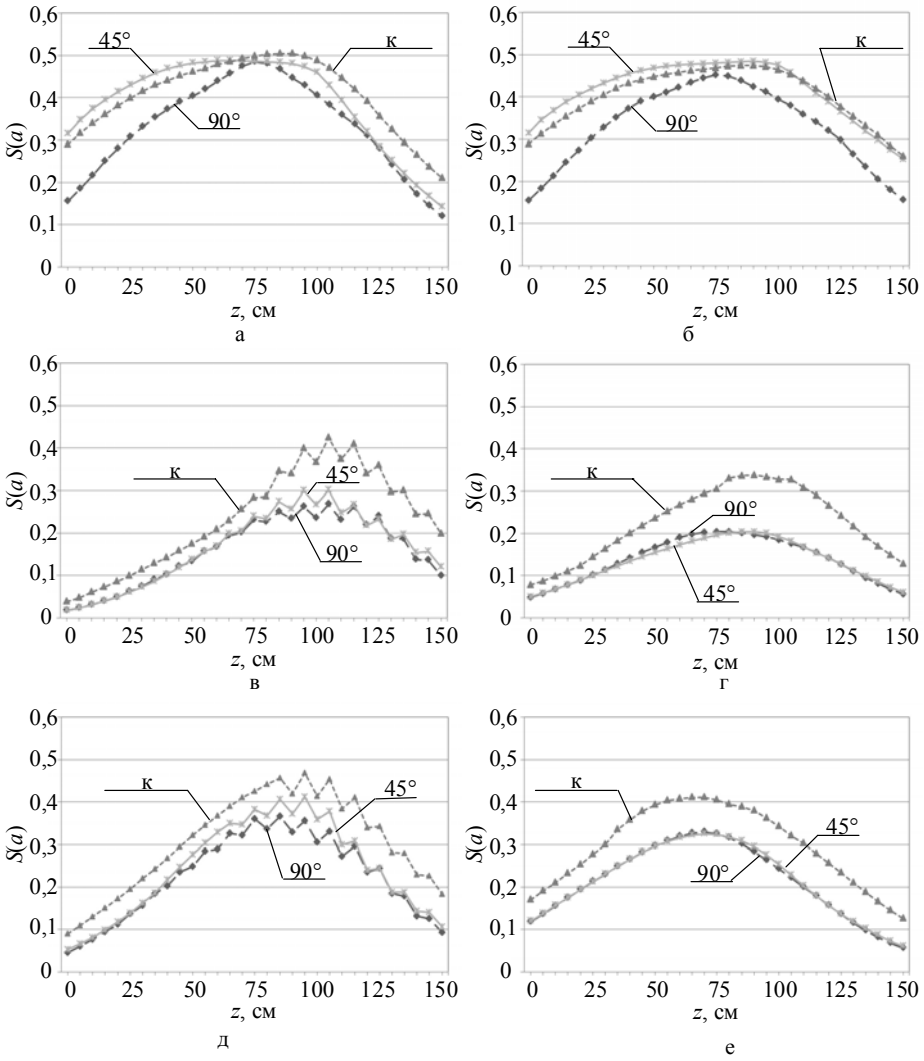


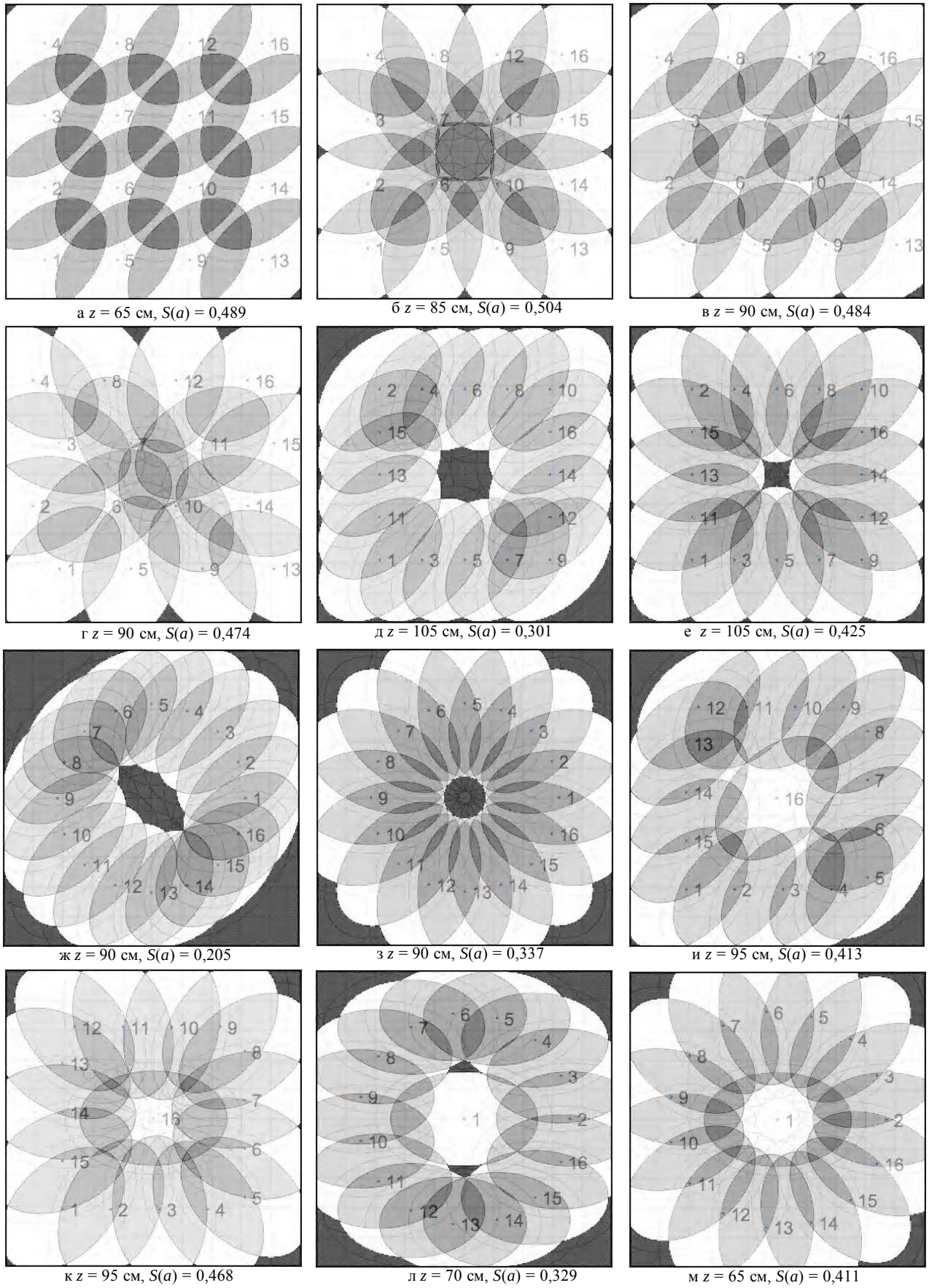
Рисунок 4 – Зависимости величины целевой функции $S(a)$ от сдвига z при различных вариантах поворота зон действия антенн и для различных паттернов размещения антенн: а – сетка, б – треугольники, в – квадрат, г – окружность, д – квадрат с антенной в центре, е – окружность с антенной в центре

ВЫВОДЫ

Разработана методика оптимизации размещения антенн для систем двумерной пространственной локализации RFID-меток при использовании комбинированного метода локализации. Используемый при этом критерий оптимальности основан на минимизации величины среднего HDOP, максимизации коэффициента покрытия зоны локализации зонами действия антенн и минимизации средней ошибки локализации алгоритма пересечений. В ходе экспериментальных исследований определено, что наилучшим среди рассмотренных паттернов размещения антенн является паттерн квадратной сетки при ориентации антенн по направлению к центру зоны локализации. В будущем планируется провести доработку методики с целью обеспечить возможность поиска оптимального размещения не только среди ряда заранее заданных паттернов, но и среди бесконечного множества нерегулярных вариантов размещения. Для этого планируется использовать эвристические алгоритмы.

СПИСОК ЛИТЕРАТУРЫ

1. Wang C. S. Goal-programming-driven genetic algorithm model for wireless access point deployment optimization / C. S. Wang, C. T. Chang // Mathematical problems in engineering. – 2012. – Vol. 2012. – P. 1–14.
2. Tak S. Modeling and performance study of AP placement framework for indoor location-awareness and network performance / S. Tak, T. Kim, D. Kim, Y. Kim // 8th International conference on ubiquitous information management and communication. Siem Reap, 9–11 January 2014 : proceedings. – Siem Reap, 2014. – P. 1–8.
3. Botero O. RFID network topology design based on Genetic Algorithms / O. Botero, H. Chaouchi // 2011 IEEE International conference on RFID-technologies and applications (RFID-TA). Sitges, 15–16 September 2011 : proceedings. – Sitges, 2011. – P. 300–305.
4. Fang S. H. A novel access point placement approach for WLAN-based location systems / S. H. Fang, T. N. Lin // 2010 IEEE Wireless communications and networking conference (WCNC), Sydney, 18–21 April 2010 : proceedings. – Sydney, 2010. – P. 1–4.
5. Ji Y. Optimal sniffers deployment on wireless indoor localization / Y. Ji, S. Biaz, S. Wu, B. Qi // 16th International conference on



а z = 65 см, $S(a)$ = 0,489 б z = 85 см, $S(a)$ = 0,504 в z = 90 см, $S(a)$ = 0,484

г z = 90 см, $S(a)$ = 0,474 д z = 105 см, $S(a)$ = 0,301 е z = 105 см, $S(a)$ = 0,425

ж z = 90 см, $S(a)$ = 0,205 з z = 90 см, $S(a)$ = 0,337 и z = 95 см, $S(a)$ = 0,413

к z = 95 см, $S(a)$ = 0,468 л z = 70 см, $S(a)$ = 0,329 м z = 65 см, $S(a)$ = 0,411

Рисунок 5 – Наилучшие варианты размещения антенн среди различных паттернов: а, б – сетка, в, г – треугольники, д, е – квадрат, ж, з – окружность, и, к – квадрат с антенной в центре, л, м – окружность с антенной в центре

computer communications and networks, 2007. ICCCN 2007, Honolulu, 13–16 August 2007 : proceedings. – Honolulu, 2007. – P. 251–256.
6. Farkas K. Optimization of Wi-Fi access point placement for indoor localization / K. Farkas, A. Huszak, G. Godor // Informatics & IT today. – 2013. – Vol. 1, No. 1. – P. 28–33.
7. Battiti R. Optimal wireless access point placement for location-dependent services / R. Battiti, M. Brunato, A. Delai // Information engineering and computer science. – 2003. – 12 p.
8. Savochkin D.A. Combinational RFID-based localization using different algorithms and measurements / D. A. Savochkin // 20th International conference on microwaves, radar, and wireless communications. Gdansk, 16-18 June 2014 : proceedings. – Gdansk, 2014. – P. 238–243.
9. Levanon N. Lowest GDOP in 2-D scenarios / N. Levanon // IEE Proceedings-radar, sonar and navigation. – 2000. – Vol. 147, No. 3. – P. 149–155.

Статья поступила в редакцию 21.01.2015.
После доработки 02.02.2015.

Савочкін Д. О.[1], Гімпілевич Ю. Б.[2]
[1]Аспірант кафедри радіотехніки та телекомунікацій Севастопольського національного технічного університету, Севастополь, Україна
[2]Д-р техн. наук, професор, завідувач кафедри радіотехніки та телекомунікацій Севастопольського національного технічного університету, Севастополь, Україна

ОПТИМІЗАЦІЯ РОЗМІЩЕННЯ АНТЕН ДЛЯ СИСТЕМ ПРОСТОРОВОЇ ДВОВИМІРНОЇ RFID-ЛОКАЛІЗАЦІЇ

Розроблено методику пошуку оптимального розміщення антен для систем просторової локалізації на основі технології радіочастотної ідентифікації (RFID). Методика заснована на мінімізації середньої помилки при реалізації комбінованого методу локалізації, що включає в себе алгоритм трилатерации, алгоритм відбитків і алгоритм перетинів. У ході проведеного обчислювального експерименту встановлено, що серед низки розглянутих патернів оптимальним для квадратної зони локалізації є варіант розміщення антен у вигляді квадратної сітки при орієнтації зон дії антен у напрямку до центру зони локалізації.

Ключові слова: RFID, локалізація, оптимізація, комбінований метод, розміщення антен.

Savochkin D. A.[1], Gimpilevich Yu. B.[2]
[1]Post-graduate student of Department of Radioelectronics and Telecommunications, Sevastopol National Technical University, Sevastopol, Ukraine
[2]Dr.Sc., Professor, Head of Department of Radioelectronics and Telecommunications, Sevastopol National Technical University, Sevastopol, Ukraine

OPTIMIZATION OF ANTENNA DEPLOYMENT FOR TWO-DIMENSIONAL RFID LOCALIZATION

This paper presents a method of search for the optimal antenna deployment for use in spatial localization systems based on the radio frequency identification technology (RFID). It involves minimizing the average error of a combinational localization method consisted of a trilateration algorithm, fingerprinting algorithm, and an intersectional algorithm. During the computational experiment we found the optimal solution among several antenna deployment patterns for a square localization field. The solution is a grid pattern with antenna interrogation zones oriented towards the field center.

Keywords: RFID, localization, optimization, combinational method, antenna deployment.

REFERENCES

1. Wang C. S., Chang C. T. Goal-programming-driven genetic algorithm model for wireless access point deployment optimization, *Mathematical problems in engineering,* 2012, Vol. 2012, pp. 1–14.
2. Tak S., Kim T., Kim D., Kim Y. Modeling and performance study of AP placement framework for indoor location-awareness and network performance, *8th International conference on ubiquitous information management and communication*. Siem Reap, 9–11 January 2014 : proceedings, Siem Reap, 2014, pp. 1–8.
3. Botero O., Chaouchi H. RFID network topology design based on Genetic Algorithms, *2011 IEEE International conference on RFID-technologies and applications (RFID-TA)*. Sitges, 15–16 September 2011, proceedings, Sitges, 2011, pp. 300–305.
4. Fang S. H., Lin T. N. A novel access point placement approach for WLAN-based location systems, *2010 IEEE Wireless communications and networking conference (WCNC)*. Sydney, 18–21 April 2010, proceedings, Sydney, 2010, pp. 1–4.
5. Ji Y., Biaz S., Wu S., Qi B. Optimal sniffers deployment on wireless indoor localization, *16th International conference on computer communications and networks*. Honolulu, 2007, ICCCN 2007, 13–16 August 2007, proceedings, Honolulu, 2007, pp. 251–256.
6. Farkas K., Huszak A., Godor G. Optimization of Wi-Fi access point placement for indoor localization, *Informatics & IT today*, 2013, Vol. 1, No. 1, pp. 28–33.
7. Battiti R., Brunato M., Delai A. Optimal wireless access point placement for location-dependent services, *Information engineering and computer science*, 2003, 12 p.
8. Savochkin D. A. Combinational RFID-based localization using different algorithms and measurements, *20th International conference on microwaves, radar, and wireless communications*. Gdansk, 16–18 June 2014, proceedings, Gdansk, 2014, pp. 238–243.
9. Levanon N. Lowest GDOP in 2-D scenarios, *IEE Proceedings-radar, sonar and navigation*. 2000, Vol. 147, No. 3, pp. 149–155.

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL AND COMPUTER MODELLING

УДК 528.85

Шама Є. О.

Аспірант кафедри радіотехніки та телекомунікацій Запорізького національного технічного університету, Запоріжжя, Україна

ЗМЕНШЕННЯ РОЗМІРНОСТІ ОЗНАКОВОГО ПРОСТОРУ РОСЛИННИХ ОБЄКТІВ ЗА ДОПОМОГОЮ ФАКТОРНОГО АНАЛІЗУ

У статті показана можливість зменшення кількості інформативних ознак для багатоспектральних даних рослинного походження за допомогою методів факторного аналізу. Для дослідження були використані реальні коефіцієнти відбиття кукурудзи, мишію та амброзії. Перед використанням метода факторного аналізу була виконана процедура перевірки відповідності початкового набору даних до нормального закону розподілу. Факторний аналіз проводився окремо для трьох рослин: кукурудзи, мишію та амброзії. Для кожної рослини встановлені розміри нового простору, а також надана якісна оцінка кореляції між елементами нового і початкового простору даних. Перехід до нового ознакового простору виконувався за умови, що рівень інформативності підтримувався на рівні не менше 80% по відношенню до початкового набору даних. Для кожного виду рослини була проведена процедура перевірки правильності вибору кількості ознак за допомогою розрахунку спільностей. Отримані результати можуть бути використані при побудові класифікатора рослинних об'єктів.

Ключові слова: факторний аналіз, ознаковий простір, спектральні коефіцієнти яскравості, ознаки, класифікатор.

НОМЕНКЛАТУРА

a_{ik} – факторні навантаження i-го значення k-го фактору;

F_k – k-й загальний фактор;

h_i^2 – спільність i-го значення;

N – кількість вимірювань, що зроблені для окремої рослини за допомогою спектрометра;

P – кількість факторів;

S_i^2 – специфічність i-го значення;

S_j – стандартне відхилення j-го значення;

$x_{\bullet j}$ – середнє значення j-ї ознаки;

X_j – j-та спостережувана змінна;

z_{ij} – компоненти матриці стандартизованих початкових даних;

U_j – j-та складова специфічних факторів;

λ_i – i-та довжина хвилі;

μ_i – i-й вимірюваний спектральний коефіцієнт відбиття;

$\mu_{i,\max}$ – i-й максимальний вимірюваний спектральний коефіцієнт відбиття;

$\mu_{i,\min}$– i-й мінімальний вимірюваний спектральний коефіцієнт відбиття;

ω_i – i-й нормований спектральний коефіцієнт відбиття.

ВСТУП

Відомо, що велика розмірність ознакового простору становить суттєву проблему для багатьох процедур розпізнавання образів, зокрема рослинних об'єктів.

Об'єктом дослідження є процес зменшення розмірності простору ознак для побудови розпізнавальних моделей.

На сьогоднішній день зростає кількість теоретичних робіт по методам зменшення розмірності ознакового простору для класифікації образів [1]. Деякі з цих методів прагнуть сформувати нові ознаки на основі лінійних комбінацій старих ознак, а інші прагнуть створити меншу підмножину початкових ознак. Основна проблема цієї теорії полягає в тому, що поділ розпізнавання образів на виділення ознак, а потім класифікацію теоретично є штучним. Повністю оптимальний видільник ознак є не чим іншим, як оптимальним класифікатором. А тому, вирішивши задачу зменшення інформативних ознак, при збереженні необхідного рівня інформативності, можна спростити задачу побудови класифікатора рослинних об'єктів.

Більшість запропонованих методів зменшення розмірності [1] забезпечують функціональне відображення, так що можливо знайти відображення довільного вектора ознак [2]. Класичними процедурами зменшення розмірності є аналіз головних компонент і факторний аналіз

DOI 10.15588/1607-3274-2015-2-2

[3], обидва з яких зменшують розмірність шляхом формування лінійних комбінацій ознак. Метою аналізу головних компонент (розкладання Кархунена-Лоева) є знаходження зображень у просторі меншої розмірності, яке враховує дисперсію ознак. Метою факторного аналізу є знаходження зображень у просторі меншої розмірності, яке враховує кореляцію між ознаками. Різні алгоритми факторного аналізу об'єднані тим, що в усіх них відбувається перехід до нового базису в початковому *n*-мірному просторі, тому у статті буде використаний факторний аналіз, так як він є загальним методом зменшення розмірності серед методів головних компонентів та факторного аналізу.

Предметом дослідження є процедура зменшення розмірності ознакового простору коефіцієнтів відбиття рослинних об'єктів за допомогою методів факторного аналізу. При цьому, початкова (перед проведенням процедури зменшення ознакового простору) розмірність коефіцієнтів відбиття рослинних об'єктів, складає 256, яка визначається роздільною здатність вимірювального приладу – спектрометра.

Мета роботи полягає у дослідженні можливості зменшення кількості інформаційних ознак за допомогою використання методів факторного аналізу.

1 ПОСТАНОВКА ЗАДАЧІ

Нехай ми маємо масив експериментально отриманих даних коефіцієнтів відбиття рослинних об'єктів X_j. Розмірність масиву коефіцієнтів відбиття рослинних об'єктів – $256 \times N$, де 256 – розмірність ознакового простору (визначається роздільною здатністю спектрометру). Тоді, задача зменшення ознакового простору рослинних об'єктів, полягає у тому, щоб отримати масив меншої розмірності даних для відповідної рослини, за допомогою якого можна зробити процедуру розпізнавання із заздалегідь встановленим рівнем вірогідності. Тобто необхідно згенерувати вихідний масив даних F_k, який був би меншим за розміром і за допомогою якого можливо було б зробити процедуру розпізнання рослини за видом. Так, якщо вихідний масив можливо описати за допомогою лінійного перетворення $F_k = \sum_{j=1}^{m} a_{jk} \cdot X_j$, то вирішувана задача в рамках цієї статті полягає в здобутті нового масиву, у якого $m << N$ і $m \to \min$.

2 ОГЛЯД ЛІТЕРАТУРИ

Найбільш відомими методами розпізнавання є метод потенціальних функцій [1, 7], дискримінантний аналіз [1, 2, 7], метод граничних спрощень [8], регресійний аналіз [2], метод групового урахування аргументів [7], нейромережеві методи [1, 2]. Порівняльний аналіз основних методів розпізнавання – кореляційних, ознакових (статистичних і детермінованих), нейромережевих [2] – дозволяє зробити висновок, що метод факторного аналізу є найбільш ефективним для вирішення завдань розпізнавання [9]. Крім того, факторні методи забезпечують більш високу стійкість роботи системи розпізнавання в умовах перешкод і дозволяють об'єднати процедуру розпізнавання і процедуру зменшення ознак.

Розпізнавання рослинних об'єктів за допомогою факторного аналізу розглянуто у ряді робот [3, 5, 9]. Але в цих роботах класифікація рослинних об'єктів проводилась лише на два класи (корисну культуру і бур'яни). Проте на практиці доцільною є класифікація на більшу кількість класів, що обумовлюється необхідністю використання різних гербіцидів для різних видів бур'янів. Більша кількість класів обумовлена тим, що для боротьби з бур'янами при хімічній обробці застосовують різні види гербіцидів. Необхідність розпізнавання рослинних об'єктів на три класи визначається також алгоритмом роботи інтелектуального обприскувача для обробки просапних культур в рослинництві. Також процедура розпізнавання в моделях на основі дискримінантного аналізу [1, 6], нейронної мережі [2, 4, 8], регресійного аналізу [7], як правило, характеризується великою надлишковістю і в них висуваються високі вимоги до обчислювальних ресурсів. Для того, щоб виключити недоліки цих методів, доцільно застосувати процедуру факторного аналізу для зменшення надлишковості даних моделі при розпізнаванні рослинних об'єктів, що скоротить час навчання та суттєво збільшить швидкість прийняття рішень і дозволить зменшити вимоги до обчислювальних ресурсів.

3 МАТЕРІАЛИ І МЕТОДИ

Факторний аналіз дозволяє розв'язати дві важливі проблеми, які виникають під час дослідження: компактно описати об'єкт виміру і в той же час всебічно. За допомогою факторного аналізу можливе виявлення прихованих змінних факторів, що відповідають за наявність лінійних статистичних кореляцій між спостережуваними змінними [4].

При проведенні факторного аналізу, в один чинник об'єднуються змінні, що сильно корелюють між собою, як наслідок відбувається перерозподіл дисперсії між компонентами і виходить максимально проста і наявна структура факторів. Після об'єднання корельованість компонентів, всередині кожного фактору між собою, буде вищим, ніж їх корельованість з компонентами з інших факторів. Ця процедура також дозволяє виділити латентні змінні, що може бути корисним при побудові класифікатора рослинних об'єктів.

Таким чином можна виділити дві мети факторного аналізу:

– визначення взаємозв'язків між змінними (класифікація змінних);

– зменшення кількості змінних необхідних для опису даних.

При факторному аналізі, коли необхідно подати початкові змінні у вигляді лінійної комбінації факторів *F*, які розраховані так, щоб найкращим способом (з мінімальною погрішністю) подати *X*, можна показати [10]:

$$X_j = \sum_{k=1}^{p} a_{jk} \cdot F_k + U_j\ ,\ \ k=1,2,\ldots,p;\ \ j=1,2\ldots,m. \qquad (1)$$

Одним з найбільш поширених методів пошуку факторів є метод головних компонент (саме цей метод будемо використовувати в межах факторного аналізу). Його основна відмінність від факторного аналізу полягає в тому,

що головні компоненти F_k пов'язані із спостережуваними змінними X_j лінійними функціями перетворення:

$$X_j = \sum_{k=1}^{p} a_{jk} \cdot F_k \text{ і } F_k = \sum_{j=1}^{m} a_{jk} \cdot X_j. \quad (2)$$

Суть цього методу полягає в заміні корельованих компонентів некорельованими чинниками. Іншою важливою характеристикою методу є можливість обмежитися найбільш інформативними головними компонентами і виключити інші з аналізу, що спрощує інтерпретацію результатів. Цінність метода головних компонентів також в тому, що він – єдиний математично обґрунтований метод факторного аналізу [5]. Але слід зауважити, що за твердженням деяких дослідників, метод головних компонентів не є методом факторного аналізу, оскільки не розщеплює дисперсію індикаторів на загальну і унікальну [6].

Метод головних компонент є простішим в розрахунках і інтерпретації, але одна з головних труднощів його використання – необхідність перетворення початкових даних, які представлені в різних одиницях виміру, в величини, які можливо співставити. Традиційним методом перетворення, але не обов'язковим, є нормування по стандартним відхиленням, коли матриця Z стандартизованих початкових даних визначається за формулою:

$$z_{ij} = \frac{x_{ij} - x_{\bullet j}}{S_j}, j=1,2\ldots,m; i=1,2\ldots,n. \quad (3)$$

Також слід зауважити, що при проведенні факторного аналізу за допомогою методу головних компонент може використовуватись така характеристика, як спільність, яка визначається формулою:

$$h_i^2 = \sum_{k=1}^{M} a_{ik}^2 + S_i^2, \quad (4)$$

Тобто спільність є дисперсією змінних, обумовленою факторами, а специфічність являє собою частину дисперсії, яка обумовлена випадковими помилками або змінними, неврахованими в моделі. Відповідно до постановки завдання необхідно шукати такі фактори, при яких сумарна спільність є максимальною, а специфічність – мінімальною.

У роботі [7] виділяють такі умови проведення факторного аналізу: всі ознаки повинні бути кількісними; кількість спостережень має бути не менше як у два рази більшою кількості змінних; вибірка має бути однорідною; факторний аналіз здійснюється по корелюючими змінними.

4 ЕКСПЕРИМЕНТИ

Виміри спектральних коефіцієнтів відбиття проводилися влітку 2013 року на полі ТОВ «Агрофірма «Матвіївка» (Запорізька область, Україна). В якості корисної культури була обрана кукурудза. Також на полі зустрічались такі бур'яни: амброзія, берізка, мишій, пирій, пастуша сумка, суріпиця та ін. Серед бур'янів, для дослідження, були відібрані мишій та амброзія, оскільки вони складали переважну долю бур'янів на полі. Вимірювання проводилися максимально наближеним до реальних польових робіт на самохідній установці: швидкість руху 0,5–1 м/с, відстань об'єктива до рослини – приблизно 0,3 м. Переміщення об'єктива приймача спектрометра від однієї рослини до іншої в продовж рядка проводилося рівномірно, без ривків. Вимірювання проводилися впродовж проходу міжряддя як в одному напрямку так і в іншому. Пристрій, за допомогою якого проводилися вимірювання спектральних коефіцієнтів відбиття, містив спектрометр Red Tide650 з волоконно-оптичним кабелем P200 – 2 – UV – VIS і лінзою 74 – VIS (фірма Ocean Optics) та веб-камеру A4Tech PK-838G.

Вимірювання спектральних характеристик рослин проводилися за допомогою приладу, який автоматично з періодом 1 с зберігав поточну інформацію (коефіцієнт відбиття і координати місця) і фотографію зовнішнього вигляду рослини.

З усіх експериментальних даних, для подальшого дослідження було відібрано 2268 спектральних коефіцієнтів відбиття рослин. Критерієм відбору була можливість візуального визначення виду рослин по зображеннях і формі спектральної кривої. Назва і кількість рослин, які були відібрані для подальшого дослідження, наведені в таблиці 1.

Таблиця 1 – Видовий склад досліджуваних рослин

Вид рослини	Латинська назва	Кількість екземплярів
Кукурудза	*Zéa máys*	699
Амброзія	*Ambrósia artemisiifólia*	882
Мишій	*Setaria viridis*	687
Усього		2268

РЕЗУЛЬТАТИ

За допомогою спектрометра експериментально були отримані спектральні коефіцієнти відбиття рослин, які були потім експериментально оброблені. Було отримано три матриці: матриця коефіцієнтів відбиття для кукурудзи – 699×256, матриця коефіцієнтів відбиття для мишію – 687×256 і матриця коефіцієнтів відбиття для амброзії – 882×256. Наступним етапом було нормування цих даних. Кожна з трьох матриць піддавалася нормуванню за лінійним законом:

$$\omega_i = \frac{\mu_i - \mu_{i,\min}}{\mu_{i,\max} - \mu_{i,\min}}. \quad (5)$$

Перед проведенням факторного аналізу необхідно перевірити генеральну сукупність даних для кожної рослини на відповідність нормальному розподілу.

Нормальність розподілу даних для кукурудзи, мишію і амброзії була перевірена у статистичному пакеті програм. Для перевірки розподілу на нормальність використовувались нормовані значення мишію. Результати перевірки наведені на рис. 1 і в табл. 2.

Таким чином, аналіз досліджуваного розподілу вибірок визначив його наближеність до нормального, на що вказують як вид гістограми, так і значення коефіцієнтів асиметрії та ексцесу: вони одного порядку по відношенню до початкових значень вибірок, а також мають малі значення стандартних помилок. Для усіх видів аналізу статистично значимими вважали відмінності при $P<0,05$.

Набір початкових ознак для кукурудзи та амброзії також піддавався аналогічній процедурі перевірки на нормальність і також було встановлено наближеність роз-

поділів цих рослин до нормального закону (що не дивно, бо природа походження і умови виміру коефіцієнтів яскравості для усіх трьох рослин є однаковими).

Наступним етапом було проведення факторного аналізу. При проведенні факторного аналізу в програмному пакеті були встановлені наступні параметри: метод виділення факторів – компонентний аналіз; максимальна кількість факторів – 256; мінімальне власне число – 0.

Факторний аналіз будемо використовувати окремо для кожного виду рослинних об'єктів: корисної культури (кукурудза), однодольного бур'яну (мишій), дводольного бур'яну (амброзія). Умовою переходу від початкового *n*-мірного простору до нового ознакового простору, буде

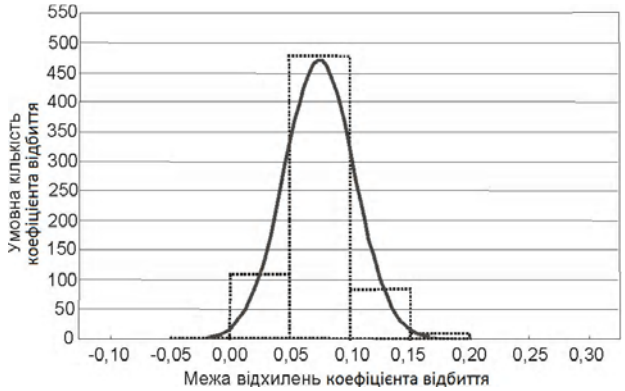


Рисунок 1 – Гістограма усередненої вибіркової сукупності значень мишію. Суцільна лінія – крива нормального закону розподілу

рівень інформативності нового простору, який повинен бути не менше 80% по відношенню до початкового.

Рівень інформативності факторного ознакового простору встановимо рівним 80% по відношенню до початкового. Це є достатнім рівнем інформативності при переході від одного ознакового простору до іншого [8]. Для визначення кількості факторів, які можна залишити і при цьому відповідати встановленому рівню інформативності визначимо вклад головних компонент для кожної рослини.

На таблицях 3 та 4 в перших стовбцях наведені оцінки власних чисел, в третьому стовбці – накопичені значення власних чисел, у другому і у четвертому стовбцях – відносний вклад кожної головної компоненти та накопичений відносний вклад відповідно.

У першому наближенні накопичений відносний вклад можна вважати рівнем інформативності нового ознакового простору по відношенню до початкового (для більш точного значення рівня інформативності необхідно зробити перерахунок через власні числа з урахуванням довірчих інтервалів) [9]. Як бачимо, для досягнення встановленого рівня інформативності, для кукурудзи можна залишити 3 фактора, для мишію – 8 факторів і для амброзії – 2 фактора.

Для інтерпретації нових ознак необхідно провести аналіз матриці факторних навантажень з урахуванням нової кількості факторів для кожної рослини.

5 ОБГОВОРЕННЯ

Як бачимо з таблиць 5 і 6, для кукурудзи перший фактор найбільш корельований з початковими ознаками, що відповідають довжинам хвиль 353–464 нм, 631–704 нм і

Таблиця 2 – Результати розрахунку критеріїв асиметрії та ексцесу для мишію

Змінна	Середнє значення	Мінімум	Максимум	Станд. відхилення	Асиметрія	Похибка асиметрії	Ексцес	Похибка ексцесу
1	0,083706	0	0,600420	0,071435	1,47213	0,093251	5,46625	0,186233
2	0,015076	0	0,124923	0,018389	1,58612	0,093251	3,48430	0,186233
3	0,013580	0	0,153521	0,018177	2,09398	0,093251	7,26450	0,186233
4	0,057173	0	0,203566	0,025330	1,11122	0,093251	3,16737	0,186233
5	0,054378	0	0,243602	0,025989	1,15261	0,093251	4,32015	0,186233
6	0,059009	0	0,225000	0,027683	1,38278	0,093251	5,47856	0,186233
7	0,054660	0	0,204996	0,026185	0,82713	0,093251	1,94301	0,186233

Таблиця 3 – Вклад головних компонент в сумарну дисперсію початкових ознак для кукурудзи і для мишію

№ фактору	Власні значення для кукурудзи				Власні значення для мишію			
	Значення	% від загального	Накопичене значення	% накопичення	Значення	% від загального	Накопичене значення	% накопичення
1	163,3985	63,827	163,3985	63,827	101,1679	39,5187	101,1679	39,518
2	29,9047	11,681	193,3032	75,509	70,7136	27,6224	171,8815	67,141
3	13,4457	5,252	206,7489	80,761	15,0487	5,8784	186,9302	73,019
4	2,9825	1,165	209,7314	81,926	9,94743	3,7009	196,4046	76,720
5	1,5078	0,588	211,1577	82,515	4,8562	1,8969	201,2607	78,617
6	0,9185	0,358	212,1577	82,874	1,5239	0,5952	202,7847	79,212
7	0,8381	0,327	212,9958	83,201	1,4546	0,5682	204,2393	79,781
8	0,7767	0,303	213,7725	83,504	1,0522	0,4110	205,2914	80,192
9	0,7139	0,278	214,4864	83,783	1,0216	0,3990	206,3130	80,591
10	0,6942	0,271	215,1806	84,057	0,9499	0,3710	207,2629	80,962
11	0,6580	0,257	215,8387	84,312	0,8506	0,3322	208,1135	81,294
12	0,6277	0,245	216,4664	84,557	0,8196	0,3201	208,9331	81,614
13	0,5955	0,232	217,0618	84,789	0,7968	0,3112	209,7298	81,925
14	0,5709	0,223	217,6327	85,012	0,7567	0,29558	210,4865	82,221
15	0,5537	0,216	218,1864	85,229	0,7094	0,2771	211,1959	82,498
16	0,5488	0,214	218,7352	85,443	0,7006	0,2773	211,8965	82,772
17	0,5163	0,201	219,2515	85,645	0,6876	0,2685	212,5841	83,040
18	0,5048	0,197	219,7563	85,842	0,6605	0,2580	213,2446	83,298
19	0,4910	0,191	220,2474	86,034	0,6506	0,2541	213,8951	83,552

Таблиця 4 – Вклад головних компонент в сумарну дисперсію початкових ознак для амброзії

№ фактору	Власні значення для амброзії			
	Значення	% від загального	Накопичене значення	% накопичення
1	145,327	56,76864	145,3277	56,7686
2	79,2253	30,94740	224,5530	87,7160
3	13,0125	5,08299	237,5655	92,7990
4	6,7203	2,62510	244,2858	95,4241
5	3,1865	1,24471	247,4722	96,6688
6	1,4380	0,56172	248,9103	97,2306
7	0,9799	0,38276	249,8901	97,6133
8	0,7389	0,28863	250,6290	97,9020
9	0,6168	0,24094	251,2458	98,1429
10	0,5851	0,22856	251,8309	98,3715
11	0,3474	0,13571	252,1783	98,5072
12	0,2324	0,09077	252,4107	98,5979
13	0,1990	0,07772	252,6097	98,6757
14	0,1852	0,07233	252,7948	98,7480
15	0,17764	0,06892	252,9713	98,8169
16	0,1617	0,06318	253,1330	98,8801
17	0,1588	0,06203	253,2918	98,9421
18	0,1416	0,05530	253,4334	98,9974
19	0,1335	0,05215	253,5669	99,0496
20	0,1194	0,04475	253,6862	99,0962
21	0,1146	0,06318	2553,8008	99,1409

Таблиця 5 – Факторні навантаження для кукурудзи та для амброзії

Змінні	Факторні навантаження (для кукурудзи)			Факторні навантаження (для амброзії)	
	Фактор 1	Фактор 2	Фактор 3	Фактор 1	Фактор 2
1	0,3154	–0,00652	0,13011	–0,5882	0,0373
2	–0,9133	0,00270	–0,02649	–0,3287	–0,20267
3	–0,8951	–0,01129	–0,02044	–0,3214	–0,17772
4	–0,9299	0,02037	0,01800	–0,8883	–0,26265
5	–0,9362	0,02430	0,06539	–0,8995	–0,28255
6	–0,9240	0,02867	0,07868	–0,8882	–0,26591
7	–0,9262	0,02683	0,00677	–0,8948	–0,27936
8	–0,9344	0,04891	0,06100	–0,8981	–0,27451
9	–0,9203	0,03827	0,05974	–0,9045	–0,27820
10	–0,9367	0,02823	0,04644	–0,9076	–0,27883
Нова 1	–0,9327	0,01571	0,15592	–0,9072	–0,27496
Нова 2	–0,9316	0,02617	0,09776	–0,9078	–0,27583
Нова 3	–0,9281	0,04715	0,16311	–0,9086	–0,27435
Нова 4	–0,9229	0,04346	0,14530	–0,9095	–0,27747
Нова 5	–0,9211	0,03981	0,18024	–0,9131	–0,27920
Нова 6	–0,9326	0,03498	0,12592	–0,9163	–0,28652
Нова 7	–0,9298	0,05401	0,13759	–0,9231	–0,28774
Нова 8	–0,9184	0,02574	0,20173	–0,9204	–0,27198
Нова 9	–0,9200	0,06748	0,18150	–0,9230	–0,28410
Нова 10	–0,9235	0,09543	0,17181	–0,9256	–0,28923
Нова 11	–0,8994	0,10865	0,22821	–0,9270	–0,28758
Нова 12	–0,8508	0,08934	0,31835	–0,9261	–0,28418
Нова 13	–0,9194	0,10535	0,08860	–0,9284	–0,29927
Нова 14	–0,9212	0,10979	0,11372	–0,9272	–0,30349

802–1050 нм, другий фактор – корельований з початковим ознаками, що відповідають довжинам хвиль 718–753 нм і третій фактор – корельований з початковим ознаками, що відповідають довжинам хвиль 507–540 нм.

Для мишію перший фактор найбільш корельований з початковими ознаками, що відповідають довжинам хвиль 414–504 нм, 547–683 нм і 712–850 нм; другий фактор – корельований з початковим ознаками, що відповідають довжинам хвиль 850–1050 нм; третій фактор – з початковим ознаками, що відповідають довжинам хвиль 510–535 нм; четвертий фактор – з початковим ознаками, що відповідають довжинам хвиль 437–480 нм; п'ятий фактор – з початковим ознаками, що відповідають довжинам хвиль 400–434 нм; шостий фактор – з початковим ознаками, що відповідають довжинам хвиль 353 нм, 370 нм і 976 нм; сьомий фактор – з початковим ознаками, що відповідають довжинам хвиль 350 нм і 996 нм; восьмий фактор – з початковим ознаками, що відповідають довжинам хвиль 350–360 нм.

Для амброзії перший фактор найбільш корельований з початковими ознаками, що відповідають довжинам хвиль 360–504 нм, 614–675 нм, 695–944 нм і 712–850 нм; другий фактор – корельований з початковим ознаками, що відповідають довжинам хвиль 495–510 нм, 535–593 нм, 678–690 нм, 947–1038 нм.

Для перевірки достатності розміру нового простору для кожної рослини, була проведена процедура розрахунку сумарної спільності і спільності по кожному фактору. Так, для амброзії сумарна спільність дорівнює 254,28, для першого фактору – 145,27, для другого фактору – 124,55. Тобто загальна спільність за двома факторами дорівнює 145,27+124,55=269,82. Оскільки загальна спільність за двома факторами (269,82) більше ніж сумарна спільність (254,28), то розмірність ознакового простору у кількості двох обрано вірно. У табл. 7 наведено результат розрахунку спільності для амброзії. Графіки власних значень для кукурудзи, мишію та амброзії наведені рис. 2–4.

Аналогічна процедура перевірки нових розмірностей була проведена для кукурудзи та мишію. В результаті проведених розрахунків, розмірність ознакових просторів для кукурудзи та мишію обрано вірно, оскільки для мишію сума спільностей по факторам дорівнює 356,87, а сумарна спільність 224,78. Для кукурудзи: сума спільностей по факторам – 256,14, а сумарна спільність – 229,24.

6 ВИСНОВКИ

В роботі розглянута задача зменшення кількості інформаційних ознак за допомогою факторного аналізу, а саме методу головних компонентів.

В результаті проведення експериментальних досліджень встановлено, що можливо перейти до нового простору з меншою кількістю ознак при збереженні рівня інформативності не менше 80%. Так для кукурудзи

Таблиця 6 – Факторні навантаження для мишію

Змінні	Факторні навантаження (для мишію)							
	Фактор 1	Фактор 2	Фактор 3	Фактор 4	Фактор 5	Фактор 6	Фактор 7	Фактор 8
1	–0,0832	0,00065	–0,13211	0,259002	0,121068	0,141987	0,782968	–0,162445
2	–0,1140	–0,44176	–0,04021	–0,238139	–0,088848	–0,385649	–0,031048	0,477626
3	–0,0672	–0,41766	–0,03178	–0,320021	–0,026030	–0,079956	–0,299183	–0,486865
4	–0,2888	–0,64547	–0,03854	–0,235601	–0,035044	–0,041635	–0,060987	–0,088411
5	–0,3005	–0,60634	–0,03809	–0,231377	–0,056948	–0,041635	0,051629	0,100092
6	–0,3465	–0,60569	–0,10694	–0,226901	–0,054001	–0,167011	0,081305	0,191632
7	–0,3274	–0,65189	–0,05375	–0,200093	–0,124574	–0,062364	0,003541	–0,067721
8	–0,3494	–0,62809	–0,14527	–0,187012	–0,091447	–0,229901	–0,024393	–0,074242
9	–0,3689	–0,56764	–0,12693	–0,208454	–0,156561	–0,083712	0,064355	0,037291
10	–0,3340	–0,61098	–0,07223	–0,163582	–0,112814	–0,254564	0,072627	0,001780
Нова 1	–0,3455	–0,69667	–0,07004	–0,145086	–0,046614	0,036432	–0,012333	–0,037291
Нова 2	–0,3741	–0,67225	–0,14621	–0,171111	–0,099000	–0,089167	0,014638	0,001780
Нова 3	–0,4215	–0,69452	–0,18345	–0,149507	–0,070722	0,185336	–0,010266	0,020758
Нова 4	–0,4024	–0,66716	–0,13470	–0,159141	–0,091824	0,131594	–0,032505	–0,019124
Нова 5	–0,3559	–0,71232	–0,10534	–0,125823	–0,070673	0,001574	0,012386	0,117445
Нова 6	–0,4320	–0,67551	–0,11532	–0,144367	–0,127146	–0,024959	0,010995	–0,012405
Нова 7	–0,4828	–0,63595	–0,14191	–0,067666	–0,123461	0,020609	0,013794	–0,040740
Нова 8	–0,4398	–0,69172	–0,17675	–0,075642	–0,076265	0,158533	–0,037117	0,022666
Нова 9	–0,4870	–0,64593	–0,14645	–0,133559	–0,135250	0,132538	0,018388	0,037193
Нова 10	–0,5379	–0,62809	–0,12343	0,021363	–0,172884	0,028273	–0,012258	–0,042859
Нова 11	–0,5832	–0,62341	–0,17869	0,034503	–0,174038	0,156002	–0,036742	0,083318
Нова 12	–0,5296	–0,66663	–0,23535	0,018286	–0,136432	0,258610	–0,062037	0,079887
Нова 13	–0,6214	–0,53139	–0,06501	0,025591	–0,274416	0,009128	–0,000690	0,083146

Таблиця 7 – Результат розрахунку спільностей для амброзії

Змінна	Спільність для амброзії		
	Від фактора 1	Від фактора 2	Значення *R*-Square
1	0,345961	0,347358	0,951460
2	0,108021	0,149095	0,811392
3	0,103269	0,134854	0,855376
4	0,789116	0,858102	0,995631
5	0,809054	0,888888	0,997044
6	0,788982	0,859689	0,997014
7	0,800586	0,878628	0,997250
8	0,806656	0,882010	0,997462
9	0,818060	0,895454	0,998322
10	0,823716	0,901460	0,998506

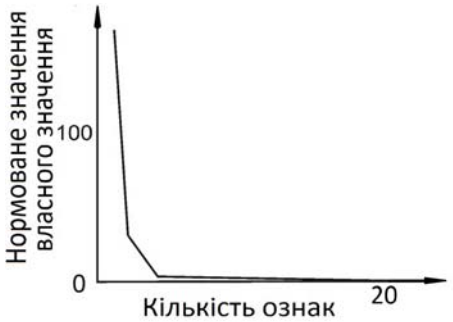


Рисунок 2 – Графік власних значень для кукурудзи

Рисунок 3 – Графік власних значень для мишію

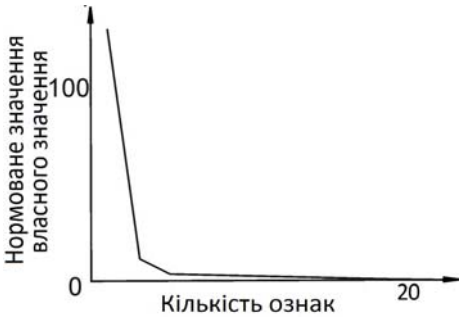


Рисунок 4 – Графік власних значень для амброзії

розмірність нового простору дорівнює три, для мишію вісім, а для амброзії два. Різницю в кількості ознак для досліджуваних рослин, можна пояснити різною геометричною структурою листків, їхньою відмінністю у морфологічному та анатомічному складу, а це в свою чергу накладає відбиток на структурі багатоспектральних даних.

Практична цінність отриманих результатів полягає в тому, що результати розрахунків можуть бути корисними при побудові класифікатора рослинних об’єктів. Як зазначалось, при використанні факторного аналізу, окремі процедури виявлення ознак і класифікація не потрібні, вони вже закладені в самому методі., а тому класифікатор, в алгоритм розпізнавання рослин якого буде закладений метод факторного аналізу, буде мати більш просту схему побудови, а це в свою чергу дасть виграш в економічному сенсі.

7 ПОДЯКИ

Стаття пов’язана з бюджетною науково-дослідною роботою Запорізького національного технічного університету «Методи, моделі та пристрої прийняття рішень в системах розпізнавання образів», номер державної реєстрації 0111U000059.

СПИСОК ЛІТЕРАТУРИ

1. Загоруйко Н.Г. Методы распознавания и их применение / Н. Г. Загоруйко. – М. : Сов. радио, 1972. – 270 с.
2. Дуда Р. Распознавание образов и анализ сцен / Р. Дуда, П. Харт. – М. : Мир, 1976. – 326 с.
3. Харман Г. Современный факторный анализ / Г. Харман. – М. : Статистика, 1972. – 488 с.
4. Бююль А. SPSS: Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей / А. Бююль, П. Цефель. – СПб.: ООО «ДиаСофтЮП», 2002. – 603 с.
5. Факторный, дискриминантный и кластерный анализ: пер. с англ. / [Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка и др.; под ред. И. С. Енюкова. – М. : Финансы и статистика, 1989. – 215 с.
6. Прикладная статистика. Классификация и снижение размерности. Справочное издание/ [С. А. Айвазян, В. М. Бухштабер, И. С. Енюков, Л. Д. Мешалкин]. – М. : Финансы и статистика, 1989. – 607 с.
7. Купер К. Индивидуальные различия / К. Купер. – М. : Аспект Пресс, 2000. – 527 с.
8. Боровиков В. STATISTICA: искусство анализа данных на компьютере. Для профессионалов / В. Боровиков. – СПб: Питер, 2001. – 656 с.
9. Митина О. В. Факторный анализ для психологов / О. В. Митина, И. Б. Михайловская. – М. : Учебно-методический колектор, Психология, 2001. – 169 с.
10. Краснопоясовський А. С. Факторний класифікаційний аналіз за методом функціонально-статистичних випробувань / А. С. Краснопоясовський // Радіоелектронні та комп'ютерні системи. – 2004. – № 4. – С. 46–50.

Стаття надійшла до редакції 01.09.2014.
Після доробки 27.01.2015.

Шама Е. А.
Аспирант кафедры радиотехники и телекоммуникации Запорожского национального технического университета, Запорожье, Украина

УМЕНЬШЕНИЕ РАЗМЕРНОСТИ ПРИЗНАКОВОГО ПРОСТРАНСТВА РАСТИТЕЛЬНЫХ ОБЪЕКТОВ С ПОМОЩЬЮ ФАКТОРНОГО АНАЛИЗА

В статье показана возможность уменьшения числа информативных признаков для многоспектральных данных растительного происхождения с помощью методов факторного анализа. Для исследования были использованы реальные коэффициенты отражения кукурузы, мышея и амброзии. Перед использованием метода факторного анализа была выполнена процедура проверки соответствия начального набора данных нормальному закону распределению. Факторный анализ проводился отдельно для трех растений: кукурузы, мышея и амброзии. Для каждого растения установлены размеры нового пространства, а также предоставлена качественная оценка корреляции между элементами нового и начального пространства данных. Переход к новому признаковому пространству производился при условии, что уровень информативности поддерживался на уровне не менее 80% по отношению к начальному набору данных. Для каждого вида растения была проведена процедура проверки правильности выбора количества признаков с помощью расчета общностей. Полученные результаты могут быть использованы при построении классификатора растительных объектов.

Ключевые слова: факторный анализ, признаковое пространство, спектральные коэффициенты яркости, признаки, классификатор.

Shama E. O.
Post-graduate Student of Department of Radio Engineering of Zaporizhzhya National Technical University, Ukraine

THE FUATURE SPACE DIMENSIONALITY REDUCTION OF PLANT OBJECTS BASED ON FACTOR ANALYSIS

A possibility of reducing the number of informative features for multispectral phytogenous data by means of the factor analysis method is shown in the article. For investigation there were used real reflection coefficients of maize, bristlegrass and ambrosia. Prior to usage of the factor analysis it was performed the procedure to verify an initial set of data for conformity with distribution by the normal distribution law. The factor analysis was conducted separately for three plants: maize, bristlegrass and ambrosia. Sizes of a new space were set for each plant, and also there was given quality estimation of correlation between elements of a new and initial space of the data. Transfer to new features was executed provided that a level of informativeness was maintained at a level of 80% minimum in comparison with the initial set of data. For each plant species the procedure of verification of the feature number choice correctness was performed by calculating of generalities. The obtained results can be used to construct plant objects classifier.

Keywords: factor analysis, feature space, spectral coefficients of brightness, feature, classifier.

REFERENCES

1. Zagoruyko N. G. Metody raspoznavaniya i ih primeneniye. Moscow, Sov. Radio, 1972, 270 p.
2. Duda R., Hart P. Raspoznavanie obrazov i analiz stsen. Moscow, Mir, 1976, 326 p.
3. Harman G. Sovremenniy faktorniy analiz. Moscow, Statistica. 1972, 488 p.
4. Byuyul A., Cefel P. SPSS: Iskustvo obrabotki informaci. Analiz statisticheskih dannuh i vosstanovlenie skrutuh zakonomernostey. SPb, OOO «DiaSoftYUP», 2002, 603 p.
5. Kim Dzh.-O., Muller Ch. U. Factorniy analiz: statisticheskir metodu i prakticheskie, *Faktorniy, diskriminantniy i klasterniy analiz: sbornik rabot*; per. s angl.; Pod. red. I. S. Enyukova. Moscow, Finansu i ststistica, 1989, 215 p.
6. Aivazyan S. A., Buhshtaber V. M., Yenyukov I. S., Meshalkin L. D. Prikladnaya statistika. Klassifikatsiya i snizheniye razmernosti. Spravochnoe izdaniye. Moscow, Finansy I statistika, 1989, 607 p.
7. Kuper K. Individualnie razlichiya. Moscow, Aspekt Press, 2000, 527 p.
8. Borovikov V. STATISTICA: iskusstvo analiza dannuh na komputere. Dlya professionalov, SPb: Piter, 2001, 656 p.
9. Mitina O. V., Mihaylovskaya I. B. Faktorniy analiz dlya psihologov. Moscow. Uchebno-metodicheskiy kolektor. Psihologiya, 2001, 169 p.
10. Krasnopoyasovskiy A. S. Faktorniy klasifikatsiynui analiz za metodom funnktsionalno-statustuchnuh vuprobuvan, *Radioelektronni ta komputerni systemu*, 2004, No.4, pp. 46–50.

УДК 681.3.012:621.1

Швачич Г. Г.[1], Соболенко О. В.[2]

[1]Д-р техн. наук, професор, завідувач кафедри прикладної математики та обчислювальної техніки Дніпропетровської металургійної академії України, Дніпропетровськ, Україна
[2]Канд. техн. наук, доцент, доцент кафедри прикладної математики та обчислювальної техніки Дніпропетровської металургійної академії України, Дніпропетровськ, Україна

МОДЕЛЮВАННЯ ШВИДКІСНИХ РЕЖИМІВ ОБРОБКИ МЕТАЛУ НА ОСНОВІ ВИКОРИСТАННЯ ВИСОКОПРОДУКТИВНИХ БАГАТОПРОЦЕСОРНИХ ОБЧИСЛЮВАЛЬНИХ СИСТЕМ

Розглянуті особливості розробки і використання багатопроцесорної обчислювальної системи з її математичним і програмним забезпеченням для моделювання режимів термічної обробки металевих заготівок. Мета роботи полягає в розробці моделі для термічної обробки довгомірного сталевого виробу, яка може бути використана для рекристалізації та сфероїдизівного відпалювання каліброваної сталі. Запропоновано застосування сучасних багатопроцесорних обчислювальних комп'ютерних технологій для збільшенні швидкодії та продуктивності обчислень, що дає змогу ефективно керувати технологічними процесами. За допомогою спеціального програмного забезпечення багатопроцесорна система здатна задавати й контролювати необхідні температурні режими на всій площині перерізу зразка при нагріванні й витримці металу, а при необхідності може контролювати тепловий режим обробки сталі в інтервалі температур відпалювання. Багатопроцесорна обчислювальна система із спеціальним програмним забезпеченням містить математичні моделі у вигляді рівняння теплопровідності. Такі рівняння розв'язуються із застосуванням методів розщеплення. Завдяки цьому підходу розв'язок двохвимірювального рівняння зводиться до послідовності інтегрування одновимірних рівнянь простішої структури. Застосування числово-аналітичного методу забезпечує використання економічних і стійких алгоритмів розв'язування задач даного типу. Проведено експерименти з дослідження властивостей сталевої заготівки. Результати експериментів дозволяють рекомендувати запропонований підхід до моделювання швидкісних режимів обробки метала для розробки нових технологічних процесів.

Ключові слова: математична модель, багатопроцесорна обчислювальна система, інформаційний двоспрямований інтерфейс, контроль температурного режиму металу.

НОМЕНКЛАТУРА

БПОС – багатопроцесорна обчислювальна система;
БК – блок керування;
ВМ ПМ – виконавчий механізм протяжного механізму;
ВМ БІВТ – виконавчий механізм блока ізотермічної витримки температури;
ВМ БП – виконавчий механізм блока підстуджування;
ВМ БН – виконавчий механізм блока нагрівання;
ОЗТ – обернена задача теплопровідності;
ХОШ – холодне об'ємне штампування;
ТО – термічна обробка;
Ac_1 – температура фазового перетворення металу;
$T_н$ – температура нагрівання поверхні зразка;
$T_к$ – контрольована температура зразка;
$T(r, t)$ – температура зразка на його площі поперечного перерізу;
W – питома потужність;
τ – критерій Фур'є;
ϑ – дійсне значення шуканого кореня.

ВСТУП

На сьогодні у світі спостерігається стрімке зростання кількості БПОС та їх сумарної продуктивності. Це викликано тим, що такі системи стали загальнодоступними і дешевими апаратними платформами для високопродуктивних обчислень. При цьому різко зріс інтерес до проблематики обчислювальних мереж і широко поширюється розуміння того, що впровадження таких мереж матиме величезний вплив на розвиток людського суспільства, порівняний із впливом на нього появи на початку століття єдиних електричних мереж. У зв'язку з цим, розглядаючи проблеми освоєння багатопроцесорних систем, слід брати до уваги і те, що вони є першою сходинкою у створенні таких обчислювальних мереж.

Крім того, сьогодні практика висуває перед ученими-прикладниками різного роду проблеми, повне вирішення яких в більшості випадків можливе лише за рахунок застосування багатопроцесорних обчислювальних комплексів. Так, наприклад, у металургійному виробництві відбувається багато найрізноманітніших і взаємопов'язаних процесів. У першу чергу, це технології виплавки й розливання залізовуглецевих сплавів, нагрівання, прокатки й термічної обробки металопродукції та ін. Виробнича практика свідчить, що ні інтенсифікація процесів металургійного виробництва, ні конструктивне вдосконалення різноманітного металургійного устаткування неможливі без вивчення й аналізу явищ тепло- та масообміну. В той же час, розв'язування зазначених задач за допомогою відомих стандартних підходів являє собою складну проблему, подолання якої можливе тільки за рахунок застосування сучасних багатопроцесорних обчислювальних комп'ютерних технологій. При цьому одна з основних особливостей застосування таких технологій полягає у збільшенні швидкодії та продуктивності обчислень. Висока продуктивність обчислень дозволяє розв'язувати багатовимірні задачі, а також задачі, які вимагають великої кількості процесорного часу. Швидкодія дає змогу або ефективно керувати технологічними процесами, або взагалі створити передумови для розробки нових перспективних технологічних процесів.

