

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

**БЕЗСОНОВ ОЛЕКСАНДР ОЛЕКСАНДРОВИЧ**

УДК 004.852 : 004.896

**ЕВОЛЮЦІЙНІ ШТУЧНІ НЕЙРОННІ МЕРЕЖІ ПРЯМОГО  
РОЗПОВСЮДЖЕННЯ: АРХІТЕКТУРИ, НАВЧАННЯ, ЗАСТОСУВАННЯ**

05.13.23 – системи та засоби штучного інтелекту

**АВТОРЕФЕРАТ**

дисертації на здобуття наукового ступеня  
доктора технічних наук

Харків – 2017

Дисертацією є рукопис.

Робота виконана в Харківському національному університеті радіоелектроніки Міністерства освіти і науки України.

**Науковий консультант:** доктор технічних наук, професор  
**Руденко Олег Григорійович,**  
Харківський національний економічний  
університет ім. С. Кузнеця МОН України,  
завідувач кафедри інформаційних систем

**Офіційні опоненти:** доктор технічних наук, професор  
**Бідюк Петро Іванович,**  
інститут прикладного системного аналізу  
Національного технічного університету України  
«КПІ» МОН України, професор кафедри матема-  
тичних методів системного аналізу;

доктор технічних наук, професор  
**Гороховатський Володимир Олексійович,**  
Харківський навчально-науковий інститут ДВНЗ  
«Університет банківської справи» МОН України,  
професор кафедри інформаційних технологій та  
вищої математики;

доктор технічних наук, професор  
**Литвиненко Володимир Іванович,**  
Херсонський національний технічний університет  
МОН України, завідувач кафедри інформатики та  
комп'ютерних наук.

Захист відбудеться «22» \_\_\_\_\_ березня \_\_\_\_\_ 2017 р. о 13-00 годині на засіданні спеціалізованої вченої ради Д 64.052.01 у Харківському національному університеті радіоелектроніки (61166, м. Харків, просп. Науки, 14).

З дисертацією можна ознайомитися у бібліотеці Харківського національного університету радіоелектроніки (61166, м. Харків, просп. Науки, 14).

Автореферат розісланий «06» \_\_\_\_\_ лютого \_\_\_\_\_ 2017 р.

Учений секретар  
спеціалізованої вченої ради,  
д.т.н., проф.

О.А. Винокурова

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** Вирішення широкого кола задач різноманітних галузей науки, техніки та економіки, таких як ідентифікація, управління, фільтрація, прогнозування тощо, пов'язане із апроксимацією деяких нелінійних функцій. Відсутність інформації щодо виду нелінійності часто призводить до неефективності традиційних методів апроксимації, а в ряді випадків – до їх непридатності. Альтернативою традиційним методам є застосування нейромережевих технологій.

Будучи універсальними апроксиматорами, деякі типи штучних нейронних мереж (ШНМ) дозволяють відновити із заданою точністю будь-яку складну безперервну нелінійну функцію.

При синтезі або використанні ШНМ виникають задачі структурної та параметричної оптимізації відповідно до вибору оптимальної топології мережі та її навчання (налаштування параметрів). Якщо задача визначення структури є дискретною оптимізаційною (комбінаторною), то пошук оптимальних параметрів здійснюється у безперервному просторі за допомогою класичних методів оптимізації.

Для навчання (оцінювання параметрів) мережі застосовуються, як правило, методи, що потребують обчислення градієнта функціоналу, який використовується (алгоритм зворотного поширення помилки (ЗП), метод сполучених градієнтів, алгоритм Гауса-Ньютона, Левенберга-Марквардта тощо). Незважаючи на популярність цих методів не тільки при навчанні ШНМ, але і при вирішенні інших задач оптимізації, вони мають ряд суттєвих недоліків.

Спроби усунути недоліки традиційних методів синтезу і функціонування ШНМ призвели до появи нового класу мереж – еволюційних ШНМ (ЕШНМ), в яких, на додаток до традиційного навчання, використовується інша фундаментальна форма адаптації – еволюція, що реалізується шляхом застосування еволюційних обчислень.

Використання в ЕШНМ цих двох форм адаптації – еволюції та навчання, що дозволяють змінювати структуру мережі, її параметри й процедури навчання без зовнішнього втручання, робить такі мережі найбільш пристосованими для роботи в нестаціонарних умовах і за наявності невизначеності щодо властивостей досліджуваного об'єкта та умов його функціонування.

Основною перевагою використання еволюційних алгоритмів (ЕА) як алгоритмів навчання є те, що багато параметрів ШНМ можуть бути закодовані в геномі та визначатися паралельно. Більш того, на відміну від більшості алгоритмів оптимізації, призначених для покрокового вирішення задачі, ЕА оперують з безліччю рішень – популяцією, що дозволяє досягти глобального екстремуму, та не застрягти в локальних. При цьому інформація про кожну особину популяції кодується в хромосомі (генотипі), а отримання рішення (фенотипу) здійснюється після еволюції (відбору, схрещування, мутації) шляхом декодування.

У зв'язку з цим актуальною задачею як з теоретичної, так і з практичної точки зору є розробка нових методів інтелектуальної обробки різноманітної інформації, які враховують властивості штучних нейронних мереж, еволюційні та

коеволюційні обчислення, генетичні алгоритми тощо. Задачі, що при цьому виникають, обумовили напрям досліджень представленої дисертаційної роботи.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційна робота виконана в рамках держбюджетних тем «Еволюційні гібридні системи обчислювального інтелекту зі змінною структурою для інтелектуального аналізу даних» (№ДР0110U000458), розділ «Еволюційні гібридні методи і моделі інтелектуального опрацювання інформації зі змінною структурою в умовах невизначеності», «Нейро-фаззі системи для поточної кластеризації та класифікації послідовностей даних в умовах їх спотворення відсутніми і аномальними спостереженнями» (№ДР0113U000361), розділ «Адаптивні методи і моделі класифікації даних і прогнозування часових рядів в умовах їх спотворення відсутніми і аномальними спостереженнями на основі штучних імунних систем», «Розробка теоретичних основ і математичного забезпечення нейро-фаззі-систем ранньої діагностики, прогнозування і моделювання в умовах апріорної і поточної невизначеності» (№ДР0101U001762), а також держдоговірної теми «Розробка та виготовлення системи збору і обробки інформації про стан трубчастих печей виробництва феросплавів» (№ДР0104U0009291), затверджених Міністерством освіти і науки України, які виконувалися в ХНУРЕ. Автор був одним із виконавців робіт за даними темами. У межах наведених тем здобувач як виконавець запропонував методи та моделі ідентифікації та управління динамічними об'єктами на основі еволюційних штучних нейронних мереж з урахуванням апріорної інформації про властивості вхідних даних.

**Мета та завдання дослідження.** Метою дисертаційної роботи є розвиток теоретичних основ (методів та реалізації) і розробки нових еволюційних ШНМ для вирішення проблеми підвищення якості інтелектуального аналізу і обробки інформації при наявності апріорної і поточної невизначеності. Для досягнення цієї мети в роботі вирішуються такі основні задачі:

1. Аналіз існуючих методів нейромережевої обробки інформації в умовах апріорної та поточної невизначеності.

2. Розробка нових та вдосконалення існуючих архітектур ШНМ з використанням еволюційного та коеволюційного підходів, орієнтованих на вирішення завдання інтелектуального аналізу даних в умовах апріорної та поточної невизначеності.

3. Розробка методів автоматичного визначення та корекції структури і параметрів ШНМ в залежності від зміни властивостей досліджуваного об'єкта.

4. Розробка методів спрощення структур та методів навчання ШНМ прямого розповсюдження з метою прискорення процесів обробки інформації при допустимій неточності.

5. Розробка Парето-еволюційних ШНМ прямого розповсюдження.

6. Розробка методів навчання еволюційних ШНМ, які володіли б підвищеною швидкістю навчання та робастністю, при наявності негаусівських завад, зокрема, з асиметричними розподілами.

7. Експериментальні дослідження властивостей та характеристик різних методів, розробка рекомендацій щодо їх застосування, вирішення тестових та

практичних завдань за допомогою розроблених ЕШНМ.

*Об'єкт дослідження* – процеси побудови інтелектуальних систем обробки інформації.

*Предмет дослідження* – еволюційні ШНМ прямого розповсюдження, призначені для інтелектуальної обробки даних за умов нестаціонарності та ап'іорної і поточної невизначеності.

*Методи дослідження* ґрунтуються на теорії обчислювального інтелекту, а саме: на методах теорії штучних нейронних мереж, яка дозволила синтезувати нейромережеві моделі та нейрорегулятори й отримати процедури їх навчання; методи теорії оптимальності, за допомогою якої були синтезовані швидкодіючі процедури навчання; методи теорії ідентифікації, на основі яких були синтезовані моделі розглянутих об'єктів управління, що настраюються; методи імітаційного моделювання, що дозволили підтвердити ефективність отриманих результатів та розробити рекомендації щодо їх практичного використання. Експериментальні дослідження проводилися в лабораторних умовах та на реальних об'єктах.

**Наукова новизна результатів дисертаційної роботи** полягає в тому, що:

*Уперше*

– запропоновано метод робастної багатокритеріальної оптимізації (Парето-оптимізації) на основі робастних фітнес-функцій та інформаційних критеріїв, які дають можливість визначати оптимальну структуру нейромережевої моделі досліджуваного об'єкта при наявності негаусівських завад;

– запропоновано узагальнений ЕА Парето-оптимізації на основі коеволюційного підходу, який дозволяє змінювати архітектуру мережі, адаптуючись до мінливого зовнішнього середовища;

– запропоновано прості в обчислювальному відношенні одно- і багатокрокові рекурентні процедури навчання, що забезпечують необхідну точність при наявності обмежених завад, та розроблені процедури адаптивної корекції параметрів;

– розроблено методи спрощення структур та методи навчання ШНМ прямого розповсюдження з метою прискорення процесів обробки інформації при допустимій неточності в умовах обмежених завад;

– запропоновано робастний метод навчання ШНМ, що дає можливість обробляти інформацію при наявності завад з негаусовськими, зокрема асиметричними, розподілами та усувати зміщення оцінок параметрів мереж, характерне для традиційних методів навчання;

– запропоновано рекурентні методи оцінювання параметрів функціоналів і завад для моделі Тьюки-Хьюбера, що дозволяє при відсутності інформації про статистичні властивості завад коригувати одержувані в процесі навчання параметри ШНМ;

– запропоновано методи апроксимації гаусівських базисних функцій в РБМ нульового та першого порядків, що дозволяють істотно спростити обчислення, супутні процесам побудови моделі досліджуваного об'єкта;

– запропоновано закони адаптивного прогнозуючого нейроуправління нелінійними нестационарними об'єктами, що функціонують в умовах невизначеності, на основі еволюційного підходу з корекцією еталонної траєкторії, що дозволяє істотно прискорити процес синтезу моделі та обчислення керуючого сигналу.

*Удосконалено*

– структуру еволюційної ШНМ, яка відрізняється від аналогів тим, що враховує еволюцію моделі завади та процедури навчання, що дозволяє вирішувати завдання обробки інформації на новому якісному рівні у порівнянні з існуючими системами.

*Отримало подальший розвиток*

– метод гібридного навчання ЕШНМ шляхом використання для остаточного тонкого налаштування параметрів мережі рекурентної робастної процедури Левенберга-Марквардта, що дозволяє підвищити якість одержуваної моделі та стійкість процесу навчання;

– еволюційні методи усунення впливу завад при визначенні структури і параметрів нейромережових моделей шляхом використання оцінок параметрів моделі завади Тьюки-Хьюбера і процедури М-навчання, що дозволяє спростити структуру хромосоми, оскільки не потребує зберігання додаткових параметрів;

– еволюційний метод багатокритеріальної оптимізації структури та параметрів ШНМ шляхом виділення загальних для еволюційного та імунного підходів операторів і використання їх для побудови мережі, що дозволяє усунути більшість труднощів та недоліків класичних методів вирішення задачі багатокритеріальної оптимізації.

