

СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

МІЖНАРОДНИЙ НАУКОВО-ТЕХНІЧНИЙ ЖУРНАЛ

№ 2

2010

ЗАСНОВАНИЙ У ЛИПНІ 2001 Р.

РЕДАКЦІЙНА КОЛЕГІЯ:

Головний редактор

М.З. ЗГУРОВСЬКИЙ, акад. НАН України

Заступник головного редактора

Н.Д. ПАНКРАТОВА, проф., д. т. н.

Члени редколегії:

П.І. АНДОН, чл.-кор. НАН України

А.В. АНІСІМОВ, проф., д. ф.-м. н.

В.Г. БОНДАРЕНКО проф., д. ф.-м. н.

Ж. БОННІН, проф., Франція

Р. ВОРОНКА, проф., США

С.М. ГРІША, проф., д. т. н.

О.Д. ГВИШАНИ, проф., д. ф.-м. н.,
Росія

В.І. КОСТЮК, проф., д. т. н.

В.М. КУНЦЕВИЧ, акад. НАН України

Г.М. ЛУЦЬКИЙ, проф., д. т. н.

О.С. МАКАРЕНКО проф., д. ф.-м. н.

О.А. МОЛЧАНОВ, проф., д. т. н.

В.В. ОСТАПЕНКО, проф., д. ф.-м. н.

О.А. ПАВЛОВ, проф., д. т. н.

В.Н. РЕДЬКО, акад. НАН України

Т.Л. СААТІ, проф., США

І.В. СЕРГІЄНКО, акад. НАН України

В.В. СКОПЕЦЬКИЙ, чл.-кор. НАН України

В.П. ТАРАСЕНКО, проф., д. т. н.

Я.С. ЯЦКІВ акад. НАН України

АДРЕСА РЕДАКЦІЇ:

03056, м. Київ,

пр. Перемоги, 37, корп. 35,

Навчально-науковий комплекс

«Інститут прикладного системного аналізу»

НТУУ «КПІ» МОН та НАН України

Тел.: **406-81-44**; Факс: **406-81-44**

E-mail: journal@mmsa.ntu-kpi.kiev.ua

http://journal.iasa.com.ua

У номері:

- **Прогресивні інформаційні технології, високопродуктивні комп'ютерні системи**
- **Проблеми прийняття рішень і управління в економічних, технічних, екологічних і соціальних системах**
- **Математичні методи, моделі, проблеми і технології дослідження складних систем**
- **Методи аналізу та управління системами в умовах ризику і невизначеності**
- **Евристичні методи та алгоритми в системному аналізі та управлінні**
- **Нові методи в системному аналізі, інформатиці та теорії прийняття рішень**

NATIONAL ACADEMY OF SCIENCES OF UKRAINE
EDUCATIONAL-SCIENTIFIC COMPLEX
«INSTITUTE FOR APPLIED SYSTEM ANALYSIS»
AT THE NATIONAL TECHNICAL UNIVERSITY OF UKRAINE «KPI»
OF MES AND NAS OF UKRAINE

SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES

INTERNATIONAL SCIENTIFIC AND TECHNICAL JOURNAL

№ 2

2010

IT IS FOUNDED IN JULY, 2001

EDITORIAL BOARD:

The editor – in – chief

M.Z. ZGUROVSKY, Academician of
NASU

Deputy editor – in – chief

N.D. PANKRATOVA, Prof., Ukraine

Associate editors:

F.I. ANDON, Correspondent
member of NASU

A.V. ANISIMOV, Prof., Ukraine

V.G. BONDARENKO Prof., Ukraine

Z. BONNIN, Prof., France

R. VORONKA, Prof., USA

S.N. GRISHA, Prof., Ukraine

A.D. GVISHIANI Prof., Russia

V.I. KOSTJUK, Prof., Ukraine

V.M. KUNTHEVICH, Academician of NASU

G.M. LUTSKY, Prof., Ukraine

A.S. MAKARENKO Prof., Ukraine

A.A. MOLCHANOV, Prof., Ukraine

V.V. OSTAPENKO, Prof., Ukraine

A.A. PAVLOV, Prof., Ukraine

V.N. REDKO, Academician of NASU

T.L. SAATY, Prof., USA

I.V. SERGIENKO, Academician of NASU

V.V. SKOPETSKY Correspondent
member of NASU

V.P. TARASENKO, Prof., Ukraine

Ya.S. YATSKIV Academician of NASU

THE EDITION ADDRESS:

03056, Kyiv,
av. Peremogy, 37, building 35,
Educational-Scientific Complex
«Institute for Applied System Analysis» at the
National Technical University of Ukraine «KPI»
of MES and NAS of Ukraine
Phone: 406-81-44; Fax: 406-81-44
E-mail: journal@mmsa.ntu-kpi.kiev.ua
http://journal.iasa.com.ua

In the issue:

• **Progressive information technologies, high-efficiency computer systems**

• **Decision making and control in economic, technical, ecological and social systems**

• **Mathematical methods, models, problems and technologies for complex systems research**

• **Methods of system analysis and control in conditions of risk and uncertainty**

• **Heuristic methods and algorithms in system analysis and control**

• **New methods in system analysis, computer science and theory of decision making**

Шановні читачі!

Навчально-науковий комплекс «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України видає міжнародний науково-технічний журнал

«СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ».

Журнал публікує роботи теоретичного і прикладного характеру в широкому спектрі проблем, пов'язаних із системними дослідженнями та інформаційними технологіями.

Провідні тематичні розділи журналу:

Теоретичні та прикладні проблеми і методи системного аналізу; теоретичні та прикладні проблеми інформатики; автоматизовані системи управління; прогресивні інформаційні технології, високопродуктивні комп'ютерні системи; проблеми прийняття рішень і управління в економічних, технічних, екологічних і соціальних системах; теоретичні та прикладні проблеми інтелектуальних систем підтримки прийняття рішень; проблемно і функціонально орієнтовані комп'ютерні системи та мережі; методи оптимізації, оптимальне управління і теорія ігор; математичні методи, моделі, проблеми і технології дослідження складних систем; методи аналізу та управління системами в умовах ризику і невизначеності; евристичні методи та алгоритми в системному аналізі та управлінні; нові методи в системному аналізі, інформатиці та теорії прийняття рішень; науково-методичні проблеми в освіті.

Головний редактор журналу — ректор Національного технічного університету України «КПІ», директор Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, академік НАН України Михайло Захарович ЗГУРОВСЬКИЙ.

Журнал «Системні дослідження та інформаційні технології» включено до переліку фахових видань ВАК України.

Статті публікуються українською, російською та англійською мовами.

Журнал можна передплатити. **Наш індекс 23918.** Якщо ви не встигли передплатити журнал, його можна придбати безпосередньо в редакції. Наша адреса: 03056, м. Київ, пр. Перемоги, 37, корп. 35.

Завідуюча редакцією С.М. Шевченко
Науковий редактор А.О. Млечко
Молодший редактор В.М. Литвин
Комп'ютерна верстка та набір А.А. Патіоха

Підписано до друку 21.06.2010. Формат 70x108/16.
Папір офс. Офс. друк. Ум. друк. арк. 14. Ціна договірна.
Свідоцтво про реєстрацію КВ № 13804–2778 ПР від 17.03.2008 р.

Видавництво та поліграфія ВПК «Політехніка»
ВПН НТУУ «КПІ»
03056, Київ–56, пр. Перемоги, 37, корп. 15

Dear Readers!

The Educational-scientific complex «Institute for applied system analysis» at the National technical university of Ukraine «KPI» of MES and NAS of Ukraine is published of the international scientific and technical journal

«SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES».

The Journal is printing works of a theoretical and applied character on a wide spectrum of problems, connected with system researches and information technologies.

The main thematic sections of the Journal are the following:

Theoretical and applied problems and methods of system analysis; theoretical and applied problems of computer science; automated control systems; progressive information technologies, high-efficiency computer systems; decision making and control in economic, technical, ecological and social systems; theoretical and applied problems of intellectual systems for decision making support; problem- and function-oriented computer systems and networks; methods of optimization, optimum control and theory of games; mathematical methods, models, problems and technologies for complex systems research; methods of system analysis and control in conditions of risk and uncertainty; heuristic methods and algorithms in system analysis and control; new methods in system analysis, computer science and theory of decision making; scientific and methodical problems in education.

The editor-in-chief of the Journal is Rector of the National Technical University of Ukraine «KPI», Director of the Educational-scientific complex «Institute for applied system analysis» at the National technical university of Ukraine «KPI» of MES and NAS of Ukraine, academician of the NASU Michael Zaharovich Zgurovsky.

The articles to be published in the Journal in Ukrainian, Russian and English languages are accepted. Information printed in the Journal is included in the Catalogue of periodicals of Ukraine.

СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

2 • 2010

ЗМІСТ

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, ВИСОКОПРОДУКТИВНІ КОМП'ЮТЕРНІ СИСТЕМИ	
<i>Згуровський М.З., Петренко А.І.</i> Е-наука на шляху до семантичного Грід. Частина 2: Семантичний Web- і семантичний Грід	7
<i>Шпортько О.В.</i> Стиснення RGB-зображень без втрат із використанням палітри	26
ПРОБЛЕМИ ПРИЙНЯТТЯ РІШЕНЬ І УПРАВЛІННЯ В ЕКОНОМІЧНИХ, ТЕХНІЧНИХ, ЕКОЛОГІЧНИХ І СОЦІАЛЬНИХ СИСТЕМАХ	
<i>Зайченко Ю.П.</i> Оценка кредитных банковских рисков с использованием не- четкой логики	37
МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ, ПРОБЛЕМИ І ТЕХНОЛОГІЇ ДОСЛІДЖЕННЯ СКЛАДНИХ СИСТЕМ	
<i>Федин С.С.</i> Прогнозирование и вероятностная оценка долговечности техни- ческих объектов	55
<i>Мальшевская Е.Н.</i> Анализ использования нейросетей для диагностики рака шейки матки по мультиспектральному изображению	64
МЕТОДИ АНАЛІЗУ ТА УПРАВЛІННЯ СИСТЕМАМИ В УМОВАХ РИЗИКУ І НЕВИЗНАЧЕНОСТІ	
<i>Крпункина Н.В., Проничкин С.В., Рыков А.С.</i> Формирование подгрупп экс- пертов с согласованными мнениями и определение групповой обобщен- ной оценки многопризнаковых объектов	72
<i>Филатов А.Г.</i> Синтез регулятора состояния для дискретной во времени сис- темы модального управления	80
<i>Zhuk S.</i> Minimax recursive state estimation for linear discrete-time descriptor systems	94
ЕВРИСТИЧНІ МЕТОДИ ТА АЛГОРИТМИ В СИСТЕМНОМУ АНАЛІЗІ ТА УПРАВЛІННІ	
<i>Песчанский А.И.</i> Календарное техническое обслуживание простой системы с учетом минимального аварийного восстановления	106
<i>Угрюмова Е.М., Трончук А.А., Афанасьевская В.Е.</i> Информационная техно- логия формирования облика элементов технических систем на основе дискретных данных об аналогах	118
НОВІ МЕТОДИ В СИСТЕМНОМУ АНАЛІЗІ, ІНФОРМАТИЦІ ТА ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ	
<i>Сявавко М.С., Пасічник Т.В., Тимків В.П.</i> Нечітка арифметика в задачах між- галузевого балансу	134
Реферати	144
Відомості про авторів	149

SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES

2 • 2010

CONTENTS

PROGRESSIVE INFORMATION TECHNOLOGIES, HIGH-EFFICIENCY COMPUTER SYSTEMS	
<i>Zgurovsky M.Z., Petrenko A.I.</i> E-Science on the way to semantic Grid. Part 2: Semantic Web- and semantic Grid	7
<i>Shportko A.V.</i> Loss-free compression of RGB-images using a palette	26
DECISION MAKING AND CONTROL IN ECONOMIC, TECHNICAL, ECOLOGICAL AND SOCIAL SYSTEMS	
<i>Zaychenko Yu.P.</i> Credit risk estimation on the base of fuzzy logic	37
MATHEMATICAL METHODS, MODELS, PROBLEMS AND TECHNOLOGIES FOR COMPLEX SYSTEMS RESEARCH	
<i>Fedin S.S.</i> Forecast and probabilistic estimations of technical object longevity	55
<i>Malyshevska K.M.</i> Analysis of using neural networks for cancer diagnosis from multisensor imagery	64
METHODS OF SYSTEM ANALYSIS AND CONTROL IN CONDITONS OF RISK AND UNCERTAINTY	
<i>Krapuhina N.V., Pronichkin S.V., Rykov A.S.</i> Formation of expert subgroups based on consensus and deriving generalized estimate of multiattribute objects	72
<i>Filatov O.G.</i> Design of state regulator for discrete-time system of modal control	80
<i>Zhuk S.</i> Minimax recursive state estimation for linear discrete-time descriptor systems	94
HEURISTIC METHODS AND ALGORITHMS IN SYSTEM ANALYSIS AND CONTROL	
<i>Peschansky A.I.</i> Calendar maintenance of a single-component system with minimal fallback renewal	106
<i>Ugryumova E.M., Tronchuk A.A., Afanasjevska V.E.</i> Information technology for generation of engineering system elements on the basis of discrete data on analogs	118
NEW METHODS IN SYSTEM ANALYSIS, COMPUTER SCIENCE AND THEORY OF DECISION MAKING	
<i>Siavavko M.S., Pasichnyk T.V., Tymkiv V.P.</i> Fuzzy arithmetic in problems of in- terindustry balance	134
References	144
Information about the authors	149

**Е-НАУКА НА ШЛЯХУ ДО СЕМАНТИЧНОГО ГРІД.
ЧАСТИНА 2: СЕМАНТИЧНИЙ WEB- І СЕМАНТИЧНИЙ ГРІД**

М.З. ЗГУРОВСЬКИЙ, А.І. ПЕТРЕНКО

Досліджено семантичний Web, заснований на метаданих, онтологіях і програмних агентах як фундамент для семантичного Грід, в якому інформація і послуги мають чітке визначення, що надає можливість комп'ютерам і людям працювати в кооперації зі знанням, а не з інформацією, як це робиться зараз. Розглянуто підходи до опису онтологій даних і Web- та Грід-сервісів (із станом і без стану), їхньої взаємодії, архітектури, а також й існуючі реалізації семантичних Грід.

ВСТУП

У першій частині даної роботи [1] було розглянуто еволюцію Грід-систем на шляху до семантичних Грід, в яких інформація та послуги дають можливість комп'ютерам і людям працювати в кооперації. Семантичний Грід як розширення сучасного Грід виявляється надзвичайно корисним для е-науки [2–6], оскільки його технології дають змогу легко, швидко та зручно працювати *зі знанням*, а не з інформацією, як то було дотепер. Додатки семантичних Грід можуть інтегрувати велику множину різнорідних джерел інформації і сервісів, яких залучено і об'єднано, а також людські й обчислювальні ресурси і сервіси передачі інформації.

1. СЕМАНТИЧНИЙ WEB

У даній частині головну увагу приділено семантичному Web як фундаменту семантичного Грід. Ідея семантичного Web — це концепція мережі, в якій кожен ресурс людською мовою забезпечено описом, зрозумілим комп'ютеру. Документи в мережі публікуються у форматі XML, який містить семантичні RDF-ствердження.

Згідно із планом, запропонованим Тімом Бернерсом-Лі, винахідником семантичного Web [7], для впровадження семантичного Web необхідно послідовно розробити:

- синтаксис для представлення знань, що використовує посилання на онтології (зроблено: RDF);
- мову опису онтологій (зроблено: OWL);

- мову опису Web-серверів (зроблено: WSDL, OWL-S);
- інструментарій розробки/читання документів семантичного Web (почато: Jena, Haystack, Protege);
- мову запитів до знань, записаних у RDF (зроблено: SPARQL);
- логічне виведення знань (повністю не зроблено);
- семантичну пошукову систему (почато: SHOE, Swoogle);
- агентів семантичного Web (поки що не зроблено).

У цьому переліку новим, у порівнянні з попереднім матеріалом [1], є поняття *онтології*, тому зупинимось на ньому детальніше. Онтологія визначає загальний словник для вчених, яким потрібно спільно використовувати інформацію в предметній галузі [7–10]. Вона включає формулювання основних понять предметної галузі й відношення між ними, що інтерпретуються комп'ютером.

Причини необхідності в розробці онтології:

- сумісне використання людьми або програмними агентами загального розуміння структури інформації;
- можливість повторного використання знань у предметній галузі;
- створення явних допущень у предметній галузі;
- відокремлення знань предметної галузі від оперативних знань;
- аналіз знань у предметній галузі.

Сумісне використання людьми або програмними агентами загального розуміння структури інформації є однією з найзагальніших цілей розробки онтологій [9]. Наприклад, нехай, декілька різних Web-сайтів містять інформацію з медицини або надають інформацію про платні медичні послуги, що сплачуються через Інтернет. Якщо ці Web-сайти спільно використовують і публікують одну і ту ж базову онтологію термінів, якими вони всі користуються, то комп'ютерні агенти можуть добувати інформацію з цих різних сайтів та накопичувати її. Агенти можуть використовувати накопичену інформацію для відповідей на запити користувачів або як вхідні дані для інших застосунків.

Забезпечення можливості використання знань у предметній галузі сприяло бурхливому розвитку у вивченні онтологій. Наприклад, для моделей багатьох різних предметних галузей необхідно сформулювати поняття часу, яке включає поняття часових інтервалів, моментів часу, відносних мір часу і т.ін. Якщо одна група вчених детально розробить таку онтологію, то інші можуть повторно використовувати її у своїх предметних галузях. Крім того, якщо нам необхідно створити велику онтологію, ми можемо інтегрувати декілька існуючих онтологій, які описують частини великої предметної галузі. Ми також можемо повторно використовувати основну онтологію, таку, як UNSPSC [9], і розширювати її для опису предметної галузі, що нас цікавить.

Створення явних допущень у предметній галузі, покладених в основу реалізації, надає можливість легко змінювати ці допущення при зміні наших знань про предметну галузь. Жорстке кодування припущень про світ мовою програмування призводить до того, що ці припущення не тільки складно знайти і зрозуміти, але й складно змінити, особливо непрограмістові. Крім

того, явні специфікації знань у предметній галузі корисні для нових користувачів, які мають засвоїти значення термінів з цієї галузі.

Відокремлення знань предметної галузі від оперативних знань — це ще один варіант загального застосування онтологій. Ми можемо описати завдання конфігурації об'єкту і його компонентів відповідно до необхідної специфікації і запровадити програму, яка формує цю конфігурацію незалежно від об'єкту і самих компонентів. Після цього ми можемо розробити онтології компонентів і характеристик комп'ютерів, на яких базується об'єкт, та застосувати цю програму для морфологічного синтезу нестандартних конфігурацій об'єктів.

Аналіз знань у предметній галузі можливий, якщо є декларативна специфікація термінів. Формальний аналіз термінів надзвичайно цінний як під час спроби повторного використання існуючих онтологій, так і при їх розширенні. Часто онтологія предметної галузі сама по собі не є метою. Розробка онтологій схожа на визначення набору даних і їх структури для використання іншими програмами. Методи розв'язання завдань, доменно-незалежні застосування і програмні агенти використовують як дані онтології і бази знань, побудовані на основі цих онтологій [10].

Для кожної предметної галузі існує безліч онтологій залежно від ступеня деталізації її структури. Прийнято говорити про *спектр онтологій*, який містить:

- *онтології малої структуризації* — тахономії (наприклад, ієрархія Yahoo, біологічна тахономія), схеми баз даних і схеми метаданих (ebXML, WSDL);
- *онтології середньої структуризації* — тезауруси (WordNet, CALL, DTIC) і концептуальні моделі (Моделі OO, UML);
- *онтології високої структуризації* — логічні концептуальні моделі (TOVE, CYC, семантичний Web).

Приклад фрагменту структури онтології з Грід-галузі наведено на рис. 1.

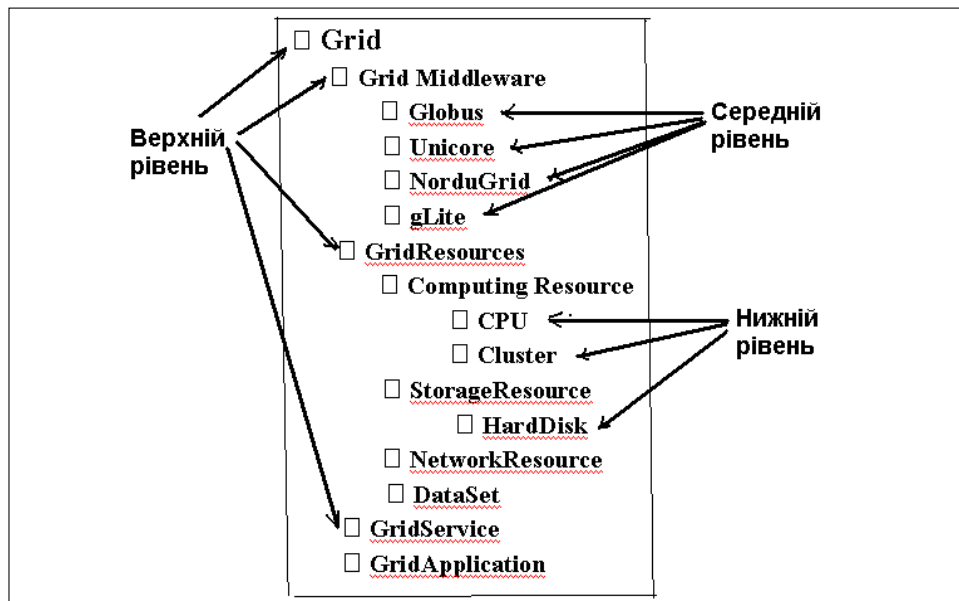


Рис. 1. Фрагмент онтології з Грід-галузі

Онтологія — це формальний явний опис понять у предметній галузі: *класів* (або *концептів*), особливостей кожного концепту (*слотів* або властивостей) і обмежень на слоти (*фасетів*, або аспектів чи ролей). Онтологія разом із набором індивідуальних зразків класів складає *базу знань*. Проте, слід пам'ятати — **для будь-якої предметної галузі не існує єдиної правильної онтології**. Проектування онтології — це творчий процес і дві онтології, розроблені різними людьми, ніколи не будуть однаковими. Потенційні додатки онтології, а також розуміння розробником предметної галузі і його точка зору на неї будуть, поза сумнівом, впливати на ухвалення рішень.

МОВИ ОПИСУ ОНТОЛОГІЙ

Для опису онтологій існують різні мови і системи, проте, найперспективнішою є мова OWL (попередня назва DALM-OIL) [11]. OWL розширює можливості з опису нових типів RDF-даних, що вже існують (наприклад, можна визначати тип класу, що є перетинанням або об'єднанням двох існуючих). OWL використовує синтаксис XML і містить команди для завдання дерева класів, команди для завдання приналежності індивідів класам, команди опису властивостей: області визначення, області значень та ін., команди опису характеристик властивостей: симетричність, транзитивність, функціональність; команди для склеювання (еквівалентності) класів.

Однією з переваг OWL-онтологій є доступність інструментів, які можуть робити *логічні висновки*. Побудова чіткої і працездатної системи логічних висновків — непросте завдання, доступніше побудувати онтологію. На сьогодні існують приклади онтологій (Protege-2000 [12], Ontolingua [13] і Chimaera [14]), побудовані у сферах традиційно науковомістких галузей, починаючи від хімічної обробки, і закінчуючи конструюванням машинобудівних підприємств. Крім того, для специфічних наукових галузей існує низка ініціатив щодо побудови великомасштабної онтології. Однією з таких галузей є *генетика*, де багато зусиль було спрямовано на створення спільної термінології та визначень, щоб дозволити вченим керувати їхніми знаннями [8]. Ці зусилля дають уявлення про те, як онтологію можуть відігравати значну роль у підтримці науковців.

Технологія SPARQL дозволяє отримувати дані з розподілених джерел і може бути як засіб інтеграції різнорідної інформації. У специфікації SPARQL відсутні недоліки, властиві традиційним мовам запитів, зокрема, не накладаються обмеження на формат даних. Завдяки цьому стає можлива взаємодія між ресурсами різного типу. Намагатися використовувати семантичну мережу без SPARQL — це все одно, що працювати з реляційною базою даних без мови структурованих запитів SQL. Тобто, SPARQL перетворює доступ до даних у деяку подібність Web-сервісу.

Кінцева амбітна мета семантичного Web полягає у створенні такого середовища, де програмні агенти можуть динамічно виявляти і опитувати ресурси, а потім взаємодіяти з ними [15]. Агенти мають справлятися з віртуальними проблемами, які виникають в інтелектуальному середовищі, виявляти нові факти і виконувати найрізноманітніші завдання, що отримуються від людини. Семантичний Web можна представити як трирівневу структуру:

Перший рівень: фундамент, що складається з унікальної глобальної ідентифікації ресурсу; метадані для підтвердження відомостей про ресурси і обслуговування запитів на такі відомості, і простої мови опису метаданих і знань.

Метадані і знання містяться в онтологіях, що забезпечують взаєморозуміння і загальноприйнятий словник для метаданих і правил виведення нових метаданих і знань.

Другий рівень: базові сервіси для виконання стандартних висновків і цілеспрямованого перегляду метаданих і онтологій, пояснення зроблених висновків.

До цього рівня відноситься також управління довірою, агенти, пошукові машини, сервери онтологій.

Третій рівень: прикладні сервіси; наприклад, сервіс певного агента (агент науковця, агент лікаря).

Це є ідеєю, яку поки що не реалізовано, але щоб здійснити щось глобальне, необхідно почати з малого. Семантичний Web — це Web для комп'ютерів. Повноцінна реалізація ідеї якого дозволила б не лише здійснювати найефективніший комплексний пошук за допомогою інтелектуальних програмних агентів, а й знаходити приховані закономірності у масштабних розподілених наборах даних, допомагати відслідковувати активність організованої злочинності та терористів, сприяти аналізу розподілених геологічних та метеорологічних даних, допомагати у розподілі та повторному використанні важливих освітніх ресурсів тощо.

Традиційні Web-сервіси описуються, використовуючи мову WSDL на синтаксичному рівні, як колекції імен і схем XML-даних. Ці описи збагачуються згодом семантичною інформацією. Виділяють чотири галузі, де можуть використовуватися онтології для семантичного опису Web-сервісів:

- *семантика даних* для визначення значення даних, тобто вхідних і вихідних даних;
- *функціональна семантика* для визначення значення операцій і процесів, за допомогою яких вхідні дані перетворюються на вихідні;
- *семантика QoS* забезпечує рівень *якості сервісів*: наявність, рівень довіри, ціну і т.ін. Ці параметри впливають на пошук (вибір) сервісу;
- *семантика виконання* забезпечує наступні параметри: початкові стани, ефекти, шаблони взаємодії з сервісами.

Співвідношення між параметрами Web-сервісів і концептів у деякій OWL-онтології можуть бути виражені у формальний засіб через анотацію WSDL-опису з посиланнями на URL в онтології або зовнішнім описом. Існують інструменти, що стандартизують такі семантичні анотації Web-сервісів, а саме OWL-S (OWL for Services), запропонований коаліцією OWL-S [16, 17], і WSDL-S, запропонований в проекті METEOR-S [18].

Структура OWL-S онтології складається з *профілю сервісу (Profile)* для залучення і пізнання сервісів, процесуального *процесу (Process)*, який підтримує побудову сервісів, і *розташування сервісу (Grounding)*, що зв'язує концепторів профілю і процесу з розташованим нижче інтерфейсом сервісу (рис. 2). Крім того, у структуру OWL-S входять *складання (Compose process)* і *складові (Atomic Process)*, які відповідають за реалізацію складеного серві-

су. Профіль сервісу (OWL-S профіль) має функціональні і нефункціональні властивості. Функціональні властивості описують входи, виходи, початкові стани й ефекти сервісу. Нефункціональні — напівструктуровану інформацію, призначену для людини-користувача (наприклад: ім'я сервісу, опис сервісу та параметри сервісу). Параметри сервісу включають наступні вимоги: безпеку, якість сервісу, географічні межі і т.ін. Реалізація сервісу забезпечує виконання конкретного Web-сервісу шляхом перекладу абстрактних концепцій профілю OWL-S і процесу в конкретні повідомлення.

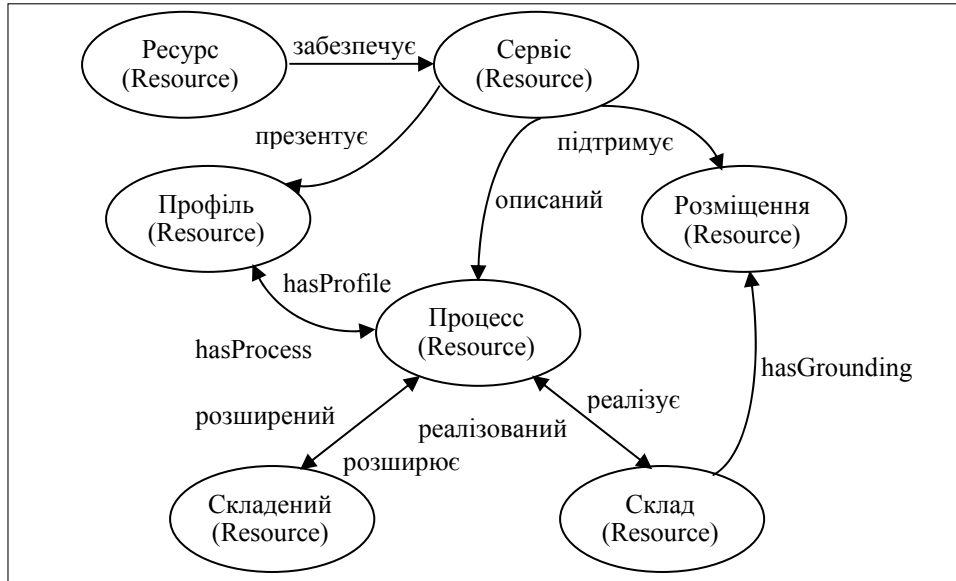


Рис. 2. Структура сервісної онтології згідно концепції OWL-S

Завдання, які вирішуються за допомогою OWL-S, включають:

- автоматичний пошук семантичного Web-сервісу;
- автоматичний виклик семантичного Web-сервісу і завантаження його параметрів;
- автоматичну композицію складу семантичного Web-сервісу і його взаємодії з іншими сервісами;
- автоматичний контроль виконання семантичного Web-сервісу.

У свою чергу, семантичні Web-сервіси можуть використовуватись з метою:

- класифікації обчислювальних ресурсів і ресурсів даних, метрик виконання, управління роботами;
- інтеграції схем описів даних і завдань та їх завантаження;
- роздрукування даних і сервісних повідомлень введення/виведення;
- вибору методів розв'язання задач і використання інтелектуальних порталів;
- інфраструктури для аутентифікації, реєстрації й управління доступом.

Базові сервіси семантичного Web можуть стати базовими Грід-сервісами. Остов семантичного Web є тим засобом, за допомогою якого в Грід можуть бути представлені метадані як відносно його інфраструктури,

так і відносно Грід-додатків. Симбіоз Грід і семантичного Web достатньо очевидний, перспективи їхнього розвитку зв'язані. Від семантичного Web Грід запозичує остов і сервіси для роботи зі своєю семантикою. Сам семантичний Web стимулюється могутнім розвитком Грід-додатків й імітує інфраструктуру Грід.

Необхідно зазначити, що й досі концепція розвитку семантичного Web не має повноцінної реалізації. Серед причин цього можна виокремити наступні:

- *Відсутність мотивації розробників.*

Необхідною умовою впровадження семантичного Web є наявність критичної кількості Web-ресурсів із коректними метаданими у відповідних форматах. Створення таких метаданих покладено на розробників Web-сайтів і є відносно складним завданням, яке не буде виправдано (зокрема фінансово) доти, доки семантичний Web не почне працювати на повну потужність.

- *Неможливість гарантувати достовірність метаданих.*

Семантичний Web надає широкі можливості для опису даних та зв'язків між ними.

- *Засоби семантичного Web є відносно складними.*

Із двох приблизно однаково ефективних підходів перевагу завжди отримує простіший. Прикладом цього є те, що формат RSS 1.0, заснований на RDF, є значно менш популярним, ніж простіший за нього RSS 2.0.

Очікується, що протягом найближчих 3–5 років ситуація докорінно зміниться. Вже функціонує програма Linking Open Data on the Semantic Web, яка збирається взагалі всю відкриту інформацію в Інтернеті перевести в RDF-формат.

СЕМАНТИЧНИЙ ГРІД

Семантичний Грід, який базується на семантичному Web, можна розглядати як спрямовану на сервіси архітектуру, в межах якої об'єкти в певному середовищі обмінюються послугами (сервісами) один із одним на певних умовах. У межах цього середовища можливе встановлення правил взаємодії об'єктів, наприклад, отримання та надання сервісів. Природно представляти власників сервісів та споживачів сервісів як автономних агентів [19].

При цьому агенти мають відповідати наступним вимогам:

- чітко ідентифікувати об'єкти, що вирішують певні задачі та мають чіткі межі й інтерфейси;
- одержувати повідомлення, пов'язані з обладнанням через датчики, і впливати на обладнання, в яке вони вбудовані, через виконавчі елементи;
- виконувати певну процедуру — у них є специфічні цілі (мета);
- бути автономними і здійснювати контроль за своїм внутрішнім станом та поведінкою;
- бути здатними гнучко розв'язувати завдання, для яких були спроектовані, тобто — реагувати своєчасно на зміни, які відбуваються в середовищі, і прораховувати свої дії з випередженням.

Обчислення, засноване на агентах, задовольняє умови динамічно змінного середовища, де автономія агентів дозволяє адаптувати обчислення до

змінних обставин [20]. Це є важливою властивістю для Грід-систем третього покоління. Однією з методик для досягнення зазначеної властивості є оперативний обмін інформацією між агентами.

Для реалізації агентної платформи проведено відповідні дослідження при допущенні, що у кожного власника сервісу будуть один або більше агентів, які діють від його особи. Ці агенти будуть керувати доступом до сервісів, за які вони відповідають, і гарантуватимуть виконання всіх погоджених правил (договорів). Ця діяльність залучає планування місцевих дій відповідно до доступних ресурсів та передбачає, що відповідні послуги буде надано вчасно й відповідно до контракту. Агенти будуть також діяти від імені користувачів сервісів. Залежно від бажаного ступеня автоматизації, це може бути прийняття договорів для отримання сервісу і одержання та представлення будь-яких отриманих результатів.

Відносини між власником і споживачем сервісу засновано на договорі надання сервісів. Цей договір визначає умови користування, за якими власник погоджується надати сервіс споживачу. Точна структура контракту буде залежати від природи обслуговування й відносин між власником і споживачем. Але, відповідні договори включають ціну за виклик обслуговування, інформацію, що споживач повинен надати власнику, очікуваний результат обслуговування, індикацію, коли цей результат може очікуватися, і штраф за невиконання контракту. Такі договори можуть укладатися як off-line, так і on-line, залежно від контексту.

Власники й виробники сервісів взаємодіють один із одним у специфічному контекстному середовищі. Воно може бути поширене на всі об'єкти Грід (мається на увазі, що всі об'єкти пропонують своє обслуговування на повністю відкритому ринку). В інших випадках це може бути закрите середовище, входом до якого можна керувати (об'єкти формують приватний клуб). Надалі це специфічне середовище буде називатись *ринком*, а об'єкт, що встановлює ринок — *власником ринку*.

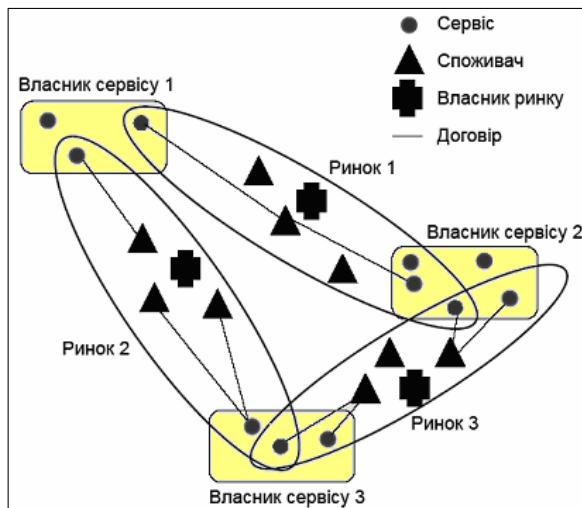


Рис. 3. Основні компоненти сервісної архітектури: округлені прямокутники — власники сервісу; заповнені кола — сервіси; заповнені трикутники — споживачі; лінії між виробниками й споживачами — договори; овали — ринок; заповнений хрест — власник ринку

На рис. 3 наведено сервісну архітектуру семантичного Грід, де власники пропонують сервіси споживачам відповідно до договорів. Кожна взаємодія власника та споживача, що має місце на цьому ринку, встановлюється власником ринку відповідно до правил. Власником ринку може бути один із об'єктів на ринку (виробник або споживач), або це може бути і нейтральна третя особа.

Сукупність ринків, кожен із яких відповідає за різні аспекти е-наукових досліджень, складає спільний науково-дослідницький

простір країни чи континенту. Дані про ринки зареєстровано у відповідній міжнародній базі даних. У цієї бази даних є власний сервіс повідомлень, користуючись яким, окремі вчені можуть реєструвати свою зацікавленість щодо певних специфічних сервісів. Коли нові сервіси створюються та з'являються на ринках, учені інформуються про це через своїх агентів.

Приклад можливої архітектури «розумної» Грід-системи, в якій поєднано програмні агенти і традиційній OGSA-підхід, приведено на рис. 4 [21]. У такій архітектурі агенти забезпечують інтелектуальні розподілені сервіси, які включають комунікацію і прийняття рішень.

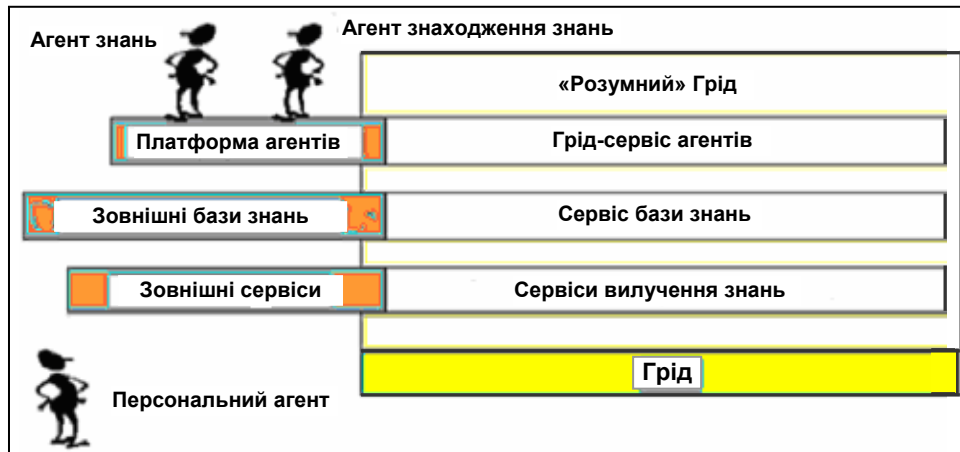


Рис. 4. Архітектура «розумного» Грід із програмними агентами

Грід-сервіс агентів відіграє головну роль при взаємодії агента і Грід-сервісу, і тому він — посередник між *платформою агентів* та іншими Грід-сервісами. Цей сервіс перетворює повідомлення агентів у точні дії й перетворює результати цих дій зворотньо в повідомлення агентів. Цей сервіс тісно співпрацює із *сервісом бази знань*, яка містить дані і їх семантику, а також інформацію про всі сервіси і ресурси, що можуть бути використані для пошуку знання.

Сервіси вилучення знань здатні виконувати пошук знань у базах даних, об'єднаних у Грід-інфраструктурі, використовуючи існуючий арсенал ефективних методів Data Mining [6].

Зовнішні сервіси розширюють систему «розумного» Грід даними, отриманими за допомогою сервісів за межами Грід-інфраструктури, наприклад, Web-сервісами.

Агент знань (A3) представляє «розумний» Грід для зовнішнього світу під час комунікації з іншими агентами. З точки зору програмного забезпечення і зовнішнього світу — це інтелектуальний агент програмного забезпечення, але з точки зору «розумного» Грід — це специфічний Грід-додаток. Фактично, цей агент працює вхідним і вихідним інтерфейсом «розумного» Грід і, тому, він — посередник (медіатор) між запитами щодо знань й інформаційними ресурсами, з іншого боку. Цей агент створено і керується Грід-сервісом агентів і може взаємодіяти з іншими агентами за допомогою повідомлень у форматі ACL [22].

Агент знаходження знання (A33) використовується «розумним» Грід, для того, щоб знайти відповідні сайти, у випадку, коли інформація, яку за-

питують, не знаходиться в *базі знань*, але її розташування відоме. Цей агент зареєстровано в Платформі агентів і керується *Грід-сервісом агентів*, який надає агентові інструкції, як запросити бажані дані, й інформацію про розташування даних. Агент може спілкуватися з іншими агентами, а також із семантичним Web, який є іншим джерелом інформації і джерелом знання всієї системи.

Персональний агент (ПА). Користувач, який запитує про знання, генерує ПА. Цей агент може адресувати питання до «розумного» Грід і передавати відповіді користувачам. Успіх цієї діяльності залежить від порозуміння агентів ПА і АЗ, а також від конструкції запиту. ПА не є частиною системи, тому що він створюється користувачем і не реєструється в Платформі агентів. Цей агент може запросити у АЗ список доменів, про які він має знання, або ПА може інформувати АЗ про онтологію, яка описує його домен, і об'єднати її разом із онтологіями АЗ, щоб обрати один загальний домен.

Розглянуті задачі можна вирішити за допомогою мови ACLs (Agent Communication Languages — мови комунікації агентів). Зокрема, консорціум FIPA (Foundation for Intelligent Physical Agents) створює програмні стандарти для різнорідних і взаємодіючих агентів і систем, заснованих на агентах. В абстрактній архітектурі FIPA [22]:

- агенти спілкуються, обмінюючись повідомленнями, які відображають динаміку змін і закодовані мовою комунікації агента;
- сервіси надають підтримку агентам, включаючи також сервіси каталогів і сервіси транспортування повідомлень;
- сервіси можуть бути здійснені у вигляді агентів або у вигляді програмного забезпечення, до якого можна звернутися через мови програмування інтерфейсів (наприклад, Java, C++ або мови опису інтерфейсів).

База знань (БЗ) містить специфічні дані про реальні об'єкти і відносини між цими об'єктами і їх властивостями. Вона складається з двох частин: перша пов'язана з добре розробленою *онтологією*, а друга — зі зразками цієї онтології.

Оскільки багатоагентну інтелектуальну платформу ще не створено, семантичний Грід сьогодні можна базувати на вживаному проміжному програмному забезпечення (middleware), додавши до базових сервісів, характерних для Грід-систем, наступні семантичні сервіси:

- *метаданих*: збирання інформації з різних джерел та розміщення її в одному місці;
- *анотацій*: відслідкування, фільтрування, аналіз та пошук сервісів, публікація інформації про них;
- *онтологій* об'єктів та ресурсів, де онтологія — це засіб опису семантики проблемної області за допомогою словника і підібраної специфікації існуючих у ній відношень та обмежень, які забезпечують інтеграцію словника;
- *композиції*: автоматична композиція сервісу як процесу відкриття, інтегрування та виконання набору пов'язаних сервісів у належній послідовності для формування змістовного та комплексного сервісу відповідно до підходів *оркестровки* та *хореографії*. При цьому оркестровка веде до централізованої архітектури, у якій механізм оркестровки контролює виконання

завдання, а хореографія ґрунтується на співробітницькому підході, в якому кожна група процедур відіграє свою роль, виконуючи дії, задані лише для неї.

При цьому, документи в мережі публікуються в форматі XML, який містить семантичні RDF-твердження. Відношення, аксіоми та твердження про об'єкти зберігаються в онтологіях, що реалізуються мовою OWL. Запросити дані з RDF можливо за допомогою SPARQL.

У результаті архітектура OGSA перетворюється в архітектуру S-OGSA (рис. 5) [23, 24].

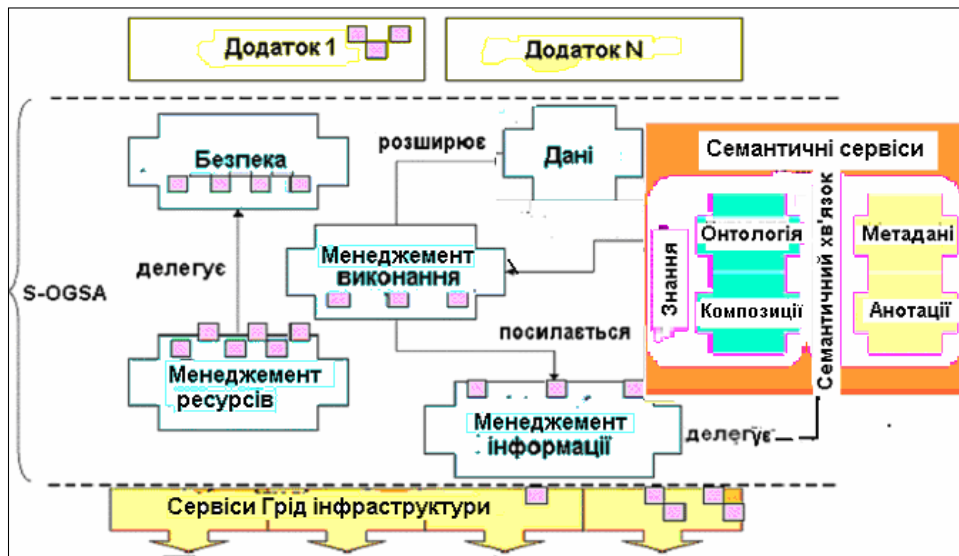


Рис. 5. Від OGSA до S-OGSA

Підхід S-OGSA намагається розширити архітектуру OGSA і забезпечити повне використання семантики разом із використанням сервісу, визначеного як *сервіс знань*. Вказаний підхід намагається реалізувати семантичний Грід методом «зверху до низу», створюючи архітектури, які мають задовольнити велику кількість додатків і вимог. Кожен Грід-сервіс переважно використовує метадані. Гнучка композиція Грід-сервісів вимагає наявності інформації про функціональність, доступність та інтерфейси самих різних сервісів. Виявлення сервісів і брокерське обслуговування використовують описи на рівні метаданих. Композиція сервісів керується і підтримується описами метаданих. Тобто, повне пов'язання Грід-сервісів із *метаданими є ключовим аспектом* архітектури S-OGSA.

На мові OWL-S можливості сервісу описуються відповідним *IOPE* (вхід, вихід, попередні стани й ефекти) [1]. Такі описи можуть бути взяті з WSDL-описів сервісу і є достатніми для сервісів, що не мають стану, сервіси зі станом, тобто WS-ресурси, складаються з сервісу і ресурсу зі станом. Властивості ресурсу (BP), як визначено в WSRF-специфікації, можуть бути динамічними. Вони можуть бути створені або знищені у часі по ходу процесу.

У процесі виконання сервісів, що мають стан, можливе використання спадкоємства властивостей ресурсу. Хоча точна ієрархія властивостей ресурсу може допомогти в генеруванні семантики, не існує стандарту, який описував би детально, як BP спадкоємство має бути реалізовано. Це може

викликати серйозні складнощі в синтаксичному аналізі сервісу і необхідно буде представити спеціальну документацію, щоб розібрати ВР-ієрархію. До того ж, властивості ресурсу можуть бути пристосовані тільки для моделювання відповідних входів і виходів сервісу, тобто сервіс приймає лише спеціальні завдання і може презентувати входи і виходи, як властивості завдання, ховаючи, таким чином, введення/виведення у властивостях ресурсів сервісу. Це необхідно враховувати при складанні сервісу.

Окрім складнощів у процесі пошуку сервісу, існує невелика різниця в процесі активізації сервісу. WS-ресурс складається з сервісу і ресурсу зі станом, тобто його визначено як ЕРР (end point reference), який описує не тільки адресу сервісу, але й ідентифікацію ресурсу. Ідентифікація сервісу в реалізації OWL-S має бути розширена в складнішу структуру. Для Грід-сервісів таке розширення надає можливість отримати декілька варіантів одного сервісу, розміщених на різних серверах. Проте, це можна зробити простіше шляхом генерації кількох OWL-S для сервісу.

У роботі [25] розроблено інструментарій для автоматичної генерації описів OWL-S Грід-сервісів, що мають стани, і для тих, що їх не мають, з відповідних описів Web-сервісів (WSDL).

Переклад починається з аналізу конфігурації і URL WSDL-документа. Перекладач аналізує WSDL-документ, відмічаючи операції типу портів, входів, виходів і властивості ресурсу. Потім перекладач генерує для кожної WSDL-операції скелет OWL-S документа, а також створює входи, виходи, початкові стани й ефекти та розв'язує елементи в онтологічні концептори, визначені в конфігурації. Якщо необхідно, він може створити онтологію, яка моделює властивості ресурсу даного сервісу.

Приклад перекладу WSDL-опису метеорологічного сервісу показано на рис. 5. Оскільки кожен сервіс має декілька операцій, семантичний опис генерується для кожної операції з метою дати можливість створювати робочі потоки сервісних операцій. Окрім WSDL-опису процес перекладу також вимагає знання конфігурації, яка описує відповідність введення/виведення WSDL онтологічним концепторам домену. Ці концептори можуть описувати інформацію про сервіс (наприклад, ім'я сервісу, провайдера), а також складні входи і виходи сервісів, такі, як географічне розташування, географічну інформацію, басейни річок і т.ін.). Далі вони зможуть ідентифікувати форми HTML, які представляються користувачеві, якщо необхідні додаткові вхідні дані.

Подібне перетворення описів WSDL в OWL-S здійснюється також у семантичному Грід, описаному в роботі [26].

Розглянуто підхід перетворення описів сервісів притаманний методології реалізації семантичного Грід «з низу до верху», яка намагається витягнути якомога більше користі з існуючих семантичних Web-сервісів. Існує також альтернативна можливість побудови семантичного Грід «з низу до верху», базуючись на підтримці OGSA-DAI (Open Grid Service Architecture — Data Access Interfaces) і створенні нового програмного забезпечення проміжного шару InteliGrid [27].

WSDL опис

```
<wsdl:portType name=«MM5ServicePortType»
  wsrp: ResourceProperties=«tns:MM5Properties»>
  <wsdl:operation name=«configureFromProperties»>
  <wsdl:input message=«tns:ConfigureInputMessage»/>
  <wsdl:output message=«types:VokiOutputMessage»/>
  </wsdl:operation>
  . . .
  <wsdl:service name=«MM5Service».>
  <wsdl:port name=«MM5ServicePortTypePort»>
  binding=«binding: MM5ServicePortTypeSOAPBinding»>
  <soap:address
  location=«http://localhost:8080/wsrt/setvices/»/>
```

Конфігурація

```
MM5Service:configureFromProperties.properties=http://gom.kw
fgrid.net/gom/ontology/DomainApplicationOntology/
  FFSC#MM5Properties
  MM5Service.configureFromProperties.voidResponse=
  Effect=
  http://gom.kwfgrid.net/gom/ontology/DomainServiceOntology/
  FFSC#MM5isConfigured
```

OWL-S опис

```
<service:Service
  rdf:ID=«configureFromProperties_MM5Service»>
  <service:presents>
  <profile:Profile
  rdf:ID=«configureFromProperties_MM5Profile»/>
  </service:presents>
  </service:describedBy>
  <process:AtomicProcess edf:ID=
  «configureFromProperties_MM5AtomicProcess»/>
  </service:describedBy>
  <service:supports>
  <grounding:WsdLGrounding
  <grounding:WsdLGrounding rdf:ID=
  «configureFromProperties_MM5Grounding»/>
  </service:supports>
  </service:Service>
  . . .
  <profile:Profile
  rdf:about=«configureFromProperties_MM5Profile»>
  <profile:hasInput
  <process:Input rdf:ID=«properties»>
  <process:parameterType>= http://gom.kwfgrid.net/gom/
  ontology/DomainApplicationOntology/FFSC#MM5Properties
  </process:parameterType>
  </process:Input>
```

Рис. 6. Приклад перекладу WSDL-опису метеорологічного сервісу в OWL-S опис

Якщо на сьогодні існують сотні успішних Грід-проектів із різних галузей науки і техніки для традиційного Грід-середовища [6], то проектів семантичного Грід поки що дуже мало. Інформація про існуючі пілотні розроблення семантичних Грід наведено у таблиці. Слід зазначити, що їх присвячено вирішенню насущних викликів людству, і значну частину з них пов'язано з проектами 6-ї Рамочної програми (FP-5). На сьогодні 7-й Рамочній Програмі подібних проектів значно більше. У найближчому майбутньому семантичні Грід використовуватимуться в складних медичних аналізах, при наркотичному контролі, фінансовому аналізі ризиків, кризовому управлінні, цифрових ринках та розробці виробів.