DOI 10.15588/1607-3274-2015-2-3

У зв'язку з цим розробка й використання багатопроцесорних обчислювальних комплексів з їх математичним та програмним забезпеченням є актуальною проблемою, що дозволяє значно скоротити кількість експериментальних досліджень і час, потрібний на їх проведення, а це дозволяє одержати необхідну інформацію для створення та впровадження різних технологічних нововведень.

1 ПОСТАНОВКА ПРОБЛЕМИ

В даній роботі розглядається проблема впровадження нових технологічних процесів термічної обробки металу. Для цього необхідно створити модель ТО металу, яку використовують при виготовленні високоміцних кріпильних виробів методом ХОШ без завершальної термічної обробки. Така модель має на меті поліпшити технологічні властивості металопрокату за рахунок забезпечення високої дисперсності й однорідності структури зразка на всій площині його перерізу. До того ж технологічний процес термічної обробки сталі повинен набувати таких переваг, як висока продуктивність, знижене енергоспоживання, поліпшення експлуатаційних характеристик. Цього можна досягти завдяки застосуванню багатопроцесорної обчислювальної системи, виконаної у вигляді окремого модуля. За допомогою спеціального програмного забезпечення багатопроцесорна система здатна задавати й контролювати необхідні температурні режими на всій площині перерізу зразка при нагріванні й витримці металу, а при необхідності може контролювати тепловий режим обробки сталі в інтервалі температур відпалювання.

Використання багатопроцесорної обчислювальної системи з її програмним забезпеченням дозволить на основі математичної моделі процесу нагрівання зразка вже у виробничих умовах контролювати нагрівання дроту до моменту його переходу в аустенітний стан та настання температури фазової перекристалізації на всій площині перерізу довгомірного сталевого виробу, а потім здійснювати контроль необхідного режиму ізотермічної витримки в інтервалі температур відпалювання теж на всій площині перерізу зразка.

Застосування установки, що забезпечує реалізацію режиму сфероїдизівного відпалювання, зумовлює рівномірний розподіл глобул цементиту у феритній матриці, а це створює необхідні механічні властивості металу, потрібні для подальшого виконання холодної деформації. Разом з тим багатопроцесорна обчислювальна система із спеціальним програмним забезпеченням як єдина база повинна включати математичні моделі у вигляді рівняння теплопровідності, тобто

$$\frac{\partial T}{\partial \tau}=\frac{\partial^2 T}{\partial z^2}+\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\cdot\frac{\partial T}{\partial r}+W, \qquad (1)$$

при цьому критерій Фур'є $\tau=\frac{at}{R^2}$, якщо $\tau>0$; W – питома потужність джерела тепла, Вт/м^2.

Крайові умови цієї задачі мають такий вигляд:

$$T(0,r,z)=f(r,\tau);$$
$$T(\tau,1,z)=\text{var};$$
$$\frac{\partial T(\tau,0,z)}{\partial r}=0;$$
$$T(\tau,0,z)\neq 0.$$

Рівняння (1) необхідно розв'язувати із застосуванням методів розщеплення, суть яких полягає в редукції складного оператора (1) до простих. Завдяки цьому підходу необхідно звести розв'язок даного рівняння до послідовності інтегрування одновимірних рівнянь простішої структури. Застосування числово-аналітичного методу повинно забезпечити використання економічних і стійких алгоритмів розв'язування задач даного типу.

Отже, основна мета даної роботи полягає в розробці моделі ТО довгомірного сталевого виробу, яка може бути використана для рекристалізації та сфероїдизівного відпалювання каліброваної сталі на основі використання багатопроцесорних обчислювальних комплексів. В основу моделі було вирішено покласти спосіб ТО заготовки з низько- й середньовуглецевих сталей, призначених для холодної висадки [1].

2 ОГЛЯД ЛІТЕРАТУРИ

Традиційна технологія сфероїдизівного відпалювання сталі передбачає використання садчикових печей (ковпакових або шахтового типу). Недоліки традиційних способів підготовки заготовок до холодної висадки з подробицями викладені в відомій роботі Долженкова І. Є. [2, 3] і глибоко проаналізовані в роботі [4].

Альтернативою способу ТО з нагріванням виробів у печі служить електротермічний спосіб, який характеризується високою швидкістю нагрівання унаслідок впливу явищ електромагнітної індукції (індукційне нагрівання) або електроопору (електроконтактне нагрівання) [5]. Запровадження індукційного нагрівання у технологічній лінії для ТО дроту вже відоме у виробничній практиці [6, 7]. Для реалізації такої технології розроблено установку для виготовлення високоміцних кріпильних виробів без завершального термозміцнення. Але в даному випадку, під час проведення ТО заготовки не здійснюється контроль температурних режимів нагрівання, витримки й охолодження у зв'язку з відсутністю засобів виміру і контролю температури металу.

Іншим підходом до реалізації електротермічного способу обробки каліброваної сталі є установка [8], в якій передбачена термокамера і терморегульовальний екран. Але процес ТО характеризується значною тривалістю режиму відпалювання, тому що ізотермічна витримка й створення необхідного режиму охолодження проводиться у термокамері і потребує довгий час. За даними авторів тривалість відпалювання дорівнює від 30 до 90 хвилин, що не дозволяє синхронізувати замкнутий цикл виготовлення трипільних виробів.

3 МАТЕРІАЛИ ТА МЕТОДИ

Щоб вирішити окреслені вище проблеми, було розроблено установку для термічної обробки довгомірного сталевого виробу [4] з застосуванням БПОС [9]. Використання БПОС з її програмним забезпеченням дозволяє на основі математичної моделі процесу нагрівання зразка вже у виробничих умовах контролювати нагрівання дроту до переходу в аустенітну область до температури фазової перекристалізації на всій площині перерізу довгомірного сталевого виробу, а потім, розв'язавши ОЗТ, здійснювати контроль необхідного режиму ізотермічної витримки в інтервалі температур відпалювання на всій площині перерізу зразка.

На рис. 1 зображено проектну схему установки для ТО довгомірного сталевого виробу, де 1 – розмотувальний пристрій; 2 – правильно-тяговий пристрій, обладнаний виконавчим механізмом 3; 4 – індуктор нагрівального пристрою; 5 – генератор з виконавчим механізмом; 6 – пірометр; 7 – камера ізотермічної витримки з виконавчим механізмом 8; 9 – пірометр; 10 – камера регламентованого підстуджування дроту з виконавчим механізмом 11 для регулювання подачі водоповітряної суміші; 12 – пірометр; 13 – камера ізотермічної витримки з виконавчим механізмом 14; 15 – пірометр; 16 – пристрій для інтенсивної сфероїдизації з виконавчим механізмом 17; 18 – пірометр; 19 – пристрій подачі дроту на подальший технологічний цикл; 20 – інформаційний двоспрямований інтерфейс збору даних з пристроїв 3, 7, 9, 12, 15, 18, приєднаний до блока керування 21 і до виконавчих механізмів (3, 5, 8, 11, 14, 17) відповідних пристроїв; 22 – інформаційний двоспрямований інтерфейс зв'язку блока керування БПОС 23.

Установка для ТО довгомірного сталевого виробу працює в описаній нижче послідовності. З розмотувального пристрою 1 через правильно-тяговий пристрій 2 дріт подається в індуктор нагрівального пристрою 4, де нагрівається до переходу в аустенітну стадію, набуваючи температури фазової перекристалізації.

Температура нагрівання контролюється пірометром 7. Підтримання температурного режиму здійснюється за допомогою блока керування 21 і БПОС 23. Сигнал з пірометра 6 через інформаційний двоспрямований інтерфейс 20 надходить у блок керування 21, а потім через інформаційний двоспрямований інтерфейс 22 – у БПОС 23, де згідно з результатом розв'язку математичної моделі прямої задачі теплопровідності, відбувається регулювання потужності генератора 5.

Потім розігрітий дріт потрапляє в камеру ізотермічної витримки 7, температура в якій регулюється виконавчим механізмом 8. При цьому сигнал з пірометра 9 через інформаційний двоспрямований інтерфейс 20 надходить у блок керування 21, а далі через інформаційний двоспрямований інтерфейс 22 у БПОС 23, де відповідно до результатів розв'язку математичної моделі ізотермічної витримки, здійснюється регулювання температурного режиму за допомогою виконавчого меха-нізму 8.

Далі дріт переходить у камеру регламентованого підстуджування 10. Залежно від режиму ТО, марки сталі й діаметра дроту задається необхідна швидкість підстуджування в інтервалі температур від 750 до 700 °С.

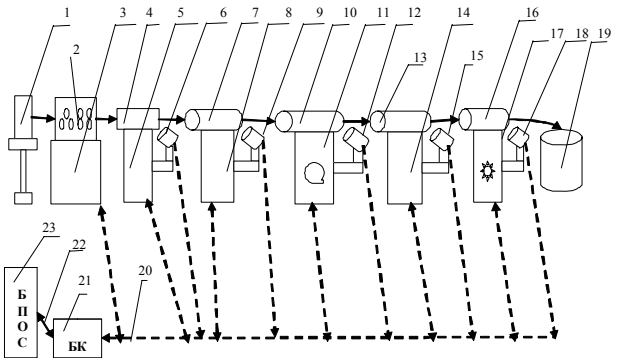


Рисунок 1 – Схема установки термічної обробки сталевого виробу

У цьому температурному інтервалі аустеніт безперервно втрачає вуглець і, досягши необхідної концентрації останнього, зазнає поліморфних перетворень, перетворившись на ферит, таким чином відбувається розпад аустеніту за анормальним механізмом. Температура підстуджування контролюється блоком керування 21 і БПОС 23, де з урахуванням температури, зафіксованої пірометром 12, і відповідно до результату розв'язку математичної моделі ОЗТ, виконавчий механізм 11 подає водоповітряну суміш, збільшуючи або зменшуючи її кількість, залежно від заданого температурного режиму підстуджування.

Потім дріт потрапляє в камеру ізотермічної витримки 13, температура в якій регулюється виконавчим механізмом 14. Під час ізотермічної витримки завершується утворення квазіевтектоїду (перліту), що включає зони з підвищеною концентрацією вуглецю і з готовими цементитними частинками. Сигнал із пірометра 15 через інформаційний двоспрямований інтерфейс 20 надходить у блок автоматичного керування 21, а далі через інформаційний двоспрямований інтерфейс 22 – у БПОС 23, де на підставі результатів розв'язку математичної моделі ізотермічної витримки через блок керування 21 здійснюється регулювання температурного режиму в камері 13.

Після цього дріт подається в пристрій інтенсивної сфероїдизації 17, де здійснюється зміна температурного режиму із швидкістю від 15 до 20 °С/хв з досягненням підкритичної температури Ac_1 (у цій температурній зоні відбувається інтенсивна сфероїдизація цементитних частинок). Сигнал із пірометра 18 через інформаційний двоспрямований інтерфейс 20 надходить у блок автоматичного керування 21, а потім через інформаційний двоспрямований інтерфейс 22 – у БПОС 23, де з урахуванням результатів розв'язку математичної моделі прямої задачі теплопровідності відбувається регулювання температурного режиму в пристрої інтенсивної сфероїдизації 16. Далі через пристрій подачі 19 дріт переходить до наступного технологічного циклу його обробки.

На рис. 2 подано блок-схему контурів системи керування установкою термічної обробки довгомірного сталевого виробу. Така система керування має у своєму розпорядженні блоки, які дозволяють отримати інформацію про поточні параметри керованих процесів. Особливість її полягає в тому, що на кожному з п'яти етапів технологічної обробки зразка розв'язується двовимірна задача теплопровідності. При цьому програмні засоби БПОС дозволяють контролювати температурні режими, як на всій площині перерізу зразка, так і по його довжині. Контроль таких температурних режимів здійснюється в центрі площини перерізу зразка.

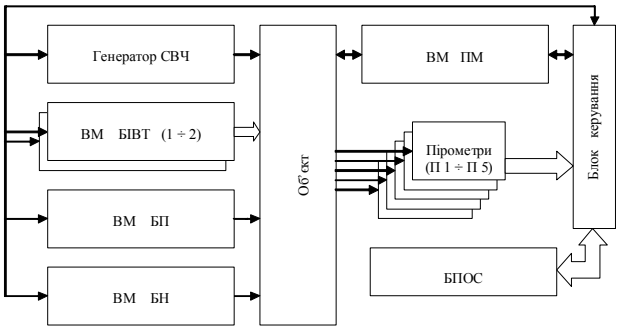


Рисунок 2 – Блок-схема контурів системи керування установкою термічної обробки довгомірного сталевого виробу

БПОС із спеціальним програмним забезпеченням як єдина база містить математичні моделі у вигляді рівняння теплопровідності (1). При цьому за координатою z граничні умови, залежно від особливостей розв'язуваної задачі, можуть бути першого, другого або третього роду. Розв'язують задачу (1) із застосуванням методів розщеплювання, суть яких полягає в редукції складного оператора (1) до простих. Цей підхід дозволяє проінтегрувати дане рівняння як послідовність інтегрування одновимірних рівнянь простішої структури. З огляду на суттєву складність математичної моделі (1), великого значення набуває розробка економічних алгоритмів для розрахунку ефектів керування функціями запропонованої установки. Процес створення зазначених алгоритмів висвітлюється в роботах [10, 11].

Зауважимо, що тут саме задача керування (як і задача синтезу) в її точній постановці відноситься до класу обернених, оскільки вона передбачає визначення керуючих функціональних параметрів на основі заздалегідь заданого, необхідного результату (обернена задача керування).

Алгоритмом розв'язування обернених задач слугує метод «вилки» з попереднім визначенням деякого початкового відрізка. Розв'язок задачі реалізується в два етапи.

На першому етапі реалізується відокремлення мінімуму нев'язки, на другому – визначається мінімум шуканої функції керування з відокремленого інтервалу. Зазначена процедура реалізується стандартно. Інакше кажучи, якщо ϑ являє собою деяке дійсне значення шуканого кореня, тобто, коли $a \le \vartheta \le b$, а $f(\vartheta)=0$, то можна обчислити число w таким, що задовольняє умові: $a \le \vartheta \le b$ та $|\vartheta - w| < \varepsilon$, тобто меншим від будь-якого наперед заданого малого числа ε. Подібна схема включена до складу математичного апарату керування БПОС. На всіх циклах ТО металу використовується математична модель (1), тому результати моделювання будуть стосуватись циклу первинного нагрівання металу.

Проблема моделювання полягає в тому, що для забезпечення необхідної точності й стійкості обчислень доводиться брати розрахункову сітку з чималою кількістю вузлів і виконувати безліч ітерацій. Унаслідок цього число арифметичних операцій, що необхідні для розрахунку температурних полів, перебуває в межах $10^7 \div 10^8$ вузлів, а коли крок за часовою ознакою становить 10^{-2} с, то загальна кількість вузлів для обчислень може досягти 10^{20} і більше. Однопроцесорні обчислювальні системи не можуть впоратись з таким навантаженням під час моделювання в реальному масштабі часу, тому найбільш виправданим буде застосування багатопроцесорних систем, що й було здійснено з метою удосконалення технології ТО довгомірного сталевого виробу.

4 ЕКСПЕРИМЕНТИ

Для випробування функцій запропонованої установки було проведено кілька експериментів, коли дріт діаметром 20 мм із сталі 20Г2Р піддавався ТО. Розглянемо один з характерних дослідів.

За початкову було взято феритно-перлітну структуру заготівки. Процес ТО матеріалу здійснювався шляхом нагрівання заготівки в межах міжкритичної зони температур. Для заданого матеріалу встановлено такі значення критичних точок: Ac_1 = 725 °C; Ac_3 = 795 °C.

Нагрівання відбувалось до такого значення: Ac_1 + (10–30 °C). Протягом наступного етапу обробки матеріалу було реалізовано процес ізотермічної витримки протягом 45 с. Далі тривало охолодження виробу зі швидкістю 20–30 °C/с до температури 620 °C з подальшою ізотермічною витримкою протягом 45 с. Нарешті, на останньому етапі обробки матеріалу зразок нагрівали зі швидкістю 15–25 °C/с до підкритичних температур. Гра-фічну інтерпретацію режиму ТО металу відображено на рис. 3.

Механічні характеристики визначалися за результатами замірів твердості зразків. Випробування на розтягування проведено на машині FU10000ez. Дослідження мікроструктури металу проводилося на світловому металографічному мікроскопі Neophot-2 з використанням структурного аналізатора «Epiquant», додатково обладнаного пристроєм «Anasonic» для цифрової реєстрації зображення.

5 РЕЗУЛЬТАТИ

За результатами експериментів одержані криві розподілу температури зразка на площині його перерізу (рис. 4), де $T_Н$ – температура нагрівання поверхні зразка,

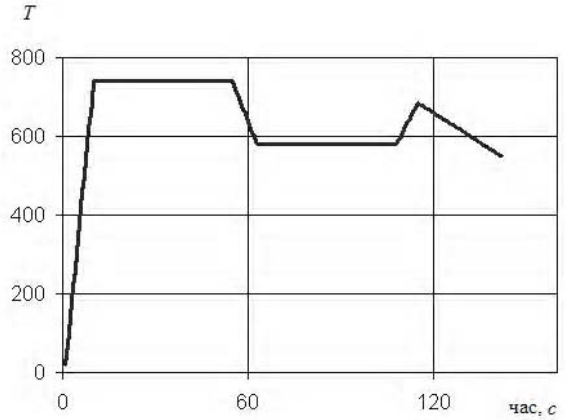


Рисунок 3 – Графік режиму термічної обробки сталі 20Г2Р

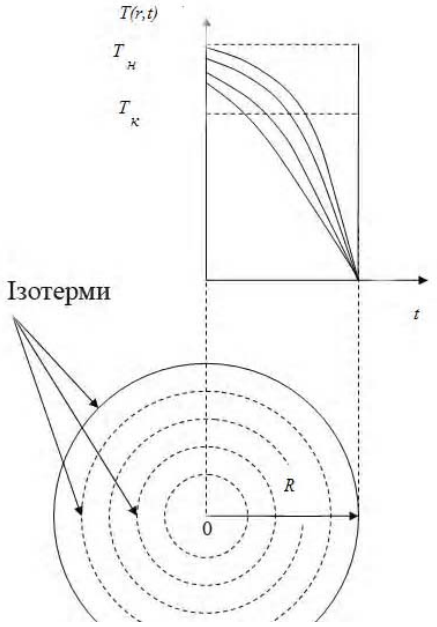


Рисунок 4 – Криві розподілу температури зразка на площині його перерізу

$T_к$ – контрольована засобами БПОС температура фазового перетворення металу (Ac_1) на площині перерізу зразка. Моделювання таких температурних полів здійснюється з урахуванням зміни теплофізичних властивостей матеріалу під час його нагрівання.

Криві розподілу температури зразка по довжині в процесі його нагрівання зображено на рис. 5, де цифрою 1 позначено температуру поверхні ($T_Н$), а цифрою 2 – температуру в центрі площини перерізу. Тут зона I відображає процес нагрівання зразка до заданої температури на його поверхні, а зона II демонструє вихід на заданий температурний режим у центрі площини перерізу зразка.

На рис. 6а,б зображено мікроструктуру зразків до і після сфероїдизації, при цьому твердість зразків після ТО набула значень 150–169 НВ.

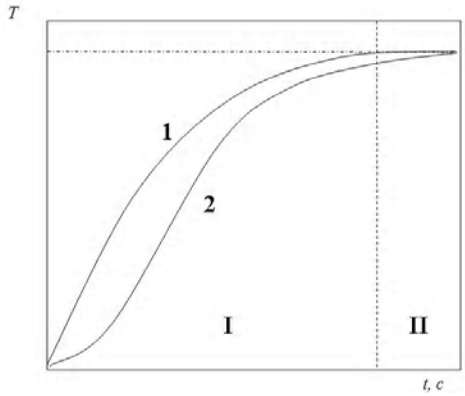
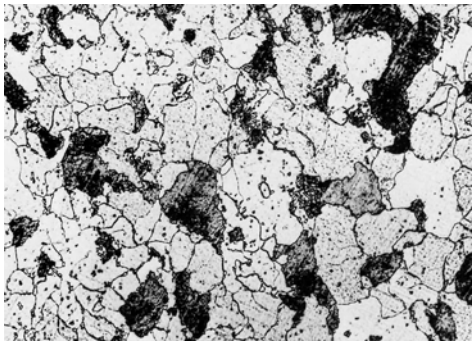
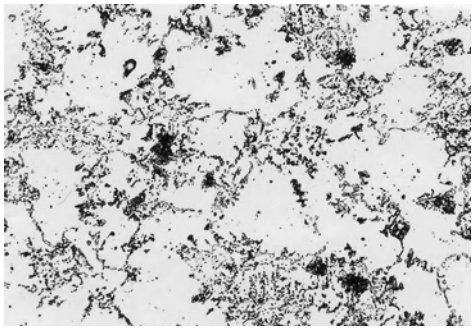


Рисунок 5 – Графіки розподілу температури зразка по його довжині в процесі нагрівання



а



б

Рисунок 6 – Мікроструктура сталі 20Г2Р:
а – початкова феритно-перлітна структура, х 500; б – структура після відпалювання – перліт зернистий (бал 5), х500

6 ОБГОВОРЕННЯ

Виконана сфероїдизація карбідної фази металу в умовах відповідних режимів ТО заготівок забезпечує надання матеріалу структури зернистого перліту. Причому швидкісна сфероїдизація зумовлює більш рівномірний розподіл глобул цементиту у феритній матриці (рис. 6б). Зразки із сталі майже однакової твердості після ТО набули дрібнодисперсної структури, що забезпечує більш високий рівень пластичності металу. Унаслідок швидкого нагрівання зразка й неповної аустенітизації сталі відбуваються певні зміни в морфології карбідної фази від пластинчастої до дрібнодисперсної глобулярної.

Технічний результат, що досягається при запровадженні запропонованої системи, полягає в тому, що забезпечується висока дисперсність й однорідність структури зразка на всій площині його перерізу, при цьому технологічний процес ТО сталі характеризується високою продуктивністю, малим енергоспоживанням, поліпшеними експлуатаційними характеристиками. Застосування установки для реалізації режиму сфероїдизівного відпалювання зумовлює рівномірний розподіл глобул цементиту у феритній матриці, а значить забезпечує необхідні механічні властивості металу для його подальшої холодної деформації.

Відзначимо, що розв’язування задач, визначених в даній роботі, зазвичай, відбувається на основі застосування апарату різницевих рівнянь, який передбачає обов’язкову заміну похідних різницевими співвідношеннями.

Виконані в даній роботі дослідження показують, що методи розв’язування задач даного класу мають бути не тільки різноманітними, але й поєднувати оцінювання кількісних показників із можливостями якісного аналізу. На сьогодні намітилися певні тенденції в розробці числово-аналітичних методів із складною логічною структурою, але вони мають порівняно з кусково-різницевими методами вищий порядок точності й можливість побудови алгоритмів, адаптованих за порядками апроксимації. З обчислювальної точки зору такий підхід відрізняється певною громіздкістю, але він дає своєрідний еталон для порівняння з іншими практичними методами. Але, зважаючи на переваги проведення обчислювального експерименту засобами багатопроцесорної системи, можна стверджувати, що обставина, яка стримувала розвиток числово-аналітичного підходу, на сьогодні втрачає свою актуальність. У зв’язку з цим, для розв’язування моделі (1), у даній роботі набула подальшого розвитку ідея розробки схем підвищеного порядку точності на основі числово-аналітичного підходу до розв’язування багатьох досліджуваних задач.

ВИСНОВКИ

Удосконалення наявних й створення нових технологічних процесів ТО металу вимагають значних витрат, пов’язаних із проведенням великої кількості натурних експериментів на лабораторному, дослідно-промисловому устаткуванні, а також у виробничих умовах. Скорочення кількості експериментальних досліджень та часу на їх проведення з одержанням необхідної інформації для побудови й упровадження технологічних розробок можна здійснити шляхом застосування багатопроцесорних об-

числювальних комплексів. У статті розв'язана актуальна задача контролю температурних режимів процесу рекристалізації та сфероїдизівного відпалювання каліброваної сталі шляхом застосування багатопроцесорної системи, що дозволяє узгоджувати часові інтервали технологічного процесу відпалювання.

Наукова новизна проведених досліджень полягає в тому, що уперше на основі багатопроцесорної обчислювальної системи створено модель швидкісної ТО довгомірного сталевого виробу в реальному часі з метою рекристалізації та сфероїдизівного відпалювання каліброваної сталі й виготовлення високоміцних кріпильних виробів методом ХОШ без завершальної ТО. Запропонований підхід дає можливість контролювати технологічні параметри в режимах ТО металу, зокрема температуру в центрі перерізу металевого виробу, що забезпечує надання матеріалу необхідних властивостей, причому на всій площині перерізу і по довжині зразка. Цього вдалося досягти за рахунок застосування багатопроцесорної обчислювальної системи, що має вигляд окремого модуля, а за допомогою спеціального програмного забезпечення вона здатна задавати й контролювати необхідні температурні режими на всій площині перерізу зразка при нагріванні й витримці металу. Порівняно з традиційними підходами, було реалізовано можливість поліпшити технологічні властивості металопрокату за рахунок забезпечення високої дисперсності й однорідності структури зразка на всій площині його перерізу.

Практична цінність отриманих результатів полягає в тому, що вдалося удосконалити технологічний процес ТО металу за рахунок використання відповідних математичних моделей та комплексу програм. Застосування математичних моделей, які обробляються на багатопроцесорній обчислювальній системі дозволяє контролювати температурне поле металу в процесі його нагрівання, витримки та охолодження і забезпечує, тим самим, швидку адаптацію виробництва металопродукції до вимог споживача.

Втілення розробленого підходу для ТО металу на основі запровадження багатопроцесорної обчислювальної системи створює проблему узгодження можливостей процесорів і мережного інтерфейсу багатопроцесорної системи. Отже, перспективними подальшими дослідженнями є шляхи вирішення зазначеної проблеми на прикладі застосування багатопроцесорних систем, що складаються із різних типів процесорів. За таких умов необхідно вивести аналітичні співвідношення для встановлення оптимального числа вузлів багатопроцесорної системи з урахуванням обчислювальних можливостей певних процесорів.

ПОДЯКИ

Роботу виконано відповідно з тематичними планами наукових досліджень Національної металургійної академії України. Дослідження проведені в рамках бюджетних і госпдоговірних робіт:

– «Розробка теоретичних основ фізико-технічних процесів обробки маловуглицевих економно легованих сталей з бейнітною структурою», номер державної реєстрації № ДР 0106U002210;

– «Математичне моделювання режимів термічної обробки при швидкісному циклічному нагріві і охолодженні довгомірного виробу», номер державної реєстрації № ДР 0107U002839.

СПИСОК ЛІТЕРАТУРИ

1. Пат. України № 36892, МПК С 21D 1/26 C21D 1/78. Спосіб термічної обробки прокату з низько- і сердньовуглецевих сталей для холодного висадження / В. П. Ковпак, А. М. Лещенко, М. О. Соболенко, Г. В. Кокашинська ; власник : Національна металургійна академія України. – № 200807153; заявл. 23.05.2008; опул. 10.11.2008, Бюл. № 21.
2. Долженков И. Е. Термическая и деформационно-термическая обработка металлопроката / И. Е. Долженков // Теория и практика металлургии. – 2002. – № 3. – С. 30 –36.
3. Долженков И. Е. Влияние пластической деформации и других предобработок на сфероидизацию карбидов в сталях / И. Е. Долженков // Теория и практика металлургии. – 2007. – № 1. – С. 66 –68.
4. Пат. 61944 Україна, МПК C21D 1/26 (2006.01), G06F 15/16 (2006.01). Установка для термічної обробки довгомірного сталевого виробу / В. П. Іващенко, Є. О. Башков, Г. Г. Швачич, М. О. Ткач, М. О. Соболенко ; власники : Національна металургійна академія України, Донецький національний технічний університет. – № u 201014225; заявл. 29.11.2010; опубл. 10.08.2011, Бюл. № 15.
5. Электротермическая обработка и теплое волочение стали / [Г. А. Хасин, А. И. Дианов, Т. Н. Попова, Л. П. Кукарцева]. – М. : Металлургия, 1984. – 152 с.
6. Бобылев М. В. Подготовка структуры при электротермообработке сталей, применяемых для высадки высокопрочных крепежных изделий / М. В. Бобылев, В. Е. Гринберг, Д. М. Закиров, Ю. А. Лавриненко // Сталь. – 1996. – № 11. – С. 54–60.
7. Бобылев М. В. Оптимизация режимов отжига с индукционным нагревом сталей 20Г2Р и 38ХГНМ / М. В. Бобылев, Д. М. Закиров, Ю. А. Лавриненко // Сталь. – 1999. – № 4. – С. 67–70.
8. Патент РФ 2137847, кл. С 21 D 1/32, С 21 D 9/60, С 21 D 11/00. Установка для термообработки калиброванной стали / Д. М. Закиров, М. В. Бобылев, Ю. А. Лавриненко, Л. П. Лебедев, В. И. Сюльдин ; Патентообладатель: Открытое акционерное общество «Автонормаль». – № 98117255/02; заявл. 16.09.1998; опубл. 20.09.1999.
9. Пат. 57663 Україна, МПК G06F 15/16 (2011.01). Модуль високоефективної багатопроцесорної системи підвищеної готовності / В. П. Іващенко, Є. О. Башков , Г. Г. Швачич, М. О.Ткач ; власники : Національна металургійна академія України, Донецький національний технічний університет. – № u 2010 09341; заявл. 26.07.2010; опубл. 10.03.2011, Бюл. № 5.
10. Швачич Г. Г. Об алгебраическом подходе в концепции распределенного моделирования многомерных систем / Г. Г. Швачич // Теория и практика металлургии. – 2007. – № 6(61). – С. 73–78.
11. Швачич Г. Г. Математическое моделирование одного класса задач металлургической теплофизики на основе многопроцессорных параллельных вычислительных систем / Г. Г. Швачич // Математичне моделювання. – 2008. – № 1 (18). – С. 60–65.

Стаття надійшла до редакції 25.11.2014.
Після доробки 19.12.2014.

Швачич Г. Г.[1], Соболенко А. В.[2]

[1]Д-р техн. наук, профессор, заведующий кафедры прикладной математики и вычислительной техники Днепропетровской металлургической академии Украины, Днепропетровск, Украина

[2]Канд. техн. наук, доцент, доцент кафедры прикладной математики и вычислительной техники Днепропетровской металлургической академии Украины, Днепропетровск, Украина

МОДЕЛИРОВАНИЕ СКОРОСТНЫХ РЕЖИМОВ ОБРАБОТКИ МЕТАЛЛА НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ МНОГОПРОЦЕССОРНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

В работе рассмотрены особенности разработки и использования многопроцессорной вычислительной системы с ее математическим и программным обеспечением для моделирования режимов термической обработки стальных заготовок. Цель работы: разработка модели режима термической обработки длинномерного стального изделия, которая может быть использована для рекристаллизации и сфероидизирующего отжига калиброванной стали. Предложено использование современных многопроцессорных вычислительных компьютерных технологий для увеличения скорости и продуктивности вычислений, что обеспечивает эффективное управление технологическим процессом. При помощи специального программного обеспечения многопроцессорная система способна задавать и контролировать необходимые температурные режимы на всей плоскости поперечного сечения образца при нагреве и выдержке металла, а при необходимости может контролировать тепловой режим обработки стали в интервале температур отжига. Многопроцессорная вычислительная система со специальным программным обеспечением включает математические модели в виде уравнения теплопроводности. Такие уравнения решаются с применением методов расщепления. Благодаря этому подходу решение двумерного уравнения сводится к последовательности интегрирования одномерных уравнений более простой структуры. Применение численно-аналитического метода обеспечивает использование экономических и устойчивых алгоритмов решения задач данного типа. Проведены эксперименты с исследованием свойств стальной заготовки. Результаты экспериментов позволяют рекомендовать предлагаемый подход к созданию моделей скоростных режимов обработки металла для разработки новых технологических процессов.

Ключевые слова: математическая модель, многопроцессорная вычислительная система, информационный двунаправленный интерфейс, контроль температурного режима металла.

Shvachych G. G.[1], Sobolenko O. V.[2]

[1]Dr. Sc., Professor, Chief of the Department of Applied Mathematics and Computer Science of the National Metallurgical Academy of Ukraine, Dnepropetrovsk, Ukraine

[2]PhD, Associate Professor of the Department of Applied Mathematics and Computer Science of the National Metallurgical Academy of Ukraine, Dnepropetrovsk, Ukraine

SIMULATION OF SPEED SCHEDULES FOR METAL ON THE BASIS OF USING THE HIGH-PERFORMANCE MULTIPROCESSOR COMPUTER SYSTEMS

The paper deals the features of the development and use of the multiprocessor computing system with mathematical and software of the latter for simulation of heat treatment of the steel billets. Objective of the work: developing the model of heat treatment of a long steel product, which can be used for recrystallization and spheroidizig annealing of calibrated steel. The use of the up-to-date multiprocessor computing technologies had been proposed for increasing speed and productivity of computations, what maintains the effective control of the technological process. Through the special software the multiprocessor system is able to set and control necessary temperature conditions on all plane of cross-sectional of standard at heating and self-control of metal, and if necessary maybe began to control the thermal mode of treatment in the interval of temperatures of annealing. The multiprocessor computer system with the special software includes mathematical models as equation of heat conductivity. Such equations decide with the use of methods of breaking up. Due to this approach the decision of two-dimensional equation is taken to the sequence of integration of one-dimensional equations of more simple structure. Application of numeral-analytical method provides the use of economic and steady algorithms of decision of problems of this type.

Experiments had been carried out with studying the properties of a steel billet. Results of the experiments allow recommending the proposed approach to creating models of high-speed schedules of metal treatment for development of new technological processes.

Keywords: mathematical model, multiprocessor computing system, information bidirectional interface, control of the temperature schedule of metal.

REFERENCES

1. Kovpak V. P., Leshtshenko A. M., Sobolenko M. O., Kokashinska G. V. Pat. Ukrainy № 36892, MPK C 21D 1/26 C21D 1/78. Sposib termichnoi obrobky prokatu z nyzko- i seredniovugletsevykh stalej dlya kholodnogo vysadzhenya ; vlasnyk : Natsionalna metalurgijna akademiya Ukrainy, №200807153; zayavl. 23.05.2008; opubl.10.11.2008, Bul. №21.
2. Dolzhenkov I. Ye. Termicheskaya i deformatsionno-termicheskaya obrabotka metalloprokata, *Teoriya i praktika metallurgii*, 2002, No. 3, pp. 30–36.
3. Dolzhenkov I. Ye. Vliyanie plasticheskoy dephormaziy I drugich predobrabotok na spheroidizaziyu karbidov v stalajch, *Teoriya i praktika metallurgii*, 2007, No. 1, pp. 66–68.
4. Ivashtshenko V. P., Bashkov Ye. O., Shvachich G. G., Tkach M. O., Sobolenko M. O. Pat. 61944 Ukraina, MPK C21D 1/26 (2006.01), GO6F 15/16 (2006.01). Ustanovka dlya termichnoi obrobki dovgomirnogo stalevogo vyrobu ; vlasnyky, Natsionalna metalurgijna akademiya Ukrainy, Donetskyj natsionalnyj technichnyj universytet, № u 201014225; zayavl. 29.11.2010; opubl. 10.08.2011, Bul. №15.
5. Khasin G. A., Dianov A. I., Popova T. N., Kukartseva L. P. Elektrotermicheskaya obrabotka i typloye volocheniye stali. Moscow, Metallurgiya, 1984, 152 p.
6. Bobylev M. V., Grinberg V. Ye., Zakirov D. M., Lavrinenko Yu. A. Podgotovka stryktyry pri elektrotermoobrabotke stalej, primenyayemykh dlya vysadki vysokoprochnykh krepyozhnykh izdelij, *Stal*, 1996, No. 11, pp. 54–60.
7. Bobylev M. V., Zakirov D. M., Lavrinenko Yu. A. Optimizatsija rejimiv otjiga s induktsionim nagrevom staltej, primenyaemich dlya visadki visokoprochnich krepejnich itsdeliyi, *Stal,* 1999, No. 4, pp. 67–70.
8. Zakirov D. M., Bobylev M. V., Lavrinenko Yu. A., Lebedev L. P., Syuldin V. I. Patent RF 2137847, kl. C 21 D 1/32, C 21 D 9/60, C 21 D 11/00. Ustanovka dlya termoobrabotki kalibrovannoj stali; Patentoobladatel: Otkrytoye aktsionernoye obshtshestvo «Avtonormal», №98117255/02; zayavl. 16.09.1998; opybl. 20.09.1999.
9. Ivashtshenko V. P., Bashkov Ye. O., Shvachych G. G., Tkach M. O. Pat.57663 Ukraina, MPK G06F 15/16 (2011.01). Modul vysokoefektyvnoi bagatoprotsesornoi systemy pidvyshtshenoi gotovnosti ; vlasnyky: Natsionalna metalurgijna akademiya Ukrainy, Donetskyj natsionalnyj tekhnicnyj universytet.№ u 2010 09341; zayavl. 26.07.2010; opubl. 10.03.2011, Bul.№5.
10. Shachich G. G. Ob algebraicheskom podkhodye v kontseptsii raspredelyonnongo modelirovaniya mnogomernykh sistem, *Teoriya i praktika metallurgii,* 2007, No.6(61), pp. 73–78.
11. Shvachich G. G. Matematicheskoye modelirovaniye odnogo klassa zadach metallurgicheskoj teplofiziki na osnovye mnogoprotsesornykh parallelnykh vychislitelnykh sistem, *Matematychne modelyuvannya*, 2008, No. 1(18), pp. 60–65.

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

УДК 004.93

Олійник А. О.

Канд. техн. наук, доцент, доцент кафедри програмних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

СИНТЕЗ НЕЙРО-НЕЧІТКИХ МОДЕЛЕЙ ДЛЯ ПРОГНОЗУВАННЯ СТАНУ ЗДОРОВ'Я ХВОРОГО НА ОСНОВІ ПАРАЛЕЛЬНИХ ОБЧИСЛЕНЬ

Вирішено завдання розробки математичного забезпечення для автоматизації процесу індивідуального прогнозування стану здоров'я хворого гіпертонічною хворобою. Об'єктом дослідження є процес синтезу моделей для медичного діагностування. Предмет дослідження становлять методи синтезу нейро-нечітких моделей для медичного діагностування. Мета роботи: підвищити ефективність процесу синтезу нейро-нечітких мереж для побудови діагностичних моделей на основі навчальних вибірок. Запропоновано стохастичний метод синтезу нейро-нечітких моделей на основі паралельних обчислень, що ґрунтується на стохастичному підході при пошуку значень параметрів, що настроюються, і полягає в розподілі найбільш ресурсномістких етапів по вузлах паралельної обчислювальної системи. Запропонований метод дозволяє скоротити час настроювання параметрів (значень вагових коефіцієнтів і параметрів функцій належності нейроелементів) синтезованих нейромоделей. Розроблено програмне забезпечення, що реалізує запропонований метод. Проведено експерименти з дослідження властивостей запропонованого методу. Результати експериментів дозволяють рекомендувати запропонований метод для використання на практиці.

Ключові слова: множина рішень, нейронна мережа, ознака, паралельне програмування, прогнозування, синтез моделі.

НОМЕНКЛАТУРА

$D\left(\chi_k^{(i)}, C_j\right)$ – відстань між рішенням $\chi_k^{(i)}$ і центром j-го кластера C_j;

$Entr(\Delta p_{mn})$ – ентропія інтервалу Δp_{mn};

$g_{lk}^{(0)}$ – значення l-го елемента (параметра) в k-му рішенні;

$G_{k,norm}^{(i,j)}$ – нормоване значення цільової функції $G_k^{(i,j)}$ для множини $R^{(i,j)}$;

$Q\big|_{p_m \in \Delta p_{mn}}$ – кількість спостережень вибірки S, значення m-х ознак p_m яких потрапили в інтервал Δp_{mn};

M – кількість атрибутів;

N_g – кількість параметрів в рішенні $\chi_k^{(i)}$;

N_{pr} – кількість процесів, одночасно виконуваних в паралельній комп'ютерній системі;

$N_{Q,p_m \in \Delta p_{mn}}$ – кількість інтервалів розбиття (класів) вихідного параметра T фрагменту вибірки $S_{p_m \in \Delta p_{mn}}$, в якому значення m-х ознак p_m потрапили в інтервал Δp_{mn};

N_χ – кількість елементів множини $R^{(0)}$ (число випадково згенерованих рішень при ініціалізації);

p_{qm} – значення m-го атрибуту q-го об'єкта;

$P = \left(p_{qm}\right)_{QM}$ – множина значень характеристик екземплярів вибірки;

Q – кількість об'єктів;

$Q\big|_{T=q;S_{p_m \in \Delta p_{mn}}}$ – кількість спостережень у множині $S_{p_m \in \Delta p_{mn}}$, в яких значення вихідного параметру T дорівнює q ($T = q$);

$R^{(i,j)}$ – j-та підмножина i-ї множини $R^{(i)}$;

t_q – значення відгуку q-го об'єкта;

T – множина значень відгуку;

$\chi_k^{(0)} = \left\{g_{1k}^{(0)}, g_{2k}^{(0)}, \ldots, g_{N_g k}^{(0)}\right\}$ – k-те рішення у множині $R^{(0)}$.

ВСТУП

Артеріальна гіпертензія є широко поширеним захворюванням, яке може загрожувати життю та здоров'ю пацієнта [1]. На характер перебігу артеріальної гіпертензії

DOI 10.15588/1607-3274-2015-2-4

впливають різні фактори (погодні і кліматичні умови, супутні захворювання, а також стан здоров'я в попередні моменти часу) [2].

Для запобігання істотних стрибків тиску пацієнта, які можуть викликати погіршення стану пацієнта, а можливо і привести до до летального наслідку, необхідно своєчасно прогнозувати розвиток гіпертензії в короткостроковому періоді часу (на найближчі половину доби або добу). Це дозволить своєчасно реалізувати профілактичні заходи, пов'язані з прийомом необхідних препаратів для запобігання очікуваних негативних наслідків.

Для прогнозування стану здоров'я хворого гіпертонічною хворобою потрібно мати модель, яка буде унікальною для кожного конкретного пацієнта. Побудова такої моделі потребує обробки великого масиву спостережень, розподілених у часі.

Перспективним засобом побудови моделей для індивідуального прогнозування є методи обчислювального інтелекту (нейро-нечіткі технології та методи стохастичною оптимізації) [3–7]. Штучні нейро-нечіткі мережі характеризуються здатністю до навчання за прецедентами, високими узагальнюючими властивостями, що обумовлює їх здатність отримувати знання з даних. Стохастичні методи, на відміну від широко використовуваних градієнтних методів оптимізації, не вимагають обчислень значень похідних цільової функції, дозволяють виходити з локальних екстремумів і синтезувати моделі на основі заданих вибірок даних.

Для врахування динаміки стану пацієнта необхідно з плином часу синтезувати нові моделі (періодично перебудовувати вже існуючі моделі) на основі постійно зростаючих масивів спостережень. Це в свою чергу вимагає використання великих обчислювальних ресурсів. Для подолання цих обмежень пропонується розпаралелити процес синтезу діагностичної моделі.

Метою даної роботи є розробка паралельного методу синтезу прогнозуючих моделей на основі нейро-нечітких мереж і стохастичного пошуку.

1 ПОСТАНОВА ЗАДАЧІ

Нехай є вибірка даних $S = < P, T >$. Тоді задача синтезу нейро-нечіткої моделі *NFN* полягає в ідентифікації її параметрів таким чином, щоб забезпечувалося прийнятне значення заданого критерію якості *G*. Як цільовий критерій *G* при навчанні нейро-нечітких моделей можуть бути використані, наприклад, похибка розпізнання або середньоквадратична помилка.

Вихідна вибірка даних про стан здоров'я хворого гіпертонічною хворобою була отримана в м. Запоріжжя (Україна). Вибірка $S = < P, T >$ містила спостереження за період з 2004 р. по 2014 р., де кожен екземпляр являв собою набір даних, що характеризують стан пацієнта в певну частину дня.

В якості об'єктивних клініко-лабораторних показників використовувалися: p_1 – спостережуваний артеріальний тиск (систолічний та діастолічний, мм. рт. ст.); p_2 – пульс (ударів на хвилину); відомості про прийом ліків (p_3 – Амло (0 – немає, 1 – є), p_4 – Егілок (0 – немає, 1 – є); p_5 – Берліприл (0 – немає, 1 – є)). В якості суб'єктивних показників використовувались характеристики самопочуття (p_6 – наявність екстросістолію (0 – немає, 1 – є), p_7 – наявність болю в голові (0 – немає, 1 – є), p_8 – наявність болю в потилиці (0 – немає, 1 – є), p_9 – наявність пульсації (0 – немає, 1 – є), p_{10} – наявність болю в лівому боці (0 – немає, 1 – є), p_{11} – наявність болю в області серця (0 – немає, 1 – є), p_{12} – нестача повітря (0 – немає, 1 – є), p_{13} – наявність болю в животі (0 – немає, 1 – є), p_{14} – загальна слабкість (0 – немає, 1 – є)). В якості метеорологічних характеристик [8] використовувалися (p_{15} – температура повітря (°C), p_{16} – атмосферний тиск (мм. рт. ст.), p_{17} – тип хмарності (0 – немає, 1 – мала, 2 – хмарно, 3 – похмуро), p_{18} – наявність грози (0 – немає, 1 – є), p_{19} – напрямок вітру (0 – штиль, 1– північний, 2 – північно–східний, 3 – східний, 4 – південно-східний, 5 – південний, 6 – південно-західний, 7 – західний, 8 – північно-західний), p_{20} – швидкість вітру (м/с), p_{21} – дані сонячної активності (Mg II індекс [9]). В якості тимчасових характеристик використовувалися: дата (рік, місяць, день), код дня неділі (p_{22}), час (година) спостереження (p_{23}), код частини дня (0 – ранок, 1 – вечір) (p_{24}).

Фрагмент вихідних даних в графічному виді представлений на рис. 1.

Отримані спостереження методом «вікон» були використані для формування вибірки для вирішення задачі якісного прогнозування стану пацієнта на найближчу другу половину доби за даними попередніх спостережень: в якості вхідних ознак використовувались дані за попередню (ранок та вечір) і поточну добу (ранок), а в якості виходу – стан пацієнта ввечері в поточну добу (0 – нормальний, 1 – погіршення стану, що супроводжується підвищенням артеріального тиску).

2 ОГЛЯД ЛІТЕРАТУРИ

У найпростішому випадку нейро-нечітка модель може розглядатися як передоброблювач, де механізм навчання штучної нейронної мережі визначає функції належності або нечіткі правила системи нечіткого виведення за навчаючими даними. Як тільки параметри системи нечіткого виведення визначено, нейронна мережа припиняє використовуватися. База правил нечіткої системи звичайно визначається на основі карт, що самоорганізуються, або методів нечіткої кластеризації. Функції належності звичайно апроксимуються нейронною мережею за навчаючими даними. Найбільш відомими прикладами паралельних нейро-нечітких систем є: нечітка асоціативна пам'ять Коско (B. Kosko), виділення нечітких правил на основі карт Педрича (W. Pedrycz), що самоорганізуються, та системи Номури (H. Nomura), здатні до навчання параметрів нечітких множин [3–5].

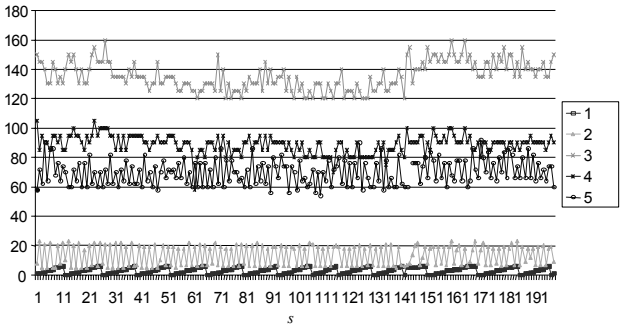


Рисунок 1 – Візуалізація фрагменту вибірки даних: 1 – день тижня, 2 – час, 3 – сістолічний тиск, 4 – діастолічний тиск, 5 – пульс

В інтегрованій моделі нейро-нечітких мереж для визначення параметрів системи нечіткого виведення використовуються методи навчання нейронних мереж. Інтегровані нейро-нечіткі системи розподіляють структури даних і подання знань. Система нечіткого виведення може використовувати людські експертні знання, зберігаючи їхні істотні компоненти в базі правил і базі даних, та виконувати нечітке виведення для одержання вихідного значення. Формування правил і відповідних функцій приналежності сильно залежить від апріорного знання щодо розглянутої системи. Однак універсальний спосіб перетворення дослідних знань людських експертів у базу знань системи нечіткого виведення є невідомим. Є також потреба в адаптованості або деяких методах навчання для одержання виходів з необхідним рівнем точності. З іншого боку, механізм навчання нейронних мереж не покладається на людську експертизу. Через однорідну структуру нейронних мереж, з них складно витягти структуроване знання. Ваги нейронної мережі являють собою коефіцієнти гіперплощини, що розподіляє вхідний простір на області з різними значеннями виходу. Якщо ми можемо візуалізувати цю структуру гіперплощини за навчаючими даними, тоді наступні процедури навчання в нейронній мережі можуть бути виключені.

Найбільш загальний спосіб застосування методу навчання до нечіткої системи полягає в тому, щоб подати її у вигляді архітектури, подібної нейронній мережі. Однак звичайні (градієнтні) методи навчання нейронних мереж не можуть безпосередньо застосовуватися до такої системи, оскільки функції, використовувані в процесі виведення звичайно є недиференційовані. Ця проблема може бути вирішена шляхом використання диференційованих функцій у системі виведення або використання стохастичних методів навчання нейромереж [5–7].

3 МАТЕРІАЛИ ТА МЕТОДИ

Оскільки навчальна вибірка характеризується високими значеннями кількості ознак і екземплярів, процес побудови моделі вимагає значних обчислювальних і часових ресурсів. Це обумовлює доцільність розпаралелювання обчислювальних операцій при синтезі розпізнавальної моделі.

Для синтезу моделі прогнозування стану здоров'я хворого гіпертонічною хворобою розроблений паралельний стохастичний метод. Запропонований метод заснований на стохастичному підході [6, 7, 10, 11], передбачає розпаралелювання основних обчислювальних операцій і використовує апріорну інформацію про значущість ознак. Це дозволяє використовувати можливості багатоядерних процесорів при синтезі нейро-нечітких моделей прогнозування стану здоров'я хворого гіпертонічною хворобою, а також скоротити час на побудову таких моделей.

У розробленому методі на етапі ініціалізація початкової множини рішень відбувається формування множини $R^{(0)} = \left\{\chi_1^{(0)}, \chi_2^{(0)}, ..., \chi_{N\chi}^{(0)}\right\}$ структур виду $\chi_k^{(0)} = \left\{g_{1k}^{(0)}, g_{2k}^{(0)}, ..., g_{N_g k}^{(0)}\right\}$, що визначають основні параметри (значення вагових коефіцієнтів нейронів і параметрів функцій активації) синтезованих нейро-нечітких моделей.

Для наближення початкових точок пошуку $\chi_k^{(0)} = \left\{g_{1k}^{(0)}, g_{2k}^{(0)}, ..., g_{N_g k}^{(0)}\right\}$ ($k = 1, 2, ..., N_\chi$) до оптимальних множину $R^{(0)}$, на відміну до відомих методів стохастичного пошуку [7, 10, 11], будемо формувати з урахуванням апріорної інформації про навчальну вибірку $S = < P, T >$. Для цього діапазон значень $\Delta p_m = \left[p_{m\min}; p_{m\max}\right]$ кожної m-ї ознаки p_m ($m = 1, 2, ..., M$) розбивається на задану кількість інтервалів N_{int}. Таким чином формуються інтервали $\Delta p_{mn} = \left[p_{mn\min}; p_{mn\max}\right]$ ($n = 1, 2, ..., N_{\text{int}}$), на підставі границь яких визначаються параметри нечітких термів.

При обчислені значень параметрів функцій належності нейро-нечіткої моделі будемо враховувати значущість n-го терму Δp_{mn} m-ї ознаки p_m для розпізнавання екземплярів вибірки $S = < P, T >$. Значущість V_{mn} n-го терму m-ї ознаки визначимо як добуток величин $\frac{Q|_{p_m \in \Delta p_{mn}}}{Q}$ та $\left(1 - Entr(\Delta p_{mn})\right)$, що характеризують, відповідно щільність розташування екземплярів множини S в інтервалі Δp_{mn} ознаки p_m і ступінь впливу терму Δp_{mn} на значення вихідного параметру T (1):

$$V_{mn} = \frac{Q|_{p_m \in \Delta p_{mn}}}{Q}\left(1 - Entr(\Delta p_{mn})\right). \quad (1)$$

Ентропія $Entr(\Delta p_{mn})$ інтервалу Δp_{mn} характеризує міру невизначеності вихідного параметра T за умови, що $p_m \in \Delta p_{mn}$, та визначається за формулою (2):

$$Entr(\Delta p_{mn}) = -\sum_{q=1}^{N_{Q, p_m \in \Delta p_{mn}}} \rho_q \log \rho_q . \quad (2)$$

Ймовірність ρ_q того, що вихідний параметр T прийме значення q в множині $S_{p_m \in \Delta p_{mn}}$ (за умови, що $p_m \in \Delta p_{mn}$) можна визначити за формулою (3):

$$\rho_q = \frac{Q|_{T=q; S_{p_m \in \Delta p_{mn}}}}{Q|_{p_m \in \Delta p_{mn}}} . \quad (3)$$

Найбільш ресурсомістким, як правило, є етап оцінювання поточної множини рішень $R^{(i)}$, на якому витрачається велика кількість комп'ютерних та часових ресурсів при обчислені значень цільової функції G для кожного k-го ($k = 1, 2, ..., N_\chi$) рішення $\chi_k^{(i)}$: $G_k^{(i)} = G\left(\chi_k^{(i)}\right)$. Оскільки даний етап має високу обчислювальну складність, виконується повільно і не вимагає обміну даними між рішеннями $\chi_k^{(i)}$, його раціонально виконувати паралельно. Оцінювання рішень $\chi_k^{(0,j)}$ з множини $R^{(0)}$ виконується шляхом перетворення $\chi_k^{(0,j)} \to NFN_k^{(0,j)}$, в результаті якого зі структури $\chi_k^{(0,j)}$ формується нейро-нечітка ме-

режа $NFN_k^{(0,j)}$. В якості критерію оцінювання $G_k^{(0,j)}$, що характеризує рішення $\chi_k^{(0,j)}$, використовується помилка розпізнавання (якщо вихід T в множині S є дискретним) або середньоквадратична помилка (якщо вихід T приймає дійсні значення). Це вимагає обчислення значень вихідного параметра $T_q\left(NFN_k^{(0,j)}\right)$ за синтезованою моделлю $NFN_k^{(0,j)}$ для кожного q-го спостереження множини $S=<P,T>$. Отримана таким чином оцінка $G\left(NFN_k^{(0,j)}\right)$ буде характеризувати пристосованість $G_k^{(0,j)}$ рішення $\chi_k^{(0,j)}$ в множині $R^{(0)}$: $G_k^{(0,j)}=G\left(\chi_k^{(0,j)}\right)=G\left(NFN_k^{(0,j)}\right)$, що надалі дозволить відбирати найбільш пристосовані рішення для формування нових множин $R^{(i+1)}$.

Важливим етапом є створення нової множини рішень $R^{(i+1)}$. З метою більш детального дослідження областей локальних оптимумів доцільним є розбиття поточної множини рішень $R^{(i)}$ на підмножині $R^{(i,j)}$ з подальшим пошуком оптимуму в кожній з них (4):

$$R^{(i)} \to \left\{R^{(i,1)}, R^{(i,2)}, ..., R^{(i,N_{pr})}\right\}. \quad (4)$$

Розбиття (4) пропонується виконувати з урахуванням апріорної інформації про розташування рішень $\chi_k^{(i)}$ в просторі елементів g_l ($l=1, 2, ..., N_g$). Такий підхід, на відміну від застосування острівної моделі еволюційного пошуку [7, 12], що припускає випадкове формування підпопуляцій $R^{(i,j)}$, дозволяє враховувати інформацію про просторове розташування рішень $\chi_k^{(i)}$ в множині $R^{(i)}$ і більш детально дослідити області можливих оптимумів. Кількість рішень $\left|R^{(i,j)}\right|$ в кожній підмножині $R^{(i,j)}$ визначається як відношення загальної кількості рішень N_χ в множині $R^{(i)}$ до кількості процесів N_{pr} (5):

$$\left|R^{(i,j)}\right| = N_\chi / N_{pr}. \quad (5)$$

Таким чином, задача розбиття (4) зводиться до необхідності формування N_{pr} підмножин $R^{(i,j)}$ (кожна з яких складається з $\left|R^{(i,j)}\right|$ рішень) в заданому просторі елементів g_l ($l=1, 2, ..., N_g$) (6):

$$R^{(i,j)} = \left\{\chi_1^{(i,j)}, \chi_2^{(i,j)}, ..., \chi_{\left|R^{(i,j)}\right|}^{(i,j)}\right\},\ j=1, 2, ..., N_{pr}. \quad (6)$$

Розбиття $R^{(i)} \to \left\{R^{(i,1)}, R^{(i,2)}, ..., R^{(i,N_{pr})}\right\}$ пропонується виконувати за допомогою модифікації методу нечітких c-середніх (FCM, Fuzzy c-means) [13–15], оскільки він передбачає можливість віднесення одного екземпляру (рішення $\chi_k^{(i)}$) до різних кластерів з різним ступенем належності, а також може бути використаний для побудови нейро-нечітких систем при визначенні нечітких множин [13, 16]. На відміну від базового методу нечітких c-середніх [15], в якому поточне розбиття на початковій ітерації отримують шляхом випадкової генерації елементів u_{kj}, що визначають належність k-го екземпляра до j-го кластера, в розробленому методі пропонується враховувати інформацію про розташування параметрів g_{lk}, що настроюються. Для визначення початкових значень величин u_{kj} обчислимо ранги $Rg\left(g_{lk}^{(i)}\right)$ рішень $\chi_k^{(i)}$ за l-ю ознакою, що відповідають номерам інтервалів розбиття діапазону значень параметра g_{lk}.