**Практичне значення отриманих результатів.** Розроблені програмні засоби, що реалізують запропоновані методи побудови еволюційних ШНМ прямого розповсюдження, які дозволяють автоматизувати процес побудови нейромережових моделей досліджуваних об'єктів, здійснити структурно-параметричний синтез в умовах апріорної і поточної невизначеності. Проведено експериментальні дослідження властивостей і характеристик розроблених методів, які підтвердили основні положення, що виносяться на захист, та показали, що запропоновані методи за рахунок використання додаткової інформації про властивості об'єкта і діючих завад дозволяють істотно скоротити час побудови моделей, а також забезпечити їх стійкість і робастність. Синтезовані в дисертації структури ШНМ, їх моделі та процедури навчання можуть бути використані при розробці систем інтелектуального аналізу даних, систем інтелектуального управління об'єктами з безперервними технологічними процесами.

**Особистий внесок здобувача.** Усі основні результати отримані автором особисто. У роботах, написаних у співавторстві, здобувачеві належать такі результати: [1] – розробка нейромережового алгоритму стиснення зображень; [2] – аналіз особливостей побудови ЕНС прямого розповсюдження, моделювання їх роботи; [3] – розробка робастного підходу до навчання вейвлет-нейронних мереж нульового та першого порядку на основі М-оцінювання; [4] – розробка і дослідження алгоритмів оцінювання параметра масштабу при робас-

тному навчанні штучних нейронних мереж в режимах on- та off-line; [5] – отримання і дослідження стійких алгоритмів навчання РБМ на основі зваженого МНК, що містить зону нечутливості; [6] – розробка методу оцінювання параметрів завад за допомогою алгоритму стохастичної апроксимації; [7] – розробка нейроконтролера на базі РБМ для керування нелінійними динамічними об'єктами з використанням кусково-лінійної апроксимації гаусівських базисних функцій; [8] – розробка методів сталого навчання радіально-базисних мереж при наявності завад вимірювання, що мають розподіл, відмінний від нормального; [9] – отримання робастного методу навчання РБМ при наявності завад вимірювання, що мають несиметричні розподіли; [10] – розробка алгоритму ідентифікації нелінійних об'єктів за допомогою еволюційної РБМ; [12] – розробка методу вибору структури еволюційної РБМ, її адаптація та навчання на основі генетичного алгоритму; [15] – визначення структури та навчання РБМ за допомогою генетичного алгоритму; [16] – отримання процедури вибору оптимальної моделі з фронту Парето за допомогою робастних інформаційних критеріїв; [17] – розробка методу боротьби з негаусівськими завадами за допомогою робастних фітнес-функцій; [18] – розробка модифікації нейронних мереж, що дозволяє використовувати додатковий вхід для нормалізації даних; [19] – розробка ефективних алгоритмів навчання мережі РБМ; [20] – використання багатокритеріального підходу до навчання еволюційних нейронних мереж прямого розповсюдження; [21] – розробка методу управління нелінійними об'єктами за допомогою еволюційних штучних нейронних мереж прямого розповсюдження; [22] – розробка алгоритму програмування з експресією генів (ПЕГ) і схем кодування синтаксичних дерев для їх подальшої лінеаризації та зберігання; [23] – прискорення процесу отримання моделі і підвищення її якості за допомогою алгоритму ПЕГ; [24] – модифікація еволюційного процесу, що використовується у ПЕГ та поліпшує властивості традиційного алгоритму; [25] – вирішення задачі нейромережевої ідентифікації нелінійного об'єкта з використанням робастних функціоналів навчання; [31] – апроксимація базисних функцій; [36] – отримання процедури робастного навчання РБМ, що забезпечує стійкість процесу навчання при наявності завад вимірювань.

**Апробація результатів дисертації.** Основні результати дисертації обговорено на: Міжнародній науково-технічній конференції "Автоматизація: проблеми, ідеї, рішення" (Севастополь, 2009 р., 2010 р., 2011 р., 2012 р., 2013 р.); на 16-й, 17-й міжнародних конференціях з автоматичного управління «Автоматика-2009», «2010-Автоматика» (Чернівці, 2009 р., Севастополь, 2010 р.); на Міжнародній науково-технічній конференції "Сучасні методи, інформаційне, програмне та технічне забезпечення систем управління організаційно-технологічними комплексами" (Київ, 2009 р.); на Міжнародній науковій конференції «Інтелектуальні системи прийняття рішень та проблеми Обчислювальна інтелекту» (Євпаторія, 2009 р.); на 1-й, 2-й Міжнародних науково-технічних конференціях «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління» (Харків-Київ, 2010 р, 2011 р.); на 1-й, 2-й Міжнародних науково-технічних конференціях "Інформаційні технології в навігації і

управлінні: стан та перспективи розвитку", (Київ, 2010 р., 2011 р.); на 1-й Міжнародній науково-технічній конференції «Обчислювальна інтелект (результати, проблеми, перспективи)» (Черкаси, 2011 р.); на 2-й, 3-й, 5-й Міжнародних науково-практичних конференціях (Смоленськ, 2012 р., 2013 р., 2015 р.); на 3-й, 4-й, 5-й, 6-й Міжнародних науково-технічних конференціях «Сучасні напрямки розвитку інформаційно-комунікаційних технологій і засобів управління» (Полтава-Белгород-Харків-Київ-Кіровоград, 2013 р., Полтава-Баку-Белгород-Кіровоград-Харків, 2014 р., Полтава-Баку-Харків-Кіровоград, 2015 р., 2016 р.); на IX Міжнародній конференції «Стратегія якості в промисловості та освіті» (Варна (Болгарія), 2013 р.); на 3-й Міжнародній науково-практичній конференції «Стратегічні рішення інформаційного розвитку економіки, суспільства та бізнесу», (Рівне, 2014 г.), на 3-ій Міжнародній науково-технічній конференції «Проблеми ін-форматизації», (Черкаси-Баку-Бельсько-Бяла-Полтава, 2015 р.); на IV Міжнародній науково-практичній конференції «Суспільство і економічна думка в ХХІ ст.: шляхи розвитку та інновації» (Воронеж, 2016 р.); на VIII Міжнародній науково-практичній конференції «Проблеми та перспективи розвитку ІТ-індустрії» (Харків, 2016 р.).

**Публікації.** Основні положення і результати дисертаційної роботи досить повно відображені у 56 опублікованих роботах (2 колективні монографії; 25 статей у виданнях, що входять до переліків наукових фахових періодичних видань України з технічних наук, серед яких 6 одноосібних, 10 індексуються у наукометричній базі SCOPUS; 28 – тези доповідей на міжнародних конференціях і форумах).

**Структура та обсяг роботи.** Дисертація складається зі вступу, шести розділів, висновків, списку використаних джерел, що містить 324 найменування на 35 сторінках, та 11 додатків на 50 сторінках. Робота містить 126 ілюстрацій, з них 83 на 44 окремих сторінках, та 16 таблиць, з них 10 на 6 окремих сторінках. Загальний обсяг роботи складає 436 сторінок, у тому числі 295 сторінок основного тексту.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність дисертаційної роботи, сформульовано основну мету і завдання досліджень, наведено відомості щодо зв'язку дисертації з планами організації, де виконана робота. Надано стислу анотацію отриманих у дисертації результатів, відзначено їх практичну цінність, наведено дані щодо використання результатів проведених досліджень.

У **першому розділі** проаналізовано стан проблеми багатовимірної оптимізації, яка включає в себе NP-повні задачі комбінаторної оптимізації, ідентифікації складних структур або багатовимірної оптимізації функцій.

Проведений аналіз засвідчив, що застосування традиційних методів дослідження, таких як градієнтні, динамічне програмування, симплекс-метод, для цих видів задач часто не дає бажаного результату, тому що обчислювальні витрати зростають експоненційно разом із розмірністю задачі. У зв'язку з цим



пропонується для вирішення практичних задач застосовувати евристичні методи, які є більш гнучкими і ефективними і потребують набагато менших обчислювальних витрат, незважаючи на те, що вони можуть і не забезпечувати досягнення глобального оптимального рішення. До таких методів оптимізації відносять метод імітації відпалу, ШНМ, еволюційні алгоритми (ЕА), мурашині алгоритми тощо. Аналіз існуючих інтелектуальних систем обробки інформації показав, що досить перспективним видається використання ШНМ і ЕА, а також їх комбінації в ЕШНМ. Зазначено, що будучи універсальними апроксиматорами, деякі типи ШНМ дозволяють відновити із заданою точністю будь-яку як загодно складну безперервну нелінійну функцію. Найбільшого поширення при вирішенні такої задачі отримали статичні ШНМ прямого розповсюдження (багатошаровий перцептрон (БП), радіально-базисна мережа (РБМ), узагальнено-регресійна мережа (УРМ), вейвлет-мережі).

Слід відзначити, що при синтезі або використанні ШНМ виникають задачі структурної та параметричної оптимізації, відповідні вибору оптимальної топології мережі та її навчання (налаштування параметрів). Задача визначення структури мережі є важливою і досить складною. Для її вирішення можуть використовуватися, наприклад, крос-валідація, статистична перевірка гіпотез, а також отримали останнім часом значного поширення інформаційні критерії.

Якщо задача визначення структури є дискретною оптимізаційною (комбінаторною), то пошук оптимальних параметрів здійснюється в безперервному просторі за допомогою класичних методів оптимізації, заснованих на МНК. Для навчання (оцінювання параметрів) мережі застосовуються, як правило, методи, що потребують обчислення градієнта функціоналу, що використовується (алгоритм зворотного поширення помилки (ЗП), метод сполучених градієнтів, алгоритм Гауса-Ньютона, Левенберга-Марквардта тощо). Незважаючи на популярність цих методів не тільки при навчанні ШНМ, але і при вирішенні інших задач оптимізації, вони мають ряд суттєвих недоліків, наприклад: отримане рішення залежить від форми функціоналу, що мінімізується, вектора початкових умов; вони застрягають в локальних екстремумах, а при наявності декількох екстремумів не забезпечують знаходження глобального; їх застосування потребує диференційованих функціоналів, що мінімізуються. Крім того, складним є визначення оптимальних параметрів алгоритмів, що забезпечують їх максимальну швидкість збіжності, точність, робастність тощо.

Зазначено, що класичні робастні методи орієнтовані на симетричність засмічення, коли викиди однаково часто з'являються як в області негативних, так і в області позитивних значень. У більш загальній ситуації довільного виду засмічення, наприклад, коли гаусівське засмічення має нульове математичне сподівання, або коли засмічення є несиметричним, оцінки, що даються цими методами, будуть зміщеними. Це викликає потребу в розробці робастних методів навчання, що забезпечують незміщеність оцінок при наявності завад, що мають асиметричний розподіл.

Спроби усунути недоліки традиційних методів синтезу та функціонування ШНМ привели до появи нового класу мереж – ЕШНМ, в яких на додаток до

традиційного навчання використовується інша фундаментальна форма адаптації – еволюція, що реалізується шляхом застосування еволюційних обчислень.

Використання в ЕШНМ двох форм адаптації – еволюції та навчання, що дозволяють змінювати структуру мережі, її параметри та алгоритми навчання без зовнішнього втручання, робить такі мережі найбільш пристосованими для роботи в нестационарних умовах і за наявності невизначеності щодо властивостей досліджуваного об'єкта та умов його функціонування.

Основною перевагою використання ЕА як алгоритмів навчання є те, що багато параметрів ШНМ можуть бути закодовані в геномі та визначатися паралельно. Більш того, на відміну від більшості алгоритмів оптимізації, призначених для покрокового вирішення задачі, ЕА оперують з безліччю рішень – популяцією, що дозволяє досягти глобального екстремуму, але не застряє в локальних. При цьому інформація про кожну особину популяції кодується в хромосомі (генотипі), а отримання рішення (фенотипу) здійснюється після еволюції (відбору, схрещування, мутації) шляхом декодування.