Таблиця. Пілотні проекти з семантичного Грід

	Назва	Опис	Призначення
1.	K-Wf Grid [25]	Розроблюється методологія компонування сервісів і додатків, підтримки паралельних обчислень, програмування Грід через визначення міжкомпонентних залежностей. Кожен компонент має декілька можливих моделей обчислень представлених у Web або Грід-сервісах, наприклад: метеорологічні методи (Aladin, MM5); гідрологічні методи (HSPF, NLC і т.ін.). Система містить автоматичний перекладач WSDL-опису сервісу в OWL-S	K-Wf Grid використовують для передбачення повеней
2.	InteliGrid [27]	Проект InteliGrid присвячено об'єднанню в Грід-системі семантичної сумісності віртуальних організацій і Грід-технологій для потреб промислових організацій у співпраці при використанні виробничої інфраструктури і комунікацій, обробці інформації. Було розроблено спеціальну платформу з різними сервісами як послугами управління документообігом. Платформа забезпечує доступ до серверів із промисловими зразками, а також забезпечує використання високопродуктивної обчислювальної інфраструктури	Архітектура InteliGrid має задовольняти велику кількість додатків і вимог
3.	myGrid [28, 8]	Адаптивна інфраструктура, яку заснувано на поняттях Грід-систем, семантичного Web і онтологій, призначена для потреб біоінформаційних і біомедичних застосувань. Вона містить сервіси та реєстратори даних, засновані на RDF; використовує RDF-метадані для експериментальних компонентів і засновані на RDF-графі походження даних; експлуатує для бази даних керовані словники на OWL і засновану на OWL-процедуру інтеграції даних	Призначено для галузі біоінформатики, зокрема, дослідження генів
4.	Insurance Grid [23]	Один із перших проектів у галузі використання технологій знань з метою збільшення та розширення архітектури можливостей Грід-систем. Демонструє вагомий приклад застосування Грід у бізнесі й забезпечує ефективні сервіси знань	Insurance Grid використовується в страхових компаніях для спрощення процесу страхових виплат й інтеграції послуг ланцюжка партнерів, залучених у процес

5.	Earth System Grid (ESG) [29]	Грід ESG є віртуальним середовищем для досліджень у галузі науки про Землю і аналізу глобальних змін клімату. Мета ESG — забезпечити пошук і безпечний доступ до величезних обсягів даних, накопичених наукою про Землю. Пошук даних за допомогою метаданих став головним аспектом ESG. Було розвинуто схеми метаданих, прототипи онтологій, сервіси пошуку і відкриття даних, засновані на семантиці	Користувачі — це вчені, які досліджують клімат у національних лабораторіях, університетах та інших урядових організаціях, забезпечуючи експертизу для Міжурядової Групи зі змін клімату
6.	Molecular GRID [30]	Властивості молекул мають дуже добре визначену семантику і дозволяють створення семантичних Грід. Було створено мову CML (Chemical Markup Language) і онтології, засновані на словниках, для підтримки широкого кола додатків, зокрема публікації даних про безпечні хімічні суміші. Велика кількість властивостей молекул може бути розрахована програмами з квантової механіки (зокрема, геометрія молекул, їх енергія, диполь, заряд, частота і т.ін.). Час моделювання для кожної молекули різниться в мільйон разів (від 0,3 сек до 4 днів)	250,000 молекул були описані, перетворені на XML, і розщеплені на 500 партій. Кожну партію оброблено одним завданням для програми Condor
7.	MONET [31]	Забезпечує формування і вирішення математичних задач у термінах домену MKMnet (Mathematical Knowledge Management NET-work — менеджмент математичних знань у мережі) за допомогою спеціальної мови OpenMath (Markup Language For Mathematics — мова розмітки для математиків); аналізує сутність математичної задачі через технологію автоматичного доказу Calculemus (Mathematical Reasoning Network — мережа математичних доказів), ідентифікує і складає замовлені сервіси з сервісного WSIL. Директорія UDDI завантажує програмні компоненти до апаратних ресурсів OGSA/Semantic Grid	Створено відповідну математичну онтологію і реалізовано мережеве використання програми Maple
8.	saBIG [32]	saBIG інтегрує дані, що збираються різноманітними постачальниками (наприклад, лабораторіями, клініками) у різноманітних форматах і з багатодисциплінарних концептуальних точок зору. Для saBIG розроблений домен моделей онтологій і метаданих, який формує основу для інтеграції даних. Послуги saBIG вирішують узгодження ресурсів (наприклад, баз даних, файлів) і даних з інтерфейсами, побудованими на Web-послугах, і дозволяють створювати черги, використовуючи стандартні мови для черг даних	Система saBIG вільно об'єднує дослідників, лікарів і пацієнтів для розповсюдження інформації про засоби запобігання і лікування раку

Інтерфейс користувача, як і в традиційних Грід, може бути забезпечено через спеціальні семантичні Web-портали знань, за допомогою яких:

- користувач створює проксі-сертифікат, щоб отримати дозвіл на роботу в Грід-системі;
- користувачу забезпечується можливість редагування при підготовці текстів, і «показчиків» Грід-сервісів і даних. При клацанні клавiші миші контекстне меню сервісів показується користувачу, за допомогою якого він контактує із директорією сервісів і поновлює сервіси, які йому доступні;



Рис. 6. Три рівні програмування в семантичному Грід

- користувач переглядає сервіси за категорією або шукає їх за ключовими словами. Після того, як користувач знайшов сервіс, надалі він здійснює пошук даних, які необхідні сервісу. Дані, що запрошуються, розміщені у файлах і відповідають вибраній RDF-онтології;
- сервіс із відповідними даними запускається на віддаленому сервері. При цьому брокер ресурсу, базуючись на розташуванні або вимогах сервісу (визначених як фрагмент із онтології сервісу), інформації про ресурс (визначену, як фрагмент онтології ресурсу) і розташування даних, знаходить оптимальне місце, щоб виконати сервіс. На екрані порталу створюється область для друку результатів або їх візуалізації;
- сервіс виконується і результати запам'ятовуються для подальшої обробки або повертаються до користувача. Користувач може дочекатися обробки обчислення, або зберегти поточний результат. Щойно обробка завершено, результати відображаються у визначеному місці екрану порталу.

Необхідно зазначити, що при розробці і використанні семантичного Грід і Web-порталу знань, програмування самих семантичних сервісів і засобів їх використання для нього, ведеться на кількох рівнях із залученням сучасного інструментарію програмування (рис. 6). На думку Білла Гейтса, у міру повсюдного проникнення Інтернету все більшу популярність набуває концепція «програмне забезпечення як сервіс». Тобто, користувачі будуть працювати з необхідним ПЗ через Web і Грід, отримуючи на свої комп'ютери готові результати. Таким чином, такі програми як PHP, ASP, JSP-скрипти, JavaBeans, COM-об'єкти та всі інші засоби програмування вже зараз можуть звертатися до будь-якої програми, яка працює на іншому сер-

вері (тобто до Web-сервісу), і використовувати відповідь, отриману від неї на своєму Web-сайті або додатку.

Відповідно, необхідність у великих локальних обчислювальних потужностях частково відпаде, що сприятиме зростанню попиту на недорогі комп'ютери з низьким енергоспоживанням.

Прототип такого Web-порталу знань зараз розроблюється в Інституті прикладного системного аналізу НАНУ і МРНУ (ІПСА) згідно із проектом № ІТ / 558–2009 від 20 липня 2009 р. Державної програми ІКТ на 2006–2010 роки.

ВИСНОВКИ

Семантичний Грід як розширення сучасного Грід надзвичайно корисний для е-науки, оскільки його технології дають змогу легко, швидко та зручно працювати зі знанням, а не з інформацією, як це робиться зараз. Це дає можливість ученим, інженерам і бізнесменам працювати за фантастичними на сьогодні сценаріями.

На сьогодні відбувається активна розробка семантичного Web, але *достатнього впровадження його сервісів поки що немає*. Семантичний Web часто представляють як глобальну базу знань, особливо у світлі нещодавно розпочатих у Web-мережі проектів побудови семантичних енциклопедій, баз знань, лексичних баз розмовних мов та інших. Ідея семантичного Web — це концепція мережі, в якій кожен ресурс на людській мові забезпечений описом, зрозумілим комп'ютеру. Синтаксична взаємодія людей і комп'ютерів, та комп'ютерів між собою, полягає в коректному синтаксичному аналізі даних і вимагає побудови відповідностей між термінами, яка, в свою чергу, вимагає аналізу контексту за допомогою формальних явно заданих специфікацій моделей доменів (*онтологій*), які визначають використані терміни та зв'язки між ними.

Семантичний (або «розумний») Грід знаходиться на початковому етапі і для своєї реалізації потребує випереджаючого розроблення *багатоагентної інтелектуальної платформи*, хоча вже зараз можна розробляти і досліджувати його архітектуру і можливі властивості. Оскільки таку агентну платформу ще не створено, перші експериментальні зразки семантичного Грід будуються на базі існуючого Грід-середовища, яке базується на вживаному проміжному програмному забезпечення (middleware). Однак у middleware типу NorduGrid і gLite3, що використовуються в Україні та всій Європі, не реалізований стандарт WSRF сумісності Грід- і Web-сервісів, тому необхідно буде в національній Грід-інфраструктурі перейти на використання американського middleware типу Globus Tools 4 (GT-4) або дочекатися поки вийде наступна версія gLite-5, яка зараз розробляється в Європі із врахуванням стандарту WSRT.

Дослідження з можливої архітектури і властивостей семантичного Грід передбачено новою Державною програмою впровадження Грід-технологій на 2009–2013 роки, зокрема, в її розділі, присвяченому «розробленню і впровадженню методів ефективного використання існуючих Грід-ресурсів за допомогою інтелектуальної обробки даних (DataMining)» [6]. Семантич-

ний Web-порталу знань, що розробляється зараз в ПСА, призначено для дослідження заходів, необхідних для побудови національних реєстраторів схем метаданих і онтологій, заснованих на міжнародних стандартах, алгоритмів вилучення знань з даних та їх подальшого ефективного використання в національній Грід-інфраструктурі й мережі Світових Центрів Даних (СЦД), в якій успішно функціонує український СЦД з геоінформатики та сталого розвитку.

ЛІТЕРАТУРА

1. Згуровський М.З., Петренко А.І. Е-наука на шляху до семантичного Грід. Частина 1: Об'єднання Web- і Грід-технологій // Системні дослідження та інформаційні технології. — 2010. — № 1. — С. 26–38.
2. Згуровський М.З., Петренко А.І. Grid-технології для е-науки і освіти // Наукові вісті НТУУ «КПІ». — 2009. — № 2. — С. 10–17.
3. De Roure D., Jennings N.R., Shadbolt N.R. The Semantic Grid: Past, Present, and Future // Proceedings of the IEEE. — 2006. — **93**, № 3. — P. 669–681.
4. Петренко А.І. Семантичний Грід для гнучкого оброблення даних. — Матеріали конф. «Системний аналіз та інформаційні технології», 26–30 травня 2009 року. — С. 16–17.
5. Петренко А.І. Застосування Грід-технологій в науці і освіті. — Київ: Політехніка, 2009. — 145 с.
6. Петренко А.І. Grid і інтелектуальна обробка даних // Системні дослідження та інформаційні технології. — 2008. — № 4. — С. 97–110.
7. Berners-Lee T., Hendler J., Lassila O. The Semantic Web // Scientific American. 2001. — May. — P. 598–602.
8. The Gene Ontology. — <http://www.geneontology.org/>.
9. Natalya F. Noy & Deborah L. McGuinness. Ontology Development 101: A Guide to Creating Your First Ontology // Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001. — 25 p.
10. Онтологія UNSPSC. — <http://www.unspsc.org/>.
11. OWL Web Ontology Language Overview. — <http://www.w3.org/TR/owl-features/>.
12. The Protégé Ontology Editor and Knowledge Acquisition System. — <http://protege.stanford.edu/>.
13. Ontolingua (1997). Ontolingua System Reference Manual. — <http://www-ksl-svc.stanford.edu:5915/doc/frame-editor/index.html>.
14. Chimaera Ontology Environment. Chimaera (2000). — <http://www.ksl.stanford.edu/software/chimaera>.
15. The Foundation for Physical Agents. — <http://www.fipa.org/>.
16. OWL-based Web Service Ontology. — <http://www.daml.org/services/owl-s/>.
17. Mindswap OWL-S API. — <http://www.mindswap.org/2004/owl-s/api/>.
18. METEOR-S / WSDL-S. — <http://lsdis.cs.uga.edu/Projects/>.
19. Semantic Grid Community Portal. — <http://www.semanticgrid.org/>.
20. Jennings N.R. On agent-based software engineering // Artificial Intelligence 117 (2000). — P. 277–296.
21. A. Min Tjoa, Brezany P. and Janciak I. Towards Grid Based Intelligent Information Systems. Vienna University of Technology, Vienna, Austria, 2006. — 12 p.
22. Foundation for Intelligent Physical Agents. Fipa agent management specification. — <http://www.fipa.org/specs/fipa00023/>, 2000.

23. *Alper P., Corcho O., Kotsiopoulos I., Missier P., Bechhofer S., Goble C.* S-OGSA as a Reference Architecture for OntoGrid and for the Semantic Grid // GGF16 Semantic Grid Workshop. Athens, Greece, February 2006. — P. 87–94.
24. *M. Antonioletti et al.* The Design and Implementation of Grid Database Services in OGSA-DAI. Concurrency and Computation: Practice and Experience, 17(2–4), 2005. — P. 357–376.
25. *Knowledge-based Workow System for Grid Applications* FP6 IST project. — <http://www.kwfgrid.net>.
26. *Mirko Niinimaki.* Grid Resources, Services and Data — Towards a Semantic Grid System. (PhD Thises). Department of Computer Science, University of Tampere, Finland, January 2006. — 106 p.
27. *Dolenc M., Klinc R., Turk Ž., Katranuschkov P., Kurowski K.* Semantic Grid Platform in Support of Engineering Virtual Organisations // Informatica. — 2008. — № 32. — P. 39–49.
28. *Сайт проекту myGrid.* — <http://www.mygrid.org.uk>.
29. *The Gene Ontology.* — <http://www.geneontology.org/>.
30. *Murray-Rust P. and Rzepa H.S.* Journal of Chemical Information and Computer Science. — 2003. — № 43. — P. 234–238.
31. *MONET: Mathematical service discovery and composition.* — <http://monet.nag.co.uk/>.
32. *Cancer Biomedical Informatics Grid.* — <http://cabig.cancer.gov/index.asp>.

Надійшла 10.09.2009

СТИСНЕННЯ RGB-ЗОБРАЖЕНЬ БЕЗ ВТРАТ ІЗ ВИКОРИСТАННЯМ ПАЛІТРИ

О.В. ШПОРТЬКО

Розглянуто та обґрунтовано можливості стиснення RGB-зображень без втрат за допомогою палітрування. Описано варіант алгоритму для реалізації такого стиснення з розбиттям результату на трендову та шумову складові. Наведено результати застосування програми, розробленої згідно із запропонованим алгоритмом для стиснення зображень набору АСТ.

У файлах, що передаються каналами зв'язку, найчастіше містяться дані одного з чотирьох основних типів: тексти, зображення, відео або звук. І якщо для зберігання однієї сторінки неформатованого тексту достатньо 4 Кб, то фотореалістичне растрове зображення 10×15 см без стиснення може займати кілька Мб, записи пісень — десятки Мб, а записи фільмів — декілька Гб. Добре, що більшість із цих файлів мають високий рівень надлишковості. Саме зменшення надлишковості дозволяє стискати файли, і тим самим підвищувати швидкість обміну інформацією по мережі та зменшувати об'єми використання дискового простору. Тому зараз, в епоху інформаційної цивілізації, коли значна частина інформації зберігається в електронному вигляді, проблема стиснення зображень залишається актуальною.

Більшість відомих алгоритмів, які дозволяють стиснути зображення в десятки разів (JPEG, фрактальних та вейвлет-перетворень), призводять до незначних втрат якості зображень, що непомітно для фотореалістичних зображень з високою роздільною здатністю, але впливає на якість зображень із невисокою роздільною здатністю з фрагментами ділової графіки чи дискретно-тоновими переходами. Крім того, існують класи зображень, для яких спотворення неприпустимі (наприклад рентгенівські знімки). Алгоритми стиснення зображень без втрат хоча й стискають зображення значно слабше, але не призводять до погіршення їхньої якості. Зараз найдинамічніше розвиваються алгоритми стиснення зображень із втратами, хоча й алгоритми стиснення зображень без втрат теж із кожним роком підвищують свої показники.

ОСНОВНІ ПІДХОДИ ДО СТИСНЕННЯ ЗОБРАЖЕНЬ БЕЗ ВТРАТ. АНАЛІЗ ОСТАННІХ ДОСЯГНЕНЬ І ПУБЛІКАЦІЙ. ПОСТАНОВКА ЗАДАЧІ

Більшість сучасних графічних форматів для збереження кольору кожного пікселя використовують три компоненти [1]. Наприклад, у кольоровій моделі RGB — це яскравість червоної (R), зеленої (G) та синьої (B) компоненти. Під кожен компоненту відводиться, як правило, 1 байт. Тому розмір файла зображення у нестисненому вигляді приблизно дорівнює потроєному добутку кількостей пікселів по вертикалі (*height*) та горизонталі (*width*).

Інший спосіб зберігання кольору пікселів зображення передбачає використання палітри. *Палітра* — це одномірний масив трибайтових елементів,

кожен з яких визначає колір [2]. При такому способі збереження зображення колір кожного пікселя задається індексом в палітрі (наприклад у форматі GIF). Зазвичай, використовуються палітри з 2, 16 чи 256 кольорів [3]; (відповідно 1, 4 чи 8 бітам на індекс кольору пікселя). Палітри з більшою кількістю кольорів хоча й можливі, але на практиці не використовуються, оскільки вимагають більше байтів для зберігання самої палітри і, головне, значно більше бітів для зберігання індексу кольору кожного пікселя.

Розмір файла палітрового зображення у нестисненому вигляді приблизно дорівнює сумі розміру палітри та добутку кількостей пікселів по горизонталі та вертикалі, помноженій на розмір індексу кольору пікселя. Наприклад, під час використання палітри з 256 кольорів розмір файла складає приблизно $256 \cdot 3 + height \cdot width$ байт, тобто майже втричі менше, ніж при зберіганні згідно з кольоровою моделлю RGB. На жаль, кількість різних кольорів палітрового зображення не може перевищувати кількості кольорів у палітрі. Тому сучасні фотореалістичні зображення, що містять сотні тисяч різних кольорів, найчастіше використовують трикомпонентні кольорові моделі (RGB, YCrCb, HLS та ін.). Проблема використання палітри для стиснення таких зображень без втрат до цього часу залишалася невирішеною. Більшість сучасних графічних форматів стиснення без втрат (BMP, PNG) дозволяють зберігати зображення як із використанням палітри, так і без неї.

Сучасні архіватори та формати графічних файлів для стиснення зображень без втрат найчастіше використовують як один із основних етапів алгоритмів словникового методу [1, 3]. Такі алгоритми намагаються замінити наступні незакодовані символи посиланням на аналогічні символи у вже закодованій частині. Фотореалістичні зображення стискаються словниковими алгоритмами неефективно, оскільки, як правило, вони містять багато подібних, але неоднакових кольорів. Словникові алгоритми належать до класу контекстно-залежних, тому що використовують опрацьовану раніше послідовність. Алгоритми цього методу покликані усувати залежності між різними послідовностями елементів.

Для покращення стиснення даних результати дії словникових алгоритмів доцільно стиснути одним із контекстно-незалежних алгоритмів (арифметичним чи Хафмана). Ідея використання цих алгоритмів полягає у заміні елементів з більшою частотою послідовностями меншої кількості бітів, ніж для елементів з меншою частотою. При цьому середня довжина коду елемента після застосування контекстно-незалежного методу згідно з теоремою Шеннона має наближатися до величини:

$$H = -\sum_i p(s_i) \log p(s_i), \quad (1)$$

де $p(s_i)$ — ймовірність появи елемента s_i , а всі логарифми тут і надалі беруться за основою 2. Цю величину також називають *ентропією джерела* [1]. Вона зменшується під час збільшення нерівномірності розподілу ймовірностей між елементами.

Для аналізу ефективності запропонованого алгоритму ми використали програму WinRar версії 4.00 із застосуванням лише основного алгоритму, що послідовно виконує LZ-кодування та стиснення Хафмана (LZ+Huff).

Основну увагу при розробці алгоритмів стиснення без втрат приділяють показникам компресії: коефіцієнту стиснення (КС), тобто відсотку зменшення початкового розміру файла або усередненому значенню ентро-

пії, тобто кількості бітів, що всередньому витрачається на кодування одного пікселя (bpp). Оскільки між цими показниками існує взаємно-однозначна відповідність, то ми скористалися лише першим, тому що він використовується для оцінки стиснення не лише зображень.

Врахуємо, що згідно з [4]: «Відхилення між значеннями сусідніх елементів у зображеннях найчастіше зумовлені двома причинами: «сильними» коливаннями, що обумовлюються зображеними об'єктами — трендом, і «слабкими» фоновими коливаннями — шумом. Тому можливі два протилежні типи моделей:

- внесок шуму незначний у порівнянні з внеском тренду;
- внесок тренду незначний у порівнянні з внеском шуму.

Око людини насамперед орієнтується на контури об'єктів (тренд) і менш чутливе до фонових коливань (шуму). Тому під час передачі даних по мережі для забезпечення прогресуючого стиснення доцільно спочатку передавати дані тренду і лише після цього — шуму.

МЕТА ДОСЛІДЖЕННЯ

Обґрунтування можливості та опис відповідного алгоритму для стиснення довільних RGB-зображень без втрат із використанням палітри та словникових методів. Основною вимогою до алгоритму є досягнення максимальних КС у практично незмінних витратах часу.

АЛГОРИТМ СТИСНЕННЯ RGB-ЗОБРАЖЕНЬ БЕЗ ВТРАТ ІЗ ВИКОРИСТАННЯМ ПАЛІТРИ

З'ясуємо спочатку глибинні причини майже потрійного стиснення дискретно-тонових зображень у разі використання палітри з 256 кольорами. У кольоровій моделі RGB значення кожної компоненти лежить у межах від 0 до 255. Отже, загалом ця модель адресує $255 \times 256 \times 256 = 16777216$ кольорів. При використанні ж палітри індексується лише 256 кольорів, тобто для зберігання індексу кольору використовується 8 бітів замість 24.

Надлишкова індексація характерна для кольорової моделі RGB і у випадку зберігання фотореалістичних зображень, адже в кожному з них використовується лише невеличка частина спектру кольорів (наприклад, в зображеннях набору Archiv Comparison Test (ACT) — менше 1% (номер 6 у табл. 1)).

Таблиця 1. Характеристики зображень набору ACT

Номер файла	Назва файла	Розмір, Кб	Розміри пікселів	Кількість різних кольорів	Використання спектру, %	Різні кольори серед пікселів, %
1	Clegg.bmp	2101	814 × 880	127696	0,76	17,83
2	Frymire.bmp	3622	1118 × 1105	3622	0,02	0,29
3	Lena.bmp	769	512 × 512	148279	0,88	56,56
4	Monarch.bmp	1153	768 × 512	78617	0,47	19,99
5	Peppers.bmp	769	512 × 512	111344	0,66	42,47
6	Sail.bmp	1153	768 × 512	75748	0,45	19,26
7	Serrano.bmp	1464	629 × 794	1313	0,01	0,26
8	Tulips.bmp	1153	768 × 512	118233	0,70	30,07

Максимальна кількість різних кольорів у зображенні у найгіршому випадку не перевищує кількості пікселів. Наприклад, у фотореалістичному зображенні *Lena.bmp* (рис. 1) використовується 148279 кольорів 262144 пікселями, тобто 0,88% кольорів від можливих. Розглядаючи навіть проєкцію кольорів пікселів цього зображення на площину RG (рис. 4), можна помітити, що використовується менше 30% індексованого простору площини.



Рис. 1. Зображення *Lena.bmp*



Рис. 2. Зображення *Lena.bmp* в палітрі з 63 кольорів



Рис. 3. Зображення *Lena.bmp* в палітрі з 256 кольорів

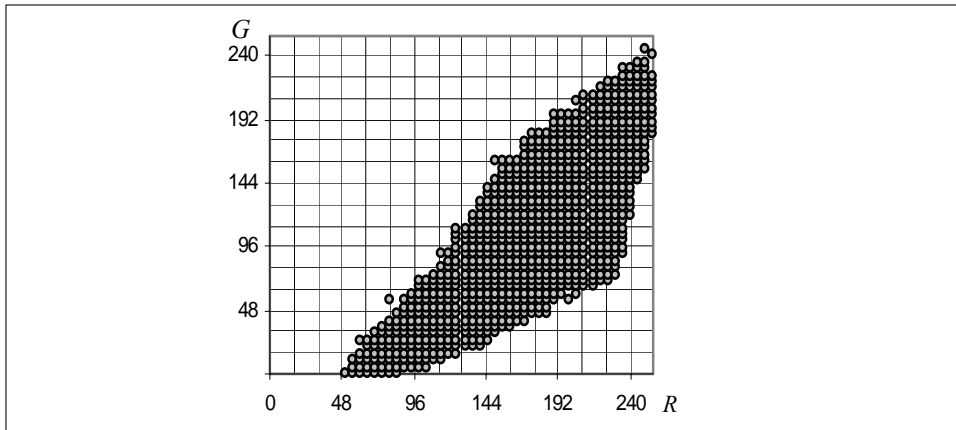


Рис. 4. RG-проєкція кольорів пікселів зображення *Lena.bmp*

Для зменшення надлишкової індексації кольоровою моделлю *сформуємо палітру тренду*, виконуючи такі кроки:

1. Розіб'ємо множину допустимих значень кольорів на сегменти максимального фіксованого розміру [5], що не перекриваються між собою (наприклад, для RGB — це множина прямокутних паралелепіпедів). Оскільки наперед невідомо, в яких сегментах містяться кольори пікселів обраного зображення, то множина сегментів має повністю покривати множину допустимих значень кольорової моделі.

2. Визначимо кількість кольорів пікселів та межі їх знаходження у кожному сегменті. Зв'язимо межі кожного сегменту до меж знаходження кольорів пікселів у ньому.

3. Поеднаємо між собою сусідні сегменти, якщо їх сукупний розмір не перевищує максимального фіксованого.

4. Збережемо в палітрі координати лише тих сегментів, які містять кольори пікселів. Тоді колір кожного пікселя можна розбити на дві складові — індекс сегменту в палітрі (тренд) та зміщення кольору всередині сегменту (шум). Оскільки у найгіршому випадку кожен вихідний сегмент може містити кольори пікселів, то початкова кількість сегментів не може перевищувати максимально допустимої кількості кольорів палітри.

5. Доповнимо палітру до максимально можливої кількості кольорів, використовуючи поділ окремих сегментів так, щоб максимально збільшити КС.

Як зазначалося вище, трендову та шумову складові доцільно зберігати окремо для забезпечення прогресивної передачі даних по мережі. Варіант програми, що реалізує запропонований алгоритм, взагалі розбиває зображення на два файли, які містять відповідно дані трендової та шумової моделі. При цьому множина всіх допустимих значень кольорів на початку програми розбивається на 256 сегментів-паралелепіпедів розміром $32 \times 32 \times 64$ значення. Саме такі сегменти будемо розглядати надалі. Кількість сегментів-паралелепіпедів, що реально використовуються окремими зображеннями, після четвертого кроку алгоритму, як правило, не перевищує 100. Наприклад, зображення *Lena.bmp* використовує 63 паралелепіпеда з 256 (рис. 2) кольорів.

Розглянемо результати застосування запропонованого алгоритму для стиснення різнотипних зображень згаданого вище набору АСТ. Файл тренду, отриманий у результаті виконання запропонованого алгоритму, під час використання палітри з 256 кольорів без стиснення займає третю частину вхідного файла. Його доцільно стиснути алгоритмом LZ+Huff. Причому розмір стиснутого файла тренду завжди буде меншим від розміру стиснутого аналогічного вхідного зображення (номери 2, 5, 8 у табл. 2). Це пов'язано як із застосуванням палітри, так і зі зведенням кольорів сегментів до єдиного кольору палітри, тобто ліквідації шумів.

Таблиця 2. Результати стиснення файлів набору АСТ (Кб) при застосуванні та без застосування палітри

Номер файла	Без поділу паралелепіпедів			З поділом паралелепіпедів			LZ+Huff файла зображення
	LZ+Huff файла тренду	Файл шуму	Всього	LZ+Huff файла тренду	Файл шуму	Всього	
1	103	1398	1501	116	1186	1302	500
2	138	2288	2426	140	1073	1213	235
3	78	510	588	186	370	556	701
4	87	766	853	216	542	758	789
5	58	508	566	140	387	527	641
6	134	767	901	289	548	837	867
7	46	863	909	60	285	345	102
8	113	767	880	232	603	835	938
Разом	757	7867	8624	1379	4994	6373	4773

Зовсім інші підходи застосовуються для стиснення файлів шумів. Використання контекстно-залежних алгоритмів типу LZ тут неефективне, оскільки повторення послідовностей практично відсутні. Файли шумів можна відразу закодувати ентропійним алгоритмом, але такий підхід несуттєво підвищує КС.

Нами пропонується спосіб зберігання шумів шляхом безпосереднього запису зміщень кольорів усередині сегменту з урахуванням його розміру. Якщо, наприклад, максимальний сегмент-паралелепіпед має розміри $32 \times 32 \times 64$ значення, то для зберігання зміщення в ньому необхідно $\lceil \log(32) \rceil + \lceil \log(32) \rceil + \lceil \log(64) \rceil = 16$ біт. Тобто, навіть у випадку максимальних розмірів усіх паралелепіпедів (для хаотичних зображень) результат палітрування перевищить розмір вхідного файла лише на опис цих сегментів. Наприклад, для опису паралелепіпедів, що не перевищують згаданого максимального, достатньо 40 бітів: 24 — для опису базового кольору паралелепіпеда в палітрі і 16 — його розмірів. Для запису зміщень в паралелепіпедах меншого розміру знадобиться менша кількість біт. Саме завдяки цьому і досягається стиснення файлів шумів. Інша перевага безпосереднього запису зміщень кольорів усередині паралелепіпеда полягає у максимальній швидкості обробки даних, оскільки при цьому не виконуються ніякі перетворення. Результати стиснення шумів зображень набору АСТ лише завдяки врахуванню розмірів сегментів-паралелепіпедів (без п'ятого кроку алгоритму) наведено в табл. 2, у номері 3.

Підвищити КС файлів шумів можна завдяки поділу окремих паралелепіпедів. Навіть безпосереднє розбиття довільного паралелепіпеда навпіл зменшує кількість значень по осі поділу вдвічі, а кількість бітів — на один для кожного зміщення паралелепіпеда. Тому розбиття таким чином паралелепіпедів із максимальною кількістю точок покращує показники стиснення. Досягнути ж кращих КС можна завдяки оптимальному розбиттю паралелепіпедів. Розглянемо, наприклад, RG-проекцію зміщень кольорів пікселів окремого умовного паралелепіпеда (рис. 5), розбиття якого посередині ребра по осі R (штрих-пунктирна лінія) створить два паралелепіпеда: перший — у діапазоні $[0; 15]$ і другий — $[16; 31]$ по цій осі. Тобто сумарного зменшення розміру від такого поділу не відбудеться. Якщо ж розбити цей паралелепіпед по значенню $R = 7$ (пунктирна лінія), то можна буде створити лівий паралелепіпед у діапазоні $[0; 5]$ та правий — $[10; 31]$, що дозволить сумарно зменшити розмір по цій осі на 4 значення.

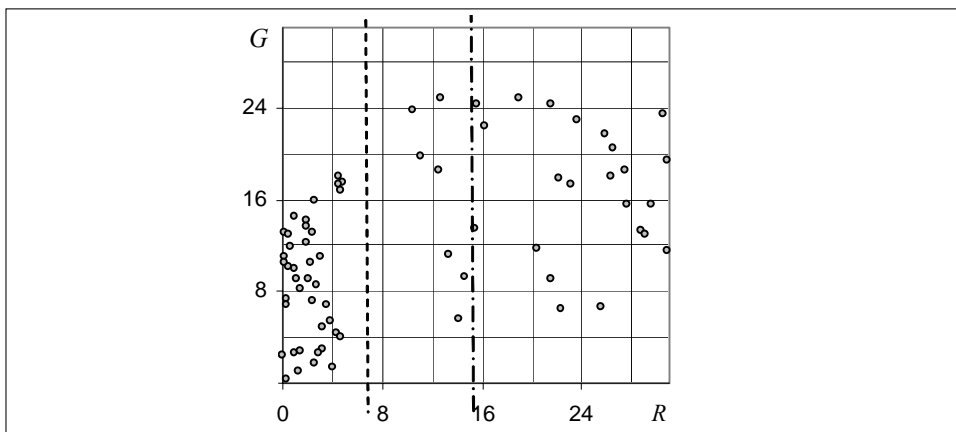


Рис. 5. RG-проекція зміщень кольорів пікселів окремого умовного паралелепіпеда

Отже, досягнути додаткового зменшення розмірів сегментів-паралелепіпедів у результаті поділу можна завдяки врахуванню положення

зміщень у ньому. Зменшення розмірів паралелепіпедів призводить до зменшення кількості бітів для збереження зміщень, тобто до покращення показників стиснення.

Виведемо формулу для підрахунку зменшення кількості бітів у записі зміщень кольорів у результаті поділу паралелепіпеда по значенню i осі R на два паралелепіпеди — правий і лівий. Нехай діапазон значень ребра паралелепіпеда до поділу по осі R знаходиться в межах $[\min R; \max R]$, по осі G — $[\min G; \max G]$, а по осі B — $[\min B; \max B]$. Позначимо countPoint загальну кількість зміщень у паралелепіпеді, а $\text{count}R(j)$, $\text{count}G(j)$, $\text{count}B(j)$ — кількості зміщень, які мають значення j відповідно по R , G та B . Очевидно, що:

$$\text{countPoint} = \sum_{j=\min R}^{\max R} \text{count}R(j) = \sum_{j=\min G}^{\max G} \text{count}G(j) = \sum_{j=\min B}^{\max B} \text{count}B(j). \quad (2)$$

Позначимо також \underline{i} праву межу лівого, а \bar{i} ліву межу правого паралелепіпедів при поділі по осі R .

$$\underline{i} = \max_{\substack{j=\min R, i \\ \text{count}R(j)>0}} j, \quad \bar{i} = \min_{\substack{j=i+1, \max R \\ \text{count}R(j)>0}} j.$$

Кількість бітів, необхідних для запису зміщень кольорів паралелепіпеда до поділу, дорівнює:

$$\text{countBit} = \text{countPoint} \left(\begin{array}{l} \lceil \log(\max R - \min R + 1) \rceil + \\ \lceil \log(\max G - \min G + 1) \rceil + \\ \lceil \log(\max B - \min B + 1) \rceil \end{array} \right). \quad (3)$$

Якщо знехтувати зменшенням розмірів лівого та правого паралелепіпедів після поділу по інших осях, то кількість бітів, яка знадобиться для запису зміщень кольорів у цих сегментах, відповідно складе:

$$\text{countLeftBit}_i = \sum_{j=\min R}^{\underline{i}} \text{count}R(j) \left(\begin{array}{l} \lceil \log(\underline{i} - \min R + 1) \rceil + \\ \lceil \log(\max G - \min G + 1) \rceil + \\ \lceil \log(\max B - \min B + 1) \rceil \end{array} \right) \quad (4)$$

та

$$\text{countRightBit}_i = \sum_{j=\bar{i}}^{\max R} \text{count}R(j) \left(\begin{array}{l} \lceil \log(\max R - \bar{i} + 1) \rceil + \\ \lceil \log(\max G - \min G + 1) \rceil + \\ \lceil \log(\max B - \min B + 1) \rceil \end{array} \right). \quad (5)$$

Тоді очевидно, що зменшення кількості бітів від поділу сегменту по значенню i осі R складе:

$$\text{price}_i = \text{countBit} - (\text{countLeftBit}_i + \text{countRightBit}_i). \quad (6)$$

Умовою оптимального розбиття сегменту-паралелепіпеда по осі R є максимальне зменшення кількості бітів:

$$\text{price}_R = \max_{i=\min R, \max R} \text{price}_i. \quad (7)$$

Підставимо послідовно (3) – (5) в (6) та (6) в (7), розкриємо дужки та, врахувавши (2), після зведення отримаємо:

$$\text{price}_R = \frac{\max}{i=\min R, \max R} \left(\begin{array}{l} \text{countPoint} \lceil \log (\max R - \min R + 1) \rceil - \\ - \sum_{j=\min R}^i \text{count} R(j) \lceil \log (i - \min R + 1) \rceil + \\ + \sum_{j=i}^{\max R} \text{count} R(j) \lceil \log (\max R - i + 1) \rceil \end{array} \right) \quad (8)$$

або

$$\text{price}_R = \text{countPoint} \lceil \log (\max R - \min R + 1) \rceil - \frac{\min}{i=\min R, \max R} \left(\begin{array}{l} - \sum_{j=\min R}^i \text{count} R(j) \lceil \log (i - \min R + 1) \rceil + \\ + \sum_{j=i}^{\max R} \text{count} R(j) \lceil \log (\max R - i + 1) \rceil \end{array} \right).$$

Формули для знаходження оптимального розбиття паралелепіпеда по осях G та B отримуються аналогічно. Умовою загального оптимального розбиття сегменту-паралелепіпеда є максимальне зменшення кількості бітів від поділу по довільній осі

$$\text{price} = \max(\text{price}_R, \text{price}_G, \text{price}_B),$$

а місцем — площа, щодо якої цей максимум досягається. Оскільки для зберігання опису паралелепіпеда витрачається 40 бітів, то поділ доцільно виконувати лише тоді, коли вигреш від оптимального розбиття перевищує це значення.

Розглянемо тепер практичні аспекти реалізації знаходження оптимального розбиття паралелепіпеда по кожній з осей, які дозволяють пришвидшити виконання обчислень, на прикладі осі R згідно з (8).

1. Зменшення кількості бітів від розбиття паралелепіпеда по довільному значенню осі не залежить від його розмірів та від значень зміщень, а залежить, насамперед, від значень проєкцій точок на аналізовану вісь.

2. Для всіх i , таких, що $\text{count} R(i) = 0$, $\text{price}_i = \text{price}_{i-1}$, проводити розрахунки зменшення кількості біт згідно з (8) не потрібно.

3. Послідовно змінюючи i , перераховувати щоразу дві внутрішні суми у (8) не потрібно, адже відносно попередньої перша сума збільшується на $\text{count} R(i)$, а друга — зменшується на цю ж величину.

У мові C алгоритм знаходження оптимального розбиття паралелепіпеда може бути реалізовано так:

```
//розрахуємо кількість проєкцій зміщень паралелепіпеда на всі значення осей
memset(countR, 0, 32 * sizeof(UBYTE4));
memset(countG, 0, 32 * sizeof(UBYTE4));
memset(countB, 0, 64 * sizeof(UBYTE4));
```

```

p = firstPoint; //індекс першої точки паралелепіпеда
//перебираємо зміщення паралелепіпеда
for (index = 0 ; index < countPoint; index++)
    {countR[imageData[3 * p] - minR] ++};
    {countG[imageData[3 * p + 1] - minG] ++};
    {countB[imageData[3 * p + 2] - minB] ++};
    p = nextPoint[p]; //перехід до наступного зміщення паралелепіпеда
    plusBit = 0; //ініціалізуємо змінну максимального зменшення кількості біт
if (maxR > minR) //якщо розбиття паралелепіпеда по осі R можливе
    {countLeft = 0; //ініціалізація першої внутрішньої суми
    {countRight = countPoint; //ініціалізація другої внутрішньої суми
    p = countPoint * countBit(maxR - minR + 1); //стале зменшуване осі
    for (index = 0; index < maxR - minR; index ++ ) //перебираємо всі значення
осі R
        {if (countR[index] > 0) //index вказує на праву межу лівого паралелепі-
педа
            {countLeft+ = countR[index];
            {countRight - = countR[index];
            index1 = index + 1; //index1 вказує на ліву межу правого паралелепіпеда
            while (countR[index1] = 0)
                index1 ++ ; //посуваємо нижню межу правого паралелепіпеда
            //підраховуємо зменшення кількості біт від чергового розбиття
            plus = p - (countLeft * countBit(index + 1) + countRight*)
                countBit(maxR - minR - index1 + 1));
            if (plus > 0)
                if (plusBit < plus) //знайдено оптимальніше розбиття
                    (plusBit = plus); //оптимальніше зменшення кількості біт
                    topMega = 1; //індекс осі з оптимальнішим розбиттям
                    mega = minR + index; //значення осі, де досягається оптимальніше
розбиття.

```

Описаний алгоритм оптимального поділу сегментів-паралелепіпедів, хоча й дозволяє цілком зменшити розміри файлів шумів, проте негативно впливає на стиснення файлів трендів (номери 5–7, у табл. 2). Це пов'язано зі створенням нових значень у палітрі, що зменшує кількості повторень і тим самим негативно впливає на стиснення алгоритмами LZ. З іншого боку, поділ сегментів зменшує нерівномірність розподілу, що призводить до збільшення ентропії (1) і негативно впливає на стиснення алгоритмом Хафмана. Загалом, використання алгоритму оптимального поділу паралелепіпедів підвищує КС по набору АСТ на 18,47% (номери 2, 3, у табл. 3). У цій же таблиці наведено КС при безпосередній компресії зображення алгоритмом LZ+Huff та кращою на сьогодні програмою для цього набору ERI 5.1

(за даними <http://www.compression.ru/arctest/act/act-tif.htm>). Крім того, збільшення кількості кольорів палітри до максимально можливої значно покращує якість трендового зображення (див. рис. 2, 3). Час палітрування та виконання програми згідно з описаним алгоритмом наведено в табл. 4.

Таблиця 3. КС файлів набору АСТ (%) при та без застосування палітри

Номер файла	КС без поділу паралелепіпедів	КС з поділом паралелепіпедів	КС LZ+Huff файла зображення	КС ERI 5,1
1	28,56	38,03	76,20	83,48
2	33,02	66,51	93,51	94,81
3	23,54	27,70	8,84	39,92
4	26,02	34,26	31,57	61,67
5	26,40	31,47	16,64	56,31
6	21,86	27,41	24,80	53,95
7	37,91	76,43	93,03	94,67
8	23,68	27,58	18,65	55,85
Разом	29,22	47,69	60,83	76,26

Таблиця 4. Час палітрування та виконання програми (С) для файлів набору АСТ (на комп'ютері з частотою процесора 300 МГц)

Номер файла	Без поділу паралелепіпедів		З поділом паралелепіпедів	
	Тривалість палітрування	Тривалість програми	Тривалість палітрування	Тривалість програми
1	2	5	2	5
2	2	8	4	10
3	1	2	2	2
4	1	2	2	3
5	1	1	1	3
6	1	2	2	4
7	1	3	2	3
8	1	3	2	3
Разом	10	26	17	33

Запропонований алгоритм дозволяє досягнути кращих КС стосовно базового алгоритму для фотореалістичних зображень із десятками тисяч різних кольорів (номери 3–6, 8, у табл. 3), хоча й значно програє на зображеннях із розсіяним спектром (номер 1), великими однаковими ділянками (номери 2, 7) і, як наслідок, загалом по набору АСТ. Виконується запропонований алгоритм майже стільки ж часу, скільки триває зчитування і запис файлів на диск.

Для подальшого вдосконалення наведеного алгоритму пропонуємо:

- забезпечити можливість необмеженого розширення палітри;
- зберігати опис та використовувати поділені сегменти-паралелепіпеди у файлах шумів для запобігання погіршення стиснення файлів тренду;
- реалізувати стиснення за допомогою палітрування з використанням предикторів;

- використовувати описаний алгоритм після LZ-алгоритму, що дозволить компактніше зберігати однакові ділянки зображення.

ВИСНОВКИ

1. Запропонований алгоритм стиснення RGB-зображень без втрат за допомогою палітрування виконує розбиття результату на трендову та шумову складові. Трендова складова містить наближену палітрову копію зображення. У шумовій — зберігаються відхилення оригіналу від палітрової копії. Описаний алгоритм дозволяє стиснути дані при створенні палітри головним чином завдяки зменшенню надлишкової індексації трикомпонентної кольорової моделі.

2. Трендова складова розбиття займає менше третини стиснених даних, і, тому може бути використана для ефективної прогресивної передачі даних по мережі. Ефективне стиснення трендової частини можливе традиційними алгоритмами (наприклад LZ+Huff).

3. Шумова складова розбиття ефективно стискується за допомогою запропонованого алгоритму оптимального розбиття сегментів-паралелепіпедів, що загалом підвищує КС у півтора рази. Таке розбиття дозволяє не тільки підвищити якість, але й збільшити розміри трендової складової результату.

4. Розроблений алгоритм пропонує кращі результати стиснення стосовно базового алгоритму на фотореалістичних зображеннях із десятками тисяч кольорів, хоча й поступається на інших типах зображень і відстає від найкращих на сьогодні алгоритмів стиснення. Ідея палітрування фотореалістичних зображень у комбінації з іншими алгоритмами стиснення дозволяє досягти кращих загальних показників компресії і буде втілюватися в життя й надалі.

ЛІТЕРАТУРА

1. Сэломон Д. Сжатие данных, изображений и звука. — М.: Техносфера, 2006. — 336 с.
2. Миано Дж. Форматы и алгоритмы сжатия изображений в действии: Учеб. пособие. — М.: Триумф, 2003. — 336 с.
3. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. — М.: ДИАЛОГ-МИФИ, 2003. — 384 с.
4. Бредихин Д.Ю. Сжатие графики без потерь качества. — 2004. — http://www.compression.ru/download/articles/i_lless/bredikhin_2004_lossless_image_compression_doc.rar.
5. Ратушняк О.А. Алгоритмы сжатия изображений без потерь с помощью сортировки параллельных блоков // Тез. докл. конф. молодых ученых по математике, мат. моделированию и информатике. — Новосибирск, 2001. — С. 48–49.

Надійшла 23.04.2008

УДК 519.8

ОЦЕНКА КРЕДИТНЫХ БАНКОВСКИХ РИСКОВ С ИСПОЛЬЗОВАНИЕМ НЕЧЕТКОЙ ЛОГИКИ

Ю.П. ЗАЙЧЕНКО

Рассмотрены проблемы оценки кредитных рисков заемщиков банков. Изложен классический метод оценки кредитных рисков на основе определения кредитного рейтинга заемщиков, а также оценки кредитного обеспечения. Предложен метод определения кредитоспособности заемщиков на основе использования систем с нечеткой логикой (алгоритм нечеткого вывода Мамдани). Сформулированы правила нечеткого вывода для оценки кредитного рейтинга и определения итогового класса заемщиков. Приведены примеры применения изложенного нечеткого метода.

ВВЕДЕНИЕ

Задача анализа кредитоспособности заемщиков банков и оценки кредитных рисков представляет значительный интерес в связи с тем, что в последние годы в Украине наблюдается бум в сфере кредитования. Поэтому, для того, чтобы соответствовать требованиям Национального Банка Украины и рекомендациям Базельского комитета по банковскому надзору (попросту говоря, не разориться самим и не разорить своих клиентов), банкам необходимо иметь тщательно разработанную политику по управлению рисками.

Одной из важнейших составляющих данной политики является набор профессионально подобранных, с учетом экономического состояния страны и специфики деятельности банка, статистических и (или) математических методов оценки ожидаемых и неожиданных потерь.

Цель работы — рассмотрение и сравнительный анализ методов оценки кредитоспособности заемщика:

- классический анализ кредитоспособности (четкий метод);
- анализ с помощью нечеткой логики, используя алгоритм Мамдани.

ЭЛЕМЕНТЫ ТЕОРИИ РИСКОВ

Определение понятия «риск»

Деятельность любого предприятия неразрывно связана с понятием «риск». Банк, в котором Вы держите свои денежные средства, может обанкротиться, деловой партнер, с которым заключена сделка, — оказаться недобросовест-

ным, а сотрудник, принятый на работу, — некомпетентным. Не стоит забывать и о стихийных бедствиях, компьютерных вирусах, экономических кризисах и других явлениях, способных нанести урон компании. Вместе с тем, рисками можно управлять так же, как процессами производства или закупки материалов.

Рассмотрим общее понятие риска.

Несмотря на то, что вся жизнь и деятельность людей происходит в условиях неопределенности, которая порождает риск, среди исследователей до сих пор не существует единственного мнения, касательно определения риска. Приведем простейшие из них.

Риск — возможная неудача.

Риск — неуверенность в возможном результате.

Риск — угроза неблагоприятного результата.

Риск — способ действия в условиях в непонятной или неопределенной обстановке.

Риск — событие, которое может произойти или не произойти.

Итак, Риск — объективно-субъективная категория, связанна с преодолением неопределенности, случайности, конфликтности в ситуации неизбежного выбора и отображающая степень достижения ожидаемого результата.

Цели управления рисками

В зависимости от сферы деятельности, деловой среды, стратегии развития и других факторов компания может сталкиваться с различными видами рисков. Тем не менее, существуют общие цели, достижению которых должен способствовать эффективно организованный процесс управления рисками.

Согласно «Методическим рекомендациям по организации и функционированию систем риск-менеджмента в банках Украины» *цель управления рисками — способствовать повышению стоимости собственного капитала банка*. Одновременно необходимо соблюдать интересы заинтересованных сторон (клиентов и контрагентов, руководителей, сотрудников, наблюдательного совета и акционеров/собственников, органов банковского контроля, рейтинговых агентств, инвесторов, кредиторов и т.д.).

Понятие кредитных банковских рисков

Кредитный риск — это возможность потерь вследствие неспособности или нежелания контрагента выполнить свои контрактные обязательства.

Данный риск присутствует во всех видах деятельности банка, где результат зависит от действий контрагента, эмитента или заемщика. Он возникает каждый раз, когда банк предоставляет (берет обязательство о предоставлении), инвестирует средства или другим образом рискует ими согласно условиям договоров независимо от того, где отображается операция — в балансе или вне его.

Для кредитора последствия невыполнения обязательств измеряются потерей основной суммы задолженности и невыплаченных процентов за вычетом суммы восстановленных денежных средств.

Кредитный риск целесообразно разделить на *индивидуальный* и *портфельный*.

Источником индивидуального кредитного риска является отдельный контрагент банка. Соответственно, оценка индивидуального кредитного риска предусматривает оценку кредитоспособности отдельного контрагента, т.е. его индивидуальную способность своевременно и в полном объеме рассчитаться по принятым обязательствам.

Портфельный кредитный риск проявляется в уменьшении стоимости активов банка. Источник портфельного кредитного риска — совокупная задолженность перед банком по операциям, которым свойственен кредитный риск (кредитный портфель, портфель ценных бумаг, портфель дебиторской задолженности и т.д.). Оценка портфельного кредитного риска предусматривает оценку концентрации и диверсификации активов банка.

Возникновение кредитного риска тесно связано с понятием кредитного события. Кредитное событие — изменение кредитоспособности заемщика или кредитного «качества» финансового инструмента, наступление которого характеризуется четко определенными условиями.

ISDA (International Swap and Derivatives Association — Международная ассоциация по свопам и производным финансовым инструментам) определяет девять различных видов кредитного события:

1. *Банкротство* (ликвидация предприятия (за исключением слияния)); неплатежеспособность (несостоятельность) предприятия; переуступка прав требования (цессия); возбуждение дела о банкротстве в суде; назначение внешнего управляющего имуществом должника; наложение ареста третьей стороной на все имущество должника).

2. *Досрочное наступление срока исполнения обязательства* (obligation acceleration — ускорение срока исполнения обязательства), которое означает объявление дефолта (отличного от невыплаты причитающейся суммы) по любому другому аналогичному обязательству данного заемщика и вступление в силу оговорки о досрочном наступлении срока исполнения данного обязательства.

3. *Дефолт по обязательству* (кросс-дефолт), означающий объявление дефолта (отличного от невыплаты причитающейся суммы) по любому другому аналогичному обязательству данного заемщика.

4. *Неплатежеспособность*, подразумевающая невыплату заемщиком определенной (превышающей оговоренный предел) суммы в срок (по истечении оговоренного льготного периода).

5. *Отказ или мораторий*, при котором контрагент отказывается от совершения платежа или оспаривает юридическую силу обязательства.

6. *Реструктуризация задолженности*, влекущая за собой односторонний отказ, отсрочку или изменение графика погашения задолженности на менее выгодных условиях.

7. *Понижение или отзыв рейтинговым агентством кредитного рейтинга заемщика* (это обычно приводит к понижению рыночной стоимости обязательств контрагента, а также к потерям в виде недополученной прибыли вследствие досрочного возврата ссуды заемщиком).

8. *Неконвертируемость валюты*, вызванная введенными государством ограничениями.

9. *Действия государственных органов*, под которыми понимаются:

- заявления или действия правительства или регулирующих органов, ставящие под угрозу юридическую силу обязательства, либо
- война или военные действия, препятствующие осуществлению деятельности правительства или банковской системы.

Оценка кредитных банковских рисков

Одно из основных требований Базельского комитета (Basel II) состоит в соответствии капитала банка его рискам, которые необходимо уметь определять, чтобы формулировать требования к капиталу, обеспечивающие банку надежность. При этом невозврат единичных кредитов не принесет ощутимого урона банку, если сможет быть компенсирован резервами, отчисляемыми под ожидаемые потери по кредитным операциям (EL — Expected Loss). Кроме того, существует шанс потерь значительной части активов в кредитном портфеле, приводящих к банкротству банка. Такие потери называют неожиданными потерями (UL — Unexpected loss).

Величина ожидаемых потерь напрямую влияет на прибыль от кредитного продукта, поскольку необходимо отчислять страховую сумму в резервный высоко-ликвидный фонд с каждого кредита. Эта сумма должна быть не менее величины EL.

Величина неожиданных потерь по портфелю также косвенно влияет на прибыль от кредитной деятельности, поскольку определяет собственный уровень надежности кредитного портфеля и банка в целом. Неождаемые потери рассматриваться не будут.

Одной из основных и нетривиальных задач в оценке риска — это задача вычисления вероятности дефолта заемщика (PD — Probability of Default). Распространены два подхода к вычислению PD. Первый основан на качественной и количественной оценке рейтинга заемщика по его внутренним финансовым показателям и особым бизнес-факторам. Второй основан на капитализации заемщика на фондовом рынке и уровне его долгов перед кредиторами. К сожалению, второй подход, хоть и является наиболее объективным, применим лишь к очень малому числу украинских открытых компаний.