Для визначення початкових значень u_{kj}, що відповідають належності k-го екземпляра до j-го кластеру, будемо використовувати формулу (7):

$$u_{kj} = \frac{N\left(Rg\left(g_{lk}^{(i)}\right)=j\right)}{N_{pr}}, \quad (7)$$

де $N\left(Rg\left(g_{lk}^{(i)}\right)=j\right)$ – кількість рангів $Rg\left(g_{lk}^{(i)}\right)$, які дорівнюють j. Формула (7) забезпечую виконання умови $\sum_{k=1}^{N_\chi} u_{kj} = 1$.

Далі для кожного j-го кластера визначається його центр $C_j = \left\{C_{1j}, C_{2j}, ..., C_{N_g j}\right\}$, де C_{lj} – l-та координата центра j-го кластера, що обчислюється за формулою (8):

$$C_{lj} = \frac{1}{\sum_{k=1}^{N_\chi}\left(u_{kj}\right)^{mp}} \sum_{k=1}^{N_\chi}\left(u_{kj}\right)^{mp} g_{jk}, \quad (8)$$

де mp – параметр методу, як правило, $mp = 2$.

Потім обчислюється значення функції $J\left(R^{(i)}, u^{(i)}, C^{(i)}\right)$ (9), що характеризує якість нечіткого розбиття:

$$J\left(R^{(i)}, u^{(i)}, C^{(i)}\right) = \sum_{k=1}^{N_\chi}\sum_{j=1}^{N_{pr}}\left(u_{kj}\right)^{mp} D^2\left(\chi_k^{(i)}, C_j\right). \quad (9)$$

Як правило, відстань $D\left(\chi_k^{(i)}, C_j\right)$ між рішенням $\chi_k^{(i)}$ і центром j-го кластера C_j обчислюється на основі Евклідової метрики. Після цього формується нове нечітке розбиття шляхом перерахунку значень u_{kj}, використовуючи формулу (10):

$$u_{kj} = \left(\sum_{t=1}^{N_{pr}}\left(\frac{D\left(\chi_k^{(i)}, C_j\right)}{D\left(\chi_k^{(i)}, C_t\right)}\right)^{\frac{2}{mp-1}}\right)^{-1}. \quad (10)$$

При цьому, якщо $D\left(\chi_k^{(i)}, C_j\right) = 0$, то $\chi_k^{(i)}$ є центром j-го кластера, тому встановлюються такі значення параметрів u_{kj}: $u_{kj} = 1$, для всіх інших кластерів: $u_{k'j} = 0, k' \neq k$. Потім, для отримання нечіткого розбиття u_{kj} ($k = 1, 2, ..., N_\chi$, $j = 1, 2, ..., N_{pr}$) використовуючи формули (8)–(9) розраховуються нові центри нечітких кластерів $C_j = \{C_{1j}, C_{2j}, ..., C_{N_g j}\}$ і значення критерію $J\left(R^{(i)}, u^{(i)}, C^{(i)}\right)$. Процес продовжується, доки кількість ітерацій не перевищить максимально допустиму або доки зміна значення критерію $J\left(R^{(i)}, u^{(i)}, C^{(i)}\right)$ носить суттєвий характер (11):

$$\left| J_{old}\left(R^{(i)}, u^{(i)}, C^{(i)}\right) - J\left(R^{(i)}, u^{(i)}, C^{(i)}\right) \right| \leq \varepsilon_J, \qquad (11)$$

де $J_{old}\left(R^{(i)}, u^{(i)}, C^{(i)}\right)$ та $J\left(R^{(i)}, u^{(i)}, C^{(i)}\right)$ – значення критерію якості розбиття J на попередній та поточній ітераціях, відповідно; ε_J – мінімально прийнятна зміна значення критерію J.

Після отримання прийнятних значень величин u_{kj}, що характеризують належність k-го екземпляра $\chi_k^{(i)}$ до j-го кластера виконується формування підмножин $R^{(i,j)}$ ($j = 1, 2, ..., N_{pr}$). При цьому пропонується враховувати як значення належності u_{kj} до кожного з кластерів $R^{(i,j)}$, так і значення цільової функції $G_k^{(i)} = G\left(\chi_k^{(i)}\right)$. Це дозволить підвищити ймовірність відбору при формуванні підмножин $R^{(i,j)}$ таких рішень $\chi_k^{(i)}$, які геометрично відповідають j-му кластеру (просторово розташовані близько до його центру) і володіють при цьому прийнятною пристосованістю $G_k^{(i)}$. Рішення $\chi_k^{(i)}$ будемо відносити в підмножини $R^{(i,j)}$ ($R^{(i,j)} = R^{(i,j)} \bigcup \chi_k^{(i)}$) при виконання умови (12):

$$G_k^{(i)} + u_{kj} \leq rand[0;2]. \qquad (12)$$

Перегляд множини $R^{(i)}$ продовжується та повторюється доки не будуть послідовно сформовані всі підмножини $R^{(i,j)}$, кожна з яких складається з N_χ / N_{pr} рішень.

Після розбиття в головному процесі множини $R^{(i)}$ на підмножини $\left\{R^{(i,1)}, R^{(i,2)}, ..., R^{(i,N_{pr})}\right\}$ виконується їх передача на інші процеси для паралельного стохастичного пошуку екстремумів.

Для формування нового набору $R^{(i',j)}$ на j-му процесі будемо відбирати рішення $\chi_k^{(i,j)} \in R^{(i,j)}$, виходячи зі значень їх цільової функції $G_k^{(i,j)} = G\left(\chi_k^{(i,j)}\right)$. Рішення $\chi_k^{(i,j)} \in R^{(i,j)}$ пропонується вводити в набір рішень $R_{selected}^{(i,j)}$, допущених для формування нового покоління $R^{(i',j)}$ при виконанні умови (13):

$$G_{k,norm}^{(i,j)} \leq rand[0;1]. \qquad (13)$$

Таким чином, чим меншою є помилка $G_k^{(i,j)}$ синтезованої мережі $NFN_k^{(i,j)} = NFN\left(\chi_k^{(i,j)}\right)$, тим вищою є ймовірність рішення $\chi_k^{(i,j)} \in R^{(i,j)}$ бути відібраним для формування нової множини $R^{(i',j)}$.

Нову множину рішень $R^{(i',j)}$ аналогічно еволюційному підходу до пошуку оптимальних рішень [11] пропонується формувати з найбільш пристосованих наборів $\chi_k^{(i,j)}$ (елітних), а також наборів, отриманих в результаті застосування еволюційних операторів схрещування та мутації. Для створення нової множини рішень пропонується використовувати оператори, розроблені в [17].

Випадковий пошук з адаптацією в кожній j-й підмножині $R^{(i,j)}$ пропонується виконувати на j-му процесі паралельної комп'ютерної системи ($j = 1, 2, ..., N_{pr}$) протягом N_{it} ітерацій. Після виконання N_{it} ітерацій випадкового пошуку в кожній підмножині $R^{(i,j)}$ виконується їх об'єднання $R^{(i')} = R^{(i',1)} \bigcup R^{(i',2)} \bigcup ... \bigcup R^{(i',N_{pr})}$ в єдину популяцію з подальшим проведенням імовірнісної оптимізації над об'єднаною множною. Це дозволяє виявити нові області, що містять локальні (можливо, і глобальний) оптимум. З метою скорочення часу імовірнісної оптимізації при роботі з об'єднаною множиною рішень пропонується обчислення значень цільової функції $G_k^{(i)}$ рішень $\chi_k^{(i)}$ виконувати на процесах в паралельній системі (розпаралелити).

Виходячи з зазначеного вище, представимо граф процесу параметричного синтезу нейро-нечітких моделей в ярусно-паралельній формі (рис. 2). Починаючи з ярусу перевірки критеріїв зупинення процес виконується ітеративно.

На рис. 2 операція $RS(R)$ представляє собою випадковий пошук в множині рішень R. Остання операція на графі – $RS(R^{(i')})$ виконується паралельно. Одна ітерація процесу випадкового пошуку $RS(R^{(i')})$ в ярусно-паралельній формі представлена на рис. 3.

Пунктирні лінії на рис. 2 і 3 відокремлюють яруси графа один від одного. Операції на одному ярусі графа виконуються паралельно. Дуги графа демонструють залежності між операціями. Так, не можна приступити до виконання будь-якої операції доки не завершаться інші операції, дуги з яких орієнтовані в вершини розглянутих операцій.

Перехід між ярусами, при якому від однієї вершини графа дуги орієнтовані в кілька вершин наступного ярусу, пов'язаний з розсилкою (Scatter) рішень по процесах. Перехід іншого типу, при якому дуги орієнтовані від декількох вершин одного ярусу до однієї вершині наступного ярусу, виконується за допомогою збору (Gather) рішень в головному процесі. Як можна побачити з рис. 2 і 3, найбільш ресурсномісткі операції, а саме ініціалізація початкової множини рішень, оцінювання поточної множини рішень, випадковий пошук в підмножині рішень, виконуються одночасно на процесах паралельної системи.

Запропонований метод синтезу нейро-нечітких мереж на основі паралельних обчислень ґрунтується на стохастичному підході при пошуку значень параметрів, що настроюються, і полягає в розподілі найбільш ресурсномістких етапів по вузлах паралельної обчислювальної системи, що дозволяє скоротити час налаштування параметрів (значень вагових коефіцієнтів і параметрів функцій належності нейроелементів) синтезованих нейромоделей.

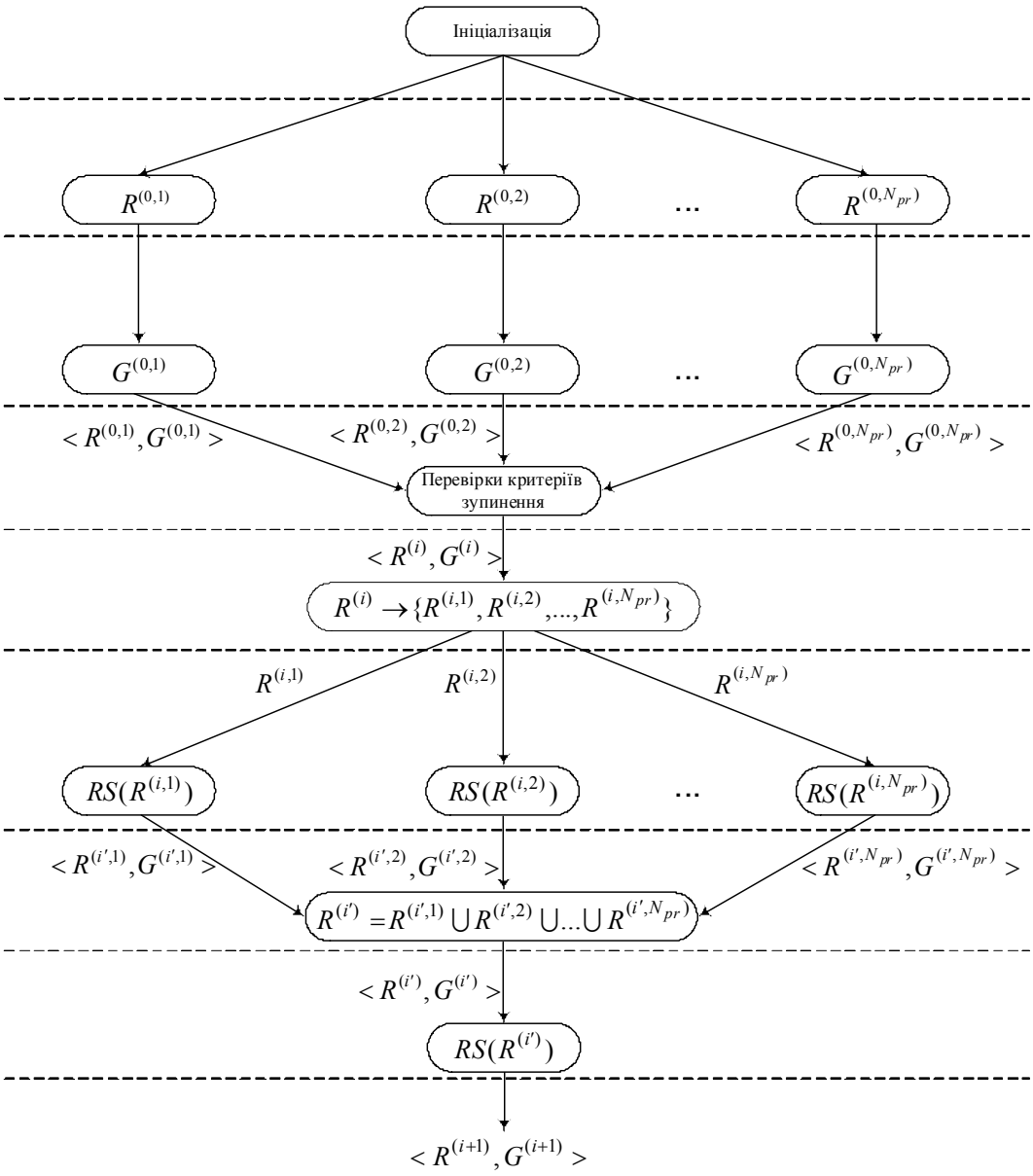


Рисунок 2 – Граф процесу параметричного синтезу нейро-нечітких мереж в ярусно-паралельній формі

4 ЕКСПЕРИМЕНТИ

Для виконання експериментального дослідження запропонованого методу синтезу нейро-нечітких мереж і вирішення задачі індивідуального прогнозування стану здоров'я хворого гіпертонічною хворобою був використаний кластер Інституту проблем моделювання в енергетику імені Г. Е. Пухова НАН України м. Київ. Було задіяно 16 вузлів наступної конфігурації: процесор – Intel Xeon 5405, оперативна пам'ять – 4×2 ГБ DDR-2 на кожен вузол, комунікаційне середовище InfiniBand 20Гб/с. На вузлах кластера встановлене middleware Torque і OMPI.

Для проведення експериментів розроблено програмне забезпечення мовою C++ із застосуванням бібліотеки MPI [18–20]. Для оцінювання витраченого на виконання методу часу використана бібліотечна функція MPI_Wtime(). Експерименти були виконані по 3 рази на одному, двох, чотирьох, восьми і шістнадцяти ядрах за-діяних процесорів.

Як аналог запропонованому методу була узята острівна модель еволюційного пошуку (Island Genetic Algorithm, IGA) [11, 12].

5 РЕЗУЛЬТАТИ

Результати експериментів наведено на рис. 4–10 та у табл. 1–2. На рис. 4 приведені усереднені часові витрати на виконання запропонованого методу і методу IGA на кластері.

Для встановлення причини нелінійного росту прискорення було виміряно комунікаційні витрати, а саме: пересилання між ядрами процесорів і синхронізацію потоків. Отримані результати наведено в таблиці 1 і на рис. 6.

Графік ефективності кластерної системи, у якій виконувалася реалізація запропонованого методу і методу IGA, представлений на рис. 7.

Запропонований метод по своїй суті добре реалізується на SIMD-системах: кожне ядро виконує ті самі інструкції для кожної реалізації хромосом, тому були проведені експерименти на графічному процесорі (GPU) NVIDIA GTX 285+, що програмувався за технологією CUDA [21]. При цьому задавалася кількість задіяних по-

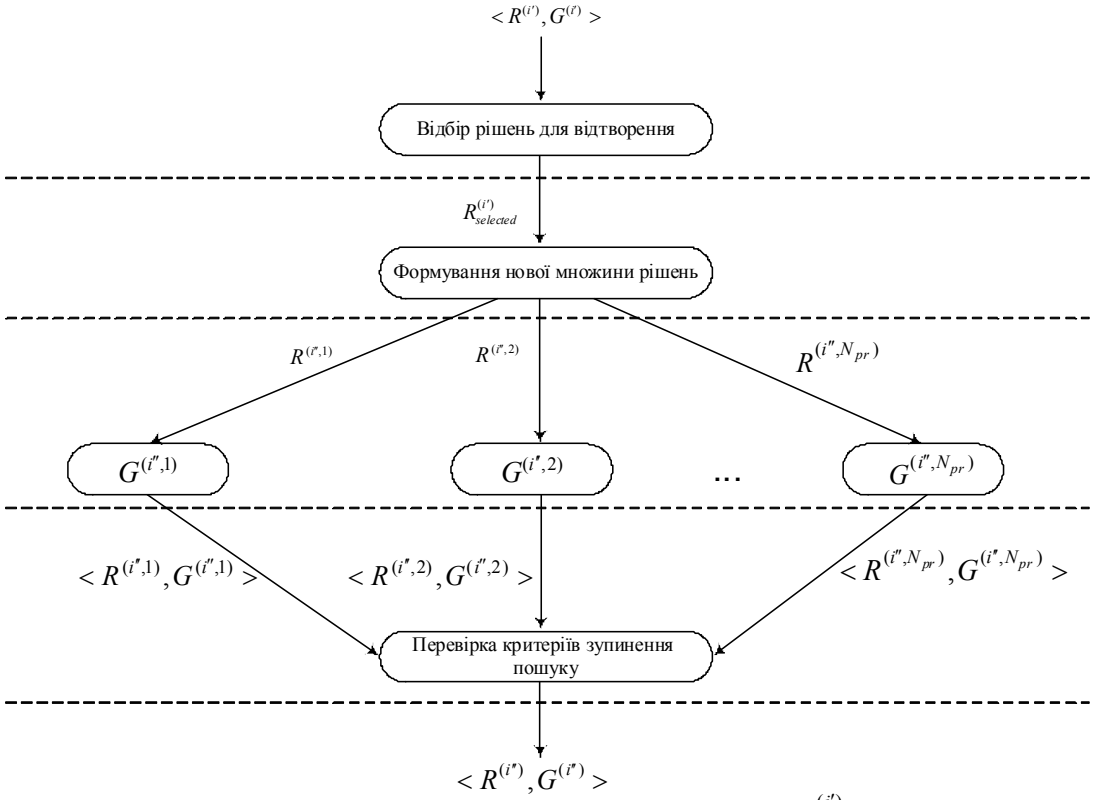


Рисунок 3 – Граф однієї ітерації випадкового пошуку $RS(R^{(i')})$

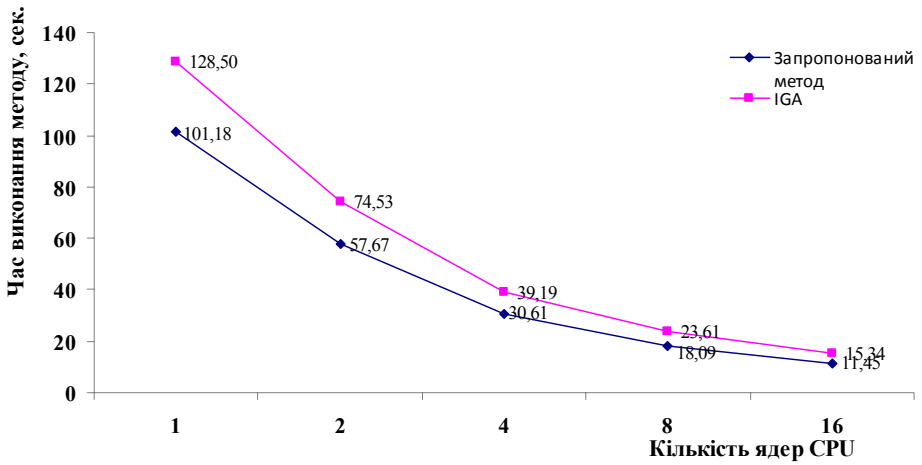


Рисунок 4 – Усереднений час виконання запропонованого методу і методу IGA на кластері

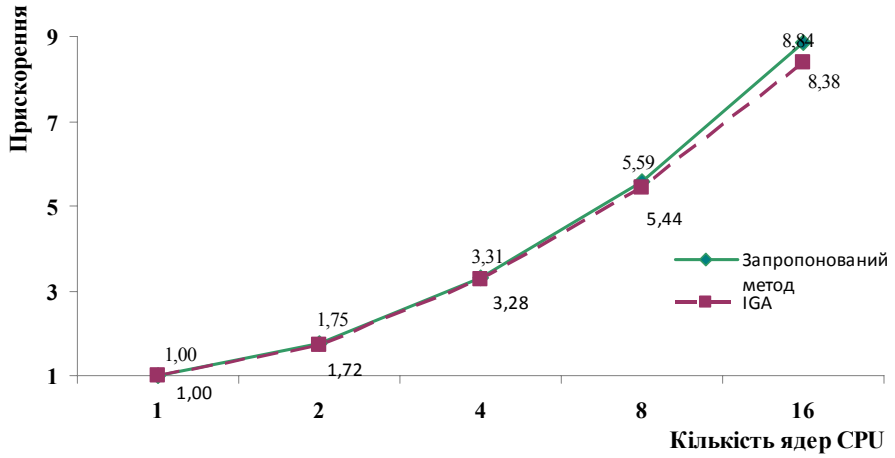


Рисунок 5 – Графік прискорення обчислювального процесу на кластері

Таблиця 1 – Комунікаційні витрати обчислювального процесу, що реалізує запропонований метод

Кількість задіяних ядер процесорів	Час пересилань та синхронізацій, с		Частка пересилань та синхронізацій	
	Запропонований метод	IGA	Запропонований метод	IGA
1	0	0	0	0
2	8,07	11,92	0,14	0,16
4	6,43	8,62	0,21	0,22
8	7,78	11,10	0,43	0,47
16	9,27	13,96	0,81	0,91

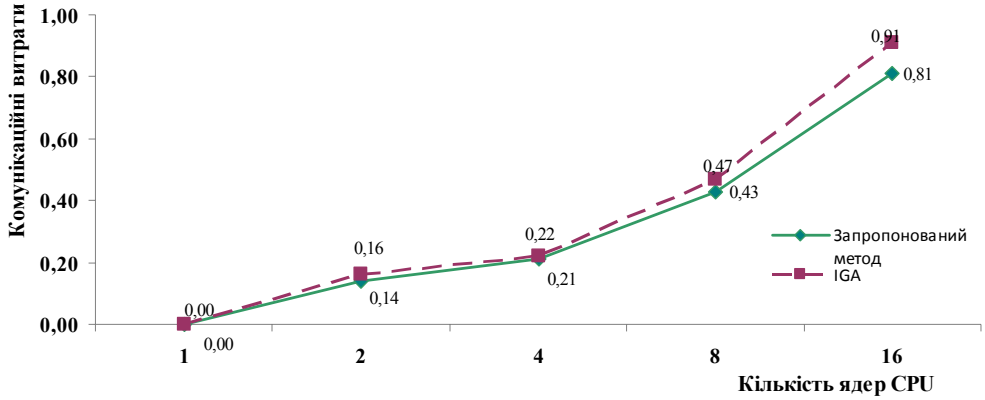


Рисунок 6 – Графік залежності комунікаційних витрат від числа задіяних ядер процесорів

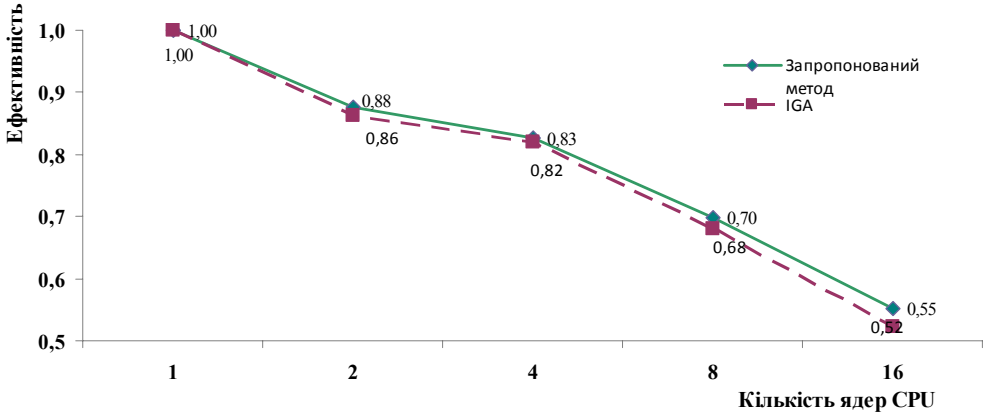


Рисунок 7 – Ефективність кластерної системи, що реалізує запропонований метод і метод IGA

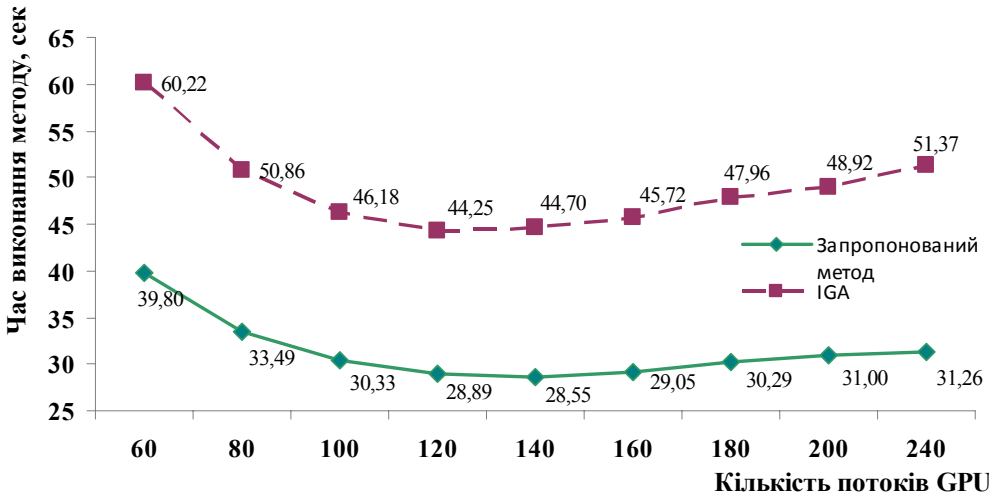


Рисунок 8 – Усереднений час виконання запропонованого методу і методу IGA на графічному процесорі

токів GPU і вимірювався час виконання методу за допомогою убудованих у пакет CUDA функцій. Експерименти для запропонованого методу і методу IGA були повторені по 2 рази на 60, 80, 100, 120, 140, 160, 180, 200 і 240 потоках. Усереднений час виконання методів відображено на рис. 8.

Графіки прискорення і комунікаційних витрат, що ілюструють роботу досліджуваних методів при використанні графічного процесора (GPU), наведено на рис. 9 і 10, відповідно.

У таблиці 2 наведено комунікаційні витрати обчислювального процесу, що реалізує запропонований метод на графічному процесорі.

6 ОБГОВОРЕННЯ

Як видно з рис. 4 і 5, запропонований метод у порівнянні з методом IGA дозволяє швидше виконати випадковий пошук, оскільки застосування модифікованих операторів створення нових рішень зменшує загальну кількість виконуваних на процесах операцій [17]. Прискорення обчислювального процесу, що реалізує запропонований метод і метод IGA, росте нелінійно зі збільшенням числа задіяних ядер процесорів.

Важливо відзначити, що обидва методи (запропонований метод і метод IGA) дозволили забезпечити побудову нейро-нечітких моделей із прийнятною точністю (при проведенні експериментів необхідно було синтезувати моделі, помилка прогнозування яких не перевищує 0,01). Помилка прогнозування на тестових даних склала 0,017 і 0,021 для запропонованого методу і методу IGA, відповідно, що також є припустимим для розв'язуваної задачі індивідуального прогнозування стану здоров'я хворого гіпертонічною хворобою.

Частка комунікаційних витрат (рис. 6, табл. 1) в обчислювальному процесі, що реалізує запропонований метод і метод IGA, істотно росте зі збільшенням кількості ядер задіяних процесорів. При цьому на більше, ніж восьми ядрах, частка комунікаційних витрат значно перевищує частку цільових обчислень. Тому при реалізації запропонованого методу на кластері доцільно застосовувати чотири-шість ядер процесорів, щоб збалансувати відношення витрат до ефективності системи. Застосування методу IGA на кластері так само є виправданим, проте, при вирішенні поставленої задачі IGA затратив більше часу на пересилання в порівнянні з запропонованим методом.

З рис. 7 видно, що обидва методи продемонстрували прийнятну ефективність у кластерній системі. Для запропонованого методу (рис. 8) збільшення числа потоків понад 140 приводить до уповільнення обчислювального процесу. Це можна пояснити ростом комунікаційних витрат, а також простоєм незадіяних потоків. Для методу IGA збільшення кількості потоків уже понад 120 сповільнює обчислювальний процес.

Дані, відображені у табл. 2 та на рис. 10, демонструють істотне зниження комунікаційних витрат при виконанні запропонованого методу і методу IGA на GPU у порівнянні з кластером, комунікаційні витрати якого подані в таблиці 1.

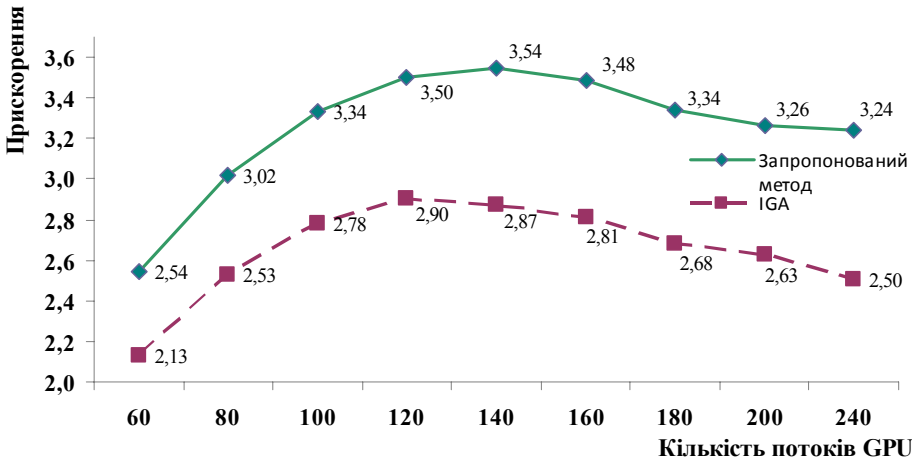


Рисунок 9 – Графік прискорення реалізації запропонованого методу і методу IGA на графічному процесорі

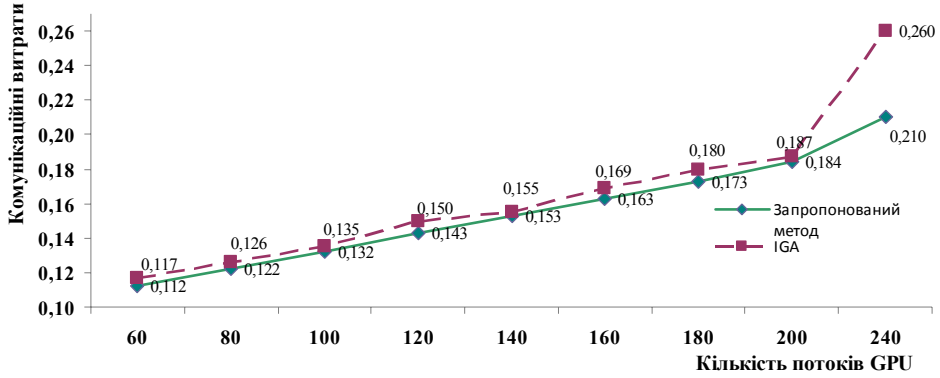


Рисунок 10 – Графік комунікаційних витрат на реалізацію запропонованого методу і методу IGA на GPU

Таблиця 2 – Комунікаційні витрати обчислювального процесу, що реалізує запропонований метод на GPU

Кількість задіяних потоків GPU	Час пересилань та синхронізацій, с		Частка пересилань та синхронізацій, с	
	Запропонований метод	IGA	Запропонований метод	IGA
60	4,46	7,05	0,112	0,117
80	4,09	6,41	0,122	0,126
100	4,00	6,23	0,132	0,135
120	4,13	6,64	0,143	0,150
140	4,37	6,93	0,153	0,155
160	4,74	7,73	0,163	0,169
180	5,24	8,63	0,173	0,180
200	5,70	9,15	0,184	0,187
240	6,56	13,36	0,210	0,260

Таким чином, застосування паралельних систем для реалізації запропонованого методу є доцільним, оскільки істотно знижується затрачуваний на виконання методу час (у 8,84 рази для 16-ти ядер кластеру та у 3,54 рази для 140 потоків GPU). При цьому SIMD-система на основі GPU є істотно дешевшою за кластер і може бути використана для одержання прийнятного прискорення реалізації запропонованого методу. У порівнянні з методом IGA запропонований метод продемонстрував кращу масштабованість як на кластері, так і на GPU. Крім того, запропонований метод дозволяє швидше одержати рішення й ефективніше задіє обчислювальні ресурси, забезпечуючи при цьому побудову моделей з прийнятною точністю прогнозування.

ВИСНОВКИ

У роботі вирішено актуальну задачу автоматизації індивідуального прогнозування стану здоров'я хворого гіпертонічною хворобою.

Наукова новизна роботи полягає в тому, що запропоновано метод синтезу нейро-нечітких моделей на основі паралельних обчислень, який ґрунтується на стохастичному підході при пошуку значень параметрів, що настроюються, і полягає в розподілі найбільш ресурсномістких етапів по вузлах паралельної обчислювальної системи. Це дозволяє скоротити час настроювання параметрів (значень вагових коефіцієнтів і параметрів функцій належності нейроелементів) синтезованих нейромоделей.

Практична цінність отриманих результатів полягає в тому, що розроблено програмне забезпечення, яке реалізує запропонований паралельний метод і дозволяє виконувати побудову нейро-нечітких мереж на основі заданих наборів спостережень для синтезу моделей індивідуального прогнозування стану здоров'я хворого гіпертонічною хворобою. Результати проведених експериментів свідчать про прийнятні значення показників ефективності запропонованого паралельного методу синтезу нейро-нечітких мереж

Перспективи подальших досліджень полягають у застосуванні запропонованого підходу до синтезу нейро-нечітких мереж з навчальних вибірок даних для вирішення практичних задач неруйнівного контролю якості.

ПОДЯКИ

Роботу виконано в рамках держбюджетної науково-дослідної теми Запорізького національного технічного університету «Інтелектуальні інформаційні технології автоматизації проектування, моделювання, керування та діагностування виробничих процесів і систем» (номер державної реєстрації 0112U005350) за підтримки міжнародного проекту «Centers of Excellence for young RESearchers» (CERES) програми «Tempus» Європейської Комісії (реєстраційний номер 544137-TEMPUS-1-2013-1-SK-TEMPUS-JPHES).

Навчальну вибірку надано професором Запорізького національного технічного університету проф. С. О. Субботіним.

СПИСОК ЛІТЕРАТУРИ

1. Beevers D. G. ABC of hypertension / D. G. Beevers, G. Y. Lip, E. O'Brien. – London : BMJ Books, 2007. – 88 p.
2. Camm A. J. The ESC Textbook of Cardiovascular Medicine / A. J. Camm, T. F. Luscher, P. Serruys. – Hoboken : Wiley-Blackwell, 2006. – 1136 p. DOI: 10.1093/med/9780199566990.001.0001.
3. Suzuki K. Artificial Neural Networks: Architectures and Applications / K. Suzuki. – New York : InTech, 2013. – 264 p. DOI: 10.5772/3409.
4. Hanrahan G. Artificial Neural Networks in Biological and Environmental Analysis / G. Hanrahan. – Boca Raton, Florida : CRC Press, 2011. – 214 p. DOI: 10.1201/b10515.
5. Субботін С. О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень : навч. посібник / С. О. Субботін. – Запоріжжя : ЗНТУ, 2008. – 341 с.
6. Baragona R. Evolutionary Statistical Procedures: An Evolutionary Computation Approach to Statistical Procedures Designs and Applications / R. Baragona, F. Battaglia, I. Poli. – Berlin : Springer, 2011. – 296 p.
7. Spall J. C. Introduction to Stochastic Search and Optimization / J. C. Spall. – Hoboken : Wiley, 2003. – 618 p. DOI: 10.1002/0471722138.
8. Дневник погоды [Электронный ресурс]. – Москва : ООО «НПЦ «Мэп Мейкер», 2014. – Режим доступа: http://www.gismeteo.ru/diary/5093.
9. Weber M. Observations of Solar Activity (Mg II Index) by GOME, SCIAMACHY, and GOME-2 [Electronic resource]. – Bremen: University of Bremen, 2014. – Access mode: http://www.iup.uni-bremen.de/gome/gomemgii.html
10. Smith S. Genetic and Evolutionary Computation: Medical Applications / S. Smith, S. Cagnoni. – Chicehster : John Wiley & Sons, 2011. – 250 p. DOI: 10.1002/9780470973134.
11. Yu X. Introduction to Evolutionary Algorithms (Decision Engineering) / X. Yu, M. Gen. – London: Springer, 2010. – 418 p. DOI: 10.1007/978-1-84996-129-5.
12. Cantu-Paz E. Efficient and Accurate Parallel Genetic Algorithms / E. Cantu-Paz. – Massachusetts : Kluwer Academic Publishers, 2001. – 162 p. DOI: 10.1007/978-1-4615-4369-5.
13. Encyclopedia of artificial intelligence / Eds.: J. R. Dopico, J. D. de la Calle, A. P. Sierra. – New York : Information Science Reference, 2009. – Vol. 1–3. – 1677 p.
14. Encyclopedia of machine learning / [eds. C. Sammut, G. I. Webb]. – New York : Springer, 2011. – 1031 p.
15. Mumford C. L. Computational Intelligence / C. L. Mumford, L. C. Jain. – Berlin : Springer-Verlag, 2009. – 732 p. DOI: 10.1007/978-3-642-01799-5.
16. Bishop C.M. Pattern recognition and machine learning / C. M. Bishop. – New York: Springer, 2006. – 738 p.
17. Oliinyk A. O. Using Parallel Random Search to Train Fuzzy Neural Networks / A. O. Oliinyk, S. Yu. Skrupsky, S. A. Subbotin // Automatic Control and Computer Sciences. – 2014. – Vol. 48, Issue 6. – P. 313–323. DOI: 10.3103/S0146411614060078.

18. Gebali F. Algorithms and Parallel Computing / F. Gebali. – New Jersey: John Wiley & Sons, 2011. – 364 p. DOI: 10.1002/9780470932025.
19. Roosta S.H. Parallel Processing and Parallel Algorithms: Theory and Computation / S. H. Roosta. – New York : Springer-Verlag, 2000. – 566 p.
20. Quinn M. J. Parallel Programming in C with MPI and OpenMP / M. J. Quinn. – New York: McGraw-Hill, 2004. – 529 p.
21. NVIDIA CUDA Compute Unified Device Architecture 5.5 / Santa Clara: NVIDIA Corporation, 2014. – 117 p.

Стаття надійшла до редакції 20.03.2015.

Олейник А. А.

Канд. техн. наук, доцент, доцент кафедры программных средств, Запорожский национальный технический университет, Запорожье, Украина

СИНТЕЗ НЕЙРО-НЕЧЕТКИХ МОДЕЛЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ СОСТОЯНИЯ ЗДОРОВЬЯ БОЛЬНОГО НА ОСНОВЕ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ

Решена задача разработки математического обеспечения для автоматизации процесса индивидуального прогнозирования состояния здоровья больного гипертонической болезнью. Объектом исследования является процесс синтеза моделей для медицинского диагностирования. Предмет исследования составляют методы синтеза нейро-нечетких моделей для медицинского диагностирования. Цель работы: повысить эффективность процесса синтеза нейро-нечетких сетей для построения диагностических моделей на основе обучающих выборок. Предложен стохастический метод синтеза нейро-нечетких моделей на основе параллельных вычислений, основанный на стохастическом подходе при поиске значений настраиваемых параметров, и заключающийся в распределении наиболее ресурсоемких этапов по узлам параллельной вычислительной системы. Предложенный метод позволяет сократить время настройки параметров (значений весовых коэффициентов и параметров функций принадлежности нейроэлементов) синтезированных нейромоделей. Разработано программное обеспечение, реализующее предложенный метод. Проведены эксперименты по исследованию свойств предложенного метода. Результаты экспериментов позволяют рекомендовать предложенный метод для использования на практике.

Ключевые слова: множество решений, нейронная сеть, признак, параллельное программирование, прогнозирование, синтез модели.

Oliinyk A.

PhD., Associate Professor, Associate Professor of Department of Software Tools, Zaporizhzhya National Technical University, Zaporizhzhya, Ukraine

SYNTHESIS OF NEURO-FUZZY MODEL FOR PATIENT HEALTH PREDICTING BASED ON PARALLEL COMPUTING

The problem of mathematical support development is solved to automate the process of individual health prediction of hypertensive patient. The object of study is the process of model synthesis for medical diagnosis. The subject of study are methods of neuro-fuzzy model synthesis for medical diagnosis. The purpose of the work is to improve the efficiency of the process of neuro-fuzzy network synthesis for constructing diagnostic models based on training samples. The stochastic method for the synthesis of neuro-fuzzy models based on parallel computing is proposed. It uses the stochastic approach for finding the values of adjustable parameters, and consists in the distribution of the most demanding stages on the nodes in parallel computing system. The proposed method can reduce the time of parameters calculation (the weighting coefficients and parameters of membership functions of neurons) of synthesized neuromodels. The software implementing proposed method is developed. The experiments to study the properties of the proposed method are conducted. The experimental results allow to recommend the proposed method for use in practice.

Keywords: solution set, neural network, feature, parallel programming, prediction, model synthesis.

REFERENCES

1. Beevers D. G., Lip G. Y., O'Brien E. ABC of hypertension. London, BMJ Books, 2007, 88 p.
2. Camm A. J., Luscher T. F., Serruys P. The ESC Textbook of Cardiovascular Medicine. AHoboken, Wiley-Blackwell, 2006, 1136 p. DOI: 10.1093/med/9780199566990.001.0001.
3. Suzuki K. Artificial Neural Networks: Architectures and Applications. New York, InTech, 2013, 264 p. DOI: 10.5772/3409.
4. Hanrahan G. Artificial Neural Networks in Biological and Environmental Analysis. Boca Raton, Florida, CRC Press, 2011, 214 p. DOI: 10.1201/b10515.
5. Subbotin S. O. Podannja j obrobka znan' u sistemah shtuchnogo intelektu ta pidtrimki prijnjattja rishen' : navch. posibnik. Zaporizhzhja, ZNTU, 2008, 341 p.
6. Baragona R., Battaglia F., Poli I. Evolutionary Statistical Procedures: An Evolutionary Computation Approach to Statistical Procedures Designs and Applications. Berlin, Springer, 2011, 296 p.
7. Spall J. C. Introduction to Stochastic Search and Optimization. Hoboken, Wiley, 2003, 618 p. DOI: 10.1002/0471722138.
8. Dnevnik pogody [Elektronnyj resurs]. Moscow, OOO «NPC «Mep Mejker», 2014, Rezhim dostupa: http://www.gismeteo.ru/diary/5093.
9. Weber M. Observations of Solar Activity (Mg II Index) by GOME, SCIAMACHY, and GOME-2 [Electronic resource]. Bremen, University of Bremen, 2014, Access mode: http://www.iup.uni-bremen.de/gome/gomemgii.html
10. Smith S., Cagnoni S. Genetic and Evolutionary Computation: Medical Applications. Chicehster, John Wiley & Sons, 2011, 250 p. DOI: 10.1002/9780470973134.
11. Yu X., Gen M. Introduction to Evolutionary Algorithms (Decision Engineering). London, Springer, 2010, 418 p. DOI: 10.1007/978-1-84996-129-5.
12. Cantu-Paz E. Efficient and Accurate Parallel Genetic Algorithms. Massachusetts, Kluwer Academic Publishers, 2001, 162 p. DOI: 10.1007/978-1-4615-4369-5.
13. Eds.: J. R. Dopico, J. D. de la Calle, A. P. Sierra. Encyclopedia of artificial intelligence. New York, Information Science Reference, 2009, Vol. 1–3, 1677 p.
14. Eds. C. Sammut, G. I. Webb Encyclopedia of machine learning. New York, Springer, 2011, 1031 p.
15. Mumford C. L., Jain L. C. Computational Intelligence. Berlin, Springer-Verlag, 2009, 732 p. DOI: 10.1007/978-3-642-01799-5.
16. Bishop C. M. Pattern recognition and machine learning. New York, Springer, 2006, 738 p.
17. Oliinyk A. O., Skrupsky S. Yu., Subbotin S. A. Using Parallel Random Search to Train Fuzzy Neural Networks, *Automatic Control and Computer Sciences,* 2014, Vol. 48, Issue 6, pp. 313–323. DOI: 10.3103/S0146411614060078.
18. Gebali F. Algorithms and Parallel Computing. New Jersey, John Wiley & Sons, 2011, 364 p. DOI: 10.1002/9780470932025.
19. Roosta S. H. Parallel Processing and Parallel Algorithms: Theory and Computation. New York, Springer-Verlag, 2000, 566 p.
20. Quinn M. J. Parallel Programming in C with MPI and OpenMP. New York, McGraw-Hill, 2004, 529 p.
21. NVIDIA CUDA Compute Unified Device Architecture 5.5 / Santa Clara. NVIDIA Corporation, 2014, 117 p.

УДК 004.032.26

Парфененко Ю. В.[1], Шендрик В. В.[2], Галіченко О. С.[3]

[1]*Асистент кафедри комп'ютерних наук, Сумський державний університет, м. Суми, Україна*
[2]*Канд. техн. наук, доцент кафедри комп'ютерних наук, Сумський державний університет, м. Суми, Україна*
[3]*Студент факультету Електроніки та інформаційних технологій, Сумський державний університет, м. Суми, Україна*

ПРОГНОЗУВАННЯ ТЕПЛОСПОЖИВАННЯ БУДІВЕЛЬ СОЦІАЛЬНО-БЮДЖЕТНОЇ СФЕРИ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ

Запропоновано метод удосконалення процесів підтримки прийняття рішень при управлінні режимами теплозабезпечення закладів соціально-бюджетної сфери шляхом розроблення моделі прогнозу споживання теплової енергії будівлею. Об'єктом дослідження є процес вибору найбільш оптимальної архітектури нейронної мережі для вирішення задачі прогнозування теплоспоживання будівлі соціально-бюджетної сфери. Предмет дослідження становлять моделі прогнозування теплоспоживання будівель соціально-бюджетної сфери з використанням штучних нейронних мереж. Метою даного дослідження є підвищення достовірності прогнозування потреб будівель соціально-бюджетної сфери у теплозабезпеченні. Розроблено моделі короткострокового прогнозування теплозабезпечення об'єктів соціально-бюджетної сфери з використанням штучних нейронних мереж, що враховують вплив погодних умов, коливання потреби в тепловій енергії в залежності від графіку роботи закладу, а також минулі значення потреб у теплозабезпеченні. Моделі побудовано на основі архітектур нелінійної мережі типу вхід-вихід, нелінійної авторегресійної мережі, нелінійної авторегресійної нейронної мережі із зовнішніми входами. Запропоновані моделі реалізовано у програмному середовищі Matlab, для навчання використано алгоритм Левенберга-Марквардта. Проведено експерименти з дослідження точності розроблених моделей, які показали, що найбільшої точності прогнозу можна досягти при використанні моделі типу NARX. Використання запропонованої моделі в рамках інформаційної системи «HeatCAM» для прогнозування теплоспоживання будівель протягом опалювального сезону дозволяє підвищити ефективність підтримки процесу прийняття рішень при управлінні режимами теплозабезпечення, що призводить до зниження споживання теплової енергії.

Ключові слова: теплоспоживання, регулювання, моделювання, інформаційна система, прогнозування, нейронні мережі, енергозбереження.

НОМЕНКЛАТУРА

$F(\cdot)$ – оператор нейромережевого відображення;

L – кількість нейронів прихованого шару;

MAE – середня абсолютна помилка прогнозу;

$MAPE$ – середня абсолютна помилка прогнозу у відсотках;

MSE – середньоквадратична помилка прогнозу;

N – потужність множини даних для перевірки достовірності прогнозу.

n – значення періоду затримки прогнозу;

NAR – модель нелінійної авторегресійної нейронної мережі;

$NARX$ – модель нелінійної авторегресійної нейронної мережі з екзогенними входами;

NIO – модель нейронної мережі типу вхід-вихід;

Q – критерій якості прогнозування;

$RMSE$ – квадратний корінь із середньоквадратичної помилки прогнозу;

t – дискретний момент часу;

$\hat{w}$ – множина значень вагових коефіцієнтів нейронної мережі;

X – множина даних моніторингу теплоспоживання;

$x(t)$ – вектор вхідних змінних;

x_1 – компонента вектору $x(t)$ температура навколишнього середовища, С;

x_2 – компонента вектору $x(t)$ швидкість вітру, м/с;

x_3 – компонента вектору $x(t)$ атмосферний тиск, мм. рт. ст.;

x_4 – компонента вектору $x(t)$ вологість повітря, %;

x_5 – компонента вектору $x(t)$ тип дня тижня;

$x(t-n)$ – значення вхідного сигналу нейронної мережі у попередні періоди часу;

Y – множина вихідних даних системи прогнозування;

$y(t)$ – фактичне значення теплоспоживання будівлі, Гкал;

$\hat{y}(t)$ – прогнозоване значення теплоспоживання будівлі, Гкал;

$y(t-n)$ – значення вихідного сигналу нейронної мережі у попередні періоди часу;

Z – множина даних теплоспоживання будівлі;

Z_{tr} – множина даних для навчання нейронної мережі;

Z_t – тестова множина даних;

Z_v – множина даних для перевірки достовірності прогнозу.

ВСТУП

Застосування інформаційних технологій у теплоенергетиці дозволяє вирішити широке коло задач проектування, експлуатації та оперативно-диспетчерського управління функціонуванням систем теплозабезпечення. Основні дослідження у цьому напрямку пов'язані з побудовою моделей для розрахунків параметрів потокорозподілу, оцінкою поточного стану та оптимізацією схем теплових мереж, розробкою алгоритмів аналізу та оптимізації усталених режимів теплозабезпечення, прогнозування режимів теплоспоживання, а також потреб у паливних ресурсах та тепловій енергії.

Прогнозування, як складова системних досліджень енергетики, є ключовим елементом управління її розвитком [1], що має на меті підвищення енергоефективності та дозволяє досягти економії енергетичних ресурсів.

Реалізація задач регулювання теплоспоживання будівель з метою забезпечення комфортних умов для проживання та роботи в опалювальних приміщеннях та усу-

DOI 10.15588/1607-3274-2015-2-5

нення понаднормового споживання теплової енергії при коливанні температури навколишнього середовища та потреб у теплозабезпеченні в залежності від графіку роботи закладу вимагає застосування прогнозних моделей з високим ступенем достовірності.

Метою даного дослідження є підвищення достовірності прогнозування потреб будівель соціально-бюджетної сфери у теплозабезпеченні. При цьому необхідно провести аналіз існуючих методів прогнозування, що застосовуються у теплоенергетиці, розробити моделі прогнозування потреб будівлі у тепловій енергії з використанням штучних нейронних мереж, реалізувати створені моделі та провести аналіз результатів моделювання.

1 ПОСТАНОВКА ЗАДАЧІ

У загальному вигляді задача нейромережевого прогнозування потреби будівлі соціально-бюджетної сфери у теплової енергії зводиться до побудови математичної моделі на основі сукупності даних щодо кліматичних умов, графіку роботи та теплоспоживання будівлі, що поділяються на множину вхідних даних X та вихідних даних Y при заданих обмеженнях. Процедура побудови математичної моделі системи теплоспоживання Z на основі даних моніторингу є задачею знаходження такої моделі нейронної мережі, яка б виконувала нелінійну апроксимацію відображення, при якому забезпечується прийнятне значення заданого критерію якості Q:

$$\hat{y}(t) = F(X(t)). \tag{1}$$

Компоненти вхідного вектора X у формулі (1) для кожного об'єкта теплоспоживання:

$$X(t) = (x(t), t). \tag{2}$$

Результуюча змінна $\hat{y}(t)$ – витрати теплової енергії на обігрів за час t кожного об'єкта теплоспоживання, Гкал.

Для оцінки достовірності моделі прогнозування та вибору найбільш оптимальної для прогнозування теплоспоживання архітектури нейронної мережі при обчисленні критерію якості Q можуть бути використані помилки прогнозу:

– середня абсолютна помилка прогнозу MAE, що обчислюється за формулою:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|y(t) - \hat{y}(t)\right|; \tag{3}$$

– середню абсолютна помилка прогнозу у відсотках MAPE, що розраховується за формулою:

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y(t) - \hat{y}(t)}{y(t)}\right| \cdot 100\%; \tag{4}$$

– квадратний корінь із середньоквадратичної помилки прогнозу RMSE, що розраховується за формулами:

$$MSE = \frac{1}{N-1}\sum_{t=1}^{N}(y(t) - \hat{y}(t))^2, \tag{5}$$

$$RMSE = \sqrt{MSE}. \tag{6}$$

2 ОГЛЯД ЛІТЕРАТУРИ

При дослідженні енергетичних систем широко застосовуються формалізовані методи, що базуються на використанні моделей та алгоритмів обробки й аналізу даних. В останні роки з'явилась значна кількість методів для вирішення проблеми прогнозування у теплоенергетиці [2]. При прогнозування потреб будівлі у тепловій енергії використовуються методи аналізу часових рядів [3] (дозволяють на основі моделі, що описує поведінку часового ряду, визначати його майбутні значення), авторегресійні моделі та методи [4–6] (базуються на побудові функції лінійної залежності параметрів часового ряду від їх попередніх значень), а також методи з використанням штучних нейронних мереж [7–9] (проводиться навчання нейронної мережі з використанням навчальної вибірки даних, яка потім застосовується при визначенні майбутніх значень вихідних параметрів).

Порівняльну характеристику методів, що використовуються при короткостроковому прогнозуванні теплоспоживання будівель, їх переваги та недоліки наведено в табл. 1.

Найпоширенішим методом прогнозування на даний час є використання штучних нейронних мереж. При цьому враховуються різні фактори впливу на потребу в тепло забезпеченні. Для короткострокового прогнозування потреб будівлі у теплозабезпеченні в роботі [8] використано багатошарову нейронну мережу з двома шарами, входами якої є добові дані теплоспоживання, а також середньодобова температура повітря навколишнього середовища. Для навчання даної нейронної мережі застосовується алгоритм Левенберга-Марквардa. В роботі [7] запропоновано враховувати при прогнозуванні потреб у теплозабезпеченні метеорологічні умови навколишнього середовища (температуру, атмосферного тиску та швидкості вітру).

Недоліками розглянутих вище моделей є те, що вони не враховують соціальний фактор. Специфіка роботи закладів соціально-бюджетної сфери полягає у тому, що потреби у теплозабезпеченні залежать не лише від кліматичних умов навколишнього середовища, а й від режи-

Таблиця 1 – Порівняльна характеристика методів прогнозування у теплоенергетиці

Метод	Переваги	Недоліки
Аналізу часових рядів	– простота; – прозорість моделювання;	– не враховує можливої взаємодії змінних;
Авторегресійні методи	– простота, прозорість моделювання; – одноманітність аналізу і проектування;	– трудомісткість і ресурсомісткість ідентифікації моделей; – неможливість моделювання нелінійних зележностей; – низька адаптивність;
Нейромережеві методи	– нелінійність моделей; – масштабованість; – висока адаптивність; – одноманітність аналізу і проектування;	– складність вибору архітектури; – жорсткі вимоги до навчальної вибірки; – складність вибору алгоритму навчання; – ресурсомісткість процесу навчання

му роботи закладу (робочий день, вихідний, канікули взимку). Таким чином у даній роботі запропоновано розглядати у якості вхідних параметрів прогнозування як дані кліматичних умов, так і режим роботи закладу. Також порівняння декількох архітектур нейронних мереж дозволить визначити найбільш прийнятний тип моделі нейронної мережі.

3 МАТЕРІАЛИ ТА МЕТОДИ

У даній роботі для вирішення задачі прогнозування потреби будівлі у тепловій енергії пропонується застосовувати методи з використанням штучних нейронних мереж. Однією з головних переваг штучних нейронних мереж порівняно з іншими методами прогнозування є можливість моделювання нелінійних залежностей. Точність прогнозування при використанні даного методу залежить від вибору топології, алгоритму навчання, що використовується для визначення вагових коефіцієнтів, а також входів нейронної мережі.

Процес побудови нейромережі для прогнозування теплоспоживання будівлі умовно можна розподілити на етап вибору типу моделі нейромережі, що визначає її структуру, та етап пошуку оптимальних параметрів нейромережі.

На основі проведеного аналізу попередніх досліджень для нейромережевого прогнозування обрано архітектури нейронної мережі типів [10]:

1) нелінійна нейромережа прямого розповсюдження типу «вхід-вихід», що враховує для прогнозування лише вхідні змінні:

$$\hat{y}(t) = F(x(t-1), x(t-2), ..., x(t-n)); \quad (7)$$

2) нелінійна авторегресійна мережа, що приймає в якості вхідних даних прогнозовані дані виходів:

$$\hat{y}(t) = F(y(t-1), y(t-2), ..., y(t-n)); \quad (8)$$

3) нелінійна авторегресійна нейромережа з екзогенними входами, яка поєднує як вхідні змінні, так і дані попередніх прогнозів:

$$\hat{y}(t) = F(y(t-1), y(t-2), ..., y(t-n), x(t-1), x(t-2), ..., x(t-n)). \quad (9)$$

На етапі налаштування параметрів моделі нейронної мережі здійснюється підбір вагових коефіцієнтів нейронної мережі $\hat{w}$ в результаті реалізації процедури навчання, яка полягає у відображенні множини даних моніторингу на множині вагів нейромережевої моделі $Z \to \hat{w}$ з метою отримання оптимального, за певним критерієм, прогнозного значення вихідного сигналу $\hat{y}(t)$.

Для навчання нейронної мережі множину даних про теплоспоживання Z слід розбити на три підмножини, що не перетинаються – множину даних для навчання Z_{tr}, тестову множину Z_t та множину для перевірки Z_v. Множини даних для навчання та перевірки формуються на основі даних моніторингу теплоспоживання будівлі закладу соціально-бюджетної сфери за декілька попередніх опалювальних сезонів. Тестування нейронної мережі, що пройшла навчання за обраним алгоритмом, відбувається на обмеженій кількості даних моніторингу, що відображають зміни теплоспоживання будівлі в залежності від кліматичних умов навколишнього середовища та графіку роботи закладу.

Для ідентифікації параметрів моделі нейроної мережі необхідно шляхом побудови нейронної мережі за даними навчальної вибірки та проведення обчислювального експерименту знайти оптимальні значення кількості нейронів прихованого шару L та множини значень вагових коефіцієнтів нейромережі $\hat{w}$, за яких похибка обчислення є найменшою. Для оцінки точності прогнозування слід обчислити помилки прогнозу та порівняти їх із допустимими значеннями.