Для запобігання ефекту перенавчання моделі необхідно контролювати її складність. Оскільки отримувана нейромережева модель, з одного боку, має бути досить простою та зручною для використання її в прикладних задачах, а з іншого – найбільш повно відображати властивості досліджуваного об'єкта, її якість визначається деяким набором критеріїв, тобто задача побудови нейромоделі є багатокритеріальною. Таким чином, виникає задача розвитку еволюційного методу багатокритеріальної оптимізації (БО) та його використання для створення Парето-еволюційних ШНМ прямого розповсюдження.

Для забезпечення можливості практичного застосування методів побудови ЕШНМ для створення інтелектуальних систем обробки інформації необхідно розробити відповідні програмні засоби, а для ефективного застосування методів і програмних засобів побудови ЕШНМ необхідно провести експериментальне дослідження їх властивостей і характеристик, а також розробити рекомендації щодо їх застосування при вирішенні практичних задач.

На основі проведеного аналізу сформульовано основні задачі наукового дослідження, вирішенню яких присвячено наступні розділи.

**У другому розділі** зазначено, що основою побудови і функціонування ЕШНМ є генетичні алгоритми (ГА), які абстрагують фундаментальні процеси дарвінівської еволюції: природного відбору та генетичних змін внаслідок рекомбінації і мутації. Однак, крім дарвінівських, ГА можуть реалізовувати механізми еволюції Ламарка, що змінюють (поліпшують) хромосоми, та Болдуїна, що покращують пристосованість хромосоми без її зміни. На відміну від дарвінізму, де еволюція є результатом конкуренції та відбору, в ламаркізмі самі організми контролюють еволюцію. Це досягається за допомогою практики, навчання та частого використання спеціальних органів. Хоча ламаркізм в природі не зустрічається, його ідеї знайшли застосування при побудові ЕШНМ в алгоритмах, що містять механізм збереження придбаних характеристик.

Аналіз робіт по застосуванню ефекту Болдуїна в ЕШНМ показав, що навчання Болдуїна можна розглядати як свого роду мінливість, яка спрямовується

фенотипом, що дозволяє збільшити дисперсію процесу відбору та робить рельєф фітнес-функції навколо оптимальних областей більш плоским, що забезпечує поліпшення процесу еволюції. Загальний ефект від навчання Болдуїна полягає в тому, що воно дозволяє знайти глобальний оптимум. Однак в деяких випадках ефект Болдуїна може призводити до неефективних гібридних алгоритмів, у той час як використання еволюції Ламарка є досить ефективним в нестационарних умовах функціонування мережі.

Аналіз робіт з побудови ЕШНМ показав, що еволюція в ЕШНМ може стосуватися архітектури мережі, її вагових параметрів, виду та параметрів базисних (активаційних) функцій (БФ), алгоритму навчання (рис.1).



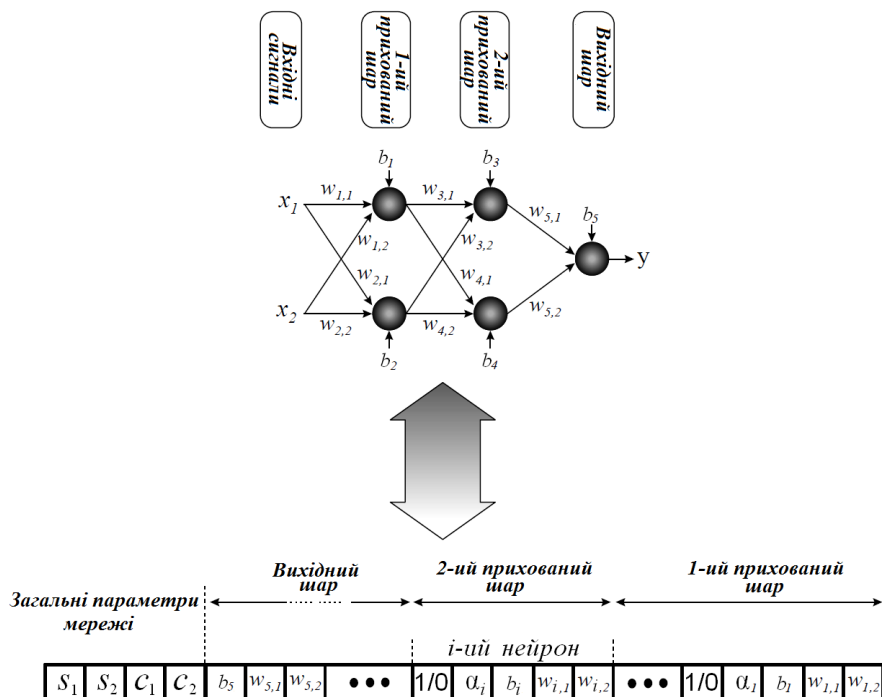
Рисунок 1 – Структура ЕШНМ

У зв'язку з цим отримала подальший розвиток архітектура еволюційної ШНМ прямого розповсюдження, яка, крім еволюції структури мережі та її параметрів, враховує також еволюцію алгоритму навчання та завади. Така ШНМ має істотно поліпшені апроксимуючі та екстраполюючі властивості, що дає можливість за рахунок адаптації структури, параметрів і процедури навчання ефективно обробляти інформацію в нестационарних умовах і за наявності невизначеності.

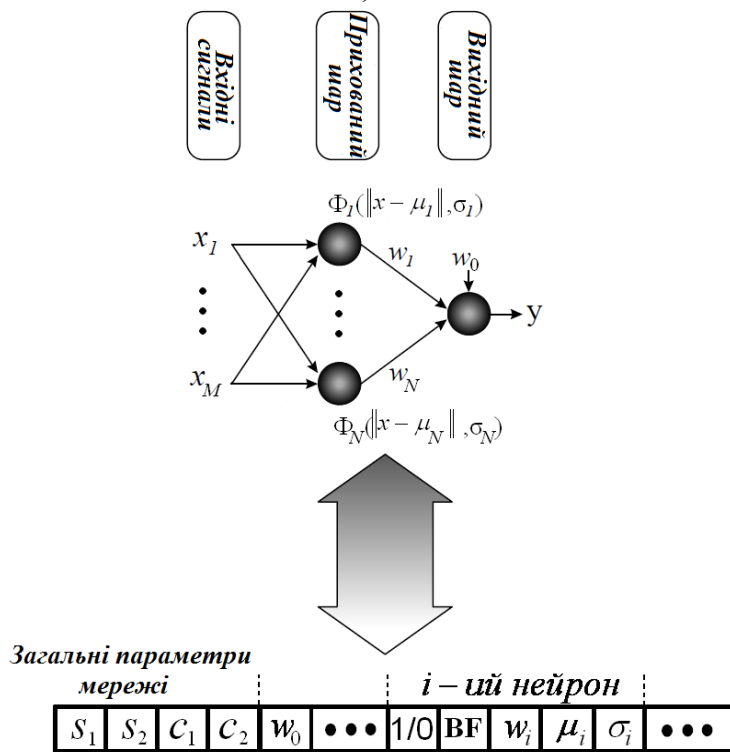
При переході від ШНМ до ЕШНМ для всіх типів мереж використовуються загальні еволюційні процедури (ініціалізація популяції, оцінка популяції, селекція, схрещування, мутації), а відмінності полягають лише в способі кодування структури та параметрів тієї або іншої ШНМ у вигляді хромосоми. У роботі запропоновано різні формати хромосом для БП та РБМ, які представлено

на рисунку 2, де використані наступні позначення:  $w_i$ ,  $\mu_i$ ,  $\sigma_i$  – вагові параметри нейронів, центри та радіуси базисних функцій відповідно;  $b_i$  – зміщення  $i$ -го нейрону;  $s_1$ ,  $s_2$ ,  $c_1$  та  $c_2$  – оцінки параметрів завади.

Отримано формули для обчислення довжини хромосоми, зокрема змінної, при кодуванні структур БП та РБМ.



а)



б)

Рисунок 2 – Формати хромосом: а) БП; б) РБМ

Внаслідок використання хромосом змінної довжини виникають особини з особливими генетичними сегментами коду (інтронами), які не беруть участь у кодуванні характеристик, але є дуже важливими для збереження цілісності хромосоми та вирівнюванні її довжини при використанні операторів схрещування.

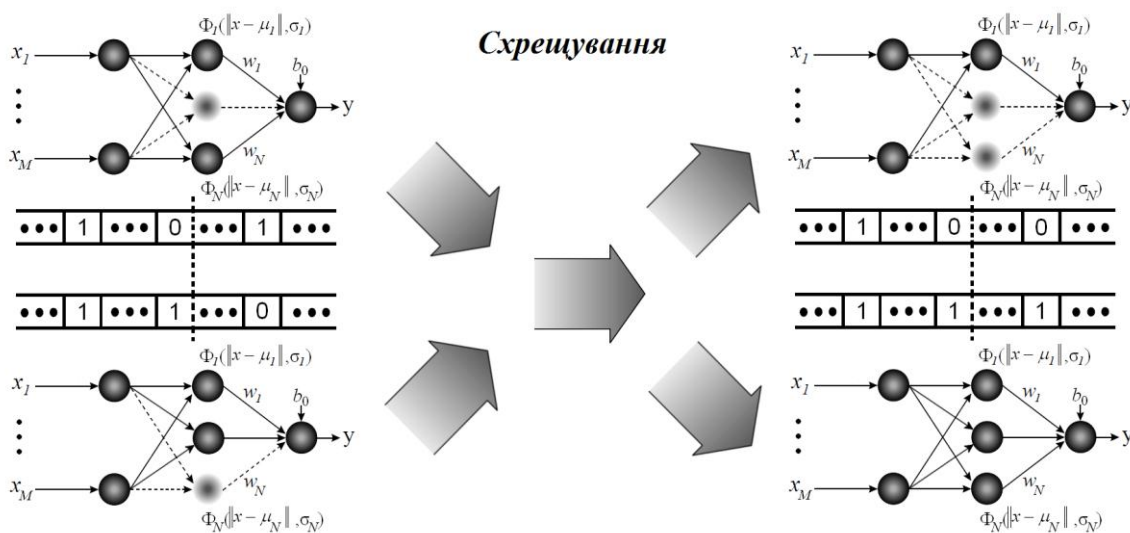
При використанні ГА для побудови ЕШНМ основними проблемами є вибір методу кодування можливого рішення та генетичних операторів; у різних задачах використовуються різні методи кодування та генетичні оператори, основні з яких розглянуто в розділі. У роботі проведено порівняльний аналіз генетичних операторів та розроблено рекомендації щодо їх застосування.

Властивості оператора селекції істотно впливають на середній час отримання рішень заданої якості та на розкид результатів при незалежних реалізаціях еволюційного процесу в ГА. У ряді випадків оператор пропорційної селекції може бути замінений іншими операторами. При цьому вибір кожної батьківської особини не обов'язково здійснюється незалежно від вибору інших особин на цій ітерації.