Рассмотрим некоторые методы первого подхода, основанного на качественной и количественной оценке рейтинга заемщика по его внутренним финансовым показателям и особым бизнес-факторам.

Оценка рейтинга заемщика по его внутренним финансовым показателям и особым бизнес-факторам

На практике банки управляют кредитными рисками, руководствуясь собственными методиками кредитного анализа и отбора заемщиков. Этот анализ заключается в определении кредитоспособности, платежеспособности и финансовой устойчивости заемщика, что, в конечном счете, приводит к формулированию оснований для предоставления кредита или отказа в нем. Основной акцент в кредитном анализе делается на готовность и способность заемщика выплатить кредит, для оценки которых тщательно изучается характер деятельности заемщика, его кредитная история, текущее финансовое состояние, его возможности и потенциал.

Под кредитоспособностью будем понимать не просто способность юридического или физического лица полностью и в срок рассчитаться по своим долговым обязательствам, а желание, соединенное с возможностью своевременно погасить выданное обязательство.

Кредитный анализ — это сложный процесс, включающий в себя следующие основные этапы:

- Анализ обоснования потребности в кредите, представленного потенциальным заемщиком.
- Анализ финансовой отчетности предприятия. Особое внимание при этом уделяется выявлению тенденций развития и изменений в деятельности предприятия за определенный период времени в прошлом для формирования ясного представления о ближайших перспективах предприятия.
- Анализ предварительной финансовой отчетности предприятия. Нетипичные или вызывающие сомнение финансовые операции могут просматриваться в предварительной финансовой отчетности, анализ которой полезен для определения целостности и достоверности раскрываемой финансовой информации.
- Рассмотрение плана движения денежных средств в целях планирования поступлений платежей и определения вероятности их отсрочки, и оценка на его основе возможности заемщика своевременно погасить ссуду.
- Сценарный анализ и оценка устойчивости заемщика к экстремальным изменениям экономической среды.
- Анализ положения предприятия на рынке по отношению к другим конкурентам в отрасли.
- Оценка высшего управленческого звена предприятия, его стратегий, методов управления и эффективности деятельности на основе достигнутых результатов.
- Оформление необходимой документации для предоставления кредита, отражающей такие существенные условия, как:
 - обеспечение возврата ссуды;
 - ограничения по размеру задолженности;
 - оценку финансового состояния и обязательств заемщика;
 - гарантии, предоставляемые материнским (дочерним) предприятием или третьим лицом;
 - условия наступления дефолта и т.д.

КЛАССИЧЕСКИЙ МЕТОД АНАЛИЗА РИСКА БАНКРОТСТВА

Рассмотрим распространенный классический подход к анализу кредитоспособности заемщика, основной целью которого является классификация потенциальных заемщиков по степени риска неплатежеспособности. Количество групп риска выбирается произвольно с учетом уровня градации, до которого необходимо осуществить распределение кредитов (например, 10, 8, 6 групп и т.п.). При этом нежелательно выбирать большое количество групп, поскольку грани между ними стираются, и возникают трудности при классификации ссуд.

Согласно этому подходу, процесс оценивания кредитного риска заемщика заключается в оценке:

- кредитоспособности заемщика;
- риска кредитного продукта.

При анализе кредитоспособности заемщика можно выделить несколько этапов:

Этап № 1 (самый важный) — комплексный анализ финансового состояния заемщика, который включает:

1) анализ структуры активов и пассивов (заемные средства предприятия группируются по степени их напряженности, а активы — по степени их ликвидности. Осуществляется оценка изменений структуры активов и пассивов в динамике);

- 2) анализ денежных потоков;
- 3) анализ финансовой устойчивости предприятия;
- 4) оценка эффективности деятельности предприятия.

Для оценки финансового состояния предприятия используются следующие основные аналитические коэффициенты:

- коэффициент текущей ликвидности или коэффициент покрытия =

$$= \frac{\text{Оборотные средства}}{\text{Краткосрочная кредиторская задолженность}}$$

(характеризует платежеспособность предприятия в период, который равен средней продолжительности одного оборота краткосрочной дебиторской задолженности);

- коэффициент быстрой ликвидности =

$$= \frac{\text{Денежные средства} + \text{Дебиторская задолженность} + \text{Прочие ликвидные активы}}{\text{Краткосрочная кредиторская задолженность}}$$

(показывает возможности предприятия по погашению текущих обязательств за счет наиболее ликвидных активов);

- коэффициент напряженности обязательств =

$$= \frac{\text{Краткосрочная кредиторская задолженность}}{\text{Дневная выручка от реализации}}$$

(указывает на долю текущих обязательств в дневном обороте заемщика);

- коэффициент финансовой независимости =

$$= \frac{\text{Собственные средства}}{\text{Заемные средства}}$$

(характеризует уровень собственных средств предприятия);

- коэффициент обеспеченности покрытия =

$$= \frac{\text{Собственные оборотные средства}}{\text{Оборотные средства}}$$

(показывает долю собственных оборотных средств в оборотном капитале);

- коэффициент рентабельности продаж =

$$= \frac{\text{Прибыль от реализации}}{\text{Выручка от реализации}}$$

(показывает эффективность основной деятельности предприятия за данный период времени);

- коэффициент оборачиваемости =

$$= \frac{\text{Всего активов}}{\text{Выручка от реализации}}$$

(показывает скорость оборота активов предприятия за определенный период времени).

Рассчитанные значения коэффициентов для потенциального заемщика сравниваются со среднеотраслевыми или нормативными значениями. По результатам сравнения делается вывод о текущем финансовом положении заемщика.

Этап № 2 — анализ деятельности высшего управленческого звена предприятия-заемщика посредством экспертной оценки уровня его профессионализма, компетентности, используемых стратегий и достигнутых результатов.

Этап № 3 — анализ отрасли, к которой принадлежит данное предприятие, а также позиции предприятия на рынке по отношению к конкурентам по отрасли.

Этап № 4. На завершающем этапе оценивается так называемый «страновой риск», т.е. проводится анализ текущей ситуации и перспектив развития в экономической и политической сферах, возможных изменений политики регулирующих органов и т.д.

На основании проведенного комплексного анализа, заемщик относится к одной из групп риска, другими словами, заемщику присваивается кредитный рейтинг.

Кредитный рейтинг представляет собой интегральную оценку финансовой устойчивости и платежеспособности страны, заемщика или отдельного кредитного продукта. Рейтинг выражает мнение агентства относительно будущей способности и намерения заемщика осуществлять выплаты кредиторам на погашение основной суммы задолженности и процентов по ней своевременно и в полном объеме.

Кредитные рейтинги отражают объективную оценку вероятности дефолта и используются для определения характера инвестиций.

После этого оценивается **группа риска кредитного продукта**, на которую влияют следующие:

1. Срок кредитного продукта — точность прогноза финансового состояния предприятия на относительно короткие периоды времени намного выше, чем на продолжительные. Следовательно, чем меньше срок до погашения кредитного продукта, тем ниже риск, и наоборот.

2. Ставка процента.

3. Условия предоставления кредита или ограничения, наложенные на заемщика при предоставлении кредита.

4. Первостепенное влияние на группу риска кредитного продукта оказывает обеспечение по кредиту (гарантирует возврат, как правило, части денежных средств).

ОЦЕНКА КРЕДИТНЫХ РИСКОВ С ПРИМЕНЕНИЕМ НЕЧЕТКОЙ ЛОГИКИ

Рассмотрим метод оценки кредитных банковских рисков с использованием аппарата нечеткой логики, а именно алгоритма нечеткого вывода Мамдани. Для решения поставленной задачи введем лингвистические переменные, описывающие кредитоспособность заемщиков банковских кредитов.

Кредитная история = (безупречная, плохая, нет данных).

Входная переменная — ответ на вопрос «Были ли просроченности у контрагента по оплате основного долга или процентов за ранее полученными кредитами, в т.ч. в других банках?». Варианты — (не было, были, нет данных). Соответствия: «не было – безупречно», «плохо – были», «нет данных – нет данных». Способ получения данных — ответ человека через диалоговое окно. Вид переменной — в данном случае — четкая. Т.е. функция принадлежности принимает значение либо 0, либо 1.

Финансово-экономическое состояние = (стабильное, негативные тенденции, позитивные тенденции).

Вообще, для оценки финансово-экономического состояния используется три показателя: «коэффициент независимости», «моментальная ликвидность», «период оборотности запасов».

Для того, чтобы определить значение лингвистической переменной «Финансово-экономическое состояние», необходимо сравнить значения выбранных коэффициентов за текущий период и за прошедший период, т.е. необходимо вычислить дельту для каждого коэффициента, причем рациональнее считать относительные изменения. Далее для простоты относительные изменения будут называться просто «дельты».

Следует обратить внимание, что при уменьшении показателей «коэффициент независимости» и «период оборотности запасов» рейтинг возрастает, а при уменьшении «моментальной ликвидности» рейтинг падает. Это учтено при составлении функций принадлежности для термов «негативные тенденции» и «позитивные тенденции».

Будем рассматривать «дельту» для каждого показателя как нечеткую переменную. Для нее необходимо ввести функцию принадлежности.

Рассмотрим дельту для каждого коэффициента в отдельности.

1. «**Коэффициент независимости**».

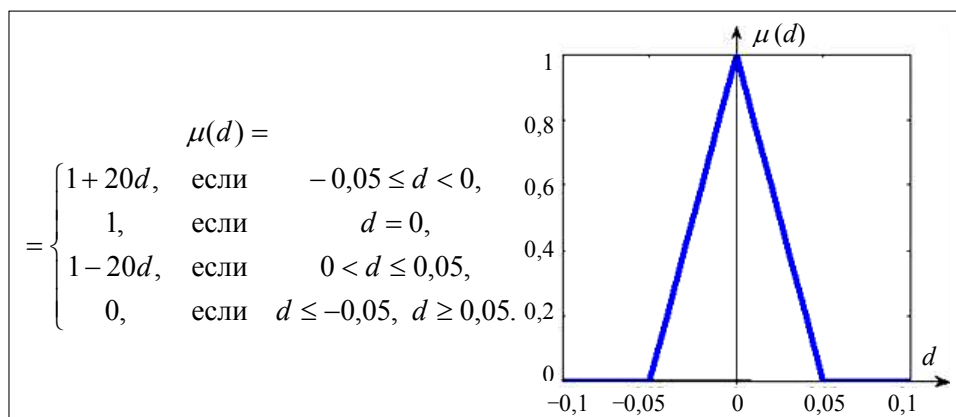


Рис. 1. ФП терма «стабильный»

Сам коэффициент изменяется от 0 до 1. Дельта изменяется от -1 до 1 . Функция принадлежности (ФП) дельты к терму «стабильный» (треугольного вида, ось абсцисс (дельта) пересекает в точке $-0,05$ и $0,05$, а ось ординат в точке 1) (рис. 1):

ФП дельты к терму «негативные тенденции» (рис. 2):

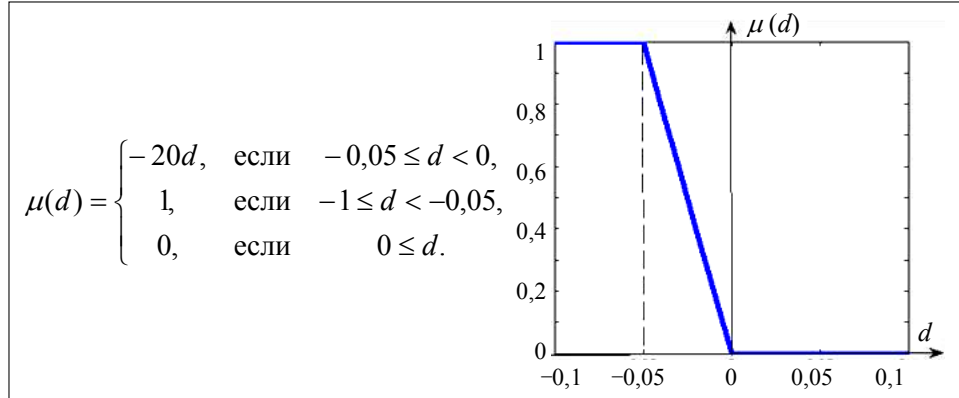


Рис. 2. ФП терма «негативные тенденции»

ФП дельты к терму «положительные тенденции» (рис. 3):

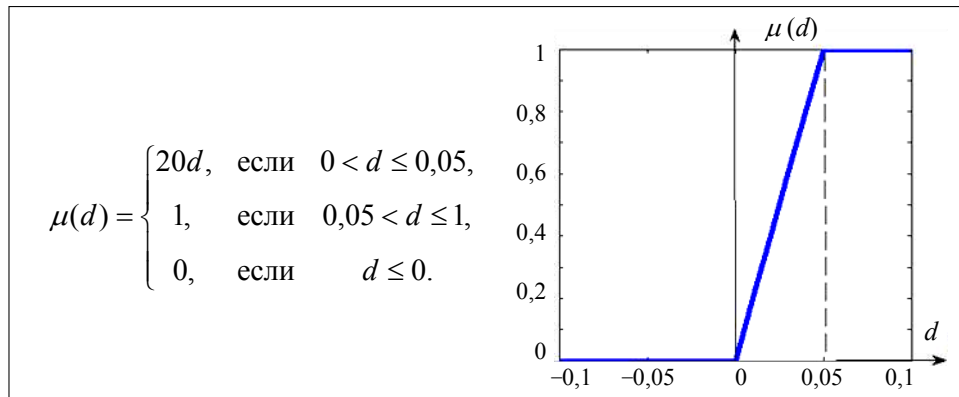


Рис. 3. ФП терма «положительные тенденции»

2. Коэффициент «мгновенной ликвидности». Сам коэффициент изменяется от 0 до 1. Дельта изменяется от -1 до 1 . ФП дельты к терму «стабильный» (если нарисовать, получим треугольник, где ось дельта пересекает в точке $-0,1$ и $0,1$, а ось ординат в точке 1):

$$\mu(d) = \begin{cases} 1+10d, & \text{если } -0,1 \leq d < 0, \\ 1, & \text{если } d = 0, \\ 1-10d, & \text{если } 0 < d \leq 0,1, \\ 0, & \text{если } d \leq -0,1 \text{ или } d \geq 0,1. \end{cases}$$

ФП дельты к терму «негативные тенденции»:

$$\mu(d) = \begin{cases} -10d, & \text{если } -0,1 \leq d < 0, \\ 1, & \text{если } -1 \leq d < -0,1, \\ 0, & \text{если } 0 \leq d. \end{cases}$$

ФП дельты к терму «позитивные тенденции»:

$$\mu(d) = \begin{cases} 10d, & \text{если } 0 < d \leq 0,1, \\ 1, & \text{если } 0,1 < d \leq 1, \\ 0, & \text{если } d \leq 0. \end{cases}$$

3. «Период оборачиваемости запасов». Сам коэффициент изменяется от 0 до 180. Дельта изменяется от -1 до 1. ФП дельты к терму «стабильный» (если нарисовать, то получим треугольник, где ось дельта пересекает в точке -0,1 и 0,1, а ось ординат в точке 1):

$$\mu(d) = \begin{cases} 1+10d, & \text{если } -0,1 \leq d < 0, \\ 1, & \text{если } d = 0, \\ 1-10d, & \text{если } 0 < d \leq 0,1, \\ 0, & \text{если } d \leq -0,1, d \geq 0,1. \end{cases}$$

ФП дельты к терму «позитивные тенденции»:

$$\mu(d) = \begin{cases} -10d, & \text{если } -0,1 \leq d < 0, \\ 1, & \text{если } -1 \leq d < -0,1, \\ 0, & \text{если } 0 \leq d. \end{cases}$$

ФП дельты к терму «негативные тенденции»:

$$\mu(d) = \begin{cases} 10d, & \text{если } 0 < d \leq 0,1 \\ 1, & \text{если } 0,1 < d \leq 1 \\ 0, & \text{если } d \leq 0. \end{cases}$$

4. **Вероятность несвоевременного погашения** = (низкая, средняя, высокая, очень высокая).

Входная переменная — дает ответ на вопрос «вероятность несвоевременного погашения». Варианты — любая вероятность от 0 до 1.

Рассматриваем эту переменную как нечеткую и введем функцию принадлежности. Вероятность изменяется от 0 до 1.

ФП дельты к терму «низкая»:

$$\mu(p) = \begin{cases} 1-20p, & \text{если } 0 \leq p < 0,05, \\ 1, & \text{если } p = 0, \\ 0, & \text{если } 0,05 \leq p. \end{cases}$$

ФП к терму «средняя»:

$$\mu(p) = \begin{cases} 3,33 - 33,33p, & \text{если } 0,07 \leq p < 0,1, \\ 1, & \text{если } 0,05 < p < 0,07, \\ 0, & \text{если } p \geq 0,1, p \leq 0,05. \end{cases}$$

ФП к терму «высокая»:

$$\mu(p) = \begin{cases} 33,33p - 2,33, & \text{если } 0,07 < p \leq 0,1, \\ -5p + 2,5, & \text{если } 0,3 < p \leq 0,5, \\ 1, & \text{если } 0,1 < p \leq 0,3, \\ 0, & \text{если } p \leq 0,07, p \geq 0,5. \end{cases}$$

ФП к терму «очень высокая»:

$$\mu(p) = \begin{cases} 5p - 1,5, & \text{если } 0,3 < p < 0,5, \\ 1, & \text{если } 0,5 \leq p, \\ 0, & \text{если } p \geq 0,3. \end{cases}$$

5. **Риск убытков** = (высокий, средний, низкий).

Входная переменная — (высокий, средний, низкий). Способ получения данных: ответ через диалоговое окно. Вид переменной — в данном случае — четкая. Т.е. функция принадлежности принимает значение либо 0, либо 1.

6. **Состояние документации** = (плохое, нормальное, хорошее).

Входная переменная: ответ на вопрос. Состояние документации (варианты) определено как — (плохое, нормальное, хорошее). Способ получения данных — ответ через диалоговое окно. Вид переменной — четкая.

7. **Банкрот** = (да, нет).

Входная переменная — ответ на вопрос «Признано предприятие банкротом».

Варианты — (да, нет). Способ получения данных — ответ через диалоговое окно. Вид переменной — четкая.

8. **Обеспечение** = (первоклассное, ликвидное, неликвидное, отсутствует, возможность проблем).

Входная переменная — ответ на вопрос «Обеспечение». Варианты — (первоклассное, ликвидное, неликвидное, отсутствует, возможность проблем). Способ получения данных — ответ через диалоговое окно. Вид переменной — четкая. Т.е. Функция принадлежности принимает значение либо 0, либо 1.

9. **Платежеспособность** = (высокая, средняя, низкая).

Входная переменная — ответ на вопрос «Платежеспособность контрагента». Варианты — (высокая, средняя, низкая). Способ получения данных — ответ через диалоговое окно. Вид переменной — в данном случае — нечеткая.

10. **Функция принадлежности для выходной переменной.**

Приведем ФП для выходной переменной — рейтинг заемщика.

$$\text{ФП к терму «класс Д»}: \mu(r) = \begin{cases} 5r + 0,25, & \text{если } -0,5 < r < 1,5, \\ 0, & \text{если } r \leq -0,5, r \geq 1,5, \\ 1, & \text{если } r = 1,5. \end{cases}$$

$$\text{ФП дельты к терму «класс Г»}: \mu(r) = \begin{cases} 5r - 0,75, & \text{если } 1,5 < r < 3,5, \\ 0, & \text{если } r \leq 1,5, r \geq 3,5, \\ 1, & \text{если } r = 3,5. \end{cases}$$

$$\begin{aligned} \text{ФП дельты к терму «класс В»}: \mu(r) &= \begin{cases} 5r - 1,75, & \text{если } 3,5 < r < 5,5, \\ 0, & \text{если } r \leq 3,5, r \geq 5,5, \\ 1, & \text{если } r = 5,5. \end{cases} \\ \text{ФП дельты к терму «класс Б»}: \mu(r) &= \begin{cases} 5r - 2,75, & \text{если } 5,5 < r < 7,5, \\ 0, & \text{если } r \leq 5,5, r \geq 7,5, \\ 1, & \text{если } r = 7,5. \end{cases} \\ \text{ФП дельты к терму «класс А»}: \mu(r) &= \begin{cases} 5r - 3,75, & \text{если } 7,5 < r < 9,5, \\ 0, & \text{если } r \leq 7,5, r \geq 9,5, \\ 1, & \text{если } r = 9,5. \end{cases} \end{aligned}$$

Соответствующий вид ФП для различных классов рейтинга приводятся на рис. 4.

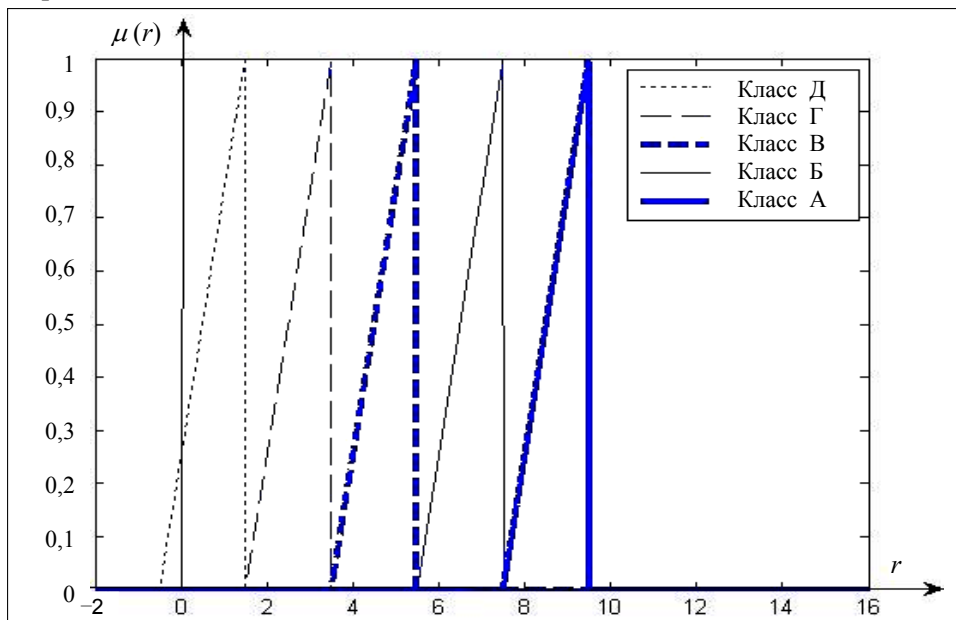


Рис. 4. ФП классов рейтинга заемщиков

ПРАВИЛА ДЛЯ ОЦЕНКИ РЕЙТИНГА. ТЕКСТОВАЯ ФОРМУЛИРОВКА ЭКСПЕРТОВ

Определение класса платежеспособности контрагента

Для определения класса платежеспособности контрагента используются следующие правила.

1. Класс контрагента (без учета уровня обеспечения по кредитной операции) определяется на основании значения итоговой рейтинговой оценки финансово-экономического состояния контрагента т.е. контрагенту присваивается класс (без учета уровня обеспечения по кредитной операции), отве-

чающий интервалу значений, в который попадает значение итоговой рейтинговой оценки контрагента, рассчитанной соответственно пункту.

2. Следующим шагом есть классификация контрагента с учетом уровня обеспечения по кредитной операции.

2.1. К классу «А» контрагент должен быть отнесен, если полностью выполняются все следующие условия:

- класс, определенный согласно пункту 1, равен «А»;
- высшее руководство контрагента имеет отличную деловую репутацию, кредитная история контрагента — безупречная;
- нет никаких свидетельств возможных задержек с возвращением основной суммы долга и / или с уплатой процентов.

2.2. К классу «Б» контрагент должен быть отнесен, если полностью выполняются все следующие условия:

- класс, определенный согласно пункту 1, равен «Б»;
- кредитная история безупречная;
- почти отсутствуют свидетельства о возможных задержках с возвращением основной суммы долга и/или с уплатой процентов;
- анализ коэффициентов финансового состояния контрагента может свидетельствовать об отрицательных тенденциях в его деятельности;
- недостатки в деятельности контрагента, которые принадлежат к классу «Б», должны быть лишь потенциальными.

2.3. К классу «В» контрагент должен быть отнесен, если полностью выполняются все следующие условия:

- класс, определенный согласно пункту 1, равен «В»;
- поступление средств и платежеспособность контрагента свидетельствуют о вероятности несвоевременного погашения кредитной задолженности в полной сумме и в сроки, предусмотренные кредитным договором, если недостатки не будут устранены. Одновременно наблюдается возможность исправления ситуации и улучшение финансового состояния контрагента;
- обеспечение кредитной операции должны быть ликвидным.

2.4. К классу «Г» контрагент должен быть отнесен, если выполняются следующие условия:

- класс, определенный согласно пункту 1, равен «Г»;
- наблюдается нестабильность финансово-экономического состояния контрагента, есть высокий риск значительного убытка, вероятность полного погашения кредитной задолженности и процентов (комиссий) по ней низкая;
- имеют место проблемы, касающиеся состояния обеспечения за кредитными операциями, необходимой документации, удостоверяющей наличие и ликвидность залога, и т.п.

2.5. К классу «Д» контрагент должен быть отнесен, если выполняются следующие условия:

- класс, определенный согласно пункту 1, равен «Д»;
- вероятности выполнения обязательств контрагентом Банка практически нет. К этому классу принадлежит контрагент Банка, который признан банкротом в установленном действующим законодательством порядком;

• обеспечение кредитной операции сомнительное (неликвидное) или вообще отсутствующее (без обеспечения).

2.6. Если уровень обеспечения (условие 3 в пунктах 2.1–2.5) по отдельной кредитной операции не отвечает условиям определенного класса, то контрагента следует отнести на класс ниже, а если обеспечение перво-классное (залог имущественных прав на денежные депозиты Банка; безусловные гарантии правительств стран категории «А», Международного банка реконструкции и развития, Европейского банка реконструкции и развития, банков с рейтингом не ниже, чем «инвестиционный класс»; обеспеченные гарантии банков Украины), то класс контрагента можно на уровень повысить.

2.7. Если Банк осуществляет долгосрочное кредитование инвестиционного проекта, расчет эффективности которого (в том числе дохода, который планируется получить от его реализации) обеспечивает погашение кредита и процентов/комиссий, за ним соответственно условиям кредитной сделки, то класс контрагента может быть повышен на один уровень.

Правила вывода

1. Для оценки класса заемщика на основе рейтинга используются следующие правила вывода:

1.1. Если $-0,50 \leq \text{рейтинг} < 1,50$, то рейтинг-класс = Д.

1.2. Если $1,50 \leq \text{рейтинг} < 3,50$, то рейтинг-класс = Г.

1.3. Если $3,50 \leq \text{рейтинг} < 5,50$, то рейтинг-класс = В.

1.4. Если $5,50 \leq \text{рейтинг} < 7,50$, то рейтинг-класс = Б.

1.5. Если $7,50 \leq \text{рейтинг} \leq 9,50$, то рейтинг-класс = А.

2. Для окончательного вывода относительно класса заемщика используются следующие правила вывода.

2.1. Если рейтинг-класс — А, а репутация руководства контрагента — безупречная, кредитная история — безупречная, несвоевременное погашение основной части долга — почти исключено, несвоевременное погашение % — почти исключено, то класс А.

2.2. Если рейтинг-класс — Б, а кредитная история — безупречная, несвоевременное погашение основной части долга — низкая, несвоевременность погашения % — почти исключено, финансово-экономическое состояние контрагента — негативные тенденции в деятельности, недочеты в деятельности контрагента — потенциальные, то класс Б.

2.3. Если рейтинг-класс — В, а несвоевременное погашение основной части долга в сроки и платежеспособность контрагента — средняя, финансово-экономическое состояние контрагента — позитивные тенденции, то класс В.

2.4. Если рейтинг-класс — Г, а финансово-экономическое состояние контрагента — негативные тенденции в деятельности, то класс Г.

2.5. Если рейтинг-класс — Г, а риск значительных убытков — высокий; вероятность полного погашения кредитной задолженности и процентов/комиссий — низкая, то класс Г.

2.6. Если рейтинг-класс — Г, а вероятность несвоевременного погашения кредитной задолженности и процентов/комиссий — высокая, то класс Г.

2.7. Если рейтинг-класс — Г, а состояние необходимой документации, которая подтверждает наличие и ликвидность залога — плохо, то класс Г.

2.8. Если рейтинг-класс — *Д*, а признанный банкротом контрагент Банка — да, то класс *Д*.

2.9. Если рейтинг-класс — *Д*, а вероятность несвоевременного погашения контрагентом — очень высокая, то класс *Д*.

3.1. Если класс — *В*, а обеспечение кредитной операции — ликвидное, то класс **В**.

3.2. Если рейтинг-класс — *Г*, а обеспечение по кредитным операциям — вероятность проблем, то класс **Г**.

3.3. Если класс — *Д*, а обеспечение кредитной операции — неликвидное, то класс *Д*.

3.4. Если класс — *Д*, а обеспечение кредитной операции — отсутствует, то класс *Д*.

3.5. Если класс — *Г*, а обеспечение кредитной операции — вероятность проблем, то класс *Д*.

3.6. Если класс — *В*, а обеспечение кредитной операции — вероятность проблем, то класс **Г**.

3.7. Если класс — *Г*, а обеспечение кредитной операции — неликвидное, то класс *Д*.

3.8. Если класс — *Г*, а обеспечение кредитной операции — отсутствует, то класс *Д*.

3.9. Если класс — *В*, а обеспечение кредитной операции — первоклассное, класс **Б**.

3.10. Если класс — *Г*, а обеспечение кредитной операции — первоклассное, то класс **В**.

3.11. Если класс — *Д*, а обеспечение кредитной операции — первоклассное, то класс **Г**.

ПРИМЕНЕНИЕ МЕТОДИКИ ОПРЕДЕЛЕНИЯ КРЕДИТОСПОСОБНОСТИ ЗАЕМЩИКА НА ОСНОВЕ НЕЧЕТКОЙ ЛОГИКИ

Рассмотрим применение вышеизложенной методики оценки кредитоспособности заемщика с использованием аппарата нечеткой логики на следующих примерах.

Оценка кредитоспособности ЗАО «АвтоЧасть»

Исходные данные.

Данное предприятие занимается оптовой поставкой автозапчастей ведущих производителей, таких как Universal Industries, Fenox, Jestic, Прамо и другие.

- В данном Банке обслуживается не полностью, более двух месяцев. Известность предприятия — локальная на уровне стран. Стабильный узкий круг потребителей в Украине, России, Эстонии.

- Опыт работы в основной сфере деятельности — 10 лет.

Данные составленные на конец октября 2005 года.

1. Обеспечение — не вызывает сомнений.
2. Коэффициент независимости (КН) — 0,36.
3. Маневренность собственных средств (коэффициент маневренности — КМ) — 0,75.
4. Сумма среднемесячной задолженности составляет 342 тыс. грн.
5. Доходы составляют 514 тыс. грн.

6. Поступление денежных средств на счета (количество раз) — 4.
7. Среднемесячные поступления денежных средств (СПДС) составляют 100 тыс. грн.
8. Среднемесячные поступления чистых денежных средств (ЧДС) составляют 80 тыс. грн.
9. Рентабельность продаж (РП) — 0,14.
10. Рентабельность активов (РА) — 0,067.
11. Коэффициент общей ликвидности (КЛ) — 0,7.
12. Коэффициент мгновенной ликвидности (КЛ1) — 0,14.
13. Коэффициент текущей ликвидности (КЛ2) — 1,6.
14. Оборачиваемость запасов (ОЗ) — 40.
15. Оборачиваемость краткосрочной дебиторской задолженности (ОДЗ) — 50.
16. Оборачиваемость краткосрочной кредиторской задолженности (ОКЗ) — 50.
17. Текущая дебиторская задолженность (ТДЗ) составляет 300 тыс. грн.
18. Краткосрочная кредиторская задолженность (ККЗ) составляет 401 тыс. грн.
19. Собственные активы (СА) составляют 1144 тыс. грн.
20. Оборотные активы (ОА) составляют 655 тыс. грн.
21. Темп роста валовой прибыли (ТРВП) — 110 %.
22. Темп роста объема реализации (ТРОР) 112 %.
23. Темп роста суммы активов (ТРСА) — 114 %.
24. Значение рейтинга за предыдущий период — 6.
25. Срок кредитной операции (СКО) — 36 мес.
26. Просрочек по оплате нет.
27. Вероятность несвоевременного погашения основной суммы (ВНПОС) — 0,05.
28. Вероятность несвоевременного погашения процентов (ВНПП) — 0,05 %.

Основные финансово-экономические показатели предприятия, кредитный рейтинг, обеспеченность кредитной операции и правило вывода итогового рейтинга предприятия приводятся в табл. 1.

Таблица 1. Результаты работы программы для оценки кредитоспособности ЗАО «АвтоЧасть»

Параметр	Значение (рейтинг)	Параметр	Значение (рейтинг)	Итоговый параметр	Значение итогового параметра
СА	1144 (10)	ОДЗ	50 (9)	Итоговая рейтинговая оценка	5,26
КН	0,36 (9)	ЧДС	45,801		
КМ	0,75 (8)	ОКЗ	50 (9)	Окончательный класс платежеспособности	В
КЛ	0,7 (3)	КДК	0,748		
КЛ1	0,14 (8)	ЗП	0 (0)	Кредитный рейтинг с учетом кредитного обеспечения	В
КЛ2	1,6 (10)	СП.Н/Р	0,19 (2)		
РП	0,14 (9)	К	0,23 (0)	Лимит безопасного кредитования на 24 мес.	352,92 тыс. грн
РА	0,067 (8)	СПДС	4 (0)		

Поскольку рейтинг-класс = *B* и кредитная история — безупречная и вероятность несвоевременного погашения кредита и вероятность несвоевременного погашения процента по кредиту — низкие, то класс заемщика = **B**.

Оценка кредитоспособности ООО «Блинок»

ООО «Блинок» — кафе быстрого обслуживания. Основной вид деятельности включает изготовление и продажу блинчиков с начинкой.

Степень известности — низкая на уровне Печерского района города Киева. Круг потребителей нестабильный.

- Обслуживание в Банке — полностью, обращается впервые.
- Опыт работы по основному виду деятельности — 0,5 лет.

Данные составлены на конец февраля 2006 года.

1. Обеспечение — отсутствует.
2. Коэффициент независимости — 0,5.
3. Маневренность собственных средств — 0,25.
4. Сумма среднемесячной задолженности составляет 42 тыс. грн.
5. Доходы составляют 114 тыс. грн.
6. Поступление денежных средств на счета (количество раз) — 1.
7. Среднемесячные поступления денежных средств составляет 15 тыс. грн.
8. Среднемесячные поступления чистых денежных средств составляет 5 тыс. грн.
9. Рентабельность продаж — 0,15.
10. Рентабельность активов — 0,06.
11. Коэффициент общей ликвидности — 0,3
12. Коэффициент мгновенной ликвидности — 0,18.
13. Коэффициент текущей ликвидности — 0,2.
14. Оборачиваемость запасов — 20 дней.
15. Краткосрочная дебиторская задолженность составляет 50 тыс. грн.
16. Краткосрочная кредиторская задолженность составляет 50 тыс. грн.
17. Текущая дебиторская задолженность составляет 10 тыс. грн.
18. Краткосрочная кредиторская задолженность составляет 10 тыс. грн.
19. Наибольший удельный вес одного кредитора составляет 80 %.
20. Собственные активы составляют 644 тыс. грн.
21. Оборотные активы составляют 305 тыс. грн.
22. Темп роста валовой прибыли составляет 105 %.
23. Темп роста объема реализации составляет 101 %.
24. Темп роста суммы активов — 101 %.
25. Начение рейтинга за предыдущий период равно 0.
26. Срок кредитной операции 24 месяца.
27. Отсрочки по оплате нет.
28. Вероятность несвоевременного погашения основной суммы 0,5.
29. Вероятность несвоевременного погашения процентов — 0,5.

Соответствующие финансово-экономические показатели и вывод итогового рейтинга для предприятия «Блинок» приводятся в табл. 2.

Таблица 2. Результаты работы программы для оценки кредитоспособности ООО «Блинок»

Параметр	Значение	Параметр	Значение	Итоговый параметр	Значение итогового параметра
СА	644 (10)	ОДЗ	50 (9)	Итоговая рейтинговая оценка	5,36
КН	0,5 (8)	ЧДС	3,27		
КМ	0,25 (3)	ОКЗ	50 (9)	Окончательный класс платежеспособности	В
КП	0,3 (1)	КДК	1 (6)		
КЛ1	0,18 (9)	ЗП	1 (5)	Кредитный рейтинг с учетом кредитного обеспечения	Г
КЛ2	0,2 (3)	СП.Н/Р	0,94 (10)		
Рп	0,05 (7)	К	0,83 (4)	Лимит безопасного кредитования на 24 мес.	59,99 тыс. грн
Ра	0,06 (8)	СПДС	1 (0)		

Поскольку *рейтинг-класс = В* и *кредитная история — безупречная и вероятность несвоевременного погашения кредита и процента по кредиту — низкие*. Поскольку Класс *В*, рейтинг-класс = *В*, а кредитное обеспечение отсутствует, то нечеткий класс = *Г*.

ЗАКЛЮЧЕНИЕ

В статье рассмотрены проблемы оценки кредитных рисков заемщиков банков. Изложен классический метод оценки кредитных рисков на основе определения кредитного рейтинга заемщиков, а также оценки кредитного обеспечения.

Предложен метод определения кредитоспособности заемщиков на основе использования систем с нечеткой логикой (алгоритм нечеткого вывода Мамдани.). Этот метод позволяет учитывать неопределенность исходной информации о заемщиках. Метод наиболее эффективен в условиях неполноты и неопределенности информации, наличия лингвистических переменных и качественных критериев.

Сформулированы правила нечеткого вывода для оценки кредитного рейтинга и общего класса заемщиков. Приводятся примеры применения изложенного нечеткого метода.

ЛИТЕРАТУРА

1. Зайченко Ю.П. Нечеткие модели в интеллектуальных системах. — Киев: Издательский дом «Слово», 2008. — 344 с.
2. Зайченко Ю.П. Анализ кредитных рисков с использованием систем с нечеткой логикой и нечетких нейронных сетей. — Матеріали XI Міжнар. науково-техн. конф. «Системний аналіз та інформаційні технології САІТ-2009». 26–30 травня 2009 р. Київ, Україна. — 305 с.

Поступила 11.06.2009

**ПРОГНОЗИРОВАНИЕ И ВЕРОЯТНОСТНАЯ ОЦЕНКА
ДОЛГОВЕЧНОСТИ ТЕХНИЧЕСКИХ ОБЪЕКТОВ**

С.С. ФЕДИН

На основе методологии марковского анализа надежности предложен вероятностный подход к прогнозированию и оценке долговечности технических объектов. Разработана математическая модель прогнозирования долговечности с использованием статистических данных, характеризующих начальное и конечное распределение продолжительности срока службы технических объектов.

ВВЕДЕНИЕ

Актуальной задачей в области оценки и обеспечения эксплуатационной надежности технических объектов (ТО) является прогнозирование их долговечности. Теоретические основы решения этой задачи рассмотрены в работах А.С. Пронилова [1], Б.В. Гнеденко [2], Р. Барлоу [3] и многих других ученых. Прогнозирование долговечности позволяет получить вероятностные оценки ожидаемой продолжительности срока службы ТО для принятия обоснованных решений при установлении регламентов технического обслуживания и планового контроля ТО.

В соответствии с ДСТУ 2860-94 долговечность характеризует свойство объекта выполнять требуемые функции до наступления предельного состояния при установленной системе технического обслуживания и ремонта [4]. Показателями долговечности являются: средний ресурс, гамма-процентный ресурс, средний срок службы и гамма-процентный срок службы. Для прогнозирования показателей долговечности в большинстве случаев применяются фактографические методы вероятностной оценки, основанные на априорной информации о характере статистического распределения исследуемой случайной величины, например, продолжительности срока службы ТО.

Одним из наиболее универсальных и эффективных методов вероятностной оценки долговечности ТО является разработка и применение математических моделей на основе методологии марковского анализа надежности. В соответствии с ДСТУ 2861-94 и ГОСТ Р 51901.15-2005 марковский анализ относится к аналитическим методам обеспечения надежности и применяется для оценки вероятностных характеристик ТО на этапе эксплуатации [5, 6]. Методы марковского анализа основаны на использовании диаграммы

состояний и переходов, применение которой позволяет моделировать зависимость изменения во времени показателей надежности ТО и является графическим представлением функционирования восстанавливаемого ТО. При этом ТО или система рассматривается как совокупность элементов, каждый из которых может существовать только в работоспособном или неработоспособном состоянии. Система в целом может существовать в различных состояниях, каждое из которых определяется определенной комбинацией работоспособного и неработоспособного состояний ее элементов. В момент отказа или восстановления элемента система последовательно переходит из одного состояния в другое.

Следует отметить, что методы марковского анализа позволяют графически отображать процесс отказов/восстановлений, который представляется в виде случайных переходов во времени от одного состояния к другому и могут применяться к ТО, в которых некоторые или все элементы являются невосстанавливаемыми [6].

Цель работы — разработка модели прогнозирования и вероятностной оценки долговечности ТО на основе методологии марковского анализа надежности.

ПОСТАНОВКА ЗАДАЧИ И РАЗРАБОТКА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ПРОГНОЗИРОВАНИЯ ДОЛГОВЕЧНОСТИ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

Методология марковского анализа надежности позволяет разработать универсальную математическую модель, которая может быть применена для формализации динамики переходов ТО от состояния к состоянию. В большинстве случаев такую модель называют моделью дискретных состояний с непрерывным временем. В соответствии с этим способом представления изменения состояний ТО применяют метод анализа пространства состояний [6]. Анализ пространства используют при исследовании надежности систем с резервированием или систем, отказ которых зависит от последовательных событий, а также для систем со сложными стратегиями технического обслуживания, к которым относят приоритетное восстановление или проблемы организации очереди. При этом необходимо отметить, что такие классические методы анализа надежности как анализ дерева неисправностей или метод структурной схемы надежности не учитывают сложные стратегии технического обслуживания, например стратегию приоритетного восстановления. Применение методов марковского анализа позволяет реализовать возможность моделирования стратегий технического обслуживания ТО с учетом последовательности возникновения многократных отказов.

При разработке марковской модели прогнозирования долговечности ТО анализируемой величиной является износ, выраженный в процентах от предельно допустимого значения (предельного состояния). Допустимый диапазон значений этой величины делится на несколько интервалов, рассматриваемых как отдельные состояния, в каждом из которых с некоторой вероятностью находится ТО, последовательно переходя из одного состояния в другое. Количество состояний определяется желаемой точностью вероятностных оценок показателей долговечности.

Так, например, если для типового состояния системы выбрать 10%-й интервал дискретизации, то состоянию X_0 будет соответствовать интервал 0% – 10% износа, состоянию X_1 — интервал 10% – 20% износа и так далее. Состояние X_n будет характеризовать 100%-й износ.

Модель износа может быть представлена ориентированным графом, в котором при непрерывном времени и определенной интенсивности износа λ и/или интенсивности восстановления μ осуществляются случайные переходы из одного состояния в другое согласно направлениям дуг (рис. 1).

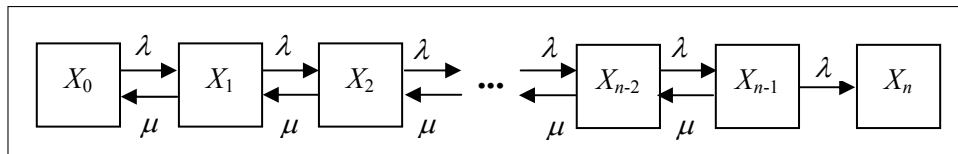


Рис. 1. Ориентированный граф: X_i ($i = 0, 1, 2, \dots, n$) — состояния; λ и μ — интенсивности потоков (λ — интенсивность износа; μ — интенсивность восстановления)

Фактически модель износа представляет собой конечную цепь из $n+1$ состояний, в которой переходы из состояния X_k ($k \neq 0, k \neq n$) возможны только в предшествующее X_{k-1} или последующее состояние X_{k+1} . Каждое состояние соответствует выбранному диапазону степеней износа ТО, выраженному в процентах от продолжительности срока службы ТО. Переход из состояния X_0 возможен только в состояние X_1 . Состояние X_n соответствует полному износу и не имеет выходов. Параметры λ характеризуют динамику износа ТО под воздействием среды. Параметры μ представляют способность ТО к восстановлению. Это свойство обычно не является необходимым, однако в общем случае для полноты модели его желательно учитывать. Поскольку состояния представляют собой диапазоны износа, выраженные в процентах от величины, измеряемой в единицах времени, интенсивности потоков λ и μ полагаются не зависящими от номера состояния. Математическое описание переходов между состояниями характеризует влияние среды на ТО с учетом следующих предположений:

- переходы между состояниями являются статистически независимыми событиями;
- интенсивность отказов λ и интенсивность восстановлений μ постоянны;
- вероятности перехода из одного состояния в другое в интервале времени Δt (при $\Delta t \rightarrow 0$) задаются в виде $\lambda \Delta t$ и/или $\mu \Delta t$.

Предполагается также, что для переходов между состояниями соответствующих каждой дуге графа, выполняются следующие два свойства пуассоновских потоков событий:

- ординарность (вероятность появления двух и более событий в течение малого интервала времени намного меньше, чем вероятность появления за это же время одного события);
- отсутствие последствия (количества событий, попадающих в два непересекающихся интервала времени, не зависят друг от друга).

Из сделанных предположений следует, что будущее состояние системы зависит только от существующего состояния, а не от того, как система оказалась в этом состоянии.

Динамика изменения во времени вероятностей нахождения системы в различных состояниях в соответствии с ориентированным графом (рис. 1) описывается системой обыкновенных дифференциальных уравнений:

$$\begin{cases} \frac{d p_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t); \\ \dots \\ \frac{d p_k(t)}{dt} = -(\lambda + \mu)p_k(t) + \lambda p_{k-1}(t) + \mu p_{k+1}(t) \quad (k = 1, 2, \dots, n-1); \\ \dots \\ \frac{d p_n(t)}{dt} = \lambda p_{n-1}(t). \end{cases}$$

Для нахождения вероятностей достижения указанных состояний необходимо выполнить интегрирование системы дифференциальных уравнений при начальных условиях:

$$p_0(0) = 1, \quad p_1(0), \dots, p_n(0); \quad \sum_{k=0}^n p_k(0) = 1,$$

а также с учетом нормализующего условия, которое должно выполняться в любой момент времени:

$$\sum_{k=0}^n p_k(t) = 1.$$

ПРОВЕДЕНИЕ ВЫЧИСЛИТЕЛЬНОГО ЭКСПЕРИМЕНТА ДЛЯ ПРОГНОЗИРОВАНИЯ ДОЛГОВЕЧНОСТИ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

При проведении имитационного вычислительного эксперимента предполагалось, что в момент начала эксплуатации ТО существует начальный износ, который описывается усеченным нормальным распределением с неотрицательным средним¹. Данное предположение обусловлено тем, что приемлемыми являются только значения износа ТО из диапазона от 0% до 100%. Для обеспечения равенства полной вероятности единице усеченное распределение нормируется [7]. Пример начального распределения износа ТО показан на рис. 2 в виде гистограммы.

В результате воздействия среды в течение определенного времени начальное распределение трансформируется в некоторое текущее распределение. При этом исходной информацией для получения вероятностных оценок являются статистические данные о продолжительности срока службы ТО, представленные на гистограмме (рис. 3).

¹ Параметры различных образцов одного и того же ТО могут отличаться из-за наличия производственных дефектов.

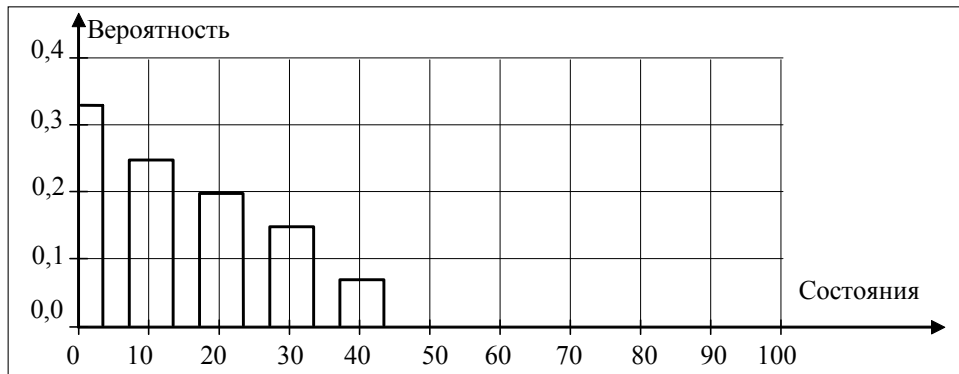


Рис. 2. Гистограмма распределения начального износа 100 однотипных ТО: ось абсцисс — состояния начального износа в %, ось ординат — вероятности нахождения ТО в определенном состоянии

За единицу измерения времени выбран час, а имитационная модель для исследования развития износа представлена цепью из 11-ти состояний, в которой каждому отдельному состоянию соответствует 10%-й интервал износа (рис. 1). Способность ТО к восстановлению в модели не рассмотрена. Динамика изменения во времени вероятностей нахождения ТО в 11-ти различных состояниях описывается системой обыкновенных дифференциальных уравнений:

$$\begin{cases}
 \frac{d p_0(t)}{dt} = -\lambda p_0(t), \\
 \frac{d p_1(t)}{dt} = -\lambda p_1(t) + \lambda p_0(t), \\
 \frac{d p_2(t)}{dt} = -\lambda p_2(t) + \lambda p_1(t), \\
 \frac{d p_3(t)}{dt} = -\lambda p_3(t) + \lambda p_2(t), \\
 \frac{d p_4(t)}{dt} = -\lambda p_4(t) + \lambda p_3(t), \\
 \frac{d p_5(t)}{dt} = -\lambda p_5(t) + \lambda p_4(t), \\
 \frac{d p_6(t)}{dt} = -\lambda p_6(t) + \lambda p_5(t), \\
 \frac{d p_7(t)}{dt} = -\lambda p_7(t) + \lambda p_6(t), \\
 \frac{d p_8(t)}{dt} = -\lambda p_8(t) + \lambda p_7(t), \\
 \frac{d p_9(t)}{dt} = -\lambda p_9(t) + \lambda p_8(t), \\
 \frac{d p_{10}(t)}{dt} = \lambda p_9(t).
 \end{cases} \quad (1)$$

Оценка параметров износа ТО сводится к вычислению интенсивности отказов λ . В качестве оценок свободного параметра λ используются значения, обеспечивающие наилучшее соответствие наблюдаемых и ожидаемых частот попадания в определенное состояние ТО в заданные моменты времени. Ожидаемая частота попадания в j -е состояние равна $p_j N$, где p_j — вероятность нахождения в этом состоянии, N — количество ТО в выборке.

Ожидаемые вероятности нахождения ТО в различных состояниях износа определены методом численного интегрирования системы дифференциальных уравнений (1)². Соответствующие наблюдаемые частоты F_j определены по распределениям продолжительностей сроков службы текущего распределения (рис. 3). Для нахождения оценки параметра интенсивности отказов λ применен метод минимума χ^2 , основанный на сравнении наблюдаемых и ожидаемых значений частотного распределения износа для выборки ТО объема N :

$$\chi^2 = \sum_{j=0}^n \frac{(F_j - p_j N)^2}{p_j N}.$$

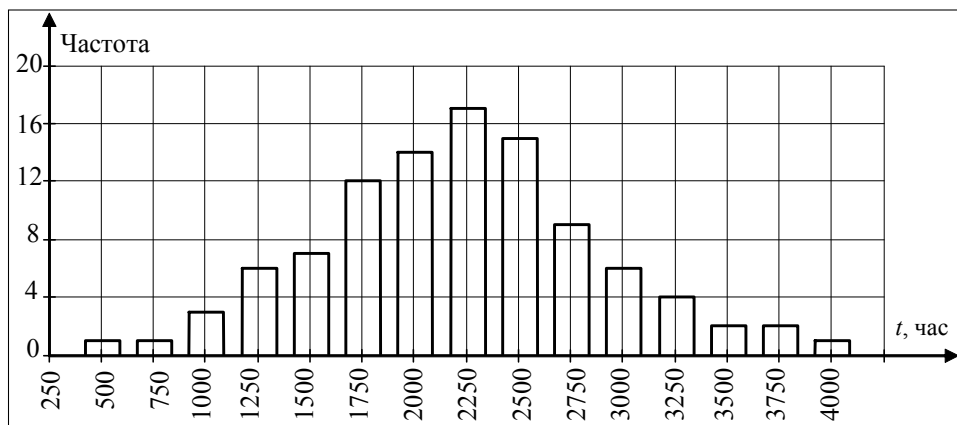


Рис. 3. Гистограмма наблюдаемых частот продолжительности срока службы 100 однотипных ТО

В результате проведенных расчетов установлено, что минимум χ^2 соответствует значению 8,069 при интенсивности отказов $\lambda = 0,00409$ (рис. 4).

Мера согласованности наблюдаемых и ожидаемых частот продолжительностей срока службы ТО определена в системе статистического анализа данных STATISTICA на основе уровня значимости, который характеризует вероятность превышения указанного значения статистики χ^2 и представляет оцененную меру уверенности в достоверности полученного результата (таблица).

² Применялся метод Рунге-Кутты 2-го порядка, реализованный в универсальной системе математического моделирования MathCAD.