4 ЕКСПЕРИМЕНТИ

Запропоновані моделі прогнозування теплоспоживання будівлі соціально-бюджетної сфери реалізовано у програмному середовищі Matlab Neural Network Toolbox. У якості об'єкта для тестування розроблених моделей прогнозування обрано систему теплоспоживання будівель Сумського державного університету. Архів даних для навчання нейронної мережі сформовано за допомогою підсистеми моніторингу інформаційно-аналітичної системи «HeatCAM» [11]. Вхідними даними для прогнозування є значення кількості теплової енергії, зібрані з лічильників, встановлених у корпусах університету, а також дані кліматичних умов навколишнього середовища та графіку роботи закладу.

У програмному середовищі Matlab Neural Network Toolbox побудовано архітектури моделей нейронних мереж типів *NIO*, *NAR* та *NARX*. Множину даних про теплоспоживання Z для навчання нейронної мережі було розбито у пропорції: Z_{tr} – 70%, Z_t – 20%, Z_v – 10%.

Шляхом проведення розрахункових експериментів встановлено кількість нейронів прихованого шару L (Number of hidden neuron) – 10, параметр затримки, що визначає число попередніх даних які будуть використовуватись у прогнозуванні n (Number of delays) – 7.

Для навчання даної нейронної мережі обрано метод нелінійної оптимізації Левенберга-Марквардта, який характеризується більш високою швидкістю збіжності у порівнянні з іншими градієнтними методами [10].

5 РЕЗУЛЬТАТИ

Результати точності прогнозу побудованих моделей для тестової вибірки даних наведено в табл. 2.

На рис. 1–3 зображено графіки фактичного значення теплоспоживання будівлі одного з корпусів Сумского державного університету та прогнозованого значення, одержаного при використанні моделей нейронної мережі типів *NIO*, *NAR* та *NARX* за три тижні січня 2013.

Результати обчислювальних експериментів з оцінки достовірності запропонованих моделей наведено в табл. 3.

6 ОБГОВОРЕННЯ

Як показують результати розрахунку точності моделей прогнозування для тестової вибірки даних, наведені в табл. 2, найбільшу точність має прогноз теплоспоживання з використанням моделі нелінійної нейронної авторегре-

Таблиця 2 – Порівняння точності моделей прогнозування типу *NAR*, *NIO* та *NARX*

Тип прогнозної моделі	Точність прогнозу, %
NAR	79,6
NIO	81,8
NARX	88,2

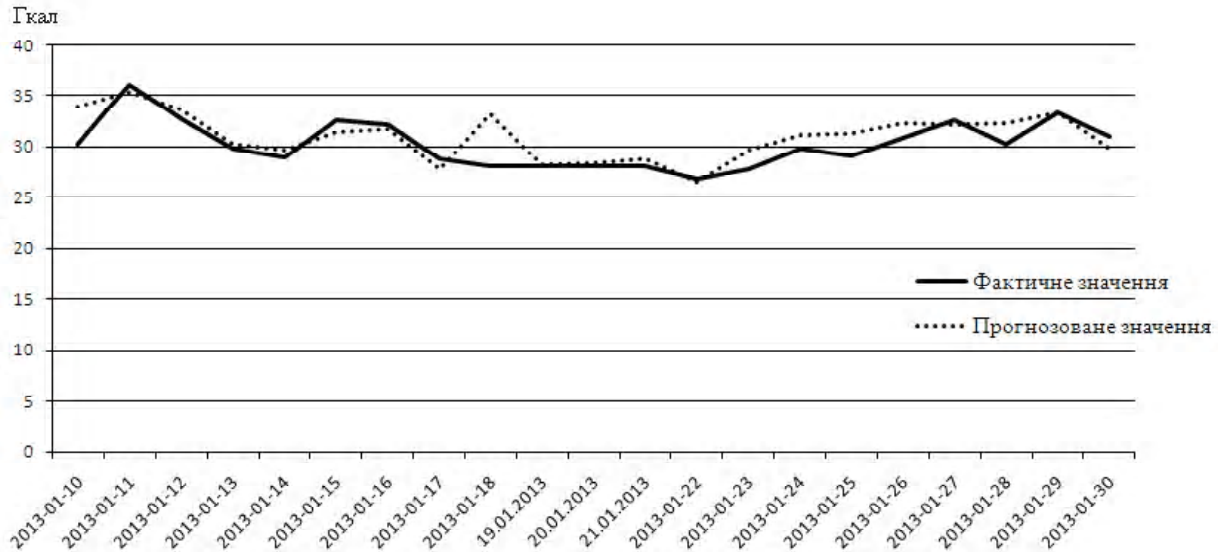


Рисунок 1 – Порівняння фактичних значень теплоспоживання будівлі та прогнозованих значень з використанням моделі нейронної мережі типу *NIO*

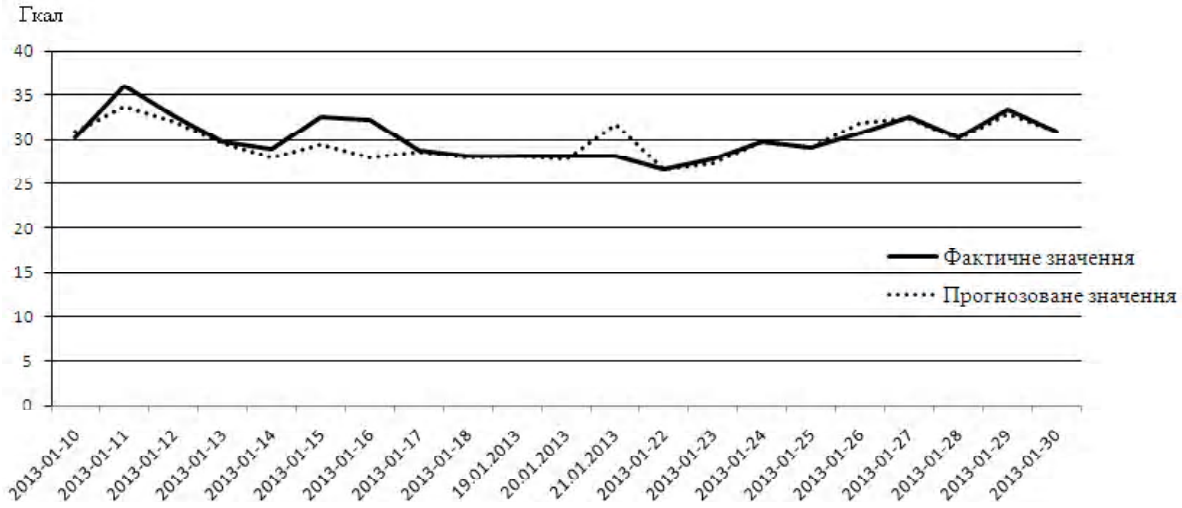


Рисунок 2 – Порівняння фактичних значень теплоспоживання будівлі та прогнозованих значень з використанням моделі нейронної мережі типу *NAR*

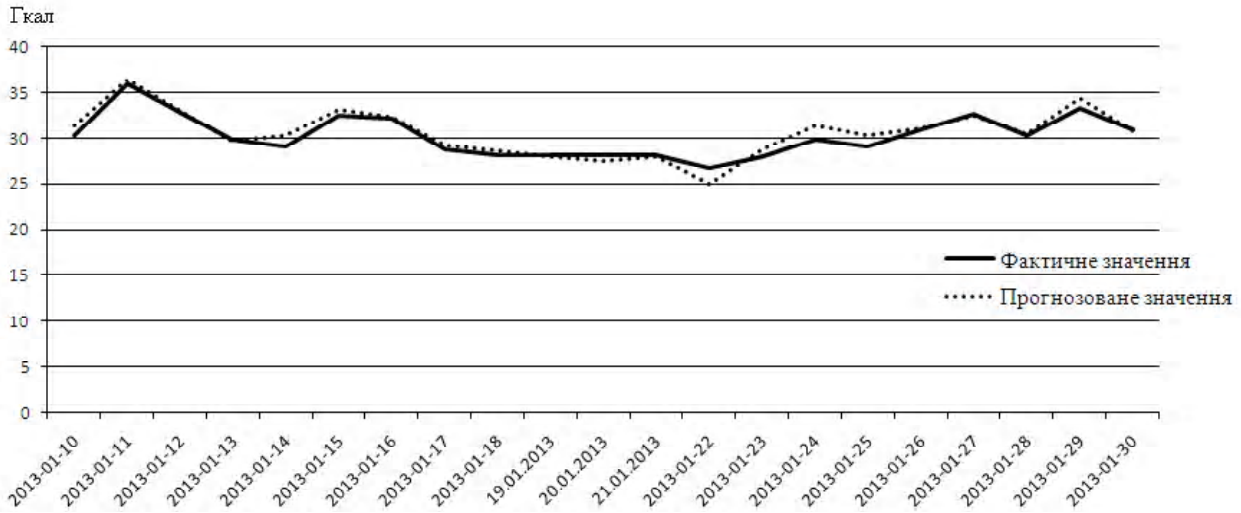


Рисунок 3 – Порівняння фактичних значень теплоспоживання будівлі та прогнозованих значень з використанням моделі нейронної мережі типу *NARX*

Таблиця 3 – Оцінка точності прогнозування теплоспоживання

Помилка	Тип моделі нейромережі	Грудень 2012 р.	Січень 2013 р.	Лютий 2013 р.
RMSE	*NIO*	6,15	1,74	0,89
	NAR	3,88	1,51	1,75
	NARX	2,19	0,88	0,66
MAE	*NIO*	4,69	1,25	0,37
	NAR	1,78	0,88	1,4
	NARX	1,06	0,68	0,47
MAPE, %	*NIO*	14,95	4,28	1,48
	NAR	5,16	2,89	5,67
	NARX	3,17	2,34	1,87

сійної мережі типу *NARX*, що враховує при прогнозуванні як результати попередніх прогнозованих значень, так і значення входів нейронної мережі. При проведенні обчислювальних експериментів з оцінки достовірності запропонованих моделей прогнозування теплоспоживання за формулами (3)–(6) обчислено помилки *RMSE*, *MAE* та *MAPE* (табл. 3). Для зими опалювального сезону 2012–2013 одержано значення середньої абсолютної помилки у відсотках *MAPE*, обчислене за формулою (4) для нейронної мережі типу *NARX* <5%, що свідчить про високу точність результатів прогнозування. Для нейронної мережі типу *NAR* помилка *MAPE* перевищує 5% для грудня 2012 р. та січня 2013 р. Значення помилки *MAPE* для грудня 2012 р. становить 14,95%, що свідчить про низьку точність даної моделі прогнозу. Як видно з графіків рис. 1–3 теплоспоживання в опалювальному сезоні 2012–2013 взимку коливається в межах від 20 до 40 Гкал на добу. Значення помилки *MAE*, обчислені за формулою (3), найменші для нейронної мережі типу *NARX* (не більше ніж 1,06 Гкал). Значення помилки *RMSE*, обчислені за формулою (6), також є найнижчим для нейронної мережі типу *NARX*. Таким чином результати проведених обчислювальних експериментів свідчать про те, що застосування нейронної мережі типу *NARX* при прогнозуванні теплоспоживання будівлі соціально-бюджетної сфери дозволяє досягти більшої точності прогнозу порівняно з нейронними мережами типів *NIO* та *NAR*.

ВИСНОВКИ

У роботі вирішено актуальну задачу прогнозування споживання теплової енергії закладами соціально-бюджетної сфери.

Наукова новизна роботи полягає в тому, що запропоновано модель прогнозування теплоспоживання будівлі закладу соціально-бюджетної сфери з використанням технології штучних нейронних мереж, яка враховує попередні значення теплоспоживання, параметри кліматичних умов навколишнього середовища та графік роботи закладу.

Проведено аналіз методів, що застосовуються при короткостроковому прогнозуванні теплоспоживання об'єктів соціально-бюджетної сфери, виявлено їх переваги та недоліки. Розроблено моделі прогнозування теплоспоживання будівель соціально-бюджетної сфери на основі нейронних мережі типу *NIO*, *NAR* та *NARX*, що враховують залежність рівня теплоспоживання від погодних умов, а також графіку роботи закладу. Моделі протестовано для прогнозування теплоспоживання одного з корпусів Сумського державого університету, проведено ідентифікацію параметрів запропонованих моделей. Для навчання нейронних мереж використано алгоритм Левенберга-Марквардта. Проведено експерименти з оцінки достовірності прогнозу теплоспоживання з використанням запропонованих моделей, які показали, що найбільшої точності прогнозу можна досягти при використанні моделі типу *NARX*. Врахування в розробленій моделі як факторів впливу кліматичних умов, так і коливання потреб у теплозабезпеченні в залежності від графіку роботи закладу дозволило підвищити точність прогнозування.

Практична цінність отриманих результатів полягає в тому, що на основі запропонованої моделі типу *NARX* розроблено підсистему прогнозування інформаційно-аналітичної системи моніторингу та прогнозування теплозабезпечення будівель «HeatCAM». Подальші дослідження пов'язано з використанням результатів прогнозування для підтримки прийняття рішень щодо управління режимами теплозабезпечення будівель закладу соціально-бюджетної сфери.

ПОДЯКИ

Роботу виконано в рамках науково-дослідної роботи Сумського державного університету за темою: «Науково-методологічні основи розробки інформаційних технологій для розподілених систем», номер державної реєстрації 0111U006119, а також комплексної цільової програми розвитку Сумського державного університету «Енергоефективний СумДУ».

СПИСОК ЛІТЕРАТУРИ

1. Системные исследования в энергетике: Ретроспектива научных направлений СЭИ-ИСЭМ / отв. ред. Н. И. Воропай. – Новосибирск : Наука, 2010. – 686 с.
2. Вороновский Г. К. Проблемы и перспективы использования искусственных нейронных сетей в энергетике. Ч. 2. Управление / Г. К. Вороновский, К. В. Махотило, С. А. Сергеев // Пробл. заг. енергетики. – 2007. – № 16. – С. 54–67.
3. Chramcov B. Heat Demand Forecasting for Concrete District Heating System / B. Chramcov, Ja. Baláte, M. Princ // International Journal of Mathematical Models and Methods in Applied Sciences. – 2010. – Vol. 4, No. 4. – P. 231–239.
4. Bacher P. Short-term heat load forecasting for single family houses / P. Bacher, H. Madsen // Energy and Buildings. – 2013. – Vol. 65. – P.101–112.
5. Kontu K. Forecasting district heating consumption based on customer measurements [Electronic resource] / K. Kontu // International association EUROHEAT & POWER. – Access mode: http://www.euroheat.org/Files/Filer/Presentations/20130527_EHP_Congress/130527_1600_1730_Pl_6_Kaisa Kontu.pdf.
6. Grosswindhager S. Online Short-Term Forecast of System heat Load in District Heating Networks / S. Grosswindhager, A. Voigt, M. Kozek // ISF 2011 : the 31st International Symposium on Forecasting, Prague, Czech Republic, 26–29 June 2011 : conference proceedings. – Lancaster : International Institute of Forecasters, 2011. – P. 1–8.
7. Dostál P. Prediction of the heat supply daily diagram via artificial neural network / P. Dostál, B. Chramcov, Ja.Baláte // East West Fuzzy Colloquium : proceedings. Zittau, Germany, 2002. – P. 178–183.
8. Buhari M. Short-Term Load Forecasting Using Artificial Neural Network / M. Buhari, S. S. Adamu // IMECS 2012 : International MultiConference of Engineers and Computer Scientist, Hong Kong, 14–16 Mar 2012 : proceedings. – Hong Kong, 2012.
9. Grzenda M. Consumer-oriented heat consumption prediction / M. Grzenda // Control and Cybernetics. – 2012. – Vol. 41, No. 1. – P. 213–240.
10. Haykin S. Neural Networks – A Comprehensive Foundation / S. Haykin. – NJ : Prentice Hal, 2005. – 823 p.

11. Parfenenko Yu. Information System for Monitoring and Forecast of Building Heat Consumption / Yu. Parfenenko, V. Shendryk, V. Nenja, S. Vashchenko // Communications in Computer and Information Science. – 2014. – Vol. 465.– P. 1–11.

Стаття надійшла до редакції 28.10.2014.
Після доробки 21.11.2014.

Парфененко Ю. В.[1], Шендрик В. В.[2], Галиченко А. С.[3]
[1]Ассистент кафедры компьютерних наук, Сумский государственный университет, г. Сумы, Украина
[2]Канд. техн. наук, доцент кафедры компьютерных наук, Сумский государственный университет, г. Сумы, Украина
[3]Студент факультета Электроники и информационных технологий, Сумский государственный университет, г. Сумы, Украина

ПРОГНОЗИРОВАНИЕ ТЕПЛОПОТРЕБЛЕНИЯ ЗДАНИЙ СОЦИАЛЬНО-БЮДЖЕТНОЙ СФЕРЫ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

Предложен метод усовершенствования процессов поддержки принятия решений при управлении режимами теплопотребления учреждений социально-бюджетной сферы путем разработки модели прогноза потребления тепловой энергии зданием. Объектом исследования является процесс выбора наиболее оптимальной архитектуры нейронной сети для решения задачи прогнозирования теплопотребления здания социально-бюджетной сферы. Предмет исследования составляют модели прогнозирования теплопотребления зданий социально-бюджетной сферы с использованием искусственных нейронных сетей. Целью данного исследования является повышение достоверности прогнозирования потребностей зданий социально-бюджетной сферы в теплообеспечении. Разработаны модели краткосрочного прогнозирования теплопотребления объектов социально-бюджетной сферы с использованием искусственных нейронных сетей, учитывающих влияние погодных условий, колебания потребности в тепловой энергии в зависимости от графика работы учереждения, а также предыдущие значения потребностей в теплообеспечении. Модели построены на основе архитектур нелинейной сети типа вход-выход, нелинейной авторегрессионной сети, нелинейной авторегрессионной нейронной сети с внешними входами. Предложенные модели реализованы в программной среде Matlab, для обучения использован алгоритм Левенберга-Марквардта. Проведены эксперименты по исследованию точности разработанных моделей, которые показали, что наибольшей точности прогноза можно достичь при использовании модели типа NARX. Использование предложенной модели в рамках информационной системы «HeatCAM» для прогнозирования теплопотребления зданий в течение отопительного сезона позволяет повысить эффективность поддержки процесса принятия решений при управлении режимами теплообеспечения, что приводит к снижению потребления тепловой энергии.

Ключевые слова: теплопотребление, регулирование, моделирование, информационная система, прогнозирование, нейронные сети, энергосбережение.

Parfenenko Yu. V.[1], Shendryk V. V.[2], Galichenko O. S.[3]
[1]Assistant of Computer Science Department, Sumy State University, Sumy, Ukraine
[2]PhD, Associate Professor of Computer Science Department, Sumy State University, Sumy, Ukraine
[3]Student of Electronics and Information Technologies Faculty, Sumy State University, Sumy, Ukraine

PREDICTION THE HEAT CONSUMPTION OF SOCIAL AND PUBLIC SECTOR BUILDINGS USING NEURAL NETWORKS

The method of improvement the process of decision support to improve management of heat supplies' modes through the development of prediction heat consumption model of the social and public sector building is proposed. The object of the study is the process of choosing the most optimal architecture of the neural network to solve goals of forecasting of heat consumption of the building of social and public sector. The subject of the study is models prediction heat consumption of buildings of social and public sector using artificial neural networks. The purpose of this study is to improve the forecasting reliability of heat energy demand of social and public sector buildings. Models of short-term prediction of heat energy demand of social and public sector buildings using artificial neural networks that take into account the influence of weather conditions, fluctuations in demand for thermal energy depending on the type of day of the week and the previous values of heat energy demand are proposed. Models are based on such architectures of neural network's as a nonlinear network input-output, a nonlinear autoregressive network, a nonlinear autoregressive neural network with external inputs. The proposed models were implemented in the programming environment Matlab, to train their the Levenberg-Marquard algorithm was used. Experiments on the study of the accuracy of the developed models, which showed that the greatest accuracy of prediction can be achieved by using a model type NARX, were carried. Using the proposed model within the information system «HeatCAM» for the prediction of building's heat consumption during the heating season can increase the effectiveness of decision support in the management of heating modes, which reduces heat consumption.

Keywords: heat consumption, management, modeling, information systems, prediction, neural networks, energy saving.

REFERENCES

1. Voropaj N. I. Sistemnye issledovanija v jenergetike: Retrospektiva nauchnyh napravlenij SEI-ISEM. Novosibirsk, Nauka, 2010, 686 p.
2. Voronovskij G. K., Mahotilo K. V., Sergeev S. A.. Problemy i perspektivy ispol'zovanija iskusstvennyh nejronnyh setej v jenergetike. Ch. 2. Upravlenie, Problemy obshhej jenergetiki, 2007, No. 16, pp. 54–67.
3. Chramcov B. Baláte Ja., Princ M. Heat Demand Forecasting for Concrete District Heating System, *International Journal of Mathematical Models and Methods in Applied Sciences,* 2010, Vol. 4, No. 4, pp. 231–239.
4. Bacher P., Madsen H. Short-term heat load forecasting for single family houses, *Energy and Buildings*, 2013, Vol. 65, pp.101–112.
5. Kontu K. Forecasting district heating consumption based on customer measurements [Electronic resource], *International association EUROHEAT & POWER*. Access mode: http://www.euroheat.org/Files/Filer/Presentations/20130527_EHP_Congress/130527_1600_1730_Pl_6_Kaisa Kontu.pdf.
6. Grosswindhager S., Voigt A., Kozek M., Online Short-Term Forecast of System heat Load in District Heating Networks, *ISF 2011 : the 31st International Symposium on Forecasting, Prague, Czech Republic, 26–29 June 2011 : proceedings*. Lancaster, International Institute of Forecasters, 2011, pp. 1–8.
7. Dostál P., Chramcov B., Baláte Ja. Prediction of the heat supply daily diagram via artificial neural network, *East West Fuzzy Colloquium : proceedings*. Zittau, Germany, 2002, pp. 178–183.
8. Buhari M., Adamu S. S. Short-Term Load Forecasting Using Artificial Neural Network, *IMECS 2012 : International MultiConference of Engineers and Computer Scientist, Hong Kong, 14–16 Mar 2012 : proceedings*. Hong Kong, 2012.
9. Grzenda M. Consumer-oriented heat consumption prediction, *Control and Cybernetics,* 2012, Vol.41, No. 1, pp. 213–240.
10. Haykin S. Neural Networks – A Comprehensive Foundation. NJ, Prentice Hal, 2005, 823 p.
11. Parfenenko Yu., Shendryk V., Nenja V., Vashchenko S. Information System for Monitoring and Forecast of Building Heat Consumption, *Communications in Computer and Information Science,* 2014, Vol. 465, pp. 1–11.

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

PROGRESSIV INFORMATICS TECHNOLOGIES

УДК 004.934

Имамвердиев Я. Н.[1], Сухостат Л. В.[2]

[1]*Канд. техн. наук, зав. отделом, Институт информационных технологий Национальной Академии Наук Азербайджана, Баку, Азербайджан*

[2]*Научный сотрудник, Институт информационных технологий Национальной Академии Наук Азербайджана, Баку, Азербайджан*

МЕТОД ОЦЕНКИ ПЕРИОДА ОСНОВНОГО ТОНА С ПРИМЕНЕНИЕМ ЭМПИРИЧЕСКОГО ВЕЙВЛЕТ ПРЕОБРАЗОВАНИЯ

Оценка периода основного тона речевого сигнала используется во многих важных областях применения речевых технологий. Однако среди существующих методов только некоторые могут работать для случая нелинейных и нестационарных сигналов. Основная причина состоит в том, что методы обнаружения периода основного тона основаны на предположении, что сам процесс речеобразования линеен, а речевые сигналы являются локально стационарными. Выбор алгоритма для оценки периода основного тона всегда состоит в нахождении компромисса между временем и разрешающей способностью по частоте, робастностью, задержкой и вычислительной сложностью. Целью данной работы является разработка метода оценки периода основного тона на основе эмпирического вейвлет преобразования. Метод построения семейства адаптированных вейвлетов состоит в предположении, что фильтры зависят от расположения информации в спектре анализируемого сигнала. Эмпирические вейвлеты определяются как полосовые фильтры на каждом отрезке речевого сигнала. Для определения периода основного тона в качестве признаков рассматривается мгновенная частота. Оператор разделения энергии Тигера-Кайзера используется для ее выделения. В работе приводится сравнение данного метода с другими алгоритмами оценки периода основного тона.

Ключевые слова: период основного тона, эмпирическое вейвлет преобразование, оператор разделения энергии Тигера-Кайзера, внутренняя модовая функция, мгновенная частота.

НОМЕНКЛАТУРА

AUTOC – Modified Autocorrelation Method;
MBSC – Multi-Band Summary Correlogram;
AMDF – Average magnitude difference function;
PEFAC – Pitch estimation Filter with Amplitude Compression;
IMF – Intrinsic Mode Function;
HHT – Hilbert-Huang Transform;
EMD – Empirical Mode Decomposition;
EWT – Empirical Wavelet Transform;
TKEO – Teager-Kaiser Energy Operator;
GPE – Gross Pitch Error;
SNR – signal-to-noise ratio;
MFPE – Mean Fine Pitch Error;
DESA – Discrete Energy Separation Algorithm;
DESA-1 – первый дискретный алгоритм разделения энергии;
DESA-2 – второй дискретный алгоритм разделения энергии;
AM-FM – Amplitude-Modulation Frequency-Modulation;
IF1 – мгновенная частота из первой IMF;
T_n – переходная фаза;
N_{FPE} – число вокализованных участков без GPE;
ω_n – границы между сегментами;
f_k – функция IMF;
$F0_{true}(i)$ – эталонное значение периода основного тона;
$F0_{est}(i)$ – полученное значение периода основного тона;
i – число фреймов;
N_{VU} – число фреймов, в которых полученное значение основного тона и эталонное значение определены как вокализованные;
$\alpha_j(t)$ – мгновенная амплитуда;
$\theta_j(t)$ – фазовая функция;
$H[c_j(t)]$ – преобразование Гильберта;
$z_j(t)$ – аналитический сигнал;
$X(t)$ – исходный сигнал;
$\mathrm{Re}\{\cdot\}$ – реальная часть комплексного выражения;

DOI 10.15588/1607-3274-2015-2-6

$H(\omega,t)$ – спектр Гильберта;

$\psi(\cdot)$ – оператор TKEO;

$x(t)$ – сигнал;

$\dot{x}(t)$ – производные первого порядка;

$\ddot{x}(t)$ – производные второго порядка;

$f(n)$ – мгновенная частота;

$|\alpha(n)|$ – мгновенная амплитуда;

F_0 – период основного тона;

N – число сегментов;

Λ_n – сегмент сигнала;

$C^k([0,1])$ – пространство k раз дифференцируемых функций на интервале [0,1];

$\beta(x)$ – функция из $C^k([0,1])$;

γ – параметр определяющий размер границ;

$\hat{\phi}_n(\omega)$ – эмпирическая масштабируемая функция;

$\hat{\psi}_n(\omega)$ – эмпирические вейвлеты;

$W_f^{\varepsilon}(0,t)$ – детализирующие коэффициенты;

$W_f^{\varepsilon}(0,t)$ – аппроксимирующие коэффициенты.

ВВЕДЕНИЕ

Начиная с 1970-х годов было разработано множество алгоритмов для выделения периода основного тона. Среди них можно выделить модифицированный автокорреляционный метод (AUTOC) [1], кепстральный метод, многополосную агрегацию коррелограмм (MBSC) [2], BaNa [3], YIN [4], YAAPT [5], среднее значение разностной функции (AMDF) [1], SWIPE' [6] и метод оценки основного тона на основе амплитудного сжатия (PEFAC) [7].

Методы оценки периода основного тона можно, в общем, разделить на временные, частотные и гибридные. Первая категория, в основном, ищет пики в автокорреляционной функции, вторая – в спектре мощности, тогда как третья группа выполняет частотно-временной анализ выхода банка полосовых фильтров. Во многих случаях алгоритмы выделяют кандидатов периода основного тона для каждого временного фрейма, а затем используют временные ограничения на непрерывность.

Однако только некоторые из них могут работать для случая нелинейных и нестационарных сигналов.

Некоторые хорошие альтернативы – методы на основе оценки параметров мгновенной частоты. Мгновенная частота была представлена в [8, 9], оценка методов предложена в [8–13]. Она необходима для понимания подробных механизмов обработки нелинейных и нестационарных процессов. На практике мгновенная частота обычно вычисляется из внутренней модовой функции (IMF) с помощью преобразования Гильберта-Хунга (HHT). HHT для получения IMF применяет эмпирическую модовую декомпозицию (EMD) [13]. В сравнении с методом автокорреляции данный подход более точен и быстрее обнаруживает период основного тона. Алгоритм EMD, применяемый при вычислении IMF, адаптивен, его вычислительная сложность, включая число извлеченных IMF и количество вычислений на этапе их получения, строго зависит от сложности самих речевых сигналов. Но метод EMD автоматически оценивает число мод, что существенно влияет на вычислительную сложность. В связи с этим в работе предлагается метод эмпирического вейвлет преобразования (EWT) [14].

Далее для получения мгновенной частоты применяется преобразование Гильберта. Однако оператор разделения энергии Тигера-Кайзера (TKEO) [15] превосходит его по вычислительной сложности и быстроте на реальных сигналах. Преобразование Гильберта и оператор TKEO могут быть применены только к монокомпонентным сигналам. В случае мультикомпонентных сигналов необходимо разбиение сигнала на простые компоненты перед применением методов. Для этого удобно использовать узкополосные фильтры [10]. Однако, в случае частотно-модулированных компонентов, это не всегда возможно в силу широкого диапазона частот.

Целью данной работы является разработка нового метода оценки периода основного тона на основе мгновенной частоты с применение EWT и нелинейного оператора TKEO. Приводятся результаты практических экспериментов на речевой базе данных Keele [16]. Оценивается робастность предложенного метода к шуму.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть рассматривается метод построения семейства вейвлетов адаптированных к обрабатываемым сигналам. Представлены реальные сигналы, где спектр симметричен относительно частоты $\omega = 0$, а также нормализованная ось Фурье, которая имеет периодичность для того, чтобы удовлетворить критерию Шеннона, и ограничивается отрезком $\omega \in [0,\pi]$.

Отрезок $[0,\pi]$ делится на N смежных сегментов. Каждый сегмент $\Lambda_n = [\omega_{n-1},\omega_n]$ (где $\omega_0 = 0$ и $\omega_N = \pi$). Вокруг ω_n определяется переходная фаза T_n шириной $2\tau_n$.

Тогда эмпирические вейвлеты [14] определяются как полосовые фильтры на каждом Λ_n. Банки фильтров строятся согласно условию $\gamma < \min_n\left(\frac{\omega_{n+1}-\omega_n}{\omega_{n+1}+\omega_n}\right)$. Далее определяются компоненты IMF.

После получения IMF компонент для вычисления мгновенной частоты и мгновенной амплитуды к каждой IMF применяется преобразование Гильберта [17] (дано: реальный сигнал с симметричным спектром относительно $\omega = 0$, надо: оценка периода основного тона на основе мгновенной частоты).

2 ЛИТЕРАТУРНЫЙ ОБЗОР

Вейвлет-анализ в настоящее время является одним из средств исследования данных. Он позволяет решать задачи моделирования, идентификации, аппроксимировать стационарные и нестационарные процессы. Преимущества вейвлетов включают частотно-временную локализацию сигналов.

Нестационарные и нелинейные сигналы требуют адаптивные пути решения. Адаптивность состоит в рассмотрении тех областей сигнала, где содержится полезная информация. Вейвлет преобразования способны собирать основную информацию о сигнале в небольшое число вейвлет коэффициентов. Полезное свойство вейвлет преобразования состоит и в том, что его можно рассматривать как применение банка фильтров.

Известен подход, называемый вейвлетами Малвара-Вильсона [18, 19], который пытается построить адаптивное представление путем сегментации сигнала во временной области, чтобы выделить временные интервалы, содержащие спектральную информацию о сигнале. В то время как данный подход интересен, временная сегментация является трудно реализуемой задачей.

В [20] авторы предлагают метод, называемый «brushlets», который строит адаптивный банк фильтров в области Фурье. Он использует идею вейвлетов Малвара-Вильсона, но сегментирует спектр Фурье вместо самого сигнала. Данная задача является конструктивно довольно сложной.

В недавней работе И. Добеши под названием «synchrosqueezed всплески» [21] предлагается подход, сочетающий в себе классический вейвлет-анализ и метод перераспределения частотно-временной информации. Этот алгоритм позволяет получить более точное частотно-временное представление.

В основу каждого метода положено классическое вейвлет преобразование, но каждый из них имеет свои индивидуальные особенности.

3 МАТЕРИАЛЫ И МЕТОДЫ

Эмпирическая масштабируемая функция и эмпирические вейвлеты определяются следующим образом

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

и

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

Функция $\beta(x)$, произвольная из C^k ([0,1]) – функция такая, что

$$\beta(x)=\begin{cases}0, & \text{если} \quad x\le 0 \text{ и } \beta(x)+\beta(1-x)=1,\ \forall x\in[0,1];\\ 1, & \text{если} \quad x\ge 1,\end{cases}$$

τ_n выбирается пропорционально ω_n: $\tau_n=\gamma\omega_n$, где $0<\gamma<1$. Следовательно, для всех $n>0$ уравнения (1) и (2) принимают вид

$$\hat{\varphi}_n(\omega)=\begin{cases}1, & \text{если} \quad |\omega|\le(1-\gamma)\omega_n;\\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}\left(|\omega|-(1-\gamma)\omega_n\right)\right)\right], & \\ & \text{если } (1-\gamma)\omega_n\le|\omega|\le(1+\gamma)\omega_n;\\ 0, & \text{в противном случае}\end{cases} \tag{3}$$

и

$$\hat{\psi}_n(\omega)=\begin{cases}1, & \text{если} \quad (1+\gamma)\omega_n\le|\omega|\le(1-\gamma)\omega_{n+1};\\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_{n+1}}\left(|\omega|-(1-\gamma)\omega_{n+1}\right)\right)\right], & \\ & \text{если } (1-\gamma)\omega_{n+1}\le|\omega|\le(1+\gamma)\omega_{n+1};\\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}\left(|\omega|-(1-\gamma)\omega_n\right)\right)\right], & \\ & \text{если } (1-\gamma)\omega_n\le|\omega|\le(1+\gamma)\omega_n;\\ 0, & \text{в противном случае.}\end{cases} \tag{4}$$

Теперь можем определить эмпирическое вейвлет преобразование $W_f^{\varepsilon}(n,t)$ также как и в случае классического вейвлет преобразования:

$$W_f^{\varepsilon}(n,t)=\langle f,\psi_n\rangle=\int f(\tau)\overline{\psi_n(\tau-t)}d\tau=\hat{f}(\omega)*\overline{\hat{\psi}_n(\omega)},$$

а аппроксимирующие коэффициенты – скалярных произведений с масштабируемой функцией

$$W_f^{\varepsilon}(0,t)=\langle f,\phi_1\rangle=\int f(\tau)\overline{\phi_1(\tau-t)}d\tau=\hat{f}(\omega)*\overline{\hat{\phi}_1(\omega)},$$

где $\hat{\psi}_n(\omega)$ и $\hat{\phi}_1(\omega)$ определяются из уравнений (3) и (4) соответственно. Обратное преобразование принимает вид

$$f(t)=W_f^{\varepsilon}(0,t)*\phi_1(t)+\sum_{n=1}^{N}W_f^{\varepsilon}(n,t)*\psi_n(t)=$$

$$=\hat{W}_f^{\varepsilon}(0,\omega)*\hat{\phi}_1(\omega)+\sum_{n=1}^{N}\hat{W}_f^{\varepsilon}(n,\omega)*\hat{\psi}_n(\omega).$$

Функция IMF f_k определяется следующим образом:

$$f_0(t)=W_f^{\varepsilon}(0,t)*\phi_1(t),$$

$$f_k(t)=W_f^{\varepsilon}(k,t)*\psi_k(t).$$

Использование преобразования Гильберта позволяет получить для каждого момента времени

$$H[c_j(t)]=\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{c_j(\tau)}{t-\tau}d\tau.$$

Теперь можно построить аналитический сигнал $z_j(t)$ как

$$z_j(t)=c_j(t)+i\,H[c_j(t)],$$

который может быть представлен в виде

$$z_j(t) = \alpha_j(t)\exp(i\omega_j(t)).$$

Мгновенная амплитуда $\alpha_j(t)$ и фазовая функция $\theta_j(t)$ определяются как

$$\alpha_j(t) = \sqrt{c_j^2(t) + H^2[c_j(t)]},$$

$$\theta_j(t) = \arctan\frac{H[c_j(t)]}{c_j(t)}.$$

Кроме того, мгновенная частота $\omega_j(t)$ может быть вычислена следующим образом:

$$\omega_j(t) = \frac{d\theta_j(t)}{dt}.$$

Так исходный сигнал может быть представлен в форме (5):

$$X(t) = \mathrm{Re}\sum_{j=1}^{n}\alpha_j(t)\exp[i\int\omega_j(t)dt], \quad (5)$$

где остаток опущен.

Выражение (5) позволяет представить мгновенную амплитуду и частоту в трехмерном пространстве, где амплитуда – это высота в частотно-временной плоскости. Это частотно-временное распределение представлено как спектр Гильберта $H(\omega,t)$

$$H(\omega,t) = \mathrm{Re}\sum_{j=1}^{n}\alpha_j(t)\exp[i\int\omega_j(t)dt].$$

ННТ удовлетворяет требованию адаптивности для анализа нестационарных сигналов. Таким образом, сигнал может быть локально и точно отображен во временной частотной области путем применения спектра Гильберта.

Оператор разделения энергии Тигера-Кайзера – нелинейный оператор, который успешно применяется во многих инженерных приложениях [15]. Он обнаруживает модуляцию энергии и определяет мгновенную частоту и мгновенную амплитуду от AM-FM сигнала [9]. Оператор TKEO $\psi(\cdot)$ для сигнала $x(t)$ определяется как

$$\psi[x(t)] = [\dot{x}(t)]^2 - x(t)\ddot{x}(t).$$

Оператор разделения энергии Тигера-Кайзера – нелинейный оператор, который успешно применяется во многих инженерных приложениях [15]. Он обнаруживает модуляцию энергии и определяет мгновенную частоту и мгновенную амплитуду от AM-FM сигнала [9]. Оператор TKEO $\psi(\cdot)$ для сигнала $x(t)$ определяется как

$$\psi[x(t)] = [\dot{x}(t)]^2 - x(t)\ddot{x}(t).$$

В дискретно-временной области оператор принимает вид

$$\psi[x(n)] = x^2(n) - x(n+1)\cdot x(n-1).$$

Мгновенная частота $f(n)$ и мгновенная амплитуда $|\alpha(n)|$ в любой момент времени для сигнала $x(n)$ даются как

$$y(n) = x(n) - x(n-1), \quad (6)$$

$$f(n) = \arccos\left(1 - \frac{\psi[y(n)] + \psi[y(n+1)]}{4\psi[x(n)]}\right), \quad (7)$$

$$|\alpha(n)| = \sqrt{\frac{\psi[x(n)]}{\sin^2[f(n)]}}, \quad (8)$$

$$f(n) = \frac{1}{2}\arccos\left(1 - \frac{\psi[x(n+1) - x(n-1)]}{2\psi[x(n)]}\right), \quad (9)$$

$$|\alpha(n)| = \frac{2\psi[x(n)]}{\sqrt{\psi[x(n+1) - x(n-1)]}}. \quad (10)$$

В общем, метод демодуляции (6)–(8) известен как DESA-1, а метод (9) и (10) – как DESA-2. Алгоритм DESA-2 только требует три значения для вычисления энергии в каждый момент времени и более прост в вычислении. Поэтому в данной работе мы рассматриваем алгоритм DESA-2.

4 ЭКСПЕРИМЕНТЫ

Для проведения экспериментов была рассмотрена речевая база данных Keele [16]. Речевые образцы получены от 10 дикторов (5 мужчин и 5 женщин). Также содержатся записи, извлеченные из ларингографа (laryngograph) с помощью алгоритма автокорреляции и сопровождаются измерениями F_0.

Для тестирования устойчивости к внешним шумам рассматриваемых алгоритмов определения периода основного тона к сигналам добавляется белый шум с различными уровнями SNR. Для генерации зашумленной речи с определенным значением SNR, энергия сигнала вычисляется только на вокализованных участках речевого сигнала, и шум усиливается или ослабевает до определенного уровня, чтобы удовлетворить значению целевого SNR.

Период основного тона был оценен с помощью автокорреляции с окном 26,5 мс и сдвигом в 10 мс.

Для сравнения производительности методов обнаружения периода основного тона, используются следующие метрики ошибок [17]:

а) процент грубых ошибок (GPE) определяет соотношение кадров, рассчитанное на основе вокализованных участков полученного периода основного тона и эталонных значений (ground truth), для которых относительная погрешность оценки выше, чем определенный порог δ (обычно 20% [22]):

$$GPE = \frac{N_{F0E}}{N_{VV}} \times 100\%,$$

где N_{F0E} – число фреймов, для которых выполняется условие

$$\left|\frac{F0_{est}(i)}{F0_{true}(i)} - 1\right| > \delta\%;$$

б) средний процент мелких ошибок (MFPE) вычисляется на вокализованных участках, где не наблюдаются ошибки GPE [23]:

$$MFPE = \frac{1}{N_{FPE}} \sum_{i=1}^{N_{FPE}} \frac{\left|F0_{true}(i) - F0_{est}(i)\right|}{F0_{true}(i)} \times 100\%,$$

где N_{FPE} – число вокализованных участков без GPE.

EWT разлагает сегменты сигнала на серию IMF функций для дальнейшего извлечения мгновенной частоты (рис. 1). Из всего набора IMF выбираем ту, чей период ближе к исходному сигналу. Из рис. 1 видно, что IMF1 содержит информацию о периоде основного тона и хорошо отображает форму сигнала. Блок-схема процесса на основе EWT-TKEO показана на рис. 2 (здесь IF1 обозначает мгновенную частоту получаемую из первой IMF).

5 РЕЗУЛЬТАТЫ

Эксперименты проводились в среде Matlab 2011b [24]. Сравнение производительности предложенного метода и других популярных алгоритмов обнаружения периода основного тона приводится в таблице 1.

Как показано в таблице 1, GPE для предложенного метода на основе EWT по сравнению с другими методами извлечения периода основного тона значительно меньше, чем у методов HHT, YIN и SWIPE' при различных уровнях шума. Предложенный метод является более робастным по сравнению с другими методами и лучше работает при высоких SNR.

6 ОБСУЖДЕНИЕ

Предложенный метод оценки периода основного тона на основе эмпирического вейвлет преобразования обеспечивает существенную точность выделения высоты. Однако данный подход требует больших временных затрат.

По сравнению с другими алгоритмами оценки высоты, такими как YAAPT, HHT и SWIPE', предложенный метод более точен. Метод был протестирован на устойчивость к внешним шумам.

Эффективность предложенного подхода может быть повышена при увеличении числа признаков. Однако это значительно влияет на скорость работы алгоритма.

Рисунок 2 – Схема оценки периода основного тона

Таблица 1 – Сравнение производительности методов извлечения периода основного тона

Метод \ SNR (dB)		–5	0	10	15
SWIPE'	GPE	20,82	14,28	11,92	8,43
	MFPE	1,77	1,21	0,90	0,83
YAAPT	GPE	13,90	23,35	9,98	6,26
	MFPE	2,10	1,94	1,51	0,81
HHT	GPE	19,30	6,02	11,34	5,10
	MFPE	0,56	0,43	0,35	0,33
EWT-TKEO	GPE	15,11	5,77	10,56	4,79
	MFPE	0,37	0,30	0,28	0,14

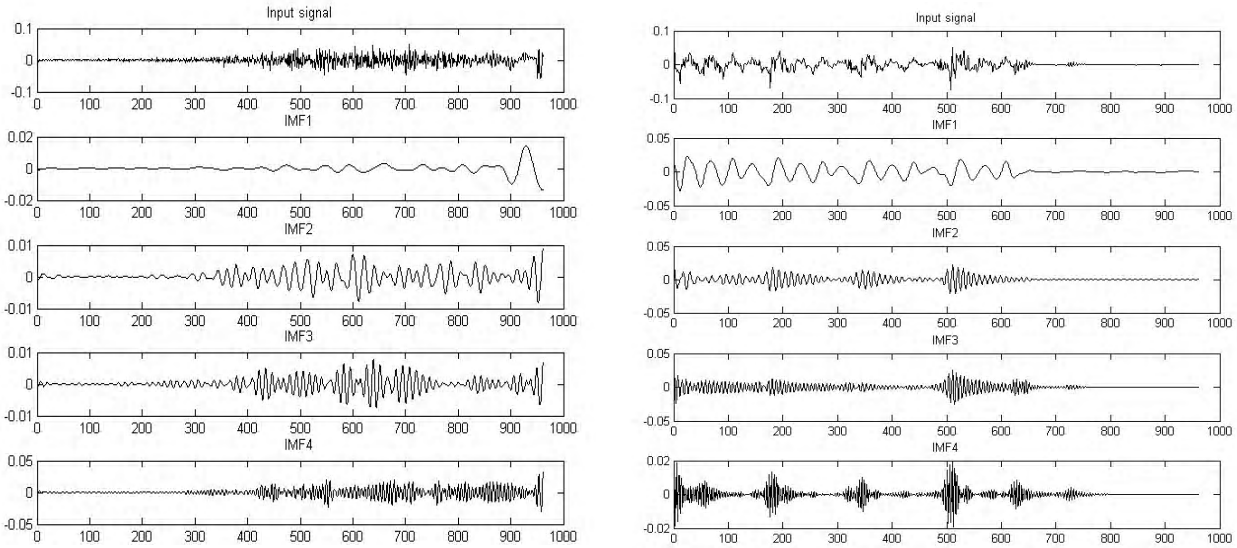


Рисунок 1 – Декомпозиция сигналов с помощью эмпирического вейвлет преобразования

ВЫВОДЫ

Решена актуальная задача разработки метода выделения периода основного тона речевого сигнала. Идея предлагаемого подхода состоит в том, что мгновенная частота содержит информацию о периоде основного тона. Был рассмотрен метод EWT. Для выделения мгновенной частоты был предложен оператор ТКЕО. Для проведения экспериментов белый шум был добавлен к речевому сигналу перед применением метода извлечения периода основного тона. Алгоритм протестирован при различных уровнях шума. Было показано, что точность алгоритма выше, чем у алгоритмов YAAPT, HHT и SWIPE' в случае быстрых модуляций высоты. Эксперименты показывают достаточную эффективность предлагаемого подхода при аддитивном шуме.

БЛАГОДАРНОСТИ

Данная работа выполнена при финансовой поддержке Фонда Развития Науки при Президенте Азербайджанской Республики – Грант № EƏF-RƏTN-MQM-2/ƏKT-2-2013-7(13)-29/18/1.

СПИСОК ЛИТЕРАТУРЫ

1. A comparative performance study of several pitch detection algorithms / [L. Rabiner, M. J. Cheng, A. E. Rosenberg, C. A. McGonegal] // IEEE Transactions on Acoustics, Speech and Signal Processing. – 1976. – № 5. – P. 399–417. DOI: 10.1109/TASSP.1976.1162846
2. Tan L. N. Multi-band summary correlogram-based pitch detection for noisy speech / L. N. Tan, A. Alwan // Speech Communication. – 2013. – Vol. 55, № 78. – P. 841–856. DOI: 10.1016/j.specom.2013.03.001
3. BaNa: a hybrid approach for noise resilient pitch detection / [H. Ba, N. Yang, I. Demirkol, W. Heinzelman] // IEEE Statistical Signal Processing Workshop. – 2012. – P. 369–372. DOI: 10.1109/SSP.2012.6319706
4. De Cheveigne A. Yin, a fundamental frequency estimator for speech and music / A. De Cheveigne, H. Kawahara // Journal of the Acoustical Society of America. – 2002. – Vol. 111, № 4. – P. 1917–1930. DOI: 10.1121/1.1458024
5. Kasi K. Yet another algorithm for pitch tracking / K. Kasi, S. A. Zahorian // Proceedings of the International Conference on Acoustics, Speech, and Signal Processing. – 2002. – Vol. 1. – P. 361–364. DOI: 10.1109/ICASSP.2002.5743729
6. Camacho A. SWIPE: a sawtooth waveform inspired pitch estimator for speech and music: thesis … doctor of philosophy / Camacho A. – Florida: University of Florida, 2007.
7. Gonzalez S. PEFAC – A Pitch Estimation Algorithm Robust to High Levels of Noise / S. Gonzalez, M. Brookes // IEEE Transactions on Audio, Speech and Language Processing. – 2011. – Vol. 22, № 2. – P. 518–530. DOI: 10.1109/TASLP.2013.2295918
8. Boashash B. Estimating and interpreting the instantaneous frequency of a signal / B. Boashash // Proceedings of the IEEE. – 1992. – Vol. 80, № 4. – P. 520–538. DOI: 10.1109/5.135376
9. Maragos P. On amplitude and frequency demodulation using energy operators / P. Maragos, J. F. Kaiser, T. F. Quatieri // IEEE Transactions on Signal Processing. – 1993. – Vol. 41, № 4. – P. 1532–1550. DOI: 10.1109/78.212729
10. Abe T. Harmonics tracking and pitch extraction based on instantaneous frequency / T. Abe, T. Kobayashi, S. Imai // Proceedings of the International Conference on Acoustics, Speech, and Signal Processing. – 1995. – Vol. 1. – P. 756–759. DOI: 10.1109/ICASSP.1995.479804
11. Abe T. Sinusoidal model based on instantaneous frequency attractors / T. Abe, M. Honda // IEEE Transactions on Audio, Speech and Language Processing. – 2006. Vol. 14, № 4. – P. 1292–1300. DOI: 10.1109/TSA.2005.858545
12. Azarov E. Estimation of the instantaneous harmonic parameters of speech / E. Azarov, A. Petrovsky, M. Parfieniuk // Proceedings of the European Signal Processing Conference. – 2008.
13. The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis / [N. E. Huang, Z. Shen, S. R. Long and other] // Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. – 1998. – Vol. 454, № 1971. – P. 903–995. DOI: 10.1098/rspa.1998.0193
14. Gilles J. Empirical Wavelet Transform / J. Gilles // IEEE Transactions on Signal Processing. – 2013. – Vol. 61, № 16. – P. 3999–4010. DOI: 10.1109/TSP.2013.2265222
15. Vakman D. On the analytic signal, the Teager–Kaiser energy algorithm, and other methods for defining amplitude and frequency / D. Vakman // IEEE Transactions on Signal Processing. – 1996. – Vol. 44, № 4. – P. 791–797. DOI: 10.1109/78.492532
16. Chu W. Reducing f0 frame error of f0 tracking algorithms under noisy conditions with an unvoiced/voiced classification frontend / W. Chu, A. Alwan // Proceedings of the International Conference on Acoustics, Speech, and Signal Processing. – 2009. – P. 3969–3972. DOI: 10.1109/icassp.2009.4960497
17. Varga A. Assessment for automatic speech recognition: II. Noisex-92: a database and an experiment to study the effect of additive noise on speech recognition systems / A. Varga, H. J. Steeneken // Speech Communication. – 1993. – Vol. 12, № 3. – P. 247–251. DOI: 10.1016/0167-6393(93)90095-3
18. Jaffard S. Wavelets: Tools for Science and Technology / S. Jaffard, Y. Meyer, R. D. Ryan. – Philadelphia: SIAM, 2001. – 256 p. DOI: 10.1063/1.1485591
19. Malvar H. S. Lapped transforms for efficien transform/subband coding / H. S. Malvar // IEEE Transactions on Acoustics, Speech and Signal Processing. – 1990. – Vol. 38, № 6. – P. 969–978. DOI: 10.1109/29.56057
20. Meyer F. G. Brushlets: A tool for directional image analysis and image compression / F. G. Meyer, R. R. Coifman // Journal of Applied and Computational Harmonic Analysis. – 1997. – Vol. 4, № 2. – P. 147–187. DOI: 10.1006/acha.1997.0208
21. Daubechies I. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool / I. Daubechies, J. Lu, H.-T. Wu / / Journal of Applied and Computational Harmonic Analysis. – 2011. –Vol. 30, № 2. – P. 243–261. DOI: 10.1016/j.acha.2010.08.002
22. Drugman T. Joint robust voicing detection and pitch estimation based on residual harmonics / T. Drugman, A. Alwan // Proceedings of Interspeech. – 2011. – P. 1973–1976.
23. Azarov E. Instantaneous pitch estimation based on RAPT framework / E. Azarov, M. Vashkevich, A. Petrovsky // Proceedings of the European Signal Processing Conference. – 2012. – P. 2787–2791.
24. Empirical Wavelet Transforms [Electronic resource]. – 2013. – Access mode: http://www.mathworks.com/matlabcentral/fileexchange/42141-empirical-wavelet-transforms

Статья поступила в редакцию 22.12.2014.
После доработки 12.01.2015.

Імамвердієв Я. Н.[1], Сухостат Л. В.[2]
[1]Канд. техн. наук, зав. відділом, Інститут інформаційних технологій Національної Академії Наук Азербайджану, Баку, Азербайджан
[2]Науковий співробітник, Інститут інформаційних технологій Національної Академії Наук Азербайджану, Баку, Азербайджан

МЕТОД ОЦІНКИ ПЕРІОДУ ОСНОВНОГО ТОНУ З ЗАСТОСУВАННЯМ ЕМПІРИЧНОГО ВЕЙВЛЕТ ПЕРЕТВОРЕННЯ

Оцінка періоду основного тону мовного сигналу використовується в багатьох важливих областях застосування мовних технологій. Однак серед існуючих методів тільки деякі можуть працювати для випадку нелінійних і нестаціонарних сигналів. Основна причина полягає у тому, що методи виявлення періоду основного тону засновані на припущенні, що сам процес мовоутворення є лінійним, а мовні сигнали є локально стаціонарними. Вибір алгоритму для оцінки періоду основного тону завжди полягає у знаход-

женні компромісу між часом і роздільною здатністю за частотою, робастністю, затримкою й обчислювальною складністю. Метою даної роботи є розробка методу оцінки періоду основного тону на основі емпіричного вейвлет перетворення. Метод побудови сімейства адаптованих вейвлетів полягає у припущенні, що фільтри залежать від розташування інформації в спектрі аналізованого сигналу. Емпіричні вейвлети визначаються як смугові фільтри на кожному відрізку мовного сигналу. Для визначення періоду основного тону як ознаки розглядається миттєва частота. Оператор поділу енергії Тігера-Кайзера використовується для її виділення. У роботі приводиться порівняння даного методу з іншими алгоритмами оцінки періоду основного тону.

Ключові слова: період основного тону, емпіричне вейвлет перетворення, оператор поділу енергії Тігера-Кайзера, внутрішня модовая функція, миттєва частота.

Imamverdiyev Y. N.[1], Sukhostat L. V.[2]

[1]PhD, Head of department, Institute of Information Technology of Azerbaijan National Academy of Sciences, Baku, Azerbaijan

[2]Research fellow, Institute of Information Technology of Azerbaijan National Academy of Sciences, Baku, Azerbaijan

PITCH PERIOD ESTIMATION METHOD USING EMPIRICAL WAVELET TRANSFORM

Pitch period evaluation of speech signal is used in many important applications of speech technology. However, among the existing methods only some can work in case of non-linear and non-stationary signals. The main reason is that the pitch detection methods are based on the assumption that speech production process is linear. Selection of pitch period estimation algorithm is always focuses on finding a compromise between time and frequency resolution, robustness, computational complexity and time delay. The aim of this paper is to develop a new method for estimating the pitch period based on empirical wavelet transformation. Method of constructing a family of adapted wavelets assumes that the filters depend on the information location in speech spectrum of the analyzed signal. Empirical wavelets are defined as bandpass filters for each segment of the speech signal. Instantaneous frequency characteristics are considered as pitch period detection features. Teager-Kaiser energy separation operator is used for its extraction. The comparison of this method with other pitch estimation algorithms is presented.

Keywords: pitch period, empirical wavelet transform, operator Teager-Kaiser energy operator, intrinsic mode function, instantaneous frequency.