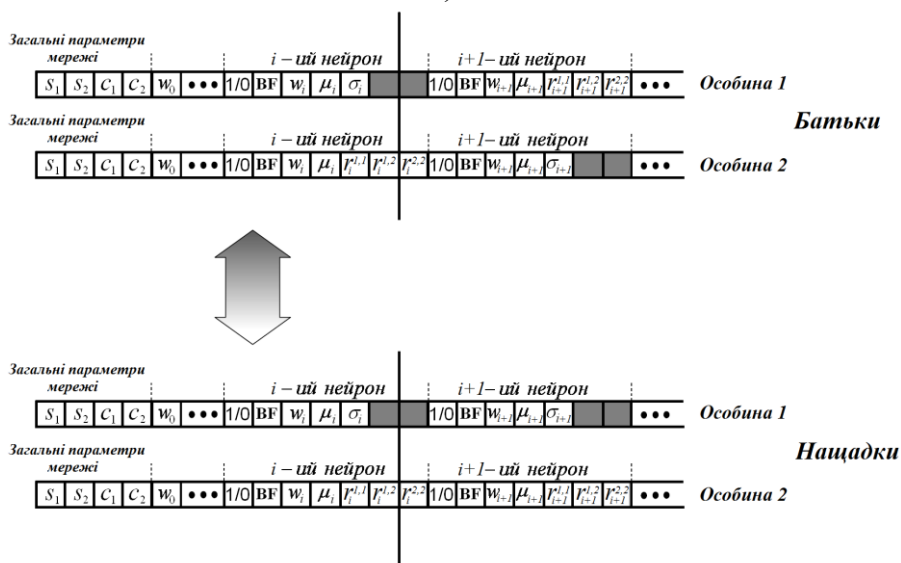
Після того, як батьківські особини було відібрано методом селекції, здійснюється їх схрещування (кросовер). Кросовер застосовується для відтворення потомства та полягає в обміні генетичною інформацією між батьківськими особинами. Було розглянуто такі оператори схрещування як однокрапковий кросовер (ОРХ), рівномірний кросовер (УХ) та його модифікації, блоковий кросовер (ВХ), кросовер із частковим відображенням (РМХ), кросовер зі збереженням дистанції (DРХ), циклічний кросовер (СХ), порядковий кросовер (ОВХ) тощо.

Приклад операції схрещування двох РБМ за допомогою оператора ОРХ показаний на рис. 3а, б. З рисунків видно, що при схрещуванні відбувається не тільки обмін інформацією про параметри мережі, а й утворення мережі з новою для популяції структурою. Також з рисунків видно, що наявність інтронів дозволяє захистити цілісність хромосом нащадків і зберегти їх логічний формат і довжину.

Після застосування оператора кросовера хромосоми піддаються мутації, яка запобігає застряганню алгоритму навчання у локальному мінімумі та дозволяє створити новий генетичний матеріал в популяції для підтримки її різноманітності. Мутація – це не що інше, як виконання довільної зміни частини хромосоми, що представляє окрему особину. Кількість мутацій в популяції регулюється параметром  $p_m$ , який визначає імовірність мутації. Зазвичай імовірність вибирається рівною  $1/L$ , де  $L$  – довжина хромосоми. Таким чином, тільки  $p_m \times L$  випадкових хромосом у популяції можуть мутувати. Імовірність мутації також може бути прив'язана до значення фітнес-функції, тобто чим гірше значення фітнес-функції, тим вище ймовірність мутації. Приклад оператора мутації в хромосомі РБМ показаний на рис.4. У роботі було ретельно досліджено ряд операторів мутацій (досліджено інверсну мутацію, мутацію обміну, реверсну мутацію, мутацію вставкою, адаптивну мутацію тощо) та стратегій їх застосування.



а)



б)

Рисунок 3 – Приклади оператора однокрапкового схрещування: а) вплив на структуру мережі; б) вплив інтронів на збереження цілісності хромосоми

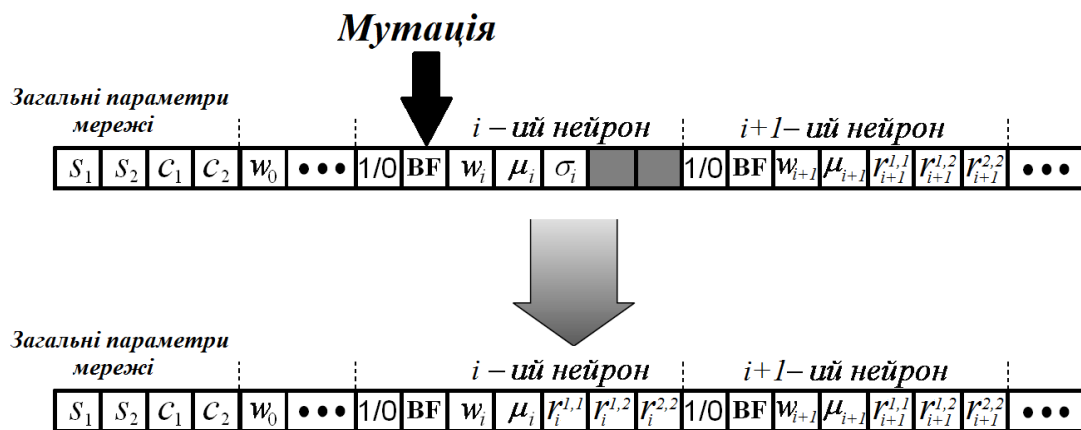


Рисунок 4 – Приклад оператора мутації

Після закінчення формування нового покоління здійснюється його оцінка. У разі, якщо виконується критерій зупинки ГА, починає роботу градієнтний алгоритм корекції вагових параметрів, який здійснює «тонке» налаштування найкращої мережі, відібраної за допомогою ГА. Як такий алгоритм може бути використана модифікація алгоритму зворотного поширення, алгоритм Гауса-Ньютона, Левенберга-Марквардта тощо.

**Третій розділ** присвячено аналізу методів навчання ЕШНМ, який показав, що на сьогодні існує велика кількість методів налаштування параметрів мережі, що відрізняються обсягом інформації, яка використовується, що впливає як на динамічні властивості алгоритмів, так і на їх обчислювальну складність. Якість навчання залежить від того, наскільки вдало було обрано функціонал помилки. Одним з основних факторів, що впливають на ефективність вибору функціоналу, є наявність інформації про властивості завади. У сьогоднішній час існує два принципово різних підходи, заснованих на тому, що:

- існує інформація або про деякі статистичні властивості завади або про приналежність її розподілу до деякого відомого класу;
- незважаючи на природу завади, вона передбачається обмеженою.

Розглянуто найбільш прості в обчислювальному відношенні однокрокові градієнтні процедури навчання, які мінімізують квадратичний функціонал і включають процедури Качмажа (Уїдроу-Хоффа), Нагумо-Ноди та Гудвіна-Ремеджа-Кейнеса, який поєднує властивості нормалізованого алгоритму Качмажа та алгоритму стохастичною апроксимації.

Розглянуто особливості застосування цих процедур та їх регуляризованих модифікацій для навчання лінійної та нелінійної Адаліни та вивчено питання їх збіжності, які зводяться до аналізу функцій Ляпунова. При навчанні багатошарового персептрону процедури корекції матриць ваг вихідного та схованого шарів ( $V$  та  $W$  відповідно) є матричними варіантами процедури Качмажа, збіжність яких досліджується також за допомогою функцій Ляпунова.

У випадку наявності інформації щодо обмеженості завад, тобто  $|\xi(k)| < \Delta$ , для навчання слід використовувати процедури, що містять зону нечутливості, яка визначається за допомогою таких виразів:

$$g(e(k), \Delta) = \begin{cases} e(k), & \text{якщо } |e(k)| \geq \Delta; \\ 0, & \text{якщо } |e(k)| < \Delta \end{cases} \quad (1)$$

або

$$g(e(k), \Delta) = \begin{cases} e(k) - \Delta(k) \operatorname{sign} e(k), & \text{якщо } |e(k)| \geq \Delta; \\ 0, & \text{якщо } |e(k)| < \Delta. \end{cases} \quad (2)$$

Введення зони нечутливості, огрублюючи процедуру навчання, забезпечує їх роботу при наявності обмежених завад. Вивчено модифіковані процедури Качмажа, які мають зону нечутливості, зокрема розглянута процедура навчання багатошарового персептрона, яка має вигляд

$$V(k) = V(k-1) + \gamma_v(k) \frac{\alpha_v(k) F(x, w(k-1)) e(k)}{\beta_v(k) + \|F(x, w(k-1))\|^2}, \quad (3)$$

$$W(k) = W(k-1) + \gamma_w(k) \frac{\alpha_w(k) \nabla f(x(k), V(k-1)) e(k) x^T(k)}{\beta_w(k) + \|\nabla f(x, V(k-1)) x(k)\|^2}, \quad (4)$$

$$\text{де } \alpha_v(k) = \begin{cases} g(e(k), \Delta_v(k)), & \text{якщо } \|e(k)\| > \Delta_v(k); \\ 0, & \text{якщо } \|e(k)\| \leq \Delta_v(k); \end{cases} \quad (5)$$

$$\alpha_w(k) = \begin{cases} g(e(k), \Delta_w(k)), & \text{якщо } \frac{\nabla f_{\min}}{\alpha} \|e(k)\| > \Delta_w(k); \\ 0, & \text{якщо } \frac{\nabla f_{\min}}{\alpha} \|e(k)\| \leq \Delta_w(k); \end{cases} \quad (6)$$

$$\nabla f_{\min} = \min[\nabla f_1(k), \nabla f_2(k), \dots, \nabla f_L(k)] > 0;$$

$\alpha$  – коефіцієнт, що визначає нахилення сігмоїди.

Підвищення обчислювальної стійкості процедур пов'язане з використанням в них параметра регуляризації, щодо ефективного вибору якого не існує рекомендацій. У зв'язку з цим розглянуто питання вибору цього параметра шляхом апроксимації регуляризованого квадратичного функціоналу з подальшим застосуванням чисельного методу або шляхом використання співвідношення сигнал-завада (критерію ENR, echo-to-noise ratio).

Результати аналізу методів навчання, що мають зону нечутливості, свідчать про те, що для монотонної збіжності процедур величина цієї зони має задовольняти відповідним умовам на кожному такті процесу навчання. Це обумовило розробку методів адаптивної корекції зони нечутливості, яка здійснюється з надходженням нової інформації.

Незважаючи на те, що однокрокові процедури є привабливими з обчислювальної точки зору, більш ефективними з точки зору навчання в нестаціонарних умовах є багатокрокові процедури, побудовані на рекурентному МНК (РМНК), який в загальному випадку має вигляд

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \gamma(k) P(k) \psi(k-1) (y(k) - \hat{\theta}^T(k-1) \psi(k-1)); \quad (7)$$

$$P^{-1}(k) = P^{-1}(k-1) + \gamma(k) \psi(k-1) \psi^T(k-1), \quad (8)$$

де  $\gamma(k) > 0$  – деякий параметр.

Обираючи різні значення параметру  $\gamma(k)$ , можна отримати різні модифікації РМНК, зокрема, при  $\gamma(k) = 1$  маємо традиційний РМНК.

Розглянуто деякі модифікації РМНК, що містять зону нечутливості, зокрема, процедури Фогеля-Хуанга, Лозано-Ліла-Ортеги та Канудаса де Віта-Карільо.

Крім того, було досліджено модифікацію РМНК, яка має вигляд

$$\hat{\theta}(k) = \begin{cases} \hat{\theta}(k-1) + \frac{\alpha P(k-1) g(e(k), \beta \delta)}{1 + (1 + \alpha g(e(k), \beta \delta)) \psi^T(k-1) P(k-1) \psi(k-1)}, & \text{якщо } |e(k)| > \beta \delta; \\ \hat{\theta}(k-1), & \text{в протилежному випадку;} \end{cases} \quad (9)$$



$$P(k) = \begin{cases} P(k-1) - \alpha g(e(k), \beta\delta) \frac{P(k-1)\psi(k-1)\psi^T(k-1)P(k-1)}{1 + (1 + \alpha g(e(k), \beta\delta))\psi^T(k-1)P(k-1)\psi(k-1)}, & \text{якщо } |e(k)| > \beta\delta; \\ P(k-1), & \text{в протилежному випадку.} \end{cases}$$

де

$$\alpha \in (0, 1]; \beta = \sqrt{1 + \alpha};$$

$$g(e(k), \beta\delta) = \begin{cases} \frac{f(e(k), \beta\delta)}{e(k)}, & \text{якщо } |e(k)| > \beta\delta; \\ 0, & \text{якщо } |e(k)| \leq \beta\delta. \end{cases} \quad (10)$$

Розглянуто проєкційні методи навчання, що займають проміжне положення між однокроковими процедурами та РМНК, мають кінцеву пам'ять та використовують ковзне вікно. Отримано рекурентні форми цих процедур, які відрізняються методом формування матриці спостережень.