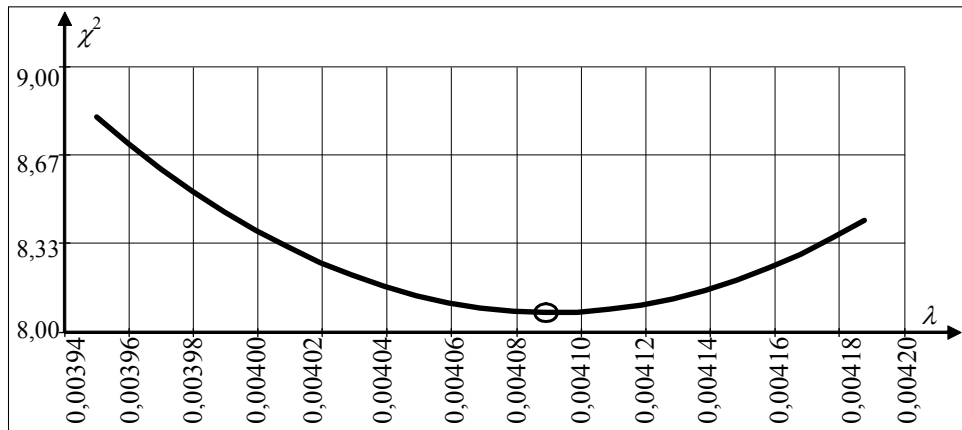


Рис. 4. График значений критерия χ^2 в зависимости от интенсивности отказов λ

Таблица. Мера согласованности и оценки параметров модели прогнозирования долговечности ТО

Модель	Статистика, χ^2	Количество степеней свободы, (df)	Уровень значимости, p	Интенсивность отказов, λ	Среднее время 10%-го износа, $T_{ср} = 1/\lambda$, час
Начальный износ ТО с нулевым средним	8,069	14	0,885	0,00409	244

Гистограмма прогнозируемых частот продолжительностей сроков службы 100 однотипных ТО показана на рис. 5.

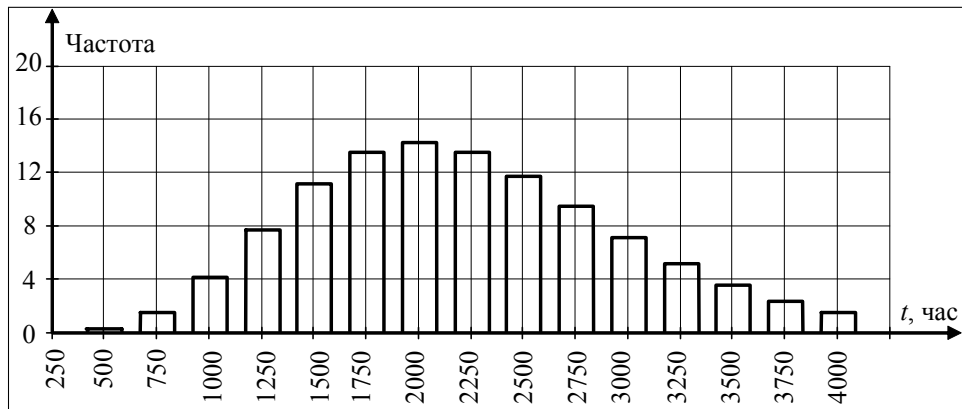


Рис. 5. Прогнозируемые частоты продолжительностей срока службы 100 однотипных ТО

На рис. 6 показаны графики вероятностей нахождения ТО в состояниях 0–10%, 10%–20%, 20%–30% износа, а также график 100%-го износа, применение которого позволяет прогнозировать ожидаемый срок службы ТО.

В качестве конечного результата получены сводные вероятностные графики, позволяющие принимать обоснованные решения относительно регламентов технического обслуживания и планового контроля ТО (рис. 7).

Сводные графики представляют суммарные вероятности продолжительности срока службы при достижении ТО состояний 50%, 70% и 90% износа от предельного состояния.

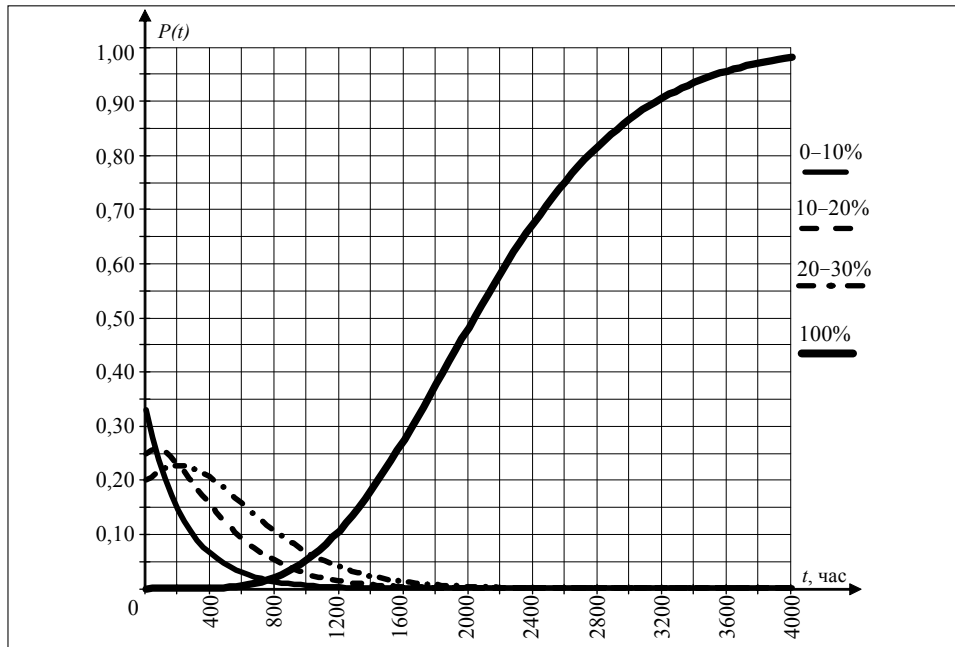


Рис. 6. Вероятности $P(t)$ нахождения ТО внутри указанных интервалов износа как функции времени

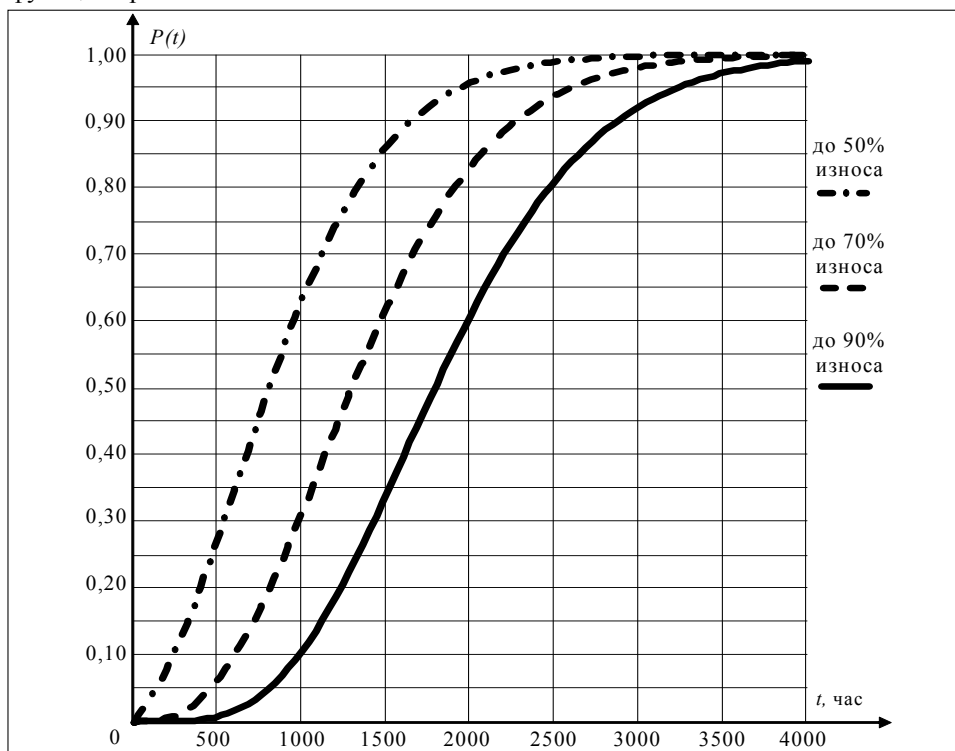


Рис. 7. Вероятности $P(t)$ долговечности изделий при достижении ТО состояний 50%, 70% и 90% износа как функции времени

Применение вероятностных графиков позволяет утверждать, что с вероятностью $P(t) = 0,95$ продолжительность срока службы ТО составит 2000 часов, 2660 часов и 3300 часов при достижении ТО состояний 50%, 70% и 90% износа соответственно.

ВЫВОДЫ

На основе теории цепей Маркова с дискретными состояниями и непрерывным временем разработана математическая модель прогнозирования и вероятностной оценки долговечности ТО, идентификация параметров которой осуществлена с использованием метода минимума χ^2 .

Достоверность полученных результатов прогнозирования подтверждена статистической оценкой меры согласованности наблюдаемых и прогнозируемых частот продолжительностей срока службы ТО.

Представленный подход может быть эффективно применен на этапе эксплуатации ТО и является инструментом для принятия обоснованных решений относительно регламентов технического обслуживания и планового контроля ТО.

ЛИТЕРАТУРА

1. *Надежность и долговечность машин и оборудования. Опыт и теоретические исследования* / Под ред. А.С. Проникова. — М.: Издательство стандартов, 1972. — 315 с.
2. *Гнеденко Б.В., Беляев Ю.К., Соловьев А.Д.* Математические методы в теории надежности. — М.: Наука, 1965. — 524 с.
3. *Барлоу Р., Прошан Ф.* Статистическая теория надежности и испытания на безотказность. — М.: Наука, 1985. — 327 с.
4. *ДСТУ 2860-94.* Надійність техніки. Терміни та визначення. Чинний від 01.01.96. — Київ: Держстандарт України, 1994. — 90 с.
5. *ДСТУ 2861-94.* Надійність техніки. Аналіз надійності. Основні положення. Чинний від 08.12.94. — Київ: Держстандарт України, 1994. — 32 с.
6. *ГОСТ Р 51901.15-2005.* Менеджмент риска. Применение марковских методов. Введ. 01.02.2006. — М.: Федеральное агентство по техническому регулированию и метрологии, 2005. — 16 с.
7. *Матвеевский В.Р.* Надежность технических систем. Учебное пособие. — М.: Московский государственный институт электроники и математики, 2002. — 113 с.

Поступила 10.12.2009

АНАЛИЗ ИСПОЛЬЗОВАНИЯ НЕЙРОСЕТЕЙ ДЛЯ ДИАГНОСТИКИ РАКА ШЕЙКИ МАТКИ ПО МУЛЬТИСПЕКТРАЛЬНОМУ ИЗОБРАЖЕНИЮ

Е.Н. МАЛЫШЕВСКАЯ

Рассмотрена проблема диагностики рака шейки матки с помощью нейронных сетей по мультиспектральному изображению. Для диагностики использованы нейронные сети с обратным распространением ошибки и сети с радиальным базисом. Приведены результаты экспериментальных исследований.

ВВЕДЕНИЕ

В данной работе рассмотрена возможность диагностики рака шейки матки. Такая задача продиктована необходимостью ранней диагностики заболевания используя компьютерную систему, которая поможет врачу определить области с большим риском возникновения раковой трансформации ткани. Данная система базируется на утверждении, что оптические свойства здоровой клетки отличаются от свойств больной и это отличие более выражено, чем отличие клеток разных людей. В медицинском университете Аризоны (США) была внедрена оптическая система, в которой кроме обычного кольпоскопического тестирования проводилось снятие мультиспектральных электронных изображений тканей шейки матки. Были сделаны мультиспектральные снимки 108 пациенткам. Параллельно, те же самые пациентки прошли осмотр у врача, который состоял в том, что была сделана биопсия из определенных участков и отослана врачу-патологу на анализ. Участки, из которых была сделана биопсия, были обозначены на снимке и результаты биопсии были сопоставлены с указанными участками. Эта информация использовалась для создания алгоритма распознавания зон риска заболевания на основе сопоставления мультиспектрального снимка, результатов биопсии и вывода врача. На основе полученной информации создана система позволяющая проводить раннюю диагностику заболевания [1].

ПОСТАНОВКА ЗАДАЧИ

По данным медицинской статистики рак шейки матки (РШМ) среди онкологических заболеваний, возникающих у женщин, занимает четвертое место (после рака желудка, кожи и молочных желез).

Первичный кольпоскопический осмотр пациентки врачом гинекологом определяет необходимость проведения биопсии и последующую консультацию онколога [3]. Рассмотренная в данной работе компьютерная система может быть использована врачом для преждевременной диагностики рака путем определения наличия определенных типов тканей в шейке матки без

проведения биопсии. Европейская Экспертная Группа, разрабатывающая Европейскую обобщенную обучающую программу, предложила следующую классификацию изменений эпителия шейки матки: SEA (squamous epithelial abnormalities) — доброкачественные изменения плоского эпителия позволяющие предположить CIN; Squamous cell changes — изменения плоского эпителия без четких признаков опухоли; CIN-I — дисплазия легкой степени; CIN-II — дисплазия умеренной степени; CIN-III — интраэпителиальная неоплазия тяжелой степени, данное понятие объединяет тяжелую дисплазию и внутриэпителиальный рак (CIS — carcinoma in situ) рак, подозрительный на инвазию, инвазивный плоскоклеточный рак.

Цель работы — создание компьютерной системы, которая сможет правильно классифицировать разные виды тканей по мультиспектральному изображению, применяя нейронные сети (НС). В данной работе рассмотрена, в первую очередь диагностика состояний дисплазии и неоплазии (CIN-I, II, III), а также наличие тканей с доброкачественными изменениями (MET). Такая задача диагностики является классической задачей для которой применяются НС. В работе применены нейронные сети с разными алгоритмами обучения.

ОСНОВНЫЕ ИДЕИ МЕТОДА ОБРАТНОГО РАСПРОСТРАНЕНИЯ ОШИБКИ

Данный метод является методом обучения искусственных нейронных сетей, который впервые был описан Полем Дж. Вербосом в 1974 г., и далее развит Дэвидом И. Румельхартом и Рональдом Дж. Вильямсом в 1986 году.

Основная идея этого метода состоит в распространении сигналов ошибки от выходов сети к ее входам, в направлении, обратном прямому распространению сигналов в обычном режиме работы.

Для возможности применения этого метода, передаточная функция нейронов должна быть дифференцируема.

Описание алгоритма

На рис. 1 изображена нейронная сеть. У сети есть входы x_1, \dots, x_n , выходы Outputs и внутренние узлы. Перенумеруем все узлы (включая входы и выходы) числами от 1 до N . Обозначим через $w_{i,j}$ вес, стоящий на ребре, соединяющем i -й и j -й узлы, а через o_i — выход i -го узла. Для m -тестовых примеров с целевыми значениями выходов $\{t_k^d\}$ $d = 1 \dots m$, $k \in \text{Outputs}$ функция ошибки выглядит так:

$$E(\{w_{i,j}\}) = \frac{1}{2} \sum_{d=1}^m \sum_{k \in \text{Outputs}} (t_k^d - o_k(x_1^d, \dots, x_n^d))^2.$$

Для модификации весов реализуется стохастический градиентный спуск, то есть подправляются веса после каждого тестового примера. Необходимо двигаться в сторону, противоположную градиенту, то есть добавлять к каждому весу $w_{i,j}$ $\Delta w_{i,j} = -\eta \frac{\partial E^d}{\partial w_{i,j}}$, где

$$E^d(\{w_{i,j}\}) = \frac{1}{2} \sum_{k \in \text{Outputs}} (t_k^d - o_k^d)^2.$$

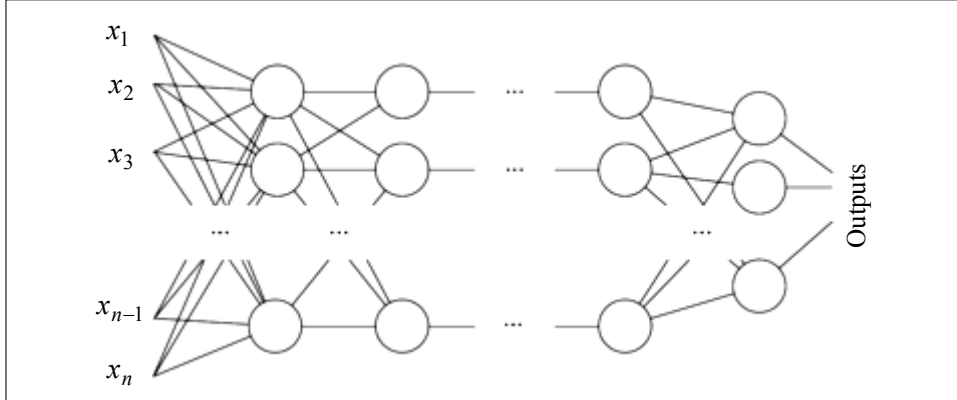


Рис. 1. Пример нейронной сети

Производная вычисляется следующим образом. Пусть сначала $j \in \text{Output}$ (интересующий вес) входит в перцептрон последнего уровня. Необходимо отметить, что $w_{i,j}$ влияет на выход перцептрона только как часть суммы

$$S_j = \sum w_{i,j} x_{i,j}, \text{ где сумма берется по входам } j\text{-го узла. Поэтому } \frac{\partial E^d}{\partial w_{i,j}} = \frac{\partial E^d}{\partial S_j} \frac{\partial S_j}{\partial w_{i,j}} = x_{i,j} \frac{\partial E^d}{\partial S_j}.$$

Аналогично, S_j влияет на общую ошибку только в рамках выхода j -го узла o_j (выход всей сети). Поэтому

$$\begin{aligned} \frac{\partial E^d}{\partial S_j} &= \frac{\partial E^d}{\partial o_j} \frac{\partial o_j}{\partial S_j} = \left(\frac{\partial}{\partial o_j} \frac{1}{2} \sum_{k \in \text{Outputs}} (t_k - o_k)^2 \right) \left(\frac{\partial \sigma(S_j)}{\partial S_j} \right) = \\ &= \left(\frac{1}{2} \frac{\partial}{\partial o_j} (t_j - o_j)^2 \right) (o_j(1 - o_j)) = -o_j(1 - o_j)(t_j - o_j). \end{aligned}$$

Если же j -й узел — не на последнем уровне, то у него есть выходы; обозначим их через $\text{Outputs}(j)$. В этом случае $\frac{\partial E^d}{\partial S_j} = \sum_{k \in \text{Outputs}(j)} \frac{\partial E^d}{\partial S_k} \frac{\partial S_k}{\partial S_j}$ и

$$\frac{\partial S_k}{\partial S_j} = \frac{\partial S_k}{\partial o_j} \frac{\partial o_j}{\partial S_j} = w_{i,j} \frac{\partial o_j}{\partial S_j} = w_{i,j} o_j(1 - o_j).$$

Здесь $\frac{\partial E^d}{\partial S_k}$ — в точности аналогичная поправка, но вычисленная для узла следующего уровня (обозначим ее через δ_k). От Δk она отличается отсутствием множителя $(-x_{i,j})$. При помощи формул для вычисления поправки для узлов последнего уровня и выражения поправки для узла более

низкого уровня через поправки более высокого уровня можно составить алгоритм. Именно из-за этой особенности вычисления поправок алгоритм называется алгоритмом обратного распространения ошибки (backpropagation). Краткое резюме рассмотренного алгоритма:

- для узла последнего уровня

$$\delta_j = -o_j(1 - o_j)(t_j - o_j);$$

- для внутреннего узла сети

$$\delta_j = -o_j(1 - o_j) \sum_{k \in \text{Outputs}} \delta_k w_{j,k};$$

- для всех узлов

$$\Delta w_{i,j} = -\eta \delta_j x_{i,j}.$$

На вход алгоритма, кроме указанных параметров, необходимо также подавать в каком-либо формате структуру сети. На практике очень хорошие результаты показывают сети достаточно простой структуры, состоящие из двух уровней нейронов — скрытого уровня (hidden units) и нейронов-выходов (output units). Каждый вход сети соединен со всеми скрытыми нейронами, а результат работы каждого скрытого нейрона подается на вход каждому из нейронов-выходов. В таком случае достаточно подать на вход количество нейронов скрытого уровня [4].

Недостатки алгоритма

Несмотря на многочисленные успешные применения обратного распространения, оно не является панацеей. Больше всего неприятностей приносит неопределенно долгий процесс обучения. В сложных задачах для обучения сети могут потребоваться дни или даже недели, она может и вообще не обучиться. Причиной может быть одна из описанных ниже.

Паралич сети. В процессе обучения сети значения весов могут в результате коррекции стать очень большими величинами. Это может привести к тому, что все или большинство нейронов будут функционировать при очень больших выходящих значениях, в области, где производная сжимающей функции очень мала. Так как посылаемая обратно в процессе обучения ошибка пропорциональна этой производной, то процесс обучения может практически остановиться. В теоретическом отношении эта проблема плохо изучена. Ее можно избежать уменьшив размер шага η , но это увеличивает время обучения. Различные эвристики используются для предохранения от паралича или для восстановления после него, но пока что они могут рассматриваться лишь как экспериментальные.

Локальные минимумы. Обратное распространение использует разновидность градиентного спуска, т.е. осуществляет спуск вниз по поверхности ошибки, непрерывно подстраивая веса в направлении к минимуму. Поверхность ошибки сложной сети сильно изрезана и состоит из холмов, долин, складок и оврагов в пространстве высокой размерности. Сеть может попасть в локальный минимум (неглубокую долину), когда рядом имеется гораздо более глубокий минимум. В точке локального минимума все направления ведут вверх, и сеть неспособна из него выбраться. Статистические методы

обучения могут помочь избежать этой ловушки, но они требуют большего затрата времени.

Размер шага. Внимательный разбор доказательства сходимости показывает, что коррекции весов предполагаются бесконечно малыми, но это неосуществимо на практике, так как ведет к бесконечному времени обучения. Размер шага должен браться конечным, и в этом вопросе приходится опираться только на опыт. Если размер шага очень мал, то сходимость слишком медленная, если же очень велик, то может возникнуть паралич или постоянная неустойчивость. Описан адаптивный алгоритм выбора шага, автоматически корректирующий размер шага в процессе обучения.

СЕТИ РАДИАЛЬНОГО БАЗИСА

Самообучающиеся сети, как известно, широко используются для преобразования данных, например при распознавании образов в пространстве очень большой размерности. В этом случае для того, чтобы процедура обучения с учителем была эффективна, необходимо сначала сжать входную информацию тем или иным способом, либо выделить значимые признаки, понизив размерность, либо произвести квантование данных. Первый путь просто понижает число входов персептрона. Второй способ требует отдельного рассмотрения, поскольку лежит в основе очень популярной архитектуры — сетей радиального базиса (radial basis functions — RBF).

Аппроксиматоры с локальным базисом

Сети радиального базиса напоминают персептрон с одним скрытым слоем, осуществляя нелинейное отображение $\mathcal{R}^d \Rightarrow \mathcal{R}^m$ $y = \sum_i h_i \phi(w_i, x)$, являю-

щееся линейной комбинацией базисных функций. Но в отличие от персептронов, где эти функции зависят от проекций на набор гиперплоскостей $\sigma(wx)$, в сетях радиального базиса используются функции (чаще всего — гауссовы), зависящие от расстояний до опорных центров: $y = \sum_i h_i \phi_i(|w_i - x|)$,

$\phi_i(z) = e^{-z^2 / \sigma_i^2}$. Все наборы базисных функций обеспечивают возможность аппроксимации любой непрерывной функции с произвольной точностью. Основное различие между ними в способе кодирования информации на скрытом слое. Если персептроны используют глобальные переменные (наборы бесконечных гиперплоскостей), то сети радиального базиса опираются на компактные шары, окружающие набор опорных центров (рис. 2).

В первом случае в аппроксимации в окрестности любой точки участвуют все нейроны скрытого слоя, во втором — лишь ближайшие. Как следствие такой неэффективности, в последнем случае количество опорных функций, необходимых для аппроксимации с заданной точностью, возрастает экспоненциально с размерностью пространства. Это основной недостаток сетей радиального базиса. Основное же их преимущество над персептронами — в простоте обучения [5].

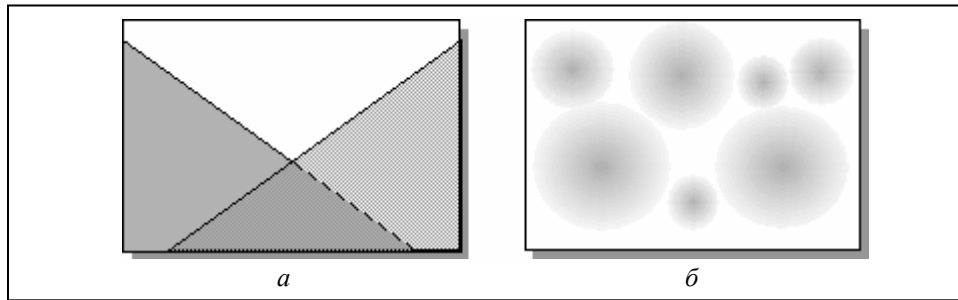


Рис. 2. Глобальная (персептроны) и локальная (сети радиального базиса) методы аппроксимации

Гибридное обучение

Относительная автономность базисных функций позволяет разделить обучение на два этапа. На первом этапе обучается первый (соревновательный) слой сети, осуществляя квантование данных. На втором этапе происходит быстрое обучение второго слоя матричными методами, т.к. нахождение коэффициентов второго слоя представляет собой линейную задачу.

Подобная возможность раздельного обучения слоев является основным достоинством сетей радиального базиса. В целом же, области применимости персептронов и сетей радиального базиса коррелируют с найденными выше областями эффективности квантования и понижения размерности.

ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ДАННЫХ

Большее количество показателей имеют очень высокую корреляцию друг с другом, что может негативно влиять на работу нейронной сети для классификации тканей на больную/здоровую. Для этого во избежание потенциальных проблем в двух случаях показатели меняются. В одном случае, для обучения нейронной сети используется только подмножество показателей, между которыми коэффициент корреляции не превышает 0,95 (определен экспериментально). В другом случае, используется метод основных компонент, показатели которого трансформируются в набор независимых некоррелируемых показателей. Эти методы должны улучшить работу нейронной сети и увеличить точность классификации.

Метод главных компонент (МГК) (в англоязычной литературе — principal component analysis — PCA) — один из наиболее распространенных методов факторного анализа. Суть метода состоит в декомпозиции матрицы данных X и представлении ее в виде произведения двух матриц: T (матрица счетов, scores) и P (матрица нагрузок, loadings): $X = TPT$, что, по сути, является переходом к новым переменным.

Одно из важнейших свойств такого преобразования — ортогональность новых переменных: $PPT = I$, $TTT = I$, где I — единичная матрица значительно упрощает дальнейшие операции с матрицами и позволяет ранжировать новые переменные в соответствии с их значимостью. Как правило, используются только значимые переменные («главные компоненты»), а незначимые переменные (соответствующие случайным погрешностям) ис-

ключаются из дальнейшего рассмотрения, что позволяет сжать данные и значительно снизить погрешность расчетов [6].

РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Имеются диагнозы двух врачей, представляющие собой процентное содержание клеток каждого из шести возможных типов тканей (Squamous, Columnar, MET, CIN1, CIN2, CIN3). В таблице приведены результаты работы нейронных сетей. При проведении экспериментов использовался метод Cross validation (скользящий контроль). Процедура скользящего контроля заключается в следующем. Фиксируется некоторое множество разбиений выборки на две части: обучающую и проверочную. Для каждого разбиения выполняется настройка алгоритма по обучающей подвыборке и вычисляется частота его ошибок на проверочной подвыборке. Оценка скользящего контроля определяется как средняя по всем разбиениям частота ошибок на проверочной. Фактически скользящий контроль непосредственно измеряет обобщающую способность метода обучения на заданой конечной выборке.

Т а б л и ц а . Сравнительный анализ работы нейронных сетей

Средне-квадратичное отклонение (СКО)	Алгоритм обратного распространения ошибки. Метод сопряженных градиентов	Алгоритм обратного распространения ошибки. Метод квази Ньютона	Алгоритм обратного распространения ошибки. Метод масштабируемых сопряженных градиентов	Сеть радиального базиса
СКО (>0,95) Squamous	0,0607	0,0593	0,0499	0,0325
СКО (PCA) Squamous	0,0502	0,0621	0,0507	0,0326
СКО (>0,95) Columnar	0,0459	0,0457	0,0449	0,0295
СКО (PCA) Columnar	0,0471	0,0480	0,0408	0,0297
СКО (>0,95) MET	0,0561	0,0530	0,0492	0,0244
СКО (PCA) MET	0,0580	0,0577	0,0582	0,0247
СКО (>0,95) CIN1	0,0462	0,0852	0,0408	0,0220
СКО (PCA) CIN1	0,0817	0,0958	0,0458	0,0224
СКО (>0,95) CIN2	0,0427	0,0497	0,0345	0,0158
СКО (PCA) CIN2	0,0392	0,0500	0,0352	0,0158
СКО (>0,95) CIN3	0,0734	0,0789	0,0739	0,0305
СКО (PCA) CIN3	0,0656	0,0755	0,0877	0,0306

В таблице представлено СКО для 4 НС, с двумя методами предварительной обработки данных (СКО(>0,95)) для подмножества показателей, между которыми коэффициент корреляции не превышает 0,95 и СКО (РСА) (данные с использованием метода основных компонент) для 6 типов тканей. Из таблицы видно, что сеть RBF дает наименьшую ошибку, при этом оба метода предварительной обработки данных равноценны.

ВЫВОДЫ

1. Результаты проведенных экспериментов показали целесообразность использования нейронных сетей для решения задач классификации типов тканей на медицинских изображениях.

2. Исходя из проведенных экспериментов видно, что сеть RBF всегда дает наилучшие результаты.

3. Использование различных способов предварительной обработки данных значительно улучшает работу нейронных сетей. После проведения ряда экспериментов оба способа оказались достаточно эффективными.

4. Разработанный подход диагностики рака шейки матки с использованием нейронных сетей может быть рекомендован для дальнейшей клинической апробации.

ЛИТЕРАТУРА

1. *Schoonmaker J. et al.* Automatic Segmentation of Uterine Cervix for in vivo Localization and Identification of Cervical Intraepithelial Neoplasia. — <http://spie.org/Documents/ConferencesExhibitions/mi07-final.pdf>.
2. *Воробьева Л.И.* Цитологический скрининг рака шейки матки // Здоровье Украины. — 2008. — № 2/1. — С. 18.
3. *Koss L.G.* The complex test for cervical cancer detection // The Journal of the American Medical Association. — 1989. — **261**, № 5. — С. 737–743.
4. *Зайченко Ю.П.* Основы проектирования интеллектуальных систем. Навчальний посібник. — Київ: Видавничий Дім «Слово», 2004. — 352 с.
5. *Domagoj Kovacevic, Sven Loncaric.* Radial Basis Function-based Image Segmentation using a Receptive Field // Computer-Based Medical Systems. — 1997. Proceedings., Tenth IEEE Symposium. — P. 126–130.
6. *Richard G. Brereton* Introduction to multivariate calibration in analytical chemistry // Analyst. — 2000. — № 125. — P. 2125–2154.

Поступила 02.07.2008

**ФОРМИРОВАНИЕ ПОДГРУПП ЭКСПЕРТОВ
С СОГЛАСОВАННЫМИ МНЕНИЯМИ
И ОПРЕДЕЛЕНИЕ ГРУППОВОЙ ОБОБЩЕННОЙ ОЦЕНКИ
МНОГОПРИЗНАКОВЫХ ОБЪЕКТОВ**

Н.В. КРАПУХИНА, С.В. ПРОНИЧКИН, А.С. РЫКОВ

Рассмотрены новые алгоритмы обработки экспертных оценок многопризнаковых объектов. Алгоритмы применены для решения практической задачи оценки компетенций студентов, оцененных несколькими экспертами по многим критериям. Применение разработанных алгоритмов позволило получить решение более адекватное предпочтениям ЛПР (лицо принимающее решение), чем в случае применения существующих алгоритмов.

ВВЕДЕНИЕ

Оценка на основе компетентностного подхода, активно используемая в настоящее время при оценке персонала предприятия, нашла свое отражение и в высшем образовании. Например, источником информации о социальных и личностных качествах (компетенциях) студентов являются преподаватели, а оценки по компетенциям используются для вычисления рейтинга студента. Рейтинговая оценка позволяет получить обобщенную оценку объекта по всем применяемым критериям, а также сравнить объекты [1]. Рейтинг вычисляется по многим критериям, но не все критерии могут быть получены в результате измерений. Источником информации по многим критериям являются эксперты.

Основная цель обработки экспертных оценок — получение обобщенных данных и выявление новой информации, содержащейся в скрытой форме в экспертных оценках. Эксперты обычно расходятся во мнениях по этому поводу. При отсутствии согласованности экспертов в существующих работах принято разбивать их на группы сходные по мнению. В связи с этим, возникают две важные задачи обработки экспертных оценок [2, 3, 4, 5]:

- выделение подгрупп экспертов с согласованными мнениями;
- определение групповой обобщенной оценки объектов на основе индивидуальных оценок экспертов в выделенных подгруппах.

Пусть t экспертов произвели оценку n объектов по s критериям, используя одну и ту же шкалу интервалов (оценка в баллах). Результаты

оценки представим в виде величин x_{ijl} , где i — номер эксперта ($i=1, \dots, m$), j — номер объекта ($j=1, \dots, n$), l — номер критерия ($l=1, \dots, s$). Величины x_{ijl} представляют собой баллы. Известна относительная важность критериев w_l , $l=1, \dots, s$, $\sum_{l=1}^s w_l = 1$.

Для получения разбиения исходной группы экспертов на подгруппы со сходными мнениями для заданного критерия принято использовать дисперсионный коэффициент конкордации [3, 6]. Данный показатель вычисляется на рангах, и позволяет оценивать согласованность мнений экспертов в группе. Так как исходные индивидуальные оценки экспертов даны в шкале интервалов, то их надо перевести в шкалу порядка, приписывая лучшим оценкам в баллах более высокие ранги, чем худшим. В результате получаются матрицы рангов $\|r_{ij}\|_l$, $i=1, \dots, m$, $j=1, \dots, n$, $l=1, \dots, s$, размерность каждой матрицы составляет $m \times n$.

Дисперсионный коэффициент конкордации при наличии связанных рангов для заданного l вычисляется по следующей формуле:

$$W = \frac{12S}{m^2(n^3 - n) - m \sum_{i=1}^m T_i}, \quad (1)$$

где $S = \sum_{j=1}^n \left(\sum_{i=1}^m r_{ij} - \bar{r} \right)^2$, $\bar{r} = \frac{\sum_{j=1}^n r_j}{n} = \frac{m(n+1)}{2}$, $r_j = \sum_{i=1}^m r_{ij}$; $T_i = \sum_{k=1}^{H_i} (h_k^3 - h_k)$ —

показатель связанных рангов в i -й ранжировке; H_i — число групп равных рангов в i -й ранжировке; h_k — число равных рангов в k -й группе связанных рангов в i -й ранжировке.

Если связанных рангов нет, то $H_i = 0$, $h_k = 0$, следовательно $T_i = 0$.

Коэффициент конкордации W равен единице, если все ранжировки экспертов одинаковы, и равен нулю, если все ранжировки различны, т.е. совершенно нет совпадения.

Коэффициент конкордации, вычисляемый по формуле (1), является оценкой истинного значения коэффициента, представляя собой случайную величину. Для оценки значимости W выбирают вероятность ошибки $p_{\text{ош}}$, определяют W_T , и если $W \geq W_T$, то W считается статистически значимым.

В работе [2] предлагается алгоритм, в котором последовательно исключаются подгруппы с согласованными мнениями из исходной группы, и тем самым исходная группа разбивается на подгруппы с согласованными мнениями для заданного критерия. Данный алгоритм не лишен недостатков:

1. Поиск подгруппы экспертов с согласованными мнениями прекращается, как только максимальное значение коэффициента конкордации для текущего числа экспертов в подгруппе оказывается значимым. Это приводит к тому, что подгруппа экспертов с согласованными мнениями, получен-

ная первой, будет больше, чем остальные подгруппы, и в общем случае разбиение исходной группы экспертов на подгруппы с согласованными мнениями не будет оптимальным, в смысле совместной максимизации коэффициентов конкордации для подгрупп экспертов с согласованными мнениями для заданного критерия.

2. Учитывается только значение дисперсионного коэффициента конкордации для подгрупп экспертов, но не учитывается диапазон изменения значений коэффициента конкордации, разброс его значений для заданного критерия, что приводит к неудовлетворительным решениям. Согласованность одной подгруппы экспертов может сильно отличаться от согласованности другой.

3. Разбиение исходной группы экспертов на подгруппы с согласованными мнениями осуществляется для заданного (одного) критерия. В общем случае получается, что для каждого объекта в зависимости от критерия набор экспертов его оценивших будет разным, причем количество экспертов в подгруппе со сходными мнениями для одного критерия может сильно отличаться от количества экспертов в подгруппе со сходными мнениями для другого критерия.

В данной работе предлагается совместно использовать значения коэффициентов конкордации для подгрупп экспертов, и стремиться к их суммарной максимизации. В то же время, желательно, чтобы согласованность в подгруппах несильно отличалась. Каждый объект оценивается каждым экспертом по многим критериям, поэтому целесообразно выделение подгрупп экспертов с согласованными мнениями по всем критериям с учетом их важности.

Для преодоления первого и второго недостатка предлагается построить целевую функцию. Проведем объединение (свертку) среднего значения коэффициентов конкордации и среднего квадратического отклонения коэффициентов конкордации для заданного критерия l . В результате получаем следующую целевую функцию:

$$f(W_l^{(t)}) = (1 - \lambda_l) \frac{\sum_{j=1}^t W_l(m_j)}{t} - \lambda_l \sqrt{\frac{\sum_{j=1}^t (\bar{W}_l(t) - W_l(m_j))^2}{t-1}}, \quad (2)$$

где $W_l^{(t)} = (W_l(m_1), \dots, W_l(m_t))$ — вектор коэффициентов конкордации полученных для разбиения исходной группы экспертов на t подгрупп $\mathcal{E}^{(t)} = (\mathcal{E}_1, \dots, \mathcal{E}_{m_1}, \dots, \mathcal{E}_{m_{t-1}+1}, \dots, \mathcal{E}_{m_t})$ для l -го критерия (количество и состав экспертов в подгруппах одинаковый для всех критериев); $W_l(m_j)$ — коэффициент конкордации, подсчитанный для оценок m_j экспертов, входящих в j -ю подгруппу для l -го критерия; $\bar{W}_l(t) = \frac{\sum_{j=1}^t W_l(m_j)}{t}$ — среднее значение коэффициентов конкордации для l -го критерия; λ_l — параметр (весовой коэффициент) для l -го критерия.

Необходимо найти такое $\mathcal{E}^{(t)}$, которому соответствует $W_l^{(t)}$ для (2), при котором среднее значение коэффициентов конкордации будет больше, а значение среднего квадратического отклонения коэффициентов конкордации будет меньше. Варьирование значений параметра $\lambda_l \in [0, 1]$ изменяет свойства целевой функции (2) для l -го критерия, позволяя в большей или в меньшей степени в зависимости от величины λ_l учитывать величины среднего значения коэффициентов конкордации и среднего квадратического отклонения коэффициентов конкордации, включенных в целевую функцию.

Для преодоления третьего недостатка предлагается построить целевую функцию. Проведем объединение (свертку) среднего значения и среднего квадратического отклонения значений целевой функции (2) с учетом важности критериев. В результате получаем следующую целевую функцию:

$$F(W^{(t)}) = (1 - \lambda) \frac{\sum_{l=1}^s w_l f(W_l^{(t)})}{s} - \lambda \sqrt{\frac{\sum_{l=1}^s w_l^2 (\bar{f}(W_l^{(t)}) - f(W_l^{(t)}))^2}{s-1}}, \quad (3)$$

где

$W^{(t)} = (W_1^{(t)}, \dots, W_s^{(t)})$ — векторы коэффициентов конкордации полученные для разбиения исходной группы экспертов на t подгрупп $\mathcal{E}^{(t)} = (\mathcal{E}_1, \dots, \mathcal{E}_{m_1}, \dots, \mathcal{E}_{m_{t-1}+1}, \dots, \mathcal{E}_{m_t})$;

$$\bar{f}(W_l^{(t)}) = \frac{\sum_{l=1}^s w_l f(W_l^{(t)})}{s} \text{ — средневзвешенное значение функций (2);}$$

w_l — важность критериев, $l = 1, \dots, s$, $\sum_{l=1}^s w_l = 1$; λ — параметр (весовой коэффициент).

Цель задачи заключается в нахождении решения $\mathcal{E}_{\text{опт}}^{(t)}$, которому соответствует $W_{\text{опт}}^{(t)}$ из условия:

$$F(W_{\text{опт}}^{(t)}) = \max_{W^{(t)} \in W_{\text{знач}}^{(t)}} F(W^{(t)}), \quad (4)$$

где $W_{\text{знач}}^{(t)}$ — множество значимых коэффициентов конкордации, $W_{\text{знач}}^{(t)} = \{W_l(m_q) | W_l(m_q) \geq W_{Tl}(m_q), q = 1, \dots, t, l = 1, \dots, s\}$.

Дадим формальное описание алгоритма решения поставленной задачи (4).

1. Задать $p_{\text{ош}}(l)$, λ_l , $l = 1, \dots, s$.

2. $\lambda = 0$, $t_1 = 1$.

3. Полные матрицы рангов $\|r_{ij}\|_l$, $i = 1, \dots, m$, $j = 1, \dots, n$, $l = 1, \dots, s$ последовательно в разных сочетаниях делить на t_1 групп строк. Строк в группе $m_q = 2, \dots, m - 2(t_1 - 1)$, $q = 1, \dots, t_1$, если $m - 2t_1 \geq 1$, то $\sum_{q=1}^{t_1} m_q = m - 1, m$,

иначе $\sum_{q=1}^{t_1} m_q = m$. Для каждого набора строк следует вычислить $W_l^{(t_1)} = (W_l(m_1), \dots, W_l(m_{t_1}))$, $l = 1, \dots, s$ и сравнить $W_l(m_q)$ с $W_{Tl}(m_q)$. Если $W_l(m_q) \geq W_{Tl}(m_q)$, $q = 1, \dots, t_1$, $l = 1, \dots, s$, то вычислить $F(W^{(t_1)})$. Найти $F(W^{(t_1)})$ с максимальным значением $F(W_{\text{опт}}^{(t_1)})$.

4. $t_1 = t_1 + 1$.

5. Проверить условие $t_1 \leq \left\lfloor \frac{m}{2} \right\rfloor$. Если условие выполнено — перейти к п. 3.

п. 3.

6. Найти $F(W_{\text{опт}}^{(t_1)})$ с максимальным значением $F(W_{\text{опт}}^{(t)})$. Запомнить λ , t , $W_{\text{опт}}^{(t)} = (W_1^{(t)*}, \dots, W_s^{(t)*})$ и соответствующие матрицы рангов $\|r_{ij}^1\|_l$, $i = 1, \dots, m_1, \dots, \|r_{ij}^t\|_l$, $i = 1, \dots, m_t$, $j = 1, \dots, n$, $l = 1, \dots, s$ с согласованными мнениями экспертов.

7. $\lambda = \lambda + 0,1$.

8. Проверить условие $\lambda \leq 1$. Если условие выполнено — перейти к п. 3.

9. В случае если решение не найдено — перейти к п. 1.

В предложенном алгоритме значения параметров λ_l , $l = 1, \dots, s$ можно задать непосредственно перед вычислением $F(W^{(t_1)})$ (см. п. 3). В качестве примера шаг изменения параметра λ выбран 0,1. Величину шага изменения параметра λ , h_λ задает ЛПР, затем ЛПР изучает полученные решения (4) при $\lambda = \{0; h_\lambda; \dots; 1\}$ и выбирает λ соответствующие его представлению о качестве решения.

Такая наглядность позволяет ЛПР изучать полученное множество решений и более ясно формулировать свои требования, сравнивая различные решения.

После решения задачи выделения подгруппы экспертов с согласованными мнениями. Необходимо решать задачу определения групповой обобщенной оценки объектов для каждой из выделенных подгрупп.

Для получения групповой оценки объекта используются средневзвешенные оценки [2, 3]:

$$x_{jl} = \sum_{i=1}^m K_{il} x_{ijl}, \quad (5)$$

где K_{il} — коэффициент компетентности i -го эксперта для l -го критерия, коэффициенты являются нормализованными величинами $\sum_{i=1}^m K_{il} = 1$, $l = 1, \dots, s$.

Выбор величин коэффициентов компетентности экспертов носит субъективный характер. В работе [2] предлагаются алгоритмы, основной идеей которых является предположение о том, что компетентность экспертов должна оцениваться по степени согласованности их индивидуальных оценок с групповой оценкой объектов по заданному критерию. Данные алгоритмы имеют ряд недостатков, поскольку каждый объект оценивается по многим критериям.

В данной работе предлагается оценивать компетентность эксперта, учитывая оценки по всем критериям. Для этого предлагаются алгоритмы, основанные на итеративной процедуре корректировки коэффициентов компетентности K_{il}^r , $l = 1, \dots, s$, $i = 1, \dots, m_u$, $r = 0, 1, 2, \dots$ — номер итерации, m_u — количество экспертов в подгруппе u ($u = 1, \dots, t$), полученной из решения задачи (4). На каждой итерации r вычисляется взвешенная групповая оценка x_{jl}^r каждого объекта ($j = 1, \dots, n$) по всем критериям ($l = 1, \dots, s$). Затем вычисляются специальные поправочные коэффициенты ΔK_{il}^r , величина которых обратно пропорциональна значению целевой функции сконструированной на основе принципа оптимальности или комбинации принципов оптимальности [3, 7]. Значение целевой функции зависит от вектора d_{il}^r , компонентами которого являются значения p_{l1} -нормы $d_{l1}(x_{i1}^r, x_{il1}, p_{l1})$ определяющей величину отклонения вектора экспертных оценок $x_{i1} = (x_{i11}, \dots, x_{in1})$ от вектора средних групповых оценок объектов $x_{l1}^r = (x_{l11}^r, \dots, x_{ln1}^r)$ по l_1 критерию ($l_1 = 1, \dots, s$) на r -й итерации (6), (7).

$$d_{l1}(x_{i1}^r, x_{il1}, p_{l1}) = \left(\sum_{j=1}^n |x_{jl1}^r - x_{ijl1}|^{p_{l1}} \right)^{\frac{1}{p_{l1}}}, \quad p_{l1} \geq 1. \quad (6)$$

$$d_{l1}(x_{i1}^r, x_{il1}, p_{l1}) = \max_j |x_{jl1}^r - x_{ijl1}|, \quad p_{l1} = \infty. \quad (7)$$

После вычисления поправочных коэффициентов с их помощью корректируются коэффициенты компетентности. Корректировка возможна в аддитивной или в мультипликативной форме [2]. После корректировки проводится нормализация коэффициентов и проверка правила останова процедуры корректировки коэффициентов компетентности. Вычисления прекращаются, когда значения коэффициентов перестают меняться.

В зависимости от выбора норм, способа корректировки коэффициента компетентности и принципа оптимальности или их комбинации порождается соответствующий вариант алгоритма.

В качестве примера в предлагаемом нами алгоритме целевая функция $D(d_{il}^r, p)$ сконструирована на основе принципа идеальной точки. Согласно принципу идеальной точки лучшим считается вектор, расположенный ближе всего (в смысле p -нормы) к идеальной точке. Идеальная точка берется формально как вектор компоненты которого равны нулю, т.е. отклонения индивидуальных оценок эксперта от групповой оценки объектов по всем

критериям ($l_1 = 1, \dots, s$) равны нулю. В зависимости от выбора нормы целевая функция принимает следующий вид (8), (9).

$$D(d_{il}^r, p) = \left(\sum_{l_1=1}^s w_{l_1}^p (d_{l_1}(x_{l_1}^r, x_{il_1}, p_{l_1l}))^p \right)^{\frac{1}{p}}, \quad p \geq 1. \quad (8)$$

$$D(d_{il}^r, p) = \max_{l_1} (w_{l_1} \cdot d_{l_1}(x_{l_1}^r, x_{il_1}, p_{l_1l})), \quad p = \infty. \quad (9)$$

Дадим формальное описание алгоритма получения групповых оценок (5).

1. Задать $p, \varepsilon > 0, \varepsilon_l \geq 0, p_{l_1l}, l_1 = 1, \dots, s, l = 1, \dots, s$.

2. $u = 1$.

3. $r = 0$, начальные значения коэффициентов компетентности:

$$K_{il}^0 = \frac{1}{m_u}, \quad i = 1, \dots, m_u, \quad l = 1, \dots, s.$$

4. $r = r + 1$.

5. Вычислить средние групповые оценки: $x_{jl}^r = \sum_{i=1}^{m_u} K_{il}^{r-1} x_{ijl}, \quad j = 1, \dots, n,$

$l = 1, \dots, s$.

6. Вычислить поправочные коэффициенты: $\Delta K_{il}^r = \frac{1}{\varepsilon + D(d_{il}^r, p)},$

$i = 1, \dots, m_u, \quad l = 1, \dots, s$.

7. Скорректировать коэффициенты компетентности для аддитивного варианта $K_{il}^{r*} = K_{il}^{r-1} + \Delta K_{il}^r, \quad i = 1, \dots, m_u, \quad l = 1, \dots, s$; для мультипликативного варианта $K_{il}^{r*} = K_{il}^{r-1} \Delta K_{il}^r, \quad i = 1, \dots, m_u, \quad l = 1, \dots, s$.

8. Нормализовать коэффициенты компетентности: $K_{il}^r = \frac{K_{il}^{r*}}{\sum_{i=1}^{m_u} K_{il}^{r*}},$

$i = 1, \dots, m_u, \quad l = 1, \dots, s$.

9. Проверить выполнение условий останова: $\max_i |K_{il}^r - K_{il}^{r-1}| \leq \varepsilon_l,$

$l = 1, \dots, s$. Если условия не выполнены, то перейти к п. 4

10. Запомнить полученные значения коэффициентов компетентности $K_{il}^r, \quad i = 1, \dots, m_u, \quad l = 1, \dots, s$ и средние групповые оценки $x_{jl}^r, \quad j = 1, \dots, n, \quad l = 1, \dots, s$.

11. $u = u + 1$.

12. Проверить условие $u \leq t$, условие выполнено, то перейти к п. 3.

Свойства алгоритма, включая его сходимость, зависят от целевой функции, параметров и способа корректировки коэффициентов компетент-

ности. Выбирая значения p_{l_l} , $l_1 = 1, \dots, s$, $l = 1, \dots, s$, можно по-разному описывать понятие расстояния между оценками.

ВЫВОДЫ

Рассмотренные алгоритмы использовались при обработке оценок компетенций студентов [8]. Источником информации о социальных и личностных качествах (компетенциях) студентов являются преподаватели. Применение разработанных алгоритмов обработки экспертных оценок многопризнаковых объектов в сочетании с алгоритмами многокритериального принятия решений [3, 4, 7] позволило вычислить рейтинг студентов более адекватный предпочтениям ЛПП, чем в случае применения существующих алгоритмов.

ЛИТЕРАТУРА

1. Федоренко З.Н. Основные подходы к созданию рейтинговых систем // Прикладная математика и информатика. Сборник научных трудов. — Петрозаводск: ПетрГУ, 2000. — С. 123–130.
2. Рыков А.А., Рыков А.С. Алгоритмы обработки экспертной информации для оценки качества информационных систем // Экономика, информационные технологии и управление в металлургии. Сборник научных трудов. — М.: МИСиС, 2003. — С. 86–90.
3. Рыков А.С. Модели и методы системного анализа: принятие решений и оптимизация: Учебное пособие для вузов. — М.: МИСиС, Издательский дом «Руда и металлы», 2005. — 352 с.
4. Орлов А.И. Теория принятия решений. — М.: «Экзамен», 2005. — 656 с.
5. Тоценко В.Г. Методы и системы поддержки принятия решений. — Киев: Наук. думка, 2002. — 381 с.
6. Прикладная статистика: Исследование зависимостей: Справ. изд. / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин; Под ред. С.А. Айвазяна. — М.: Финансы и статистика, 1985. — 487 с.
7. Черноруцкий И.Г. Методы принятия решений. — СПб.: БХВ-Петербург, 2005. — 416 с.
8. Разработка организационного, методического и математического обеспечения независимой системы оценки и мониторинга качества образования. Отчет о НИР (заключит.) / Московский институт стали и сплавов (МИСиС). — М.: МИСиС, 2007. — 1240 с.

Поступила 02.07.2008

СИНТЕЗ РЕГУЛЯТОРА СОСТОЯНИЯ ДЛЯ ДИСКРЕТНОЙ ВО ВРЕМЕНИ СИСТЕМЫ МОДАЛЬНОГО УПРАВЛЕНИЯ

А.Г. ФИЛАТОВ

Предложена процедура синтеза регулятора состояния для дискретной во времени системы управления, основанная на теории модального управления и обеспечивающая произвольную степень устойчивости замкнутой системы в пространстве переменных состояния и их средних интегральных значений. Полученные результаты проиллюстрированы примером.

ВВЕДЕНИЕ

В работе [1] рассмотрен метод синтеза достаточно простого в реализации модального регулятора состояния в дискретной во времени системе управления электрическими нагрузками электроэнергетических систем, обеспечивающий произвольную степень устойчивости системы в пространстве переменных состояния и их средних интегральных значений. Однако этот метод применим только для квазидинамических дискретных систем, характеризуемых единичной матрицей состояния. Ниже изложено развитие этого метода для дискретной во времени системы модального управления с произвольной матрицей состояния.