REFERENCES

1. Rabiner L., Cheng M. J., Rosenberg A. E., McGonegal C. A. A comparative performance study of several pitch detection algorithms, *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1976, No. 5, pp. 399–417. DOI: 10.1109/TASSP.1976.1162846
2. Tan L. N., Alwan A. Multi-band summary correlogram-based pitch detection for noisy speech, *Speech Communication*, 2013, Vol. 55, No. 78, pp. 841–856. DOI: 10.1016/j.specom.2013.03.001
3. Ba H., Yang N., Demirkol I., Heinzelman W. BaNa: a hybrid approach for noise resilient pitch detection, *IEEE Statistical Signal Processing Workshop*, 2012, pp. 369–372. DOI: 10.1109/SSP.2012.6319706
4. De Cheveigne A., Kawahara H. Yin, a fundamental frequency estimator for speech and music, *Journal of the Acoustical Society of America*, 2002, Vol. 111, No. 4, pp. 1917–1930. DOI: 10.1121/1.1458024
5. Kasi K., Zahorian S. A. Yet another algorithm for pitch tracking, *Proceedings of the International Conference on Acoustics, Speech, and Signal,* Vol. 1, 2002, pp. 361–364. DOI: 10.1109/ICASSP.2002.5743729
6. Camacho A. SWIPE: a sawtooth waveform inspired pitch estimator for speech and music: thesis … doctor of philosophy. Florida, University of Florida, 2007.
7. Gonzalez S., Brookes M. PEFAC – A Pitch Estimation Algorithm Robust to High Levels of Noise, *IEEE Transactions on Audio, Speech and Language Processing,* 2011, Vol. 22, No. 2, pp. 518–530. DOI: 10.1109/TASLP.2013.2295918
8. Boashash B. Estimating and interpreting the instantaneous frequency of a signal, *Proceedings of the IEEE,* 1992, Vol. 80, No. 4, pp. 520–538. DOI: 10.1109/5.135376
9. Maragos P., Kaiser J. F., Quatieri T. F. On amplitude and frequency demodulation using energy operators, *IEEE Transactions on Signal Processing,* 1993, Vol. 41, No. 4, pp. 1532–1550. DOI: 10.1109/78.212729
10. Abe T. Kobayashi T., Imai S. Harmonics tracking and pitch extraction based on instantaneous frequency, *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, 1995, Vol. 1, pp. 756–759. DOI: 10.1109/ICASSP.1995.479804
11. Abe T., Honda M. /Sinusoidal model based on instantaneous frequency attractors, *IEEE Transactions on Audio, Speech and Language Processing*, 2006,Vol. 14, No. 4, pp. 1292–1300. DOI: 10.1109/TSA.2005.858545
12. Azarov E., Petrovsky A., Parfieniuk M. Estimation of the instantaneous harmonic parameters of speech, *Proceedings of the European Signal Processing Conference,* 2008.
13. Huang N. E., Shen Z., Long S. R., Wu M. L., Shih H. H., Zheng Q., Yen N. C., Tung C. C., Liu H. H. The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 1998, Vol. 454, No. 1971, P. 903–995. DOI: 10.1098/rspa.1998.0193
14. Gilles J. Empirical Wavelet Transform, *IEEE Transactions on Signal Processing,* 2013, Vol. 61, No. 16, pp. 3999–4010. DOI: 10.1109/TSP.2013.2265222
15. Vakman D. On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency, *IEEE Transactions on Signal Processing,* 1996, Vol. 44, No. 4, pp. 791–797. DOI: 10.1109/78.492532
16. Chu W., Alwan A. Reducing f0 frame error of f0 tracking algorithms under noisy conditions with an unvoiced/voiced classification frontend, *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, 2009, pp. 3969–3972. DOI: 10.1109/icassp.2009.4960497
17. Varga A., Steeneken H. J. Assessment for automatic speech recognition: II. Noisex-92: a database and an experiment to study the effect of additive noise on speech recognition systems, *Speech Communication,* 1993, Vol. 12, No. 3, pp. 247–251. DOI: 10.1016/0167-6393(93)90095-3
18. Jaffard S., Meyer Y., Ryan R. D. Wavelets: Tools for Science and Technology, *Philadelphia: SIAM, 2001,* 256 p. DOI: 10.1063/1.1485591
19. Malvar H. S. Lapped transforms for efficien transform/subband coding, *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1990, Vol. 38, No. 6, pp. 969–978. DOI: 10.1109/29.56057
20. Meyer F. G., Coifman R. R. Brushlets: A tool for directional image analysis and image compression, *Journal of Applied and Computational Harmonic Analysis,* 1997, Vol. 4, No. 2, pp. 147–187. DOI: 10.1006/acha.1997.0208
21. Daubechies I., Lu J., Wu H.-T. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool, *Journal of Applied and Computational Harmonic Analysis*, 2011, Vol. 30, No. 2, pp. 243–261. DOI: 10.1016/j.acha.2010.08.002
22. Drugman T., Alwan A. Joint robust voicing detection and pitch estimation based on residual harmonics, Proceedings of Interspeech, 2011, pp. 1973–1976.
23. Azarov E., Vashkevich M., Petrovsky A. Instantaneous pitch estimation based on RAPT framework, *Proceedings of the European Signal Processing Conference,* 2012, pp. 2787–2791.
24. Empirical Wavelet Transforms [Electronic resource], 2013, Access mode: http://www.mathworks.com/matlabcentral/fileexchange/42141-empirical-wavelet-transforms

УДК 004.891.3: 004.3

Красій А. В.[1], Говорущенко Т. О.[2]

[1]Аспірант кафедри системного програмування Хмельницького національного університету, Хмельницький, Україна
[2]Канд. техн. наук, доцент, старший науковий співробітник, доцент кафедри системного програмування Хмельницького національного університету, Хмельницький, Україна

МЕТОД ПРОГНОЗУВАННЯ УСПІШНОСТІ ПРОГРАМНИХ ПРОЕКТІВ НА ОСНОВІ АНАЛІЗУ ПРОГНОЗОВАНИХ ЗНАЧЕНЬ ХАРАКТЕРИСТИК ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Вирішено завдання розроблення методологічного забезпечення для прогнозування успішності програмних проектів. Об'єктом дослідження є процес прогнозування успішності програмних проектів на ранніх етапах життєвого циклу. Предметом дослідження є метод прогнозування успішності програмних проектів на основі аналізу прогнозованих значень характеристик програмного забезпечення. Мета роботи – підвищити достовірність вибору програмних проектів та успішність програмних проектів. Запропоновано інтелектуальний метод прогнозування успішності програмних проектів, який дозволяє обчислювати інтегративні показники успішності різних програмних проектів на основі прогнозованих (штучною нейронною мережею) значень характеристик розроблюваного за проектом програмного забезпечення, а також обчислювати ймовірності успішності цих програмних проектів. Метод відрізняється від відомих тим, що дозволяє прогнозувати успішність програмних проектів, порівнювати програмні проекти комплексно за основними характеристиками проекту і розроблюваного програмного забезпечення та прогнозованим значенням ймовірності успішності (а не тільки за вартістю та тривалістю, як відбувається наразі) та виконувати обґрунтований вибір програмного проекту замовником і розробником для подальшої реалізації. Проведені експерименти з дослідження роботоздатності методу. Результати експериментів дозволяють рекомендувати запропонований метод для використання на практиці.

Ключові слова: програмне забезпечення, програмний проект, специфікація вимог до програмного забезпечення, інтегративний показник проекту, успішність програмного проекту.

НОМЕНКЛАТУРА

ПЗ – програмне забезпечення;

ЖЦ ПЗ – життєвий цикл програмного забезпечення;

ШНМ – штучна нейронна мережа;

Cp – кросплатформність програмного забезпечення;

Cs – вартість програмного забезпечення;

Cx – складність програмного забезпечення;

Dsp – тривалість програмного проекту;

$Ecef$ – економічна ефективність програмного забезпечення;

$Ipup_{\mathrm{Max}_{best}}$ – максимальне значення інтегративного показника успішності програмного проекту за характеристиками ПЗ, які вимагають максимізації;

$Ipup_{\mathrm{Max}_{Sp}}$ – інтегративний показник успішності програмного проекту за характеристиками ПЗ, які вимагають максимізації;

$Ipup_{\mathrm{Min}_{bad}}$ – максимальне значення інтегративного показника успішності програмного проекту за характеристиками ПЗ, які вимагають мінімізації;

$Ipup_{\mathrm{Min}_{Sp}}$ – інтегративний показник успішності програмного проекту за характеристиками ПЗ, які вимагають мінімізації;

$P_{\mathrm{Max}\,Sp}$ – ймовірність успішності програмного проекту за інтегративним показником $Ipup_{\mathrm{Max}_{Sp}}$ – успішності програмного проекту;

$P_{\mathrm{Min}\,Sp}$ – ймовірність успішності програмного проекту за інтегративним показником $Ipup_{\mathrm{Min}_{Sp}}$ – успішності програмного проекту;

P_{Sp} – ймовірність успішності програмного проекту;

Qs – якість програмного забезпечення;

$R1$ – множина показників розділу 1 специфікації вимог до ПЗ;

$R2$ – множина показників розділу 2 специфікації вимог до ПЗ;

$R3$ – множина показників розділу 3 специфікації вимог до ПЗ;

$R4$ – множина показників розділу 4 специфікації вимог до ПЗ;

Rs – надійність програмного забезпечення;

$SCH_{\mathrm{Max}_{Sp}}$ – множина характеристик програмного проекту та розроблюваного ПЗ, збільшення значень яких викликають збільшення успішності програмного проекту;

$SCH_{\mathrm{Min}_{Sp}}$ – множина характеристик програмного проекту та розроблюваного ПЗ, збільшення значень яких викликають зменшення успішності програмного проекту;

SCH_{Sp} – множина основних характеристик програмного проекту;

$Slcm$ – тип моделі життєвого циклу програмного забезпечення;

Sp – програмний проект;

SRS – специфікація вимог до ПЗ;

Ub – зручність використання програмного забезпечення.

ВСТУП

Програмний проект – це комплекс взаємозв'язаних заходів, спрямованих на досягнення поставлених задач з чітко визначеними цілями протягом заданого періоду часу та при встановленому бюджеті [1]. Наразі великі програмні проекти часто виконуються з відставанням від графіка або з перевищенням кошторису витрат, розроблений продукт часто не має необхідних функціональних можливостей [2].

DOI 10.15588/1607-3274-2015-2-7

Сьогодні, коли кількість високобюджетних програмних проектів стрімко зростає, дуже важливим і актуальним є вміння оцінити можливу успішність програмного проекту на ранніх етапах життєвого циклу, а також допомогти замовнику і розробнику обрати ймовірно успішний програмний проект з множини альтернативних програмних проектів.

Статистика успішності програмних проектів за даними The Standish Group International (Chaos reports) [3] наведена у табл. 1.

Таблиця 1 – Успішність програмних проектів

Категорія проекту	1994	2000	2006	2012
Успішні, %	16	28	35	39
Проблемні, %	53	49	46	43
Провальні, %	31	23	19	18

Аналіз даних табл. 1 дав можливість побачити, що частка проблемних проектів є досить сталою величиною і складає мінімум 43 %. При цьому лише невелика кількість (максимум 39 % за 1994–2012 роки) проектів має необхідні функціональні можливості при існуючих обмеженнях за вартістю та терміном [3].

Програмні проекти часто зазнають невдач через помилки на ранніх етапах життєвого циклу ПЗ, а саме [2]: 1) неадекватне формулювання вимог; 2) невдале проектування або неефективне планування; 3) невірне розуміння або недостатній аналіз специфікації; 4) нереалістичні проектні плани; 5) невірно обрану модель життєвого циклу.

Помилки формулювання вимог та проектування складають 25–55% всіх помилок, причому чим більший обсяг ПЗ, тим більше помилок вноситься саме на ранніх етапах [3]. Слід врахувати й той факт, що вартість виправлення помилки проектування в два-чотири рази вища вартості виправлення помилки конструювання [3].

1 ПОСТАНОВКА ЗАДАЧІ

Нехай програмний проект задано специфікацією вимог до ПЗ [4] у наступному формалізованому вигляді [5]: $SRS =< R1, R2, R3, R4 >$, де $R1$ – множина показників розділу 1 специфікації вимог до ПЗ, $R2$ – множина показників розділу 2, $R3$ – множина показників розділу 3, $R4$ – множина показників розділу 4.

Тоді задача оцінювання успішності програмного проекту полягає у:

1) прогнозуванні (на основі множин показників специфікації $R1$, $R2$, $R3$, $R4$) характеристик програмного проекту та розроблюваного за ним ПЗ – множини значень $SCH_{Sp} = \{Cs, Dsp, Ecef, Cp, Cx, Ub, Qs, Rs, Slcm\}$;

2) інтерпретації отриманих відносних значень характеристик проекту та ПЗ-критеріями для такої інтерпретації є інтегративні показники успішності $Ipup_{\mathrm{Min}_{Sp}}$ (за характеристиками ПЗ $SCH_{\mathrm{Min}_{Sp}} = \{Cs, Dsp, Cx\}$, $SCH_{\mathrm{Min}_{Sp}} \subset SCH_{Sp}$) та $Ipup_{\mathrm{Max}_{Sp}}$ (за характеристиками ПЗ $SCH_{\mathrm{Max}_{Sp}} = \{Ecef, Ub, Cp, Qs, Rs\}, SCH_{\mathrm{Max}_{Sp}} \subset SCH_{Sp}$);

3) оцінюванні ймовірності успішності програмного проекту P_{Sp}-критеріями є ймовірність успішності $P_{\mathrm{Min}_{Sp}}$ за інтегративним показником $Ipup_{\mathrm{Min}_{Sp}}$ та ймовірність успішності $P_{\mathrm{Max}_{Sp}}$ за інтегративним показником $Ipup_{\mathrm{Max}_{Sp}}$.

2 ОГЛЯД ЛІТЕРАТУРИ

Аналіз основних характеристик ПЗ [6–8] зробив очевидним факт, що існуючі математичні інструменти та методи визначення основних характеристик ПЗ не придатні для оцінки їх значень на ранніх етапах життєвого циклу (на етапах формулювання вимог та проектування), оскільки вони орієнтовані на готовий програмний код, а не на наявну на етапі проектування специфікацію вимог до програмного забезпечення. Саме аналіз специфікації вимог до ПЗ може надати різнопланову інформацію для подальшого розрахунку тривалості програмного проекту, вартості, кросплатформності, ефективності, складності, якості та надійності розроблюваного за проектом ПЗ, а також для вибору прийнятної моделі життєвого циклу ПЗ.

Аналіз специфікації вимог до ПЗ [6, 9, 10] показав, що вже на етапі проектування, коли розробник надав готову специфікацію, аналіз її вимог дозволяє сформувати множини кількісних та якісних показників, на основі яких замовник і розробник можуть отримати прогнозовані кількісні значення характеристик розроблюваного за програмним проектом ПЗ, які дозволяють отримати прогноз успішності даного програмного проекту.

З проведеного аналізу [6, 11, 12] відомих автоматизованих засобів аналізу специфікацій та оцінювання характеристик програмного забезпечення зрозуміло, що всі вони призначені для роботи з вимогами та специфікаціями, причому деякі з них спрямовані на контроль за реалізацією вимог, але жоден з них не визначає прогнозованих значень характеристик ПЗ за специфікацією. Очевидно тоді, що існуючі засоби не прийнятні для кількісного оцінювання характеристик ПЗ на основі аналізу специфікацій.

3 МАТЕРІАЛИ ТА МЕТОДИ

На характеристики тривалості програмного проекту, типу моделі життєвого циклу, вартості, економічної ефективності, складності, зручності використання, кросплатформності, якості та надійності ПЗ впливають певні показники специфікації [5], але невідомі функції (формули, залежності), за якими можна обчислити значення тієї чи іншої характеристики ПЗ на основі множини впливових показників специфікації – всі наявні формули та методики оцінювання характеристик ПЗ орієнтовані на готовий програмний код, а не на специфікацію вимог.

Теорема Хехт-Нільсена [13] доводить можливість розв'язку задачі представлення багатовимірної функції довільного вигляду на ШНМ, тому для реалізації невідомих функцій залежності характеристик ПЗ від показників специфікації використовуватимемо саме ШНМ. Важкоформалізованою задачею прогнозування характеристик ПЗ є визначення ваг та взаємовпливу показників специфікації в межах кожної характеристики ПЗ. Ця задача може бути вирішена за допомогою використання навченої ШНМ.

У [14] розроблено ШНМ, яка опрацьовує показники специфікації, здійснює апроксимацію показників та надає прогнозовані кількісні оцінки характеристик ПЗ, на основі яких можна зробити прогноз успішності програмного проекту та здійснити вибір моделі життєвого циклу. Концепцію прогнозування основних характеристик ПЗ на основі аналізу специфікації вимог із використанням ШНМ представлено на рис. 1.

Отже, за рисунком 1, ШНМ прогнозування характеристик програмного забезпечення на основі аналізу специ-фікацій видає множину $SCH_{Sp} = \{Cs, Dsp, Ecef, Cp, Cx, Ub, Qs, Rs, Slcm\}$ значень основних характеристик програмного проекту Sp та розроблюваного за ним ПЗ. ШНМ прогнозування характеристик ПЗ на основі аналізу специфікацій навчена так, що всі значення характеристик ПЗ (результуючі дані ШНМ) належатимуть інтервалу (0;1]. На основі отриманих з ШНМ значень основних характеристик ПЗ замовнику складно оцінити успішність проекту, оскільки складно вірно інтерпретувати одержані значення характеристик. Крім цього, успішність програмного проекту та розроблюваного за ним ПЗ слід оцінювати, приймаючи до уваги всі отримані значення характеристик в комплексі та враховуючи які значення характеристик викликають зменшення успішності програмного проекту, а які значення – її збільшення. Така задача є ще більш важкою та незрозумілою для замовника. Тому варто визначити метод, який допомагатиме замовнику оцінити успішність програмного проекту, інтерпретуючи та інтегруючи одержані значення характеристик ПЗ.

Збільшення значень деяких характеристик $SCH_{\mathrm{Min}_{Sp}} = \{Cs, Dsp, Cx\}$ $(SCH_{\mathrm{Min}_{Sp}} \subset SCH_{Sp})$ викликають зменшення успішності програмного проекту. Тоді значення характеристик множини $SCH_{\mathrm{Min}_{Sp}}$, близькі до 0, будуть вказувати на високу успішність програмного проекту. Збільшення значень інших характеристик $SCH_{\mathrm{Max}_{Sp}} = \{Ecef, Ub, Cp, Qs, Rs\}(SCH_{\mathrm{Max}_{Sp}} \subset SCH_{Sp})$ викликають збільшення успішності програмного проекту. Тоді значення характеристик множини $SCH_{\mathrm{Max}_{Sp}}$, близькі до 1, вказуватимуть на високу успішність програмного проекту.

Визначення 1. Інтегративний показник успішності програмного проекту Sp – це кількісний показник успішності програмного проекту на основі множини значень прогнозованих характеристик $SCH_{Sp} = \{Cs, Dsp, Ecef, Cp, Cx, Ub, Qs, Rs, Slcm\}$.

Оскільки, як показано вище, значення різних характеристик по-різному впливають на успішність програмного проекту, то введемо два інтегративні показники успішності.

Визначення 2. Інтегративний показник успішності $Ipup_{\mathrm{Min}_{Sp}}$ програмного проекту Sp – це кількісний показник успішності програмного проекту на основі множини значень прогнозованих характеристик $SCH_{\mathrm{Min}_{Sp}} = \{Cs, Dsp, Cx\}$.

Визначення 3. Інтегративний показник успішності $Ipup_{\mathrm{Max}_{Sp}}$ програмного проекту Sp – це кількісний показник успішності програмного проекту на основі множини значень прогнозованих характеристик $SCH_{\mathrm{Max}_{Sp}} = \{Ecef, Ub, Cp, Qs, Rs\}$.

Для отримання $Ipup_{\mathrm{Min}_{Sp}}$ проекту Sp створимо графік в системі координат, яка має три основих вісі (для трьох характеристик – $Cs \in (0;1]$, $Dsp \in (0;1]$, $Cx \in (0;1]$) – рис. 2. Для отримання $Ipup_{\mathrm{Max}_{Sp}}$ проекту Sp створимо графік в системі координат, яка має п'ять основих осей (для п'яти характеристик – $Ecef \in (0;1]$, $Ub \in (0;1]$, $Cp \in (0;1]$, $Qs \in (0;1]$, $Rs \in (0;1]$) – рис. 3.

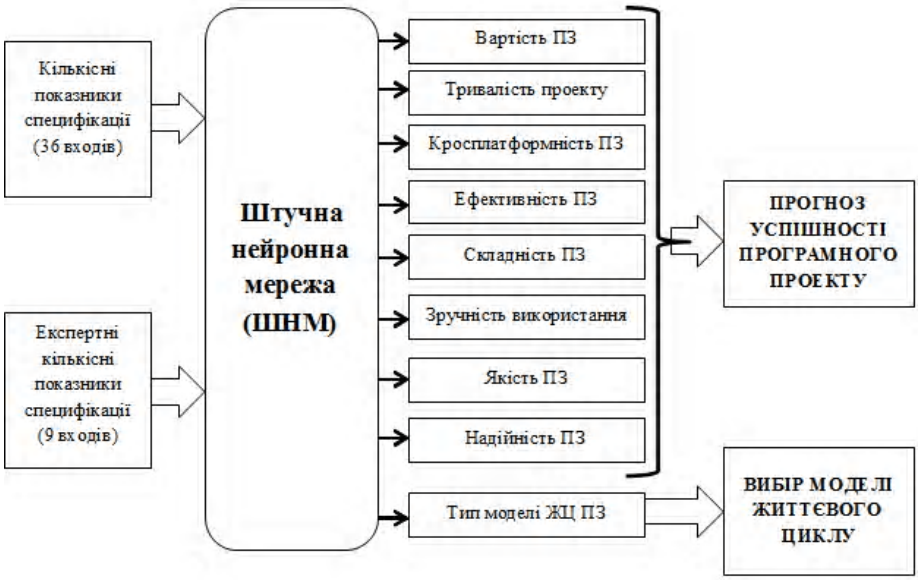


Рисунок 1 – Концепція прогнозування характеристик ПЗ на основі аналізу специфікацій із використанням нейромережних інформаційних технологій

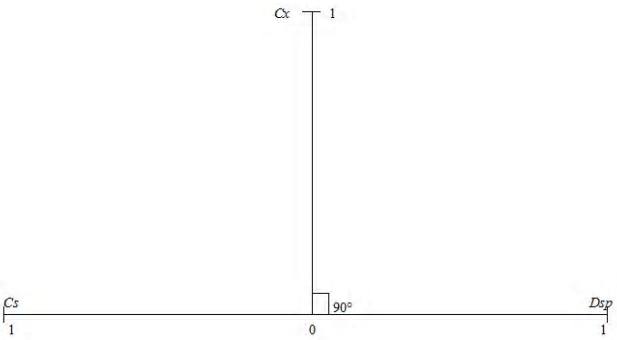


Рисунок 2 – Система координат для $Ipup_{\mathrm{Min}_{Sp}}$

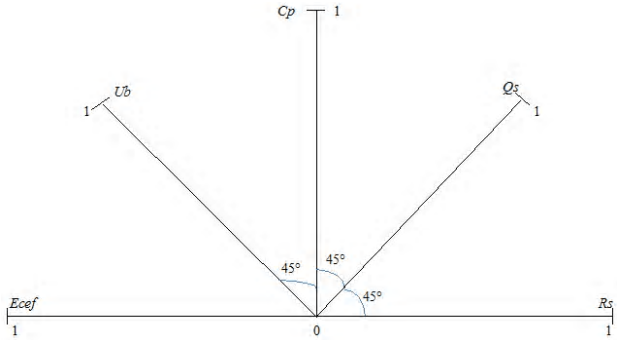


Рисунок 3 – Система координат для $Ipup_{\mathrm{Max}_{Sp}}$

Нехай ШНМ прогнозування характеристик ПЗ на основі аналізу специфікацій надала нам наступні значення характеристик для Sp: $Cs_{ANN} \in (0;1]$, $Dsp_{ANN} \in (0;1]$, $Cx_{ANN} \in (0;1]$; $EcEf_{ANN} \in (0;1]$, $Ub_{ANN} \in (0;1]$, $Cp_{ANN} \in (0;1]$, $Qs_{ANN} \in (0;1]$, $Rs_{ANN} \in (0;1]$. Відкладемо тоді отримані значення у відповідних системах координат і отримаємо наступні графічні представлення $Ipup_{\mathrm{Min}_{Sp}}$ (рис. 4) та $Ipup_{\mathrm{Max}_{Sp}}$ (рис. 5) проекту Sp.

Інтегративним показником успішності $Ipup_{\mathrm{Min}_{Sp}}$ програмного проекту Sp є площа виділеного суцільною жирною лінією на рис. 4 трикутника $Cs_{ANN}Dsp_{ANN}Cx_{ANN}$, а інтегративним показником успішності $Ipup_{\mathrm{Max}_{Sp}}$ проекту Sp є площа виділеного суцільною жирною лінією на рис. 5 п'ятикутника $Ecef_{ANN}Ub_{ANN}Cp_{ANN}Qs_{ANN}Rs_{ANN}$.

Тоді для визначення $Ipup_{\mathrm{Min}_{Sp}}$ проекту Sp (визначення площі трикутника $Cs_{ANN}Dsp_{ANN}Cx_{ANN}$) застосуємо формулу площі трикутника за відомою стороною та висотою:

$$Ipup_{\mathrm{Min}_{Sp}} = \frac{1}{2} \cdot (Cs_{ANN} + Dsp_{ANN}) \cdot Cx_{ANN} . \qquad (1)$$

Визначальною характеристикою щодо прогнозування успішності програмного проекту (серед характеристик тривалості проекту, вартості та складності ПЗ) є характеристика складності, оскільки постійне зростання складності функцій ПЗ неминуче призводить до збільшення трудомісткості створення програмних проектів, кількості помилок у програмному коді, а також до зменшення успішності програмних проектів [3]. Тому система координат для $Ipup_{\mathrm{Min}_{Sp}}$ (рис. 2) побудована таким чином, що при визначенні $Ipup_{\mathrm{Min}_{Sp}}$ як площі трикутника $Cs_{ANN}Dsp_{ANN}Cx_{ANN}$ саме характеристика складності є визначальною і не підлягає компенсації іншими характеристиками.

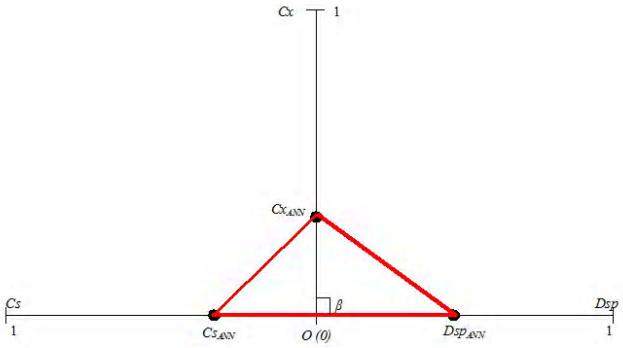


Рисунок 4 – Графічне подання $Ipup_{\mathrm{Min}_{Sp}}$ програмного проекту Sp

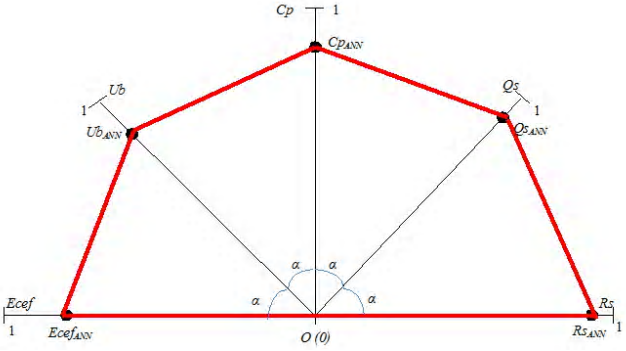


Рисунок 5 – Графічне подання $Ipup_{\mathrm{Max}_{Sp}}$ програмного проекту Sp

Для визначення $Ipup_{\mathrm{Max}_{Sp}}$ проекту Sp (площі п'ятикутника $Ecef_{ANN}Ub_{ANN}Cp_{ANN}Qs_{ANN}Rs_{ANN}$) розіб'ємо п'ятикутник на чотири трикутники – $Ecef_{ANN}OUb_{ANN}$, $Ub_{ANN}OCp_{ANN}$, $Cp_{ANN}OQs_{ANN}$, $Qs_{ANN}ORs_{ANN}$ і знайдемо площу для кожного з трикутників, застосовуючи формулу площі трикутника за відомими двома сторонами та кутом між ними:

$$\begin{aligned} Ipup_{Max_{Sp}} &= S_{Ecef_{ANN}OUb_{ANN}} + S_{Ub_{ANN}OCp_{ANN}} + \\ &+ S_{Cp_{ANN}OQs_{ANN}} + S_{Qs_{ANN}ORs_{ANN}} = \\ &= \frac{1}{2} \cdot Ecef_{ANN} \cdot Ub_{ANN} \cdot \sin 45° + \frac{1}{2} \cdot Ub_{ANN} \cdot Cp_{ANN} \cdot \sin 45° + \\ &+ \frac{1}{2} \cdot Cp_{ANN} \cdot Qs_{ANN} \cdot \sin 45° + \frac{1}{2} \cdot Qs_{ANN} \cdot Rs_{ANN} \cdot \sin 45° = \\ &= 0{,}3536 \cdot (Ecef_{ANN} \cdot Ub_{ANN} + Ub_{ANN} \cdot Cp_{ANN} + \\ &+ Cp_{ANN} \cdot Qs_{ANN} + Qs_{ANN} \cdot Rs_{ANN}). \end{aligned} \qquad (2)$$

Інтегративні показники успішності програмного проекту дають змогу визначити ймовірність його успішності.

Для визначення ймовірності успішності програмного проекту за інтегративним показником успішності $Ipup_{\mathrm{Min}_{Sp}}$ слід знайти максимальне значення інтегративного показника успішності $Ipup_{\mathrm{Min}_{bad}}$ (площу максимально можливого трикутника $CsDspCx$, окресленого пунктирною лінією на рис. 6).

Оскільки ШНМ прогнозування характеристик ПЗ на основі аналізу специфікацій навчена так, що максимальні значення характеристик вартості, тривалості та складності, отримані з неї, становлять 1, то:

$$Ipup_{\mathrm{Min}_{bad}} = \frac{1}{2}\cdot(Cs + Dsp)\cdot Cx = \frac{1}{2}\cdot(1+1)\cdot 1 = 1. \quad (3)$$

Для визначення ймовірності успішності програмного проекту за інтегративним показником успішності $Ipup_{\mathrm{Max}_{Sp}}$ слід знайти максимальне значення інтегративного показника успішності $Ipup_{\mathrm{Max}_{best}}$ (площу максимально можливого п'ятикутника $EcefUbCpQsRs$, окресленого пунктирною лінією на рис. 7).

Оскільки ШНМ прогнозування характеристик ПЗ на основі аналізу специфікацій навчена так, що максимальні значення характеристик економічної ефективності, зручності використання, кросплатформності, якості та надійності, отримані з неї, становлять 1, то:

$$Ipup_{\mathrm{Max}_{best}} = 0{,}3536\cdot(Ecef_{ANN}\cdot Ub_{ANN} + Ub_{ANN}\cdot Cp_{ANN} + $$
$$+Cp_{ANN}\cdot Qs_{ANN} + Qs_{ANN}\cdot Rs_{ANN}) =$$
$$= 0{,}3536\cdot(1\cdot 1 + 1\cdot 1 + 1\cdot 1 + 1\cdot 1) = 1{,}4144. \quad (4)$$

Оскільки характеристики вартості, тривалості та складності вимагають мінімізації, то максимальне значення $Ipup_{\mathrm{Min}_{Sp}}$, яке складає $Ipup_{\mathrm{Min}_{bad}} = 1$ (згідно формули (3)), – це найгірше значення $Ipup_{\mathrm{Min}_{Sp}}$ програмного проекту.

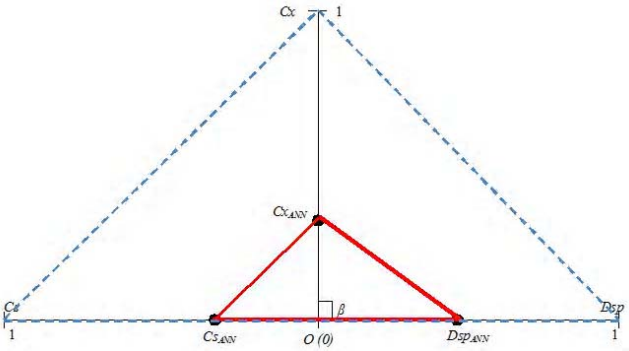


Рисунок 6 – Графічне подання інтегративного показника успішності $Ipup_{\mathrm{Min}_{Sp}}$ та максимального значення інтегративного показника успішності $Ipup_{\mathrm{Min}_{bad}}$

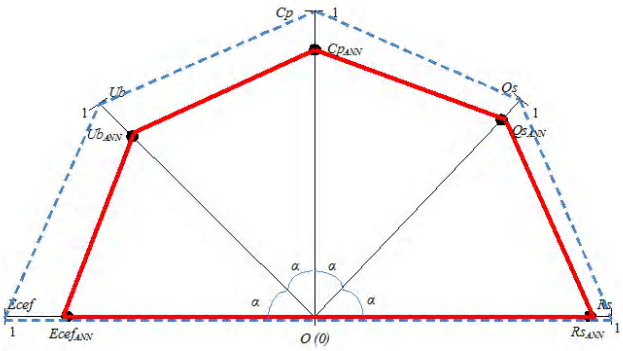


Рисунок 7 – Графічне подання інтегративного показника успішності $Ipup_{\mathrm{Max}_{Sp}}$ та максимального значення інтегративного показника успішності $Ipup_{\mathrm{Max}_{best}}$

Тоді за $Ipup_{\mathrm{Min}_{Sp}}$ ймовірність успішності $P_{\mathrm{Min}\,Sp}$ програмного проекту Sp складає:

$$P_{\mathrm{Min}\,Sp} = 1 - \frac{Ipup_{\mathrm{Min}_{Sp}}}{Ipup_{\mathrm{Min}_{bad}}} = 1 - \frac{Ipup_{\mathrm{Min}_{Sp}}}{1} = 1 - Ipup_{\mathrm{Min}_{Sp}}. \quad (5)$$

Характеристики економічної ефективності, зручності використання, кросплатформності, якості та надійності навпаки вимагають максимізації, тому максимальне значення $Ipup_{\mathrm{Max}_{Sp}}$, яке складає $Ipup_{\mathrm{Max}_{best}} = 1{,}4144$ (згідно формули (4)), – це найкраще значення $Ipup_{\mathrm{Max}_{Sp}}$ програмного проекту.

Тоді за $Ipup_{\mathrm{Max}_{Sp}}$ ймовірність успішності $P_{\mathrm{Max}\,Sp}$ програмного проекту Sp складає:

$$P_{\mathrm{Max}\,Sp} = \frac{Ipup_{\mathrm{Max}_{Sp}}}{Ipup_{\mathrm{Max}_{best}}} = \frac{Ipup_{\mathrm{Max}_{Sp}}}{1{,}4144} = 0{,}7070\cdot Ipup_{\mathrm{Max}_{Sp}}. \quad (6)$$

Оскільки $Ipup_{\mathrm{Min}_{Sp}}$ проекту Sp, на основі якого обчислюється $P_{\mathrm{Min}\,Sp}$ проекту, залежить від трьох основних характеристик ПЗ, а $Ipup_{\mathrm{Max}_{Sp}}$ проекту Sp, на основі якого обчислюється $P_{\mathrm{Max}\,Sp}$ проекту, залежить від п'яти основних характеристик ПЗ, тоді усереднене значення ймовірності успішності P_{Sp} програмного проекту Sp на основі двох отриманих значень $P_{\mathrm{Min}\,Sp}$ та $P_{\mathrm{Max}\,Sp}$ обчислюється за формулою:

$$P_{Sp} = \frac{3\cdot P_{\mathrm{Min}_{Sp}} + 5\cdot P_{\mathrm{Max}_{Sp}}}{8}. \quad (7)$$

Формалізація даного методу прогнозування успішності програмного забезпечення на етапі проектування матиме наступний вигляд:

$$SCH_{Sp} = \{SCH_{\mathrm{Min}_{Sp}}, SCH_{\mathrm{Max}_{Sp}}\} \Rightarrow \begin{cases} Ipup_{\mathrm{Min}_{Sp}} = f_1(SCH_{\mathrm{Min}_{Sp}}) \\ Ipup_{\mathrm{Max}_{Sp}} = f_2(SCH_{\mathrm{Max}_{Sp}}) \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} P_{Min_{Sp}} = f_3(Ipup_{\mathrm{Min}_{Sp}}, Ipup_{\mathrm{Min}_{bad}}) \\ P_{Max_{Sp}} = f_4(Ipup_{\mathrm{Max}_{Sp}}, Ipup_{\mathrm{Max}_{best}}) \end{cases} \Rightarrow P_{Sp} = f_5(P_{\mathrm{Min}_{Sp}}, P_{\mathrm{Max}_{Sp}}). \quad (8)$$

де функція $f_1()$ обчислюється за формулою (1), функція $f_2()$ – за формулою (2), функція $f_3()$ – за формулою (5), функція $f_4()$ – за формулою (6), функція $f_5()$ – за формулою (7); показники $Ipup_{\mathrm{Min}_{bad}}$ та $Ipup_{\mathrm{Max}_{best}}$ обчислюються за формулами (3) та (4) відповідно.

4 ЕКСПЕРИМЕНТИ

Виконаємо експериментальне дослідження методу прогнозування успішності проектів. Для цього розглянемо чотири альтернативні програмні проекти, розроблені різними софтверними компаніями для вирішення однієї задачі. Нехай для розглядуваних чотирьох альтернативних проектів було отримано наступні прогнозовані значення характеристик ПЗ – таблиця 2.

5 РЕЗУЛЬТАТИ

Згідно методу прогнозування успішності програмних проектів (формула (8)), обчислимо $Ipup_{\mathrm{Min}_{Sp}}$ і $Ipup_{\mathrm{Max}_{Sp}}$ для чотирьох альтернативних програмних проектів, а також $P_{\mathrm{Min}\,Sp}$, $P_{\mathrm{Max}\,Sp}$, P_{Sp} цих проектів. Результати обчислень представлені у таблиці 3.

Таблиця 2 – Прогнозовані значення характеристик ПЗ для чотирьох альтернативних програмних проектів

Характеристика	Проект1	Проект2	Проект3	Проект4
1	2	3	4	5
Cs	0,2	0,78	0,61	0,41
Dsp	0,1	0,81	0,59	0,43
Cx	0,15	0,69	0,63	0,38
Ecef	0,8	0,13	0,52	0,52
Ub	0,85	0,15	0,5	0,56
1	2	3	4	5
Cp	0,87	0,21	0,47	0,57
Qs	0,89	0,17	0,49	0,61
Rs	0,91	0,24	0,41	0,49

Таблиця 3 – Інтегративні показники та ймовірності успішності чотирьох альтернативних програмних проектів

Величина успішності	Проект1	Проект2	Проект3	Проект4
$Ipup_{\mathrm{Min}_{Sp}}$	0,0225	0,5486	0,3780	0,1596
$Ipup_{\mathrm{Max}_{Sp}}$	1,0621	0,0451	0,3275	0,4445
$P_{\mathrm{Min}\,Sp}$	0,9775 (97,75 %)	0,4514 (45,14 %)	0,6220 (62,20 %)	0,8404 (84,04 %)
$P_{\mathrm{Max}\,Sp}$	0,7509 (75,09 %)	0,0319 (3,19 %)	0,2316 (23,16 %)	0,3143 (31,43 %)
P_{Sp}	0,8359 (83,59 %)	0,1892 (18,92 %)	0,3780 (37,80 %)	0,5116 (51,16 %)

6 ОБГОВОРЕННЯ

Так, за результатами, наведеними у таблиці 3, проект1 має найбільшу ймовірність успішності (майже 84%), а проект2 – найменшу ймовірність успішності (близько 19 %). Враховуючи отримані результати, замовнику для вирішення поставленої задачі пропонується обрати проект1, оскільки саме він має найкращі характеристики та найвищу ймовірну успішність.

Наразі вибір програмного проекту як замовником, так і розробником здійснюється на основі лише прогнозованих у специфікації характеристик вартості, тривалості та власної інтуїції. Але не завжди розробнику вдається вірно спрогнозувати орієнтовану вартість та тривалість при розробленні специфікації вимог до ПЗ. Як було доведено у [6], вартість і тривалість слід оцінювати комплексно з врахуванням всіх впливових вимог (показників) специфікації.

Прогнозовані в специфікації значення вартості та тривалості чотирьох розглядуваних програмних проектів наведено у таблиці 4.

Характеристики програмних проектів з таблиці 4 свідчать, що для всіх чотирьох програмних проектів у специфікаціях зазначено приблизно однакові прогнозовані вартість та тривалість розроблення, але в таблиці 2 показано, що вони мають суттєво різні відносні оцінки всіх характеристик ПЗ, в тому числі вартості та тривалості, обчислені комплексно, з врахуванням всіх значущих показників специфікації. Так, відносна вартість коливається в межах від 0,2 (для проекту1) до 0,78 (для проекту2); відносна тривалість – в межах від 0,1 (для проекту1) до 0,81 (для проекту2). Отже, якщо оцінювати вартість та тривалість, враховуючи всі значущі показники специфікації, то очевидно, що їх значення не будуть однаковими для чотирьох розглядуваних програмних проектів. Тому на основі лише прогнозованих у специфікації вартості та часу розроблення як замовник, так і розробник може прийняти хибний висновок щодо вибору програмного проекту.

Крім цього, успішність програмних проектів залежить не лише від вартості та тривалості, але й від решти основних характеристик ПЗ – складності, ефективності, зручності використання, кросплатформності, якості та надійності, які в явному кількісному вигляді взагалі не зазначені у специфікації вимог до ПЗ. Тому саме оцінки всіх основних характеристик, надані ШНМ, та опрацьовані згідно запропонованого методу, допоможуть зробити вірний вибір і реалізовувати проект, який має кращі значення характеристик та найвищу ймовірність успішності (серед розглядуваних чотирьох проектів – це проект1).

Отже, результати роботи запропонованого методу дають можливість замовнику обрати програмний проект з точки зору його основних характеристик та прогнозованої успішності.

ВИСНОВКИ

Вперше запропоновано метод прогнозування успішності програмних проектів на основі аналізу прогнозованих значень характеристик ПЗ, який базується на опрацюванні вихідних функціоналів нейромережної моделі процесу прогнозування характеристик ПЗ, які відповідають кількісним значенням основних характеристик ПЗ і дають можливість оцінити сумарний вплив показників специфікації на характеристики розроблюваного за специфікацією ПЗ. Опрацювання вихідних функціоналів полягає у обчисленні двох інтегративних показників успішності (за характеристиками ПЗ, які вимагають мінімізації, та за характеристиками ПЗ, які вимагають максимізації) для будь-якого описаного специфікацією програмного проекту, на основі яких обчислюються значення ймовірностей успішності програмного проекту за обома інтегративними показниками та усередненої ймовірності успішності програмного проекту.

Метод відрізняється від відомих тим, що дозволяє прогнозувати успішність програмних проектів та порівнювати програмні проекти комплексно – за основними ха-

Таблиця 4 – Прогнозовані в специфікації значення вартості та тривалості для чотирьох альтернативних програмних проектів

Характеристика	Проект1	Проект2	Проект3	Проект4
Прогнозована вартість, вказана у специфікації	87000 грн.	89000 грн.	85000 грн.	86500 грн.
Прогнозована тривалість, вказана у специфікації	200 роб.днів	210 роб.днів	198 роб.днів	203 роб.днів

рактеристиками проекту і розроблюваного ПЗ та прогнозованими значеннями ймовірності успішності (а не тільки за вартістю та тривалістю, як відбувається наразі).

Практичне значення розробленого методу полягає у його використанні для виконання обґрунтованого вибору програмного проекту замовником і розробником для подальшої реалізації.

Перспективою для подальших досліджень авторів є підвищення достовірності роботи даного методу за рахунок підвищення достовірності роботи ШНМ прогнозування основних характеристик ПЗ на основі аналізу специфікацій, для чого необхідним є збір великої кількості показників та значень характеристик ПЗ зі специфікацій вже розроблених проектів для побудови навчальної вибірки ШНМ (для достовірної на 100% роботи ШНМ необхідний об'єм навчальної вибірки становить 20250 векторів).

СПИСОК ЛІТЕРАТУРИ

1. Futrell R. T. Quality software project management / R. T. Futrell, D. F. Shafer, L. I. Shafer. – New York : Prentice Hall PTR, 2003. – 1136 p.
2. Pomorova O. Intelligent assessment and pediction of software characteristics at the design stage / O. Pomorova, T. Hovorushchenko // American journal of software engineering and applications (AJSEA). – 2013. – Vol. 2. – P. 25–31.
3. CHAOS Manifesto: think big, act small, 2013 [Electronic resource] – Access mode: http://www.versionone.com/assets/img/files/CHAOSManifesto2013.pdf
4. IEEE 830-1998. Recommended practice for software requirements specifications. – New York : IEEE, 1998 – 31 p.
5. Говорущенко Т. О. Математичне моделювання специфікації вимог та характеристик програмного забезпечення / Т. О. Говорущенко, А. В. Красій // Радіоелектронні і комп'ютерні системи. – 2014. – № 5. – С. 34–39.
6. Говорущенко Т. О. Визначення характеристик та вибір моделі життєвого циклу програмного забезпечення на основі аналізу специфікацій / Т. О. Говорущенко, А. В. Красій // Вісник Хмельницького національного університету. – 2013. – № 6. – С. 201–208.
7. Maedche A. Software for people: fundamentals, trends and best practices (Management for professionals) / A. Maedche, A. Botzenhardt, L. Neer. – Berlin : Springer-Verlag Berlin Heidelberg, 2012. – 293 p.
8. Jones C. The economics of software quality / C. Jones, O. Bonsignour. – Boston : Pearson Education, 2012. – 588 p.
9. Wiegers K. Software requirements: 3-rd edition / K. Wiegers, J. Beatty. – Washington : MS Press, 2013. – 640 p.
10. Wiegers K. More about software requirements: Thorny issues and practical advice / K. Wiegers. – Washington : MS Press, 2006. – 203 p.
11. Chen A. Visual models for software requirements / A. Chen, J. Beatty. – Washington : MS Press, 2012. – 444 p.
12. Open Source Requirements Management Tool [Electronic resource] – Access mode: http://sourceforge.net/projects/osrmt/
13. Callan R. The essence of neural networks / R. Callan. – London : Prentice Hall Europe, 2003. – 288 p.
14. Красій А. В. Моделювання процесу прогнозування характеристик програмного забезпечення на основі аналізу специфікацій / А. В. Красій // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. – 2014. – № 2. – С. 66–76.

Стаття надійшла до редакції 09.12.2014.
Після доробки 19.12.2014.

Красий А. В.[1], Говорущенко Т. А.[2]

[1]Аспирант кафедры системного программирования Хмельницкого национального университета, Хмельницкий, Украина

[2]Канд. техн. наук, доцент, старший научный сотрудник, доцент кафедры системного программирования Хмельницкого национального университета, Хмельницкий, Украина

МЕТОД ПРОГНОЗИРОВАНИЯ УСПЕШНОСТИ ПРОГРАММНЫХ ПРОЕКТОВ НА ОСНОВЕ АНАЛИЗА ПРОГНОЗИРОВАННЫХ ЗНАЧЕНИЙ ХАРАКТЕРИСТИК ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Решена задача разработки методологического обеспечения для прогнозирования успешности программных проектов. Объектом исследования является процесс прогнозирования успешности программных проектов на ранних этапах жизненного цикла. Предметом исследования является метод прогнозирования успешности программных проектов на основе анали за прогнозируемых значений характеристик програмного обеспечения. Цель работы – повысить достоверность выбора программных проектов и успешность программных проектов. Предложен интеллектуальный метод прогнозирования успешности программных проектов, который позоляет вычислять интегративные показатели успешности различных программных проектов на основе прогнозируемых (искусственной нейронной сетью) значений характеристик разрабатываемого по проекту програмного обеспечения, а также вычислять вероятности успешности этих программных проектов. Метод отличается от известных тем, что позволяет прогнозировать успешность программных проектов, сравнивать программные проекты комплексно по основным характеристикам проекта и разрабатываемого программного обеспечения и по прогнозированным значениям вероятности успешности (а не только по стоимости и продолжительности, как происходит сейчас) и выполнять обоснованный выбор програмного проекта заказчиком и разработчиком для дальнейшей реализации. Проведены эксперименты по исследованию работоспособности метода. Результаты экспериментов позволяют использовать предложенный метод на практике.

Ключевые слова: программное обеспечение, программный проект, спецификация требований к программному обеспечению, интегративный показатель проекта, успешность программного проекта.

Krasiy A. V.[1], Hovorushchenko T. O.[2]

[1]Post-graduate student of System Programming Department of Khmelnitsky National University, Khmelnitsky, Ukraine

[2]PhD, Associate Professor, Senior Researcher, Associte Professor of System Programming Department of Khmelnitsky National University, Khmelnitsky, Ukraine

THE METHOD OF PREDICTION OF SOFTWARE PROJECTS SUCCESS BASED ON THE ANALYSIS OF PREDICTED VALUES OF SOFTWARE CHARACTERISTICS

The task of development of methodological support for predicting of the software projects success is solved. The object of research is the process of predicting of software projects success at the early lifecycle stages. The subject of the research is a method of prediction of software projects success based on the analysis of predicted values of software characteristics. The purpose of the work is to increase the accuracy of selection of software projects and to increase the success of software projects. The intelligent method of prediction of software projects success

is proposed. It provides to calculate the integrative indicators of success of various software projects based on the predicted (by artificial neural network) values of the characteristics of the software developed for the project. It also provides to calculate the probability of success of software projects. The method differs from the known methods, that provides to predict the success of software projects, compare software projects on the main characteristics of the project and the software and on the predicted value of the probability of success (and not just on the cost and duration, as is happening now), and perform the grounded choice of software project by customer and by developer for further realization. Experiments on investigations of the method operability were performed The experimental results provide to recommend the proposed method for use in practice.

Keywords: software, software project, software requirements specification, integrative indicator of software project, success of software project.

REFERENCES

1. Futrell R. T., Shafer D. F., Shafer L. I. Quality software project management. New York, Prentice Hall PTR, 2003, 1136 p.
2. Pomorova O., Hovorushchenko T. Intelligent assessment and pediction of software characteristics at the design stage, *American journal of software engineering and applications (AJSEA),* 2013, Vol. 2, pp. 25–31.
3. CHAOS Manifesto : think big, act small, 2013 [Electronic resource]. Access mode: http://www.versionone.com/assets/img/files/CHAOSManifesto2013.pdf
4. IEEE 830-1998. Recommended practice for software requirements specifications. New York, IEEE, 1998, 31 p.
5. Hovorushchenko T. O., Krasiy A. V. Matematichne modeluvannya spetsifikacii vimog ta kharakteristik programnogo zabezpechennya, *Radioelektronni i komputerni sistemi,* 2014, No. 5, pp. 34–39.
6. Hovorushchenko T. O. Krasiy A. V. Viznachennya kharakteristik ta vibir modeli jittevogo tsiklu programnogo zabezpechennya na osnovi analizu spetsifikatsiy, *Visnik Khmelnitskogo natsionalnogo universitetu*, 2013, No. 6, pp. 201–208.
7. Maedche A. Botzenhardt A., Neer L. Software for people: fundamentals, trends and best practices (Management for professionals). Berlin, Springer-Verlag Berlin Heidelberg, 2012, 293 p.
8. Jones C., Bonsignour O. The economics of software quality. Boston, Pearson Education, 2012, 588 p.
9. Wiegers K., Beatty J. Software requirements: 3-rd edition. Washington, MS Press, 2013, 640 p.
10. Wiegers K. More about software requirements: Thorny issues and practical advice. Washington, MS Press, 2006, 203 p.
11. Chen A., Beatty J. Visual models for software requirements. Washington, MS Press, 2012, 444 p.
12. Open Source Requirements Management Tool [Electronic resource]. Access mode: http://sourceforge.net/projects/osrmt/
13. Callan R. The essence of neural networks. London, Prentice Hall Europe, 2003, 288 p.
14. Krasiy A. V. Modeluvannya protsesu prognozuvannya kharakteristik programnogo zabezpechennya na osnovi analizu spetsifikaciy, *Komputerno-integrovani tekhnologii: osvita, nauka, virobnictvo,* 2014, No. 2, pp. 66–76.

УДК 004.93

Молчанова В. С.

Ст. преподаватель кафедры информатики ГВУЗ «Приазовский государственный технический университет», Мариуполь, Украина

АДАПТИВНЫЙ ПОРОГОВЫЙ МЕТОД БИНАРИЗАЦИИ РАСТРОВЫХ ИЗОБРАЖЕНИЙ ТЕХНИЧЕСКИХ ЧЕРТЕЖЕЙ

В статье предлагается метод адаптивной пороговой бинаризации растровых изображений технических чертежей. Проанализированы специфические особенности изображений технических чертежей и их влияние на результат бинаризации известными универсальными методами. Выявлены артефакты, возникающие на результирующих бинарных изображениях и предложены способы их устранения.

В основу предложенного метода положено соображение о том, что очень маленькие фрагменты изображения чертежа напоминают букву «С», которая в особых случаях вырождается в прямую линию. Дальнейшие рассуждения сводятся к выбору способа определения соответствия объекта на изображении форме «С». Отдельное влияние уделено вопросам настройки порога бинаризации в зависимости от яркости изображения.

В работе представлена математическая модель предложенного метода, его алгоритмическое описание. Проведены эксперименты по исследованию качественных и количественных показателей эффективности. Качественная оценка выполнялась путем сравнения результата бинаризации с эталонным черно-белым изображением. В качестве критериев количественного оценивания рассмотрены такие критерии, как время выполнения алгоритма, полнота и *f*-мера. Представлены результаты проведенных экспериментов, показывающие превосходство предложенной в работе методики бинаризации растровых изображений технических чертежей, как количественно, так и качественно.

Ключевые слова: растровое изображение, бинаризация, технический чертеж, порог бинаризации, яркость, оттенок серого, полезный сигнал, фон, шум, ошибка

НОМЕНКЛАТУРА

B – матрица, задающая полученное бинарное изображение;

b_{xy} – элемент матрицы B, определяющий цвет в точке (x, y) получаемого в ходе бинаризации черно-белого изображения

BO – число точек, принадлежащих фону на «идеальном» бинарном изображении;

BT – число точек, принадлежащих фону на бинарном изображении;

F – F-мера результатов бинаризации;

FO – число точек, принадлежащих объекту на «идеальном» бинарном изображении;

FR – число точек, одновременно принадлежащих объекту на «идеальном» бинарном изображении и изображении, полученном в ходе преобразования;

FT – число точек, принадлежащих объекту на бинарном изображении;

H – длина исходного изображения;

i – индекс точки, принадлежащей стороне «окна» вокруг точки (x, y);

$L(P)$ – булевая функция, определяющая превышает ли разность среднего значения серого в окрестности и значения серого в некоторой точке, порог T;

m – смещение по оси x;

ME – ошибка классификации;

n – смещение по оси y;

P – матрица, задающая исходное изображение в градациях серого;

$\overline{P}$ – среднее значение серого в окрестности точки;

p_{xy} – элемент матрицы P, определяющий тон серого в точке (x, y) исходного изображения;

P_i – точка, принадлежащая окрестности точки (x, y);

P_i' – точка, принадлежащая окрестности точки (x, y), расположенная напротив точки P_i;

$P_{\max}$ – максимальное значение серого в окрестности точки;

$P_{\min}$ – минимальное значение серого в окрестности точки;

Q – правило преобразования исходного изображения в градациях серого в черно-белый формат;

SW – длина стороны локального окна;

T – локальный порог бинаризации;

$T1$ – глобальный порог бинаризации;

W – ширина исходного изображения;

x – координата точки по оси x;

y – координата точки по оси y;

α – коэффициент поправки;

ΔT – корректировка глобального порога бинаризации.

ВВЕДЕНИЕ

Изображения технических чертежей следует сканировать в цветной палитре или в палитре оттенков серого, так как в этом случае получается изображение более высокого качества, чем при использовании черно-белого варианта. Данное утверждение относится, прежде всего, к документам, выполненным вручную чертежными шрифтами, и к многократным копиям документов. В тоже время распознавание, а также предварительную фильтрацию и скелетизацию проще выполнять для монохромного изображения. Кроме того, обработка монохромных изображений требует меньших затрат времени и ресурсов вычислительной системы. Таким образом, становится актуальной проблема преобразования цветного или полутонового изображения в его монохромный эквивалент. Такая задача называется бинаризацией и по своей сути представляет отделение полезного сигнала от фона.

DOI 10.15588/1607-3274-2015-2-8

Объект исследования представляет собой процесс бинаризации цветных или полутоновых изображений технических чертежей. Предметом исследования в работе являются специфические особенности изображений технических чертежей, влияющие на качество результирующего изображения, а также способы устранения недостатков, полученных в ходе обработки изображений стандартными алгоритмами бинаризации.

Цель работы состоит в разработке и реализации эффективного алгоритма бинаризации изображений, адаптированного под специфические особенности технических чертежей. Для достижения поставленной цели необходимо решить следующие задачи:

– определить специфические особенности изображений технических чертежей, влияющие на результат бинаризации;

– предложить способ устранения недостатков обработки изображений универсальными алгоритмами;

– реализовать алгоритм бинаризации растровых изображений технических чертежей;

– проанализировать полученные результаты.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть некоторое исходное изображение задано в виде прямоугольной матрицы P размерностью $W \times H$ (1):

$$P = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0,W-1} \\ p_{10} & p_{11} & \cdots & p_{1,W-1} \\ \cdots & \cdots & \cdots & \cdots \\ p_{H-1,0} & p_{H-1,1} & \cdots & p_{H-1,W-1} \end{pmatrix}. \quad (1)$$

Каждый элемент p_{xy} матрицы P определяет тон серого в соответствующей точке изображения и может принимать любые целые значения от 0, что соответствует черному цвету в точке (x,y) исходного изображения, до 255 – белому [1–3]. Необходимо определить правило Q, такое, что матрица B (2), полученная как $P \xrightarrow{Q} B$, определяет бинарное представление исходного изображения, при этом значение 0 будет соответствовать белому цвету, а значение 1 – черному:

$$B = \begin{pmatrix} b_{00} & b_{01} & \cdots & b_{0,W-1} \\ b_{10} & b_{11} & \cdots & b_{1,W-1} \\ \cdots & \cdots & \cdots & \cdots \\ b_{H-1,0} & b_{H-1,1} & \cdots & b_{H-1,W-1} \end{pmatrix}. \quad (2)$$

Таким образом, правило Q представляет собой способ классификации точек исходного изображения на полезный сигнал и фон на основании некоторого критерия, называемого порогом T [4], метод вычисления которого определяет Q (3):

$$b_{xy} = \begin{cases} 1, & \text{если } p_{xy} < T : Q; \\ 0, & \text{иначе.} \end{cases} \quad (3)$$

В качестве целевого критерия Q при построении продукционных правил может выступать:

– ошибка классификации [5], которая рассчитывается по формуле: $ME = 1 - \dfrac{|BO \cap BT| + |FO \cap FT|}{|BO| + |FO|}$;

– F-мера, представляющая собой гармоническое среднее полноты и точности [5], рассчитывается по формуле: $F = \dfrac{2 \times \text{полнота} \times \text{точность}}{\text{полнота} + \text{точность}}$, где полнота и точность вычисляется соответственно по формулам: $\text{Полнота} = \dfrac{FR}{FO}$, $\text{Точность} = \dfrac{FR}{FT}$.

2 ОБЗОР ЛИТЕРАТУРЫ

На сегодняшний день существует большое число готовых реализаций бинаризации изображений, которые описаны, например в работах [5–11]. Все они сводятся к сравнению яркости каждой точки P_{xy} с некоторым пороговым значением T и в зависимости от результата сравнения отнесения ее к «белым» или «черным». Принципиально эти методы отличаются лишь способом вычисления порога яркости и принципом построения пороговой поверхности (методы локальной и глобальной бинаризации) [6, 12–13].

Оба подхода имеют некоторые недостатки. Они обусловлены тем, что методы глобальной бинаризации не учитывают характеристики точек в их окрестности, в то время как локальные не учитывают характеристики всего изображения в целом, что, в конечном счете, влияет на качество бинаризации. Устранение этих недостатков предложено в работе [14], где описывается гибридный метод бинаризации, предполагающий использование нескольких пороговых величин. Этот метод показал хорошие результаты при обработке текстовых документов.

В [14] представлен метод бинаризации изображений с адаптивным порогом на основе анализа гистограмм в локальных регионах и автоматическом определении окрестности с центром в заданной точке (рис. 1). Порог бинаризации для каждого локального фрагмента вычисляется индивидуально.