Вивчено рекурентні процедури Гауса-Ньютона та Левенберга-Марквардта та запропоновано їх модифікації, що містять зону нечутливості.

Так алгоритм навчання Гауса-Ньютона із зоною нечутливості має вигляд

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{\alpha P(k-1) \nabla \hat{f}(k) e(k, \theta)}{1 + \nabla^T \hat{f}(k) P(k-1) \nabla \hat{f}(k)}, \quad (11)$$

$$P(k) = P(k-1) - \frac{\alpha P(k-1) \nabla \hat{f}(k) \nabla^T \hat{f}(k) P(k-1)}{1 + \nabla^T \hat{f}(k) P(k-1) \nabla \hat{f}(k)}, \quad (12)$$

$$\alpha = \begin{cases} 1, & \text{якщо } |e(k)| > \Delta; \\ 0, & \text{якщо } |e(k)| \leq \Delta, \end{cases} \quad (13)$$

$$\text{де } \nabla f(l, \hat{\theta}_i) = \frac{\partial f(l, \theta)}{\partial \theta}.$$

Рекурентну ж процедуру Левенберга-Марквардта із зоною нечутливості можна записати у такий спосіб:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \alpha P(k) \nabla \hat{f}(k, \theta) e(k, \theta), \quad (14)$$

$$P(k) = \frac{1}{\lambda} \left[ P(k-1) - \alpha P(k-1) \nabla f^*(k, \theta) K^{-1}(k) \nabla^T f^*(k, \theta) P(k-1) \right], \quad (15)$$

де  $\alpha$  обчислюється відповідно з (17), а  $K(k)$  – за допомогою виразу

$$K(k) = \nabla^T f^*(k, \theta) P(k-1) \nabla f^*(k, \theta) + \lambda \Lambda^*(k). \quad (16)$$

**У четвертому розділі** розглянуто робастні методи навчання ЕШНМ. У зв'язку з тим, що методи навчання, засновані на мінімізації квадратичних функціоналів, не є стійкими, якщо у вимірах присутні викиди або завади мають розподіл, відмінний від гаусівських, як альтернативу для забезпечення робастності цільову функцію модифікують таким способом, щоб обмежити вплив найбільших вимірів. Основним наслідком цього є, як правило, більш низька швидкість збіжності алгоритмів оптимізації.

Розглянуто деякі типи класів розподілів, що зустрічаються при вирішенні практичних задач:  $P_1$  – клас невивіржених розподілів,  $P_2$  – клас розподілів з

обмеженою дисперсією,  $P_5$  – клас фінітних розподілів (завада обмежена по абсолютній величині, а будь-які відомості про щільність її розподілу відсутні),  $P_3$ ,  $P_4$  та  $P_6$  – класи наближено нормальних, наближено рівномірних і наближено фінітного розподілів відповідно, що описуються моделлю Тьюки-Хьюбера

$$p(x) = (1 - \varepsilon)p_0(x) + \varepsilon q(x), \quad (17)$$

де  $p_0(x)$  – щільність відповідного основного розподілу;  $q(x)$  – щільність розподілу, що засмічує (довільного);  $\varepsilon \in [0,1]$  – параметр, що характеризує ступінь засмічення основного розподілу.

Якщо інформація про приналежність завади деякому певному класу розподілів відома, то шляхом мінімізації оптимального критерію, який представляє собою узятий з оберненим знаком логарифм функції розподілу завади, може бути отримана оцінка максимальної правдоподібності (М-оцінка).

Якщо ж такої інформації немає, то для оцінювання шуканого вектора параметрів слід застосувати будь-який неквадратичний критерій, що забезпечує робастність одержуваної оцінки.

Через те, що модульний критерій дозволяє отримати оцінку, менш чутливу до хвостів розподілу завади, ніж МНК-оцінка, у роботі досліджено такі різновиди модульного критерію, як функціонал А. Форсайта та змішаний функціонал, що використовує одночасно зважені квадратичну та модульну норми помилки.

Проведено аналіз однокрокових робастних процедур навчання. Крім того, було досліджено процедуру

$$\theta(k) = \theta(k-1) + \gamma(k)[\lambda 2e(k) + (1 - \lambda)\text{sign } e(k)]x(k), \quad (18)$$

де  $\lambda \in [0,1]$ , яка мінімізує комбінований функціонал, поєднує властивості МНК і МНМ, та отримано умови її збіжності в середньому та середньоквадратичному.

Розглянуто властивості робастних багатокрокових проекційних процедур навчання.

Запропоновано нову процедуру М-навчання нейромереж, в якій для вибору структури мережі було застосовано робастний інформаційний критерій, для оцінки параметрів – досить відомий робастний алгоритм Гауса-Ньютона або запропонована в роботі робастна процедура Левенберга-Марквардта, яка має вигляд

$$\hat{\theta}(k+1) = \hat{\theta}(k) + P(k)\nabla\hat{f}(k)\psi(e(k)); \quad (19)$$

$$P(k) = P(k-1) - P(k-1)\nabla\hat{f}^*(k)K^{-1}(k)\nabla^T\hat{f}^*(k)P(k-1)\psi'(k), \quad (20)$$

де

$$K(k) = \psi'(k)\nabla^T\hat{f}^*(k)P(k-1)\nabla\hat{f}^*(k) + \Lambda^*(k);$$

$$\Lambda^{*-1}(k) = \begin{pmatrix} 1 & 0 \\ 0 & \delta N \end{pmatrix}; \quad \psi(k) = \frac{\partial\rho(e(k))}{\partial e(k)} - \text{функція впливу};$$

$\rho(e(i))$  – деяка функція втрат, що залежить від закону розподілу завад;

$$\nabla \hat{f}^*(k) = \begin{pmatrix} \nabla^T \hat{f}^*(k) \\ 0 \dots 0 \quad 1 \dots 0 \end{pmatrix}^T.$$

↑

позиція =  $(k \bmod N + 1)$

Слід зазначити, що другі похідні деяких функціоналів за певних значень аргументів можуть бути від'ємними, що призводить до нестійкої роботи процедур навчання. Якщо ж замість  $\psi'(k)$  в цих процедурах використовувати вагову функцію  $\omega(k) = \psi'(e(k)) / e(k)$ , ця проблема зникає.

Класичні робастні методи орієнтовані на симетричність засмічення, коли викиди однаково часто з'являються як в області негативних, так і в області позитивних значень. У більш загальній ситуації довільного виду засмічення, наприклад, коли гаусівський розподіл, що засмічує, має ненульове математичне сподівання або коли розподіл, що засмічує, є несиметричним, оцінки, що даються цими методами, будуть зміщеними.

Необхідність урахування асиметрії розподілів обумовлює доцільність вибору асиметричних функціоналів навчання. Принципово можуть бути отримані асиметричні функціонали (АФ) різних видів. При цьому базою для них служать відповідні традиційні симетричні функціонали робастного М-навчання. Наявність інформації про вид асиметричного розподілу даних та завод є основою для вибору параметрів функціоналів, здійснюючи значний вплив на вигляд АФ.

У роботі отримано більш загальний вид асиметричних функціоналів, їх перших, других похідних та вагових функцій. Слід зазначити, що параметри, які входять до виразів як симетричних, так і асиметричних функціоналів, впливають на їх вигляд та залежать від властивостей завод. Необхідно мати на увазі наступне. У випадку нормального розподілу оптимальні оцінки, одержувані за допомогою методу найменших квадратів, інваріантні до параметру масштабу  $\sigma_\xi^2$ . При М-оцінюванні, внаслідок неоднорідності функціоналу, одержувані оцінки не зберігають властивості інваріантності. Для того, щоб властивість інваріантності масштабу виконувалася, у функціоналах замість помилки  $e(i, \theta)$  слід брати  $\tilde{e}(i, \theta) = (e(i, \theta) - m) / S$ , де  $S$  – заводостійка оцінка параметра масштабу або міра розсіювання залишкових різниць (у разі нормального розподілу  $S$  є оцінкою  $\sigma_\xi$ );  $m$  – математичне сподівання засмічуючої заводи  $q(x)$  в моделі (17), яке в загальному випадку є відмінним від нуля.

При нульовому математичному сподіванні заводи оцінка параметра масштабу  $S$  використовуються як константа, що входить в розглянуті функціонали та функції впливу.

Розроблено нові робастні методи навчання ЕШНМ, які дають можливість ефективно обробляти інформацію при наявності завод з асиметричними розподілами. Ці методи використовують додаткові процедури оцінювання параметрів функціоналів і завод, що дозволяє здійснювати корекцію одержуваних оцінок і усувати їх зміщення. Приклади асиметричності модифікованих функціоналів

Хьюбера, Хемпеля тощо наведено на рис. 5.

Розроблено нові процедури корекції параметрів, використовуваних при навчанні функціоналів і оцінюванні параметрів завади, що описується моделлю Тьюки-Хьюбера. Це дозволяє при відсутності апріорної інформації про статистичні властивості завод адаптивно коригувати параметри ЕШНМ.

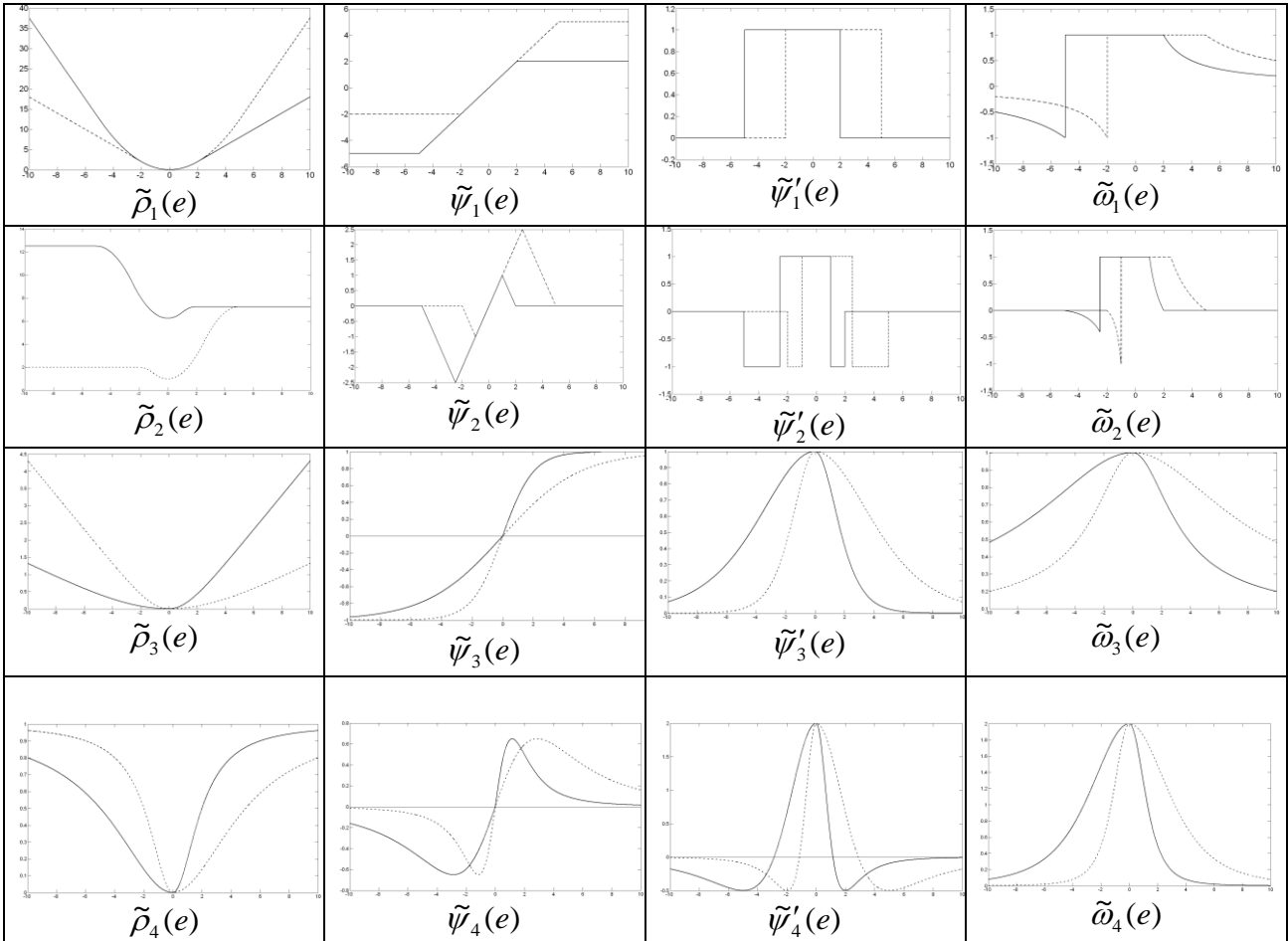


Рисунок 5 – Приклади асиметричності функціоналів

Отримав подальший розвиток еволюційний метод усунення впливу завод при визначенні структур та параметрів нейромережкових моделей за рахунок використання оцінок параметрів моделі завади Тьюки-Хьюбера в процедурі М-навчання. Такий підхід дозволяє поліпшити фільтруючі властивості використовуваних процедур, спростити структуру хромосоми та скоротити процес отримання моделі ЕШНМ.