ПОСТАНОВКА ЗАДАЧИ

Рассмотрим дискретную во времени систему управления, представленную следующими уравнениями состояния:

$$x(k+1) = Ax(k) + Bu(k), \quad (1)$$

где $x(k)$ — n -мерный вектор состояния в момент времени t_k ; $x(k+1)$ — n -мерный вектор состояния в момент времени t_{k+1} ; $u(k)$ — n -мерный вектор управления в момент времени t_k ; A и B — $n \times n$ -матрица состояния и матрица управления соответственно с постоянными действительными элементами. Предположим, что пара матриц A, B является полностью управляемой, а все собственные числа $\lambda_1, \lambda_2, \dots, \lambda_n$ матрицы A — различны. Будем также полагать, что элементы вектора состояния x являются кусочно-постоянными величинами на интервалах времени $\Delta t_k = t_k - t_{k-1}$, $k = 1, 2, 3, \dots$ и представлены их непосредственными измерениями или их оценками.

Как известно [2–4], в системе (1) при формировании управляющих сигналов с помощью обратной связи (ОС) по состоянию (регулятора состояния)

$$u(k) = Kx(k), \quad (2)$$

где K — матрица коэффициентов ОС размерностью $n \times n$, собственные числа матрицы $(A + BK)$ замкнутой системы могут быть выбраны произвольным образом, если пара A, B полностью управляема. Соответствующим выбором собственных чисел замкнутой системы можно обеспечить желаемый характер свободного движения переменных состояния. Для уменьшения установившейся ошибки системы и противодействия неконтролируемым возмущениям в законы управления вводят интегральную составляющую. Однако введение интегральной составляющей, реализация которой в дискретных системах может быть достигнута различными способами, приводит к снижению степени устойчивости системы и усложняет устройство управления [4].

В данной работе интегральная составляющая формируется на основе средних интегральных значений переменных состояния объекта на текущий момент времени k с начала некоторого периода контроля $\Delta T_k \gg \Delta t_k$. Это позволяет непосредственно включить в контур модального управления не только переменные состояния процесса, но и их средние интегральные значения на текущий момент времени. В итоге возникает возможность задания произвольного спектра замкнутой системы управления.

Под средним интегральным значением i -й переменной состояния на каждом k -м шаге управления будем понимать величину, определяемую следующим уравнением:

$$\delta_i(k) = \frac{1}{k} \sum_{j=1}^k x_i(j), \quad i = 1, \dots, n, \quad (3)$$

где k — количество интервалов (шагов) управления $\Delta t_k = t_k - t_{k-1}$, ($k = 1, 2, 3, \dots$) с начала некоторого периода контроля ΔT_k на текущий момент времени t_k . В дальнейшем величину (3) будем именовать средней интегральной ошибкой управления. Совокупность величин (3) образует вектор $\Omega(k) = [\delta_1(k), \delta_2(k), \delta_3(k), \dots, \delta_n(k)]^T$.

Закон управления с пропорциональной и интегральной составляющими в цепи обратной связи (дискретный регулятор состояния с ПИ-законом управления) можно сформировать в следующем виде:

$$u(k) = K_1 x(k) + K_2 \Omega(k), \quad (4)$$

где K_1 и K_2 — $n \times n$ -матрицы коэффициентов усиления пропорциональной и интегральной обратных связей соответственно.

Затем можно определить нашу задачу, как задачу определения способа вычисления матриц K_1 и K_2 коэффициентов усиления обратных связей таким образом, чтобы собственные числа системы (1), замкнутой обратной связью (4), имели желаемые наперед заданные значения.

СИНТЕЗ РЕГУЛЯТОРА СОСТОЯНИЯ

Преобразуем систему (1) к каноническому диагональному виду. Пусть P — невырожденная матрица собственных векторов матрицы A , преобразующая вектор состояния $x(k)$ в вектор $y(k)$, т.е.

$$x(k) = Py(k), \quad (5)$$

$$y(k) = P^{-1}x(k). \quad (6)$$

Сделаем замену переменных в (1) с помощью (5), получим

$$Py(k+1) = APy(k) + Bu(k). \quad (7)$$

Умножая (7) слева на P^{-1} , получаем

$$y(k+1) = P^{-1}APy(k) + P^{-1}Bu(k). \quad (8)$$

Напомним теперь определение собственных чисел и собственных векторов. Если Λ — диагональная $n \times n$ матрица собственных чисел квадратной $n \times n$ матрицы A :

$$\Lambda = \text{diag} \{ \lambda_i, i = 1, \dots, n \}, \quad (9)$$

то справедливо следующее соотношение:

$$P\Lambda = AP. \quad (10)$$

Умножая (10) слева на P^{-1} получим:

$$\Lambda = P^{-1}AP. \quad (11)$$

С учетом (11) уравнение (8) примет следующий канонический диагональный вид:

$$y(k+1) = \Lambda y(k) + Gu(k), \quad (12)$$

где

$$G = P^{-1}B \quad (n \times n). \quad (13)$$

Аналогично запишем уравнение (3) для переменных $y(k)$ линейно преобразованного вектора $x(k)$, т.е. средняя интегральная ошибка управления для переменных $y_i(k)$, $i = 1, \dots, n$ будет иметь вид:

$$\theta_i(k) = \frac{1}{k} \sum_{j=1}^k y_i(j), \quad i = 1, \dots, n. \quad (14)$$

Величины $\theta_i(k)$, $i = 1, \dots, n$ будут линейной комбинацией наблюдаемых по предположению переменных состояния $x_i(k)$, $i = 1, \dots, n$.

Рассмотрим динамическое уравнение, определяющее среднюю интегральную ошибку $\theta_i(k+1)$ на $(k+1)$ -м шаге в зависимости от номера шага управления и от величины ошибки $\theta_i(k)$ на k -м шаге или величины ошибки $\theta_i(k-1)$ на $(k-1)$ -м шаге:

$$\begin{aligned} \theta_i(k+1) &= \frac{1}{k+1} \sum_{j=1}^{k+1} y_i(j) = \frac{1}{k+1} y_i(k+1) + \frac{k}{k+1} \theta_i(k) = \\ &= \frac{1}{k+1} y_i(k+1) + \frac{1}{k+1} y_i(k) + \frac{k-1}{k+1} \theta_i(k-1), \quad i = 1, \dots, n. \end{aligned} \quad (15)$$

Необходимым требованием к системе управления (1) или (12) является требование асимптотической устойчивости. Этому требованию эквивалентно условие:

$$\lim_{k \rightarrow \infty} y(k) = 0. \quad (16)$$

Воспользуемся условием (16) и будем считать, что система управления функционирует таким образом, чтобы обеспечить отклонение состояния на $(k+1)$ -м шаге управления равным нулю, т.е. примем $y_i(k+1) = 0$, $i = 1, \dots, n$. Тогда уравнение (15) примет следующий вид:

$$\theta_i(k+1) = \frac{1}{k+1} y_i(k) + \frac{k-1}{k+1} \theta_i(k-1), \quad i = 1, \dots, n. \quad (17)$$

Запишем уравнение (17) в матричном виде:

$$\Theta(k+1) = Dy(k) + F\Theta(k), \quad (18)$$

где $\Theta(k+1)$ и $\Theta(k-1)$ вектора-столбцы вида:

$$\Theta(k+1) = [\theta_1(k+1), \theta_2(k+1), \dots, \theta_n(k+1)]^T, \quad (19)$$

$$\Theta(k-1) = [\theta_1(k-1), \theta_2(k-1), \dots, \theta_n(k-1)]^T. \quad (20)$$

D и F — $n \times n$ диагональные матрицы:

$$D = \begin{bmatrix} \frac{1}{k+1} & 0 \\ & \ddots \\ 0 & \frac{1}{k+1} \end{bmatrix}, \quad F = \begin{bmatrix} \frac{k-1}{k+1} & 0 \\ & \ddots \\ 0 & \frac{k-1}{k+1} \end{bmatrix}. \quad (21)$$

Сформируем расширенный вектор состояния системы управления в следующем виде:

$$z(k+1) = \begin{bmatrix} y(k+1) \\ \Theta(k+1) \end{bmatrix}, \quad z(k) = \begin{bmatrix} y(k) \\ \Theta(k-1) \end{bmatrix}. \quad (22)$$

Таким образом, расширенный вектор состояния (22) включает в себя линейно-преобразованные вектора переменных состояния $y(k)$ и средних интегральных ошибок управления $\Theta(k-1)$.

Далее, объединив уравнения (12) и (18), получим расширенное уравнение состояния системы управления:

$$\begin{bmatrix} y(k+1) \\ \Theta(k+1) \end{bmatrix} = \begin{bmatrix} \Lambda & O \\ D & F \end{bmatrix} \begin{bmatrix} y(k) \\ \Theta(k-1) \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} u(k) \quad (23)$$

или

$$z(k+1) = \hat{A}z(k) + \hat{B}u(k), \quad (24)$$

где

$$\hat{A} = \begin{bmatrix} \Lambda & O \\ D & F \end{bmatrix}, \quad (25)$$

$$\hat{B} = \begin{bmatrix} G \\ 0 \end{bmatrix}. \quad (26)$$

Вектор управляющих сигналов $u(k)$ будем искать в виде следующей обратной связи по расширенному вектору состояния $z(k)$, т.е.

$$u(k) = B^{-1}P[C_1; C_2] \begin{bmatrix} y(k) \\ \Theta(k-1) \end{bmatrix}, \quad (27)$$

где C_1 и C_2 — $n \times n$ -матрицы коэффициентов усиления соответственно пропорциональной ОС по состоянию и ОС по средней интегральной ошибке управления. Множитель $B^{-1}P$ введен для компенсации взаимодействий между управляющими сигналами.

Подставив уравнение (27) в уравнение (23) и учитывая уравнение (13), получим расширенное уравнение замкнутой системы управления:

$$\begin{aligned} \begin{bmatrix} y(k+1) \\ \Theta(k+1) \end{bmatrix} &= \begin{bmatrix} \Lambda & O \\ D & F \end{bmatrix} \begin{bmatrix} y(k) \\ \Theta(k-1) \end{bmatrix} + \begin{bmatrix} P^{-1}B \\ 0 \end{bmatrix} B^{-1}P[C_1; C_2] \begin{bmatrix} y(k) \\ \Theta(k-1) \end{bmatrix} = \\ &= \begin{bmatrix} \Lambda + C_1 & C_2 \\ D & F \end{bmatrix} \begin{bmatrix} y(k) \\ \Theta(k-1) \end{bmatrix}. \end{aligned} \quad (28)$$

Окончательно задача синтеза регулятора состояния формулируется следующим образом: найти такие $(n \times n)$ -матрицы коэффициентов усиления ОС C_1 и C_2 , чтобы при управлении (27) корни характеристического уравнения расширенной замкнутой системы (28) размещались в наперед заданных точках комплексной плоскости. Это позволит обеспечить асимптотическую устойчивость (сходимость) системы (28). Кроме того, матрицы C_1 и C_2 будем искать в виде диагональных $(n \times n)$ -матриц, что значительно упрощает реализацию управляющего устройства, т.е.

$$C_1 = \begin{bmatrix} C_{11} & \dots & 0 \\ 0 & \dots & C_{1n} \end{bmatrix}, \quad C_2 = \begin{bmatrix} C_{21} & \dots & 0 \\ 0 & \dots & C_{2n} \end{bmatrix}. \quad (29)$$

Асимптотическая устойчивость замкнутой дискретной системы (28) будет обеспечена, если собственные числа ее матрицы состояния

$$H = \begin{bmatrix} \Lambda + C_1 & C_2 \\ D & F \end{bmatrix} \quad (30)$$

будут лежать внутри единичного круга комплексной плоскости.

Для нахождения собственных чисел рассмотрим характеристическое уравнение системы (28):

$$\det \left(\begin{bmatrix} \Lambda + C_1 & C_2 \\ D & F \end{bmatrix} - \hat{\Lambda} \right) = \det [H - \hat{\Lambda}] = 0, \quad (31)$$

где $\hat{\Lambda}$ — диагональная $(2n \times 2n)$ -матрица:

$$\hat{\Lambda} = \begin{bmatrix} \lambda & & 0 \\ & \dots & \\ 0 & & \lambda \end{bmatrix}. \quad (32)$$

Произведем $2(n-1)$ перестановок строк и столбцов в матрице $[H - \hat{\Lambda}]$ размерности $2n \times 2n$ так, чтобы в результате получить следующую блочно-диагональную матрицу:

$$[H - \hat{\Lambda}] = \text{blokdiag}_{i=1, \dots, n} \begin{bmatrix} \lambda_1 + C_{1i} - \lambda & C_{2i} \\ \frac{1}{k+1} & \frac{k-1}{k+1} - \lambda \end{bmatrix}. \quad (33)$$

Так как четное число перестановок строк и столбцов матрицы не изменяет её определителя, то характеристическое уравнение (31) можно записать так:

$$\det [H - \hat{\Lambda}] = \prod_{i=1}^n \det \begin{bmatrix} \lambda_1 + C_{1i} - \lambda & C_{2i} \\ \frac{1}{k+1} & \frac{k-1}{k+1} - \lambda \end{bmatrix} =$$

$$= \prod_{i=1}^n \left(\lambda^2 - \left(\frac{\lambda_i k + \lambda_i + C_{1i} k + C_{1i} + k - 1}{k+1} \right) \lambda + \frac{(\lambda_i + C_{1i})(k-1) - C_{2i}}{k+1} \right) = 0. \quad (34)$$

Из (34) следует, что для нахождения коэффициентов усиления C_{1i} и C_{2i} ($i=1, \dots, n$) матриц обратных связей C_1 и C_2 достаточно определить способ нахождения коэффициентов C_{1i} и C_{2i} одного из сомножителей, который представляет собой полином второго порядка:

$$\lambda^2 - \left(\frac{\lambda_i k + \lambda_i + C_{1i} k + C_{1i} + k - 1}{k+1} \right) \lambda + \frac{(\lambda_i k - \lambda_i + C_{1i} k - C_{1i} - C_{2i})}{k+1} = 0. \quad (35)$$

$$i = 1, \dots, n.$$

Коэффициенты C_{1i} и C_{2i} определим так, чтобы корни характеристического уравнения замкнутой системы, т.е. корни всех полиномов второго порядка (35), имели наперед заданные значения $\tilde{\lambda}_{1i}, \tilde{\lambda}_{2i}$, $i = 1, \dots, n$. Пусть желаемый i -й характеристический полином второго порядка равен $\lambda^2 + \alpha_{1i} \lambda + \alpha_{2i} = 0$. Коэффициенты α_{1i}, α_{2i} этого полинома можно определить по заданным значениям корней $\tilde{\lambda}_{1i}, \tilde{\lambda}_{2i}$ с помощью формул Виета, или же путем раскрытия скобок в произведении $(\lambda - \tilde{\lambda}_{1i})(\lambda - \tilde{\lambda}_{2i})$, т.е.

$$(\lambda - \tilde{\lambda}_{1i})(\lambda - \tilde{\lambda}_{2i}) = \lambda^2 - (\tilde{\lambda}_{1i} + \tilde{\lambda}_{2i}) \lambda + \tilde{\lambda}_{1i} \tilde{\lambda}_{2i} = 0, \quad i = 1, \dots, n. \quad (36)$$

Приравняв коэффициенты при одинаковых степенях полиномов (35) и (36) для i -го полинома второго порядка, получим:

$$\tilde{\lambda}_{1i} + \tilde{\lambda}_{2i} = \frac{\lambda_i k + \lambda_i + C_{1i} k + C_{1i} + k - 1}{k+1}, \quad (37)$$

$$\tilde{\lambda}_{1i} \tilde{\lambda}_{2i} = \frac{\lambda_i k - \lambda_i + C_{1i} k - C_{1i} - C_{2i}}{k+1}. \quad (38)$$

Теперь из уравнений (37) и (38) можно выразить неизвестные коэффициенты C_{1i} и C_{2i} обратных связей через остальные известные величины:

$$C_{1i} = \tilde{\lambda}_{1i} + \tilde{\lambda}_{2i} - \lambda_i - \frac{k-1}{k+1}, \quad (39)$$

$$C_{2i} = \left(\tilde{\lambda}_{1i} + \tilde{\lambda}_{2i} - \frac{k-1}{k+1} \right) (k-1) - \tilde{\lambda}_{1i} \tilde{\lambda}_{2i} (k+1), \quad i = 1, \dots, n. \quad (40)$$

Таким образом, найденные уравнения (39) и (40) позволяют вычислить коэффициенты усиления C_{1i} и C_{2i} ($i = 1, \dots, n$) регулятора состояния (27) по заданным желаемым собственным числам $\tilde{\lambda}_{1i}$ и $\tilde{\lambda}_{2i}$ ($i = 1, \dots, n$) замкнутой системы (28). Очевидно, что коэффициенты C_{1i} и C_{2i} ($i = 1, \dots, n$) являются переменными во времени коэффициентами, т. к. зависят от номера шага управления k . Смысл этого состоит в том, что коэффициенты C_{1i} и C_{2i} формируют динамическую обратную связь системы управления.

Задав желаемые корни характеристического уравнения дискретной системы в начале координат комплексной плоскости, т.е. задав $\tilde{\lambda}_{1i} = \tilde{\lambda}_{2i} = 0$, $i = 1, \dots, n$, получим из (39) и (40) коэффициенты обратных связей, обеспечивающие максимальную степень устойчивости дискретной системы (23), т.е.

$$C_{1i} = -\left(\lambda_i + \frac{k-1}{k+1} \right), \quad i = 1, \dots, n, \quad (41)$$

$$C_{2i} = -\frac{(k-1)^2}{k+1}, \quad i = 1, \dots, n. \quad (42)$$

В этом случае матрицы обратных связей C_1 и C_2 уравнения регулятора состояния (27) будут иметь следующий вид:

$$C_1 = \begin{bmatrix} -\left(\lambda_1 + \frac{k-1}{k+1} \right) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & -\left(\lambda_n + \frac{k-1}{k+1} \right) \end{bmatrix}, \quad (43)$$

$$C_2 = \begin{bmatrix} -\frac{(k-1)^2}{k+1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & -\frac{(k-1)^2}{k+1} \end{bmatrix}. \quad (44)$$

Очевидно, что использование регулятора состояния в виде (27) приводит к разложению замкнутой системы (28) на независимые матричные подсистемы второго порядка:

$$\begin{bmatrix} y_i(k+1) \\ \theta_i(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_i + C_{1i} & C_{2i} \\ \frac{1}{k+1} & \frac{k-1}{k+1} \end{bmatrix} \begin{bmatrix} y_i(k) \\ \theta_i(k-1) \end{bmatrix}, \quad i = 1, \dots, n. \quad (45)$$

Изменяя коэффициенты C_{1i} и C_{2i} согласно уравнениям (39) и (40), можно как угодно менять собственные числа каждой замкнутой подсистемы (45) независимо друг от друга, т.е. изменение коэффициентов C_{1i} и C_{2i} будет влиять только на i -ю переменную y_i и соответствующую ей среднюю интегральную ошибку θ_i .

Возвращаясь в уравнении (28) к исходным переменным состояния, получим следующее уравнение состояния замкнутой системы:

$$\begin{bmatrix} x(k+1) \\ \Omega(k+1) \end{bmatrix} = \begin{bmatrix} P(\Lambda + C_1)P^{-1} & PC_2P^{-1} \\ PDP^{-1} & PFP^{-1} \end{bmatrix} \begin{bmatrix} x(k) \\ \Omega(k-1) \end{bmatrix}. \quad (46)$$

Аналогично получим и уравнение регулятора состояния (27) для исходных переменных:

$$\begin{aligned} u(k) &= B^{-1}P[C_1; C_2]P^{-1} \begin{bmatrix} x(k) \\ \Theta(k-1) \end{bmatrix} = \\ &= B^{-1}[PC_1P^{-1}x(k) + PC_2P^{-1}\Theta(k-1)] = K_1x(k) + K_2\Theta(k-1), \end{aligned} \quad (47)$$

где

$$K_1 = B^{-1}PC_1P^{-1},$$

$$K_2 = B^{-1}PC_2P^{-1}.$$

Матрица состояния замкнутой системы (28)

$$A_y = \begin{bmatrix} \Lambda - C_1 & C_2 \\ D & F \end{bmatrix} \quad (48)$$

полностью определяется характеристическим полиномом, который остается неизменным при эквивалентных преобразованиях. Поэтому матрица состояния системы (46)

$$A_x = \begin{bmatrix} P(\Lambda + C_1)P^{-1} & PC_2P^{-1} \\ PDP^{-1} & PFP^{-1} \end{bmatrix} \quad (49)$$

будет иметь такой же характеристический полином, что и матрица (48). Отсюда следует, что изменяя коэффициенты усиления C_{1i} и C_{2i} в матрицах C_1 и C_2 можно как угодно менять и собственные числа замкнутой системы (46) с первоначальными переменными состояния \bar{x} и их средними интегральными ошибками Ω .

Следует отметить, что совокупность всех собственных векторов, отвечающих заданному собственному числу, образует линейное подпространство в пространстве всех числовых векторов размерности n . Однако, в рассматриваемом случае, когда все собственные числа системы (1) предполагаются различными, каждое из этих подпространств одномерное, т.е. для каждого собственного числа соответствующий собственный вектор определен с точностью до числового множителя [5]. Эта одномерность вытекает из того, что ненулевые собственные векторы, отвечающие различным собственным числам, обязательно линейно независимы, а в n -мерном пространстве числовых векторов не может быть более n линейно независимых векторов. Поэтому, каноническое преобразование $P^{-1}AP = \Lambda$ в данном случае не зависит от численных значений элементов матрицы собственных векторов P и является однозначным. Отсюда следует, что результат предлагаемого способа вычисления матриц обратной связи K_1 и K_2 , при вычислении которых используется умножение на прямую и обратную матрицу собственных векторов

P , также является однозначным и независимым от численных значений матрицы P .

Следует также отметить, что если размерность вектора управления m меньше размерности вектора состояния n (матрица управления B имеет размерность $n \times m$), то изложенный метод модального управления может быть применен к первым m собственным числам матрицы A , а значит управляемыми будут только первые m переменных состояния.

Блок-схема системы управления вида (46) с регулятором состояния вида (47) приведена на рис. 1. На схеме выделены те вычисления, которые можно провести вне контура управления (в режиме off-line). Начальная установка индекса дискреты времени $k = 1$, используемая в блок-схеме системы управления (рис. 1), может осуществляться по условию

$$\|x(k)\| \leq \delta_{\text{доп1}}, \quad \|\Omega(k)\| \leq \delta_{\text{доп2}}, \quad (50)$$

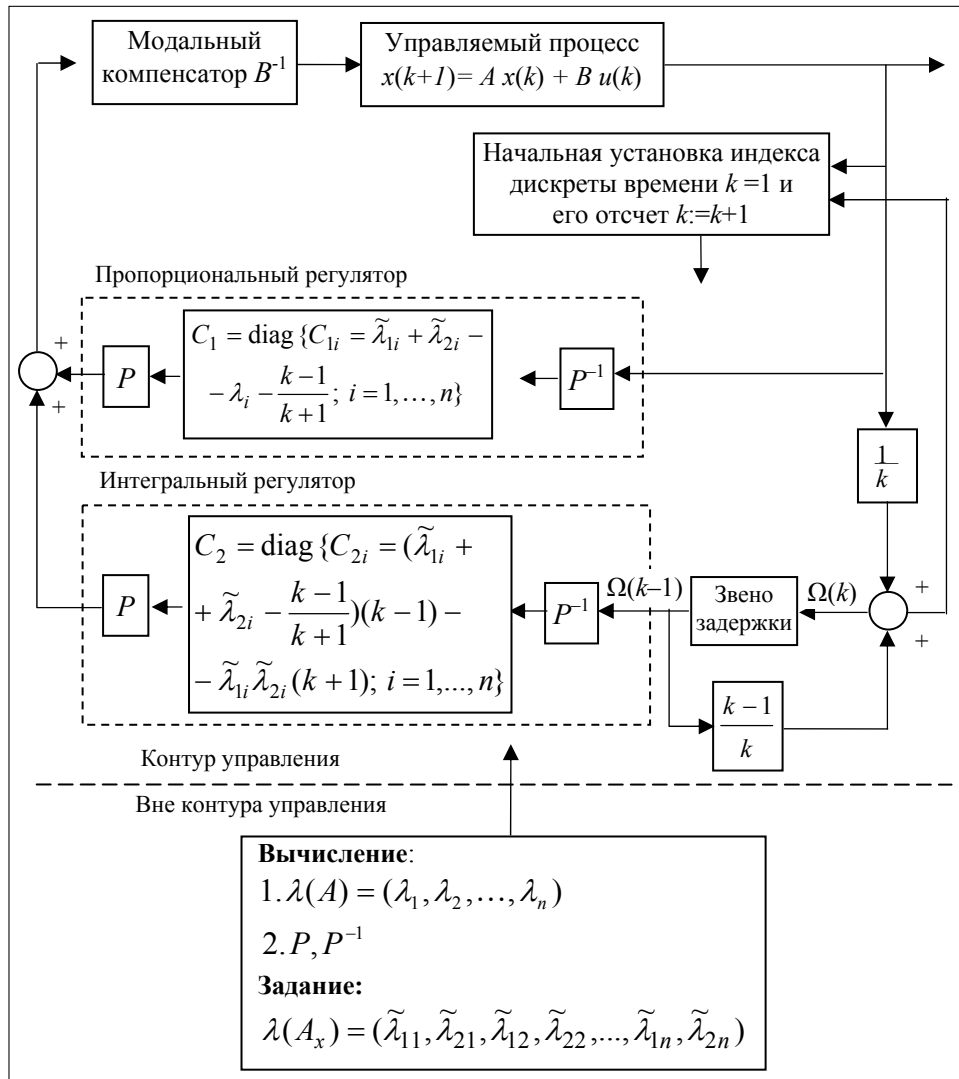


Рис. 1. Блок-схема модальной дискретной во времени системы управления

где $\delta_{\text{доп1}}$, $\delta_{\text{доп2}}$ — заданные допустимые значения нормы векторов состояния и средних интегральных ошибок состояния.

Пример. Для иллюстрации изложенного метода рассмотрим дискретную систему

$$x(k+1) = Ax(k) + Bu(k), \text{ где } A = \begin{bmatrix} 0,6 & 0,233 \\ -0,466 & -0,097 \end{bmatrix}; B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Пара матриц $[A; B]$ является управляемой. Задача состоит в определении матриц K_1 и K_2 таких, чтобы при наличии обратной связи вида $u(k) = K_1x(k) + K_2\Omega(k-1)$, собственные значения замкнутой системы располагались в точках $\tilde{\lambda}_{11} = \tilde{\lambda}_{12} = \tilde{\lambda}_{21} = \tilde{\lambda}_{22} = 0$.

Так как матрицы K_1 и K_2 согласно (43), (44) и (47) являются функциями шага управления k . Рассмотрим вычисление этих матриц на нескольких последовательных шагах управления, начиная с $k=1$ при $x(1) = [1; 1]^T$ и $\Omega(0) = [0; 0]^T$, т.е. рассмотрим вычисление переходной функции замкнутой системы управления. Одновременно на каждом шаге покажем вычисление переменных $u(k)$, $x(k)$ и $\Omega(k)$.

Собственные числа матрицы состояния A равны $\lambda_1 = 0,138$ и $\lambda_2 = 0,365$. Соответствующая матрица собственных векторов

$$P = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \text{ и } P^{-1} = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix}.$$

При этом значения собственных чисел и собственных векторов округлены, что дает несущественную погрешность (около 1–3%) в приведенных ниже вычислениях.

Найдем выражения для матриц обратной связи K_1 и K_2 , воспользовавшись формулой (47):

$$\begin{aligned} K_1 &= B^{-1}PC_1P^{-1} = \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} -0,138 - \left(\frac{k-1}{k+1}\right) & 0 \\ 0 & -0,365 - \left(\frac{k-1}{k+1}\right) \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix} = \\ &\cong \begin{bmatrix} -0,592 - \left(\frac{k-1}{k+1}\right) & -0,227 \\ 0,454 & 0,089 - \left(\frac{k-1}{k+1}\right) \end{bmatrix}; \\ K_2 &= B^{-1}PC_2P^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} -\frac{(k-1)^2}{k+1} & 0 \\ 0 & -\frac{(k-1)^2}{k+1} \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix} = \end{aligned}$$

$$= \begin{bmatrix} -\frac{(k-1)^2}{k+1} & 0 \\ 0 & -\frac{(k-1)^2}{k+1} \end{bmatrix}.$$

Тогда управляющий сигнал можно вычислить по формуле (47):

$$u(k) = \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} = K_1 \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + K_2 \begin{bmatrix} \delta_1(k-1) \\ \delta_2(k-1) \end{bmatrix}.$$

Вектор состояния согласно формуле (46) можно вычислить так:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 0,6 & 0,233 \\ -0,466 & -0,097 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}.$$

Средние интегральные ошибки управления, составляющие вектор $\Omega(k) = [\delta_1(k), \delta_2(k)]^T$ вычисляются по формуле (3).

Матрица состояния замкнутой системы согласно формуле (49) имеет вид:

$$A_x = \begin{bmatrix} P(\Lambda + C_1)P^{-1} & PC_2P^{-1} \\ PDP^{-1} & PFP^{-1} \end{bmatrix} =$$

$$= \begin{bmatrix} -\left(\frac{k-1}{k+1}\right) & 0 & -\frac{(k-1)^2}{k+1} & 0 \\ 0 & -\left(\frac{k-1}{k+1}\right) & 0 & -\frac{(k-1)^2}{k+1} \\ \frac{1}{k+1} & 0 & \frac{k-1}{k+1} & 0 \\ 0 & \frac{1}{k+1} & 0 & \frac{k-1}{k+1} \end{bmatrix}.$$

Последовательные во времени вычисления:

$k = 1$

$$K_1(1) = \begin{bmatrix} -0,592 & -0,227 \\ 0,454 & 0,089 \end{bmatrix}, \quad K_2(1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$u(1) = \begin{bmatrix} -0,819 \\ 0,543 \end{bmatrix}, \quad x(1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \Omega(1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$A_x(1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0,5 & 0 & 0 & 0 \\ 0 & 0,5 & 0 & 0 \end{bmatrix}, \quad \lambda(A_x) \cong \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$k = 2$

$$K_1(2) = \begin{bmatrix} -0,925 & -0,227 \\ 0,454 & -0,244 \end{bmatrix}, \quad K_2(2) = \begin{bmatrix} -0,333 & 0 \\ 0 & -0,333 \end{bmatrix},$$

$$u(2) = \begin{bmatrix} -0,34 \\ -0,322 \end{bmatrix}, \quad x(2) = \begin{bmatrix} 0,014 \\ -0,02 \end{bmatrix}, \quad \Omega(2) = \begin{bmatrix} 0,507 \\ 0,49 \end{bmatrix},$$

$$A_x^{(2)} = \begin{bmatrix} -0,333 & 0 & -0,333 & 0 \\ 0 & -0,333 & 0 & -0,333 \\ 0,333 & 0 & 0,333 & 0 \\ 0 & 0,333 & 0 & 0,333 \end{bmatrix}, \quad \lambda(A_x) \cong \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$k = 3$

$$K_1(3) = \begin{bmatrix} -1,092 & -0,227 \\ 0,454 & -0,411 \end{bmatrix}, \quad K_2(3) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$u(3) = \begin{bmatrix} -0,005 \\ -0,508 \end{bmatrix}, \quad x(3) = \begin{bmatrix} -0,336 \\ -0,327 \end{bmatrix}, \quad \Omega(3) = \begin{bmatrix} 0,226 \\ 0,218 \end{bmatrix},$$

$$A_x(3) = \begin{bmatrix} -0,5 & 0 & -1 & 0 \\ 0 & -0,5 & 0 & -1 \\ 0,25 & 0 & 0,5 & 0 \\ 0 & 0,25 & 0 & 0,5 \end{bmatrix}, \quad \lambda(A_x) \cong \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$k = 4$

$$K_1(4) = \begin{bmatrix} -1,192 & -0,227 \\ 0,454 & -0,511 \end{bmatrix}, \quad K_2(4) = \begin{bmatrix} -1,8 & 0 \\ 0 & -1,8 \end{bmatrix},$$

$$u(4) = \begin{bmatrix} 0,073 \\ -0,378 \end{bmatrix}, \quad x(4) = \begin{bmatrix} -0,328 \\ -0,32 \end{bmatrix}, \quad \Omega(4) = \begin{bmatrix} 0,087 \\ 0,083 \end{bmatrix},$$

$$A_x(4) = \begin{bmatrix} -0,6 & 0 & -1,8 & 0 \\ 0 & -0,6 & 0 & -1,8 \\ 0,2 & 0 & 0,6 & 0 \\ 0 & 0,2 & 0 & 0,6 \end{bmatrix}, \quad \lambda(A_x) \cong \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$k = 5$

$$K_1(5) = \begin{bmatrix} -1,259 & -0,227 \\ 0,454 & -0,578 \end{bmatrix}, \quad K_2(5) = \begin{bmatrix} -2,667 & 0 \\ 0 & -2,667 \end{bmatrix},$$

$$u(5) = \begin{bmatrix} 0,071 \\ -0,199 \end{bmatrix}, \quad x(5) = \begin{bmatrix} -0,198 \\ -0,194 \end{bmatrix}, \quad \Omega(5) = \begin{bmatrix} 0,03 \\ 0,028 \end{bmatrix},$$

$$A_x(5) = \begin{bmatrix} -0,667 & 0 & -2,667 & 0 \\ 0 & -0,667 & 0 & -2,667 \\ 0,167 & 0 & 0,667 & 0 \\ 0 & 0,167 & 0 & 0,667 \end{bmatrix}, \quad \lambda(A_x) \cong \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

$k = 6$

$$K_1(6) = \begin{bmatrix} -1,306 & -0,227 \\ 0,454 & -0,625 \end{bmatrix}, \quad K_2(6) = \begin{bmatrix} -3,571 & 0 \\ 0 & -3,571 \end{bmatrix},$$

$$u(6) = \begin{bmatrix} 0,034 \\ -0,087 \end{bmatrix}, \quad x(6) = \begin{bmatrix} -0,093 \\ -0,088 \end{bmatrix}, \quad \Omega(6) = \begin{bmatrix} 0,01 \\ 0,009 \end{bmatrix},$$

$$A_x(6) = \begin{bmatrix} -0,714 & 0 & -3,571 & 0 \\ 0 & -0,714 & 0 & -3,571 \\ 0,143 & 0 & 0,714 & 0 \\ 0 & 0,143 & 0 & 0,714 \end{bmatrix}, \quad \lambda(A_x) \cong \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

На основании полученных в примере результатов для шести расчетных шагов управления, построены соответствующие графики переходных процессов по каждому из двух каналов управления. На рис. 2 приведены пере-

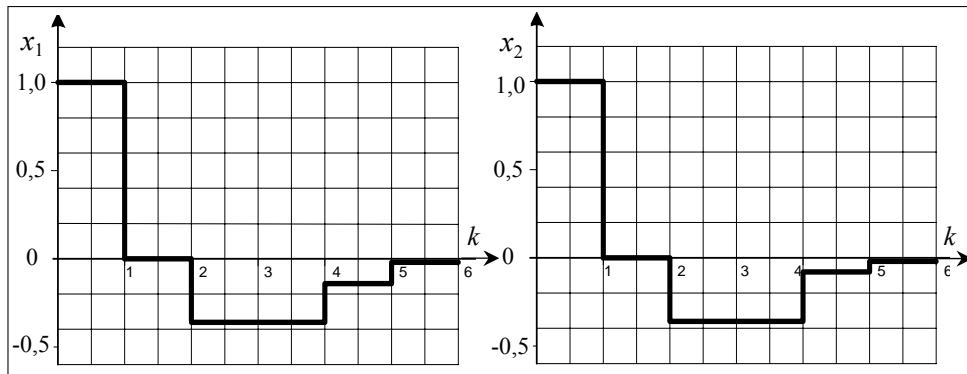


Рис. 2. Переходные дискретные функции системы

ходные дискретные функции системы для переменных состояния x_1 и x_2 . На рис. 3 приведены дискретные процессы изменения управляющих сигналов u_1 и u_2 . На рис. 4 приведены дискретные графики изменения средних интегральных ошибок управления δ_1 и δ_2 . Анализ полученных в примере результатов показывает, что замкнутая модальная система управления под

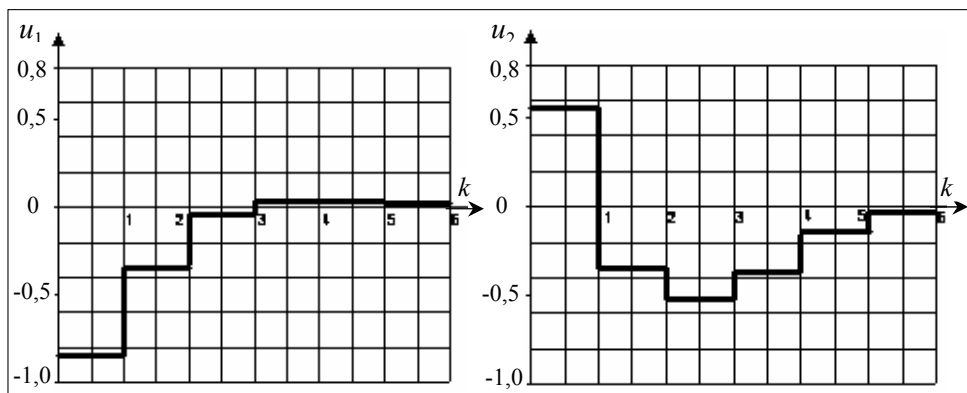


Рис. 3. Дискретные процессы изменения управляющих сигналов

действием управлений u_1 и u_2 имеет заданный спектр и небольшое перерегулирование при экспоненциальном затухании средних интегральных ошибок управления δ_1 и δ_2 .

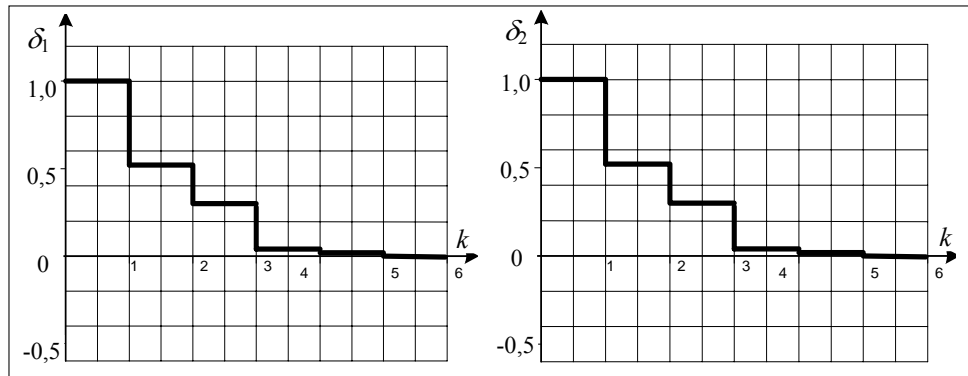


Рис. 4. Дискретные процессы изменения средних интегральных ошибок управления

ВЫВОДЫ

Для дискретной во времени системы управления разработан метод синтеза модального регулятора состояния, который позволяет обеспечить произвольную, в том числе и максимальную степень устойчивости замкнутой системы в пространстве переменных состояния и их средних интегральных значений. Полученный закон управления позволяет обеспечить экспоненциальный характер переходной функции для средних интегральных ошибок управления системы при наименьшем перерегулировании переменных состояния и минимально возможном при этом времени сходимости переходной функции для переменных состояния к установившейся ошибке. Регулятор имеет достаточно простую структуру, что упрощает его физическую реализацию, а использование средних интегральных ошибок управления системы в цепи обратной связи позволяет использовать потенциально полезное влияние случайных возмущений в реальных условиях работы на процесс поддержания их заданных значений.

ЛИТЕРАТУРА

1. Филатов А.Г. Стабилизация электрических нагрузок в электроэнергетических системах // Технічна електродинаміка. Тематичний випуск «Проблеми сучасної електротехніки», частина 5. — 2006. — С. 3–8.
2. Воронов А.А. Введение в динамику сложных управляемых систем. — М.: Наука, 1985. — 352 с.
3. Рей У. Методы управления технологическими процессами. — М.: Мир, 1983. — 368 с.
4. Куо Б. Теория и проектирование цифровых систем управления. — М.: Машиностроение, 1986. — 448 с.
5. Мышкис А. Д. Лекции по высшей математике. — М.: Наука, 1973. — 640 с.

Поступила 02.06.2008

MINIMAX RECURSIVE STATE ESTIMATION FOR LINEAR DISCRETE-TIME DESCRIPTOR SYSTEMS

S. ZHUK

This paper describes an approach to the online state estimation of systems described by a general class of linear noncausal time-varying difference descriptor equations subject to uncertainties. An approach is based on the notions of a linear minimax estimation and an index of causality introduced here for singular difference equations. The online minimax observer is derived by the application of the dynamical programming and Moore's pseudoinverse theory to the minimax estimation problem.

INTRODUCTION

There is a number of physical and engineering objects most naturally modelled as systems of differential and algebraic equations (DAEs) or descriptor systems: microwave circuits [1], flexible-link planar parallel platforms [2] and image recognition problems (noncasual image modeling) [3]. DAEs arise in economics [4]. Also nonlinear differential-algebraic systems are studied with help of linear DAEs by linearization: a batch chemical reactor model [5].

On the other hand there are many papers devoted to the mathematical processing of data, obtained from the measuring device during an experiment. In particular, the problem of observer design for continuous-time DAEs was considered in [7] and discrete-time case was studied in [8]–[9]. The minimax state estimation for uncertain linear dynamical systems was investigated in [10]. Other approaches to state estimation with set-membership description of uncertainty were discussed in [12]–[14].

In [6] authors derive a so-called «3-block» form for the optimal filter and a corresponding 3-block Riccati equation using a maximum likelihood approach. A filter is obtained for a general class of time-varying descriptor models. Measurements are supposed to contain a noise with Gaussian distribution. The obtained recursion is stated in terms of the 3-block matrix pseudoinverse.

In [8] the filter recursion is represented in terms of a deterministic data fitting problem solution. The authors introduce an explicit form of the 3-block matrix pseudoinverse for a descriptor system with a special structure, so their filter coincides with obtained in [6].

In this paper we study an observer design problem for a general class of linear noncasual time-varying descriptor models with no restrictions on system structure. Suppose we are given an exact mathematical model of some real process and the vector x_k describes the system output at the moment k in the corresponding state space of the system. Also successive measurements $y_0 \dots y_k \dots$ of the system output x_k are supposed to be available with the noise $g_0 \dots g_k \dots$ of an uncertain nature. (For instance we do not have a-priory infor-

mation about its distribution.) Further assume that the system input f_k , start point q and noise g_k are arbitrary elements of the given set G . The aim of this paper is to design a minimax observer $k \mapsto \hat{x}_k$ that gives an online guaranteed estimation of the output x_k on the basis of measurements y_k and the structure of G . In [9] minimax estimations were derived from the 2-point boundary value problem with conditions at $i=0$ (start point) and $i=k$ (end point). Hence a recalculation of the whole history $\hat{x}_0 \dots \hat{x}_k$ is required if the moment k changes. Here we derive the observer $(k, y_k) \mapsto \hat{x}_k$ by applying dynamical programming methods to the minimax estimation problem similar to the posed one in [9]. We construct a map \hat{x} that takes (k, y_k) to \hat{x}_k making it possible to assign a unique sequence of estimations $\hat{x}_0 \dots \hat{x}_k \dots$ to given sequence of observations $y_0 \dots y_k \dots$ in the real time. A resulting filter recursion is stated in terms of pseudoinverse of positive semi-defined $n \times n$ -matrices.

Minimax estimation problem

Assume that $x_k \in \mathbf{R}^n$ is described by the equation

$$F_{k+1}x_{k+1} - C_k x_k = f_k, \quad k=0,1,\dots, \quad (1)$$

with initial condition

$$F_0 x_0 = q, \quad (2)$$

and y_k is given by

$$y_k = H_k x_k + g_k, \quad k=0,1,\dots, \quad (3)$$

where F_k, C_k are $m \times n$ -matrices, H_k is $p \times n$ -matrix. Since we deal with descriptor system we see that for any k there is a set of vectors $x_1^0 \dots x_k^0$ satisfying (1) while $f_i = 0, q = 0$. Thus the undefined inner influence caused, by $x_1^0 \dots x_k^0$, may appear in the system's output. Also we suppose the initial condition q , input $\{f_k\}$ and noise $\{g_k\}$ to be unknown elements of the given set. (Here and after (\cdot, \cdot) denotes an inner product in an appropriate Euclidean space, $\|\mathbf{x}\| = (x, x)^{1/2}$.)

$$\begin{aligned} \Gamma = & \left\{ (q, \{f_k\}, \{g_k\}) : G(q, \{f_k\}, \{g_k\}) = \right. \\ & \left. = (Sq, q) + \sum_0^\infty (S_k f_k, f_k) + (R_k g_k, g_k) \leq 1 \right\}, \quad (4) \end{aligned}$$

where S, S_k, R_k are some symmetric positive-defined weight matrices with appropriate dimensions. The trick is to fix any N -partial sum of (4) so that $(q, \{f_k\}, \{g_k\})$ belongs to

$$\mathbf{G}^N := \left\{ (q, \{f_k\}, \{g_k\}) : (Sq, q) + \sum_{k=0}^{N-1} (S_k f_k, f_k) + \sum_{k=0}^N (R_k g_k, g_k) \leq 1 \right\}. \quad (5)$$

Then we derive the estimation $\hat{x}_N = v(N, y_N, \hat{x}_{N-1})$ considering a minimax estimation problem for \mathbf{G}^N . Let us denote by \mathbf{N} a set of all $(\{x_k\}, q, \{f_k\})$ such that (1) is held. The set \mathbf{G}_y^N is said to be *a-posteriori set*, where

$$\mathbf{G}_y^N := \{ \{x_k\} : (\{x_k\}, q, \{f_k\}) \in \mathbf{N}, (q, \{f_k\}, \{y_k - H_k x_k\}) \in \mathbf{G}^N \}. \quad (6)$$

It follows from the definition that \mathbf{G}_y^N consists of all possible $\{x_k\}$, causing an output $\{y_k\}$, while $(q, \{f_k\}, \{g_k\})$ runs through \mathbf{G}^N . Thus, it's naturally to look for estimation x_N of **only** among the elements of $P_N(\mathbf{G}_y^N)$, where P_N denotes the projection that takes $\{x_0 \dots x_N\}$ to x_N .

Definition 1. A linear function (ℓ, \hat{x}_N) is called a minimax a-posteriori estimation if the following condition holds:

$$\inf_{\{\tilde{x}_k\} \in \mathbf{G}_y^N} \sup_{\{x_k\} \in \mathbf{G}_y^N} |(\ell, x_N) - (\ell, \tilde{x}_N)| = \sup_{\{x_k\} \in \mathbf{G}_y^N} |(\ell, x_N) - (\ell, \hat{x}_N)|.$$

The non-negative number

$$\hat{\sigma}(\ell, N) = \sup_{\{x_k\} \in \mathbf{G}_y^N} |(\ell, x_N) - (\ell, \hat{x}_N)|$$

is called a minimax a-posteriori error in the direction ℓ . A map

$$N \mapsto I_N = \dim \{ \ell \in \mathbf{R}^n : \hat{\sigma}(\ell, N) < +\infty \}$$

is called an index of causality for the pair of systems (1)–(3).

Now we say that a minimax estimation problem is to construct an a-posteriori linear minimax estimation (ℓ, \hat{x}_N) for the system (1) on the basis of the measurements (3) and a-posteriori set \mathbf{G}_y^N . A solution of the minimax estimation problem in the form of a recursive map $k \mapsto (\ell, \hat{x}_N)$ is presented in the next section.

Minimax online observer

Denote by $k \mapsto Q_k$ a recursive map that takes each natural number k to the matrix Q_k , where

$$\begin{aligned} Q_k &= H'_k R_k H_k + F'_k [S_{k-1} - S_{k-1} C_{k-1} W_{k-1}^+ C'_{k-1} S_{k-1}] F_k, \\ Q_0 &= F'_0 S F_0 + H'_0 R_0 H_0, W_k = Q_k + C'_k S_k C_k. \end{aligned} \quad (7)$$

Let $k \mapsto r_k$ be a recursive map that takes each natural number k to the vector $r_k \in \mathbf{R}^n$, where

$$r_k = F'_k S_{k-1} C_{k-1} W_{k-1}^+ r_{k-1} + H'_k R_k y_k, \quad r_0 = H'_0 R_0 y_0 \quad (8)$$

and to each natural number $i \in$ assign a number α_i , where

$$\alpha_i = \alpha_{i-1} + (R_i y_i, y_i) - (W_{i-1}^+ r_{i-1}, r_{i-1}), \quad \alpha_0 = (S g, g) + (R_0 y_0, y_0). \quad (9)$$

The main result of this paper is formulated in the next theorem.

Theorema (minimax recursive estimation). Suppose we are given a natural number N and a vector $\ell \in \mathbf{R}^n$. Then a necessary and sufficient condition for a minimax a-posteriori error $\hat{\sigma}(\ell, N)$ to be finite is that

$$Q_N^+ Q_N \ell = \ell. \quad (10)$$

Under this condition we have

$$\hat{\sigma}(\ell, N) = [1 - \alpha_N + (Q_N^+ r_N, r_N)]^{\frac{1}{2}} (Q_N^+ \ell, \ell)^{\frac{1}{2}} \quad (11)$$

and

$$(\ell, \hat{x}_N) = (\ell, Q_N^+ r_N). \quad (12)$$

Corollary 1. The index of causality I_N for the pair of systems (1)–(3) can be represented as $I_N = \text{rank}(Q_N)$.

Corollary 2 (minimax observer). The online minimax observer is given by $k \mapsto \hat{x}_k = Q_k^+ r_k$ and (we assume here that $1/0 = +\infty$.)

$$\begin{aligned} \hat{\rho}(N) &= \min_{\{x_k\} \in \mathbf{G}_y^N} \max_{\{\tilde{x}_k\} \in \mathbf{G}_y^N} \|x_N - \tilde{x}_N\|^2 = \\ &= \max_{\{x_k\} \in \mathbf{G}_y^N} \|x_N - \hat{x}_N\|^2 = \frac{[1 - \alpha_N + (Q_N \hat{x}_N, \hat{x}_N)]}{\min_i \{\lambda_i(N)\}} \end{aligned} \quad (13)$$

where $\lambda_i(N)$ are eigenvalues of Q_N . In this case all possible realisations of the state vector x_N of (1) fill the ellipsoid $P_N(\mathbf{G}_y^N) \subset \mathbf{R}^n$, where

$$P_N(\mathbf{G}_y^N) = \{x: (Q_N x, x) - 2(Q_N \hat{x}_N, x) + \alpha_N \leq 1\}. \quad (14)$$

Remark 1. If $\lambda_{\min}(H_k' R_k H_k)$ grows for $k = i, i+1, \dots$ then the minimax estimation error $\hat{\rho}(k)$ becomes smaller causing \hat{x}_k to get closer to the real state vector x_k .

In [8] Kalman's filtering problem for descriptor systems was investigated from the deterministic point of view. Authors recover Kalman's recursion to the time-variant descriptor system by a deterministic least square fitting problem over the entire trajectory: find a sequence $\{\hat{x}_{0|k}, \dots, \hat{x}_{k|k}\}$ that minimises the following fitting error cost

$$\begin{aligned} J_k(\{x_{i|k}\}_0^k) &= \|F_0 x_{0|k} - g\|^2 + \|y_0 - H_0 x_{0|k}\|^2 + \\ &+ \sum_{i=1}^k \|F_i x_{i|k} - C_{i-1} x_{i-1|k}\|^2 + \|y_i - H_i x_{i|k}\|^2 \end{aligned}$$

assuming that the $\text{rank} \begin{bmatrix} F_k \\ H_k \end{bmatrix} \equiv n$. According to [8] the successive optimal estimates $\{\hat{x}_{0|k}, \dots, \hat{x}_{k|k}\}$ resulting from the minimisation of J_k can be found from the recursive algorithm

$$\begin{aligned} \hat{x}_{k|k} &= P_{k|k} F_k' (E + C_{k-1} P_{k-1|k-1} C_{k-1}')^{-1} C_{k-1} \hat{x}_{k-1|k-1} + \\ &+ P_{k|k} H_k' R_k y_k, \hat{x}_{0|0} = P_{0|0} (F_0' q + H_0' y_0), \\ P_{k|k} &= (F_k' (E + C_{k-1} P_{k-1|k-1} C_{k-1}')^{-1} F_k + H_k' H_k)^{-1}, \\ P_{0|0} &= (F_0' F_0 + H_0' H_0)^{-1}. \end{aligned} \tag{15}$$

Corollary 3 (Kalman's filter recursion). Suppose the rank $\begin{bmatrix} F_k \\ H_k \end{bmatrix} \equiv n$, and let $k \mapsto r_k$ be a recursive map that takes each natural number k to the vector $r_k \in \mathbf{R}^n$, where

$$\begin{aligned} r_k &= H_k' y_k + F_k' C_{k-1} (C_{k-1}' C_{k-1} + Q_{k-1})_{k-1}^+ r_{k-1}, \\ r_0 &= F_0' q + H_0' y_0. \end{aligned} \tag{16}$$

Then $Q_k^+ r_k = \hat{x}_{k|k}$ for each $k \in \mathbf{N}$, where $\hat{x}_{k|k}$ is given by (15) and $I_k = n$.

Acknowledgements. It is a pleasure to thank Prof. A. Nakonechniy and Dr. V. Pichkur for insightful discussions about the key ideas presented in this paper.