В процессе бинаризации все точки обрабатываются последовательно, а результат определяется формулой [4]:

$$b_{xy} = \begin{cases} 1, \text{ если } \vee_{i=0}^{3} \left[L(P_i) \wedge L(P_i') \wedge L(P_{i+1}) \wedge L(P_{i+1}') \right] = \text{истина}; \\ 0, \text{ в противном случае.} \end{cases}$$

$$P_i' = P_{(i+4) \bmod 8}, \quad \text{при } i = 0, 1, ..., 3,$$

$$L(P) = \overline{P} - p_{xy} > T,$$

$$\overline{P} = \frac{1}{(2 \cdot SW + 1)^2} \times \sum_{-SW \le m \le SW} \sum_{-SW \le n \le SW} p_{x+m, y+n}. \quad (4)$$

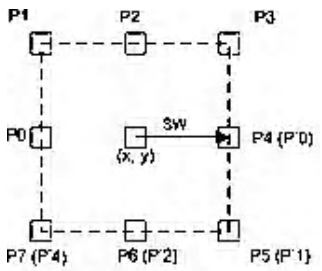


Рисунок 1 – Соседние локальные регионы вокруг точки

Этот способ позволяет получить достаточно хороший результат, но преимущественно он ориентирован на изображения текстовых документов, поэтому при попытке обработки таким методом изображений технических чертежей зачастую возникает проблема нарушения исходной топологии объектов, что в дальнейшем оказывает существенное влияние на результат последующих операций над изображением и, следовательно, быть использован не может.

В работе [15] обнаружено, что изображения технических чертежей обычно состоят из прямых линий и дуг окружности. Благодаря чему сделан вывод: очень маленькие фрагменты изображения имеют форму, близкую к форме буквы «С». Таким образом, несмотря на то, что изолированные точки имеют более низкую шкалу оттенков серого, они не образуют форму фигуры «С» в совокупности со своими соседями и, следовательно, не могут быть бинаризованы как точки объектов.

В качестве примера рассмотрим рисунок 2. Типичная С-форма создается точками *P*0, *P*1, *P*2 или *P*2, *P*3, *P*4 или *P*4, *P*5, *P*6 или *P*6, *P*7, *P*0. Точка p_{xy} бинаризуется как точка объекта, если в локальной области с заданными геометрическими размерами образуется по крайней мере одна область в форме буквы «С». Исключительной ситуацией С-формы является ситуация, когда пиксели с низким оттенком серого образуют прямую линию. В этом случае пиксели также принимаются как пиксели объекта.

Для локальных областей, уровень яркости которых отличается от яркости изображения в целом, глобальный порог бинаризации не будет соответствовать значению полученному для всего изображения и приведет к неверному результату [12]. Для корректировки порога бинаризации в таких ситуациях вводится коэффициент поправки α, значение которого определяется исходя из следующего выражения.

$$\alpha = \begin{cases} 1, & \text{если} \quad p_{xy} < 90; \\ 0{,}33, & \text{если} \quad 90 \le p_{xy} \le 170; \\ 0{,}1, & \text{если} \quad p_{xy} > 170, \end{cases}$$

$$T1 \to T + \Delta T \ ,$$

где ΔT – небольшое целое число в пределах [–20; 20]. Т. к. $0 < \alpha \le 1$ и чаще всего $\alpha = 0{,}33$, то параметр α может быть использован для корректировки порога бинаризации на величину в интервале [–6; 6].

3 МАТЕРИАЛЫ И МЕТОДЫ

Примеры инженерных чертежей представлены на рисунке 3. а) синька; б) чертеж, выполненный от руки; в) печатный чертеж.

Изображения инженерных чертежей имеют ряд специфических особенностей.

1) изображение состоит преимущественно из прямых линий, текстовых блоков и дуг окружности;

2) линии на чертеже имеют различные геометрические размеры;

3) наличие фонового шума;

4) большой диапазон яркости для изображений технических чертежей различных типов.

Из рисунка 3 видно, что различные виды инженерных чертежей имеют различные показатели качества, яркости и контрастности. Наиболее подходящим для обработки такого типа изображений является метод, описанный в [4]. В случае выбора локальной области надлежащего размера (*SW*=3), а также порога бинаризации в пределах от 20 до 120 данный метод дает весьма неплохие результаты. Исключение составляют линии единичной толщины, а также точки пересечения линий и дуг (рисунок 4). По сути это ситуация, когда рассматривается вырожденная форма «С» в виде прямой линии. Выделенная точка на рисунке 4 принадлежит объекту, однако с точки зрения метода с-форм она будет бинаризована как точка фона, что приведет к нарушению целостности объекта. Уменьшение размера окна до (*SW*=1) не позволяет существенно улучшить результат, что свидетельствует о целесообразности доработки данного метода.

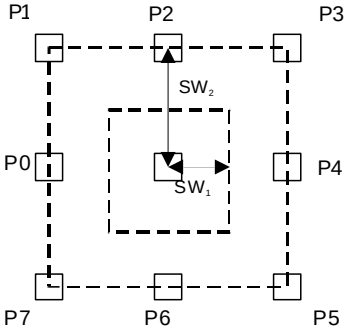
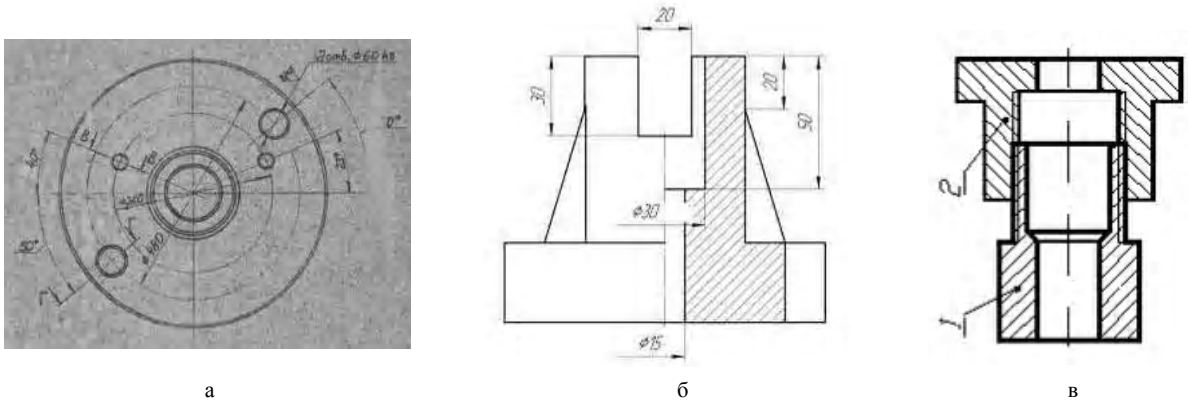


Рисунок 2 – Соседние локальные регионы для окна со сторонами *W*1 и *W*2



а б в

Рисунок 3 – Изображения инженерных чертежей: а – на синьке, б – ручной чертеж на ватмане в – распечатанный чертеж

Для решения проблемы «потери» тонких линий предлагается:

1) производить поиск С-форм для локальных областей различных размеров (*SW*=3) и (*SW*=1);

2) включить в поиск возможность исключительный ситуации, когда *c*-форма вырождается в прямую линию;

3) рассмотреть ситуацию, когда происходит обработка конечной точки линии или дуги.

Данное утверждение может быть реализовано с помощью шаблонов путем поточечного сравнения исходного изображения с каждым эталонным вариантом, а также с помощью математических уравнений. Наиболее удобным вариантом для дальнейшей реализации является вариант решения задачи математическим путем.

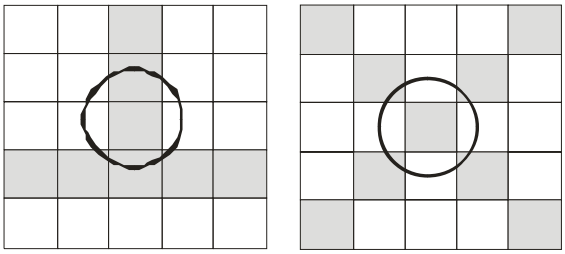


Рисунок 4 – Примеры линий объекта не образующих форму «С»

Так, например, для проверки соответствия вырождается ли форма *c* в прямую линию необходимо формулу (1) переписать в следующем виде:

$$b_{xy} = \begin{cases} 1, & \text{если } \vee_{i=0}^{4} \left[L(P_i) \wedge L(P_i') \right] = \text{истина}; \\ 0, & \text{в противном случае.} \end{cases}$$

А далее воспользоваться ранее полученной методикой.

Модифицированный таким образом метод, позволит выполнить корректную бинаризацию точек объекта в особых ситуациях, аналогичных приведенным на рис. 4.

Методика, описанная в [4] предполагает расчет пороговых значений на неидеально белой бумаге. В тоже время большинство чертежей выполняются на белом ватмане. Следовательно, фон таких изображений будет иметь высокую яркость (большую 170). Таким образом, для большей части изображения коэффициент α будет составлять 0,1. Поэтому с целью достижения корректного различения фона и объектов диапазон корректировки необходимо увеличить до 20–120.

Предложенный алгоритм бинаризации реализуется следующей последовательностью шагов (рис. 5).

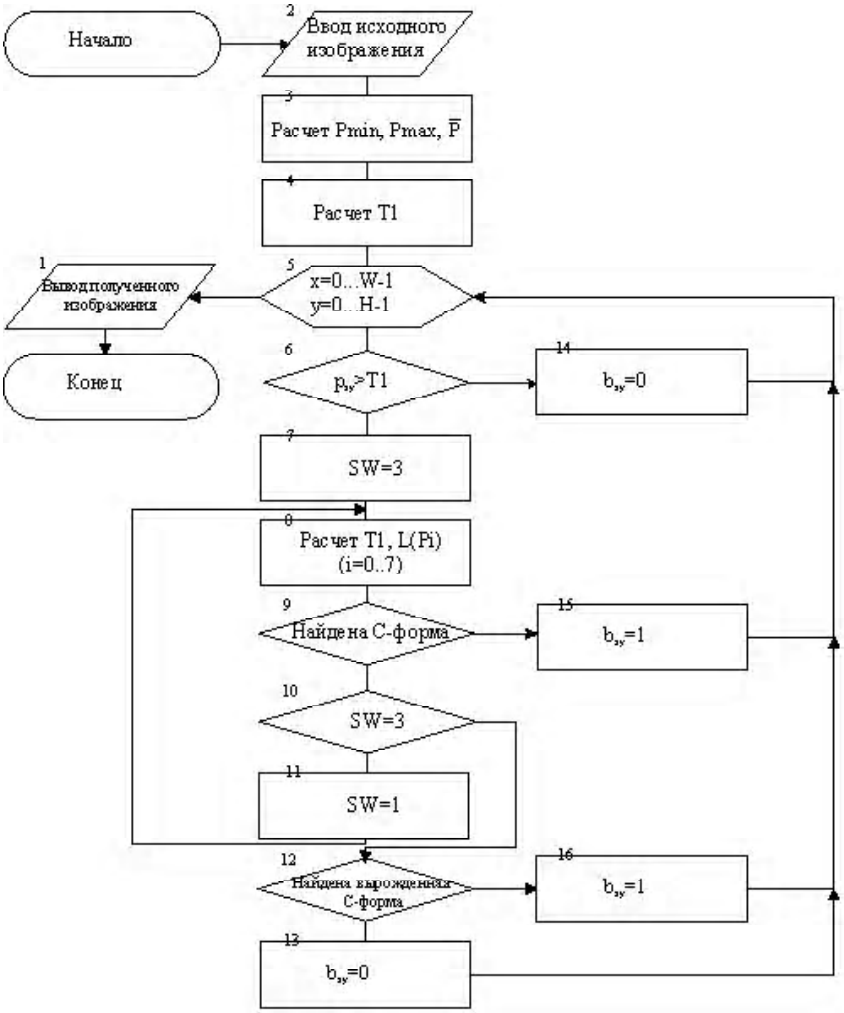


Рисунок 5 – Схема предложенного алгоритма

1. Рассчитываем минимальное, максимальное, среднее значения яркости цвета для всех точек изображения. Полученные результаты позволяют определить класс исходного изображения (ватман, синька).

2. Исходя из полученных результатов, определяем поправочный коэффициент и глобальный порог бинаризации T_1.

3. Просматриваем все изображение и бинаризуем каждую точку, исходя из того, что объект=1, фон=0.

Если яркость обрабатываемой точки больше глобального порога $T1$, то после бинаризации данная точка однозначно определяется как точка фона b_{xy}=0.

В противном случае обрабатываемая точка берется как центр квадрата размером $(2SW+1)\times(2SW+1)$, рассчитывается среднее значение яркости в данной области и динамический локальный порог $T = \alpha \times \Delta T$, ΔT – диапазон поправочный коэффициент порога бинаризации.

Определяем яркость 8 соседних точек. Если для каких-либо 3-х последовательных, либо лежащих на одной прямой точек, выполняется условие $|\overline{P}-p_{xy}|>T$ обрабатываемая точка бинаризуется как b_{xy}=1, т.е точка, принадлежащая объекту.

В противном случае обрабатываемая точка бинаризуется как b=0.

4. Если имеются необработанные точки возвращаемся к шагу 2.

5. Выводим результирующее двухградационное изображение.

4 ЭКСПЕРИМЕНТЫ

Предложенный алгоритм реализован в среде C++ Builder XE5. Разработанный программный продукт позволяет выполнить бинаризацию изображения каждым из рассмотренных в работе методом и наглядно оценить результат (рис. 6).

Для выполнения количественной оценки результата предусматривается возможность сравнения результата бинаризации с идеальным черно-белым образцом и расчета основных оценочных критериев, на основании значений которых можно сделать вывод об эффективности каждого метода. Программа снабжена гибкой системой настроек каждого метода.

5 РЕЗУЛЬТАТЫ

В тестовой выборке участвовало 20 образцов полутоновых изображений, площадью от 500 до 50000 точек. Сравнение результатов бинаризации каждым методом осуществлялось по критериям качество и время. Наиболее сложно обстоит вопрос с определением качества бинаризации, т.к. задача определения точки как точки фона или точки объекта не является тривиальной и в некоторых случаях ее затруднительно решить даже при ручном разборе изображения.

На международном конкурсе DIBCO [16–18] алгоритмов бинаризации для оценки качества получаемых результатов предлагаются следующие метрики: ошибка классификации (misclassification error, *ME*), полнота (recall), точность (precision) и F-мера (F-measure). Чем выше значение F-меры, тем лучше производительность алгоритма

Значения этих метрик для ряда тестовых образцов, после их обработки логическим пороговым методом, методом *c*-форм и нашим методом представлены в таблице 1.

Наглядно результаты тестирования методов бинаризации показаны на графиках (рис. 7–9).

6 ОБСУЖДЕНИЕ

Результаты анализа данных, представленных в табл. 1, показывают, что наилучшее качество результирующего бинарного изображения обеспечивает метод, предложенный в работе. Так, например, этот метод дает самое

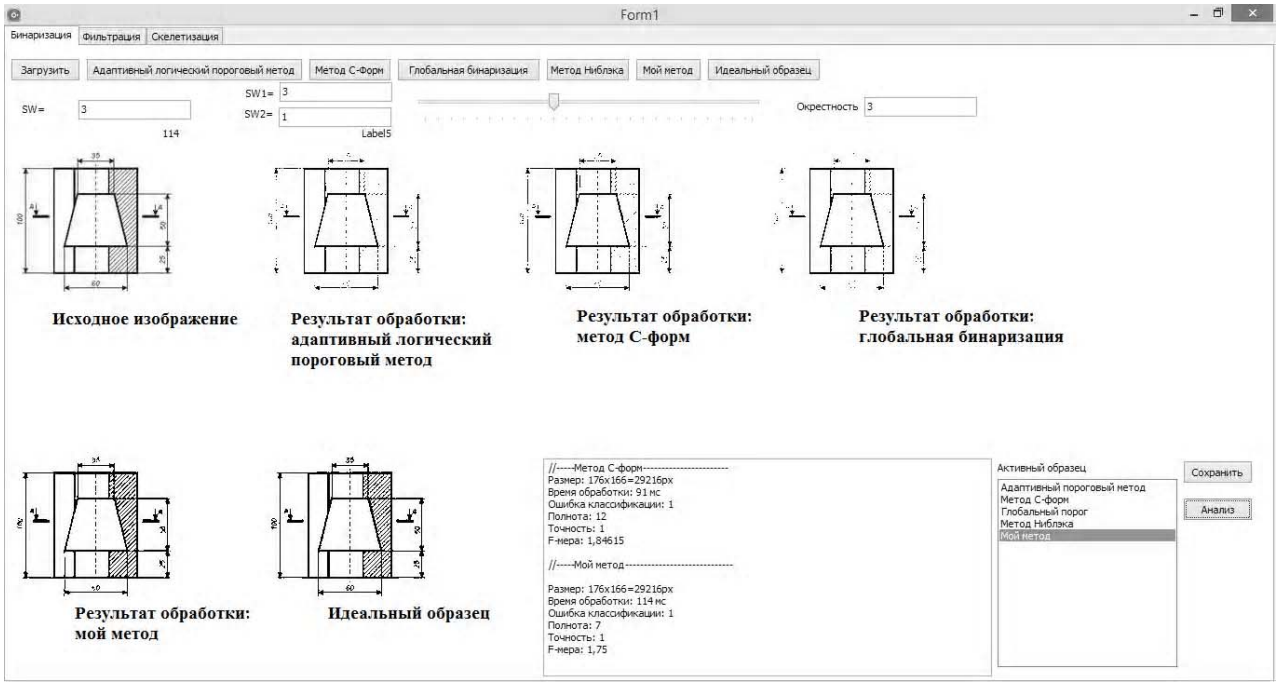


Рисунок 6 – Экранная форма программы, используемой при проведении экспериментов

высокое значение F-меры для всех без исключения тестовых образцов.

Качественный анализ изображений, полученных в ходе обработки рассмотренными в работе методами бинаризации, показал, что в результате обработки изображений логическим пороговым методом образуется шум, особенно вблизи границ объектов, который значительно ухудшает качество изображения, кроме того, имеет место потеря некоторых тонких линий. Метод *c*-форм дает хорошие результаты на изображениях, толщина линий которых приблизительно одинакова, а линии единичной толщины отсутствуют, однако корректно обработать тонкие линии данный метод не способен. Все эти недостатки отсутствуют в изображениях, полученных в результате бинаризации рассмотренным в работе методом.

Таблица 1 – Результаты тестирования методов бинаризации

Номер изображения	Размер	Логический пороговый метод			Метод с-форм			Глобальная бинаризация			Новый метод		
		Время	Полнота	F-мера	Время	Полнота	F-мера	Время	Полнота	F-мера	Время	Полнота	F-мера
1	130×135	34	18	1,89	114	16	1,88	1	22	1,91	126	9	1,8
2	176×166	36	14	1,86	72	12	1,84	1	17	1,89	90	7	1,75
3	147×192	58	14	1,87	166	13	1,86	1	18	1,89	141	8	1,77
4	244×321	137	7	1,75	254	7	1,75	1	9	1,8	303	6	1,71
5	247×321	419	8	1,77	620	7	1,75	1	12	1,84	401	6	1,74
6	170×190	290	12	1,85	148	10	1,82	1	13	1,86	304	8	1,78
7	133×123	30	8	1,78	46	8	1,78	1	12	1,85	70	6	1,71
8	131×190	26	86	1,98	101	73	1,97	1	415	1,99	89	22	1,91
9	267×195	215	37	1,95	456	27	1,93	1	229	1,99	564	20	1,9
10	144×321	64	45	1,96	162	42	1,95	1	83	1,98	187	18	1,89
11	267×226	104	37	1,95	334	33	1,94	1	48	1,96	356	16	1,88
12	267×205	96	11	1,83	143	11	1,83	1	16	1,88	190	9	1,8
13	161×200	75	8	1,78	151	8	1,78	1	18	1,89	187	6	1,71
14	104×113	38	28	1,93	149	23	1,92	1	45	1,96	66	7	1,75
15	267×321	291	38	1,95	705	17	1,88	1	31	1,94	905	7	1,75
16	207×214	46	1655	1,99	123	459	1,99	1	599	1,99	146	19	1,9
17	213×321	92	539	1,99	247	439	1,99	1	1514	1,99	246	3	3
18	267×147	44	8	1,78	100	8	1,78	1	10	1,82	127	5	1,65
19	236×162	43	28	1,93	103	21	1,91	1	71	1,97	30	15	1,89
20	117×96	19	26	1,93	31	22	1,91	1	30	1,94	37	7	1,75

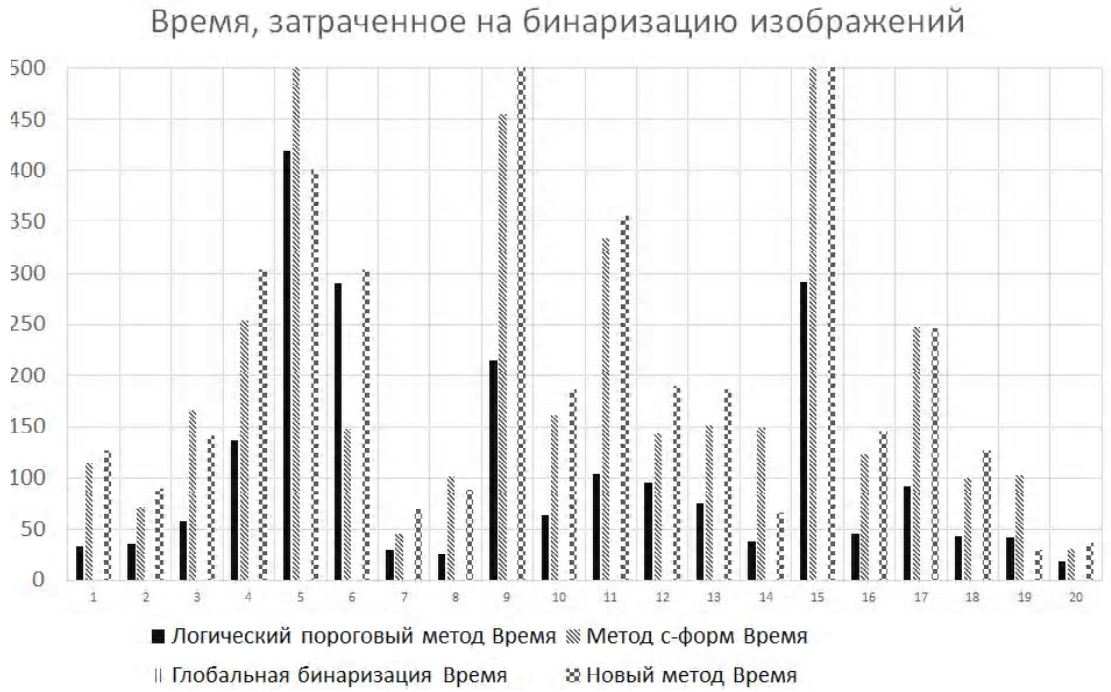


Рисунок 7 – Гистограмма времени, затраченного на бинаризацию

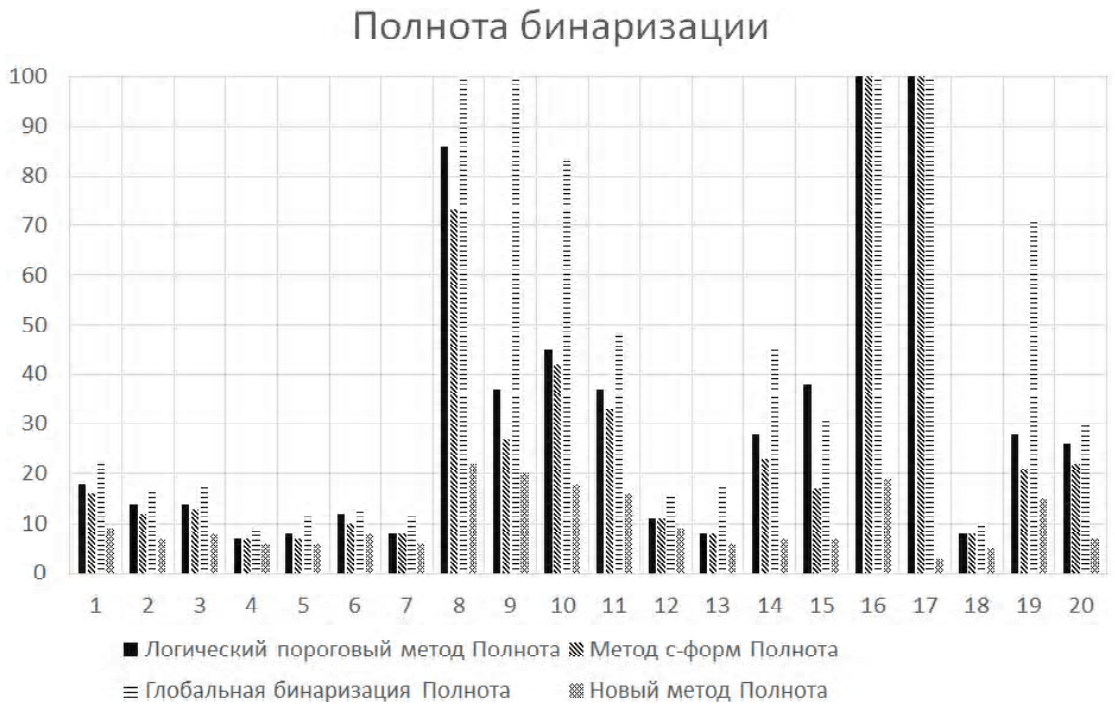


Рисунок 8 – Полнота бинаризации

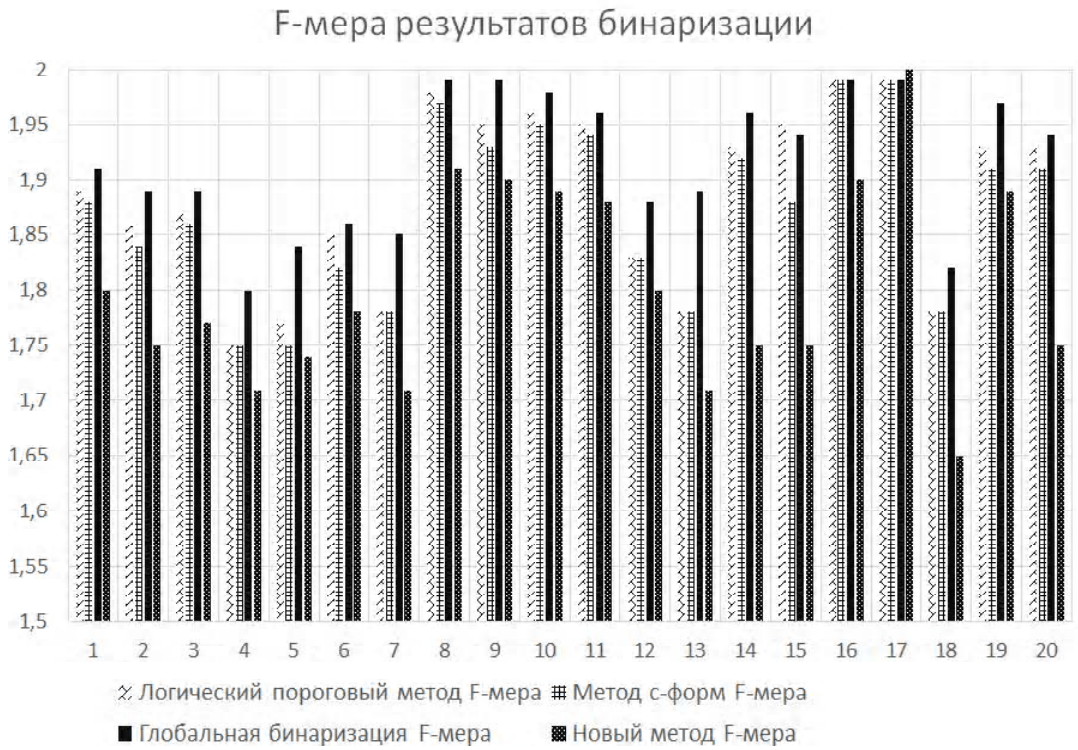


Рисунок 9 – F-мера результатов бинаризации

Вторым критерием оценки качества работы алгоритма является время его выполнения соизмеримо с размерами обрабатываемой области. Характеристики скорости выполнения сравниваемых алгоритмов также представлены в таблице 1.

Анализ времени обработки каждого тестового образца показывает, что наиболее быстрым является метод, использующий глобальный порог, что объясняется наиболее простым алгоритмом обработки каждой точки изображения, однако результаты, получаемые таким методом, непригодны для дальнейшей обработки. Среди методов, дающих удовлетворительный результат, наиболее быстрым является метод *c*-форм. Необходимость больших временных затрат для выполнения предложен-

ного метода объясняется выполнением дополнительных проверок на наличие «особых ситуаций», результат которых оказывает значительное влияние на формируемое в процессе бинаризации изображение. Таким образом, для образцов, не содержащих «особых ситуаций», время выполнения обоих методов одинаково либо отличается незначительно.

ВЫВОДЫ

В работе решена актуальная задача преобразования изображений технических чертежей, представленных в градациях серого, в черно-белый формат, удобный для дальнейшей обработки таких изображений.

Научная новизна работы состоит в том, что получила дальнейшее развитие идея применения метода с-форм для бинаризации изображений технических чертежей. Данный метод позволяет получить хорошие результаты при обработке таких сложных ситуаций как тонкие линии и точки пересечения линий, в тоже время метод «отсеивает» случайные шумы в изображении. Второй отличительной особенностью предложенного метода, является автоматическая корректировка порога яркости, что позволяет обрабатывать, как изображения, представленные на белом ватмане, так и на «синьке».

Предложен ряд критериев оценки бинаризации изображений технический чертежей: время выполнения, полнота и F-мера. В качестве альтернативных методов рассматривается логический пороговый метод, метод с-форм, метод глобальной бинаризации. Показано, что наилучшие результате дает предложенный в работе метод. Исключением является время выполнения метода, которое в некоторых ситуациях большее, чем время выполнения других, рассматриваемых в качестве альтернативы вариантов. Это объясняется необходимостью выполнения дополнительных проверок на наличие специфических элементов на изображении и их корректной обработки.

Практическая ценность полученных результатов состоит в том, что разработано программное обеспечение, которое позволяет выполнить бинаризацию растровых изображений технических чертежей, представленных в градациях серого.

Перспективы дальнейших исследований состоят в разработке способа векторизации изображений технических чертежей и реализации на его основе полномасштабной системы обработки и распознавания растровых изображений технических чертежей и их дальнейшего преобразования в формат современных CAD-систем.

БЛАГОДАРНОСТИ

Автор выражает благодарность научному руководителю И. С. Грунскому за внимание к работе и полезные замечания.

СПИСОК ЛИТЕРАТУРЫ

1. Гонсалез Р. Цифровая обработка изображений в среде Matlab / Р. Гонсалез, Р. Вудс, С. Эддинс. – М. : Техносфера, 2006. – 616 с.
2. Абламейко С. В. Обработка изображений: технология, методы, применение / С. В. Абламейко, Д. М. Лагуновский. – Минск: Амалфея, 2000. – 304 с.
3. Красильников Н. Н. Цифровая обработка 2D- и 3D-изображений : учебное пособие / Н. Н. Красильников. – СПб. : БХВ-Петербург, 2011. – 608 с.
4. Методы компьютерной обработки изображений / под ред. В. А. Сойфера. – М. : Физматлит, 2003. – 784 с.
5. Грузман И. С. Цифровая обработка изображений в информационных системах : учебное пособие / И. С. Грузман, В. С. Киричук, В. П. Косых и др. – Новосибирск : Издательство НГТУ, 2002. – 352 с.
6. Gonzalez R. Digital image processing / R. C. Gonzalez, R. E.Woods. – Prentice Hall, 2007. – 976 p.
7. Rogers D. Matematical elements for computer graphics / David F. Rogers, J. Alan Adams. – New York : McGraw-Hill, 2002. – 630 p.
8. Otsu N. A threshold selection method from grey level histogram / N. Otsu // IEEE Transactions on Systems, Man and Cybernetics. – 1979. – Vol. 9, № 1. – P. 62–66.
9. Niblack W. An introduction to digital image processing / Niblack W. – Prentice Hall, Englewood Cliffs, 1986. – 231 p.
10. Sauvola J. Adaptive Document Binarization / J. Sauvola, M. Pietikainen // Pattern Recognition. – 2000. – № 33. – P. 225–236.
11. Анисимов Б. В. Распознавание и цифровая обработка / Б. В. Анисимов, В. Д. Курганов, В. К. Злобин. – М. : Высшая школа, 1983. – 95 с.
12. Navon E. Color image segmentation based on adaptive local thresholds / E. Navon, O. Miller, A. Averbuch // Image and Vision Computing. – 2012. – № 23. – P. 69–85.
13. O'Gorman L. Binarization and multithresholding of document images using connectivity // CVGIP: Graphical Models and Image Processing. – 2004. – Vol. 5, № 6. – P. 494–506.
14. Yang Y. An adaptive logical method for binarization of degraded document images / Y. Yang, H. Yan // Pattern Recognition. – 2010. – Vol. 33. – P. 787–807.
15. Zhao M. An adaptive thresholding method for binarization of blueprint images / M. Zhao, Y. Yang, and H. Yan // Pattern Recognition Letters. – 2010. – Vol. 21. – P. 927–943.
16. Gatos B. Document image binarization contest / B. Gatos, K. Ntirogiannis, I. Pratikakis // DIBCO 2009: 10th International Conference on Document Analysis and Recognition, Spain, Barcelona, 26–29 July, 2009. – Barcelona : Universitat Autrnoma, 2009. – P. 1375–1382.
17. Pratikakis I. Document Image Binarization Contest (DIBCO 2011) / I. Pratikakis, B. Gatos, K. Ntirogiannis // ICDAR 2011: 11th International Conference on Document Analysis and Recognition, China, Beijing, 18–21 September, 2011. – Beijing, 2011. – P. 1506–1510.
18. Pratikakis I. Document Image Binarization Contest (ICDAR 2013) / I. Pratikakis, B. Gatos, K. Ntirogiannis // ICDAR 2013: 12th International Conference on Document Analysis and Recognition, USA, Washington, 25–28 August, 2013. – Washington, 2013. – P. 1471–1476.

Статья поступила в редакцию 04.11.2014.
После доработки 05.11.2014.

Молчанова В. С.
Ст. викладач кафедри інформатики ДВНЗ «Приазовський державний технічний університет», Маріуполь, Україна

АДАПТИВНИЙ ПОРОГОВИЙ МЕТОД БІНАРИЗАЦІЇ РАСТРОВИХ ЗОБРАЖЕНЬ ТЕХНІЧНИХ КРЕСЛЕНЬ

У статті пропонується метод адаптивної порогової бінаризації растрових зображень технічних креслень. Проаналізовано специфічні особливості зображень технічних креслень та їх вплив на результат бінаризації відомими універсальними методами. Виявлено артефакти, що виникають на результуючих бінарних зображеннях та запропоновано способи їх усунення.

В основу запропонованого методу покладено ідею про те, що дуже маленькі фрагменти зображення креслення нагадують букву «С», яка в особливих випадках вироджується у пряму лінію. Подальші роздуми зводяться до вибору способу визначення відповідності об'єкта на зображенні формі «С». Окремий вплив приділено питанням Налаштування порога бінарізації залежно від яскравості зображення.

В роботі подана математична модель запропонованого методу, його алгоритмічний опис. Проведені експерименти з дослідження якісних і кількісних показників ефективності. Якісна оцінка виконувалася шляхом порівняння результату бінарізації з еталонним чорно-білим зображенням. В якості критеріїв кількісного оцінювання розглянуті такі критерії, як час виконання алгоритму, повнота і *f*-міра. Представлені результати проведених експериментів, що показують перевагу запропонованої в роботі методики бінарізації растрових зображень технічних креслень, як кількісно, так і якісно.

Ключові слова: растрове зображення, бінаризація, технічне креслення, поріг бінаризація, яркость, відтінок сірого, корисний сигнал, фон, шум, помилка.

Molchanova V. S.

Senior Lecturer of Department of Informatics of Azov State Technical University, Mariupol, Ukraine

THE ADAPTIVE THRESHOLD BINARIZATION METHOD OF RASTER IMAGES OF TECHNICAL DRAWINGS

The paper proposes a method of adaptive threshold binarization of technical drawings raster images. The specific image features of technical drawings and their influence on the result of binarization of known universal methods are analysed. The artifacts that arise on the resulting binary images are identified and ways to address them are suggested.

The proposed method is based on the idea that very small parts of the image of the drawing resembles the letter «C», which in special cases degenerates into a straight line. Further arguments boil down to the choice of method for determining the conformity of the object in the image form «C». Separate effect given to the issues of binarization threshold settings depending on the brightness of the image.

The paper presents a mathematical model of the proposed method, its algorithmic description and experimental qualitative and quantitative evidence of efficacy. The quantitative evaluation was performed by comparing the results with the reference binarization black and white image. As evaluation criteria considered criteria such as time of the algorithm, completeness and f-measure. The results of the conducted experiments demonstrating the superiority of the proposed methodology in the raster image binarization technical drawings, both quantitatively and qualitatively.

Keywords: bitmap, binarization, technical drawings, threshold binarization, brightness, shade of gray, useful signal, background, noise, error.

REFERENCES

1. Gonsalez R., Vuds R., Jeddins S. Cifrovaja obrabotka izobrazhenij v srede Matlab. Moscow, Tehnosfera, 2006, 616 p.
2. Ablamejko S. V. Lagunovskij D. M. Obrabotka izobrazhenij: tehnologija, metody, primenenie. Minsk, Amalfeja, 2000, 304 p.
3. Krasil'nikov N. N. Cifrovaja obrabotka 2D- i 3D-izobrazhenij: uchebnoe posobie. SPb, BHV-Peterburg, 2011, 608 p.
4. Metody komp'juternoj obrabotki izobrazhenij / pod red. V. A. Sojfera. Moscow, Fizmatlit, 2003, 784 p.
5. Gruzman I. S. Kirichuk V. S., Kosyh V. P. i dr. Cifrovaja obrabotka izobrazhenij v informacionnyh sistemah: uchebnoe posobie. Novosibirsk, Izdatel'stvo NGTU, 2002, 352 p.
6. Gonzalez R., Woods R. E. Digital image processing. Prentice Hall, 2007, 976 p.
7. Rogers D., Adams J. A. Matematical elements for computer graphics. New York, McGraw-Hill, 2002, 630 p.
8. Otsu N. A threshold selection method from grey level histogram, *IEEE Transactions on Systems, Man and Cybernetics*, 1979, Vol. 9, No. 1, pp. 62–66.
9. Niblack W. An introduction to digital image processing. Prentice Hall, Englewood Cliffs, 1986, 231 p.
10. Sauvola J., Pietikainen M. Adaptive Document Binarization, *Pattern Recognition*, 2000. No. 33, pp. 225–236.
11. Anisimov B. V., Kurganov V. D., Zlobin V. K. Raspoznavanie i cifrovaja obrabotka izobrazhenij. Moscow, Vysshaja shkola, 1983, 295 p.
12. Navon E., Miller O., Averbuch A. Color image segmentation based on adaptive local thresholds, *Image and Vision Computing,* 2012, No. 23, pp. 69–85.
13. O'Gorman L. Binarization and multithresholding of document images using connectivity, *CVGIP: Graphical Models and Image Processing*, 2004, Vol. 5, No. 6, pp. 494–506.
14. Yang Y., Yan H. An adaptive logical method for binarization of degraded document images, *Pattern Recognition,* 2010, Vol. 33, pp. 787–807.
15. Zhao M., Yang Y., and Yan H. An adaptive thresholding method for binarization of blueprint images, *Pattern Recognition Letters*, 2010, Vol. 21, pp. 927–943
16. Gatos B., Ntirogiannis K., Pratikakis I. Document image binarization contest, *DIBCO 2009: 10th International Conference on Document Analysis and Recognition, Spain, Barcelona, 26–29 July, 2009*. Barcelona, Universitat Autrnoma, 2009, pp. 1375–1382.
17. Pratikakis I., Gatos B., Ntirogiannis K. Document Image Binarization Contest (DIBCO 2011), *ICDAR 2011: 11th International Conference on Document Analysis and Recognition, China, Beijing, 18–21 September, 2011*. Beijing, 2011, pp. 1506–1510.
18. Pratikakis I., Gatos B., Ntirogiannis K. Document Image Binarization Contest (ICDAR 2013), *ICDAR 2013: 12th International Conference on Document Analysis and Recognition, USA, Washington, 25–28 August, 2013*. Washington, 2013, pp. 1471–1476.

УДК 004.056.54:932.2

Прогонов Д. А.[1], Кущ С. Н.[2]

[1]Магистр, аспирант кафедры физико-технических средств защиты информации, Физико-технический институт, Национальный технический университет Украины «Киевский политехнический институт», Киев, Украина
[2]Канд. техн. наук, доцент, доцент кафедры физико-технических средств защиты информации, Физико-технический институт, Национальный технический университет Украины «Киевский политехнический институт», Киев, Украина

СПЕКТРАЛЬНЫЙ АНАЛИЗ СТЕГАНОГРАММ

Для повышения стойкости стеганоалгоритмов к известным методам стегоанализа широко применяются методы скрытия сообщений в области преобразования цифровых изображений. Использование при формировании стеганограмм стандартных методов обработки изображений-контейнеров и стегоданных – двумерных дискретных косинусного и вейвлет преобразований – позволяет минимизировать искажения скрытых сообщений при компрессии изображений.

В работе для выявления факта встраивания стегоданных с применением стандартных преобразований контейнера и стегоданных предложено использовать методы спектрального анализа изображений: анализ спектра сингулярностей и мультифрактальный анализ. Преимуществом данного подхода, по сравнению со статистическими методами стегоанализа, является возможность исследования одновременно нескольких компонент изображения-контейнера, что позволяет локализировать область встраивания.

Анализ эффективности применения методов спектрального анализа был проведен для случая использвания одноэтапных и двухэтапных методов формирования стеганограмм. Показано, что скрытие данных приводит к изменениям всех компонент сингулярного разложения изображения-контейнера, а также появлению характерных искажений мультифрактальных характеристик стеганограмм – увеличения мощности множества монофрактальных компонент изображений. Выявленные демаскирующие признаки наличия встроенных сообщений являются устойчивыми в широком диапазоне степеней заполнения контейнера стегоданными, что упрощает задачу обнаружения сообщений, скрытых в области преобразования изображений.

Ключевые слова: цифровые изображения, пассивный стегоанализ, анализ спектра сингулярностей, мультифрактальный анализ.

НОМЕНКЛАТУРА

J+SRM – JPEG & Spatial Rich Model;

PSRM – Projection Spatial Rich Model;

RGB – аддитивная цветовая модель цифровых изображений (Red/Green/Blue);

UHD – Ultra High Definition;

YCbCr – цветоразностная цветовая модель цифрового изображения (Y – компонента яркости; Cb/Cr – синяя/красная цветоразностные компоненты);

YIQ – цветоразностная цветовая модель цифрового изображения (Y – компонента яркости; I/Q – синфазная/квадратурная цветоразностные компоненты);

AMBC – адаптивные методы встраивания стегоданных;

АСС – анализ спектра сингулярностей;

ДДВП – двумерное дискретное вейвлет-преобразования;

ДКП – дискретное косинусное преобразование;

КМ – корреляционная матрица;

МФА – мультифрактальный анализ;

МФС – мультифрактальный спектр;

ПА – преобразование Арнольда;

СОФР – спектр обобщенных фрактальных размерностей;

ЦИ – цифровое изображение;

a – средняя «шероховатость» текстуры цифрового изображения;

C – корреляционная матрица;

D_1 – информационная размерность мультифрактала;

D_2 – корреляционная размерность мультифрактала;

$D_{+\infty}\left(D_{-\infty}\right)$ – граничная обобщенная фрактальная размерность мультифрактала;

D_q – обобщенная фрактальная размерность мультифрактала;

$D_{x,y}$ – стегоданные (полутоновое цифровое изображение);

$D[\cdot]$ – дисперсия значений;

$E[\cdot]$ – среднее значение;

$f(\alpha)$ – функция мультифрактального спектра;

$F_{\vartheta}(\cdot)$ – преобразование цифрового изображения в базисе функций $\{\vartheta_i\}_{i=1}^{N_{\vartheta}}$;

G – весовой коэффициент для коэффициентов разложения стегоданных в заданном базисе;

$H^{\Re^{x\times y}}$ – оператор приведения квадратной матрицы к ганкелевому виду (обработка блоками $x\times y$ (элементов));

$I_{x,y}$ – изображение-контейнер;

IQR – интерквартильный размах;

$J\left(D_{x,y}\right)$ – преобразование стегоданных $D_{x,y}$ перед встраиванием в контейнер;

k_{st} – стегоключ, используемый при скрытии данных в изображении-контейнере;

K_{steg} – коэффициенты преобразования стеганограммы;

M_{BH} – блочно-ганкелева матрица M;

M_i – i-я тройка сингулярного разложения;

M^T – транспонированная матрица M;

n_s – относительная высота отдельного блока разбиения матрицы (размер блока – $s\times s$ пикселей);

N_A – количество итераций преобразования Арнольда;

$N(s)$ – количество блоков разбиения матрицы (размер блока – $s\times s$ пикселей);

DOI 10.15588/1607-3274-2015-2-9

P – количество ненулевых элементов вектора собственных чисел;

$p(n_s)$ – распределение вероятностей значений относительных высот отдельных блоков разбиения матрицы (размер блока – $s \times s$ пикселей);

q – масштабирующий показатель степени обобщенной статистической суммы $Z(q,s)$;

$Q_1(Q_3)$ – первый (третий) квартили распределения значений;

$R_{x,y}$ – скользящее окно обработки цифрового изображения;

$S_{x,y}$ – стеганограмма;

$T_C^i\left(T_S^i\right)$ – i-я спектральная характеристика контейнера (стеганограммы);

U – ортонормированная система факторных векторов сингулярного разложения матрицы $\mathrm{M}_{BH}\mathrm{M}_{BH}^T$;

V – ортонормированная система собственных векторов сингулярного разложения матрицы $\mathrm{M}_{BH}\mathrm{M}_{BH}^T$;

$Z(q,s)$ – обобщенная статистическая сумма;

$\alpha(q)$ – вероятность заполнения ячеек разбиения монофрактального подмножества мультифрактала при фиксированном значении масштабирущего показателя степени q обобщенной статистической суммы $Z(q,s)$;

λ_i – i-е собственное число матрицы $\mathrm{M}_{BH}\mathrm{M}_{BH}^T$;

σ – среднеквадратическое отклонение значений яркости пикселей полутонового цифрового изображения;

$\tau(q)$ – функция распределения масштабирующих показателей степени q обобщенной статистической суммы $Z(q,s)$;

Δ_α – ширина мультифрактального спектра;

Δ_b – шаг выборки блоков цифрового изображения;

Δ_f – размах мультифрактального спектра;

Δ_q – шаг выборки значений параметра q;

Λ – подмножество спектральных характеристик ЦИ, значения которых наиболее изменяются при встраивании стегоданных.

ВВЕДЕНИЕ

В современных методах стеганографии ЦИ широко используются АМВС, особенностью которых является локализация области встраивания – скрытие сообщений производится в шумоподобных текстурах изображения-контейнера, что позволяет существенно повысить стойкость получаемых стеганограмм к известным методам пассивного стегоанализа.

Наиболее часто АМВС применяются для встраивания стегоданных в ЦИ, представленные в графических форматах сжатия без потерь. Повышение устойчивости стеганоалгоритмов на основе АМВС, при компрессии или фильтрации изображений-контейнеров, обеспечивается за счет использования при формировании стеганограмм стандартных преобразований, как изображения-контейнера, так и стегоданных – ДКП и ДДВП.

Для обнаружения стеганограмм, сформированных с использованием ДКП, в работе [1] был предложен эффективный метод статистического стегоанализа. На основе данного метода в работе [2] был построен универсальный стегодетектор, позволяющий выявлять стегоданные, встроенные, как в пространственной, так и частотной областях изображений. Существенным ограничением практического применения предложенного стегодетектора является чрезвычайно высокая размерность используемого пространства признаков – 35263 характеристики ЦИ, по совокупности которых, с высокой вероятностью можно определить наличие встроенных в контейнер данных.

Поэтому актуальной является задача поиска новых методов обнаружения стеганограмм, сформированных на основе АМВС, обеспечивающих высокую вероятность распознавания независимо от области скрытия сообщений при использовании пространства признаков сравнительно небольшой размерности.

Одним из подходов к стегоанализу стеганограмм, сформированных на основе АМВС является использование методов структурного и спектрального анализов ЦИ. Необходимо отметить, что локализация области встраивания стегоданных при использовании АМВС существенно ограничивает эффективность применения классических методов спектрального анализа, например, преобразования Фурье. Поэтому, для выявления характерных изменений спектральных компонент стеганограмм, обусловленных встраиванием стегоданных в контейнер, представляет интерес использование комплексных методов спектрального анализа.

Объектом исследований является пассивный стегоанализ стеганограмм, сформированных путем встраивания исходных, либо предварительно обработанных сообщений в области преобразования ЦИ для общего случая применения одно- и многоэтапных АМВС.

Предметом исследований являются структурные и спектральные методы пассивного стегоанализа, включая АСС и МФА цифровых изображений.

Целью работы является исследование возможностей использования методов АСС и МФА для выявления факта скрытия сообщений для общего случая проведения преобразований, как изображения-контейнера, так и стегоданных.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть мы имеем стеганографическую систему скрытой передачи сообщений, основанную на встраивании изображений-стегоданных $D_{x,y}$ в области преобразования полноцветного изображения-контейнера $I_{x,y}$. При формировании стеганограмм проведены преобразования, как изображения-контейнера, так и скрываемого сообщения. Коэффициенты разложения стеганограмм в базисе преобразования $\{\psi_i\}_{i=1}^{N_\psi}$ определены как сумма соответствующих коэффициентов разложения контейнера $K_{\text{cont}} = F_\psi\left(I_{x,y}, k_{st}\right)$ и, взятых с весовым коэффициентом G, скрываемого сообщения $K_{data} = F_\psi\left(D_{x,y}, k_{st}\right)$:

$$K_{\text{steg}} = K_{\text{cont}} + G \cdot K_{\text{data}}. \quad (1)$$

Рассматривается наиболее сложный для проведения стегоанализа случай, когда в качестве стегоключа k_{st} выбраны позиции коэффициентов K_{cont}, измененных при скрытии сообщений, а также параметры преобразования стегоданных $J\left(D_{x,y}\right)$.

Тогда для выявления стеганограмм необходимо выделить подмножество спектральных характеристик ЦИ, которые наиболее сильно изменяются при встраивании стегоданных в области преобразования изображений-контейнеров согласно АМВС. Данная задача сводится к задаче минимизации целевой функции:

$$\frac{1}{\sum_{i\in\Lambda}\left|T_C^i - T_S^i\right|} \to \min.$$

2 ОБЗОР ЛИТЕРАТУРЫ

Локальный характер изменений параметров ЦИ, в случае применения адаптивных методов скрытия сообщений, существенно усложняет проведение стегоанализа с использованием стандартных методов, основанных на использовании простых статистических моделей ЦИ, например, χ^2-теста, RS-анализа.

Для повышения эффективности пассивного стегоанализа стеганограмм в случае применения АМВС был предложен метод, основанный на использовании статистических моделей ЦИ [3], включающих наборы (кластеры) статистических характеристик изображений. При формировании кластеров использовались наиболее общие параметры ЦИ, в частности, гистограммы распределений и матрицы сочетаемости яркости пикселей. Дальнейшее увеличение точности распознавания АМВС было достигнуто путем объединения нескольких кластеров характеристик ЦИ (например, J+SRM-модель [2]), и анализа результатов свертки шумовых компонент ЦИ со случайно выбранными векторами в PSRM-модели [4]. Платой за высокую точность обнаружения стеганограмм являлось существенное возрастание требований к объему тестовой выборки ЦИ и времени настройки стегодетектора.

В работах [5, 6] для исследования изменений параметров ЦИ, обусловленных встраиванием стегоданных с применением ДКП и ДДВП, было предложено использовать специальные методы структурного анализа ЦИ. Было показано, что скрытие сообщений в контейнерах приводит к существенным изменениям мультифрактальных характеристик шумовых компонент ЦИ [5], вид которых зависит от типа использованных преобразований контейнера и наличия/отсутствия этапа предварительной обработки стегоданных [7]. В ряде случаев практическое использование данных методов ограничивает их высокая вычислительная сложность и возможность анализа только шумовых компонент ЦИ. Поэтому представляет интерес использование методов спектрального анализа, позволяющих проводить комплексное исследование изменений параметров ЦИ, для определения факта встраивания стегоданных с применением ДКП и ДДВП.

В работах [5, 6] для исследования изменений параметров ЦИ, обусловленных встраиванием стегоданных с применением ДКП и ДДВП, было предложено использовать специальные методы структурного анализа ЦИ. Было показано, что скрытие сообщений в контейнерах приводит к существенным изменениям мультифрактальных характеристик шумовых компонент ЦИ [5], вид которых зависит от типа использованных преобразований контейнера и наличия/отсутствия этапа предварительной обработки стегоданных [7]. В ряде случаев практическое использование данных методов ограничивает их высокая вычислительная сложность и возможность анализа только шумовых компонент ЦИ. Поэтому представляет интерес использование методов спектрального анализа, позволяющих проводить комплексное исследование изменений параметров ЦИ, для определения факта встраивания стегоданных с применением ДКП и ДДВП.

3 МАТЕРИАЛЫ И МЕТОДЫ

В работе исследованы одноэтапные и двухэтапные стеганографические методы, использующие преобразования, как контейнеров, так и стегоданных. Этапы обработки контейнеров и стегоданных согласно методам Дея [9], Элайона [8] и Гунджаля [10] представлены в табл. 1.

На предварительном этапе обработки ЦИ, согласно методам Элайона и Гунджаля, производилось изменение системы цвета изображения-контейнера (табл. 1): с RGB на YCbCr/YIQ, что позволило минимизировать искажения цвета изображения-контейнера при встраивании стегоданных.

Формирование стеганограмм согласно методам Дея и Элайона производилось согласно формуле (1) с выборов в качестве базисных функций ДДВП вейвлета Хаара и соответствующей ему скейлинг-функции. Минимальное и максимальное значения параметра G были опре-

Таблица 1 – Этапы обработки контейнера и стегоданных при формировании стеганограмм согласно рассмотренным стеганографическим методам

Метод	Обработка контейнера			Обработка стегоданных		Весовой коэффициент G
	Изменение системы цвета	Первый этап	Второй этап	Первый этап	Второй этап	
Дея	–	ДДВП	–	ДДВП	–	0,02; 0,04; 0,06; 0,08
Элайона	RGB → YCbCr (скрытие в Y-канале)	ДДВП	–	Преобразование Арнольда	ДДВП	1; 4; 8; 12
Гунджаля	RGB → YIQ (скрытие в I-канале)	ДДВП	ДКП	Преобразование Арнольда	ДКП	$G_0 = 3$; $G_1 \in \{5,8,11,14\}$

делены из условий возможности извлечения стегоданных без искажений и появления видимых изменений изображения-контейнера при встраивании сообщений.

Для дополнительного уменьшения вероятности выявления стеганограмм, при их формировании, к стегоданным, представленным в виде полутонового изображения $D_{x,y}$ с размерами $W_x \times W_y \left(W_x = W_y\right)$ (пикселей), применялись N_A итераций ПА [11]. Особенностью ПА является периодичность – после конечного количества итераций исходное и полученное изображения совпадают. Для предотвращения такой ситуации в работе количество итераций ПА было выбрано равным $N_A = 23$, которое не является делителем размеров встраиваемых стегоданных.

В случае двухэтапного метода Гунджаля, к детализирующим коэффициентам ДДВП, полученным на первом этапе обработки изображения-контейнера, применялось ДКП. Встраивание сообщений, полученных после ПА, производилось аналогично формуле (1) – путем суммирования коэффициентов ДКП контейнера и стегоданных. Отметим, что величина весового параметра G изменялись в зависимости от значения скрываемого бита (табл. 1) – единицы ($G = G_1$) либо нуля ($G = G_0$).

АСС ЦИ является двумерным расширением известного метода исследований временных рядов Singular Spectrum Analysis [12]. Особенностью данного метода является использование сингулярного разложения для разделения информативных и шумовых компонент рассматриваемого сигнала.

АСС изображений проводился в четыре этапа [12]. На первом этапе, путем обработки заданного полутонового изображения $I_{x,y}$ ($N_x \times N_y$ пикселей) с использованием скользящего окна $R_{x,y}$ ($L_x \times L_y$ пикселей), была сформирована блочно-ганкелева матрица M_{BH}:

$$M_{BH} = \begin{pmatrix} B_1 & B_2 & B_3 & \cdots & B_{K_y} \\ B_2 & B_3 & B_4 & \cdots & B_{K_y+1} \\ B_3 & B_4 & \iddots & \iddots & \vdots \\ \vdots & \vdots & \iddots & \iddots & \vdots \\ B_{L_y} & B_{L_y+1} & \cdots & \cdots & B_{N_y} \end{pmatrix},$$

$$B_j = \begin{pmatrix} I_{1,j} & I_{2,j} & \cdots & I_{K_x,j} \\ I_{2,j} & I_{3,j} & \cdots & I_{K_x+1,j} \\ \vdots & \vdots & \ddots & \vdots \\ I_{L_x,j} & I_{L_x+1,j} & \cdots & I_{N_x,j} \end{pmatrix}. \qquad (2)$$

где $K_x = N_x - L_x + 1, K_y = N_y - L_y + 1$.

На втором этапе было выполнено сингулярное разложение матрицы M_{BH}:

$$\mathrm{M}_{BH} = \sum_{i=1}^{P} M_i = \sum_{i=1}^{P} \sqrt{\lambda_i} U_i V_i^T. \qquad (3)$$

На третьем этапе анализа производилась группировка троек M_i (3) в L групп:

$$\mathrm{M}_{BH} = \sum_{i=1}^{P} M_i = \sum_{j=1}^{L} \left(\sum_{q=1}^{L_j} M_q \right), \sum_j L_j = P. \qquad (4)$$

На последнем этапе АСС для восстановления исходного ЦИ в пространственной области проводилась нормировка групп собственных троек (4) по норме Фробениуса и проектирование их на множество блочно-ганкелевых матриц вида (2), используя операторы $\mathrm{H}_{\omega}^{\Re^{1\times1}}$ и $\mathrm{H}^{\Re^{L_x\times K_x}}$ [12], усредняющие элементы внутри блоков матрицы M:

$$\tilde{\mathrm{M}} = \mathrm{H}^{\Re^{L_x\times K_x}} \circ \mathrm{H}_{\omega}^{\Re^{1\times1}}\left(\mathrm{M}\right) = \mathrm{H}_{\omega}^{\Re^{1\times1}} \circ \mathrm{H}^{\Re^{L_x\times K_x}}\left(\mathrm{M}\right).$$

Обычно порядок группировки троек M_i для чистого изображения-контейнера неизвестен. Поэтому анализ интегральных изменений компонент АСС-разложения ЦИ (групп собственных троек) был проведен на основе исследования параметров КМ (табл. 2), элементы которой рассчитывались согласно формуле:

$$C_{ij} = \frac{E\left[\left(M_i - E\left[M_i\right]\right)\left(M_j - E\left[M_j\right]\right)\right]}{\sqrt{D\left[M_i\right]}\sqrt{D\left[M_j\right]}}, i \le P, j \le P.$$

Однородность значений элементов КМ (табл. 2) является мерой близости положения элементов с большими значениями к главной диагонали КМ, а энтропия позволяет оценить степень «хаотичности» значений элементов корреляционной матрицы C.