Якщо розподіл завади асиметричний, то нульове математичне сподівання обумовлює зміщення оцінок, яке може бути усунуто лише при наявності інформації про вид розподілу. Якщо ж такої інформації немає, можна використовувати деяку апроксимацію розподілу, наприклад, моделлю (17) з  $m_1 \neq 0$  та  $m_2 \neq 0$ . При цьому слід мати на увазі, що основною метою апроксимації є не адекватне відображення властивостей завади, а компенсація можливого зміщення оцінок.

У цьому випадку оцінювання цього параметра можна здійснити за допо-

могою наступного алгоритму стохастичної апроксимації:

$$\hat{m}_1(k) = \begin{cases} \hat{m}_1(k-1), & \text{якщо } |e(k)| > 3S_1(k-1); \\ \hat{m}_1(k-1) + \frac{1}{k^\lambda} (y(k) - \hat{m}_1(k-1)) & \text{в протилежному випадку;} \end{cases} \quad (21)$$

$$\hat{m}_2(k) = \begin{cases} \hat{m}_2(k-1), & \text{якщо } |e(k)| \leq 3S_1(k-1); \\ \hat{m}_2(k-1) + \frac{1}{k^\lambda} (y(k) - \hat{m}_2(k-1)) & \text{в протилежному випадку,} \end{cases}$$

де  $\lambda \in (0,1]$  – деякий коефіцієнт.

При використанні моделі засмічення (17) необхідно оцінювати величини  $\sigma_1^2$  та  $\sigma_2^2$ , а також врахувати ці оцінки в процедурі навчання. Якщо  $\sigma_1^2$  та  $\sigma_2^2$  не змінюються в часі, то таке оцінювання також може бути здійснене методом стохастичної апроксимації, коли оцінки дисперсій  $\sigma_1^2$  та  $\sigma_2^2$  ( $S_1^2$  та  $S_2^2$  відповідно) обчислюються наступним чином:

$$S_1^2(k) = \begin{cases} S_1^2(k-1) + \frac{1}{l_1(k)} (\tilde{e}^2(k) - S_1^2(k-1)), & \text{якщо } |\tilde{e}(k)| \leq 3S_1(k-1); \\ S_1^2(k-1) & \text{в протилежному випадку;} \end{cases} \quad (22)$$

$$S_2^2(k) = \begin{cases} S_2^2(k-1) + \frac{1}{l_2(k)} (\tilde{e}^2(k) - S_2^2(k-1)), & \text{якщо } |\tilde{e}(k)| > 3S_1(k-1); \\ S_2^2(k-1) & \text{в протилежному випадку,} \end{cases}$$

де

$$l_1(k) = k - l_2(k);$$

$$l_2(k) = \begin{cases} 0, & \text{якщо } |\tilde{e}(k)| \leq 3S_1(k-1); \\ l_2(k-1) + 1 & \text{в протилежному випадку.} \end{cases} \quad (23)$$

Загальна дисперсія завади обчислюється за формулою

$$S^2(k) = \begin{cases} S_1^2(k), & \text{якщо } |\tilde{e}(k)| \leq 3S_1(k-1); \\ S_2^2(k) & \text{в протилежному випадку} \end{cases} \quad (24)$$

та може бути використана як параметр зважування квадратичного функціонала, тобто  $\rho(\tilde{e}(k), S) = 0.5e^2(k)S^{-2}(k)$ . У цьому випадку оцінка, що отримана шляхом його мінімізації, буде більш стійкою.

Результати імітаційного моделювання розроблених моделей і процедур навчання підтвердили їх працездатність та ефективність.

Зокрема розглядалася задача апроксимації при наявності завод з різними розподілами в вихідних сигналах. Функція, що апроксимувалася, описувалася таким рівнянням:

$$y = 0.5 + \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{1 + 0.001(x_1^2 + x_2^2)} + \xi(k), \quad (25)$$

де  $x_1, x_2$  – стаціонарні випадкові послідовності з рівномірним законом розподілу в інтервалі  $[-10, 10]$ , що генеруються датчиком випадкових чисел;  $\xi$  – завада вимірювань.

Популяція складалася із 128 особин (РБМ), максимально допустима кількість нейронів у кожній мережі була обмежена 100 нейронами. На рис. 6а показана поверхня, що описується рівнянням (25), а на рис. 6б приведена поверхня, відновлена за допомогою ЕРБМ, що використовує асиметричну модифікацію фітнес-функції Хьюбера.

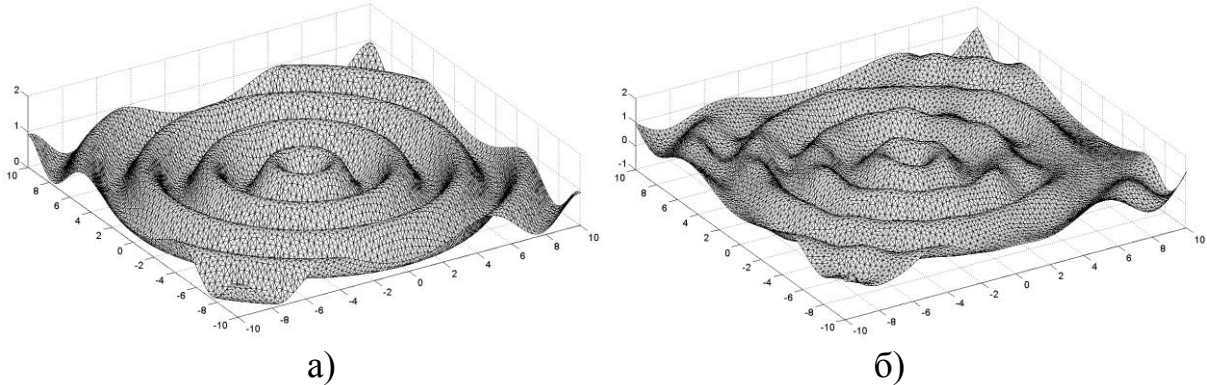


Рисунок 6 – Поверхня, що описується рівнянням (25): а) еталонна поверхня; б) її апроксимація за допомогою ЕРБМ

Використання моделі Тьюки-Хьюбера для апроксимації завади досліджувалася на задачі апроксимації функції

$$y(k) = 0.725 \sin \left( \frac{16x_1 + 8x_2}{3 + 4x_1^2 + 4x_2^2} \right) + 0.2x_1 + 0.2x_2 + \xi(k), \quad (26)$$

де  $x_1, x_2$  – вхідні сигнали, які являли собою стаціонарні випадкові послідовності з рівномірним законом розподілу в інтервалі  $[-1, 1]$ , що генеруються датчиком випадкових чисел;  $\xi(k) = (1 - \varepsilon)q_1(k) + \varepsilon q_2(k)$  – засмічена випадкова завада, що розподілена за законом Релея з дисперсією  $\sigma = 1.6$ .

Апроксимація проводилася за допомогою РБМ із застосуванням процедур (7)–(8) та (19)–(20). Зашумлена поверхня, її перетини та результати апроксимації представлені на рис. 7. Лінією 1 позначено перетин зашумленої поверхні, 4 – перетин еталонної поверхні, 2 – перетин поверхні, відновленої без урахування математичного сподівання завади, 3 – з урахуванням оцінки математичного сподівання завади, отриманої за допомогою процедури (22)–(24).

Використання моделі (17) з апроксимацією її параметрів (22)–(24) і корекцією результатів є досить ефективним при наявності негаусівських завад. Зокрема це було продемонстровано на задачі апроксимації функції, що описувалась рівнянням:

$$y(k) = x_1 \sin(4\pi x_1) - x_2 \sin(4\pi x_2 + \pi) + 1 + \xi(k). \quad (27)$$

У цьому експерименті завада  $\xi(k)$  представляла собою суміш двох нормально розподілених завад  $q_1(k) \sim N(0;0.6)$  та  $q_2(k) \sim N(1;6)$  при  $\varepsilon = 0.1$ . Гістограма такої завади наведена на рис. 8а.

На кожному кроці процесу навчання здійснювалась оцінка параметрів завади з використанням процедур (22)–(24). На останньому кроці навчання було отримано такі оцінки:  $S_1 = 0.6022$ ;  $S_2 = 6.3530$ ;  $\hat{m}_1 = 0.0231$  и  $\hat{m}_2 = 1.2539$ , які використовувались для корекції результатів. На рис. 8б показано еталонну поверхню, що описується рівнянням (27), а на рис. 8в – відновлену за допомогою нейронної мережі поверхню.