Proof of Theorem. By definition, put

$$\mathbf{F} = \begin{pmatrix} F_0 & 0_{mn} & 0_{mn} & \cdots & 0_{mn} \\ 0_{mn} & & & & \\ -C_0 & F_1 & 0_{mn} & \cdots & 0_{mn} \\ 0_{mn} & & & & \\ 0_{mn} & -C_1 & F_2 & \cdots & 0_{mn} \\ 0_{mn} & & & & \\ \vdots & \vdots & \vdots & \cdots & \\ 0_{mn} & 0_{mn} & 0_{mn} & \cdots & -C_{N-1} \\ F_N \end{pmatrix},$$

$$\mathbf{H} = \begin{pmatrix} H_0 & 0_{pn} & \cdots & 0_{pn} \\ 0_{pn} & H_1 & \cdots & 0_{pn} \\ \vdots & \vdots & \cdots & \vdots \\ 0_{pn} & 0_{pn} & \cdots & H_N \end{pmatrix}, \mathbf{X} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \mathbf{F} = \begin{bmatrix} q \\ f_0 \\ f_1 \\ \vdots \\ f_{N-1} \end{bmatrix}, \mathbf{G} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \\ g_N \end{bmatrix}.$$

By direct calculation we obtain $(\ell, x_N) = (\mathbf{L}, \mathbf{X})$,

$$\mathbf{G}_y^N = \{ \mathbf{X} : \|\mathbf{F}\mathbf{X}\|_1^2 + \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_2^2 \leq 1 \},$$

where $\|\mathbf{F}\|_1^2 = (Sq, q) + \sum_0^{N-1} (S_k f_k, f_k)$, $\|\cdot\|_2$ is induced by R_k on the same way. This implies

$$\sup_{\{x_k\} \in \mathbf{G}_y^N} |(\ell, x_N - \tilde{x}_N)| = \sup_{\mathbf{X} \in \mathbf{G}_y^N} |(\mathbf{L}, \mathbf{X}) - (\mathbf{L}, \tilde{\mathbf{X}})|.$$

Denote by \mathbf{M} the set $R[\mathbf{F}' \mathbf{H}']$. We obviously get

$$\mathbf{L} \in \mathbf{M} \Leftrightarrow \sup_{\mathbf{X} \in \mathbf{G}_y^N} |(\mathbf{L}, \mathbf{X}) - (\mathbf{L}, \tilde{\mathbf{X}})| < +\infty.$$

The application of Corollary 4 yields (10). Consider a vector $\mathbf{L} \in \mathbf{M}$. Clearly

$$\inf_{\mathbf{X} \in \mathbf{G}_y^N} (\mathbf{L}, \mathbf{X}) \leq (\mathbf{L}, \mathbf{X}) \leq \sup_{\mathbf{X} \in \mathbf{G}_y^N} (\mathbf{L}, \mathbf{X}), \quad \mathbf{X} \in \mathbf{G}_y^N.$$

Let c denotes $\frac{1}{2}(\sup_{\mathbf{X} \in \mathbf{G}_y^N} (\mathbf{L}, \mathbf{X}) + \inf_{\mathbf{X} \in \mathbf{G}_y^N} (\mathbf{L}, \mathbf{X}))$. Therefore

$$\sup_{\mathbf{X} \in \mathbf{G}_y^N} |(\mathbf{L}, \mathbf{X}) - (\mathbf{L}, \tilde{\mathbf{X}})| = \frac{1}{2}(s(\mathbf{L} | \mathbf{G}_y^N) + s(-\mathbf{L} | \mathbf{G}_y^N)) + |c - (\mathbf{L}, \tilde{\mathbf{X}})|$$

hence

$$\hat{\sigma}(\ell, N) = \frac{1}{2}(s(\mathbf{L} | \mathbf{G}_y^N) + s(-\mathbf{L} | \mathbf{G}_y^N)), \quad (\ell, \hat{x}_N) = \frac{1}{2}(s(\mathbf{L} | \mathbf{G}_y^N) - s(-\mathbf{L} | \mathbf{G}_y^N)), \quad (17)$$

where $s(\cdot | \mathbf{G}_y^N)$ denotes the support function of \mathbf{G}_y^N . Clearly, \mathbf{G}_y^N is a convex closed set. Hence the equality $(\mathbf{L}, \tilde{\mathbf{X}}) = (\ell, \hat{x}_N)$ is held for some $\tilde{\mathbf{X}} \in \mathbf{G}_y^N$. Thus, to conclude the proof we have to calculate $s(\mathbf{L}, \mathbf{G}_y^N)$. Let

$$\mathbf{G}_0^N = \{\mathbf{X} : \|\mathbf{F}\mathbf{X}\|^2 + \|\mathbf{H}\mathbf{X}\|^2 \leq \beta_N\}, \quad (18)$$

where $\beta_N = 1 - \alpha_N + (Q_N^+ r_N, r_N) \geq 0$.

Lemma 1.

$$s(\mathbf{L}, \mathbf{G}_y^N) = (\ell, Q_N^+ r_N) + s(\mathbf{L} | \mathbf{G}_0^N). \quad (19)$$

It follows from the definition of \mathbf{G}_0^N that $s(\mathbf{L} | \mathbf{G}_0^N) = s(-\mathbf{L} | \mathbf{G}_0^N)$ hence (17) implies

$$(\ell, \hat{x}_N) = (\ell, Q_N^+ r_N), \quad \hat{\sigma}(\ell) = s(\mathbf{L} | \mathbf{G}_0^N).$$

The application of Lemma 2 completes the proof.

Lemma 2.

$$s(\mathbf{L} | \mathbf{G}_0^N) = \begin{cases} \sqrt{\beta_N} (Q_N^+ \ell, \ell)^{\frac{1}{2}}, & [E - Q_N^+ Q_N] \ell = 0, \\ +\infty, & [E - Q_N^+ Q_N] \ell \neq 0. \end{cases} \quad (20)$$

Let r_k denote \mathbf{R}^n — valued recursive map

$$\begin{aligned} r_k &= F_k' (S_{k-1} - S_{k-1} C_{k-1} P_{k-1}^+ C_{k-1}' S_{k-1}) f_{k-1} + F_k' S_{k-1} C_{k-1} W_{k-1}^+ r_{k-1} + H_k' R_k y_k, \\ r_0 &= F_0' S q + H_0' R_0 y_0, \quad P_k = C_k' S_k C_k + Q_k, \end{aligned} \quad (21)$$

and set

$$\mathbf{J}(\{x_k\}) = \|F_0 x_0 - g\|_S^2 + \|y_0 - H_0 x_0\|_0^2 + \sum_{k=1}^N \|F_k x_k - C_{k-1} x_{k-1} - f_{k-1}\|_{k-1}^2 + \|y_k - H_k x_k\|_k^2,$$

where $\|g\|_S^2 = (Sg, g)$, $\|f_k\|_k^2 = (S_k f_k, f_k)$, $\|y_i\|_i^2 = (R_i y_i, y_i)$.

Lemma 3. Let $x \mapsto \hat{x}_k$ be a recursive map that takes any $k \in \text{natural}$ to $\hat{x}_k \in \mathbf{R}^n$, where

$$\hat{x}_k = P_k^+(C'_k S_k (F_{k+1} \hat{x}_{k+1} - f_k) + r_k), \quad \hat{x}_N = Q_N^+ r_N. \quad (22)$$

Then

$$\min_{\{x_k\}} \mathbf{J}(\{x_k\}) = \mathbf{J}(\{\hat{x}_k\}).$$

Proof. By definition put $\Phi(x_0) := \|F_0 x_0 - g\|_S^2 + \|y_0 - H_0 x_0\|_0^2$, $\Phi_i(x_i, x_{i+1}) := \|F_{i+1} x_{i+1} - C_i x_i - f_i\|_i^2 + \|y_{i+1} - H_{i+1} x_{i+1}\|_{i+1}^2$.

Then we obviously get

$$\mathbf{J}(\{x_k\}) = \Phi(x_0) + \sum_{i=0}^{N-1} \Phi_i(x_i, x_{i+1}). \quad (23)$$

Let us apply a modification of Bellman's method (so-called «Kyivsky vinyk» method) to the nonlinear programming task

$$\mathbf{J}(\{x_k\}) \rightarrow \min_{\{x_k\}}.$$

By definition put

$$\ell_1(x_1) := \min_{x_0} \{\Phi(x_0) + \Phi_0(x_0, x_1)\}.$$

Using (7) and (21) one can get

$$\Phi(x_0) = (Q_0 x_0, x_0) - 2(r_0, x_0) + \alpha_0 \geq 0, \quad \alpha_0 := \|g\|_S^2 + \|y_0\|_0^2.$$

On the other hand it's clear that

$$\ell_1(x_1) = \Phi(\hat{x}_0) + \Phi_0(\hat{x}_0, x_1) = (Q_1 x_1, x_1) - 2(r_1, x_1) + \alpha_1 \geq 0,$$

where $\hat{x}_0 = P_0^+(r_0 + C'_0 S_0 (F_1 x_1 - f_0))$

$$\alpha_1 := \alpha_0 + \|y_1\|_1^2 + \|f_0\|_0^2 - (P_0^+(r_0 - C'_0 S_0 f_0), r_0 - C'_0 S_0 f_0).$$

Considering $\ell_1(x_1)$ as an induction base and assuming that

$$\begin{aligned} \ell_{i-1}(x_{i-1}) &= \min_{x_{i-2}} \{\Phi_{i-2}(x_{i-2}, x_{i-1}) + \ell_{i-2}(x_{i-2})\} = \\ &= (Q_{i-1} x_{i-1}, x_{i-1}) - 2(r_{i-1}, x_{i-1}) + \alpha_{i-1} \end{aligned}$$

now we are going to prove that

$$\ell_i(x_i) = \min_{x_{i-1}} \{\Phi_{i-1}(x_{i-1}, x_i) + \ell_{i-1}(x_{i-1})\} = (Q_i x_i, x_i) - 2(r_i, x_i) + \alpha_i. \quad (24)$$

Note that [11] for any convex function $(x, y) \mapsto f(x, y)$

$$y \mapsto \min\{f(x, y) \mid (x, y) : P(x, y) = y\}, \quad P(a, b) = b$$

is convex. Thus taking into account the definition of $\ell_1(x_1)$ one can prove by induction that ℓ_{i-1} is convex and

$$\Phi_{i-1}(x_{i-1}, x_i) + \ell_{i-1}(x_{i-1}) \geq 0.$$

Hence (the function $x \mapsto (Ax, x) - 2(x, q) + c$ is convex if $A = A' \geq 0$) $Q_{i-1} \geq 0$, the set of global minimums Ψ_{i-1} of the quadratic function.

$$x_{i-1} \mapsto \Phi_{i-1}(x_{i-1}, x_i) + (Q_{i-1} x_{i-1}, x_{i-1}) - 2(r_{i-1}, x_{i-1}) + \alpha_{i-1}$$

is non-empty and $\hat{x}_{i-1} \in \Psi_{i-1}$, where (The vector \hat{x}_{i-1} has the smallest norm among other points of the minimum.)

$$\hat{x}_{i-1} = (Q_{i-1} + C'_{i-1} S_{i-1} C_{i-1})^+ (C'_{i-1} S_{i-1} (F_i x_i - f_{i-1}) + r_{i-1}).$$

This implies

$$\ell_i(x_i) = \Phi_{i-1}(\hat{x}_{i-1}, x_i) + \ell_{i-1}(\hat{x}_{i-1}) = (Q_i x_i, x_i) - 2(r_i, x_i) + \alpha_i,$$

where

$$\begin{aligned} \alpha_i &= \alpha_{i-1} + (R_i y_i, y_i) + (S_{i-1} f_{i-1}, f_{i-1}) - \\ &\quad - (P_{i-1}^+ (r_{i-1} - C'_{i-1} S_{i-1} f_{i-1}), r_{i-1} - C'_{i-1} S_{i-1} f_{i-1}). \end{aligned}$$

Therefore, we obtain

$$\min_{x_N} \ell_N(x_N) = \ell_N(\hat{x}_N) = \alpha_N - (r_N, Q_N^+ r_N), \quad \hat{x}_N = Q_N^+ r_N$$

so that $\min_{\{x_k\}} \mathbf{J}(\{x_k\}) = \mathbf{J}(\{\hat{x}_k\})$.

Corollary 4. Suppose $\mathbf{L} = [0 \dots \ell]$; then

$$\mathbf{L} \in \mathbf{R}[\mathbf{F}' \mathbf{H}'] \Leftrightarrow [E - Q_N^+ Q_N] \ell = 0$$

and

$$\|[\mathbf{F}' \mathbf{H}']^+ \mathbf{L}\|^2 = (Q_N^+ \ell, \ell).$$

Proof. Suppose $S_k = E$, $R_k = E$ for a simplicity. If $\mathbf{L} \in \mathbf{R}[\mathbf{F}' \mathbf{H}']$ then

$$F'_N z_N + H'_N u_N = \ell, \quad F'_k z_k + H'_k u_k - C'_k z_{k+1} = 0 (*),$$

for some $z_k \in \mathbf{R}^m$, $u_k \in \mathbf{R}^p$. Let's find the projection $\{(\hat{z}_k, \hat{u}_k)\}_{k=0}^N$ of the vector $\{(z_k, u_k)\}_{k=0}^N$ onto the range of the matrix $\begin{bmatrix} \mathbf{F} \\ \mathbf{H} \end{bmatrix}$. Lemma 3 implies

$$\hat{z}_0 = F_0 \hat{x}_0, \quad \hat{z}_k = F_k \hat{x}_k - C_{k-1} \hat{x}_{k-1}, \quad \hat{u}_k = H_k \hat{x}_k, (**)$$

where

$$\begin{aligned} \hat{x}_k &= P_k^+(C'_k F_{k+1} \hat{x}_{k+1} + r_k - C'_k z_{k+1}), \quad \hat{x}_N = Q_N^+ r_N, \\ r_k &= F'_k C_{k-1} P_{k-1}^+ r_{k-1} + F'_k (E - C_{k-1} P_{k-1}^+ C'_{k-1}) z_k + \\ &+ H'_k u_k, r_0 = F'_0 z_0 + H'_0 u_0, P_k = C'_k C_k + Q_k \end{aligned}$$

(*) implies $r_k = C'_k z_{k+1}$, $k = 0, \dots, N-1$, $r_N = \ell$ thus $\hat{x}_N = Q_N^+ \ell$, $\hat{x}_k = P_k^+ C'_k F_{k+1} \hat{x}_{k+1}$ or $\hat{x}_k = \Phi(k, N) Q_N^+ \ell$,

$$\Phi(k, N) = P_k^+ C'_k F_{k+1} \Phi(k+1, N), \Phi(s, s) = E.$$

Combining this with (**) we obtain

$$\begin{aligned} \hat{z}_k &= (F_k \Phi(k, N) - C_{k-1} \Phi(k-1, N)) Q_N^+ \ell, \\ \hat{u}_k &= H_k \Phi(k, N) Q_N^+ \ell, \quad \hat{z}_0 = F_0 \Phi(0, N) Q_N^+ \ell. \end{aligned} \quad (25)$$

By definition, put $U(0) = Q_0$,

$$U(k) = \Phi'(k-1, k) U(k-1) \Phi(k-1, k) + H'_k H_k + F_k (E - C_{k-1} P_{k-1}^+ C'_{k-1})^2 F_k.$$

It now follows that

$$\|[\mathbf{F}' \mathbf{H}]^+ \mathbf{L}\|^2 = \sum_0^N \|\hat{z}_N\|^2 + \|\hat{u}_N\|^2 = (U(N) Q_N^+ \ell, Q_N^+ \ell).$$

It's easy to prove by induction that $Q_k = U(k)$.

Since

$$\mathbf{L} \in R[\mathbf{F}' \mathbf{H}']$$

we obtain by substituting \hat{z}_k, \hat{u}_k into (*)

$$F'_N \hat{z}_N + H'_N \hat{u}_N = \ell.$$

On the other hand (7) and (25) imply

$$F'_N \hat{z}_N + H'_N \hat{u}_N = \ell \Rightarrow [E - Q_N^+ Q_N] \ell = 0.$$

Suppose that $[E - Q_N^+ Q_N] \ell = 0$. To conclude the proof we have to show that

$$(\ell, x_N) = (Q_N^+ \ell, Q_N x_N) = 0, \quad \forall [x_0 \dots x_N] \in \mathbf{N}[\mathbf{F}' \mathbf{H}'].$$

By induction, fix $N = 0$. If $F_0 x_0 = 0$, $H_0 x_0 = 0$, then $Q_0 x_0 = 0$. We say that $[x_0 \dots x_k] \in \mathbf{N}[\mathbf{F}' \mathbf{H}']$ if

$$F_0 x_0 = 0, \quad H_0 x_0 = 0, \quad F_s x_s = C_{s-1} x_{s-1}, \quad H_s x_s = 0.$$

Suppose $Q_{k-1} x_{k-1} = 0$, $\forall [x_0 \dots x_{k-1}] \in \mathbf{N}[\mathbf{F}' \mathbf{H}']$ and fix any $[x_0 \dots x_k] \in \mathbf{N}[\mathbf{F}' \mathbf{H}']$. Then $F_k x_k = C_{k-1} x_{k-1}$, $H_k x_k = 0$. Combining this with (7) we obtain

$$Q_k x_k = F'_k (E - C_{k-1} P_{k-1}^+ C'_{k-1}) C_{k-1} x_{k-1} (*).$$

We show that $Q_k \geq 0$ in the proof of Theorem 1. One can see that

$$\begin{bmatrix} C_{k-1} \\ \frac{1}{Q_{k-1}^2} \end{bmatrix}^+ = [(C'_{k-1}C_{k-1} + Q_{k-1})^+ C'_{k-1}, (C'_{k-1}C_{k-1} + Q_{k-1})^+ \frac{1}{Q_{k-1}^2}].$$

Since

$$\begin{bmatrix} C_{k-1} \\ \frac{1}{Q_{k-1}^2} \end{bmatrix} \begin{bmatrix} C_{k-1} \\ \frac{1}{Q_{k-1}^2} \end{bmatrix}^+ \begin{bmatrix} C_{k-1} \\ \frac{1}{Q_{k-1}^2} \end{bmatrix} x_{k-1} = \begin{bmatrix} C_{k-1} \\ \frac{1}{Q_{k-1}^2} \end{bmatrix} x_{k-1}$$

we obviously get

$$C_{k-1}(C'_{k-1}C_{k-1} + Q_{k-1})^+ C'_{k-1}C_{k-1}x_{k-1} = C_{k-1}x_{k-1} \Rightarrow Q_{k-1}x_{k-1} = 0$$

as it follows from (*). This completes the proof.

Proof of Lemma 1. Taking into account the definitions of the matrices \mathbf{F} , \mathbf{H} and (6) we clearly have

$$\mathbf{G}_y^N = \{\mathbf{X} : \|\mathbf{FX}\|^2 + \|\mathbf{Y} - \mathbf{HX}\|^2 \leq 1\}.$$

Let $\hat{\mathbf{X}}$ be a minimum of the quadratic function $\mathbf{X} \mapsto \|\mathbf{FX}\|^2 + \|\mathbf{Y} - \mathbf{HX}\|^2$. It now follows that

$$\mathbf{G}_y^N = \hat{\mathbf{X}} + \mathbf{G}_0^N \Rightarrow s(\mathbf{L} | \mathbf{G}_0^N) = (\mathbf{L}, \hat{\mathbf{X}}) + s(\mathbf{L} | \mathbf{G}_0^N).$$

The application of Lemma 3 yields

$$(\mathbf{L}, \hat{\mathbf{X}}) = (\ell, Q_N^+ r_N).$$

This completes the proof.

Proof of Lemma 2. Suppose the function $f: \mathbf{R}^n \rightarrow R^1$ is convex and closed. Then [11] the support function $s(\cdot | \{x : f(x) \leq 0\})$ of the set $\{x : f(x) \leq 0\}$ is given by

$$s(z | \{x : f(x) \leq 0\}) = \text{cl} \inf_{\lambda \geq 0} \left\{ \lambda f^* \left(\frac{z}{\lambda} \right) \right\}.$$

To conclude the proof it remains to compute the support function of \mathbf{G}_0^N according to this rule and then apply Corollary 4.

Proof of Corollary 3. The proof is by induction on k . For $k=0$, there is nothing to prove. The induction hypothesis is $P_{k-1|k-1} = Q_{k-1}^{-1}$. Suppose S is $n \times n$ -matrix such that $S = S' > 0$, A is $m \times n$ -matrix; then

$$A(S^{-1} + A'A)^{-1} = (E + ASA')^{-1} AS. \quad (26)$$

Using (26) we get

$$ASA' = [E + ASA']A[A'A + S^{-1}]^{-1}A'. \quad (27)$$

Combining (27) with the induction assumption we get the following

$$E + C_{k-1}P_{k-1|k-1}C'_{k-1} = E + [E + C_{k-1}P_{k-1|k-1}C'_{k-1}] \times$$

$$\times C_{k-1}[Q_{k-1} + C'_{k-1}C_{k-1}]^{-1}C'_{k-1}.$$

By simple calculation from the previous equality follows

$$E - C_{k-1}(Q_{k-1} + C'_{k-1}C_{k-1})^{-1}C'_{k-1} = (E + C_{k-1}P_{k-1|k-1}C'_{k-1})^{-1}.$$

Using this and (7), (15) we obviously get $Q_k^{-1} = P_{k|k}$.

It follows from the definitions that $Q_0^{-1}r_0 = \hat{x}_{0|0}$. Suppose that $Q_{k-1}^{-1}r_{k-1} = \hat{x}_{k-1|k-1}$. The induction hypothesis and (26) imply

$$(E + C_{k-1}P_{k-1|k-1}C'_{k-1})^{-1}C_{k-1}\hat{x}_{k-1|k-1} = C_{k-1}(C'_{k-1}C_{k-1} + Q_{k-1})_{k-1}^{-1}r_{k-1}.$$

Combining this with (15), (16) and using $Q_k^{-1} = P_{k|k}$ we obtain

$$\hat{x}_{k|k} = Q_k^{-1}(F'_k C_{k-1}(C'_{k-1}C_{k-1} + Q_{k-1})_{k-1}^{-1}r_{k-1} + H'_k y_k).$$

This concludes the proof.

Proof of Corollary 2. If $I_k < n$ then $\text{rank}(Q) < n$ hence $\lambda_{\min}(Q_k) = 0$. In this case there is a direction $\ell \in \mathbf{R}^n$ such that $\hat{\sigma}(\ell, k) = +\infty$. So $\hat{\rho}(k) = +\infty$.

If $I_k = n$ then it follows from formula (11) that

$$\begin{aligned} \min_{\{x_k\} \in \mathbf{G}_y^N} \max_{\{\tilde{x}_k\} \in \mathbf{G}_y^N} \|x_N - \tilde{x}_N\|^2 &= \min_{\{x_k\} \in \mathbf{G}_y^N} \max_{\{\tilde{x}_k\} \in \mathbf{G}_y^N} \left\{ \max_{\|\ell\|=1} |(\ell, x_N - \tilde{x}_N)| \right\}^2 = \\ &= \left\{ \min_{\|\ell\|=1} \max_{\{x_k\} \in \mathbf{G}_y^N} |(\ell, x_N - \tilde{x}_N)| \right\}^2 \geq \\ &\geq \left\{ \max_{\|\ell\|=1} \min_{\{x_k\} \in \mathbf{G}_y^N} \max_{\{\tilde{x}_k\} \in \mathbf{G}_y^N} |(\ell, x_N - \tilde{x}_N)| \right\}^2 = \\ &= [1 - \alpha_N + (Q_N^+ r_N, r_N)] \max_{\|\ell\|=1} (Q_N^+ \ell, \ell) = \frac{[1 - \alpha_N + (Q_N^+ r_N, r_N)]}{\min_i \{\lambda_i(N)\}}. \end{aligned} \quad (28)$$

On the other hand formula (11) implies

$$\begin{aligned} \max_{\{\tilde{x}_k\} \in \mathbf{G}_y^N} \|\hat{x}_N - \tilde{x}_N\|^2 &= \left\{ \max_{\|\ell\|=1} \max_{\{\tilde{x}_k\} \in \mathbf{G}_y^N} |(\ell, x_N - \tilde{x}_N)| \right\}^2 = \\ &= \left\{ \max_{\|\ell\|=1} [1 - \alpha_N + (Q_N^+ r_N, r_N)]^2 (Q_N^+ \ell, \ell)^{\frac{1}{2}} \right\}^2. \end{aligned} \quad (29)$$

Using (28)–(29), we get (13).

Since (29) we see that the condition $I_N = n$ implies \mathbf{G}_y^N is a bounded set.

On the other hand $I_N = n$ implies $[E - Q_N^+ Q_N] = 0$ for the given N . It follows from Lemmas 1, 2 that

$$s(\ell | P_N(\mathbf{G}_y^N)) = s(P_N' \ell | \mathbf{G}_y^N) = s(\mathbf{L} | \mathbf{G}_y^N) = (\ell, Q_N^+ r_N) + \sqrt{\beta_N} (Q_N^+ \ell, \ell)^{\frac{1}{2}}, \quad (30)$$

for any $\ell \in \mathbf{R}^n$. By Young's theorem [11], (30), so that

$$\begin{aligned} P_N(\mathbf{G}_y^N) &= \{x \in \mathbf{R}^n : (x, \ell) \leq s(\ell | P_N(\mathbf{G}_y^N)), \forall \ell \in \mathbf{R}^n\} = \\ &= \{x \in \mathbf{R}^n : \sup_{\ell} \{(x, \ell) - (\ell, \hat{x}_N) - \sqrt{\beta_N} (Q_N^+ \ell, \ell)^{\frac{1}{2}}\} \leq 0\} = \\ &= \{x \in \mathbf{R}^n : (Q_N x, x) - 2(Q_N \hat{x}_N, x) + \alpha_N \leq 1\}. \end{aligned}$$

This completes the proof.

REFERENCES

1. Favini A., Vlasenko L. On solvability of degenerate nonstationary differential-difference equations in Banach spaces.— Journal of Differential and Integral Equations. — 2001. — **14**, № 7. — P. 83–896.
2. James K. Mills Dynamic modelling of a a flexible-link planar parallel platform using a substructuring approach. — Mechanism and Machine Theory. — 2006. — № 41. — P. 671–687.
3. Hasan M.A. Noncausal image modelling using descriptor approach // IEEE Transactions on Circuits and Systems II. — 1995. — **2**, № 42. — P. 36–540.
4. Luenberger D., Arbel A. Singular dynamic Leontief systems // Econometrica. — 1977. — **45**, № 4. — C. 12–24.
5. Becerra V.M., Roberts P.D., Griffiths G.W. Applying the extended Kalman filter to systems described by nonlinear differential-algebraic equations // Control Engineering Practice 9 (2001). — P. 267–281.
6. Nikoukhah R., Campbell S.L. and Delebecque F. Kalman filtering for general discrete-time linear systems // IEEE Transactions on Automatic Control. — 1999. — № 44. — P. 1829–1839.
7. Biehn N., Campbell S., Nikoukhah R., Delebecque F. Numerically constructible observers for linear time-varying descriptor systems // Automatica. — 2001. — № 37. — P. 445–452.
8. Ishihara J.Y., Terra M.H., Campos J.C.T. Optimal recursive estimation for discrete-time descriptor systems // International Journal of System Science. — 2005. — **36**, № 10. — P. 1–22.
9. Zhuk S. Minimax estimations for linear descriptor difference equations. — <http://arxiv.org/abs/math/0609709>, 2006.
10. Grygorov A., Nakonechniy A. State estimation for discrete-time systems with inner noise // Journal of applied and computational mathematics. — 1973. — № 2 — P. 20–26.
11. Rockafellar R. Convex analysis // Princeton University Press. — 1970. — 465 c.
12. Bakan G. Analytical synthesis of guaranteed estimation algorithms of dynamic process states // Journal of automation and information science. — 2003. — № 35(5). — P. 12–20.
13. Kurzhanski A., Valyi I. Ellipsoidal calculus for estimation and control birkhauser, 1997. — 190 c.
14. Kuntsevich V., Lychak M. Guaranteed estimates, adaptation and robustness in control system // Springer-Verlag, 1992. — 250 c.

Received 12.02.2008

From the Editorial Board: the article corresponds completely to submitted manuscript.

**КАЛЕНДАРНОЕ ТЕХНИЧЕСКОЕ ОБСЛУЖИВАНИЕ
ПРОСТОЙ СИСТЕМЫ С УЧЕТОМ МИНИМАЛЬНОГО
АВАРИЙНОГО ВОССТАНОВЛЕНИЯ**

А.И. ПЕСЧАНСКИЙ

Рассмотрена стратегия минимальных восстановлений с периодически полными заменами и учетом времени восстановления. Для системы с простой структурой введен альтернирующий процесс минимального восстановления. Исследовано его асимптотическое поведение с помощью тауберовых теорем. Определены оптимальные сроки проведения полных замен.

ВВЕДЕНИЕ

Одной из стратегий технического обслуживания простой системы является стратегия минимальных восстановлений с периодически полными заменами [1]. Данная стратегия предполагает полное обновление системы только в определенные моменты времени. Если же система отказывает на интервале между двумя последовательными полными восстановлениями, то производится лишь минимальное восстановление. Минимальное восстановление означает, что наработка восстановленной системы, проработавшей к моменту отказа время s , имеет следующую функцию распределения:

$$F_s(t) = \frac{F(s+t) - F(t)}{\overline{F}(s)}. \quad (1)$$

В работе [1] оптимальное планирование профилактических замен получено в предположении, что все мероприятия по восстановлению осуществляются за пренебрежительно малое время. Поэтому, при построении математической модели функционирования системы время восстановления считается равным нулю. Однако, на практике это предположение нередко не выполняется и приходится учитывать время на восстановительные мероприятия.

Цель статьи — исследование стратегии минимальных восстановлений с периодически полными заменами и учетом времени восстановления.

ПОСТАНОВКА ЗАДАЧИ

Рассмотрим систему, состоящую из одного элемента, в которой возможно проведение планового технического обслуживания (ТО) и внеплановых ава-

рийно-профилактических ремонтов. Установим следующую очередность проведения восстановительных работ. Время безотказной работы элемента в нулевой момент времени СВ α с ФР $F(t) = P(\alpha \leq t)$. После каждого отказа элемента, который обнаруживается мгновенно, начинается его минимальное аварийное восстановление (МАВ) [1]. Продолжительность МАВ–СВ β с ФР $G(t) = P(\beta \leq t)$. По определению, после МАВ элемент, проработавший время s , имеет «остаточную наработку» с ФР (1).

Таким образом, МАВ делает элемент работоспособным, но по его окончании интенсивность отказов такая же, как непосредственно перед отказом. После следующего отказа и МАВ ФР «остаточной наработки» определяется по-прежнему формулой (1), в которой s — суммарное время работы элемента с начала его эксплуатации и т.д. Кроме внеплановых аварийно-профилактических ремонтов в системе проводится предупредительное ТО. Предупредительное ТО элемента планируется через время τ после начала его работы, независимо от того в работоспособном или отказовом состоянии он находится. Продолжительность ТО–СВ β^p с ФР $G^p(t) = P(\beta^p \leq t)$. В результате ТО элемент полностью обновляет свои надежность характеристики и очередное ТО планируется через время τ после обновления элемента. Временная диаграмма функционирования элемента изображена на рис. 1.

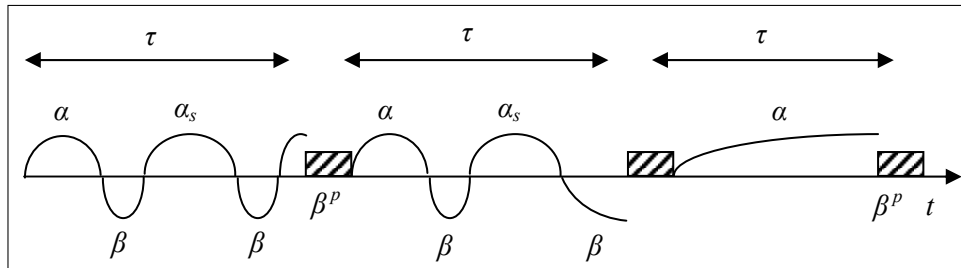


Рис. 1. Временная диаграмма функционирования системы

Предполагается, что СВ α , β и β^p независимы, имеют абсолютно непрерывные ФР и конечные математические ожидания $M\alpha$, $M\beta$, $M\beta^p$. Отключение и включение элемента происходит мгновенно. Доход за единицу времени исправного функционирования, плата за единицу времени аварийного восстановления и плата за единицу времени ТО элемента соответственно равны c^0 , c и c^p .

Необходимо определить следующие показатели качества функционирования системы: стационарный коэффициент технического использования $K_u(\tau)$, среднюю удельную прибыль $S(\tau)$, приходящуюся на единицу календарного времени и средние удельные затраты $C(\tau)$, приходящиеся на единицу времени исправного функционирования системы. Определить промежутки времени τ между окончанием предыдущего и началом последующего ТО элемента, для которых указанные показатели качества функционирования системы имели бы оптимальные значения.

ПОСТРОЕНИЕ МАТЕМАТИЧЕСКОЙ МОДЕЛИ

Элемент может находиться в одном из трех физических состояний: в работоспособном состоянии, в состоянии восстановления или пребывать в состоянии ТО. Случайный процесс, описывающий эволюцию системы во времени является регенерирующим. Точками регенерации являются моменты обновления системы после ТО. Тогда коэффициент технического использования (КТИ) K_u элемента можно определить по формуле $K_u = \frac{MX^{(1)}}{MX}$ [2],

где MX среднее время между точками регенерации, а $MX^{(1)}$ среднее время работы элемента за период между точками регенерации. В нашем случае $MX = \tau + M\beta^P$. Для нахождения $MX^{(1)}$ подготовим определенные теоретические основы, для чего введем и исследуем альтернирующий процесс минимального восстановления.

Рассмотрим следующий случайный процесс. Элемент отказывает спустя случайную наработку α_1 и осуществляется его минимальное восстановление по прошествии случайного времени β_1 . Восстановленный элемент работает время α_2 , затем наступает отказ и новое минимальное восстановление через время β_2 и т.д. (рис. 2). Моменты времени $T_1 = \alpha_1, T_2 = \alpha_1 + \beta_1 + \alpha_2, \dots$, в которые элемент отказывает, назовем моментами 0-восстановлений. Моменты времени $S_1 = \alpha_1 + \beta_1, S_2 = \alpha_1 + \beta_1 + \alpha_2 + \beta_2, \dots$, в которые заканчиваются восстановления, назовем моментами 1-восстановления.

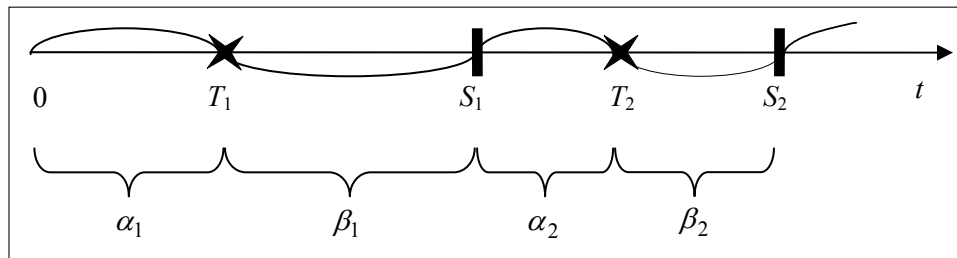


Рис. 2. Реализация альтернирующего процесса минимального восстановления

Каждая СВ из последовательности $\{\beta_n, n \geq 1\}$ имеет ФР $G(t)$. СВ α_1 из последовательности $\{\alpha_n, n \geq 1\}$ имеет ФР $F(t)$, а ФР всех остальных СВ определяются формулой (1). Последовательность $\{\alpha_n, n \geq 1\}$ порождает нестационарный пуассоновский процесс с параметром $\Lambda(s) = -\ln \bar{F}(s)$ [1].

Определение. Последовательность $\{(\alpha_n, \beta_n), n \geq 1\}$, так же как и последовательность $\{(T_n, S_n), n \geq 1\}$, назовем альтернирующим процессом минимального восстановления.

Альтернирующий процесс минимального восстановления можно эквивалентным образом описать процессом $\{Z(t), t \geq 0\}$ с помощью соотношения

$$Z(t) = \begin{cases} 0, & t \in [T_n, S_n), \\ 1, & t \notin [T_n, S_n). \end{cases} \quad (2)$$

По определению процесс $Z(t)$ задает состояние элемента в момент t : $Z(t) = 1$, если элемент работает в момент t , и $Z(t) = 0$, если элемент в момент t восстанавливается (рис. 3).

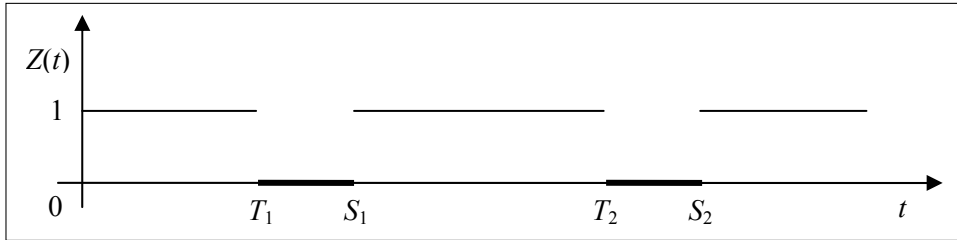


Рис. 3. Реализация процесса $\{Z(t)\}$

Отметим, что в случае экспоненциального распределения наработки альтернирующий процесс минимального восстановления есть обычный альтернирующий процесс восстановления.

Через $N_0(t)$ обозначим случайное число 0-восстановлений, а через $N_1(t)$ — случайное число 1-восстановлений на интервале $(0, t]$. Очевидно, $N_0(t)$ и $N_1(t)$ являются считающими альтернирующими процессами, для которых справедливы соотношения

$$\begin{aligned} P(N_0(t) \geq k) &= P(T_k \leq t) = \\ &= \int_0^t G^{*(k-1)}(s) \frac{[\Lambda(t-s)]^{k-1}}{(k-1)!} f(t-s) ds, \quad (G^{*(0)}(s) \equiv 1), \quad k \in N, \\ P(N_1(t) \geq k) &= P(S_k \leq t) = \int_0^t G^{*(k)}(s) \frac{[\Lambda(t-s)]^{k-1}}{(k-1)!} f(t-s) ds, \quad k \in N. \end{aligned}$$

Здесь $\Lambda(t)$ — накопленная интенсивность отказов: $\Lambda(t) = \int_0^t \lambda(s) ds$, $\lambda(t)$ — интенсивность отказов. Отсюда средние значения 0- и 1-восстановлений на интервале $[0, t]$ задаются такими функциями восстановления:

$$\begin{aligned} \tilde{H}^{(0)}(t) &= M[N_0(t)] = F(t) + \sum_{n=1}^{\infty} \int_0^t G^{*(n)}(s) \frac{[\Lambda(t-s)]^n}{n!} f(t-s) ds, \\ \tilde{H}^{(1)}(t) &= M[N_1(t)] = \sum_{n=1}^{\infty} \int_0^t G^{*(n)}(s) \frac{[\Lambda(t-s)]^{n-1}}{(n-1)!} f(t-s) ds. \end{aligned}$$

Плотности 0- и 1-восстановлений определяются формулами:

$$\tilde{h}^{(0)}(t) = \frac{d}{dt} \tilde{H}^{(0)}(t) = f(t) + \sum_{n=1}^{\infty} \int_0^t g^{*(n)}(s) \frac{[\Lambda(t-s)]^n}{n!} f(t-s) ds,$$

$$\tilde{h}^{(1)}(t) = \frac{d}{dt} \tilde{H}^{(1)}(t) = \sum_{n=1}^{\infty} \int_0^t g^{*(n)}(s) \frac{[\Lambda(t-s)]^{n-1}}{(n-1)!} f(t-s) ds .$$

Плотности восстановлений имеют следующую вероятностную интерпретацию: $\tilde{h}^{(0)}(t)dt$ ($\tilde{h}^{(1)}(t)dt$). Она представляет собой вероятность того, что 0-восстановление (1-восстановление) произойдет в интервале времени $(t, t + dt]$.

Особый интерес представляет вероятность $P(Z(t) = 1, \tilde{V}_u^{(1)} > t)$, где $\tilde{V}_u^{(1)}$ означает остаточную наработку. Очевидно $P(Z(t) = 1, \tilde{V}_u^{(1)} > t)$ означает вероятность того, что исправный к моменту u элемент не откажет на следующем интервале времени $(u, u + t]$.

С учетом соотношения (2) по формуле полной вероятности получаем

$$\begin{aligned} \bar{V}^{(1)}(u, t) &= P(Z(t) = 1, \tilde{V}_u^{(1)} > t) = \\ &= \bar{F}(u+t) + \sum_{n=1}^{\infty} \int_0^u \bar{F}(u-s+t) \frac{[\Lambda(u-s)]^n}{n!} g^{*(n)}(s) ds . \end{aligned} \quad (3)$$

Выражение $\bar{V}^{(1)}(u, t) = P(Z(t) = 1, \tilde{V}_u^{(1)} > t)$ называется нестационарным коэффициентом оперативной готовности для альтернирующего процесса минимального восстановления. Нестационарный коэффициент готовности определяется соотношением $K(u) = P(Z(u) = 1) = M(Z(u))$. Он равен вероятности того, что элемент работает в момент u . Если положить в соотношении (3) $t = 0$, то для этого частного случая

$$K(u) = \bar{F}(u) + \sum_{n=1}^{\infty} \int_0^u \bar{F}(u-s) \frac{[\Lambda(u-s)]^n}{n!} g^{*(n)}(s) ds .$$

Найдем среднее время работы элемента за период между точками регенерации $MX^{(1)}$. Суммарное время безотказной работы $U(\tau)$ элемента за время $(0, \tau]$: $U(\tau) = \int_0^{\tau} Z(t) dt$. Очевидно, что $U(\tau)$ равняется сумме наработок α_i (интерпретирующиеся здесь как рабочие периоды) до момента τ , включая, возможно, неполный рабочий период, непосредственно примыкающий к моменту τ .

$$\begin{aligned} MX^{(1)} &= M[U(\tau)] = M \left[\int_0^{\tau} Z(u) du \right] = \int_0^{\tau} M[Z(u)] du = \int_0^{\tau} K(u) du = \\ &= \int_0^{\tau} \bar{F}(u) du + \sum_{n=1}^{\infty} \int_0^{\tau} \bar{F}(u-s) \frac{[\Lambda(u-s)]^n}{n!} G^{*(n)}(s) ds . \end{aligned}$$

С учетом соотношения $\bar{F}(t) \frac{[\Lambda(t)]^n}{n!} = \int_0^t f(s) \frac{[\Lambda(s)]^{n-1}}{(n-1)!} ds - \int_0^t f(s) \frac{[\Lambda(s)]^n}{n!} ds$,

выражение в правой части последнего равенства приводится к виду:

$$MX^{(1)} = \tau - \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du .$$

Следовательно, КТИ определяется формулой

$$K_u(\tau) = \frac{\tau - \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du}{\tau + M\beta^p} . \quad (4)$$

Средний удельный доход $S(\tau)$, приходящийся на единицу календарного времени, и средние удельные затраты $C(\tau)$, приходящиеся на единицу времени исправного функционирования элемента, определяются соотношениями [2]:

$$S(\tau) = \frac{c^0 MX^{(1)} - cMX^{(0)} - c^p MX^{(2)}}{MX}, \quad C(\tau) = \frac{cMX^{(0)} + c^p MX^{(2)}}{MX^{(1)}}, \quad (5)$$

где $MX^{(0)}$, $MX^{(2)}$ среднее суммарное время АВ и среднее время ТО элемента за период регенерации соответственно. В нашем случае эти формулы принимают вид:

$$S(\tau) = \frac{c^0 \tau - (c^0 + c) \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du - c^p M\beta^p}{\tau + M\beta^p},$$

$$C(\tau) = \frac{c \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du + c^p M\beta^p}{\tau - \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du} .$$

ОПТИМИЗАЦИЯ ПОКАЗАТЕЛЕЙ КАЧЕСТВА

Определение наилучших показателей качества функционирования системы сводится к отысканию точек экстремума функций (4) и (5). Приравнявая к нулю производные этих функций, получаем уравнения:

$$(\tau + M\beta^p) (\tilde{H}^{(0)}(\tau) - \tilde{H}^{(1)}(\tau)) - \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du = M\beta^p,$$

$$(\tau + M\beta^p) (\tilde{H}^{(0)}(\tau) - \tilde{H}^{(1)}(\tau)) - \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du = \frac{c^0 + c^p}{c^0 + c} M\beta^p, \quad (6)$$

$$\tau (\tilde{H}^{(0)}(\tau) - \tilde{H}^{(1)}(\tau)) - \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du +$$

$$+ \frac{c^p}{c} M\beta^p (\tilde{H}^{(0)}(\tau) - \tilde{H}^{(1)}(\tau)) = \frac{c^p}{c} M\beta^p.$$

Для доказательства существования решений этих уравнений определим множество значений непрерывных функций в левых частях уравнений. Функции в точке $\tau = 0$ равны нулю. Для определения поведения при $\tau \rightarrow \infty$ исследуем асимптотическое поведение функций $\tilde{H}^{(0)}(t) - \tilde{H}^{(1)}(t)$ и $\Psi(t) \equiv t(\tilde{H}^{(0)}(t) - \tilde{H}^{(1)}(t)) - \int_0^t (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du$ при $t \rightarrow \infty$ с помощью

тауберовых теорем.

Теорема 1. Если существует $\lim_{t \rightarrow \infty} (\tilde{H}^{(0)}(t) - \tilde{H}^{(1)}(t))$, тогда

$$\lim_{t \rightarrow \infty} (\tilde{H}^{(0)}(t) - \tilde{H}^{(1)}(t)) = \frac{\lambda(\infty)M\beta}{1 + \lambda(\infty)M\beta}, \quad 0 \leq \lambda(\infty) \leq \infty. \quad (7)$$

Для доказательства этого утверждения воспользуемся предельным соотношением [3]

$$\lim_{t \rightarrow +\infty} \frac{\Gamma(\nu + 1)}{t^\nu} \varphi(t) = \lim_{p \rightarrow +0} p^{\nu+1} \hat{\varphi}(p), \quad \nu > -1, \quad (8)$$

где $\hat{\varphi}(p)$ — преобразование Лапласа функции $\varphi(t)$: $\hat{\varphi}(p) \equiv L[\varphi(t)](p) = \int_0^\infty \varphi(t) e^{-pt} dt$.

Имеем

$$\begin{aligned} L[\tilde{H}^{(0)}(t) - \tilde{H}^{(1)}(t)](p) &= L \left[F(t) - \sum_{n=1}^\infty \int_0^t g^{*(n)}(t-s) \bar{F}(s) \frac{[\Lambda(s)]^n}{n!} ds \right] (p) = \\ &= \frac{\hat{f}(p)}{p} - \sum_{n=1}^\infty [\hat{g}(p)]^n \int_0^\infty e^{-pt} \bar{F}(t) \frac{[\Lambda(t)]^n}{n!} dt = \frac{1}{p} - \int_0^\infty e^{-pt} e^{(1-\hat{g}(p))\Lambda(t)} dt. \end{aligned}$$

Вычислим предел в правой части соотношения (8) в случае $\nu = 0$:

$$\lim_{p \rightarrow +0} pL[\tilde{H}^{(0)}(t) - \tilde{H}^{(1)}(t)](p) = 1 - \lim_{p \rightarrow +0} \int_0^\infty e^{-s} e^{-(1-\hat{g}(p))\Lambda\left(\frac{s}{p}\right)} ds.$$

Равномерная сходимость относительно p позволяет совершить предельный переход под знаком интеграла. Учитывая равенство $\lim_{p \rightarrow 0} (1 - \hat{g}(p))\Lambda\left(\frac{s}{p}\right) = \lambda(\infty)sM\beta$, приходим к формуле (7).

Следствие. Если ТО элемента не проводится, тогда стационарный КТИ равен

$$K_u(\infty) = \frac{1}{1 + \lambda(\infty)M\beta}.$$

Действительно,

$$\lim_{\tau \rightarrow \infty} \frac{\tau - \int_0^{\tau} (\tilde{H}^{(0)}(u) - \tilde{H}^{(1)}(u)) du}{\tau + M\beta^p} = \lim_{\tau \rightarrow \infty} \left[1 - (\tilde{H}^{(0)}(\tau) - \tilde{H}^{(1)}(\tau)) \right] = \frac{1}{1 + \lambda(\infty)M\beta}.$$

Таким образом, если интенсивность отказов $\lambda(t)$ стремится к нулю, тогда КТИ системы стремится к единице и ТО системы проводить нецелесообразно. В случае неограниченного возрастания интенсивности отказов КТИ системы стремится к нулю, поэтому естественно проводить ТО системы.

Далее исследуем асимптотическое поведение функции $\Psi(t)$.

Теорема 2. Если $\lambda(t) = O(t^\varepsilon)$, $\varepsilon > 0$, тогда $\Psi(t) = O(t^{\frac{1}{\varepsilon+1}})$ при $t \rightarrow \infty$.

Доказательство. Используем предельное соотношение (8) в случае $\nu = \frac{1}{\varepsilon + 1}$. Заметим, что

$$\Psi(t) = \int_0^t u (\tilde{h}^{(0)}(u) - \tilde{h}^{(1)}(u)) du,$$

$$L[\Psi(t)](p) = \frac{1}{p} \int_0^\infty e^{-pt - (1 - \hat{g}(p))\Lambda(t)} (1 - pt + p\hat{g}'(p)\Lambda(t)) dt.$$

Имеем

$$\begin{aligned} \lim_{p \rightarrow 0} p^{\frac{1}{\varepsilon+1}+1} \hat{\psi}(p) &= \lim_{p \rightarrow 0} p^{\frac{1}{\varepsilon+1}} \int_0^\infty e^{-pt - (1 - \hat{g}(p))\Lambda(t)} (1 - pt + p\hat{g}'(p)\Lambda(t)) dt = \\ &= \left[t = \left(\frac{x}{1 - \hat{g}(p)} \right)^{\frac{1}{\varepsilon+1}} \right] = \lim_{p \rightarrow 0} p^{\frac{1}{\varepsilon+1}} \int_0^\infty \left[1 - p \left(\frac{x}{1 - \hat{g}(p)} \right)^{\frac{1}{\varepsilon+1}} + \right. \\ &\left. + p\hat{g}'(p)\Lambda \left[\left(\frac{x}{1 - \hat{g}(p)} \right)^{\frac{1}{\varepsilon+1}} \right] \right] \frac{e^{-p \left(\frac{x}{1 - \hat{g}(p)} \right)^{\frac{1}{\varepsilon+1}} - (1 - \hat{g}(p))\Lambda \left[\left(\frac{x}{1 - \hat{g}(p)} \right)^{\frac{1}{\varepsilon+1}} \right]}}{(\varepsilon + 1)(1 - \hat{g}(p))^{\frac{1}{\varepsilon+1}}} x^{\frac{1}{\varepsilon+1}-1} dx. \end{aligned}$$

Совершим предельный переход под знаком интеграла и учтем, что

$$\begin{aligned} \lim_{p \rightarrow 0} p \left(\frac{x}{1 - \hat{g}(p)} \right)^{\frac{1}{\varepsilon+1}} &= 0, \quad \lim_{p \rightarrow 0} (1 - \hat{g}(p))\Lambda \left[\left(\frac{x}{1 - \hat{g}(p)} \right)^{\frac{1}{\varepsilon+1}} \right] = \\ &= \frac{Kx}{1 + \varepsilon}, \quad \lim_{p \rightarrow 0} p\Lambda \left[\left(\frac{x}{1 - \hat{g}(p)} \right)^{\frac{1}{\varepsilon+1}} \right] = \frac{Kx}{(1 + \varepsilon)M\beta}. \end{aligned}$$

Получаем

$$\begin{aligned} \lim_{p \rightarrow 0} p^{\frac{1}{\varepsilon+1}+1} \hat{\psi}(p) &= \frac{1}{(\varepsilon+1)(M\beta)^{\frac{1}{\varepsilon+1}}} \int_0^{\infty} e^{-\frac{K}{\varepsilon+1}x} \left(1 - \frac{K}{\varepsilon+1}x\right) x^{\frac{1}{\varepsilon+1}-1} dx = \\ &= \frac{1}{(\varepsilon+1)} \left(\frac{\varepsilon+1}{KM\beta}\right)^{\frac{1}{1+\varepsilon}} \int_0^{\infty} e^{-s} (1-s) s^{\frac{1}{\varepsilon+1}-1} ds = \frac{\varepsilon}{(\varepsilon+1)^2} \left(\frac{\varepsilon+1}{KM\beta}\right)^{\frac{1}{1+\varepsilon}} \Gamma\left(\frac{1}{1+\varepsilon}\right). \end{aligned}$$

Следовательно,

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\Gamma\left(\frac{1}{1+\varepsilon} + 1\right) \psi(t)}{t^{\frac{1}{1+\varepsilon}}} &= \frac{\varepsilon}{(\varepsilon+1)^2} \left(\frac{\varepsilon+1}{KM\beta}\right)^{\frac{1}{1+\varepsilon}} \Gamma\left(\frac{1}{1+\varepsilon}\right), \\ \psi(t) &\sim \frac{\varepsilon}{\varepsilon+1} \left(\frac{\varepsilon+1}{KM\beta}\right)^{\frac{1}{1+\varepsilon}} t^{\frac{1}{1+\varepsilon}}, \quad \psi(t) = O\left(t^{\frac{1}{1+\varepsilon}}\right). \end{aligned}$$

Заметим, что условия этой теоремы удовлетворяют СВ, имеющие усеченное слева нормальное распределение, а также распределенные по закону

Вейбулла-Гнеденко с плотностью $f(t) = \frac{\lambda}{\theta} \left(\frac{t}{\theta}\right)^{\lambda-1} e^{-\left(\frac{t}{\theta}\right)^\lambda}$, $\theta > 0$, $\lambda > 1$.