Для выявления и исследования фрактальных свойств временных рядов широко используется МФА [13]. Отличие МФА от известных методов оценки показателя Херста, например, R/S-анализа, состоит в представлении исследуемого сигнала в виде суперпозиции набора фракталов с единственным значением фрактальной размерности (монофракталов), что позволяет исследовать мультифрактальные свойства отдельных компонент сигнала.

Таблица 2 – Исследованные параметры корреляционной матрицы компонент АСС-разложения стеганограмм

Параметр корреляционной матрицы	Формула для расчета
Равномерность значений элементов	$R_1 = \sum_{i=1}^{P}\sum_{i=1}^{P} C_{ij}^2$
Однородность значений элементов КМ	$R_2 = \sum_{i=1}^{P}\sum_{i=1}^{P}\left[C_{ij}/\left(1+\lvert i-j\rvert\right)\right]$
Средние значения элементов КМ	$R_3 = \left(1/P^2\right)\sum_{i=1}^{P}\sum_{i=1}^{P} C_{ij}$
Параметр корреляционной матрицы	Формула для расчета
Энтропия значений элементов КМ	$R_4 = -\sum_{i=1}^{P}\sum_{i=1}^{P}\left(p_{ij}\times\log_2 p_{ij}\right)$, $p_{ij} = C_{ij}/\sum_{i=1}^{P}\sum_{i=1}^{P}\left(C_{ij}\right)$

Классические методы МФА позволяют проводить анализ только бинарных ЦИ и дают существенную погрешность при оценке мультифрактальных параметров, обусловленную наличием краевых эффектов [14]. Поэтому в работе был использован усовершенствованный метод оценки мультифрактальных характеристик ЦИ [15], который выдержал следующие этапы:

1. Фрагментирование исходного ЦИ $I_{x,y}$, представленного в градациях серого цвета, на перекрывающиеся блоки размером $s \times s$ пикселей. Область пересечения блоков по горизонтали (вертикали) составляла одну строку (столбец).

2. Вычисление относительной высоты n_s каждого блока разбиения ЦИ:

$$n_s = \begin{cases} \left\lceil (z_{\max} - z_{\min}) \cdot (1 + 2a\sigma) / (s-1) \right\rceil, & \text{если } z_{\max} \neq z_{\min}, \\ 1, & \text{если } z_{\max} = z_{\min}. \end{cases}$$

Согласно рекомендациям [15], значение параметра a было выбрано равным 3, что позволило минимизировать погрешность оценки мультифрактальных параметров ЦИ при анализе шумоподобных текстур.

3. Вычисление распределения вероятностей $p(n_s) = n_s / \sum_s n_s$ для высот n_s отдельных блоков разбиения $I_{x,y}$ и заданного размера блока s.

4. Вычисление обобщенной статистической суммы $Z(q,s)$ и спектра обобщенных фрактальных размерностей D_q:

$$Z(q,s) = \begin{cases} \sum_{i=1}^{N(s)} p_i^q(s), & \text{если } q \in \mathbb{R} \setminus \{0\}; \\ \sum_{i=1}^{N(s)} \left[-p_i(s) \times \log(p_i(s)) \right], & \text{если } q = 0, \end{cases}$$

$$D_q = \frac{\tau(q)}{q-1}, \tau(q) = \lim_{r \to 0} \frac{\ln Z(q,r)}{\ln(r)}.$$

5. Вычисление мультифрактального спектра:

$$\alpha(q) = \frac{d}{dq}\left[(q-1) \cdot D_q \right], f(\alpha) = q \cdot \alpha(q) - \tau(q).$$

Для повышения точности оценки значений обобщенных фрактальных размерностей этапы № 1–3 повторялись не менее трех раз при разных значениях размера блока разбиения s. При проведении исследований мультифрактальных параметров стеганограмм были использованы следующие характеристики СОФР и МФС [13]: информационная (D_1), корреляционная (D_2) и граничные обобщенные фрактальные размерности ($D_{-\infty}, D_{+\infty}$); ширина $(\Delta_\alpha = \alpha_{\max} - \alpha_{\min})$ и размах $(\Delta_f = f_{\max} - f_{\min})$ МФС.

4 ЭКСПЕРИМЕНТЫ

Исследование было проведено с использованием тестового пакета из ста полноцветных ЦИ с разрешением UHD-4K. В качестве стегоданных были выбраны три полноцветных изображения с различной степенью детализации: портрет, чертеж и карта. Характеристики тестовых изображений и скрываемых сообщений представлены в табл. 3.

Анализ стеганограмм проводился при изменении степени заполнения контейнера стегоданными (отношения измененных коэффициентов разложения ЦИ к общему количеству коэффициентов разложения изображения-контейнера) от 5% до 25%, с шагом 5%, и от 25% до 95%, с шагом 10%.

В работе приведены результаты исследования АСС стеганограмм, сформированных согласно методу Дея при размерах скользящего окна $R_{x,y}$ равных 50×50 (пикселей) и шаге выборки блоков ЦИ по строкам/столбцам $\Delta_b = 500$ пикселей. Отметим, что АСС характеризуется высокой вычислительной сложностью и для его реализации необходим большой объем памяти. Это связано с тем, что размеры матрицы $M_{BH}(2)$ нелинейно зависят от размеров анализируемого ЦИ и скользящего окна $R_{x,y}$:

$$M_{(L_x L_y) \times (K_x K_y)} = M_{(L_x L_y) \times \left[(N_x - L_x + 1)(N_y - L_y + 1) \right]}.$$

По результатам предварительных исследований было установлено, что заданная точность определения характеристик МФА ЦИ при минимальном времени расчета достигается при выборе диапазона значений параметра $q \in [-20; 20]$ и шаге выборки $\Delta_q = 0,5$. Зависимости D_q, $\alpha(q)$ и $f(\alpha)$ получены при использовании сплайн-интерполяции данных.

5 РЕЗУЛЬТАТЫ

Результаты АСС стеганограмм, сформированных согласно методу Дея, представлены на рис. 1–2 в виде графика «ящик с усами». Линией в «ящике» и кружком отображаются, соответственно, медиана и среднее значение параметров КМ; концы «усов» рассчитывались согласно следующим выражениям: $L_{up} = Q_3 + 1,5 \cdot (Q_3 - Q_1)$, $L_{down} = Q_1 - 1,5 \cdot (Q_3 - Q_1)$, где первый ($Q_1$) и третий ($Q_3$) квартили распределения значений параметров КМ соответствуют нижней /верхней границе «ящика».

Скрытие стегоданных согласно методу Дея приводит к существенным изменениям медиан значений равномерности (рис. 1а) и энтропии (рис. 2) элементов КМ, что дает возможность распознавать наличие скрытых сообщений в ЦИ. Одновременное уменьшение интерквартильного размаха $IQR = Q_3 - Q_1$ значений анализируемых параметров свидетельствует об устойчивости данного демаскирующего признака.

Таблица 3 – Характеристики тестовых изображений и стегоданных

Параметр	Тестовые изображения	Стегоданные		
		Портрет	Чертеж	Карта
Разрешение, пикселей	3840×2160	565×850	567×463	800×800
Формат представления	JPEG	BMP	BMP	BMP
Система цвета	RGB	RGB	RGB	RGB
Глубина цвета, $[\text{бит}/(\text{пиксель} \times \text{канал})]$	8	8	8	8

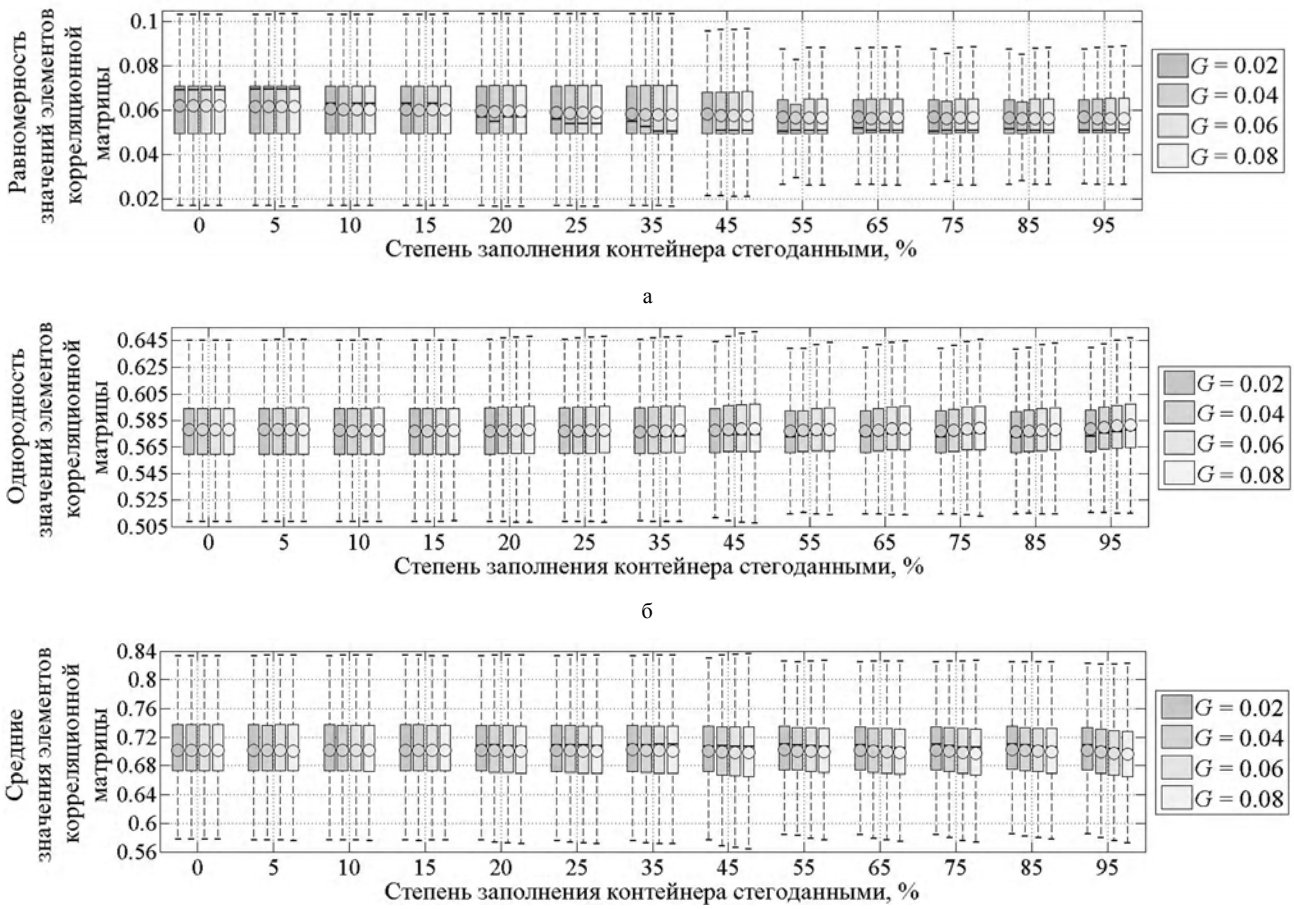


Рисунок 1 – Результаты АСС стеганограмм, сформированных согласно методу Дея: а – равномерность значений элементов КМ; б – однородность значений элементов КМ; в – средние значения элементов КМ

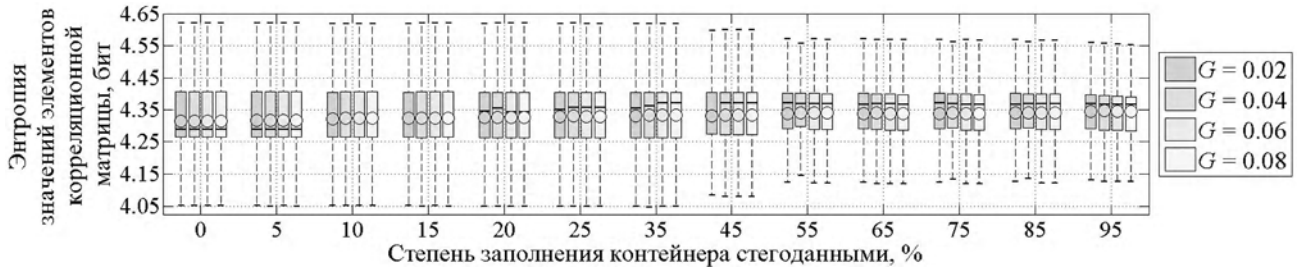
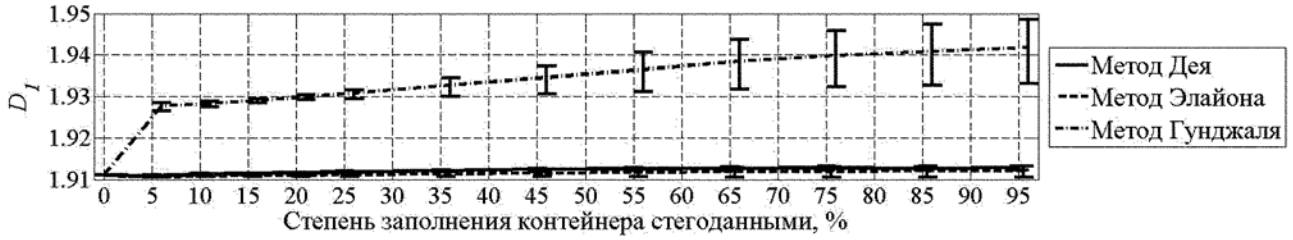


Рисунок 2 – Результаты АСС стеганограмм, сформированных согласно методу Дея – энтропия значений элементов КМ

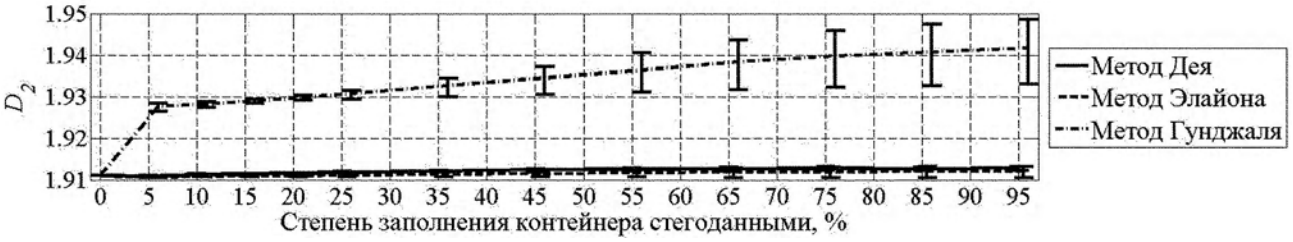
Для оценки изменений мультифрактальных параметров компонент ЦИ при встраивании стегоданных был проведен МФА стеганограмм. Результаты МФА представлены на рис. 3–4, где минимальные значения параметров получены при $G = G_{\min}$, а максимальные – при $G = G_{\max}$ (табл. 1).

Изменения информационной (рис. 3а), корреляционной (рис. 3б) и граничной обобщенной фрактальной размерности $D_{+\infty}$ (рис. 4б) для одноэтапных методов встраивания стегоданных близки к нулю. Это свидетельствует о незначительных изменениях энтропии яркости пикселей всего изображения и корреляции значений яркости соседних пикселей при формировании стеганограмм согласно методам Дея и Элайона.

Изменения параметров СОФР и МФС ЦИ (рис. 3–4) в случае применения двухэтапного метода Гунджаля существенно больше изменений соответствующих параметров для одноэтапных методов Дея и Элайона, что позволяет распознавать применение одно- и двухэтапных методов скрытия сообщений при проведении стегоанализа ЦИ.



а



б

Рисунок 3 – Результаты МФА стеганограмм: а – информационная размерность D_1; б – корреляционная размерность D_2; в – граничная обобщенная фрактальная размерность $D_{-\infty}$

6 ОБСУЖДЕНИЕ

Предложенный метод пассивного стегоанализа, по сравнению с методами, рассмотренными в работах [2–4], при заданной вероятности распознавания стеганограмм, требует существенно меньших затрат времени и ресурсов памяти, даже для случая многоэтапного формирования стеганосистемы.

Формирование стеганограмм согласно методу Дея практически не изменяет усредненных значений элементов КМ (рис. 1в), но приводит к возрастанию энтропии их значений (рис. 2), что соответствует росту дисперсии значений элементов КМ. С другой стороны, однородность значений элементов КМ остается практически неизменной (рис. 1б), что свидетельствует об одинаковом уровне искажений, внесенных в компоненты АСС-разложения ЦИ при скрытии стегоданных.

К недостаткам предложенного метода пассивного стегоанализа можно отнести то, что диапазон изменений характеристик АСС-разложения стеганограмм сравнительно мал (порядка 10^{-3}). Вследствие этого (существенно) усложняется алгоритм настройки стегодетектора, однако использование интегральных характеристик АСС-разложения стеганограмм позволяет достичь вероятности распознавания стеганограмм сопоставимой со случаем применения J+SRM-модели [2] при существенном сокращении размерности используемого пространства признаков.

Отметим, что при проведении МФА был получен a priori неочевидный результат. Двухэтапный метод Гуджаля встраивания стегоданных, который за счет уменьшения уровня демаскирующих признаков показал высокую робастность к стандартным методам стегоанализа, в случае применения МФА оказался менее устойчивым, чем одноэтапные методы. Существенное увеличение значений обобщенных фрактальных размерностей $D_{-\infty}$ ($D_{+\infty}$) для данного метода (рис. 4а–4б), может быть интерпретировано, как соответствующее увеличение фрактальных размерностей шумовых и детерминированных, например, контуров объектов, компонент ЦИ. Возрастание ширины Δ_α и уменьшение размаха Δ_f МФС (рис. 4в–4г) свидетельствует об увеличении мощности множества монофрактальных компонент ЦИ при сужении диапазона значений фрактальных размерностей данных компонент.

Также было выявлено, что предварительная обработка стегоданных в случае одноэтапных методов (метод Элайона) позволяет минимизировать искажения мультифрактальных параметров ЦИ (рис. 3–4) в широком диапазоне значений степени заполнения контейнера.

ВЫВОДЫ

В работе решена задача выявления характерных искажений спектральных характеристик стеганограмм, сформированных согласно АМВС, с использованием АСС и МФА для общего случая обработки, как изображения-контейнера, так и скрываемых сообщений.

Научная новизна результатов, полученных в статье, состоит в том, что впервые выдвинута и экспериментально подтверждена гипотеза о возможности использования спектральных методов анализа ЦИ для распознавания факта скрытия сообщений в области преобразования контейнера.

В результате проведенных исследований спектральных характеристик стеганограмм, сформированных согласно одно- и двухэтапным методам встраивания стегоданных, установлено следующее.

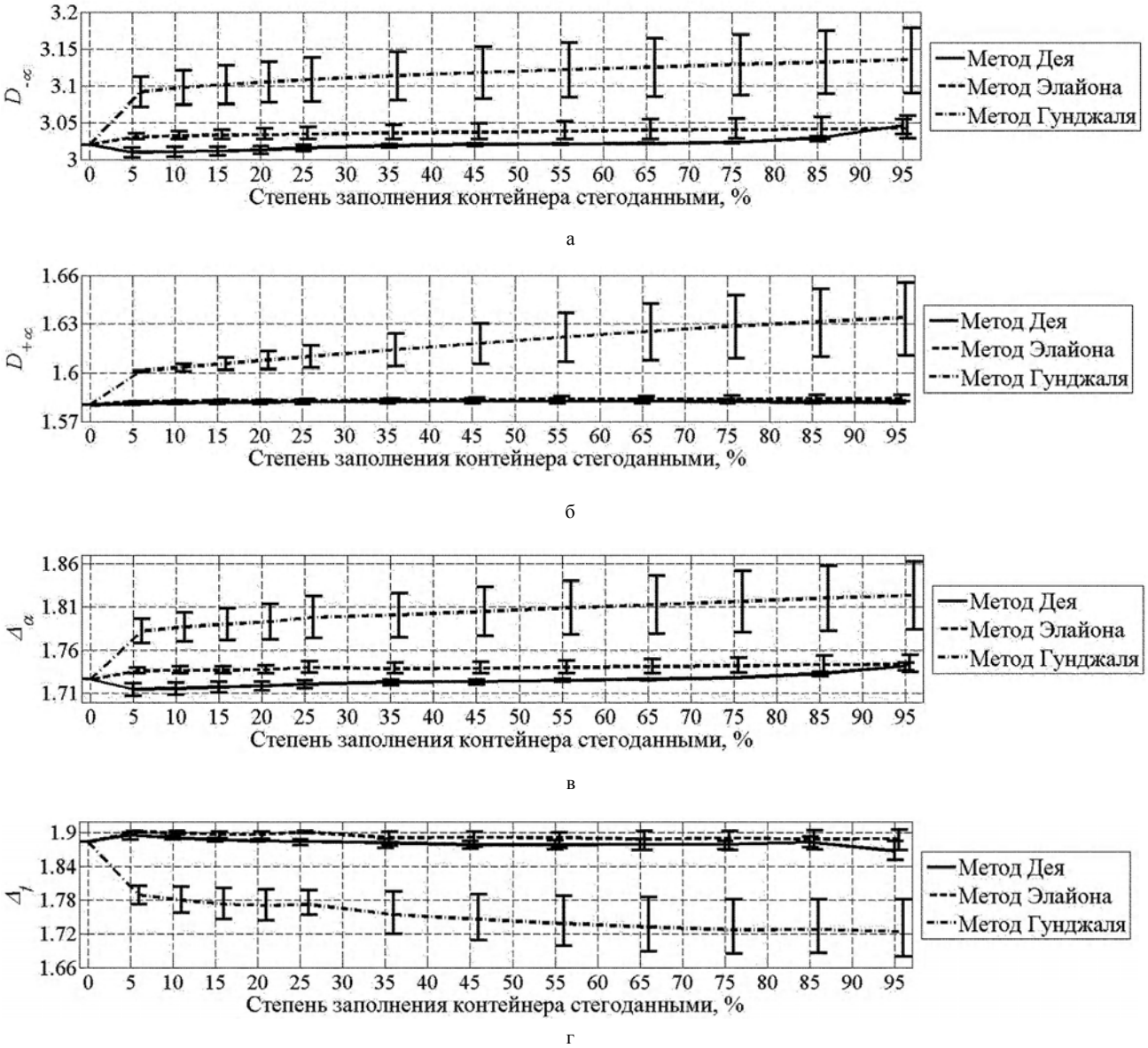


Рисунок 4 – Результаты МФА стеганограмм: а – граничная обобщенная фрактальная размерность $D_{-\infty}$; б – граничная обобщенная фрактальная размерность $D_{+\infty}$; в – ширина МФС Δ_α; г – размах МФС Δ_f

1. Встраивание стегоданных в контейнеры с применением одноэтапных методов приводит к значительным изменениям энтропии значений элементов КМ компонент АСС-разложения ЦИ. Данный демаскирующий признак является устойчивым в широком диапазоне значений степени заполнения контейнера стегоданными, что представляет особый интерес в связи с постоянным увеличением объемов контейнеров, например от Full HD к UHD-4K и UHD-8K.

2. Характерные изменения результатов МФА стеганограмм, сформированных двухэтапными методами – возрастание мощности множества монофрактальных компонентов ЦИ и уменьшение диапазона изменений фрактальных размерностей данных компонент, позволяет обнаруживать факт использования данных стеганографических методов с высокой точностью.

3. Не смотря на то, что при формировании стеганограмм на основе одно и двухэтапных методов используются преобразования одного вида, для повышения эффективности пассивного стегоанализа, для каждого метода необходима разработка специализированного стегодетектора.

Практическая значимость работы состоит в том, что:

– разработан программный комплекс, позволяющий проводить пассивный стегоанализ стеганограмм, сформированных на основе одно- и многоэтапных АМВС;

– применение выявленных характерных изменений спектральных характеристик ЦИ при скрытии сообщений согласно АМВС позволило существенно уменьшить время и вычислительную сложность процедуры настройки стегодетектора;

Полученные результаты могут быть использованы для повышения точности современных стегодетекторов и разработки целевых методов активного стегоанализа двухэтапных комплексных методов скрытия сообщений в ЦИ.

Дальнейший интерес представляет исследование возможности применения МФА в случае использования статистических AMBC (например, методов HUGO, S-UNIWARD).

БЛАГОДАРНОСТИ

Статья написана по материалам, полученным в рамках выполнения научно-исследовательской темы «Стеганография многомерных сигналов кафедры физико-технических средств защиты информации Физико-технического института Национального технического университета Украины «Киевский политехнический институт».

СПИСОК ЛИТЕРАТУРЫ

1. Kodovsky J. Calibration revisited / J. Kodovsky, J. Fridrich // 11th Multimedia and Security Workshop, New York, 2009 : proceedings. – New York : ACM Press, 2009. DOI 10.1145/1597817.1597830.
2. Fridrich J. Rich models for steganalysis of digital images / J. Fridrich, J. Kodovsky // Transactions on Information Forensics and Security. – 2012. – Vol. 7, Issue 3. – P. 868–882. DOI 10.1109/tifs.2012.2190402.
3. Liu Q. Steganalysis of DCT–embedding based adaptive steganography and YASS / Q. Liu // 13th Multimedia & Security Workshop, New York, 2011 : proceedings. – New York : ACM Press, 2011. – P. 77–86. DOI 10.1145/2037252.2037267.
4. Holub V. Random Projections of Residuals for Digital Image Steganalysis / V. Holub, J. Fridrich // Transactions on Information Forensics and Security. – 2013. – Vol. 8, Issue 12. – P. 1996–2006. DOI 10.1109/tifs.2013.2286682.
5. Прогонов Д. О. Виявлення стеганограм з даними, прихованими в області перетворення цифрових зображень / Д. О. Прогонов, С. М. Кущ // Вісник Національного технічного університету України «Київський політехнічний інститут». Серія «Радіотехніка. Радіоапаратобудування». – 2014. – № 57. – С. 128–142.
6. Прогонов Д. Варіограмний аналіз стеганограм / Д. Прогонов, С. Кущ // Захист інформації і безпека інформаційних систем: III Міжнародна науково-технічна конференція, Львів, 5–6 червня 2014 р.: матеріали конференції. – Львів : Українська академія друкарства, 2014. – С. 84–85.
7. Прогонов Д. А. Определение типа преобразований, использованных для скрытия данных в цифровых изображениях / Д. А. Прогонов, С. Н. Кущ // System analysis and information technologies : 16-я Международная научно-практическая конференция, Киев, 26–30 мая 2014 г. : материалы конференции. – К. : УНК «ИПСА» НТУУ «КПИ», 2014. – С. 435–436.
8. Elahian A. Improved robust DWT-watermaking in YCbCr color space / A. Elahian, M. Khalili, S. B. Shokouhi // Global journal of computer application and technology. – 2011. – Vol. 1, Issue 3. – P. 300–304.
9. Dey N. A novel approach of color image hiding using RGB color planes and DWT / N. Dey, A. B. Roy, S. Dey // International journal of computer application. – 2011. – Vol. 36, Issue 5. – P. 19–24.
10. Gunjal B. L. Secured color image watermarking technique in DWT-DCT domain / B. L. Gunjal, S. N. Mali // International Journal of Computer Science, Engineering and Information Technology (IJCSEIT). – 2011. – Vol. 1, Issue 3. – P. 36–44.
11. Nance J. Periods of discretized Arnold Cat Map and its extension to dimensions [Electronic resource] / J. Nance. – Cornell University Library, e-print archive. – Acceess mode: http://arxiv.org/abs/1111.2984.
12. Golyandina N. E. 2D-extention of singular spectrum analysis: algorithm and elements of theory / N. E. Golyandina, K. D. Usevich // World scientific. – P. 449–473. – DOI 10.1142/9789812836021_0029.
13. Божокин С.В. Фракталы и мультифракталы: монография / С. В. Божокин, Д. А. Паршин. – Ижевск : НИЦ. Регулярная и хаотическая динамика, 2001. – 128 с.
14. Lopes R. Fractal and multifractal analysis: A review / R. Lopes, N. Betrouni // Medical Image Analysis. – 2009. – Vol. 13, Issue 4. – P. 634–649. DOI 10.1016/j.media.2009.05.003.
15. Li J. An improved box-counting method for image fractal dimension estimation / J. Li, Q. Du, C. Sun // Pattern Recognition. – 2009. – Vol. 42, Issue 11. – P. 2460–2469. DOI 10.1016/j.patcog.2009.03.001.

Статья поступила в редакцию 09.12.2014.
После доработки 12.12.2014.

Прогонов Д. О.[1], Кущ С. М.[2]

[1]Магістр, аспірант кафедри фізико-технічних засобів захисту інформації, Фізико-технічний інститут, Національний технічний університет України «Київський політехнічний інститут», Київ, Україна

[2]Канд. техн. наук, доцент, доцент кафедри фізико-технічних засобів захисту інформації, Фізико-технічний інститут, Національний технічний університет України «Київський політехнічний інститут», Київ, Україн*а*

СПЕКТРАЛЬНИЙ АНАЛІЗ СТЕГАНОГРАМ

Для підвищення стійкості стеганоалгоритмів до відомих методів стегоаналізу широко використовуються методи приховання повідомлень в області перетворення цифрових зображень. Використання при формуванні стеганограм стандартних методів обробки зображень-контейнерів та стегоданих – двовимірних дискретних косинусного та вейвет перетворень – дозволяє мінімізувати спотворення вбудованих повідомлень при стисненні зображень.

В роботі для виявлення факту вбудовування стегоданих з використанням стандартних перетворень контейнеру та стегоданих запропоновано використовувати методи спектрального аналізу зображень: аналіз спектру сингулярностей та мультифрактальний аналіз. Перевагою даного підходу у порівнянні зі статистичними методами стегоаналізу є можливість дослідження одночасно декількох компонент зображення-контейнеру, що дозволяю локалізувати область приховання повідомлень.

Аналіз ефективності застосування методів спектрального аналізу був проведений для випадку використання одноетапних та багатоетапних методів формування стеганограм. Показано, що вбудовування повідомлень призводить до змін усіх компонент сингулярного розкладу зображення-контейнеру, а також появі характерних спотворень мультифрактальних характеристик стеганограм – підвищення потужності множини монофрактальних компонент зображень. Виявлені демаскуючі ознаки наявності прихованих повідомлень є стійкими в широкому діапазоні змін ступеня заповнення контейнеру стегоданими, що спрощує задачу виявлення повідомлень, прихованих в області перетворення зображень.

Ключові слова: цифрові зображення, пасивний стегоаналіз, аналіз спектру сингулярностей, мультифрактальний аналіз.

Progonov D. O.[1], Kushch S. M.[2]

[1]M.Sc., post-graduate student of Faculty of Information Security, The Institute of Physics and Technology, National technical University of Ukraine «Kyiv Polytechnic Institute», Kyiv, Ukraine

[2]Ph.D, Associate Professor, Associate Professor of Faculty of Information Security, The Institute of Physics and Technology, National technical University of Ukraine «Kyiv Polytechnic Institute», Kyiv, Ukraine

SPECTRAL ANALYSIS OF STEGANOGRAMS

The data embedding algorithms in transformation domain of digital images are widely used for increasing the robustness of stegoalgorithms to known steganalysis methods. Applying of standard transformations the digital images by steganograms forming – two-dimensional discrete cosine and wavelet transforms – allows minimizing the garbling the hidden messages by image compression.

We propose to use the spectral analysis methods – singular spectrum analysis and multifractal analysis – for revealing the fact of message embedding with usage of standard transform of cover image and stegodata. The advantage of such approach in comparison with statistical steganalysis methods is possibility to simultaneously investigation of several components the cover image, which allows localizing the domain of data hiding.

Research of effectiveness of the spectral analysis methods was provided for the case of usage the one-stage and multi-stage embedding methods. It is shown that stegodata embedding leads to alteration the all components of singular value decomposition of cover images as well as appearance the peculiar alteration of multifractal features the steganograms – increasing the cardinality of monofractal components the image. Revealed distinctive features of steganograms are stable in wide range of cover image payloads, which is simplifying of task of stegodata revealing, hidden in transformation domain of digital images.

Keywords: digital images, passive steganalysis, singular spectrum analysis, multifractal analysis.

REFERENCES

1. Kodovsky J., Fridrich J. Calibration revisited, *11th Multimedia and Security Workshop, New York, 2009: proceedings*. New York, ACM Press, 2009. DOI 10.1145/1597817.1597830.
2. Fridrich J., Kodovsky J. Rich models for steganalysis of digital images, Transactions on Information Forensics and Security, 2012, Vol. 7, Issue 3, pp. 868–882. DOI 10.1109/tifs.2012.2190402.
3. Liu Q., Steganalysis of DCT-embedding based adaptive steganography and YASS, *13th Multimedia & Security Workshop, New York, 2011, proceedings*. New York, ACM Press, 2011, p. 77–86. DOI 10.1145/2037252.2037267.
4. Holub V., Fridrich J. Random Projections of Residuals for Digital Image Steganalysis, *Transactions on Information Forensics and Security*, 2013, Vol. 8, Issue 12, pp. 1996–2006. DOI 10.1109/tifs.2013.2286682.
5. Progonov D. O., Kushh S. M. Vyjavlennja steghanoghram z danymy, prykhovanymy v oblasti peretvorennja cyfrovykh zobrazhenj, *Vîsnyk Nacionaljnogho tekhnichnogho universytetu Ukrajiny «Kyjivsjkyj politekhnichnyj instytut»*. Serija – Radiotekhnika. Radioaparatobuduvannja, 2014, Vol. 57, pp.128–142.
6. Progonov D., Kushh S. Variogramnyj analiz steganogram, *Zahyst informacii' i bezpeka informacijnyh system: III Mizhnarodna naukovo-tehnichna konferencija*, L'viv, 5–6 chervnja 2014 r.: materialy konferencii'. L'viv, Ukrai'ns'ka akademija drukarstva, 2014, pp. 84–85.
7. Progonov D. A., Kushh S. N. Opredelenie tipa preobrazovanij, ispol'zovannyh dlja skrytija dannyh v cifrovyh izobrazhenijah, *System analysis and information technologies: 16-taja Mezhdunarodnaja nauchno-prakticheskaja konferencija*, Kiev, 26–30 maja 2014 g.: materialy konferencii. Kiev, UNK «IPSA» NTUU «KPI», 2014, pp. 435–436.
8. Elahian A., Khalili M., Shokouhi S. B. Improved robust DWT-watermaking in YCbCr color space, *Global journal of computer application and technology,* 2011, Vol. 1, Issue 3, pp. 300–304.
9. Dey N., Roy A. B., Dey S. A novel approach of color image hiding using RGB color planes and DWT, International journal of computer application, 2011, Vol. 36, Issue 5, pp. 19–24.
10. Gunjal B. L., Mali S. N. Secured color image watermarking technique in DWT-DCT domain, *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT)*, 2011, Vol. 1, Issue 3, pp. 36–44.
11. Nance J. Periods of discretized Arnold Cat Map and its extension to dimensions [Jelektronnyj resurs]. Cornell University Library, e-print archive. Rezhim dostupa http://arxiv.org/abs/1111.2984. Zagl. s jekrana
12. Golyandina N. E., Usevich K. D. 2D-extention of singular spectrum analysis: algorithm and elements of theory, *WORLD SCIENTIFIC*, pp. 449–473. DOI 10.1142/9789812836021_0029
13. Bozhokin S. V., Parshin D. A. Fraktaly i mul'tifraktaly: monografija. Izhevsk, NIC, Reguljarnaja i haoticheskaja dinamika, 2001, 128 p.
14. Lopes R., Betrouni N. Fractal and multifractal analysis: A review, *Medical Image Analysis*, 2009, Vol. 13, Issue 4, pp. 634–649. DOI 10.1016/j.media.2009.05.003.
15. Li J., Du Q., Sun C. An improved box-counting method for image fractal dimension estimation, Pattern Recognition, 2009, Vol. 42, Issue 11, pp. 2460–2469. DOI 10.1016/j.patcog.2009.03.001.

УДК 004.932.2:519.652

Савков С. О.[1], Мороз В. В.[2]

[1]Аспірант кафедри обчислювальної математики Одеського національного університету ім. І. І. Мечникова, Одеса, Україна

[2]Канд. техн наук, професор кафедри обчислювальної математики Одеського національного університету ім. І. І. Мечникова, Одеса, Україна

АНАЛІЗ МЕТОДІВ СКЕЛЕТИЗАЦІЇ НЕРЕГУЛЯРНИХ ТЕКСТУР В ЗАДАЧІ ПОШУКУ УНІКАЛЬНИХ ДЕСКРИПТОРІВ ЗОБРАЖЕННЯ

Розглянута задача аналізу, ідентифікації та розпізнавання текстур зображень. Об'єктом дослідження є процес скелетизації напівтонових зображень. Предмет дослідження становлять ітераційні та неітераційні методи скелетизації текстур. Мета роботи полягає у виборі придатного методу скелетизації неоднорідних нерегулярних текстур для пошуку унікальних дескрипторів зображень у задачі перевірки автентичності об'єкта.

Розглянуто типи текстур та класифікацію алгоритмів скелетизації, переваги та недоліки алгоритмів різних типів. Проведено обчислювальний експеримент та зроблено порівняльний аналіз ітераційних алгоритмів Жанг-Сюенга, Го-Хола та неітераційного методу серединного осьового перетворення Блюма, які застосовувались до зразків зображень неоднорідних нерегулярних текстур паперу з мікроволокнами. Вирішено задачу порівняльного аналізу алгоритмів скелетизації для зображень неоднорідних нерегулярних текстур. Розроблено критерій якості, який використовує інформацію про кількість відгалужень скелету текстури, що, у свою чергу, залежить від кількості граничних точок скелету. Розроблено програмне забезпечення, що реалізує розглянуті методи скелетизації та дозволяє ранжувати їх згідно з обраним критерієм.

Результати експерименту дозволяють рекомендувати алгоритм Жанг-Сюенга до практичного застосування для вирішення задачі скелетизації зображень неоднорідних нерегулярних текстур.

Ключові слова: розпізнавання текстур, ідентифікація, скелетизація, витончення, серединне осьове перетворення, алгоритм Жанг-Сюенга, алгоритм Го-Хола.

НОМЕНКЛАТУРА

$A(P_1)$ – кількість патернів-переходів 01 у впорядкованій множині $P_2, P_3, P_4, \ldots, P_8, P_9$ восьми сусідів пікселя P_1;

B – бінарне зображення текстури;

$B(P_1)$ – кількість ненульових сусідів P_1;

$C(P)$ – кількість різних 8-зв'язних компонент які мають значення 1 в околі пікселя;

G – багатокутник;

Gs – напівтонове зображення текстури;

i – координати пікселя по вісі абсцис;

$I(i, j)$ – повнокольорове зображення текстури;

j – координати пікселя по вісі ординат;

$M(G)$ – серединна вісь;

N – кількість граничних точок скелету;

p – піксель бінарного зображення;

R – радіус-функція;

S – скелет зображення текстури;

T – матриця пікселів;

$\{q\}$ – множина точок внутрішніх для G;

β – бінарне перетворення зображення;

γ – напівтонове перетворення зображення;

σ – перетворення скелетизації бінарного зображення;

$\vee$ – логічне АБО.

ВСТУП

Ідентифікація текстур відноситься до важливих напрямків розвитку обробки зображень. Вона дозволяє вирішувати задачі у багатьох сферах життєдіяльності людини: сільському господарстві, картографії, безпеці, машинобудуванні.

Однією з ключових задач є перевірка автентичності об'єкта, у якій зацікавлені як виробники, що намагаються гарантувати якість, так і споживачі, які відстоюють своє право на якісне обслуговування.

Ця задача може бути розв'язана за допомогою використання додаткового маркування: водяних знаків, голограм, RFID-міток, паперу зі спеціальними волокнами. Ці методи мають різноманітні недоліки, такі як висока ціна, вразливість до пошкоджень, необхідність використання спеціальних приладів перевірки достовірності. Альтернативою може бути використання текстур матеріалів у якості унікального дескриптора об'єкта.

Метою роботи є огляд, порівняльний аналіз та пошук придатних методів скелетизації бінарних зображень неоднорідних нерегулярних текстур.

1 ПОСТАНОВКА ЗАДАЧІ

Нехай ми маємо $I(i, j)$. Отримаємо з нього $Gs(i, j)$, використовуючи перетворення $Gs(i, j) = \gamma(I(i, j))$. До $Gs(i, j)$ застосуємо перетворення β та отримаємо $B(i, j) = \beta(Gs(i, j))$.

Передній план бінарного зображення може бути поданий у вигляді тонких ліній $S(i, j)$ за допомогою перетворення $\sigma(B(i, j))$. Необхідно зробити порівняльний аналіз методів скелетизації, провести експеримент для зразків нерегулярних текстур та пошук придатних методів обробки текстур заданого класу. Як критерій якості будемо використовувати кількість N отриманого $S(i, j)$.

2 ОГЛЯД ЛІТЕРАТУРИ

Історія алгоритмів скелетизації починається з робіт Рутовіца [1], який у 1966 році представив паралельний алгоритм видалення пікселів. Цей алгоритм послужив основою для багатьох алгоритмів скелетизації. Блюм розглядав інший підхід до вирішення задачі скелетизації та у 1967 році запропонував неітераційний метод – серединне осьове перетворення [2], якому, на відміну від ітераційного методу Рутовіца, не потрібно було вивчати усі пікселі зображення. Неітераційні методи отримали розвиток у роботі Баруха, який запропонував використання віконного перетворення та ліній слідування.

DOI 10.15588/1607-3274-2015-2-10

Розенфельд (1975) повернувся до паралельних алгоритмів та вперше оцінив необхідні та достатні умови для збереження топології при видаленні граничних пікселів. Таким чином він вирішив проблему ерозії при витонченні діагональних ліній [3]. У подальшому у співавторстві з Даєром він представив алгоритм скелетизації напівтонових зображень, заснований на ідеї зв'язності пікселів [4]. Павлідіс у 1982 році запропонував комбінацію паралельного та послідовного алгоритмів, яка зменшувала обчислювальний час, необхідний для отримання скелету [5–6].

Різноманітні методи скелетизації отримали широке застосування у системах оптичного розпізнавання символів [7–8], анімації [9], задачах комп'ютерного зору [10], картографії [11]. Огляд та порівняння методів скелетизації у задачі розпізнавання символів представлені у роботах [12] та [13]. Методи скелетизації також широко використовуються для класифікації текстур [14].

3 МАТЕРІАЛИ ТА МЕТОДИ

У різних наукових напрямках термін «текстура» має свої специфічні означення. Історично він відноситься до ткацького мистецтва (лат. textura – тканина). У сучасних словниках текстура визначається як будова твердої речовини, гірської породи, природний малюнок деревини в розрізі. У якості означення приймемо «особливості будови речовини, зумовлені характером та розташуванням її складових частин».

Текстури можна класифікувати за наявністю відносини порядку між елементами. Якщо текстурні елементи розміщені у якомусь порядку, текстура буде називатися регулярною, в іншому випадку – нерегулярною. Також текстури можна розділити на однорідні, якщо вони містять текстурні елементи тільки одного типу, та неоднорідні, якщо вони складаються з елементів багатьох типів.

Термін «скелет» у загальному розумінні обробки зображень означає подання об'єкту множиною тонких ліній, яка описує його форму. Одним з найпростіших пояснень терміну скелет є наочний приклад «пожежа у степу». Зображення розглядається як степ, рівномірно вкритий сухою травою. Передбачається, що по межі області одночасно спалахує вогонь, який поширюється усередині неї з усіх напрямків та з постійною швидкістю. Точками скелета будуть ті очки області, в яких сходяться два (або більше) вогняні фронти. Вони будуть рівновіддаленими від найближчих точок загоряння на кордоні.

Якщо цей підхід поширити на растрові бінарні зображення, то підмножина чорних пікселів растру буде розглядатися як дискретний образ певної замкнутої області, граничні точки растрової множини – як образ кордону області, а безперервне поширення вогню буде моделюватися дискретним процесом послідовного «спалення» чорних пікселів растру. Також можуть використовуватися терміни «серединна вісь» для позначення геометричного місця точок фігури, рівновіддалених від її границь.

У деяких наукових працях «витончення» та «скелетизація» використовуються як синоніми, а термін «скелет» – для позначення результату, незалежно від форми вихідного малюнка або методу, що використовується [12].

Алгоритми скелетизації можна класифікувати як ітераційні та неітераційні. Неітераційні методи будують скелет зображення за один прохід, без вивчення всіх окремих пікселів. Одним з найперших неітераційних методів є метод серединного осьового перетворення Блюма, який розглядає скелетизацію як перетворення для опису форми. Результатом перетворення є множина точок, рівновіддалених від границь фігури [12].

Барух запропонував неітераційний непіксельний однопрохідний алгоритм, який будує скелет по лінії слідування. Лінія проходиться вікнами різного розміру, скелет будується як зв'язані центри послідовних вікон. Цей метод менш чутливий до шуму, ніж звичайні алгоритми проріджування [15].

Ітераційні методи витончення будують скелет, досліджуючи та видаляючи контурні пікселі за допомогою неітераційного процесу, послідовного або паралельного (табл. 1).

Паралельний алгоритм скелетизації Рутовіца розглядає пікселі контуру. Ті з них, які мають не менш двох сусідів кольору об'єкту та мають зв'язність чотири, позначаються для видалення. Недоліками запропонованого методу були надмірна ерозія та розташування скелету не по центру, в зв'язку з асиметричною природою його умов.

Паралельні алгоритми також можуть бути додатково класифіковані відповідно до їх продуктивності на алгоритми з чотирма, двома або одним підциклом [16]. Паралельні алгоритми з одним підциклом отримали більш значну увагу, оскільки вони скорочують час обчислення на кількість ітерацій, тому їх іноді називають одинопрохідні або повністю паралельні алгоритми [17].

У послідовних алгоритмах точки контуру для видалення розглядаються у заздалегідь визначеному порядку. Це може бути досягнуто за допомогою растрового сканування або обходу зображення по контурних пікселях. Наприклад, послідовний алгоритм Хілдча (1968) сканує зображення зліва направо та зверху вниз, позначаючи пікселі для видалення [18].

У паралельних алгоритмах пікселі для видалення розглядаються на підставі результатів, отриманих тільки з попередньої ітерації. Тому паралельні алгоритми проріджування підходять для реалізації на обладнанні, що використовує паралельні процесори.

Серединне осьове перетворення, запропоноване Блюмом, для багатокутника G визначається наступним чином: $M(G)$ будується з множини точок $\{q\}$ внутрішніх для G. При цьому є принаймні дві точки на кордоні об'єкта, що рівновіддалені від $\{q\}$ і є найближчими до $\{q\}$ (рис. 1). З серединною віссю пов'язана радіус-функція R, яка визначає відстань до кордону багатокутника для кожної точки осі.

Одним з недоліків серединного осьового перетворення є висока чутливість до шуму.

Метод скелетизації Жанг-Сюенга [19] розглядає зображення як матрицю T, де усі пікселі $T(i, j)$ мають значення 1 або 0. Патерн текстури складається з тих точок, які мають значення 1.

Таблиця 1 – Класифікація алгоритмів скелетизації

Методи скелетизації				
Ітераційні		Неітераційні		
Послідовні	Паралельні	Серединне осьове перетворення	Метод ліній слідування	Інші

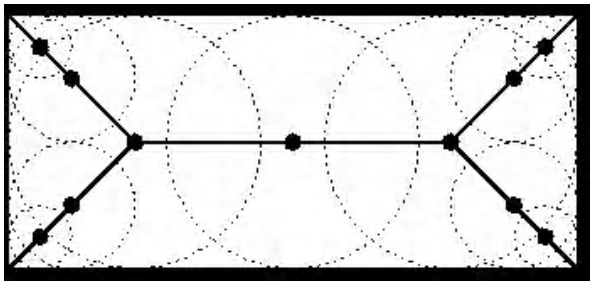


Рисунок 1 – Скелетизація методом серединного осьового перетворення
(http://www.cs.ru.nl/~ths/rt2/col/h9/medialaxisprin.gif)

Ітераційні перетворення застосовуються до T, точка за точкою, відповідно до значень невеликого набору сусідніх пікселів:

$$\begin{matrix} p_9(i-1,j-1), & p_2(i-1,j), & p_3(i-1,j+1); \\ p_8(i,j-1), & p_1(i,j), & p_4(i,j+1); \\ p_7(i+1,j-1), & p_6(i+1,j), & p_5(i+1,j+1). \end{matrix}$$

Метод реалізує видалення всіх контурних точок зображення, крім тих точок, що належать до скелету. Для того, щоб зберегти зв'язність скелета, кожна ітерація ділиться на дві під-ітерації.

У першій під-ітерації точка p_1 видаляється, якщо вона задовольняє таким умовам:

$$2 \le B(P_1) \le 6,\ 2 \le B(P_1) \le 6, \quad (1)$$

$$A(P_1) = 1, \quad (2)$$

$$P_2 * P_4 * P_6 = 0, \quad (3)$$

$$P_4 * P_6 * P_8 = 0, \quad (4)$$

$$I = (x,y),$$

де $B(P_1) = P_2 + P_3 + P_4 + \ldots + P_8 + P_9$.

Якщо будь-яка умова не виконується, тоді піксель P_1 не видаляється.

На другій під-ітерації умови (3) та (4) змінюються на наступні умови:

$$P_2 * P_4 * P_8 = 0, \quad (3')$$

$$P_2 * P_6 * P_8 = 0. \quad (4')$$

Алгоритм Го та Хола (1989) відноситься до методів паралельного витончення [20]. Він використовує окіл 3x3 для визначення сусідів оброблюємого пікселя *P*. Вертикальні та горизонтальні сусіди позначаються парними індексами, діагональні – непарними:

$$\begin{matrix} p_1(i-1,j-1), & p_2(i-1,j+1), & p_3(i-1,j+1); \\ p_8(i,j-1), & P, & p_4(i,j+1); \\ p_7(i+1,j-1), & p_6(i+1,j), & p_5(i+1,j+1). \end{matrix}$$

Тоді $N(p) = MIN\left[N_1(p), N_2(p)\right]$,

де $N_1(p) = (p_1 \vee p_2) + (p_3 \vee p_4) + (p_5 \vee p_6) + (p_7 \vee p_8)$,

$N_2(p) = (p_2 \vee p_3) + (p_4 \vee p_5) + (p_6 \vee p_7) + (p_8 \vee p_1)$.

$C(p) = 1;$

$2 \le N(p) \le 3,$

$$\left[\begin{array}{ll} (p_2 \vee p_3 \vee p_5) \vee p_4 = 0, & \text{для непарних ітерацій,} \\ (p_6 \vee p_7 \vee p_1) \vee p_8 = 0, & \text{для парних ітерацій.} \end{array}\right.$$

Процес витончення зупиняється, якщо більше неможливо видалити жодного пікселя.

4 ЕКСПЕРИМЕНТИ

У даній роботі методи скелетизації серединне осьове перетворення, Жанг-Сюенга та Го-Хола були протестовані на двох зразках зображень текстури паперу з волокнами, розміром 255x255 пікселів (рис. 2). Текстура зразків має нерівномірний фон та мікроволокна розміщені у випадковому порядку. Таким чином, зразки можуть бути класифіковані як нерегулярні неоднорідні текстури. Обробка текстур ускладнюється присутнім на зображеннях цифровим шумом. У якості методу бінарізації використовується метод Саувола з наступними параметрами: розмір вікна 25x25 пікселів, доповнення – симетричне, порогове значення – 0,15. Експеримент проводився у програмному середовищі Matlab 7.12.0 та OpenCV 2.4.9.

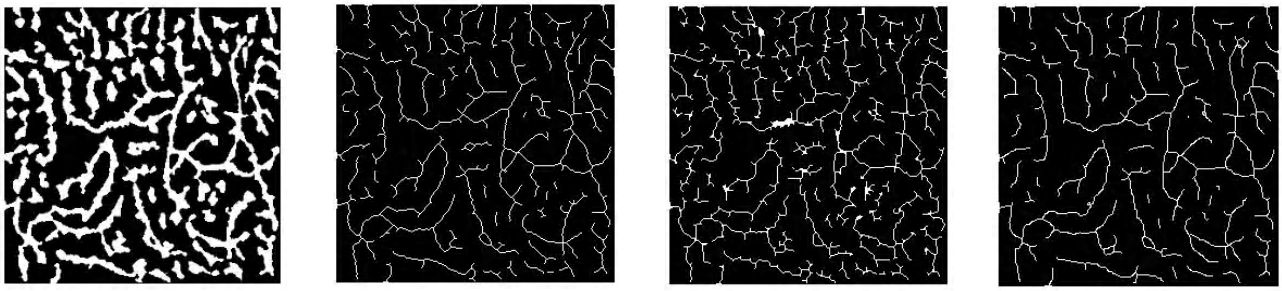
5 РЕЗУЛЬТАТИ

Результати застосування перетворення скелетизації до тестових зразків текстур приведені на рис. 3 та рис. 4.

а

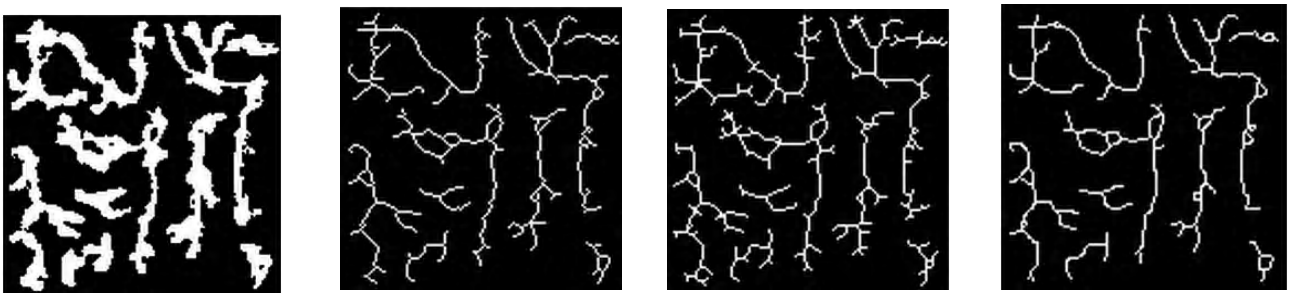
б

Рисунок 2 – Фрагменти текстур паперу з мікроволокнами:
а – зразок 1; б – зразок 2



а б в г

Рисунок 3 – Результати експерименту для зразку 1:
а – бінарізоване зображення; б – метод Го-Хола; в – метод середньої вісі; г – метод Жанг-Сюенга



а б в г

Рисунок 4 – Результати експерименту для зразку 2:
а – бінарізоване зображення; б – метод Го-Хола; в – метод середньої вісі; г – метод Жанг-Сюенга

6 ОБГОВОРЕННЯ

З отриманих результатів видно, що скелет, отриманий за допомогою методу середньої вісі, має розриви та багато відгалужень; алгоритм Го-Хола більш чутливий до форми патерну текстури та дає більш гіллястий скелет; метод Жанг-Сюенга дає найменшу кількість відгалужень та дає більш гладкий скелет.

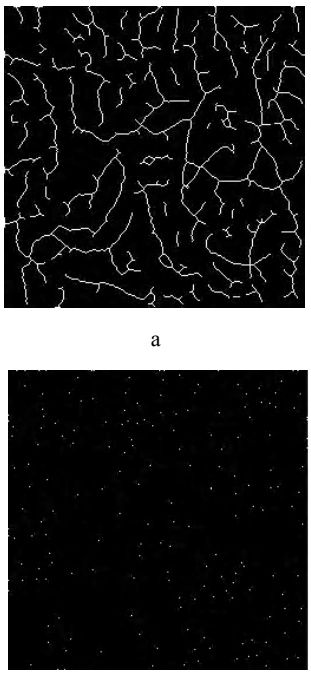
Через те, що надлишкова кількість відгалужень може заважати розпізнаванню текстури, як критерій якості будемо використовувати число N, які містяться у $S(i,j)$ (рис. 5а,б). Точка p буде граничною, якщо число $B(p)$ ненульових сусідніх пікселів дорівнює 1. Алгоритм з мінімальним N буде вважатися найкращим. Кількість граничних точок скелетів тестових зображень наведена у таблиці 2.

Найменшу кількість граничних точок дає алгоритм Жанг-Сюенга, що співпадає з візуальною оцінкою результатів скелетизації.

Результати порівняльного аналізу показали, що менша кількість граничних точок (більш гладкий скелет та відсутність шумових відгалужень) дають можливість зробити висновок про більшу придатність алгоритму Жанг-Сюенга для обробки нерегулярних неоднорідних текстур.

Таблиця 2 – Кількість граничних точок тестових зразків

Алгоритм скелетизації	Кількість граничних точок	
	Зразок 1	Зразок 2
Метод Го-Хола	244	65
Метод середньої вісі	690	113
Метод Жанг-Сюенга	183	44



а

б

Рисунок 5 – Обчислення граничних точок скелету:
а – скелетизація зображення за методом Го-Хола; б – граничні точки скелету

ВИСНОВКИ

У роботі вирішено актуальну задачу скелетизації бінаризованих зображень нерегулярних неоднорідних текстур, яка є одним з етапів задачі пошуку унікальних дескрипторів зображення.

Наукова новизна роботи полягає у тому, що вперше було зроблено порівняльний аналіз ітераційних та неітераційних алгоритмів скелетизації зображень нерегулярних неоднорідних текстур.

Практична цінність отриманих результатів полягає в тому, що обраний метод скелетизації використовується у якості одного з етапів пошуку унікальних дескрипторів зображення текстури у задачі ідентифікації об'єктів.

Перспективи подальших досліджень полягають оптимізації методу скелетизації для багатопроцесорних та мобільних систем та у використанні алгоритму скелетизації Жанг Сюенга для пошуку унікальних дескрипторів зображення текстури.

ПОДЯКИ

Роботу виконано в рамках науково-дослідної роботі №229 «Дослідження деяких класів крайових задач диференційних рівнянь, їх апроксимації та методів обробки зображень» кафедри обчислювальної математики Одеського національного університету ім. І. І. Мечникова.