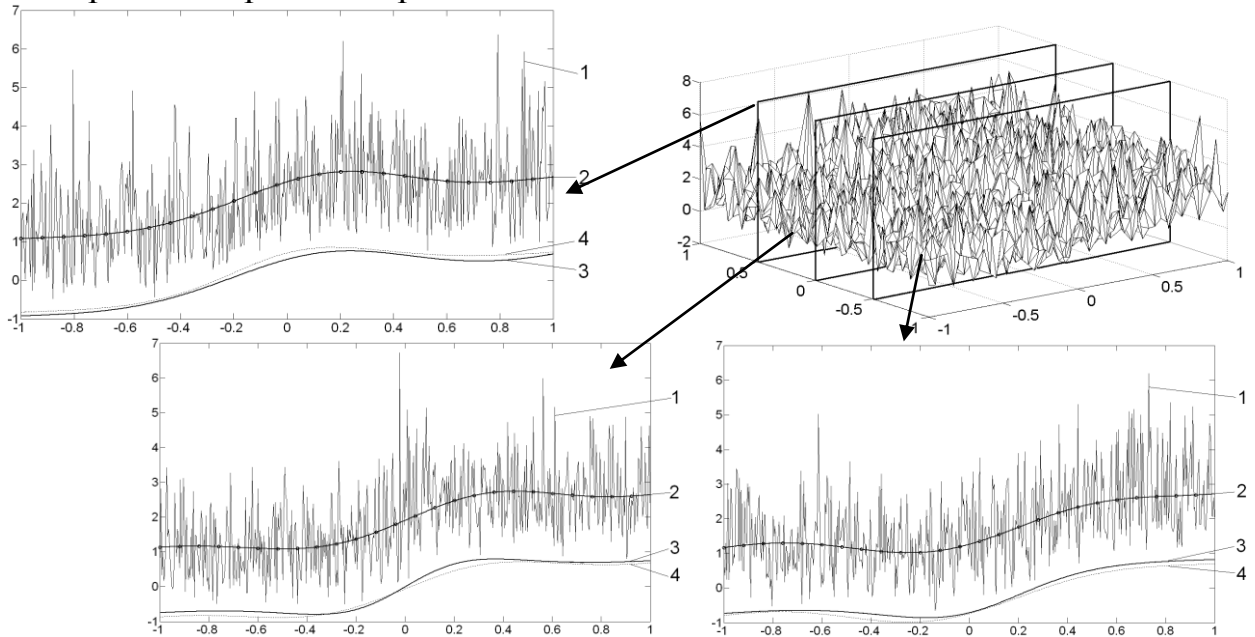
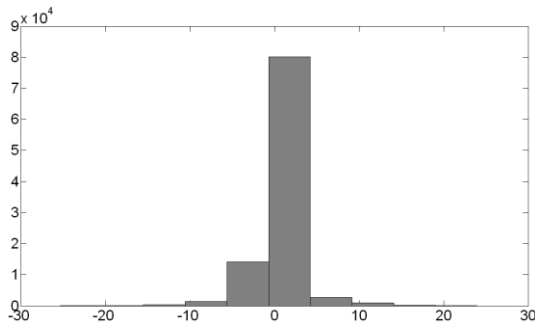
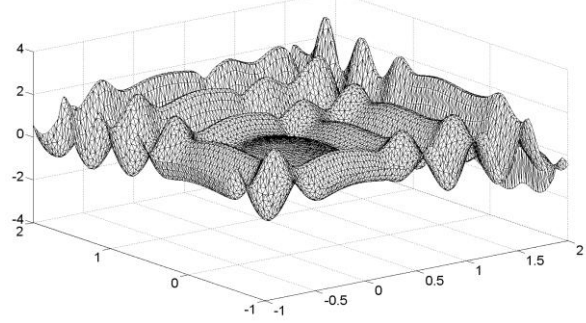


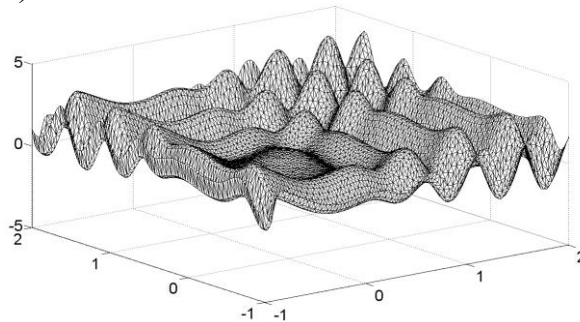
Рисунок 7 – Результати апроксимації функції (26)



а)



б)



в)

Рисунок 8 – Результати апроксимації функції (27): а) гістограма завади; б) еталонна поверхня; в) відновлена за допомогою нейронної мережі поверхня

**П'ятий розділ** присвячено багатокритеріальній оптимізації еволюційних нейромережових моделей. Зазначено, що використання ШНМ потебує вирішення задач структурної та параметричної оптимізації, що відповідає вибору оптимальної топології мережі та її навчання. На відміну від задачі визначення структури, яка є дискретною оптимізаційною (комбінаторною), пошук оптимальних параметрів здійснюється в безперервному просторі за допомогою класичних методів оптимізації. Для навчання мереж прямого розповсюдження з учителем застосовуються, як правило, алгоритми, що оптимізують деяку цільову функцію. Однак традиційно до уваги береться лише одна ціль як вартісна функція або кілька цілей об'єднуються в одну скалярну функцію виду

$$f = F(e) + \lambda\Omega, \quad (28)$$

де  $F$  – функція помилки;  $\Omega$  – міра складності моделі, наприклад, кількість вільних параметрів моделі;  $\lambda > 0$  – деякий вільно обраний параметр. Одержуваний при цьому алгоритм навчання здатний оптимізувати дві цілі, хоча функція цілі (28) є скалярною.

Однак використанню скаляризованих цільових функцій для багатоцільової оптимізації притаманні два основних недоліки: по-перше, нетривіальною є задача визначення оптимального параметра  $\lambda$ ; по-друге, при цьому може бути отримано тільки одне рішення, яке в ряді випадків може бути неефективним.

Більш потужним порівняно з навчанням на основі скалярної вартісної функції є багатоцільове навчання на основі підходу Парето, коли мінімізується векторна цільова функція, що забезпечує отримання певної кількості Парето-оптимальних рішень.

Так, скаляризована двіцільова проблема навчання (28) може бути сформульована як багатокритеріальна оптимізація на основі Парето таким чином:

$$\min \{f_1, f_2\}; \quad (29)$$

$$f_1 = F(e); \quad (30)$$

$$f_2 = \Omega. \quad (31)$$

Найбільш часто в якості  $f_1$  вибирається квадратичний функціонал, а в якості  $f_2$ , що служить для оцінювання складності нейромережової моделі, сума квадратів вагових параметрів:

$$\Omega = \sum_{i=1}^M w_i^2 \quad (32)$$

або сума їх абсолютних значень

$$\Omega = \sum_{i=1}^M |w_i|, \quad (33)$$

відомі як регуляризатори Гауса і Лапласа відповідно. Тут  $w_i, i=1, \dots, M$  – вагові параметри нейромережової моделі;  $M$  – загальна кількість нейронів мережі (для РБМ) або загальна кількість зв'язків (для БП).

У зв'язку з тим, що нейромережева модель, з одного боку, повинна бути досить простою і зручною для використання її в прикладних задачах, а з іншого – найбільш повно відображати властивості досліджуваного об'єкта, її якість визначається деяким набором критеріїв, тобто задача побудови нейромоделі є ба-



гатокритеріальною.

Розглянуто математичне формулювання задачі багатокритеріальної оптимізації за Парето та запропоновано узагальнений еволюційний алгоритм її вирішення, що включає ГА та алгоритм, заснований на штучних імунних системах. Слід зазначити, що ГА для вирішення задачі БО багато в чому схожий з процедурою, заснованою на штучних імунних системах. Тому узагальнений алгоритм включає ці два алгоритми та містить наступні основні кроки:

1. Створення початкової популяції.
  - 1.1. Ініціалізація хромосоми кожної особини.
  - 1.2. Оцінювання початкової популяції.
2. Етап еволюції – побудова нового покоління.
  - 2.1. Відбір (селекція) кандидатів на схрещування (клонування).
  - 2.2 Схрещування (клонування).
  - 2.3. Мутація.
  - 2.4. Оцінювання нової популяції (отриманих клонів).
3. Побудова фронту Парето на основі обраних критеріїв навчання.

4. Вибір єдиного рішення з отриманого набору оптимальних рішень за допомогою деякого інформаційного критерію після досягнення критерію зупини.

Слід зазначити, що найбільш важливим кроком процедури оптимізації ШНМ є вибір оптимального рішення з фронту Парето. Тому доцільним для цього є використання інформаційного критерія, що містить інформацію про помилку навчання, складність та параметри моделі.

Розглянуто питання синтезу коеволюційних ШНМ, в яких реалізується адаптація до мінливого зовнішнього середовища та використовуються значення фітнес-функцій щодо деяких опонентів, а не однієї особини. Аналіз двох основних форм коеволюційних систем (конкуренції та кооперації) дозволив визначити їх переваги та недоліки та реалізувати багатокритеріальні методи навчання ШНМ на основі кооперативної та конкурентної коеволюції. При цьому показано, що для всіх типів мереж у всіх суб-популяціях використовуються загальні еволюційні процедури (селекція, схрещування, мутація і т. д.), а відмінності полягають лише в способі кодування структури і параметрів ШНМ.

Докладно розглянуто багатокритеріальний коеволюційний метод навчання нейронних мереж на основі кооперації. При такому підході кожна нейронна мережа навчається з урахуванням декількох цілей (наприклад, точність і складність мережі). Рішення, які є прийнятними для обох цілей, переносяться та зберігаються. При кооперативній коеволюції кращі особини використовуються як представники своєї суб-популяції, а особини, що залишилися всередині суб-популяції, ранжируються у відповідності до значень їх фітнес-функцій.

Кооперативна оцінка особини в суб-популяції здійснюється шляхом об'єднання даної особини з випадковими недомінуючими особинами в інших суб-популяціях. Потім особина оцінюється для кожної заданої мети. Після оцінювання всіх особин в межах суб-популяцій оцінюється вся суб-популяція та визначаються її домінуючі особини. Під час першої ітерації, коли домінуючі

особини невідомі, особини оцінюються шляхом об'єднання з випадковими особинами з інших суб-популяцій. Цикл завершується, коли всі суб-популяції еволюціонують. Після цього відбувається їх переоцінка для проведення оцінки нових особин, що виникли в процесі еволюції.

Слід зазначити, що для БП і РБМ підходи до розбиття мережі на суб-популяції дещо відрізняються. Детально процес розбиття глобальної задачі багатокритеріального навчання БП на підзадачі, які вирішуються в окремих суб-популяціях, і оцінки їх придатності пояснюється на рис. 9. На цьому рисунку представлено загальний вигляд всіх суб-популяцій в рамках коєволюції на основі кооперації. На початку роботи алгоритму кожна особина в суб-популяції прагне мінімізувати кожен з цілей фітнес-функції. Для оцінки придатності особини вона об'єднується з випадковими невідоміючими особинами з інших суб-популяцій. Потім всі особини об'єднуються для формування спільного рішення. Обрані особини з суб-популяцій об'єднуються, формуючи загальну хромосому (генотип) і відображаються в нейронну мережу (фенотип). Отримана ШНМ потім оцінюється з урахуванням різних цілей. Після цього значення фітнес-функції отриманої результуючої мережі бере участь в обчисленні фітнес-функцій тих особин (представників суб-популяцій), які брали участь у формуванні цієї мережі. Потім весь процес починається знову для наступної особини.

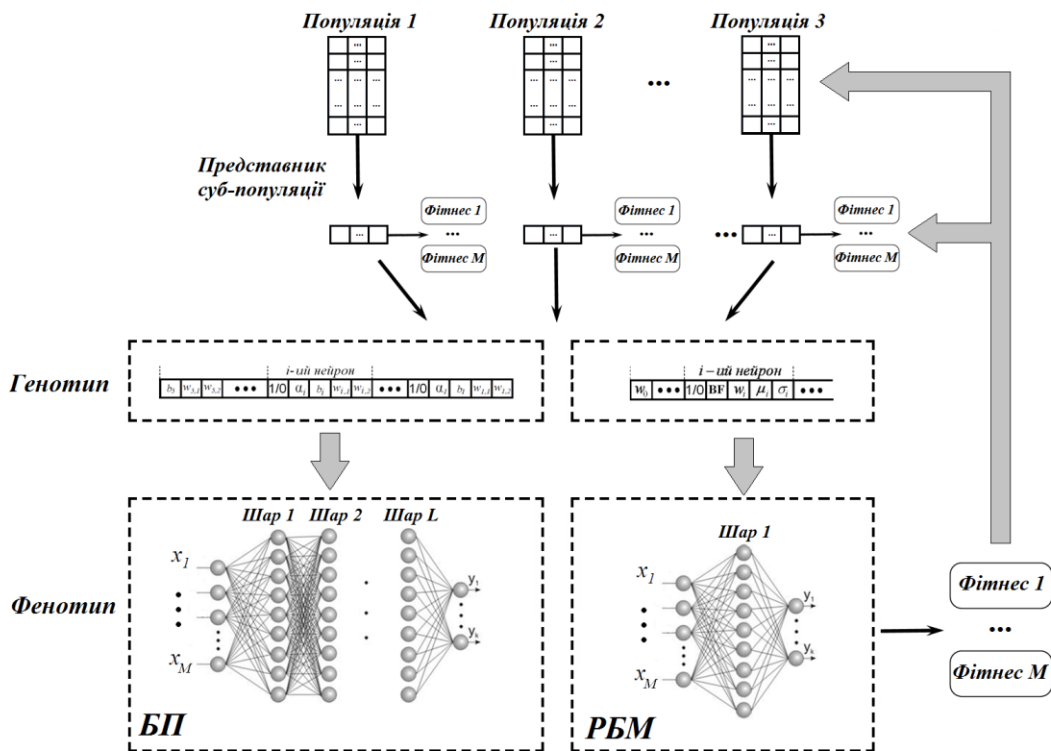


Рисунок 9 – Процес розбиття глобальної задачі багатокритеріального навчання ШНМ на підзадачі

Також розглянуто дещо інший підхід до застосування механізму коєволюції до налаштування ШНМ, який отримав у літературі назву COVNET. COVNET є кооперативною коєволюційною моделлю, в якій кілька видів спільно коєволюціонують. Кожен вид при цьому є підмережею і являє собою част-

кове вирішення проблеми. Комбінація декількох особин з різних видів становить ШНМ, яка повинна бути застосована до конкретної проблеми. Популяція підмереж, які називаються нодулами, складається з декількох суб-популяцій, які еволюціонують незалежно одна від одної. Кожна з таких суб-популяцій є окремим видом, а ключовим елементом запропонованої моделі є комбінація особин з цих різних суб-популяцій, що коеволюціонують разом.