Для исследования асимптотического поведения функции $\Psi(t)$ в случае ограниченного возрастания функции интенсивности $\lambda(t)$ нам понадобится следующая лемма.

Лемма. Если $\lim_{t \rightarrow +\infty} \frac{\varphi(t)}{\ln t}$ существует, то существует $\lim_{p \rightarrow +0} \frac{-p}{\ln(\gamma p)} \hat{\varphi}(p)$,

где $\gamma = e^C$, $C = -\Gamma'(1)$, $C = 0,577215\dots$ — постоянная Эйлера, и эти пределы равны.

Доказательство. Пусть существует $\lim_{t \rightarrow +\infty} \frac{\varphi(t)}{\ln t} = A$. В силу условия теоремы для любого положительного числа ε существует такое $N > 1$, что

$\left| \frac{\varphi(t)}{\ln t} - A \right| < \frac{\varepsilon}{2}$ для всех $t > N$. Тогда из равенства

$$\begin{aligned} \frac{p}{-\ln(\gamma p)} \int_0^{\infty} \varphi(t) e^{-pt} dt - A &= \frac{p}{\Gamma'(1) - \ln p} \int_0^{\infty} \varphi(t) e^{-pt} dt - \\ &- A \frac{\Gamma'(1) - \ln p}{\Gamma'(1) - \ln p} = \frac{p}{\Gamma'(1) - \ln p} \int_0^{\infty} \varphi(t) e^{-pt} dt - \\ &- \frac{A}{\Gamma'(1) - \ln p} \left(\int_0^{\infty} e^{-s} \ln s ds - \ln p \int_0^{\infty} e^{-s} ds \right) = \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{\Gamma'(1) - \ln p} \left(p \int_0^{\infty} \varphi(t) e^{-pt} dt - A \int_0^{\infty} \ln \left(\frac{s}{p} \right) e^{-s} ds \right) = \\
 &= \frac{1}{\Gamma'(1) - \ln p} \left(p \int_0^{\infty} \varphi(t) e^{-pt} dt - Ap \int_0^{\infty} \ln t e^{-pt} dt \right) = \\
 &= \frac{p}{-\ln(\gamma p)} \left(\int_0^N (\varphi(t) - A \ln t) e^{-pt} dt + \int_N^{\infty} e^{-pt} \left(\frac{\varphi(t)}{\ln t} - A \right) \ln t dt \right)
 \end{aligned}$$

имеем

$$\begin{aligned}
 \left| \frac{p}{-\ln(\gamma p)} \int_0^{\infty} \varphi(t) e^{-pt} dt - A \right| &\leq \frac{p}{|\ln(\gamma p)|} \int_0^N |\varphi(t) - A \ln t| dt + \frac{p}{|\ln(\gamma p)|} \int_N^{\infty} e^{-pt} \left| \frac{\varphi(t)}{\ln t} - A \right| \ln t dt \leq \\
 &\leq \frac{p}{|\ln(\gamma p)|} \int_0^N |\varphi(t) - A \ln t| dt + \frac{p}{|\ln(\gamma p)|} \int_N^{\infty} e^{-pt} \ln t dt \frac{\varepsilon}{2}.
 \end{aligned}$$

Выберем $\delta > 0$ настолько малым, чтобы для всех $0 < p < \delta$ выполнялись неравенства

$$\frac{p}{|\ln(\gamma p)|} \int_0^N |\varphi(t) - A \ln t| dt < \frac{\varepsilon}{2}, \quad \frac{p}{|\ln(\gamma p)|} \int_N^{\infty} e^{-pt} \ln t dt \leq 1.$$

В таком случае получаем

$$\left| \frac{p}{-\ln(\gamma p)} \int_0^{\infty} \varphi(t) e^{-pt} dt - A \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \text{ при } 0 < p < \delta.$$

Теорема 3. Пусть $\lim_{t \rightarrow \infty} \lambda(t) = \lambda$ и $t\lambda(t) - \Lambda(t) = O(\ln t)$, тогда $\Psi(t) = O(\ln t)$ при $t \rightarrow \infty$.

Для доказательства этого утверждения используем лемму. Имеем

$$\begin{aligned}
 \lim_{p \rightarrow 0} \frac{-p}{\ln(\gamma p)} \hat{\psi}(p) &= - \lim_{p \rightarrow 0} \frac{1}{\ln(\gamma p)} \int_0^{\infty} e^{-pt - (1 - \hat{g}(p))\Lambda(t)} (1 - pt + p\hat{g}'(p)\Lambda(t)) dt = \\
 &= [pt = x] = - \lim_{p \rightarrow 0} \frac{1}{p \ln(\gamma p)} \int_0^{\infty} e^{-x - (1 - \hat{g}(p))\Lambda\left(\frac{x}{p}\right)} \left(1 - x + p\hat{g}'(p)\Lambda\left(\frac{x}{p}\right) \right) dx = \\
 &= - \lim_{p \rightarrow 0} \frac{1}{1 + \ln(\gamma p)} \int_0^{\infty} e^{-x - (1 - \hat{g}(p))\Lambda\left(\frac{x}{p}\right)} \left[\hat{g}''(p)p\Lambda\left(\frac{x}{p}\right) + \right. \\
 &\quad \left. + x\lambda\left(\frac{x}{p}\right) \frac{p\hat{g}'(p) - \hat{g}(p) + 1}{p^2} \left(1 - x + p\hat{g}'(p)\Lambda\left(\frac{x}{p}\right) \right) \right] dx +
 \end{aligned}$$

$$+ \lim_{p \rightarrow 0} \frac{1}{1 + \ln(\gamma p)} \int_0^{\infty} e^{-x - (1 - \hat{g}(p))\Lambda\left(\frac{x}{p}\right)} \left[\hat{g}'(p) \left(\frac{x}{p} \lambda \left(\frac{x}{p} \right) - \Lambda\left(\frac{x}{p}\right) \right) \left(2 - x + p \hat{g}'(p) \Lambda\left(\frac{x}{p}\right) \right) \right] dx.$$

При вычислении этих пределов учтем, что

$$\lim_{p \rightarrow 0} p \Lambda\left(\frac{x}{p}\right) = \lambda x, \quad \lim_{p \rightarrow 0} (1 - \hat{g}(p)) \Lambda\left(\frac{x}{p}\right) = \lambda M \beta x,$$

$$\lim_{p \rightarrow 0} \frac{p \hat{g}''(p) - \hat{g}'(p) + 1}{p^2} = \frac{\hat{g}''(0)}{2}, \quad - \lim_{p \rightarrow 0} \frac{\frac{x}{p} \lambda \left(\frac{x}{p} \right) - \Lambda\left(\frac{x}{p}\right)}{1 + \ln(\gamma p)} = K,$$

где K — константа.

Имеем

$$\lim_{p \rightarrow 0} \frac{-p}{\ln(\gamma p)} \hat{\psi}(p) = 0 + K M \beta \int_0^{\infty} e^{-x - \lambda M \beta x} (2 - x + \lambda M \beta x) dx = \frac{K M \beta}{1 + \lambda M \beta}.$$

Следовательно, $\Psi(t) \sim \frac{K M \beta \ln t}{1 + \lambda M \beta}$ при $t \rightarrow \infty$.

Отметим, что условиям теоремы удовлетворяют СВ, распределенные по закону Эрланга.

Если выполняются условия теорем 2 и 3, тогда функции в правых частях уравнений (6) неограниченно возрастают и, следовательно, эти уравнения имеют решения. Тогда оптимальные показатели качества функционирования системы определяются формулами

$$K_u^{\max} = 1 - \left(\tilde{H}^{(0)}(\tau^u) - \tilde{H}^{(1)}(\tau^u) \right),$$

$$S^{\max} = c^0 - \left(c^0 + c \right) \left(\tilde{H}^{(0)}(\tau^s) - \tilde{H}^{(1)}(\tau^s) \right),$$

$$C^{\min} = \frac{c \left(\tilde{H}^{(0)}(\tau^c) - \tilde{H}^{(1)}(\tau^c) \right)}{1 - \left(\tilde{H}^{(0)}(\tau^c) - \tilde{H}^{(1)}(\tau^c) \right)},$$

где τ^u, τ^s, τ^c — точки абсолютных экстремумов соответственно функций $K(\tau), S(\tau), C(\tau)$.

Приведем пример применения формул (4) и (5). Пусть наработка на отказ имеет распределение Вейбулла-Гнеденко с плотностью $f(t) =$

$$= \frac{\lambda}{\theta} \left(\frac{t}{\theta} \right)^{\lambda-1} e^{-\left(\frac{t}{\theta}\right)^{\lambda}},$$

а время аварийного восстановления имеет гамма-

распределение с плотностью $g(t) = \mu \frac{(\mu t)^{\nu-1}}{\Gamma(\nu)} e^{-\mu t}.$

Исходные данные и результаты расчетов приводятся в таблицах 1 и 2.

Таблица 1. Исходные данные в примере

№	λ	θ	$M\alpha$	μ	ν	$M\beta$	$M\beta^p$	c^0	c	c^p
1	2	5	13,395	0,5	0,2	0,4	0,2	3	2	1
2	1,5	15	33,272	1,2	0,8	0,667	0,1	3	2	1
3	3	20	72,512	2	1,3	0,65	0,3	4	3	2

Таблица 2. Результаты расчетов

τ_u	T_+^u	T_-^u	K_u^{\max}	τ_s	T_+^s	T_-^s	S^{\max}	τ^c	C^{\min}	T_+^c	T_-^c
2,096	1,771	0,252	0,875	1,935	1,677	0,24	2,467	1,685	0,176	1,513	0,225
2,262	2,571	0,143	0,947	2,541	2,389	0,134	2,773	2,163	0,07	2,067	0,121
7,922	7,472	0,371	0,953	7,627	7,241	0,36	3,706	7,203	0,109	6,898	0,347

ВЫВОДЫ

Найдены стационарные надежностные и экономические характеристики однокомпонентной восстанавливаемой системы в случае минимальных аварийных восстановлений после ее отказов и полных обновлений после календарного ТО на основании введенного альтернирующего процесса минимальных восстановлений. Решены задачи по определению оптимальной периодичности проведения ТО системы с учетом найденных надежностных и экономических критериев. Установлены достаточные условия существования конечных решений этих задач.

Предполагаемое направление для дальнейших исследований — перенесение рассмотренной стратегии технического обслуживания на системы со сложной структурой и решение задач оптимизации надежностных показателей системы при ограничении на стоимость показатели и наоборот.

ЛИТЕРАТУРА

1. Байхельт Ф., Франкен П. Надежность и техническое обслуживание. Математический подход. — М.: Радио и связь, 1988. — 392 с.
2. Каиштанов В.А., Медведев А.И. Теория надежности сложных систем (теория и практика). — М.: Европейский центр по качеству, 2002. — 470 с.
3. Черкесов Г.Н. Надежность аппаратно-программных комплексов. Учебное пособие. — СПб.: Питер, 2005. — 479 с.

Поступила 19.11.2008

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ФОРМИРОВАНИЯ ОБЛИКА ЭЛЕМЕНТОВ ТЕХНИЧЕСКИХ СИСТЕМ НА ОСНОВЕ ДИСКРЕТНЫХ ДАННЫХ ОБ АНАЛОГАХ

Е.М. УГРЮМОВА, А.А. ТРОНЧУК, В.Е. АФАНАСЬЕВСКАЯ

Предложены методология и компьютерная система поддержки принятия решений для формирования облика элементов технических систем. Осуществлено построение формальных математических моделей систем с помощью обучаемых искусственных нейронных сетей. Определены управляющие переменные, соответствующие желаемым критериям качества на основе применения эволюционного метода с регуляризацией.

ВВЕДЕНИЕ

Современный этап развития сложных технических систем (СТС) характеризуется значительным усложнением самой техники и соответственно большими абсолютными затратами времени и ресурсов на их проектирование и создание. Снижение риска появления ошибок разработчиков, особенно на начальных этапах создания СТС, которые могут приводить к неоправданным затратам, возможно за счет внедрения в практику проектирования СТС современных методов и средств информационной технологии решения задач системной оптимизации.

Под техническим обликом систем обычно понимается набор оптимальных в заданном смысле характеристик (параметров) систем для решения отдельных расчетных задач или их совокупности. При структуризации задач формирования технического облика систем на основе дискретных данных об аналогах возникает ряд неопределенностей [1]:

- концептуальная неопределенность — неопределенность выбора типа исходной либо формальной математической модели (ФММ), используемой в дальнейшем для определения оптимальных параметров, ее структуры и параметров;
- методологическая неопределенность — неопределенность выбора постановки задачи синтеза оптимальных параметров, сводящейся, как правило, к многокритериальной задаче параметрической оптимизации (МЗПО), метода и алгоритма ее численного решения.

Процесс раскрытия этих неопределенностей сводится к разработке методологии решения задачи структурно-параметрической оптимизации, включающей методы аппроксимации векторной функции векторной переменной и построения решений МЗПО. Следует отметить, что задачи аппроксимации данных и синтеза оптимальных параметров СТС в общем случае относятся к числу некорректных задач [2]. Таким образом, возникает необходимость разработки устойчивых к возмущениям входных данных и погрешностям вычислений математических методов для их решения.

Существующие методы решения задачи формирования технического облика систем [3, 4] основаны, как правило, на многошаговом поиске решения на базе прямых задач; на использовании одноцелевой процедуры, согласно которой параметры и управления оптимизируются при условиях, отражающих по существу одно характерное (расчетное или номинальное) задание. Недостатком рассмотренных методов является их высокая информационная сложность. Следует также отметить, что в большинстве работ, посвященных решению поставленной задачи, отсутствует анализ устойчивости решений к возмущениям входных данных и погрешностям вычислений.

Вытекающие из современных тенденций и особенностей развития СТС, требования к качеству и срокам их создания в условиях экономии сырьевых и ограничения финансово-производственных ресурсов, обуславливают актуальность разработки эффективных математических методов и алгоритмов численного решения задач формирования технического облика перспективной системы, а также реализующих их компьютерных систем поддержки принятия решений.

В данной работе использован подход к решению задачи формирования технического облика функциональных элементов СТС путем сведения этой задачи к задаче модификации [5] на основе концепции обратных задач. Представлена методология решения задачи модификации для случая, когда известны дискретные данные об аналогах. Предлагаемая методология включает методы:

- построения формальных математических моделей (ФММ) рассматриваемых процессов на основе применения обучаемых искусственных нейронных сетей (ИНС). Для формализации представления ФММ использованы однонаправленные многослойные (ОМС) и радиально-базисные (РБС) ИНС. Обучение ИНС осуществлено на основе метода стохастической аппроксимации с регуляризацией;
- решения задачи модификации — определение управляющих переменных, соответствующих желаемым критериям качества рассматриваемой технической системы. Синтез квазирешения задачи модификации осуществляется путем регуляризации поиска экстремума сглаживающего функционала с использованием метода А.Н. Тихонова. Выбор параметра регуляризации осуществляется в соответствии с обобщенным принципом невязки. Разработан эволюционный метод решения задачи модификации, основанный на использовании генетического алгоритма (ГА).

Разработана реализующая предложенную методологию интерактивная компьютерная система «Concept_Pro[®]», ориентированная на широкий круг пользователей, работающих над проблемами совершенствования технических систем.

В качестве примера реализации предлагаемой методологии рассмотрено решение задачи о выборе технического облика радиального вентилятора с загнутыми назад лопатками рабочего колеса.

ПОСТАНОВКА ЗАДАЧИ И МЕТОДОЛОГИЯ ЕЕ РЕШЕНИЯ

Задачу формирования и обоснования (оптимизации) технического облика СТС можно свести по существу, к задаче модификации. Рассмотрим содер-

жательную постановку задачи модификации СТС. Задача модификации является частным случаем задачи реконструкции [5].

Известны следующие данные, представленные в формализованном виде: описание объекта исследования, общие характеристики и свойства СТС, условия функционирования и основные требования к ее тактико-техническим и технико-экономическим показателям; структура СТС; цели модификации; сведения об аналогах и прототипе, классе допустимых управлений (способов и реализующих их устройств), критериях качества проектных решений.

Цель работы — определить оптимальные значения параметров конструкции функциональных элементов (ФЭ) для реализации желаемых критериев качества модифицированной СТС из условия рационального компромисса заданных требований.

Общая методология решения задачи модификации на основе дискретных данных об аналогах включает методы аппроксимации векторной функции векторной переменной и построения решений МЗПО.

В данной работе решение задачи аппроксимации векторной функции векторной переменной с целью построения ФММ рассматриваемых систем осуществлено на основе теории обучаемых искусственных нейронных сетей [6, 7].

Для аппроксимации данных использованы однонаправленные многослойные и радиально-базисные ИНС. Исходными данными для аппроксимации данных с помощью ИНС являются:

- входные параметры и управляющие переменные опытных образцов (аналогов) $\{Y_{ph}^{(0)}\}$;
- выходные параметры $\{d_{pi}\}$.

Обычно все исходные данные предварительно приводятся к безразмерному виду. В нашем случае использовано прямое преобразование:

$$f^\circ = \frac{2(f - \langle f \rangle)}{(f_{\max} - f_{\min})}, \text{ где } \langle f \rangle = (f_{\max} + f_{\min})/2, f^\circ \in [-1, 1],$$

и обратное:

$$f = [(f_{\max} - f_{\min})f^\circ + (f_{\max} + f_{\min})]/2.$$

Простейшая ИНС с одним скрытым слоем ($K=1$) изображена на рис. 1. Здесь $\{Y_{ph}^{(0)}\}$ — множество входных данных, $\{Y_i^{(k)}\}$ — множество выходных данных k -го слоя; k — номер слоя, $k=1 \dots (K+1)$; K — число скрытых слоев, $p=1 \dots P$, P — число аналогов; $\{W_{ij}^{(k)}\}$ — множество весов k -ого слоя; i — элемент k -го слоя; j — элемент $(k-1)$ -го слоя. Здесь введены следующие обозначения: H_0 — количество входов сети; H_1 — количество нейронов скрытого слоя; H_2 — количество выходов сети.

Если наблюдаемые в выбранных точках $\vec{Y}_p^{(0)} \in D_Y$, $p=1 \dots P$, значения $d_{pi}(\vec{Y}_p^{(0)}) \in D_i$ являются случайными величинами, а $\vec{Y}_p^{(0)}$ заданы плотно-

стями розподілення вероятностей $\rho_i(\vec{Y}^{(0)}) = \rho(\vec{Y}^{(0)} / D_i)$, то в качестве функции выбора рационального решения задачи аппроксимации данных для каждого выхода обучаемой искусственной нейронной сети принимается математическое ожидание некоторой выпуклой функции отклонения аппроксимирующей функции $Y_i^{(K+1)}(\vec{Y}^{(0)})$ от аппроксимируемой $d_i(\vec{Y}^{(0)})$ в виде:

$$e_i = \int_{\vec{Y}^{(0)} \in D_Y} f_{fit}(\Delta_i^2(\vec{Y}^{(0)})) \rho(\vec{Y}^{(0)} / D_i) d\vec{Y}^{(0)}, \quad i = 1 \dots H_{K+1}, \quad (1)$$

где f_{fit} — выпуклая функция (Fitness function), $\Delta_i = Y_i^{(K+1)}(\vec{Y}^{(0)}) - d_i(\vec{Y}^{(0)})$. Поскольку в представленное выше выражение входит неизвестная плотность распределения вероятностей, решение — аппроксимирующую функцию вида $Y_i^{(K+1)}(\vec{Y}^{(0)})$ — можно искать методом стохастической аппроксимации. Следуя (1), в дальнейшем в качестве функции выбора при обучении ОМС используется функция вида:

$$E = \frac{1}{2} \sum_{p=1}^P \gamma^{P-p} \sum_{i=1}^{H_{K+1}} f_{fit}(\Delta_{pi}^2) + \frac{1}{2} \beta \sum_{k=1}^{K+1} \sum_{i=1}^{H_k} \left(\frac{\sum_{j=0}^{H_{k-1}} (w_{ij}^{(k)})^2}{1 + \sum_{n=0}^{H_{k-1}} (w_{in}^{(k)})^2} \right), \quad (\gamma = 0,95 \dots 0,99, \beta = 10^{-7}).$$

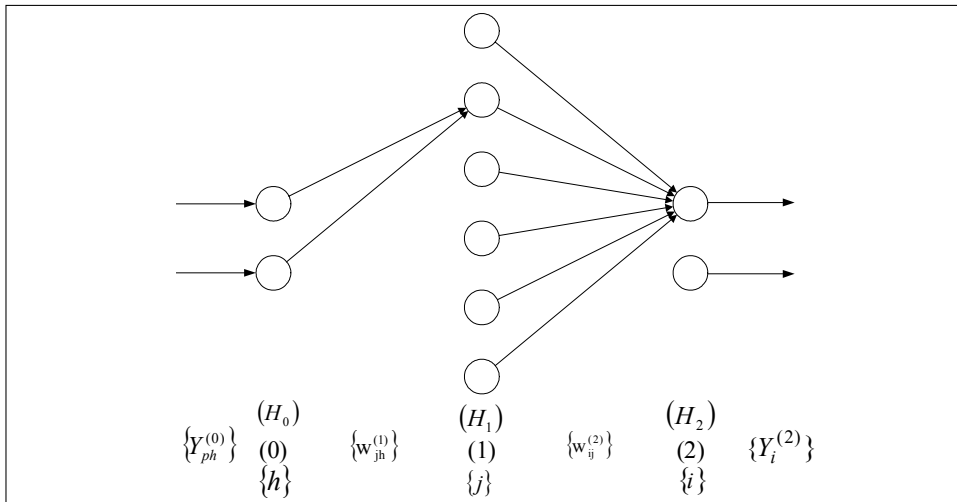


Рис. 1. Структура ИНС

Обучение ОМС осуществляется на основе алгоритма с обратным распространением ошибки. Коррекция весов связей осуществляется по следующей формуле (представлен рекуррентный алгоритм обучения, соответствующий методу стохастической аппроксимации, обеспечивающий сходимость $w_{ij}^{(k)}(t) \xrightarrow{t \rightarrow \infty} \hat{w}_{ij}^{(k)}$ с вероятностью $P = 1$):

$$w_{ij}^{(k)}(t+1) = w_{ij}^{(k)}(t) + \mu(t) \{ \eta_{ij}^{(k)}(t) r_{ij}^{(k)}(t) + \alpha_{ij}^{(k)}(t) [w_{ij}^{(k)}(t) - w_{ij}^{(k)}(t-1)] \} + \tilde{w}_{ij}^{(k)}(t+1), \quad (2)$$

где

$$\mu(t) = \frac{\mu(0)}{\left[1 + \left(1 + \frac{t}{H_0 + H_1 + 1}\right)^{-\lambda}\right]^t} \text{ — коэффициент обучения (опреде-}$$

ляемый в соответствии с условиями Дворецкого: $\mu(t) > 0$, $\sum_{t=1}^{\infty} \mu(t) = \infty$,

$\sum_{t=1}^{\infty} \mu^2(t) < \infty$; $\lambda < 1$ (здесь принимается $\mu(0) = 1$, $\lambda = 0,5$); $t = 1 \dots PT$, T —

количество эпох обучения;

$$r_{ij}^{(k)}(t) = S_{ij}^{(k)}(t) + \frac{(S_i^{(k)}(t))^T (S_i^{(k)}(t) - S_i^{(k)}(t-1))}{(S_i^{(k)}(t-1))^T S_i^{(k)}(t-1)} r_{ij}^{(k)}(t-1) \text{ — проекции}$$

вектора направления поиска, определяемые в соответствии с методом сопряженных градиентов (Полака–Рибьера), $r_{ij}^{(k)}(0) = 0$;

$$S_{ij}^{(k)}(t) = -\frac{\partial E}{\partial w_{ij}^{(k)}} = -\gamma^{P-p} f'_{fit, \Delta_i} \delta_j^{(k)} Y_j^{(k-1)} - \beta w_{ij}^{(k)} \frac{(1 + \sum_{n=0}^{H_{k-1}} w_{in}^2)}{(1 + \sum_{n=0}^{H_{k-1}} w_{in}^2)^2} \text{ —}$$

проекция градиента;

$$\tilde{w}_{ij}^{(k)}(t+1) = \nu(t) w_{ij}^{(k)}(t) \Delta w^\circ (2\xi_{ij} - 1) \text{ — «аддитивный шум» (здесь } \nu(t) = \mu(t) \sqrt{\frac{2}{\ln(2+t)}} \text{; принимается } \Delta w^\circ = 0,2 \text{; } \xi_{ij} \in [0,1] \text{ — нормально распре-}$$

деленные случайные величины, имеющие нулевое значение среднего и единичную дисперсию). «Аддитивный шум» подключается к обучению, согласно представленной выше формуле, в двух случаях:

- в конце g эпох обучения (так называемый алгоритм глобальной

«встряски» весов), если выполняется условие: $\Delta E_g^\circ < \varepsilon^g$, где $\Delta E_g^\circ = \prod_{\tau=1}^g \Delta E_\tau^\circ$,

$$\Delta E_\tau^\circ = \text{abs}(E_\tau^\circ - 1), E_\tau^\circ = \sqrt{E_\tau / E_{\tau-1}}, \tau = 1 \dots T;$$

- если $\text{abs}(\Delta w_{ij}^{(k)}(t)) < (1 - \varepsilon) \text{abs}(\Delta w_{ij}^{(k)}(t-1))$ (так называемый алгоритм локальной «встряски» весов). Принимается $\varepsilon = 0,03$.

В рассматриваемом случае используется $f_{fit}(\Delta_{pi}^2) = 1 - \exp((-L_{fit} / 4) \Delta_{pi}^2)$, $L_{fit} \geq 4$.

При обучении используется регуляризирующий алгоритм, реализующий прерывания в итерационном процессе в случаях накопления ошибок вычислений (верхний индекс T в представленных ниже формулах означает операцию транспонирования вектора в строку):

$$\text{if } \frac{(S_i^{(k)}(t))^T (S_i^{(k)}(t) - S_i^{(k)}(t-1))}{(S_i^{(k)}(t-1))^T S_i^{(k)}(t-1)} \geq r_{\max} \text{ then } r_{ij}^{(k)}(t-1) = 0;$$

$$\text{if } (S_i^{(k)}(t-1))^T S_i^{(k)}(t) \geq r_{\min} (S_i^{(k)}(t))^T S_i^{(k)}(t) \text{ then } r_{ij}^{(k)}(t-1) = 0$$

(приймається: $r_{\max} = 5, r_{\min} = 0,2$);

$$\delta_j^{(k)} = \varphi_j^{(k)} \sum_{i=1}^{H_{k+1}} \delta_i^{(k+1)} w_{ij}^{(k+1)}; \delta_i^{(K+1)} = -\Delta_i \varphi_j^{(K+1)};$$

$$\varphi_i^{(k)} = b(1 - f^2(Y_i^{(k)})); Y_i^{(k)} = f\left(w_{io}^{(k)} + \sum_{j=1}^{H_{k-1}} w_{ij}^{(k)} Y_i^{(k-1)}\right);$$

$k = 1 \dots (K+1); i = 1 \dots H_k; j = 1 \dots H_{k-1}; H_k$ — число елементів в k -ом слое;

$$f(s) = \frac{e^{bs} - e^{-bs}}{e^{bs} + e^{-bs}} \text{ — передаточна функція; } \varphi(s) \text{ — производная передаточной функции.}$$

Коефіцієнти обучения и момента находятся по следующим зависимостям:

Коефіцієнти обучения и момента находятся по следующим зависимостям:

$$\eta_{ij}^{(k)}(0) = \eta_{\max}; \tilde{\alpha}_{ij}^{(k)}(t) = \frac{S_{ij}^{(k)}(t)}{S_{ij}^{(k)}(t-1) - S_{ij}^{(k)}(t)};$$

$$\text{if } E_{\tau}^{\circ} > K_w \text{ and } S_{ij}^{(k)}(t) S_{ij}^{(k)}(t-1) < 0 \text{ then } \eta_{ij}^{(k)}(t) = \rho_d \eta_{ij}^{(k)}(t-1)$$

$$\text{else } \eta_{ij}^{(k)}(t) = \rho_i \eta_{ij}^{(k)}(t-1), \alpha_{ij}^{(k)}(t) = \tilde{\alpha}_{\text{mid}};$$

$$\text{if } \eta_{ij}^{(k)}(t) < \eta_{\min} \text{ then } \eta_{ij}^{(k)}(t) = \eta_{\min}.$$

Реализация условия: $E_{\tau+1} > E_{\tau}, E_{\tau} < E_{\tau-1}$ — представлена в форме:

$$\text{if } E_{\tau}^{\circ} \leq K_w \text{ and } S_{ij}^{(k)}(t) (w_{ij}^{(k)}(t) - w_{ij}^{(k)}(t-1)) \tilde{\alpha}_{ij}^{(k)}(t) < 0 \text{ then } \alpha_{ij}^{(k)}(t) = \tilde{\alpha}_{\max}$$

$$\text{else } \alpha_{ij}^{(k)}(t) = \tilde{\alpha}_{ij}^{(k)}(t);$$

$$\text{if } \alpha_{ij}^{(k)}(t) > \tilde{\alpha}_{\max} \text{ then } \alpha_{ij}^{(k)}(t) = \tilde{\alpha}_{\max}.$$

Приймається: $\rho_d = 0,7, \rho_i = 1,05, \eta_{\min} = 0,01, \eta_{\max} = 0,4 \dots 0,8, \tilde{\alpha}_{\text{mid}} = 0,01, \tilde{\alpha}_{\max} \leq 1,75, K_w = 1,04$.

При переходе к новой эпохе обучения порядок предъявления обучающих пар $(\bar{Y}_p^{(0)}, \bar{d}_p)$, $p = 1 \dots P$, в рекуррентном алгоритме обновлялся (использовался генератор случайных чисел на интервале $[1 \dots P]$).

Инициализация начальных значений весов осуществляется в диапазоне указанных ниже значений с помощью генератора случайных чисел: для скрытого слоя: $\pm H_0 \sqrt{H_1}$; для выходного слоя: $\pm 0,5$.

Структура РБС с одним скрытым слоем ($K = 1$) аналогична представленной структуре на рис. 1. Для формализации представления РБС введем следующие обозначения:

$\vec{Y}^{(k)} = [Y_1^{(k)}, \dots, Y_{H_k}^{(k)}]^T$, $k = 0, 1, 2$ — вектор входных данных k -го слоя;

$\vec{c}_j = [c_{j1}, c_{j2}, \dots, c_{jH_0}]^T$, $j = 1 \dots H_1$ — вектор координат центров активационной функции для нейронов скрытого слоя;

$\vec{\sigma}_j = [\sigma_{j1}, \sigma_{j2}, \dots, \sigma_{jH_0}]^T$, $j = 1 \dots H_1$ — вектор, задающий ширину окна активационной функции j -ого нейрона скрытого слоя;

$\varphi_j(\vec{Y}_p^{(0)}, \vec{c}_j, \vec{\sigma}_j) = \exp \left[-\frac{1}{2} \sum_{h=1}^{H_0} \frac{(Y_{ph}^{(0)} - c_{jh})^2}{\sigma_{jh}^2} \right] \equiv \varphi_{pj}$ — радиально-базисная

активационная функция нейрона скрытого слоя;

w_{ij} — вес связи между i -м нейроном выходного слоя и j -м нейроном скрытого слоя (имеется ввиду, что $w_{jh}^{(1)} = e_{jh} = 1$, $w_{ij}^{(2)} \equiv w_{ij}$).

В случае многомерной интерполяции с помощью РБС принимается, что количество нейронов скрытого слоя равно количеству тренировочных шаблонов. В этом случае центры активационных функций нейронов скрытого слоя размещаются в точках пространства входных сигналов сети, которые входят в набор тренировочных шаблонов сети: $\vec{c}_j = \vec{Y}_j^{(0)}$, $j = 1 \dots P$. Веса нейронов выходного слоя сети $\vec{w}_i = [w_{i1}, w_{i2}, \dots, w_{iH_1}]^T$, $i = 1 \dots H_2$, $j = 1 \dots P$ находятся путем решения соответствующей системы уравнений методом квадратного корня [8].

Для обучения РБС в случае, когда количество обучающих пар значительно превышает количество нейронов в скрытом слое $P \gg H_1$, применяется гибридный алгоритм [6]. В нем процесс обучения разделяется на два этапа:

- подбор линейных параметров сети (весов выходного слоя) на основе использования метода псевдоинверсии [6, 9];
- адаптация нелинейных параметров активационных функций (центров \vec{c}_j и ширины $\vec{\sigma}_j$ этих функций).

На втором этапе в качестве функции выбора при обучении РБС используется функция вида:

$$E = \frac{1}{2} \sum_{p=1}^P \gamma^{P-p} \sum_{i=1}^{H_2} f_{fit}(\Delta_{pi}^2).$$

Начальные значения матрицы координат центров $C = \{c_{jh}\}$ активационных функций для нейронов скрытого слоя задаются в первом приближении случайным образом (если предварительно все исходные данные приведены к безразмерному виду, то $c_{jh} \in [-1, 1]$). Для уточнения значений

$C = \{c_{jh}\}$ применим алгоритм K -средних ($K = H_1$) [6], что позволяет сократить длительность обучения РБС.

При известных значениях центров радиальных базисных функций, определяются значения элементов ковариационной матрицы $\sigma = \{\sigma_{jh}\}$. В качестве начального приближения выбирается: $\sigma_{jh} = \frac{\rho}{\sqrt{2H_1}}$, где ρ — максимальное расстояние между центрами \bar{c}_j , $j = 1 \dots H_1$, $h = 1 \dots H_0$;

$$\rho = \max_j \max_k \left(\sqrt{\sum_{h=1}^{H_0} (c_{jh} - c_{kh})^2} \right), \quad k = 1 \dots H_1 \quad (\text{если предварительно все исходные данные приведены к безразмерному виду, то } \rho = 2\sqrt{H_0}, \sigma_{jh} = \sqrt{\frac{2H_0}{H_1}}).$$

Для уточнения значений $\sigma = \{\sigma_{jh}\}$ применяется алгоритм формирования «области охвата» радиально-базисными функциями с учетом K «соседей»: $\sigma_{jh}^2 = \Sigma_j = \frac{1}{K} \sum_{k=1}^K \sum_{h=1}^{H_0} (c_{jh} - c_{kh})^2$, $k = 1 \dots K$, что позволяет сократить длительность обучения РБС [6]. Принимается $K \in [3, 5]$.

Последующее уточнение элементов ковариационной матрицы $\sigma = \{\sigma_{jk}\}$ и координат центров $C = \{c_{jh}\}$ завершает очередной цикл обучения. Коррекция элементов ковариационной матрицы осуществляется по формуле, аналогичной (2):

$$\begin{aligned} \sigma_{jh}(t+1) &= \sigma_{jh}(t) + \mu(t)\{\eta_{jh}(t)r_{jh}(t) + \\ &+ \alpha_{jh}(t)[\sigma_{jh}(t) - \sigma_{jh}(t-1)]\} + \tilde{\sigma}_{jh}(t+1), \end{aligned}$$

где проекции вектора направления поиска $r_{jh}(t)$, определяются в соответствии с методом сопряженных градиентов (Полака–Рибьера), а проекции градиента функции выбора находятся по формулам:

$$S_{jh}(t) = - \sum_{p=1}^P \gamma^{P-p} \sum_{i=1}^{H_2} \left\{ f'_{fit, \Delta_i} \Delta_{pi} w_{ij} \exp \left[- \frac{1}{2} \sum_{h=1}^{H_0} \frac{(Y_{ph}^{(0)} - c_{jh})^2}{\sigma_{jh}^2} \right] \frac{(Y_{ph}^{(0)} - c_{jh})^2}{\sigma_{jh}^3} \right\}.$$

Аналогичным образом осуществляется уточнение элементов матрицы координат центров $C = \{c_{jh}\}$. Проекция градиента функции выбора в этом случае находятся по формулам:

$$S_{jh}(t) = - \sum_{p=1}^P \gamma^{P-p} \sum_{i=1}^{H_2} \left\{ f'_{fit, \Delta_i} \Delta_{pi} w_{ij} \exp \left[- \frac{1}{2} \sum_{h=1}^{H_0} \frac{(Y_{ph}^{(0)} - c_{jh})^2}{\sigma_{jh}^2} \right] \frac{(Y_{ph}^{(0)} - c_{jh})}{\sigma_{jh}^2} \right\}.$$

В завершении осуществляется восстановление аналитического представления искомых функций $Y_{pi}^{(2)} = w_{io} + \sum_{j=1}^{H_1} w_{ij} \varphi_j(\bar{Y}_p^{(0)}, \bar{c}_j, \bar{\sigma}_j)$, $i = 1 \dots H_2$, в начале в безразмерной форме, а затем с использованием обратного преобразования — в размерной форме.

Многочисленное повторение обоих этапов ведет к полному и быстрому обучению сети, особенно когда начальные значения параметров радиально-базисных функций близки к оптимальным значениям.

На практике выделенные этапы в разной степени влияют на адаптацию параметров. Как правило, быстрее функционирует метод псевдоинверсии (он за один шаг находит локальный минимум функции). Для выравнивания этой диспропорции одно уточнение линейных параметров сопровождается обычно несколькими циклами адаптации нелинейных параметров.

В завершении обучения ИНС осуществляется восстановление аналитического представления искомых функций вначале в безразмерной форме, а затем с использованием обратного преобразования — в размерной форме.

Для проверки значимости (качества) предсказания данных ОМС вычисляются следующие — усредненные, исходя из результатов 10 независимых запусков — величины:

- энергия среднеквадратичной ошибки: $E_{av} = \frac{1}{2P} \sum_{p=1}^P \sum_{i=1}^{H_{K+1}} \Delta_{pi}^2(M)$;
- средние относительные погрешности: $\delta_i^\circ = \frac{1}{P} \sum_{p=1}^P \left| \frac{\Delta_{pi}}{d_{pi}} \right|$.

Рассмотрим особенности поиска подмножества рациональных альтернатив \hat{V} и лучшей альтернативы $\hat{v}^\circ \in \hat{V}$ как процесса построения квазирешений обратных задач.

Прямая задача анализа является корректно поставленной, если задан оператор отображения $\Phi : (D_\Pi, D_U) \rightarrow D_\Phi$, при этом отображение Φ определено и является непрерывным на всем $Q_{inp} = (D_\Pi, D_U)$, $Q_{inp} \subset Q$, $\|\Phi\| < M$, $M > 0$ [2]. Здесь область $D_\Pi \subset \Pi$ является областью имеющих физический смысл режимов, область $D_U \subset U$ — областью допустимых управлений, $Q = (D_\Pi, D_U, D_\Phi)$ — подмножеством корректности, а D_Φ — областью достижимых значений фазовых переменных (рис. 2).

Построение квазирешений обратной задачи в общем случае можно рассматривать как процедуру, реализующую отображение $F_\beta : (D_\Pi, \Phi^*) \rightarrow \hat{Q}_{out}$, $\hat{Q}_{out} = (D_U, \hat{\Phi})$, $\hat{Q}_{out} \subset \hat{Q}$, $\hat{Q} = (D_\Pi, D_U, \hat{\Phi})$ — подмножество рациональных альтернатив; $F_\beta(\Pi^\circ, \Phi^\circ, \beta)$ — регуляризирующий оператор, аппроксимирующий значения Φ^{-1} на элементах D_Φ ; β — параметр регуляризации. Обратная задача считается корректно поставленной, если регуляризирующий оператор, аппроксимирующий значения Φ^{-1} на элементах

D_Φ , $\hat{Q}_{inp} = (D_\Pi, \Phi^*)$, $\hat{Q}_{inp} \subset \hat{Q}$, $\|F_\beta\| < M$, а также метрика в пространстве фазовых переменных такая, что $\forall (\hat{\Phi}_i^\circ, \hat{\Phi}^\circ) \in \Phi: \rho(\hat{\Phi}_i^\circ, \hat{\Phi}^\circ) = \|\hat{\Phi}_i^\circ - \hat{\Phi}^\circ\|_{L_2} < M$, $M > 0$. В случае предлагаемой условно-корректной постановки обратной задачи — $\hat{\Phi} \subseteq D_\Phi$ (рис. 2).

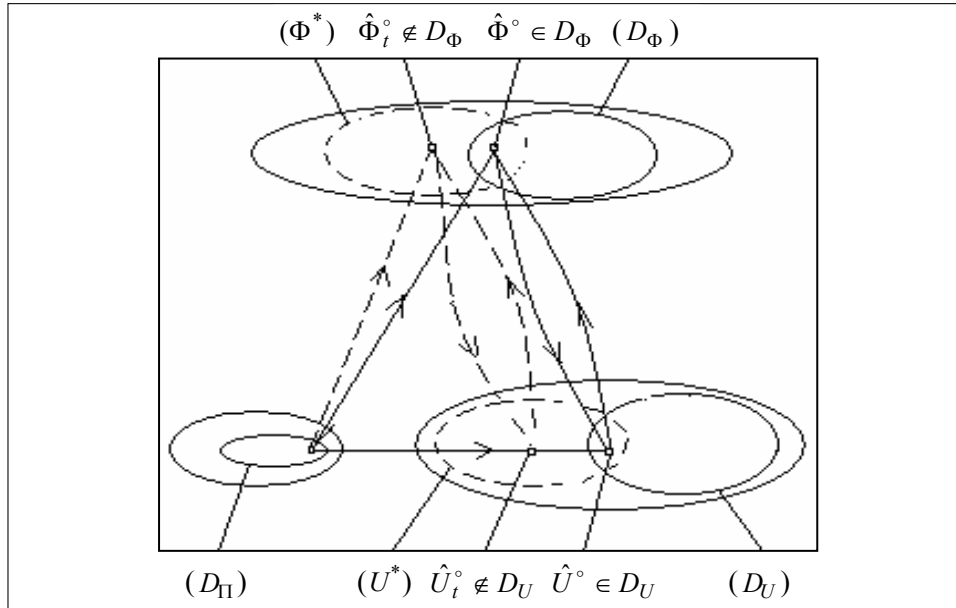


Рис. 2. Расчетная схема построения квази-решений обратной задачи

В данной работе синтез квази-решений задач модификации предлагается осуществить на основе регуляризации поиска минимума сглаживающих функционалов с использованием метода А.Н. Тихонова — для обеспечения единственности и устойчивости решений относительно малых вариаций входных данных.

Пусть $Q = \{q_p^\circ\}$, $(q_p^\circ = (\Pi_p^\circ, U_p^\circ, \Phi_p^\circ))$, $p = 0 \dots P$ — конечное множество допустимых проектных решений (подмножество корректности). С точки зрения лица, принимающего решения (ЛПР), качество любого решения $q_p^\circ \in Q$ определяется относительно критериев $W = \{w_n\}$, $n = 1 \dots N$, $N \leq H_{K+1}$. Пусть для каждого решения существует отображение $A: \Phi_p^\circ \rightarrow W_p$, тогда значение $A_n(\Phi_p^\circ) = w_{np}$ — оценка решения $q_p^\circ \in Q$ по n -му критерию $w_{np} \in W$.

Пусть для критериев $W = \{w_n\}$ существует отображение $B: W \rightarrow W^\circ$, $W^\circ = \{w_n^\circ\}$, тогда значение $B_n(w_{np}) = w_{np}^\circ$ — оценка критерия $w_{np} \in W$ по нормированному критерию $w_{np}^\circ \in W^\circ$.

Выберем критерии качества решения задачи модификации в виде:

$$\Delta w_n^\circ = w_n^\circ(U^\circ) - w_n^*, \quad w_n^\circ \in [-1, 1],$$

где w_n^* — желаемые значения рассматриваемых критериев качества, приведенные к безразмерному виду.

Аналогичным образом введем приведение к безразмерному виду всех управляющих переменных:

$$\Delta u_m^\circ = \frac{2(u_m - u_{m,0})}{u_m'' - u_m'}, \quad u_m^\circ \in [-1, 1], \quad m = 1 \dots M, \quad M \leq H_0,$$

где $u_{m,0}$ — значения переменных, соответствующие прототипу, M — число управляющих переменных в рассматриваемой задаче.

Введем отображение $G : \{w_n^\circ\} \rightarrow \hat{w}^\circ$, определяющее обобщенную функцию выбора:

$$\hat{w}^\circ(U^\circ, \beta) = \sum_{n=1}^N f_{fit}[(\Delta w_n^\circ)^2] + \beta \sum_{m=1}^M (\Delta u_m^\circ)^2, \quad (3)$$

где G — система предпочтений лица, принимающего решения, β — параметр регуляризации.

Задача (3) относится к классу существенно некорректных задач [2]. Квазирешение поставленной задачи (нормальное решение) может быть найдено методом регуляризации А.Н. Тихонова:

$$\hat{U}_r^\circ = \arg \min_{U^\circ \in D_U} \hat{w}^\circ(U^\circ, \beta_r),$$

$$D_U = \left\{ U^\circ = (u_1, \dots, u_m, \dots, u_M) : (\forall m \in [1 \dots M]) u_m' \leq u_m \leq u_m'' \right\}. \quad (4)$$

Параметр β_r выбирается ($\beta_{r+1} = \beta_r / 10$, $r = 0, 1, 2, \dots$) в соответствии с обобщенным принципом навязки [2]:

$$\left| \Delta w_n^\circ \right| \leq \xi_n + h \left\| \Delta u_m^\circ \right\|,$$

где ξ_n — погрешность определения Δw_n^* , h — погрешность определения \hat{w}° .

Поиск решения задачи (4) осуществлялся с помощью разработанного авторами эволюционного метода.

В качестве контрольных переменных и настроек генетического алгоритма (ГА) при расчетах использовались следующие параметры: размер популяции — 20, схема отбора родительских особей — рулетка в совокупности с элитарным отбором; тип кроссовера — стандартный одноточечный, вероятность кроссовера — 0,9; тип мутации — гауссовский, вероятность мутации — 0,3; максимальное количество итераций — 100.

Для предохранения популяции от доминирования неоптимальной хромосомы и тем самым для предотвращения преждевременной сходимости ГА используется масштабирование функции приспособленности:

$$\text{Fitness} = 1 - \exp(-C \hat{w}^\circ), \quad C > 1.$$

Для повышения точности нахождения экстремума разработана процедура адаптации. Суть этой процедуры заключается в следующем: последовательно производится запуск ГА с уменьшающейся областью определения управляющих переменных и увеличивающейся точностью до тех пор, пока не будет выполнено условие остановки. Процедура адаптации позволяет снизить информационную сложность алгоритма.

ПРИМЕР РЕШЕНИЯ ЗАДАЧИ О ВЫБОРЕ ТЕХНИЧЕСКОГО ОБЛИКА РАДИАЛЬНОГО ВЕНТИЛЯТОРА

В качестве примера реализации предлагаемой методологии рассмотрено решение задачи о выборе технического облика радиального вентилятора с загнутыми назад лопатками рабочего колеса на основе дискретных данных об аналогах. Расчеты выполнены с использованием реализующей предложенную методологию интерактивной компьютерной системы поддержки принятия решений «Concept_Pro®» [10], разработанной авторами.

Первоначально решается задача построения аэродинамических характеристик радиальных вентиляторов на основе применения обучаемых искусственных нейронных сетей (ИНС). Для формализации представления характеристик используются ОМС и РБС. Выборка взята из $P = 16$ значений соответствующих параметров аэродинамических характеристик радиальных вентиляторов с загнутыми назад лопатками рабочего колеса [11].

В качестве входных данных для ИНС задаются значения: густоты решетки (τ), относительной вогнутости профиля (\bar{f}), безразмерной хорды профиля (\bar{l}), безразмерного диаметра входа решетки (\bar{D}_1).

В качестве выходных данных заданы: безразмерные параметры расхода ($y_1 \equiv \bar{c}_{2r}^* = \frac{\varphi^*}{4\bar{b}_2}$) и полного давления ($y_2 = 0,5\psi^* = \bar{c}_{2u}^* \eta_g$) для режима максимального КПД, где \bar{c}_{2r}^* — коэффициент расхода, \bar{c}_{2u}^* — относительная окружная составляющая абсолютной скорости на выходе, φ^* — коэффициент производительности, ψ^* — коэффициент полного давления, \bar{b}_2 — относительная ширина колеса на выходе, η_g — гидравлический КПД рабочего колеса.

Для решения задачи аппроксимации данных используются ОМС и РБС с разнообразной структурой. Эти же сети были построены и обучались с помощью математического пакета MATLAB 7.0.1.

Введем следующие обозначения для выделения метода обучения ОМС и РБС: СГ1 — обучение методом сопряженных градиентов ОМС с одним скрытым слоем, СГ2 — обучение методом сопряженных градиентов ОМС с двумя скрытыми слоями, ЛМ1 — обучение методом Левенберга-Маквардта ОМС с одним скрытым слоем, ЛМ1Р — обучение ОМС с одним скрытым слоем методом ЛМ1 на основе байесовской регуляризации, ЛМ2 — обучение методом Левенберга-Маквардта ОМС с двумя скрытыми слоями, СГ1

(ХАИ) — обучение предложенным методом ОМС с одним скрытым слоем, РБС (ХАИ) — обучение предложенным методом РБС, МНК — обучение методом наименьших квадратов мультипликативной модели.

Результаты обучения ИНС для этой выборки представлены в таблице. В качестве примера на рис. 3 приведен график изменения E_{av} в процессе обучения ОМС методом СГ1 (ХАИ) за $T = 300$ эпох. Из анализа этих данных следует, что процесс обучения ОМС и РБС, предложенным методам быстро сходится.

Таблица. Результаты обучения ИНС ($P = 16, H_0 = 4, T = 300$)

Метод	δ_ϕ^0	δ_ψ^0	E_{av}	Параметры ИНС
МНК	0,3509	0,0830	0,0205	—
СГ1	0,0020	0,0015	3,8e-5	4-10-2
СГ2	0,0023	0,0016	8,1e-5	4-10-5-2
ЛМ1	4,8e-7	2,8e-7	3,4e-12	4-10-2
ЛМ1Р	0,0241	0,0125	0,0743	4-10-2
ЛМ2	1,4e-6	2,6e-6	1,9e-10	4-10-5-2
СГ1(ХАИ)	1,7e-3	0,0164	6,2e-3	4-10-2
РБС	1,0e-16	4,3e-17	2,1e-32	4-15-2
РБС(ХАИ)	0,0138	3,0e-3	6,2e-6	4-15-2

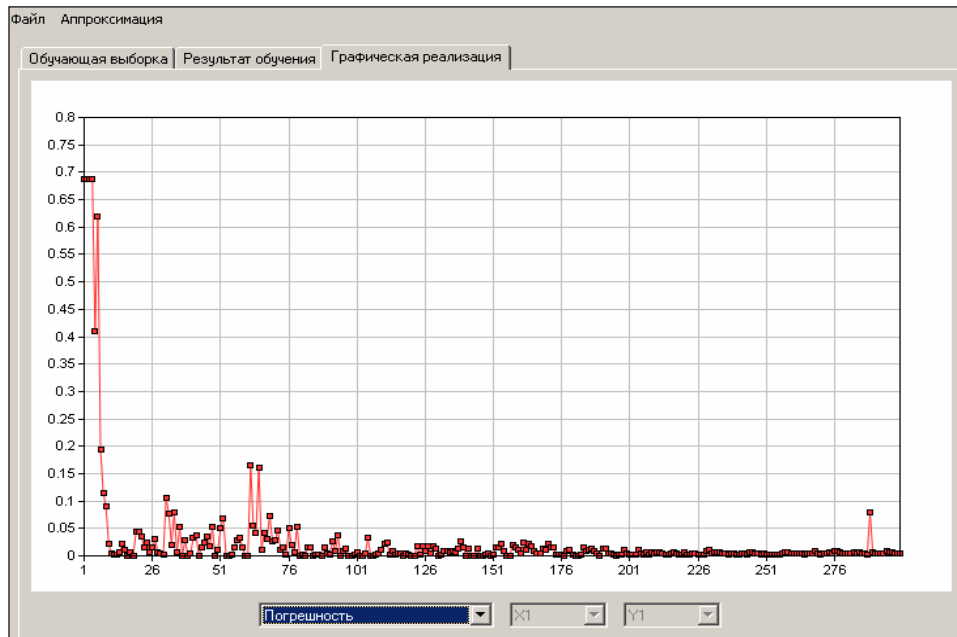


Рис. 3. Изменение E_{av} в процессе обучения ОМС ($P = 16, H_0 = 4, T = 300$)

Общий анализ данных показывает, что ОМС с двумя скрытыми слоями обучается лучше, чем с одним скрытым слоем. ЛМ1 и ЛМ2 имеют эффективную реализацию в системе MATLAB, однако эффективность этих методов ухудшается в случае подключения байесовской регуляризации (см. ЛМ1Р в сравнении с ЛМ1). Таким образом, разработанные методы СГ1

(ХАИ) и РБС (ХАИ) обучения ИНС на основе предложенного метода стохастической аппроксимации с регуляризацией обеспечивают построение робастных ФММ процессов, а также достаточную с практической точки зрения точность аппроксимации данных.

Далее решается задача модификации — определение управляющих переменных, соответствующих желаемым критериям качества рассматриваемой технической системы. В качестве прототипа выбран радиальный вентилятор с параметрами: густота решетки — $\tau = 2,50$, относительная вогнутость профиля — $\bar{f} = 0,095$, безразмерная хорда профиля — $\bar{l} = 0,37$, безразмерный диаметр входа решетки — $\bar{D}_1 = 0,508$. В качестве желаемых критериев выбраны: параметр расхода — $y_1 = 0,2$, параметр полного давления — $y_2 = 0,5$.