СПИСОК ЛІТЕРАТУРИ

1. Rutovitz D. Pattern recognition / D. Rutovitz // Journal of Royal Statistical Society. – 1966. – Vol. 129, no. 4. – P. 504–530.
2. Blum H. A transformation for extracting new descriptors of shape / H. Blum // Models for the Perception of Speech and Visual Form. – Cambridge : MITPress, 1967. – P. 362–380.
3. Rosenfeld A. A characterization of parallel thinning algorithms / A. Rosenfeld // Information and Control. – 1975. – Volume 29. – Issue 3. – P. 286–291.
4. Dyer C. A. Thinning algorithms for gray-scale pictures / C. Dyer, A. Rosenfeld // IEEE Transactions on Pattern Analysis and Machine Intelligence. –1979. – № 1. – P. 88–89.
5. Pavlidis T. Algorithms for Graphics and Image Processing / T. Pavlidis. – Berlin-NY. : Computer Science Press, Springer–Verlag, 1982. – 447 p.
6. Pavlidis T. An asynchronous thinning algorithm / T. Pavlidis // Computer Graphics Image Processing. – 1982. – № 20. – P. 133–157.
7. Ning L. An Implementation of OCR System Based on Skeleton Matching / L. Ning. – Canterbury : University of Kent, 1993. – 66 p.
8. Mohamed A. An efficient thinning algorithm for arabic OCR systems / A. Mohamed // Signal & Image Processing : An International Journal. – June 2012. – Vol. 3, Issue 3. – P. 31.
9. Porawat V. The Skeleton Pruning-Smoothing Algorithm for Realistic Character Animation / V. Porawat, K. Prachumrak // JMMT : Journal of Man, Machine and Technology. – 2013. – Vol. 2, no. 1. – P. 21–34,
10. Nicu D. Curve-Skeleton Properties, Applications and Algorithms / D. Nicu, C. Silver, D. Silver // IEEE Transactions on Visualization and Computer Graphics. – 2007. – Vol. 13, Issue 3. – P. 530–548.
11. Gold C. Map generalization by skeleton retraction / C. Gold and D. Thibault // Proceedings of the 20th International Cartographic Conference BeiJing, China, August 6–10, 2001 (ICC 2001). – 2003. – P. 2072–2081.
12. Lam L. Thinning Methodologies – A Comprehensive Survey / L. Lam, S.-W. Lee, C. Y. Suen // IEEE Trans. Pattern Anal. Mach. Intell. – 14, 9 (September 1992). – P. 869–885.
13. Couprie M. Note on fifteen 2D parallel thinning algorithms : Internal Report / M. Couprie. – Université de Marne-la Vallée, 2006. – 21 p.
14. Kumar V. Skeleton Primitive Extraction Method on Textures with Different Nonlinear Wavelets, Texture Classification Based on Extraction of Skeleton Primitives Using Wavelets / V. Kumar, U.S.N. Raju, P. Premchand, A. Suresh // Journal of Computer Science. – 2008. – Vol. 4, Issue 7. – P. 591.
15. Baruch O. Line thinning by line following / O. Baruch // Pattern Recognition Letters. – 1988. – Vol. 8, no. 4. – P. 271–276.
16. Daya B. Parallelization of Two-Dimensional Skeletonization Algorithms / B. Daya // Journal of Undergraduate Research. – University of Florida. – 2008. – Vol. 9, Issue 4. – P. 1–23.
17. Saeed K. K3M a Universal algorithm for image skeletonization and a review of thinning techniques / K. Saeed, M. Tabкdzki, M. Rybnik, M. Adamski // International Journal of Applied Mathematics and Computer Science. – July 2010. – Vol. 20, Issue 2. – P. 317–335.
18. Hilditch J. An application of graph theory in pattern recognition / J. Hilditch // Machine Intelligence. – Edinburgh. – 1968. – Vol. 3. – P. 325–347.
19. Zhang T. Y. A fast parallel algorithm for thinning digital patterns / T. Y. Zhang, C. Y. Suen // Communications of the ACM. – March 1984. – Vol. 27, Issue 3. – P. 236–239.
20. Guo Z. Parallel thinning with two-subiteration algorithms / Z. Guo, R. W. Hall // Communications of the ACM. – March 1989. – Vol. 32, Issue 3. – P. 359–373.

Стаття надійшла до редакції 06.10.2014.
Після доробки 10.11.2014.

Савков С. А.[1], Мороз В. В.[2]
[1]Аспирант кафедры вычислительной математики Одесского национального университета им. И. И. Мечникова
[2]Канд. техн наук, профессор кафедры вычислительной математики Одесского национального университета им. И. И. Мечникова, Одесса, Украина

АНАЛИЗ МЕТОДОВ СКЕЛЕТИЗАЦИИ НЕРЕГУЛЯРНЫХ ТЕКСТУР В ЗАДАЧЕ ПОИСКА УНИКАЛЬНЫХ ДЕСКРИПТОРОВ ИЗОБРАЖЕНИЯ

Рассмотрена задача анализа, идентификации и распознавания текстур изображений. Объектом исследования является процесс скелетизации полутоновых изображений. Предмет исследования составляют итерационные и неитерационные методы скелетизации текстур.

Цель работы заключается в выборе подходящего метода скелетизации неоднородных нерегулярных текстур для поиска уникальных дескрипторов изображений в задаче проверки подлинности объекта. Рассмотрены типы текстур и классификация алгоритмов скелетизации, преимущества и недостатки алгоритмов различных типов. Проведен вычислительный эксперимент и сделан сравнительный анализ итерационных алгоритмов Жанг-Сюенга, Го-Холла и неитерационного метода срединного осевого преобразования Блюма, применявшихся к образцам изображений неоднородных нерегулярных текстур бумаги с микроволокнами. Решена задача сравнительного анализа алгоритмов скелетизации для изображений неоднородных нерегулярных текстур. Разработан критерий качества, который использует информацию о количестве ответвлений скелета текстуры, что, в свою очередь, зависит от количества граничных точек скелета. Разработано программное обеспечение, которое реализует рассмотренные методы скелетизации и позволяет ранжировать их согласно выбранному критерию.

Результаты эксперимента позволяют рекомендовать алгоритм Жанг-Сюенга к практическому применению для решения задачи скелетизации изображений неоднородных нерегулярных текстур.

Ключевые слова: распознавание текстур, идентификация, скелетизация, утончение, срединное осевое преобразование, алгоритм Жанг-Сюенга, алгоритм Го-Хола.

Savkov S. O.[1], Moroz V. V.[2]
[1]Post-graduate student of Computational Mathematics Department of I. I. Mechnikov Odessa National University
[2]PhD, Professor of Computational Mathematics Department of I. I. Mechnikov Odessa National University

ANALYSIS OF SKELETONIZATION METHODS FOR IRREGULAR TEXTURE IN UNIQUE IMAGE DESCRIPTORS SEARCHING PROBLEM

The problem of texture analysis, identification and recognition was investigated. The object of study is the process of grayscale image skeletonization. The subject of research is the iterative and non-iterative texture skeletonization methods for searching unique image descriptors in the problem of authentication object.

The purpose of the work is to select the appropriate method for inhomogeneous irregular texture skeletonization. A review of texture types and skeletonization algorithms classification, advantages and disadvantages of various types of algorithms was given. Computational experiments and comparison of iterative Zhang-Suen and Guo-Hall algorithms and non-iterative method of the Blum's median axis transformation that's applied to the sample images inhomogeneous irregular paper textures with microfibers were made. The problem of the comparative analysis of texture skeletonization algorithms for images of inhomogeneous irregular textures was solved. Was developed quality criterion which uses information about the number of skeleton texture branches, which depends on the number of boundary points of the skeleton. Software that implements the discussed skeletonization algorithms and allows to rank them according to selected criteria was developed.

The experimental results allow us to recommend the algorithm of Zhang-Suen for practical use to solve the problem of image skeletonization of inhomogeneous irregular textures.

Keywords: texture recognition, identification, skeletonization, medial axis transformation, Zhang-Suen method, Guo-Hall method.

REFERENCES

1. Rutovitz D. Pattern recognition, *Journal of Royal Statistical Society,* 1966, Vol. 129, no. 4, pp. 504–530.
2. Blum H. A transformation for extracting new descriptors of shape, *Models for the Perception of Speech and Visual Form.* Cambridge, MITPress, 1967, pp. 362–380.
3. Rosenfeld A. A characterization of parallel thinning algorithms, *Information and Control,* 1975, Volume 29, Issue 3, pp. 286–291.
4. Dyer C., Rosenfeld A. Thinning algorithms for gray-scale pictures, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* No. 1 (1–1979), pp. 88–89.
5. Pavlidis T. Algorithms for Graphics and Image Processing. Berlin-NY, Computer Science Press, Springer-Verlag, 1982, 447 p.
6. Pavlidis T. An asynchronous thinning algorithm, *Computer Graphics Image Processing,* 1982, No. 20, pp. 133–157.
7. Ning L. An Implementation of OCR System Based on Skeleton Matching. Canterbury, University of Kent, 1993, 66 p.
8. Mohamed A. An efficient thinning algorithm for arabic OCR systems, Signal & Image Processing : An International Journal, June 2012, Vol. 3, Issue 3, P. 31.
9. Porawat V., Prachumrak K. The Skeleton Pruning-Smoothing Algorithm for Realistic Character Animation, *JMMT : Journal of Man, Machine and Technology*, 2013, Vol. 2, no. 1, pp. 21–34.
10. Nicu D., Silver C., Silver D. Curve-Skeleton Properties, Applications and Algorithms, IEEE *Transactions on Visualization and Computer Graphics*, 2007, Vol. 13, Issue 3, pp. 530–548.
11. Gold C., Thibault D. Map generalization by skeleton retraction, *Proceedings of the 20th International Cartographic Conference BeiJing*. China, August 6–10, 2001 (ICC 2001), 2003, pp. 2072–2081.
12. Lam L., Lee S.-W., Suen C. Y. Thinning Methodologies – A Comprehensive Survey, *IEEE Trans. Pattern Anal. Mach. Intell,* 14, 9 (September 1992), pp. 869–885.
13. Couprie M. Note on fifteen 2D parallel thinning algorithms : Internal Report. Université de Marne-la Vallée, 2006, 21 p.
14. Kumar V., Raju U.S.N., Premchand P., Suresh A. Skeleton Primitive Extraction Method on Textures with Different Nonlinear Wavelets, Texture Classification Based on Extraction of Skeleton Primitives Using Wavelets, *Journal of Computer Science,* 2008, Vol. 4, Issue 7, pp. 591.
15. Baruch O. *Line thinning by line following*, *Pattern Recognition Letters, 1988,* Vol. 8, no. 4, pp. 271–276.
16. Daya B. Parallelization of Two-Dimensional Skeletonization Algorithms, *Journal of Undergraduate Research.* University of Florida, 2008, Vol. 9, Issue 4, pp. 1–23.
17. Saeed K., Tabкdzki M., Rybnik M., Adamski M. K3M a Universal algorithm for image skeletonization and a review of thinning techniques, International Journal of Applied Mathematics and Computer Science, July 2010, Vol. 20, Issue 2, pp. 317–335.
18. Hilditch J. An application of graph theory in pattern recognition, *Machine Intelligence*. Edinburgh. 1968, Vol. 3, pp. 325–347.
19. Zhang T. Y., Suen C. Y. A fast parallel algorithm for thinning digital patterns, *Communications* of the *ACM,* March 1984, Vol. 27, Issue 3, pp. 236–239.
20. Guo Z., Hall R. W. Parallel thinning with two-subiteration algorithms, *Communications of the ACM*, March 1989, Vol. 32, Issue 3, pp. 359–373.

УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

CONTROL IN TECHNICAL SYSTEMS

УДК 519.248 : 621.311.238

Шитикова Е. В.[1], Табунщик Г. В.[2]

[1]*Аспирант кафедры программных средств Запорожского национального технического университета, Запорожье, Украина*

[2]*Канд. техн. наук, доцент, доцент кафедры программных средств Запорожского национального технического университета, Запорожье, Украина*

МЕТОД УПРАВЛЕНИЯ НЕОПРЕДЕЛЕННОСТЬЮ В УСЛОВИЯХ ОГРАНИЧЕННЫХ РЕСУРСОВ

Рассмотрены вопросы надежности и функциональной безопасности газотурбинных установок наземного применения. Проведена формализация ранее разработанных моделей рисков газотурбинных установок наземного применения и их процессов. Предложен метод управления неопределенностью сложных технических систем и их процессов в условиях ограниченных ресурсов, основанный на методе оценки рисков и методе выбора контрмер, что позволяет сформировать область критичных рисков для исследуемой сложной технической системы. Усовершенствован метод выбора контрмер из альтернативных вариантов, основанный на методе анализа иерархий и методе приемлемого риска ALARP, поскольку применение контрмер осуществляется за счет реально имеющихся (ограниченных) ресурсов. Проведена классификация контрмер по времени применения и по объекту воздействия. Выделенные типы послужили критерием сравнения при выборе контрмер из альтернативных вариантов. Создана сводная таблица соотношения контрмер различных типов. Метод выбора контрмер из альтернативных вариантов позволяет сделать вывод в условиях ограниченных ресурсов и дать экономически обоснованную оценку при принятии решения об уровне приемлемого риска. Полученные методы апробированы для оценки рисков процесса испытаний газотурбинных установок наземного применения.

Ключевые слова: газотурбинные установки наземного применения, модель рисков, метод управления неопределенностью, метод выбора контрмер.

НОМЕНКЛАТУРА

ALARP – «as low as reasonably practicable» – принцип приемлемого риска;

ГТУ НП – газотурбинная установка наземного применения;

СТС – сложная техническая система;

A – множество атрибутов гиперкуба;

A_{di} – множество атрибутов измерения d_i;

$A_{di\ \mathrm{macro}}$ – множество атрибутов измерения d_i на макроуровне;

$A_{di\ \mathrm{micro}}$ – множество атрибутов измерения d_i на микроуровне;

$A'_{di\ \mathrm{macro}}$ – множество атрибутов измерения d_i для среза $M'_{R\ \mathrm{macro}}$;

$A'_{di\ \mathrm{micro}}$ – множество атрибутов измерения d_i для среза $M'_{R\ \mathrm{micro}}$;

$A'_{d3\ \mathrm{alt}}$ – множество, содержащее только группы альтернативных контрмер из множеств $A'_{d3\ macro}$ и $A'_{d3\ \mathrm{micro}}$;

$A'_{d3\ mod}$ – модифицированное множество $A'_{d3\ alt}$, содержащее перечень контрмер, рекомендованных к применению;

a_g^i – g-й элемент множества атрибутов измерения d_i;

Asset – множество активов, представляющих ценность для СТС;

D – множество измерений гиперкуба;

d_i – i-е измерение гиперкуба;

$H(D, A)$ – гиперкуб данных;

$H_{\mathrm{macro}}(D_{\mathrm{macro}}, A_{\mathrm{macro}})$ – гиперкуб данных макроуровня;

$H_{\mathrm{micro}}(D_{\mathrm{micro}}, A_{\mathrm{micro}})$ – гиперкуб данных микроуровня;

h – булева переменная для атрибута a_g^3, равная 0, если контрмера не имеет альтернативных вариантов, и равная 1, если имеет;

k – относительные веса каждой из контрмер множества $A'_{d3\ alt}$;

k_g – относительные веса g-й контрмеры для атрибута a_g^3;

k' – степень компенсации риска от применения контрмеры;

k'_g – степень компенсации риска от применения g-й контрмеры для атрибута a_g^3;

DOI 10.15588/1607-3274-2015-2-11

$L(a_g^2)$ – возможный ущерб от наступления g-го риска для атрибута a_g^2;

$L(a_g^3)$ – затраты на реализацию g-й контрмеры для атрибута a_g^3;

M_R – модель рисков СТС;

M_{Rang} – матрица рисков;

$M_{R\ \mathrm{macro}}$ – макромодель рисков;

$M'_{R\ \mathrm{macro}}$ – срез макромодели рисков;

$M_{R\ \mathrm{micro}}$ – микромодель рисков;

$M'_{R\ \mathrm{micro}}$ – срез микромодели рисков;

na_g^2 – наименование g-го риска для атрибута a_g^2;

na_g^3 – наименование g-й контрмеры для атрибута a_g^3;

nO – наименование СТС, для которой проводится оценка риска;

O – множество, характеризующее исследуемую СТС;

O' – множество, характеризующее исследуемую СТС с учетом управления неопределенностью в условиях ограниченных ресурсов;

R – интегрированная модель рисков;

R_{crit} – интегрированная модель критичных рисков;

$type_g$ – тип g-й контрмеры для атрибута a_g^3;

γ – множество рангов риска;

γ_g – ранг g-го риска для атрибута a_g^2.

ВВЕДЕНИЕ

Обеспечение высокого качества, надежности и безопасности функционирования СТС является одной из важнейших научно-технических проблем. В современных условиях, характеризующихся напряженной экономической ситуацией, стратегической линией решения этой проблемы является дальнейшее развитие научных принципов комплексного повышения качества и надежности новой техники, внедрение перспективных технологий автоматизированного проектирования процесса производства с максимальным использованием их потенциальных возможностей и рационального использования всех видов ресурсов.

Готовность организаций и предприятий, разрабатывающих и эксплуатирующих различные СТС, выполнять анализ их надежности и безопасности является обязательным условием государственной и международной сертификации. Главная цель такого анализа – своевременное получение достоверной информации о свойствах надежности и безопасности СТС, необходимой для выработки, обоснования и реализации эффективных проектных и эксплуатационных решений. В настоящее время общепризнанно, что надежность и безопасность СТС является одним из важнейших факторов их конкурентоспособности.

При обеспечении надежности СТС и их процессов немаловажное место занимает риск-ориентированная методология. При этом риск представляет собой комплекс неопределенностей, т.е. разного рода непредвиденных событий, нарушающих нормальную работу системы, связанных с состоянием техники, климатическими условиями, настроением и здоровьем людей – теми факторами, на которые мы не всегда в силах повлиять. Повышение надежности за счет снижения риска имеет большое значение в производственной практике [1–4].

Процесс управления риском охватывает различные аспекты работы с риском, от идентификации и анализа риска до оценки его допустимости и определения потенциальных возможностей снижения риска посредством выбора, реализации и контроля соответствующих управляющих действий.

Объектом исследования данной работы является система надежности СТС, а предметом исследования – модели рисков ГТУ НП.

На сегодняшний день отсутствуют стандарты или другие нормативные документы, которые бы регламентировали оценку и управление рисками для таких СТС как ГТУ НП с учетом их специфики. Ранее авторами были разработаны модели рисков для ГТУ НП и для процесса испытаний данных установок [5, 6], однако формализованы данные модели не были. Поэтому актуальной является задача формализации моделей рисков ГТУ НП для дальнейшего их использования в процессе управления испытаниями ГТУ НП.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть модель рисков СТС будет представлена в виде гиперкуба данных, где измерения гиперкуба – это компоненты моделей рисков. Под гиперкубом данных будем понимать множество ячеек, которым соответствуют множество измерений D и множество атрибутов A:

$$M_R = H(D, A), \tag{1}$$

где $D = \{d_1, d_2, \ldots, d_i\}$ – например, для ГТУ НП измерений три, по числу компонентов в каждой из моделей рисков: d_1 – «Опасности», d_2 – «Риски», d_3 – «Контрмеры» [2, 3]:

$$A = A_{d1} \cup A_{d2} \cup \ldots \cup A_{di};$$

$$A_{di} = \left\{a_1^i, a_2^i, \ldots a_g^i\right\}, \text{ где } i = 1, 2, \ldots, i. \tag{2}$$

Для ГТУ НП каждой ячейке гиперкуба данных соответствует единственно возможный набор атрибутов измерений, содержащий данные об опасностях, рисках и контрмерах.

Интегрированная модель рисков [2, 3], опираясь на формулу (1), может быть представлена в виде следующего отношения:

$$R = M_{R\ \mathrm{macro}} \cup M_{R\ \mathrm{micro}},$$

где $M_{R\ \mathrm{macro}} = H_{\mathrm{macro}}(D_{\mathrm{macro}}, A_{\mathrm{macro}})$; $M_{R\ \mathrm{micro}} = H_{\mathrm{micro}}(D_{\mathrm{micro}}, A_{\mathrm{micro}})$.

Пусть задаче исследования соответствует некое множество O, характеризующее СТС, которое можно представить в виде:

$$O = \{nO, Asset, R\}. \tag{3}$$

При этом под активами $Asset$ будем понимать все доступные ресурсы, т.е. не только материальные ценности, но и временные ресурсы, а также нематериальные ресурсы, к которым можно отнести персонал и окружающую среду.

Однако не все риски несут одинаковую угрозу нанесения ущерба. Поэтому в условиях реально имеющихся (ограниченных) ресурсов возникает задача выделить такое отношение R_{crit}, содержащее только те ячейки гиперкубов, которым соответствуют критичные риски:

$$R_{\text{crit}} = M'_{R\ \text{macro}} \cup M'_{R\ \text{micro}}.$$

Срезы $M'_{R\ \text{macro}}$ и $M'_{R\ \text{micro}}$ формируются в результате фиксации ячеек соответствующих гиперкубов, содержащих наборы атрибутов измерений, в состав которых входят критичные риски.

Тогда, модифицировав множество (3), получим множество O', характеризующее СТС с учетом управления неопределенностью в условиях ограниченных ресурсов:

$$O' = \{nO, Asset, R_{\text{crit}}\}. \quad (4)$$

2 ОБЗОР ЛИТЕРАТУРЫ

Начальный импульс к созданию численных методов оценки надежности был дан авиационной промышленностью. После первой мировой войны в связи с увеличением интенсивности полетов и авиационных катастроф были выработаны критерии надежности для самолетов и требования к уровню безопасности [7, 8]. В дальнейшем методы надежности и оценки рисков разрабатывались по инициативе различных организаций и структур. Пионерами в данной области можно назвать работы по исследованию ракетных комплексов в Германии во время Второй мировой войны. Также велось развитие методов надежности для военной техники со стороны Министерства обороны США, в Великобритании разрабатывались анализ опасности (Trevor Kletz – ICI Ltd) и методы надежности (Green и Bourne – the UKAEA) [7]. Интенсивная работа по оценке риска, связанного с эксплуатацией атомных электростанций, была организована Комиссией по атомной энергии США и завершилась в 1977 г. выпуском отчета «WASH-1400. Анализ безопасности реактора» [7, 8]. Эти работы и многие другие сыграли свою роль и достойны внимания. Они способствовали повышению уровня информированности широкой общественности о потенциальных опасностях и подчеркнули необходимость в совершенствовании методов обеспечения надежности СТС и повышению их безопасности.

Задачи надежности и безопасности сложных технических систем рассматриваются в работах И. А. Рябинина [9], А. С. Можаева [10], В. С. Викторовой [11], K. Kołowrocki [1] и многих других. Значительный вклад в области анализа, оценки и управления рисками сложных технических систем сделали Е. Д. Соложенцев [12], И. А. Рябинин [9], А. С. Можаев [10], E. I. Henly и H. Kumamoto [8], J. D. Andrews и T. R. Moss [7] и другие.

На сегодняшний день принципы надежности, безопасности и анализа рисков технических систем регламентированы в многочисленных нормативных документах. Однако анализ стандартов в области надежности [5] показал отсутствие нормативной документации в системе стандартов Украины по управлению рисками для ГТУ НП. В связи с этим авторами были разработаны модели рисков для ГТУ НП и для их процесса испытаний [5–6].

Согласно стандарту [13] общий процесс оценки риска (risk assessment) состоит из анализа риска (risk analysis) и оценивания риска (risk evaluation). При этом управление риском (risk control) включает в себя оценку риска и снижение/контроль риска. Данная последовательность действий в той или иной формулировке присутствует во многих нормативных документах [14, 15], поэтому будет принята за основу разрабатываемого метода.

3 МАТЕРИАЛЫ И МЕТОДЫ

Для формирования множества (4) необходимо разработать метод управления неопределенностью.

На первом этапе идентифицируем исследуемую СТС nO и активы $Asset$, представляющие ценность для данной системы.

На втором этапе, используя разработанные модели рисков [5, 6], определим опасности и риски, представляющие угрозу для СТС и для ее активов как на макро- так и микроуровне. Для этого идентифицируем атрибуты множеств $A_{d1\ \text{macro}}$, $A_{d1\ \text{micro}}$, $A_{d2\ \text{macro}}$ и $A_{d2\ \text{micro}}$, занеся в них наименования опасностей и рисков.

Далее необходимо определить критичность каждого из выделенных рисков измерения d_2 для определения целесообразности применения контрмер в условиях реально имеющихся (ограниченных) ресурсов.

Для этого на этапе три производится качественное ранжирование выделенных рисков множеств $A_{d2\,\text{macro}}$ и $A_{d2\,\text{micro}}$. В виду недостатка статистических данных об отказах ГТУ НП и их процессов оценивание рисков целесообразно провести с применением экспертных методов оценивания. Были определены и описаны категории рисков – ранги $\gamma = \{\text{I, II, III, IV}\}$, которые находятся в диапазоне от неприемлемого риска до пренебрежимо малого риска [16].

На основе выделенных рангов рисков γ была построена матрица рисков M_{Rang}, где значение каждой из ячеек матрицы соответствует рангу, присвоенному исходя из частоты опасных событий и последствий опасных событий [16]. Используя матрицу рисков M_{Rang}, каждому из выделенных рисков присваивается соответствующий ранг.

На четвертом этапе по результатам качественного ранжирования риски с рангами I и II признаются наиболее критичными и формируется отношение R_{crit}, путем фиксации срезов $M'_{R\ \text{macro}}$ и $M'_{R\ \text{micro}}$.

Принимая во внимание вышеизложенное, прослеживается необходимость сохранения сопутствующих данных о рисках. Поэтому каждый из элементов множества атрибутов A_{d2}, описанного формулой (2), представим в виде кортежа:

$$a_g^2 = \left\langle na_g^2; \gamma_g; L(a_g^2) \right\rangle, \text{ где } g = 1, 2, \ldots, g. \quad (5)$$

На пятом этапе для всех критичных рисков отношения R_{crit} определяются контрмеры для их компенсации или предотвращения, путем идентификации множеств атрибутов $A'_{d3\ \text{macro}}$ и $A'_{d3\ \text{micro}}$.

На сегодняшний день общепризнано, что концепция «абсолютной» безопасности неадекватна внутренним законам природы, которые имеют вероятностный характер. Поскольку даже использование самых эффективных систем безопасности, самых современных методов контроля над технологическими процессами не обеспечивает – и в принципе не может обеспечить – абсолютную надежность работы, исключаю-

щую аварийные ситуации [12]. Поэтому при определении мер по снижению риска (контрмер) мы руководствовались принципом приемлемого риска ALARP. Данный принцип подразумевает, что реализация контрмер не должна требовать неоправданно высоких затрат или неоправданно больших усилий, т.е. применение контрмер осуществляется за счет реально имеющихся (ограниченных) ресурсов и целесообразно только в случае, если затраты на реализацию контрмер меньше размера ущерба от наступления рискового события.

Если предположить, что для одного и того же риска могут быть выделены несколько альтернативных контрмер, то с учетом принципа ALARP возникает задача выбора наиболее целесообразного варианта.

На шестом этапе необходимо выбрать контрмеры из имеющихся альтернативных вариантов. Для этого необходимо разработать подметод выбора контрмер.

Важным аспектом метода выбора контрмер является определение критериев сравнения. Одним из таких критериев может служить тип контрмеры.

В ходе работы были выделены следующие типы контрмер.

1. По времени применения контрмеры могут разделяться на упреждающие и контрмеры реакции на риск. Упреждающие контрмеры, как видно из названия, способны полностью предупредить наступление возможного риска, при этом их применение осуществляется до наступления рискового события. А контрмеры реакции на риск предназначены для компенсации, как правило, частичной, риска и применяются после наступления нежелательных событий.

При этом затраты по реализации упреждающих контрмер будут взыматься вне зависимости от того наступило или нет событие неуспеха в отличие от контрмер реакции на риск, затраты на реализацию которых требуются только в случае наступления риска. С другой стороны, применение последних не всегда полностью может компенсировать риск, что, в конечном счете, может привести к увеличению суммарного риска осуществления данного события.

2. По объекту воздействия контрмеры могут подразделяться на следующие типы:

– сокращение или исключение первичных угроз, дефектов и ошибок в системе исследования, обусловленных недостатками их проектирования, разработки, модификации или планирования, отражающихся на рисках функциональной пригодности или характеристиках системы;

– сокращение уязвимости компонентов системы при воздействии на них угроз, дефектов и ошибок, путем введения средств защиты для блокирования их возможного негативного воздействия на риски функционирования;

– непосредственное изменение и сокращение последствий проявления рисков путем их оперативного обнаружения и ликвидации ущерба.

Необходимо отметить, что исключение первичных угроз и сокращение уязвимостей ведет в конечном итоге к улучшению самой системы, тогда как сокращение последствий рисков происходит, как правило, при сохранении вызывающих их первичных источников и причин.

Исходя из вышеизложенных характеристик, была создана сводная таблица 1, в которой приведено соотношение контрмер различных типов. В ячейках данной таблицы приведены ранги критических рисков, для которых применение контрмеры данного типа предпочтительнее.

Выделенные типы контрмер будут использованы как критерии сравнения при выборе контрмер из альтернативных вариантов, исходя из рекомендаций таблицы 1.

Рассмотрим метод выбора контрмер. Как было сказано выше, для одного и того же риска могут быть выделены несколько альтернативных контрмер. Выбор наиболее эффективного варианта и есть цель метода выбора контрмер.

Входными данными для данного метода будут служить сформированные ранее множества атрибутов $A'_{d3\,\mathrm{macro}}$ и $A'_{d3\,\mathrm{micro}}$, содержащие контрмеры для интегрированной модели критичных рисков R_{crit}.

Шаг 1. Определим некое множество $A'_{d3\,\mathrm{alt}}$, которое содержало бы только группы альтернативных контрмер из множеств $A'_{d3\,\mathrm{macro}}$ и $A'_{d3\,\mathrm{micro}}$.

Шаг 2. Для каждой из групп альтернативных контрмер множества $A'_{d3\,\mathrm{alt}}$ определим следующие критерии сравнения: затраты на реализацию контрмер; k' – степень компенсации события неуспеха от применения контрмеры; тип контрмер – как по времени применения, так и по объекту воздействия.

Шаг 3. Оценим относительные веса k каждой из контрмер множества $A'_{d3\,\mathrm{alt}}$ по отношению к альтернативным вариантам с использованием метода анализа иерархий [6]. Данный метод позволяет найти такой вариант, который наилучшим образом согласуется с пониманием сути проблемы и требованиями к ее решению, а также сравнить и выполнить количественную оценку вариантов.

Метод анализа иерархий состоит в декомпозиции проблемы на более простые составляющие части и дальнейшей обработке последовательности суждений лица, принимающего решение, с помощью матриц парных сравнений. В результате работы с матрицами рассчитываются относительные веса элементов в иерархии, и выбирается наилучшая с точки зрения цели альтернатива.

Шаг 4. На основании относительных весов k каждой из контрмер, модифицируя множество $A'_{d3\,\mathrm{alt}}$, определим множество $A'_{d3\,\mathrm{mod}}$, содержащее перечень контрмер, рекомендованных к применению.

Шаг 5. Согласно множеству $A'_{d3\,\mathrm{mod}}$ уточним исходные множества $A'_{d3\,\mathrm{macro}}$ и $A'_{d3\,\mathrm{micro}}$, исключив контрмеры, применение которых не целесообразно: $A'_{d3\,\mathrm{macro}} / \overline{A'_{d3\,\mathrm{mod}}}$; $A'_{d3\,\mathrm{micro}} / \overline{A'_{d3\,\mathrm{mod}}}$.

Для хранения описанных данных определим элементы множества A_{d3}, описанного формулой (2), как

$$a_g^3 = \left\langle na_g^3; L(a_g^3); type_g; k'_g; k_g; h \right\rangle, \qquad (6)$$

где $g = pl$, где $p = 1, 2, \ldots$; p – номер кортежа в множестве атрибутов A_{d2}, содержащего информацию о риске, для которого была выделена рассматриваемая контрмера;

Таблица 1 – Соотношение контрмер различных типов

	По объекту воздействия		
По времени применения	Исключение первичных угроз	Сокращение уязвимости	Сокращение последствий
Упреждающие контрмеры	I	I	–
Контрмеры реакции на риск	–	II	II

$l = 0$ в случае если для данного риска была выделена одна контрмера, $l = 1, 2, \ldots n$, если для данного риска было выделено n контрмер, в том числе альтернативных;

Схема метода выбора контрмер из альтернативных вариантов показана на рисунке 1.

На седьмом заключительном этапе метода управления неопределенностью производится формирование итогового множества O' (4). Данное множество содержит информацию об области критичных рисков и контрмерах, применение которых ведет к снижению общего уровня риска для исследуемой СТС.

Если оценка рисков СТС используется для обеспечения непрерывного процесса управления риском, то ее необходимо выполнять таким образом, чтобы она могла корректироваться на протяжении всего жизненного цикла СТС. Оценка должна обновляться по мере поступления новой информации и в соответствии с потребностями процесса управления.

Схема метода управления неопределенностью показана на рисунке 2.

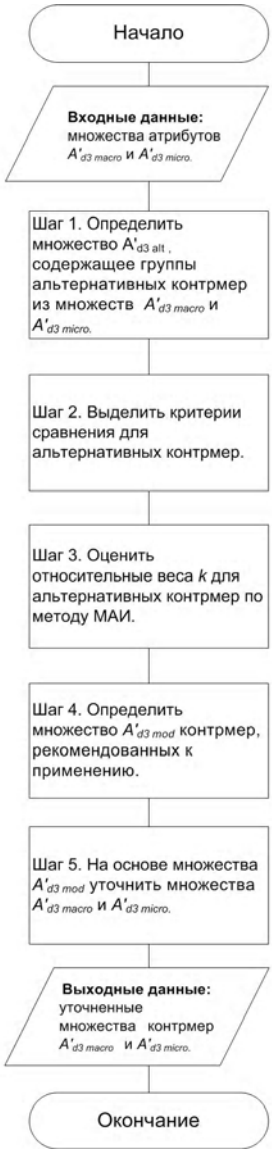


Рисунок 1 – Схема метода выбора контрмер из альтернативных вариантов

Рисунок 2 – Схема метода управления неопределенностью

4 ЭКСПЕРИМЕНТЫ

Одной из центральных проблем для ГТУ НП является обеспечение высокой надежности, безопасности и большого ресурса работы. Подтверждение данных параметров осуществляется на испытаниях различных видов, проводимых практически на всех стадиях жизненного цикла ГТУ НП: разработки, производства, эксплуатации. Поэтому апробацию разрабатываемого метода управления неопределенностью проведем для процесса испытаний ГТУ НП, общая модель основных этапов которого представлена на рисунке 3 [17].

Для наглядности представления разработанного метода более подробно остановимся на этапе «Планирование» процесса испытаний ГТУ НП [17], поскольку ему необходимо уделить особое внимание, т.к. эффективное управление всем процессом испытаний ГТУ НП в первую очередь зависит от рационального планирования всех работ и оптимального распределения ресурсов.

5 РЕЗУЛЬТАТЫ

Активами для всего процесса испытаний являются материальные и нематериальные ресурсы ГТУ НП:

– материальная база, обеспечивающая проведение процесса испытаний;

– газотурбинная установка, включая все входящие функциональные системы и оборудование;

– участники испытаний и окружающая среда;

– результаты испытаний (информация о работе установки, данные результатов испытаний, отчеты и техническая документация).

Следует отметить, что для этапа «Планирование» не характерны опасности связанные с материальными ресурсами, обеспечивающими проведение испытаний и относящимися к самой ГТУ НП, а также опасности, связанные с причинением вреда окружающей среде. Также для данного этапа в виду его специфики отсутствует микромодель рисков, отражающая риски, влияющие на выполнение функций безопасности, что также упрощает описание данного примера и способствует его наглядности. Поэтому в названиях переменных будут опущены индексы macro и micro.

В таблице 2 представлены выделенные для данного этапа опасности, риски и результаты ранжирования. Как видно, не было выделено ни одного риска ранга I, поэтому R_{crit} будет включать только риски ранга II.

Далее для каждого риска, входящего в R_{crit}, были определены контрмеры A'_{d3}, как видно из таблицы 3 некоторые из них содержали альтернативные варианты. Все альтернативные пары были выделены во множество $A'_{d3\,alt}$.

Для определения целесообразности применения каждой из альтернативных контрмер были рассчитаны относительные веса k.

Пример выбора относительных весов k для этапа «Планирование» процесса испытаний ГТУ НП приведен в таблице 4.

Как видно из таблиц 3 и 4, согласно относительным весам k каждой из альтернативных контрмер, риски na'^{2}_{1} и na'^{2}_{3} целесообразно компенсировать упреждающими контрмерами, которые предназначены для устранения первичных угроз, тогда как риск na'^{2}_{5} – контрмерой реакции на риск, предназначенной для сокращения последствий. Поэтому множество $A'_{d3\,\mathrm{mod}}$ примет вид $A'_{d3\,\mathrm{mod}} = \{na'^{3}_{11\,alt}; na'^{3}_{31\,alt}; na'^{3}_{52\,alt}\}$, согласно которого будет уточнено множество A'_{d3} и сформировано итоговое множество O'.

6 ОБСУЖДЕНИЕ

Применение метода управления неопределенностью в условиях ограниченных ресурсов для процесса испытаний ГТУ НП позволило сократить время на проведение оценки рисков для данного процесса за счет снижения количества обрабатываемых данных. Анализ количества обрабатываемых элементов модели рисков приведен в таблице 5.

Также полученный метод позволяет выбрать те контрмеры, которые наиболее эффективно снижают общий уровень риска для данного процесса, что в конечном счете позволяет снизить затраты на ликвидацию последствий от возможного риска. Полученная формализованная модель рисков ГТУ НП и их процессов может быть использована при автоматизации процесса управления неопределенностью ГТУ НП в условиях ограниченных ресурсов.

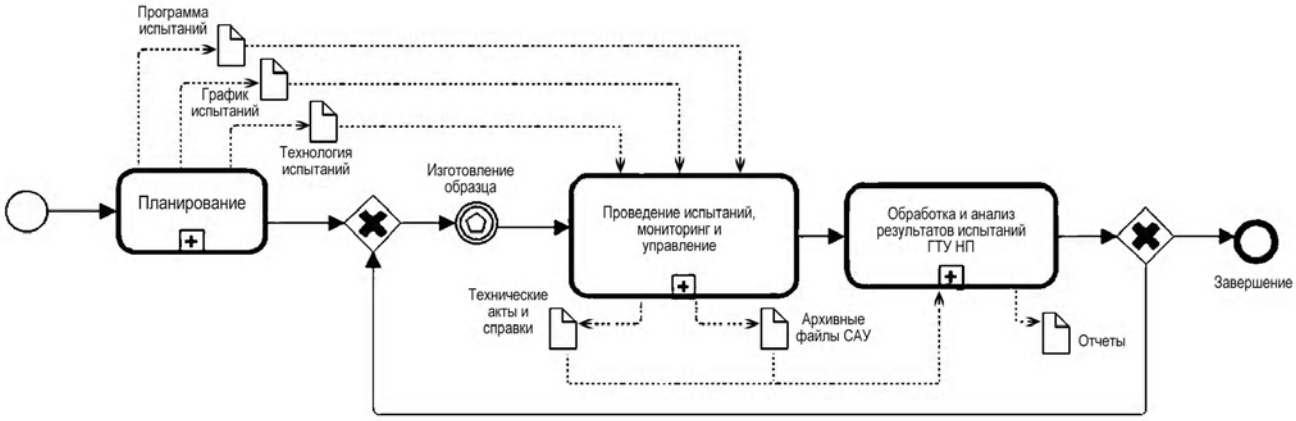


Рисунок 3 – Общая модель основных этапов процесса испытаний ГТУ НП

Таблица 2 – Оценка риска этапа «Планирование» процесса испытаний ГТУ НП

g	Опасности множества A_{d1}	Наименования рисков множества A_{d2}, na_g^2	Ранг, γ_g
1	a_1^1 – Ошибки при создании программы испытаний (ПИ) по причине человеческого фактора или недостатка достоверной исходной информации	na_1^2 – Риск создания некорректной ПИ	II
2		na_2^2 – Риск нарушения сроков создания ПИ	III
3	a_2^1 – Ошибки при создании технологии испытаний (ТИ) по причине человеческого фактора или недостатка достоверной исходной информации	na_3^2 – Риск создания некорректной ТИ	II
4		na_4^2 – Риск нарушения сроков создания ТИ	III
5	a_3^1 – Ошибки при создании графика испытаний (ГИ) по причине человеческого фактора или недостатка достоверной исходной информации	na_5^2 – Риск нарушения сроков создания ГИ	II
6		na_6^2 – Риск создания некорректного ГИ	III
7	a_4^1 – Опасности связанные с персоналом	na_7^2 – Риски, связанные с угрозами для персонала	IV

Таблица 3 – Компоненты отношения R_{crit} для этапа «Планирование» процесса испытаний ГТУ НП

Наименования рисков множества A'_{d2}, na'^2_g	Наименование контрмеры множества A'_{d3}, na'^3_x
na'^2_1 – Риск создания некорректной программы испытаний (ПИ)	na'^3_{11} – Содержание квалифицированного персонала на постоянной основе
	na'^3_{12} – Временное привлечение квалифицированных кадров
	na'^3_{13} – Корректная организация работ
na'^2_3 – Риск создания некорректной технологии испытаний (ТИ)	na'^3_{31} – Содержание квалифицированного персонала на постоянной основе
	na'^3_{32} – Временное привлечение квалифицированных кадров
	na'^3_{33} – Корректная организация работ
na'^2_5 – Риск нарушения сроков создания графика испытаний (ГИ)	na'^3_{51} – Обеспечение изготовления образца ГТУ НП
	na'^3_{52} – Предусмотреть резерв материальных и нематериальных ресурсов для компенсации неуспеха постфактум

Таблица 4 – Компоненты множества $A'_{d3\,alt}$ для этапа «Планирование» процесса испытаний ГТУ НП

Наименование контрмеры множества $A'_{d3\,alt}$, $na'^3_{g\,alt}$	Относительные веса, k_g	Степень компенсации события неуспеха, k'_g
$na'^3_{11\,alt}$ – Содержание квалифицированного персонала на постоянной основе	0,67	0,17
$na'^3_{12\,alt}$ – Временное привлечение квалифицированных кадров	0,33	0,83
$na'^3_{31\,alt}$ – Содержание квалифицированного персонала на постоянной основе	0,67	0,17
$na'^3_{32\,alt}$ – Временное привлечение квалифицированных кадров	0,33	0,83
$na'^3_{51\,alt}$ – Обеспечение изготовления образца ГТУ НП	0,43	0,14
$na'^3_{52\,alt}$ – Предусмотреть резерв материальных и нематериальных ресурсов для компенсации неуспеха постфактум	0,57	0,86

Таблица 5 – Анализ количества элементов модели рисков для этапа «Планирование» процесса испытаний ГТУ НП

Наименование измерений гиперкуба	Количество элементов (атрибутов) модели риска	
	до применения метода управления неопределенностью	после применения метода управления неопределенностью
Опасности	4	4
Риски	7	3
Контрмеры	12	5
Итого:	23	12

ВЫВОДЫ

В работе решена задача формализации модели рисков ГТУ НП и их процессов. Что позволило разработать метод управления неопределенностью в условиях ограниченных ресурсов, основанный на методе оценки рисков и усовершенствованный методом выбора контрмер и позволяющий сформировать область критичных рисков для исследуемой СТС.

Усовершенствован метод выбора контрмер из альтернативных вариантов, основанный на методе приемлемого риска ALARP и методе анализа иерархий. Данный метод позволяет дать экономически обоснованную оценку уровня приемлемого риска.

Согласно разработанному методу управления неопределенностью выбор контрмер производится только для наиболее критичных рисков, что исключает необходимость проводить анализ вариантов компенсации для всех возможных рисков. Данное преимущество позволяет в значительной степени экономить время для проведения оценки рисков и средства для их компенсации.

Практическая ценность данной работы заключается в проведенной апробации разработанного метода управления неопределенностью в условиях ограниченных ресурсов для процесса испытаний ГТУ НП, в частности, для этапа «Планирование». Данная апробация показала, что решение задачи определения области критичных рисков позволило снизить количество обрабатываемых элементов модели рисков на 48%.

По результатам проведенной работы можно сделать вывод, что риск с большей долей вероятности и приносящий больший ущерб целесообразно компенсировать до наступления этого события путем исключения первичных угроз или сокращения уязвимости компонентов исследуемой СТС. Тогда как риск с большей долей неопределенности необходимо компенсировать постфактум за счет сокращения последствий проявления рисков путем их оперативного обнаружения и ликвидации ущерба с помощью резервных фондов материальных и нематериальных ресурсов, формирование которых и должно происходить, исходя из приведенного анализа.

БЛАГОДАРНОСТИ

Авторы выражают благодарность за оказанную помощь и содействие руководству Управления энергетических установок и газоперекачивающих агрегатов АО «Мотор Сич» в лице Главного конструктора – начальника управления В. И. Морозова, Заместителя главного конструктора – начальника управления В. П. Митина, Главного конструктора ГПА В.Н. Разношинского.

Данная работа выполнена на кафедре программных средств Запорожского национального технического университета в рамках темы № ДР0112U005351 «Разработка методов и информационных технологий для автоматизации процессов управления сложными объектами и системами» при поддержке международного проекта Темпус 530278-TEMPUS-1-2012-1-DE-TEMPUS-JPHES «Промышленное сотрудничество и креативное инженерное образование на основе удаленного дистанционного инженерного и виртуального инструментария» («Industrial Cooperation and Creative Engineering Education based on Remote Engineering and Virtual Instrumentation» [ICo-op]). Также работа выполнена в рамках договора о научно-техническом сотрудничестве № 8940/10-Д (УЭУиГПА) от 26.11.2010 (продлен дополнительным соглашением № 1 от 13.12.12) между Запорожским национальным техническим университетом и АО «Мотор Сич» «Разработка и программная реализация планирования испытаний газотурбинном установок наземного применения» (Исполнители от ЗНТУ: Табунщик Г. В., Шитикова Е. В.).

СПИСОК ЛИТЕРАТУРЫ

1. Kołowrocki K. Reliability and Safety of Complex Technical Systems and Processes: Modeling-Identification-Prediction-Optimization / K. Kołowrocki, J. Soszycska-Budny. – London : Springer, 2011. – 405 p. DOI: 10.1007/978-0-85729-694-8.
2. Kuo W. An annotated overview of system-reliability optimization / W. Kuo, V. R. Prasad // IEEE Transactions on Reliability. – 2000. – Vol. 49(2). – P. 176–187.
3. Kuo W. Optimal Reliability Modeling: Principles and Applications / W. Kuo, M. J. Zuo. – Hoboken : John Wiley & Sons, 2003. – 559 p.
4. Vercellis S. Business Intelligence: Data Mining and optimization for decision making / S. Vercellis. – Hoboken : John Wiley & Sons. 2009. – 417 p. DOI: 10.1002/9780470753866.
5. Шитикова Е. В. Анализ рисков газотурбинных установок наземного применения / Е. В. Шитикова, Г. В. Табунщик // Вісник двигунобудування. – 2012. – №1. – С. 54–59.
6. Шитикова Е. В. Моделирование рисков процесса испытаний газотурбинных установок наземного применения / Е. В. Шитикова, Г. В. Табунщик // Наукові праці Донецького національного технічного університету, серія: «Проблеми моделювання та автоматизації проектування». – 2012. – № 1(10)–2(11). – С. 79–89.
7. Andrews J. D. Reliability and Risk Assessment. Second edition / J. D. Andrews, T. R. Moss. – London and Bury St Edmunds : Professional Engineering Publishing, 2002. – 540 p.
8. Henley E. I. Probabilistic Risk Assessment and Management for Engineers and Scientists / E. I. Henley, H. Kumamoto. – New York : Wiley-IEEE Press, 1996. – 620 p.
9. Рябинин И. А. Надежность и безопасность структурно-сложных систем / И.А. Рябинин. – СПб. : Политехника, 2000. – 248 с.
10. Можаев А. С. Применение автоматизированного структурно-логического моделирования для проектного расчета надежности АСУ / А. С. Можаев, М. С. Скворцов, А. В. Струков // Нефть. Газ. Новации. – 2010. – № 9. – С. 72–78.
11. Викторова В. С. Комплекс программ для анализа надежности, безопасности и эффективности технических систем / В. С. Викторова, А. С. Степанянц // Приборы и системы управления. – 1998. – № 6. – С. 11–17.
12. Соложенцев Е. Д. Сценарное логико-вероятностное управление риском в бизнесе и технике. Изд. 2-е. / Е. Д. Соложенцев. – СПб. : Издательский дом : Бизнес-пресса, 2006. – 530 с.
13. Менеджмент риска. Анализ риска технологических систем : ГОСТ Р 51901.1-2002. – [Дата введения 2003-09-01]. – М. : ИПК Издательство стандартов, 2002. – 28 с. – (Государственный стандарт Российской Федерации).
14. Безпечність машин. Принципи оцінювання ризику (EN 1050:1996, IDT) : ДСТУ EN 1050:2003. – [Чинний від 2004-10-01]. – К. : Держспоживстандарт України, 2005. – 22 с. – (Національний стандарт України).
15. Безопасность машин и оборудования. Оценка риска : ГОСТ Р 54124-2010. – [Дата введения 2012-06-01]. – М. : Стандартинформ, 2013. – 102 с. – (Государственный стандарт Российской Федерации).
16. Ранжирование рисков газотурбинных установок наземного применения / [Е. В. Шитикова, Г. В. Табунщик, В. П. Митин, В. И. Морозов] // Моделирование и Анализ Безопасности и Риска в Сложных Системах : Международная научная школа МАБР-2011, Санкт-Петербург, 28 июня – 02 июля 2011 г. : труды школы. – СПб. : ГУАП, 2011. – С. 277–279.
17. Шитикова Е. В. Информационная модель процесса испытаний газотурбинных установок наземного применения / Е. В. Шитикова, Г. В. Табунщик // Радіоелектроніка, інформатика, управління. – 2012. – № 1(28). – С. 101–108.

Статья поступила в редакцию 24.11.2014.
После доработки 10.12.2014.

Шитікова О. В.[1], Табунщик Г. В.[2]
[1]Аспірант кафедри програмних засобів Запорізького національного технічного університету, Запоріжжя, Україна
[2]Канд. техн. наук, доцент, доцент кафедри програмних засобів Запорізького національного технічного університету, Запоріжжя, Україна

МЕТОД УПРАВЛІННЯ НЕВИЗНАЧЕНІСТЮ В УМОВАХ ОБМЕЖЕНИХ РЕСУРСІВ

Розглянуті питання надійності та функціональної безпеки газотурбінних установок наземного використання. Проведена формалізація раніше розроблених моделей ризиків газотурбінних установок наземного використання та їх процесів. Запропоновано метод управління невизначеністю складних технічних систем і їх процесів в умовах обмежених ресурсів, заснований на методі оцінки ризиків і методі вибору контрзаходів, що дозволяє сформувати область критичних ризиків для досліджуваної складної технічної системи. Удосконалено метод вибору контрзаходів з альтернативних варіантів, заснований на методі аналізу ієрархій і методі прийнятного ризику ALARP, оскільки застосування контрзаходів здійснюється за рахунок фактично наявних (обмежених) ресурсів. Проведено класифікацію контрзаходів за часом застосування та за об'єктом впливу. Виділені типи були використані як один з критеріїв порівняння при виборі контрзаходів із альтернативних варіантів. Створена зведена таблиця співвідношення контрзаходів різних типів. Метод вибору контрзаходів з альтернативних варіантів дозволяє зробити висновок в умовах обмежених ресурсів і дати економічно обґрунтовану оцінку при ухваленні рішення щодо рівня прийнятного ризику. Отримані методи апробовані для оцінки ризиків процесу випробувань газотурбінних установок наземного використання.

Ключові слова: газотурбінні установки наземного використання, модель ризиків, метод управління невизначеністю, метод вибору контрзаходів.

Shitikova Yel.[1], Tabunshchik G.[2]
[1]Postgraduate student of Software Tools Department, Zaporizhzhya National Technical University, Zaporizhzhya, Ukraine
[2]PhD, Associate professor, Associate professor of Software Tools Department, Zaporizhzhya National Technical University, Zaporizhzhya, Ukraine

METHOD OF MANAGING UNCERTAINTY IN RESOURCE-LIMITED SETTINGS

The article deals with the questions of reliability and functional safety of the gas turbine units for terrestrial usage. Previously developed risk models of the gas turbine units for terrestrial usage and processes are formalized. Authors propose the method of managing uncertainty of complex technical systems and their processes in resource-limited settings. It is based on the method of risk assessment and the method for choice of alternative countermeasures. It allows to form the risk critical area for the complex technical systems. The method for choice of alternative countermeasures improved. It is based on the Analytic Hierarchy Process and the method of acceptable risk ALARP, since the application of countermeasures realize at the expense actually available (limited) resources. Countermeasures were classified by time of application and impact object. Dedicated types served as a compare criterion of the method for choice of alternative countermeasures. A summary table was created for ratio of different types of countermeasures. The method for choice of alternative countermeasures allows to make a conclusion in resource-limited settings and provides economically reasonable estimate of the decision about the level of acceptable risk. The derived methods are tested for the GTU tests process risk assessment.

Keywords: gas turbine unit for terrestrial usage, risk model, method of managing uncertainty, method for choice of countermeasures.

REFERENCES

1. Kołowrocki K., Soszycska-Budny J. Reliability and Safety of Complex Technical Systems and Processes: Modeling-Identification-Prediction-Optimization. London, Springer, 2011, 405 p. DOI: 10.1007/978-0-85729-694-8.
2. Kuo W., Prasad V. R. An annotated overview of system-reliability optimization, *IEEE Transactions on Reliability*, 2000, Vol. 49(2), pp. 176–187.
3. Kuo W., Zuo M. J. Optimal Reliability Modeling: Principles and Applications. Hoboken, John Wiley & Sons, 2003, 559 p.
4. Vercellis S. Business Intelligence: Data Mining and optimization for decision making. Hoboken, John Wiley & Sons, 2009, 417 p. DOI: 10.1002/9780470753866.
5. Shitikova E. V., Tabunshhik G. V. Analiz riskov gazoturbinnyh ustanovok nazemnogo primenenija, *Visnyk dvygunobuduvannja*, 2012, No. 1, pp. 54–59.
6. Shitikova E. V., Tabunshhik G. V. Modelirovanie riskov processa ispytanij gazoturbinnyh ustanovok nazemnogo primenenija, *Naukovi praci Donec'kogo nacional'nogo tehnichnogo universytetu, serija: «Problemy modeljuvannja ta avtomatyzacii' proektuvannja»*, 2012, No. 1(10)–2(11), pp. 79–89.
7. Andrews J. D., Moss T. R. Reliability and Risk Assessment. Second edition. London and Bury St Edmunds, Professional Engineering Publishing, 2002, 540 p.
8. Henley E. I., Kumamoto H. Probabilistic Risk Assessment and Management for Engineers and Scientists. New York, Wiley-IEEE Press, 1996, 620 p.
9. Rjabinin I. A. Nadezhnost' i bezopasnost' strukturno-slozhnyh sistem. SPb, Politehnika, 2000, 248 p.
10. Mozhaev A. S., Skvorcov M. S., Strukov A. V. Primenenie avtomatizirovannogo strukturno-logicheskogo modelirovanija dlja proektnogo rascheta nadezhnosti ASU, *Neft'. Gaz. Novacii,* 2010, No. 9, pp. 72–78.
11. Viktorova V. S., Stepanjanc A. S. Kompleks programm dlja analiza nadezhnosti, bezopasnosti i jeffektivnosti tehnicheskih sistem, *Pribory i sistemy upravlenija*, 1998, No. 6, pp. 11–17.
12. Solozhencev E. D. Scenarnoe logiko-verojatnostnoe upravlenie riskom v biznese i tehnike. Izd. 2-e. SPb, Izdatel'skij dom, Biznes-pressa, 2006, 530 p.
13. Menedzhment riska. Analiz riska tehnologicheskih sistem : GOST R 51901.1-2002. – [Data vvedenija 2003-09-01], Moscow, IPK Izdatel'stvo standartov, 2002, 28 p. – (Gosudarstvennyj standart Rossijskoj Federacii).
14. Bezpechnist' mashyn. Pryncypy ocinjuvannja ryzyku (EN 1050:1996, IDT) : DSTU EN 1050:2003. – [Chynnyj vid 2004-10-01]. Kyiv, Derzhspozhyvstandart Ukrai'ny, 2005, 22 p. – (Nacional'nyj standart Ukrai'ny).
15. Bezopasnost' mashin i oborudovanija. Ocenka riska : GOST R 54124-2010. – [Data vvedenija 2012-06-01]. Moscow, Standartinform, 2013, 102 p. – (Gosudarstvennyj standart Rossijskoj Federacii).
16. Shitikova E. V., Tabunshhik G. V., Mitin V. P., Morozov V. I. Ranzhirovanie riskov gazoturbinnyh ustanovok nazemnogo primenenija, Modelirovanie i Analiz Bezopasnosti i Riska v Slozhnyh Sistemah, Mezhdunarodnaja nauchnaja shkola MABR-2011, Sankt-Peterburg, 28 ijunja– 02 ijulja 2011 g. : trudy shkoly. SPb, GUAP, 2011, pp. 277–279.
17. Shitikova E. V., Tabunshhik G. V. Informacionnaja model' processa ispytanij gazoturbinnyh ustanovok nazemnogo primenenija, *Radioelektronika, informatyka, upravlinnja,* 2012, No. 1(28), pp. 101–108.

Наукове видання

Радіоелектроніка, інформатика, управління

№ 2/2015

Науковий журнал

Головний редактор – д-р фіз.-мат. наук В. В. Погосов

Заст. головного редактора – д-р техн. наук С. О. Субботін

Комп'ютерне моделювання та верстання — С. В. Зуб
Редактор англійських текстів — С. О. Субботін

Оригінал-макет підготовлено у редакційно-видавничому відділі ЗНТУ

Свідоцтво про державну реєстрацію
КВ № 6904 від 29.01.2003.

Підписано до друку 15.06.2015. Формат 60×84/8.
Папір офс. Різогр. друк. Ум. друк. арк. 11,16.
Тираж 300 прим. Зам. № 633.

69063, м. Запоріжжя, ЗНТУ, друкарня, вул. Жуковського, 64

Свідоцтво суб'єкта видавничої справи
ДК № 2394 від 27.12.2005.