В результате решения задачи модификации получены следующие значения: густоты решетки — $\tau = 2,73$, относительной вогнутости профиля — $\bar{f} = 0,131$, безразмерной хорды профиля — $\bar{l} = 0,45$, безразмерного диаметра входа решетки — $\bar{D}_1 = 0,454$ и, соответствующих им, достигнутые значения выходных данных — безразмерных параметров: расхода — $y_1 = 0,205$ и полного давления — $y_2 = 0,528$.

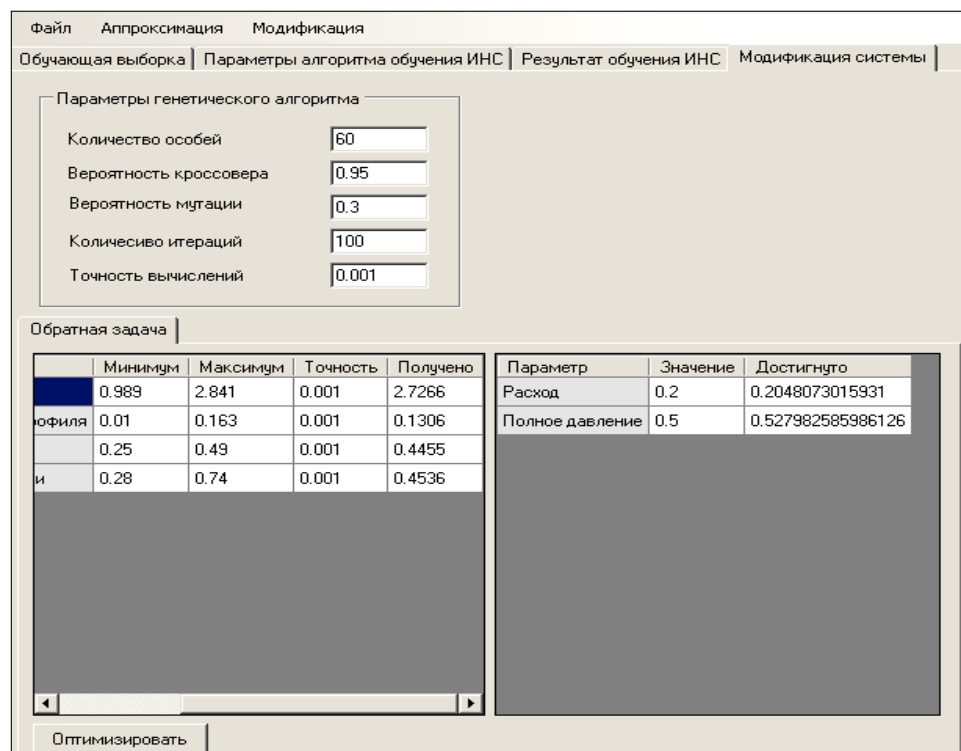


Рис. 4. Пример решения задачи модификации

Экранная форма результатов решения, найденного с использованием компьютерной интерактивной системы поддержки принятия решений при формировании облика элементов технических систем «Concept_Pro®», представлена на рис. 4.

Таким образом, получены результаты решения задачи формирования облика радиального вентилятора с загнутыми назад лопатками рабочего колеса.

ВЫВОДЫ

Разработана методология формирования технического облика функциональных элементов СТС для случая, когда известны дискретные данные об аналогах, путем сведения этой задачи к задаче модификации на основе концепции обратных задач. Разработаны методы:

- построения формальных математических моделей рассматриваемых процессов на основе применения обучаемых искусственных нейронных сетей. Для формализации представления ФММ использованы однонаправленные многослойные и радиально-базисные ИНС. Обучение ИНС осуществлено на основе метода стохастической аппроксимации с регуляризацией;
- решения задачи модификации — определения управляющих переменных, соответствующих желаемым критериям качества рассматриваемой технической системы. Синтез квазирешения задачи модификации осуществлен путем регуляризации поиска экстремума сглаживающего функционала с использованием метода А.Н. Тихонова. Выбор параметра регуляризации осуществлен в соответствии с обобщенным принципом навязки. Разработан эволюционный метод решения задачи модификации, основанный на использовании генетического алгоритма.

Предложенные методы обеспечивают устойчивость решений к возмущениям входных данных и погрешностям вычислений.

Разработана реализующая предложенную методологию интерактивная компьютерная система «Concept_Pro[®]», ориентированная на широкий круг пользователей, работающих над проблемами совершенствования технических систем.

ЛИТЕРАТУРА

1. Згуровский М.З., Панкратова Н.Д. Системный анализ: проблемы, методология, приложения. — Киев: Наук. думка, 2005. — 743 с.
2. Тихонов А.Н., Гончарский А.В. Численные методы решения некорректных задач / В.В. Степанов, А.Г. Ягола. — М.: Наука, 1990. — 232 с.
3. Стоянов Ф.А. Оптимальное автоматизированное проектирование проточных частей осевых турбин. — Киев: Наук. думка, 1989. — 176 с.
4. Бойко А.В., Говорущенко Ю.Н. Аэродинамический расчет и оптимальное проектирование проточной части турбомашин / С.В. Ершов, А.В. Русанов, С.Д. Северин. — Х.: НТУ «ХПИ», 2002. — 356 с.
5. Угрюмова Е.М., Волков С.Г., Угрюмов М.Л. Совершенствование сложных технических систем методом обратных задач // *Авиационно-космическая техника и технология*. — 2006. — № 1 (27). — С. 91–95.
6. Осовский С. Нейронные сети для обработки информации. — М.: Финансы и статистика, 2002. — 344 с.

7. Рутковская Д., Пилинский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. — М.: Горячая линия – Телеком, 2004. — 452 с.
8. Самарский А.А., Гулин А.В. Численные методы. — М.: Наука, 1989. — 432 с.
9. Голуб Дж., Ван Лоун Ч. Матричные вычисления. — М.: Мир, 1999. — 534 с.
10. Комп'ютерна програма «Комп'ютерна інтерактивна система підтримки прийняття рішень при формуванні вигляду елементів складних технічних систем «Сонсерт_Pro»»: Свідоцтво про реєстрацію авторського права на твір № 28975 / Е.М. Угрюмова, О.А. Трончук, В.Є. Афанасьєвська (Україна). — Дата реєстрації 29.05.2009.
11. Соломахова Т.С., Беляновский Е.С. Применение аэродинамических характеристик круговых решеток тонких профилей при проектировании радиальных вентиляторов // Промышленная аэродинамика. — Вып. 1/33. — М.: Машиностроение, 1986.— С. 63–70.

Поступила 27.05.2009

УДК 303.725.34:510.22;62-50

**НЕЧІТКА АРИФМЕТИКА В ЗАДАЧАХ
МІЖГАЛУЗЕВОГО БАЛАНСУ**

М.С. СЯВАВКО, Т.В. ПАСІЧНИК, В.П. ТИМКІВ

Запропоновано використання теорій нечітких мір для задач міжгалузевого балансу. Показано як методи нечіткої логіки надають можливість кількісної інтерпретації якісних факторів. Пропонується використати дробово-раціональний алгоритм для знаходження нормального розв'язку некоректно поставлених задач або погано обумовлених систем. Наведено приклади розв'язування нечітких задач.

Для задач управління складними системами нечіткі рівняння набувають значної ваги. Вони виникають при прийнятті рішень, медичній діагностиці, економіці та в інших задачах, де параметри визначено нечітко, або їх потрібно розмивати, а інколи їх визначено суб'єктивно.

Для розв'язання нечітких рівнянь, необхідно, перш за все, провести аналіз арифметичних операцій над нечіткими числами (НЧ). Цей аналіз вперше було розглянуто в роботі [1]. Було показано, що:

- нечітке число не має протилежного і оберненого чисел;
- додавання і множення комутативні, асоціативні, але в загальному випадку недистрибутивні.

Тому розв'язання таких рівнянь можливе завдяки введенню додаткових операцій віднімання та ділення нечітких чисел, через апроксимацію нечітких чисел за системою рівневих множин, або через використання L - R нечітких чисел [2, 3].

У роботі використано останній підхід. Для випадку L - R нечітких чисел рівняння з НЧ можна розв'язати, одержавши відповідну скобкову форму. Слід також підкреслити, що α -рівневий розклад опуклих нечітких підмножин дозволяє здійснити подальший аналіз задач із НЧ за допомогою методів інтервального аналізу.

**ЗАГАЛЬНИЙ ПІДХІД РОЗВ'ЯЗУВАННЯ НЕЧІТКИХ ОПЕРАТОРНИХ
РІВНЯНЬ ІЗ L - R НЕЧІТКИМИ ЧИСЛАМИ**

Нехай E — розширення числової осі, L — множина неспадних, неперервних справа функцій $L: E \rightarrow [0, 1]$ із $L(-\infty) = 0$, $L(+\infty) = 1$; R — множина

незростаючих, неперервних зліва функцій $R: E \rightarrow [0,1]$ із $R(-\infty)=1$, $R(+\infty)=0$. Надалі множини таких функцій позначимо через \mathbf{L} і \mathbf{R} .

Один із методів розв'язання нечітких рівнянь тісно пов'язаний із квазі-оберненими функціями L^q , M^q , R^q і S^q .

Означення 1. Функція $L^q: [0,1] \rightarrow E: L^q(\alpha) = \inf \{x \in E \mid L(x) \geq \alpha\}$ називається квазіоберненою до функції $L \in \mathbf{L}$.

Функція $M^q: E \rightarrow [0,1]: M^q(x) = \sup \{\alpha \in [0,1] \mid M(\alpha) > x\}$ називається квазіоберненою до функції $M \in \mathbf{M}$.

Означення 2. Функція $R^q: [0,1] \rightarrow E: R^q(\alpha) = \sup \{x \in E \mid R(x) \geq \alpha\}$ називається квазіоберненою до функції $R \in \mathbf{R}$.

Функція $S^q: E \rightarrow [0,1]: S^q(x) = \inf \{\alpha \in [0,1] \mid S(\alpha) < x\}$ називається квазіоберненою до функції $S \in \mathbf{S}$.

Тут \mathbf{M} — множина неспадних, неперервних зліва функцій $M: [0,1] \rightarrow E$ із $M: [0,1] = -\infty$, а \mathbf{S} — множина не зростаючих, неперервних зліва функцій $S: [0,1] \rightarrow E$ із $S(0) = +\infty$.

Зауважимо, що оператор переходу до квазіоберненої функції суттєво залежить від того, до якого класу функцій належить задана функція. Наприклад, можна переконатись у тому, що у випадку суворо монотонних неперервних функцій квазіобернена функція співпадає з оберненою. До того ж у загальному випадку $(L^q)^q = L$ і $(R^q)^q = R$.

Надалі нам необхідні наступні поняття.

L - R — нечітке число, яким називатимемо нечітку множину A в E , функція належності якої має вигляд

$$\mu_A(x) = \min(L_A(x), R_A(x)), \quad (1)$$

де $L_A \in \mathbf{L}$, $R_A \in \mathbf{R}$.

L - R нечіткі числа, що володіють наступними властивостями:

- функція належності такого числа напівнеперервна зверху. Крім того, існує $x^* \in E$ для якого

$$\mu_A(x^*) = \sup_{x \in E} \mu_A(x) = h_A;$$

де A — L - R нечітке число, а h_A — його висота;

- всі множини рівняння $y \in (0, h_A)$ L - R нечіткого числа A мають вигляд

$$A_y = \{x \in E \mid L_A^q(y) \leq x \leq R_A^q(y)\},$$

де $L_A^q \in \mathbf{M}$, $R_A^q \in \mathbf{S}$ — квазіобернені функції.

Нехай задано звичайну функцію $\varphi: E^n \rightarrow E$.

Значенням нечіткої функції $\Phi(x_1, \dots, x_n)$ від нечітких чисел x_1, \dots, x_n називають нечітку множину в E з функцією належності

$$\mu_{\Phi}(t) = \begin{cases} \sup_{\varphi(x_1, \dots, x_n)=t} \min(\mu_{x_1}(x_1), \dots, \mu_{x_n}(x_n)), & \text{якщо } \varphi^{-1}(t) \neq \emptyset, \\ 0, & \text{якщо } \varphi^{-1}(t) = \emptyset, \end{cases}$$

де \emptyset — пуста множина.

Цей принцип розширення встановлює формальний апарат для перенесення операцій (арифметичних, алгебраїчних) із звичайних множин у нечіткі.

Згідно [4] нечітким рівнянням називатимемо співвідношення

$$\Phi(X, A_1, \dots, A_n) \subseteq B, \quad (2)$$

де $\Phi(X, A_1, \dots, A_n)$ — значення нечіткої функції від нечітких чисел X, A_1, \dots, A_n , що одержана на засадах принципу розширення.

В (2) X — невідома величина, а $\Phi: E^{n+1} \rightarrow E$ — одержана з неперервної та монотонної за всіма змінними функції $\varphi(x, a_1, \dots, a_n)$ за принципом розширення.

Нечітке число X вважаємо розв'язком (2), якщо

$$\mu_{\Phi}(t) = \sup_{\Phi(x, a_1, \dots, a_n)=t} \min(\mu_x, \mu_{A_1}(a_1), \dots, \mu_{A_n}(a_n)) \leq \mu_B(t), \quad (3)$$

для всіх тих t , для яких $\varphi^{-1}(t) \neq \emptyset$.

У (3) під символом μ розуміємо функцію належності. Не конкретизуючи певних деталей у роботі [4] встановлено, що розв'язок рівняння (2) еквівалентний розв'язанню, відповідно в класі \mathbf{M} і \mathbf{S} , двох (незалежних один від одного) звичайних (чітких) рівнянь

$$\begin{aligned} \varphi(L_X^q(y), L_{A_1}^q(y), \dots, L_{A_n}^q(y)) &= L_B^q(y), \\ \varphi(R_X^q(y), R_{A_1}^q(y), \dots, R_{A_n}^q(y)) &= R_B^q(y). \end{aligned} \quad (4)$$

Розв'язавши (4) із значень $L_X^q(y)$ і $R_X^q(y)$, встановимо прямі значення $L_X(x), R_X(x)$. Тоді розв'язок нечіткого рівняння (2) має вигляд L - R нечіткого числа (1).

Приклад 1 [4]. Розв'язати нечітке рівняння $X + A \subseteq B$ за умов

$$\begin{aligned} L_A^q(y) &= \frac{5}{8}y, \quad R_A^q(y) = \frac{5}{8}, \\ L_B^q(y) &= \begin{cases} \frac{13}{8}y + \frac{3}{8}, & \text{якщо } 0 \leq y < \frac{1}{2}, \\ \frac{1}{8}y + \frac{9}{8}, & \text{якщо } \frac{1}{2} \leq y < \frac{3}{4}, \\ \frac{13}{8}y, & \text{якщо } \frac{3}{4} \leq y \leq 1, \end{cases} \quad R_B^q(y) = \frac{13}{8}. \end{aligned}$$

Функції $L_A^q(y), R_A^q(y), L_B^q(y), R_B^q(y)$ є ядром функцій, які визначають систему (4).

Розв'язання. Для даного випадку система (4) виглядає так:

$$\begin{cases} L_X^q + L_A^q = L_B^q, \\ R_X^q + R_A^q = R_B^q. \end{cases}$$

Звідси маємо

$$L_X^q = \begin{cases} y + \frac{3}{8}, & \text{якщо } 0 \leq y < \frac{1}{2}, \\ -\frac{1}{2}y + \frac{9}{8}, & \text{якщо } \frac{1}{2} \leq y \leq \frac{3}{4}, \\ y, & \text{якщо } \frac{3}{4} \leq y \leq 1, \end{cases} \quad R_X^q = 1.$$

Тому

$$\mu_X(x) = \min(L_X(x), R_X(x))$$

із

$$L_X(x) = \begin{cases} 0, & \text{якщо } x \leq 3/8, \\ x - 3/8, & \text{якщо } 3/8 \leq x < 7/8, \\ x, & \text{якщо } 7/8 \leq x < 1, \\ 1, & \text{якщо } x \geq 1, \end{cases} \quad R_X(x) = \begin{cases} 0, & \text{якщо } x \leq 1, \\ 1, & \text{якщо } x > 1. \end{cases}$$

Важливість дослідження рівнянь вигляду (2) підкреслює наступний приклад.

Приклад 2 [5]. Для визначення кількості води, необхідної для зрошення деякої сільськогосподарської культури (зрошувальної норми культури), агроном розв'язує рівняння

$$x + a + b = c,$$

де x — шукана зрошувальна норма, a — кількість опадів вегетації культури, b — використані запаси води з кореневого шару ґрунту, а c — сумарна потреба води на один га заданої культури.

Агроному заздалегідь невідомі конкретні значення параметрів a і b . Але він може вказати інтервали A і B , у яких містяться значення цих параметрів. Агроном має інформацію про цільові значення сумарної потреби води, за якої відбувається нормальний розвиток рослини. Іншими словами, йому відомий інтервал C , в який повинна потрапити сума $x + a + b$.

Отже, маємо наступну задачу: визначити таке значення зрошувальної норми x , що при будь-яких значеннях параметрів $a \in A$ і $b \in B$ сума $x + a + b \in C$.

Визначивши тут суму двох інтервалів як $A + B = \{a + b \mid a \in A, b \in B\}$, одержимо інтервальне рівняння

$$X + A + B \subseteq C,$$

методи розв'язання якого слід шукати в монографії [6]. В інтервальному рівнянні максимальний за включенням розв'язок і є множиною всіх тих значень x , для яких із $a \in A$ і $b \in B$ сума $x + a + b \in C$.

Замінивши у цьому рівнянні інтервали з чіткими межами на інтервали з розмитими межами, приходимо до рівнянь (2).

Оскільки балансові моделі, власне кажучи, лінійні та скінченновимірні, надалі в роботі основну увагу досліджень буде скеровано на системи лінійних алгебраїчних рівнянь, в основному, погано обумовлених, некоректно поставлених. Для цього нагадаємо один алгоритм розв'язання чітких систем лінійних алгебраїчних рівнянь

$$BX = C \tag{5}$$

із регуляризуючими властивостями.

В (5) $B = (b_{kj})_{\substack{k=\overline{1,m} \\ j=\overline{1,n}}}$, $C = (c_k)$ — відповідно прямокутна матриця і вектор чітких чисел.

У разі некоректно поставлених задач система (5) є або погано обумовленою, або виродженою, або матриця B у ній прямокутна. Для таких систем мінімальний многочлен матриці $A = -B^T B$ має розміри $s \leq r = \text{rang } A$.

Згідно [3] нормальний розв'язок X^+ системи (5) має вигляд

$$X^+ = \frac{1}{d_s} B_{s-1} f,$$

де $f = B^T C$, а значення d_s і B_{s-1} визначаються за алгоритмом

$$\begin{aligned} d_1 &= -\text{Sp } A, & B_0 &= I, \\ d_2 &= -\frac{1}{2} \text{Sp } (B_1 A), & B_1 &= B_0 A + d_1 I, \\ &\dots & & \dots \\ d_n &= -\frac{1}{2} \text{Sp } (B_{n-1} A), & B_{n-1} &= B_{n-2} A + d_{n-1} I, \end{aligned}$$

де $\text{Sp } A$ — слід матриці A .

Цей алгоритм надалі буде використано.

У роботах [2, 3] вказано й інші стійкі (регуляризуючі) алгоритми розв'язання некоректно поставлених задач.

НЕОБХІДНІСТЬ «РОЗМИВАННЯ» ПАРАМЕТРІВ МОДЕЛІ МІЖГАЛУЗЕВОГО БАЛАНСУ

Починаючи з двадцятих років минулого століття модель міжгалузевого балансу широко впроваджувалась для розробки народногосподарських планів різноманітних рівнів та горизонтів планування. На теперішній час зацікавленість цією моделлю не тільки відроджується, але досить часто її включають у різноманіття систем моделей для проведення як теоретичних досліджень, так і практичних розрахунків, пов'язаних із оцінкою альтернатив розвитку економіки України.

Застосування міжгалузевого балансу в наукових та прогнозуючих розрахунках базується на гіпотезі стійкості значень коефіцієнтів прямих витрат

навіть за вельми значних змін правої частини (а отже, і розв'язку) балансової задачі.

У роботі [5] досліджено питання про накопичення помилок у зведеному балансі. Очевидно за таких умов доцільно досліджувати та впроваджувати балансові моделі за нечіткої вхідної інформації. При побудові елементів зведеної матриці балансу використовують нормативні витрати регіональних матриць, а це проводить через операцію осереднення до абсолютної помилки виконання відповідних рівнянь зведеної балансової задачі.

Сучасний науковий стан побудови, аналізу та застосування балансових моделей, які ґрунтуються на моделі Леонт'єва, дозволяє переконатись у реальній змістовності та корисності цих моделей при розв'язанні задач економічного аналізу та прогнозування. Однак, у більшості випадків ці моделі є детермінованими, тобто з високим ступенем ідеалізації відображають реальну ситуацію. Для того, щоб усунути розбіжність між результатами моделі і станом реальної проблеми, необхідно використати один із перевірених часом методів, а саме метод «розмивання» параметрів досліджуваної проблеми.

Підвищення рівня системності математичних моделей можна досягнути завдяки впровадженню в моделі слабкоформалізованих аспектів проблемних ситуацій, опису погано визначених, неоднозначно зрозумілих ситуацій, об'єктів, понять.

На сьогодні, розв'язуючи задачі аналізу складних систем за умов невизначеності, широко використовуються методи теорії ймовірності та математичної статистики. Ці методи припускають імовірнісну інтерпретацію експериментальних даних стосовно параметрів системи та одержання на їх підставі статистичних висновків. Однак, коли невизначеність відносного стану об'єкта дослідження втрачає риси статистичної невизначеності, застосування класичної ймовірності як характеристики масових процесів стає неможливим.

Ймовірнісна міра володіє властивістю адитивності. Але доведено, що реальна поведінка людини найчастіше суперечить припущенню про адитивність. Тому при побудові більш реальних моделей, слід користуватись нечіткими мірами. Нечітка міра вільна від вимог адитивності, що є дуже привабливим при розв'язанні низки задач, у яких присутня невизначеність, що має вигляд нечіткості.

Поняття ймовірнісної міри є звуженням більш загального поняття — нечіткої міри. Незавжди переконатись у тому, що поняття густини ймовірності та функції належності, якими характеризують нечітку множину є порівняльними. Якщо ймовірнісна міра є шкалою для виміру невизначеності типу випадковості, тоді нечіткі міри є об'єктивними шкалами для нечіткості. Таким чином, у теорії ймовірності розглядають статистичну невизначеність, наприклад ймовірність попадання в ціль дорівнює 0,9. Теорія ж нечітких множин (міра можливості) дозволяє опрацювати лінгвістичну невизначеність, наприклад «влучний стрілець». Ймовірнісна міра є частковим випадком нечітких мір довір'я або правдоподібності. Крім того, міру можливості можна побудувати через функцію належності нечіткої множини.

Отже, теорія нечітких мір дозволила скерувати в одне русло весь спектр понять невизначеності, як статистичної, так і лінгвістичної, нечіткої та розпливчастої.

Нечіткі відношення та міри дозволяють при прийнятті рішень моделювати плавну, поступову зміну властивостей, а також невідомі функціональні залежності, що припускають використання нечітких інструкцій, притаманних різноманітним сферам людських дій. Вони дозволяють описати наближені міркування і, отже, корисні як інструмент при прийнятті рішень для тих систем і процесів, що є надто складними, якщо користуватись традиційними кількісними методами.

Теорія нечітких множин та мір — це крок на шляху до зближення точності класичної математики з просякнутим неточністю реальним світом. Більшість класів (понять) реального світу на противагу класам або множинам класичної математики не мають чітких меж, які б відокремлювали об'єкти, що входять у цей клас, від об'єктів, які не входять до нього. До того ж застосування нечітких чисел до прогнозу параметрів вимагає від експерта не формувати миттєві ймовірнісні оцінки, а задавати розрахунковий інтервал значень прогнозованих параметрів. Тоді очікуваний ефект оцінюється експертом також як нечітке число зі своїм розрахунковим розкидом (ступенем нечіткості). На таких засадах дослідник оперує нечіткими мірами та інтегралами. Крім того, методи нечіткої логіки надають можливість кількісної інтерпретації якісних факторів, виражених у термінах природної мови, поєднуючи таким чином переваги кількісного та якісного аналізу.

Аналіз складних систем, який побудовано на підставі теорій нечітких множин, нечітких мір та нечітких інтегралів, дозволяє дати коректний опис розпливчастих тверджень, реалізуючи таким чином спробу подолати лінгвістичний бар'єр між людиною (судження і оцінки якої є наближеними та нечіткими), і машинами, які можуть виконувати тільки чіткі інструкції.

НЕЧІТКИЙ ВАРІАНТ СТАЦІОНАРНОЇ ЛІНІЙНОЇ МОДЕЛІ МІЖГАЛУЗЕВОГО БАЛАНСУ

За умов $y_i, x_i \geq 0$ стаціонарну систему

$$x_i = y_i + \sum_{j=1}^n a_{ij} x_j \quad (6)$$

називають стандартною системою міжгалузевого балансу.

У (6) вектори y, x визначають можливі значення правої частини та розв'язку балансової задачі.

Перш ніж перейти до нечіткого аналога системи (6) вважатимемо, що всі розглянені нижче нечіткі числа є опуклими і нормальними, або вони задовольняють умові

$$\mu_a(x) = \min(L_a(x), R_a(x)),$$

де $L_a \in \mathbf{L}$, $R_a \in \mathbf{R}$.

Кожне таке число \tilde{a} можна зобразити через α -рівневий розклад

$$\tilde{a} = \bigcup_{\alpha \in [0,1]} (\underline{a}_\alpha, \bar{a}_\alpha), \quad (7)$$

де \underline{a}_α (\bar{a}_α) — нижня (верхня) межі нечіткого числа \tilde{a} на α -рівні.

Розглянемо нечітку модель міжгалузевого балансу

$$\tilde{X} = \tilde{Y} + \tilde{A}\tilde{X}, \quad (8)$$

яка в координатній формі має вигляд

$$\tilde{x}_i = \tilde{y}_i + \sum_{j=1}^n \tilde{a}_{ij} \tilde{x}_j, \quad i = \overline{1, n}.$$

Нехай нечіткі параметри \tilde{a}_{ij} , \tilde{y}_i системи (8) допускають зображення (7) із $\underline{a}_{ij}^{\alpha_s}$ ($\bar{a}_{ij}^{\alpha_s}$) та $\underline{y}_i^{\alpha_s}$ ($\bar{y}_i^{\alpha_s}$), де $s = \overline{1, k}$.

Тоді операторні рівняння (4) перетворюються у скобкові системи

$$\underline{x}_i^s = \underline{y}_i^{\alpha_s} + \sum_{j=1}^n \underline{a}_{ij}^{\alpha_s} \underline{x}_j^s \quad (9)$$

та

$$\bar{x}_i^s = \bar{y}_i^{\alpha_s} + \sum_{j=1}^n \bar{a}_{ij}^{\alpha_s} \bar{x}_j^s. \quad (10)$$

Запишемо тепер кожному нечітку множину \tilde{a}_{ij} та \tilde{y}_i із $i = \overline{1, n}$ згідно об'єднанню (7), а це через виконання умов

$$(\underline{a}_{\alpha_k}, \bar{a}_{\alpha_k}) \leq (\underline{a}_{\alpha_{k-1}}, \bar{a}_{\alpha_{k-1}}) \leq \dots \leq (\underline{a}_{\alpha_1}, \bar{a}_{\alpha_1})$$

дозволить для кожного $x_i = z$ побудувати функції належності (рисунок).

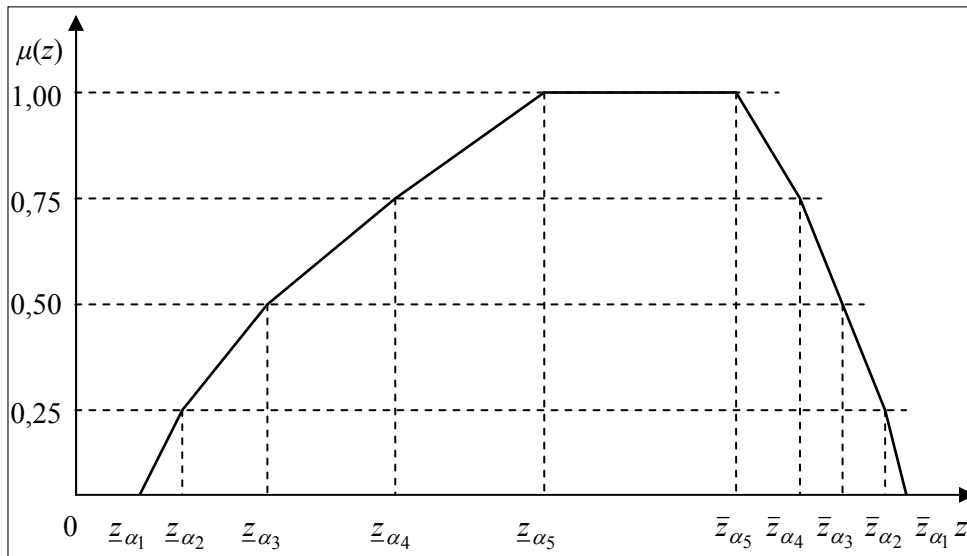


Рисунок. Код функції належності

Приклад. Розглянемо модель (8) із

$$\tilde{A} = \begin{pmatrix} 0,3 & 0,1 \\ 0,4 & 0,5 \end{pmatrix}, \quad \tilde{Y} = \begin{pmatrix} 200 \\ 100 \end{pmatrix},$$

де

$$\begin{aligned} 0,3 &= \frac{0,1}{0} + \frac{0,15}{0,5} + \frac{0,2}{1} + \frac{0,35}{1} + \frac{0,375}{0,5} + \frac{0,4}{0}; \\ 0,1 &= \frac{0}{0} + \frac{0,025}{0,5} + \frac{0,05}{1} + \frac{0,2}{1} + \frac{0,25}{0,5} + \frac{0,3}{0}; \\ 0,4 &= \frac{0,2}{0} + \frac{0,25}{0,5} + \frac{0,31}{1} + \frac{0,5}{1} + \frac{0,525}{0,5} + \frac{0,55}{0}; \\ 0,5 &= \frac{0,3}{0} + \frac{0,35}{0,5} + \frac{0,4}{1} + \frac{0,55}{1} + \frac{0,55}{0,5} + \frac{0,55}{0}, \end{aligned}$$

а

$$\begin{aligned} 200 &= \frac{160}{0} + \frac{170}{0,5} + \frac{180}{1} + \frac{210}{1} + \frac{215}{0,5} + \frac{220}{0}; \\ 100 &= \frac{80}{0} + \frac{85}{0,5} + \frac{90}{1} + \frac{110}{1} + \frac{120}{0,5} + \frac{130}{0}. \end{aligned}$$

Тоді, позначивши

$$\begin{aligned} \tilde{x}_1 &= \frac{x_1^{(0)}}{0} + \frac{x_1^{(0,5)}}{0,5} + \frac{x_1^{(1)}}{1} + \frac{x_1^{(1)}}{1} + \frac{x_1^{(0,5)}}{0,5} + \frac{x_1^{(0)}}{0}, \\ \tilde{x}_2 &= \frac{x_2^{(0)}}{0} + \frac{x_2^{(0,5)}}{0,5} + \frac{x_2^{(1)}}{1} + \frac{x_2^{(1)}}{1} + \frac{x_2^{(0,5)}}{0,5} + \frac{x_2^{(0)}}{0} \end{aligned}$$

та розв'язавши відповідні системи (9), (10), одержимо нечіткі розв'язки

$$\begin{aligned} \tilde{x}_1 &= \frac{178}{0} + \frac{206,2}{0,5} + \frac{605,2}{1} + \frac{828,3}{1} + \frac{215}{0,5} + \frac{1314,3}{0}, \\ \tilde{x}_2 &= \frac{165}{0} + \frac{210,1}{0,5} + \frac{444,9}{1} + \frac{916,9}{1} + \frac{1233}{0,5} + \frac{1895,3}{0}. \end{aligned}$$

Дефазифікуючи ці значення згідно центру ваги, одержимо розв'язки

$$x_1^{ЦВ} = \frac{1375,25}{3} = 458,4; \quad x_2^{ЦВ} = \frac{2083,35}{3} = 694,45.$$

Тут можливий і варіант відшукування нормального розв'язку розмитої системи (8). Наприклад, коли шуканий X і постійний Y вектори — чіткі, а розмитою є матриця A , тоді нормальний розв'язок задачі матиме вигляд

$$X^+ = \begin{pmatrix} 248,724 \\ 285,594 \end{pmatrix}.$$

Аналогічно, коли ж розмитою є \tilde{Y} , то

$$X^+ = \begin{pmatrix} 223,312 \\ 258,185 \end{pmatrix}.$$

Актуальним є також аналог оптимізаційних моделей міжгалузевого балансу. Звернемо увагу на оптимізаційну модель міжгалузевого еколого-економічного балансу. За нечіткої вхідної інформації вона має вигляд

$$\langle \tilde{P}, X \rangle \Rightarrow \max \quad (11)$$

$$HX \leq \tilde{Q}, \quad X \geq 0$$

із

$$H = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad \tilde{Q} = \begin{pmatrix} \tilde{R} \\ \tilde{Z} \end{pmatrix}, \quad \tilde{P} = \begin{pmatrix} c^{(1)} \\ -c^{(2)} \end{pmatrix}, \quad X = \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix},$$

що несуть певне еколого-економічне навантаження.

Методи розв'язання задач типу (11) див. в роботах [2, 3].

ВИСНОВКИ

Використання теорії нечітких множин, мір та інтегралів дозволяє по новому розглянути задачі міжгалузевого балансу, більш гнучко оцінити результати досліджень, всесторонньо використати знання та рекомендації експертів у процесі моделювання, ефективно враховувати у задачах якісні показники.

ЛІТЕРАТУРА

1. Mizumoto M., Tanaka K. Algebraic properties of Fuzzy numbers // Proc.of the IEEE Intern. Conf. on Cybernetics and Society. — Washington: IEEE, 1976. — P. 559–563.
2. Рибицька О.М., Сявакко М.С. Математичні аспекти відновлення інформації. — Львів: Растр-7, 2008. — 320 с.
3. Сявакко М.С., Рибицька О.М. Математичне моделювання за умов невизначеності. — Львів: Українські технології, 2000. — 320 с.
4. Гвоздик А.А. Решение нечетких уравнений // Изв. АН СССР. Техн. Кибернетика. — 1984. — № 5. — С. 176–183.
5. Борисов А.Н., Крумберг О.А., Федоров И.П. Принятие решений на основе нечетких моделей. Примеры использования. — Рига: Знание, 1990. — 184 с.
6. Альфельд Г., Херцбергер Ю. Введение в интервальные вычисления. — М.: Мир, 1987. — 360 с.
7. Медвицкий В.Г., Медвицкий Ю.В. О зависимости значений элементов балансовых матриц от цен и технологии производства // Экономика и мат.методы. — 2004. — 40, № 1. — С. 90–104.

Надійшла 01.10.2008

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ,
ВИСОКОПРОДУКТИВНІ КОМП'ЮТЕРНІ
СИСТЕМИ

PROGRESSIVE INFORMATION TECHNOLOGIES,
HIGH-EFFICIENCY COMPUTER SYSTEMS

УДК 519.7.007.52

Е-наука на пути к семантическому Грид. Часть 2: Семантический Web- и семантический Грид / Згуровский М.З., Петренко А.И. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 7–25.

Исследован семантический Web, основанный на метаданных, онтологиях и программных агентах как фундамент для семантического Грид, в котором информация и услуги имеют четкое определение, предоставляющее возможность компьютерам и людям работать в кооперации со знанием, а не с информацией, как это делается в настоящее время. Рассмотрены подходы к описанию онтологии данных и Web- и Грид-сервисов (с состоянием и без состояния), их взаимодействия, архитектуры, а также существующие реализации семантических Грид. Рис.: 6. Табл.: 1. Библиогр.: 32 назв.

UDC 519.7.007.52

E-Science on the way to semantic Grid. Part 2: Semantic Web- and semantic Grid / Zgurovsky M.Z., Petrenko A.I. // System Research and Information Technologies. — 2010. — № 2. — P. 7–25.

The semantic Web, based on metadata, ontology, and software agents, is investigated as a foundation for the semantic Grid, in which information and services have clear determination, which gives computers and people the possibility to cooperate rather with knowledge than with information (as that is done at present). Approaches to ontology and Web- and Grid-services with and without state, their cooperation, architecture, and the existent realizations of the semantic Grid are considered. Figs: 6. Tabl.: 1. Refs: 32 titles.

УДК 004.043

Сжатие RGB-изображений без потерь с использованием палитры / Шпортко А.В. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 26–36.

Рассмотрены и обоснованы возможности сжатия RGB-изображений без потерь с помощью палитрования. Описан вариант алгоритма для реализации такого сжатия с разбиением результата на трендовую и шумовую составляющие. Приведены результаты применения программы, разработанной согласно предложенного алгоритма для сжатия изображений набора АСТ. Рис.: 5. Табл.: 4. Библиогр.: 5 назв.

UDC 004.043

Loss-free compression of RGB-images using a palette / Shportko A.V. // System Research and Information Technologies. — 2010. — № 2. — P. 26–36.

The possibilities of RGB-image compression without losses using a palette are considered. A version of the algorithm for realization of the compression with partition of the result into a trend and noise components is described. The results of using the program developed according to the offered algorithm for compression of the images of the ACT set are presented. Figs: 5. Tabl.: 4. Refs: 5 titles.

ПРОБЛЕМИ ПРИЙНЯТТЯ РІШЕНЬ І
УПРАВЛІННЯ В ЕКОНОМІЧНИХ, ТЕХНІЧНИХ,
ЕКОЛОГІЧНИХ І СОЦІАЛЬНИХ СИСТЕМАХ

DECISION MAKING AND CONTROL IN
ECONOMIC, TECHNICAL, ECOLOGICAL AND
SOCIAL SYSTEMS

УДК 519.8

Оцінка кредитних банківських ризиків із використанням нечіткої логіки / Зайченко Ю.П. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 37–54.

Розглянуто проблеми оцінки кредитних ризиків позичальників у банках. Викладено класичний метод оцінки кредитних ризиків на основі визначення кредитного рейтингу позичальників, а також оцінки забезпечення кредитної операції. Запропоновано новий метод визначення кредитоспроможності на основі використання систем із нечіткою логікою (алгоритм нечіткого висновку Мамдані). Сформульовано правила нечіткого висновку для оцінки кредитного рейтингу та загального класу позичальників. Наведено приклади застосування викладеного нечіткого методу. Рис.: 4. Табл.: 2. Бібліогр.: 2 назви.

UDC 519.8

Credit risk estimation on the base of fuzzy logic / Zaychenko Yu.P. // System Research and Information Technologies. — 2010. — № 2. — P. 37–54.

The problem of credit risk estimation is considered. The classical method of credit risk estimation based on borrower's credit rating determination is presented. A novel method of credits rating determination based on fuzzy logic systems (Mamdani algorithm) is suggested. Rules for fuzzy credit rating determination and general class of borrower's are formulated. Some applications of the suggested fuzzy method are presented. Figs: 4. Tabl.: 2. Refs: 2 titles.

МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ, ПРОБЛЕМИ І
ТЕХНОЛОГІЇ ДОСЛІДЖЕННЯ СКЛАДНИХ
СИСТЕМ

MATHEMATICAL METHODS, MODELS,
PROBLEMS AND TECHNOLOGIES FOR
COMPLEX SYSTEMS RESEARCH

УДК 519.2:621

Прогнозування та ймовірнісна оцінка довговічності технічних об'єктів / Федін С.С. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 55–63.

На основі методології марківського аналізу надійності запропоновано ймовірнісний підхід до прогнозування та оцінки довговічності технічних об'єктів. Розроблено математичну модель прогнозування довговічності з використанням статистичних даних, що характеризують початковий та кінцевий розподіл тривалості терміну служби технічних об'єктів. Рис.: 7. Табл.: 1. Бібліогр.: 7 назв.

UDC 519.2:621

Forecast and probabilistic estimations of technical object longevity / Fedin S.S. // System Research and Information Technologies. — 2010. — № 2. — P. 55–63.

Based on the methodology of the Markov analysis, a probabilistic approach to forecast and estimation of the longevity of technical objects is offered. A mathematical model of the longevity forecast is developed with the use of statistical data characterizing the initial and final distribution of the service term of technical objects. Figs: 7. Tabl.: 1. Refs: 7 titles.

УДК 004.93(015.7)

Аналіз використання нейронних мереж для діагностики раку шийки матки по мультиспектральному зображенню / Малишевська К.М. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 64–71.

Розглянуто проблему діагностики раку шийки матки за допомогою нейронних мереж по мультиспектральному зображенню. Для діагностики використано нейронні мережі зі зворотним розповсюдженням помилки та мережі з радіальним базисом. Приведено результати експериментальних досліджень. Рис.: 2. Табл.: 1. Бібліогр.: 6 назв.

UDC 004.93(015.7)

Analysis of using neural networks for cancer diagnosis from multisensor imagery / Malyshevska K.M. // System Research and Information Technologies. — 2010. — № 2. — P. 64–71.

The problem of cervical cancer diagnosis using neural networks from multisensor imagery is investigated. The radial basis function networks and error backpropagation networks are used for the diagnosis. The results of experiments are presented. Figs: 2. Tabl.: 1. Refs: 6 titles.

МЕТОДИ АНАЛІЗУ ТА УПРАВЛІННЯ
СИСТЕМАМИ В УМОВАХ РИЗИКУ І
НЕВИЗНАЧЕНОСТІ

METHODS OF SYSTEM ANALYSIS AND
CONTROL IN CONDITIONS OF RISK AND
UNCERTAINTY

УДК 519.8

Формування підгруп експертів із узгодженими думками й визначення групової узагальненої оцінки багатознакових об'єктів / Крапухіна Н.В., Пронічкін С.В., Риков О.С. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 72–79.

Розглянуто нові алгоритми обробки експертних оцінок багатоозначових об'єктів. Алгоритми застосовано для вирішення практичного завдання оцінки компетенцій студентів, які оцінені декількома експертами за багатьма критеріями. Застосування розроблених алгоритмів дозволило отримати рішення адекватніше перевагам ОПР (особа, яка приймає рішення), ніж у випадку вживання існуючих алгоритмів. Бібліогр.: 8 назв.

UDC 519.8

Formation of expert subgroups based on consensus and deriving generalized estimate of multiattribute objects / Krapuhina N.V., Pronichkin S.V., Rykov A.S. // System Research and Information Technologies. — 2010. — № 2. — P. 72–79.

New algorithms for handling expert estimates of multiattribute objects are considered and used for estimation of the competences of students estimated by several experts using many criteria. The application of the developed algorithms helps to make a more adequate decision to preferences of the PMD compared to the existing algorithms. Refs: 8 titles.

УДК 621. 34-52

Синтез регулятора стану для дискретної в часі системи модального управління / Філатов О.Г. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 80–93.

Запропоновано процедуру синтезу регулятора стану для дискретної в часі системи керування, яка базується на теорії модального керування та забезпечує довільний і максимальний ступінь стійкості замкненої системи у просторі змінних стану та їх середніх інтегральних значень. Отримані результати проілюстровано на прикладі. Рис.: 4. Бібліогр.: 5 назв.

UDC 621. 34-52

Design of state regulator for discrete-time system of modal control / Filatov O.G. // System Research and Information Technologies. — 2010. — № 2. — P. 80–93.

A procedure for design of a state regulator for a discrete-time control system is proposed. It is based on the theory of modal control and provides arbitrary and maximal degree of stability for a closed system in the space of variable states and their average integral values. The results obtained are illustrated by an example. Figs: 4. Refs: 5 titles.

УДК 517.926:681.518.2

Мінімаксна рекурсивна оцінка стану лінійних дескрипторних систем із дискретним часом / Жук С.М. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 94–105.

Розглянуто підхід до оцінювання стану системи, що описується дескрипторним рівнянням із дискретним часом за спостереженнями, що надходять

у реальному часі. Підхід базується на понятті лінійної мінімаксної оцінки та індексу причинності, що вводяться у статті для сингулярних різницевих рівнянь. Рекурсивний оцінювач стану будується шляхом застосування методу «Київського віника» та теорії псевдоінверсних матриць до проблеми мінімаксного оцінювання. Библиогр.: 14 назв.

УДК 517.926:681.518.2

Минимаксная рекурсивная оценка состояния линейных дескрипторных систем с дискретным временем / Жук С.М. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 94–105.

Рассмотрен подход к оцениванию состояния системы, описываемой дескрипторным уравнением с дискретным временем по наблюдениям, поступающим в реальном времени. Подход основан на понятии линейной минимаксной оценки и индекса причинности, введенных в статье для сингулярных разностных уравнений. Рекурсивный оценщик строится путем применения метода «Киевского веника» и теории псевдообратных матриц к проблеме минимаксного оценивания. Библиогр.: 14 назв.

ЕВРИСТИЧНІ МЕТОДИ ТА АЛГОРИТМИ В
СИСТЕМНОМУ АНАЛІЗІ ТА УПРАВЛІННІ

HEURISTIC METHODS AND ALGORITHMS IN
SYSTEM ANALYSIS AND CONTROL

УДК 519.873

Календарне технічне обслуговування простої системи з урахуванням мінімального аварійного поновлення / Песчанський О.І. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 106–117.

Розглянуто стратегію мінімальних поновлень із періодично повними змінами й урахуванням часу поновлення. Для системи з простою структурою введено альтернуючий процес мінімального поновлення. Досліджено його асимптотичну поведінку за допомогою тауберових теорем. Визначено оптимальні терміни проведення повних замін. Рис.: 3. Табл.: 2. Библиогр.: 3 назви.

UDC 519.873

Calendar maintenance of a single-component system with minimal fallback renewal / Peschansky A.I. // System Research and Information Technologies. — 2010. — № 2. — P. 106–117.

Strategy for minimal fallback renewal with periodical total replacement is considered taking into account the time of renewal. For single-component systems, an alternating process of minimal renewal is introduced and its asymptotic behavior is investigated using tauberian theorems. Optimal terms of total replacement are determined. Figs: 3. Tabl.: 2. Refs: 3 titles.

УДК 681.518.2:519.816

Інформаційна технологія формування вигляду елементів технічних систем на основі дискретних даних щодо аналогів / Угрюмова К.М., Трончук О.А., Афанасьєвська В.Є. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 118–133.

Запропоновано методологію та комп'ютерну систему підтримки прийняття рішень щодо формування вигляду елементів технічних систем. Здійснено побудову формальних математичних моделей систем за допомогою штучних нейронних мереж, які можна навчати. Визначено керівні змінні, які відповідають бажаним критеріям якості на основі застосування еволюційного методу з регуляризацією. Рис.: 4. Табл.: 1. Библиогр.: 11 назв.

UDC 681.518.2:519.816

Information technology for generation of engineering system elements on the basis of discrete data on analogs / Ugryumova E.M., Tronchuk A.A., Afanasjevska V.E. // System Research and Information Technologies. — 2010. — № 2. — P. 118–133.

Methodology and a computer decision support system for engineering system element habit generation are offered. Formal mathematical models of the system have been constructed using artificial neural networks. Control variables that are consistent with the desired quality criteria are determined on the basis of the evolutionary method with regularization. Figs: 4. Tabl.: 1. Refs: 11 titles.

НОВІ МЕТОДИ В СИСТЕМНОМУ АНАЛІЗІ,
ІНФОРМАТИЦІ ТА ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ

NEW METHODS IN SYSTEM ANALYSIS, COMPUTER
SCIENCE AND THEORY OF DECISION MAKING

УДК 303.725.34:510,22;62-50

Нечеткая арифметика в задачах межотраслевого баланса / Сявавко М.С., Пасичник Т.В., Тымкив В.П. // Системні дослідження та інформаційні технології. — 2010. — № 2. — С. 134–143.

Предложено использование теории нечетких мер для задач межотраслевого баланса. Показано как методы нечеткой логики предоставляют возможность количественной интерпретации качественных факторов. Предлагается использовать дробно-рациональный алгоритм для нахождения нормального решения некорректно поставленных задач или плохо обусловленных систем. Приведены примеры решения нечетких задач. Рис.: 1. Библиогр.: 7 назв.

UDC 303.725.34:510,22;62-50

Fuzzy arithmetic in problems of interindustry balance / Siavavko M.S., Pasichnyk T.V., Tymkiv V.P. // System Research and Information Technologies. — 2010. — № 2. — P. 134–143.

To solve problems of interindustry balance, the use of fuzzy-step theory is proposed. Fuzzy logic methods are shown to enable quantitative interpretation of qualitative factors. To find a normal solution to incorrectly-assigned tasks or ill-conditioned systems, the use of the fractionally rational algorithm is offered. Examples of solutions to fuzzy problems are given. Figs: 1. Refs: 7 titles.

ВІДОМОСТІ ПРО АВТОРІВ

- Афанасьєвська Вікторія Євгенівна,**
аспірантка Національного аерокосмічного університету ім. М.Є. Жуковського «ХАІ»,
Україна, Харків
- Жук Сергій Миколайович,**
кандидат фізико-математичних наук, молодший науковий співробітник Київського
національного університету ім. Тараса Шевченка, Україна, Київ
- Зайченко Юрій Петрович,**
професор, доктор технічних наук, професор Навчально-наукового комплексу «Інсти-
тут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ
- Згуровський Михайло Захарович,**
академік НАН України, професор, доктор технічних наук, ректор Національного тех-
нічного університету України «КПІ», директор Навчально-наукового комплексу
«Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ
- Крапухіна Ніна Володимирівна,**
професор, кандидат технічних наук, директор Московського державного інституту
сталі та сплавів, завідувач кафедри інженерної кібернетики, Росія, Москва
- Малишевська Катерина Миколаївна,**
аспірантка Навчально-наукового комплексу «Інститут прикладного системного аналізу»
НТУУ «КПІ» МОН та НАН України, Київ
- Пасічник Тимофій Васильович,**
доцент, кандидат фізико-математичних наук, завідувач кафедри інформаційних тех-
нологій Львівського національного аграрного університету, Україна, Львів
- Песчанський Олексій Іванович,**
кандидат фізико-математичних наук, доцент кафедри вищої математики Севасто-
польського національного технічного університету, Україна, Севастополь
- Петренко Анатолій Іванович,**
професор, доктор технічних наук, завідувач кафедри систем автоматизованого проєк-
тування Навчально-наукового комплексу «Інститут прикладного системного аналізу»
НТУУ «КПІ» МОН та НАН України, Київ
- Пронічкін Сергій Васильович,**
аспірант Московського державного інституту сталі та сплавів, Росія, Москва
- Риков Олександр Семенович,**
академік РАН, доктор технічних наук, професор кафедри автоматизованих систем
управління Московського державного інституту сталі та сплавів, Росія, Москва
- Сявавко Мар'ян Степанович,**
професор, доктор фізико-математичних наук, завідувач кафедри обчислюваль-
ної математики та моделювання Інституту новітніх технологій та управління
ім. В. Чорновола, Україна, Львів
- Тимків Володимир Пилипович,**
завідувач відділення обліку та фінансів Коломийського економіко-правового коледжу
(відділення Київського національного торговельно-економічного університету),
Україна, Коломия
- Трончук Олексій Адамович,**
аспірант Національного аерокосмічного університету ім. М.Є. Жуковського «ХАІ»,
Україна, Харків
- Угрюмова Катерина Михайлівна,**
аспірантка Національного аерокосмічного університету ім. М.Є. Жуковського «ХАІ»,
Україна, Харків
- Федін Сергій Сергійович,**
кандидат технічних наук, доцент Київського національного університету технологій
та дизайну, Україна, Київ
- Філатов Олександр Григорович,**
кандидат технічних наук, доцент факультету електроенерготехніки та автоматики
Національного технічного університету України «КПІ», Україна, Київ
- Шпортко Олександр Володимирович,**
аспірант кафедри інформатики та прикладної математики Рівненського державного
гуманітарного університету, Україна, Київ

ПРАВИЛА ОФОРМЛЕННЯ СТАТЕЙ

для журналу

«СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ»

1. Стаття з анотаціями подається у двох варіантах: **роздрукована** (на папері формату А4 через 1,5 інтервали у двох примірниках, розмір шрифту 12 пт., поля зліва та справа по 3 см) та в **електронному** вигляді.

Назва статті, прізвище та ініціали автора (авторів), а також анотація (від 700 до 1000 символів) подаються українською, російською та англійською мовами.

2. Текст статті набирається в редакторі Microsoft Word 2003 шрифтом Times New Roman із вирівнюванням по ширині.

3. На початку статті вказується індекс УДК, назва статті, ініціали та прізвище автора (авторів).

4. Список літератури подається в порядку посилання після тексту статті.

5. Набір формул здійснюється в редакторі формул «Equations».

6. Символи у формулах (крім особливих випадків) набираються курсивом. Допускаються прямі символи, але їх треба розмітити або включити до списку спеціальних символів.

7. Нумерувати необхідно тільки ті формули (винесені на середину рядків), на які є посилання в тексті.

8. Таблиці та рисунки виконуються в одному стилі, їм надається номер, назва і розташовуються після посилань у тексті. Не слід повторювати в статті підписи до рисунків та таблиць.

9. Стаття разом із таблицями, рисунками і списком літератури має бути не більше 15 сторінок. Оглядові статті обсягом до 25 сторінок приймаються за домовленістю із редколегією.

10. До рукопису додаються відомості про кожного автора: країна, місто, місце роботи, службова посада, вчений ступінь та звання, сфера наукових досліджень, контактні телефони, адреса та e-mail.