Слід зазначити, що генотип нодулів має взаємно-однозначне відображення у фенотип. ШНМ при цьому можна визначити як сукупність нодулів. На практиці для спрощення алгоритму зазвичай вважають, що всі мережі популяції повинні мати однакову фіксовану протягом усього процесу еволюції кількість нодулів.

Розглянуто особливості рішення задачі багатокритеріального навчання при використанні РБМ. Для оцінки ролі кожного нейрону (БФ) у кооперативно-конкурентному середовищі потрібен деякий механізм їх оцінки. Для цього пропонується визначати такі три параметри: вклад –  $a_i$ , помилку –  $e_i$  та перекриття –  $o_i$ .

Вклад  $a_i$  для  $i$ -ї БФ  $\varphi_i$  ( $i = 1 \dots m$ ), визначається з урахуванням її вагового параметра  $w_i$ , та кількості зразків навчаючої вибірки, яка знаходиться усередині її рецептивного поля, що визначається радіусом  $\sigma_i$ . БФ з невеликим значенням вагового параметру та невеликою кількістю вхідних сигналів усередині цього рецептивного поля буде відповідно вносити неістотний вклад у загальний результат мережі:

$$a_i = \begin{cases} |w_i|, & \text{якщо } \sigma_i > q; \\ |w_i|(\sigma_i/q), & \text{в протилежному випадку,} \end{cases} \quad (34)$$

де  $q$  – середнє значення усіх  $\sigma_i$  мінус їх стандартне відхилення.

Вимірювання параметра помилки  $e_i$  кожної БФ (нейрона) засновано на підрахунку всіх помилок при обробці навчальних пар всередині її рецептивного поля:

$$e_i = \frac{1}{N} \sum_{j=1}^N (y_j^* - y_j)^2, \quad (35)$$

де  $y_j^*$  та  $y_j$  – відповідно бажане та реальне значення вихідного сигналу мережі;  $N$  – кількість навчальних пар всередині рецептивного поля  $i$ -ї БФ (нейрона).

Ступінь перекриття  $i$ -ої БФ  $\varphi_i$  іншими БФ характеризується параметром  $o_i$ , що обчислюється таким чином:

$$o_i = \sum_{j=1}^M o_{ij}; \quad (36)$$

$$o_{ij} = \begin{cases} \left( 1 - \frac{\|\varphi_i - \varphi_j\|}{d_i} \right), & \text{якщо } \|\varphi_i - \varphi_j\| < d_i; \\ 0, & \text{в протилежному випадку,} \end{cases}$$

де  $o_{ij}$  визначає ступінь перекриття  $i$ -ої БФ  $\varphi_i$  з  $j$ -ою  $\varphi_j$ .

Слід зазначити, що прагнення мінімізувати ступінь перекриття БФ може, з одного боку, призводити до підвищення точності роботи мережі, але з іншого – до збільшення кількості нейронів прихованого шару.

У роботі запропоновано метод генерації тестових функцій для перевірки алгоритмів вирішення задач БО, які використовують сферичні функції, функції Еклі тощо.

Розглянуто можливість реалізації багатокритеріального навчання ШНМ з використанням алгоритмів кластеризації, зокрема, оптимізованого методу векторного квантування (OLVQ). Як приклад було вирішено задачу апроксимації функції (25) з  $\xi(k)=0$  за допомогою кооперативно-конкурентної стратегії, що полягає в паралельній еволюції множини популяцій, які вирішують різні підзадачі апроксимації. Для цього простір вхідних сигналів розбивалося на кластери за допомогою додаткової популяції, яка відповідала за налаштування центрів кластерів за допомогою ЕА. Було задано 25 кластерів, початкові значення центрів яких задавалися двома способами – випадковим розподілом і рівномірним. Задачу апроксимації в кожному кластері вирішувала окрема популяція РБМ. Після 500 епох навчання кінцеве рішення було сформоване за допомогою особин-переможців в кожному кластері. Результати апроксимації функції (25) з різним підходом до вибору початкових значень центрів кластерів представлені на рис. 10. На рис. 10а наведено відновлені поверхні, а на рис. 10б – результуюче розбиття простору вхідних сигналів на кластери. Як видно з результатів експериментів, для цього випадку початковий рівномірний розподіл простору вхідних сигналів на кластери дає кращі результати, однак в загальному випадку рекомендується все ж використовувати випадкове розбиття, тому що найчастіше вид функції, що апроксимується, є невідомим.

Вивчалися особливості застосування коеволюційного підходу до вирішення задачі апроксимації нелінійних багатовимірних функцій: чотиривимірної

$$F(x_1, x_2, x_3, x_4) = x_1 + \sin(\pi x_1) \cdot \cos(\pi x_2) \cdot \sin(\pi x_3) \cdot (\sin(\pi x_4)^2 - 1) \quad (37)$$

та шестивимірної

$$F(x_1, x_2, x_3, x_4, x_5, x_6) = \sin(\pi x_1) \cdot \sin(\pi x_2) \cdot \sin(\pi x_3) + \sin(\pi x_4) \cdot \sin(\pi x_5) \cdot \sin(\pi x_6) + \exp(-(x_1^2 + x_2^2)) + \exp(-(x_3^2 + 6x_3x_4 + 10x_4^2)). \quad (38)$$

Апроксимація функцій (37)–(38) здійснювалася за допомогою коеволюційного підходу на основі конкуренції. За отримання права надання найкращого рішення задачі апроксимації змагалися дві популяції еволюційних ШНМ, що склалися з РБМ і БП. Обидві популяції містили по 150 особин, і після кожної епохи навчання за результатами обчислення значень фітнес-функції (28), вимірювання часу симуляції і визначення кількості параметрів, що настроюються, для кожної особини будувався результуючий фронт Парето для обох популяцій. Всі можливі рішення після 1000 епох навчання та результуючий фронт Парето для функції (37) представлено на рис. 11а. Аналогічно для функції (38) результати моделювання наведені на рис. 11б. Рішення, отримані за допомогою БП, на

рисунках позначено зірочками, а місце їх зосередження обведено суцільною жирною лінією. Рішення ж, отримані за допомогою РБМ, показано на рисунках кружечками, а область їх скупчення обведена пунктирною лінією. Фронт Парето будувався загальний для обох популяцій і з нього вибиралося найкраще рішення відповідно до критерію Акаїке. Краще рішення могло належати як до популяції, що складається як з БП, так і з РБМ.

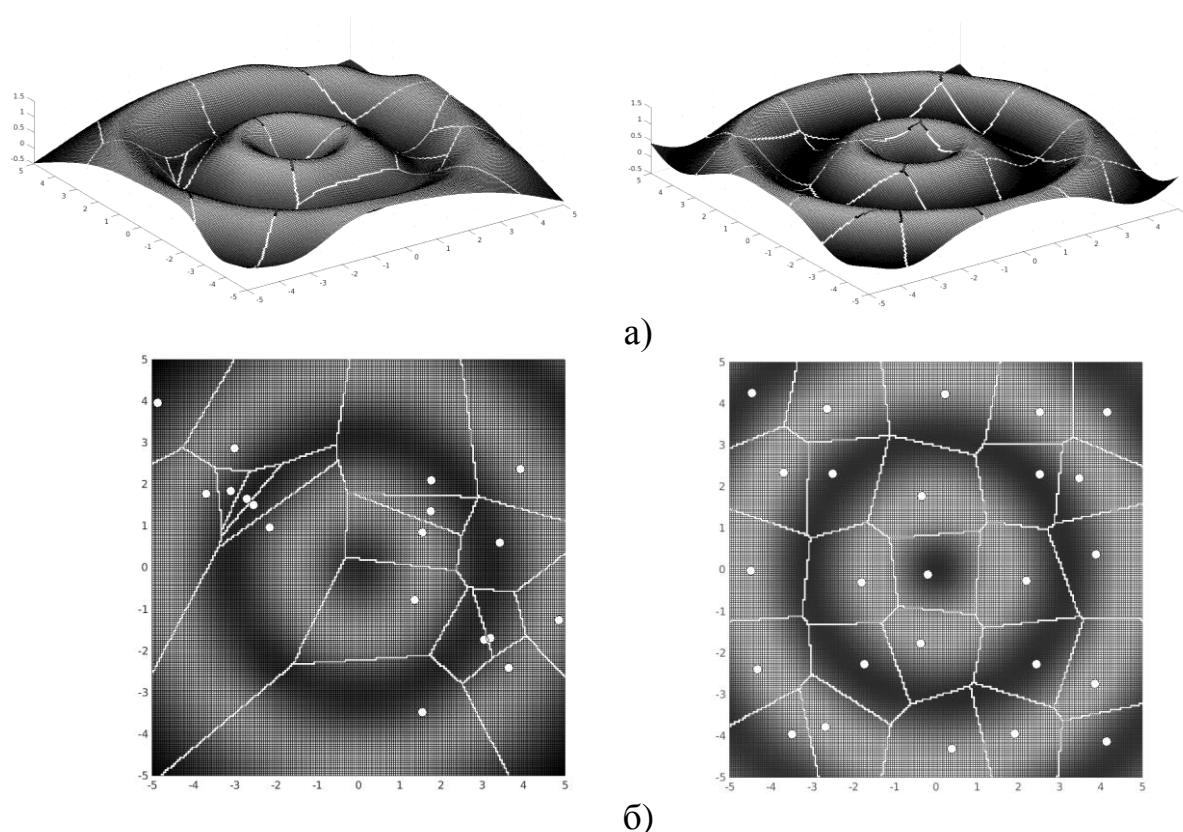


Рисунок 10 – Результати апроксимації функцій (37)-(38): а) відновлені поверхні; б) результуюче розбиття простору вхідних сигналів на кластери

Як видно з результатів моделювання, для чотиривимірного об'єкта популяція РБС дає кращі результати за всіма критеріями, що мінімізуються, а для шестивимірного об'єкта популяція МП дає кращі рішення по складності моделі і часу симуляції, дещо програючи, проте, по точності ідентифікації.

**Шостий розділ** присвячено імітаційному моделюванню та вирішенню практичних задач ідентифікації, управління, прогнозування та обробки зображень за допомогою ЕШНМ. Показано, що серед сучасних нейросимуляторів, таких як *Scilab*, *Torch 7*, *Theano*, *Caffe* та *NeurophStudio*, останній надає найбільш повні можливості щодо дослідження властивостей алгоритмів навчання ШНМ. У зв'язку з цим моделювання ШНМ здійснювалося у середовищі *NeurophStudio*.

Вирішувались задачі ідентифікації стаціонарних і нестаціонарних нелінійних динамічних одно- і багатовимірних об'єктів та адаптивного управління такими об'єктами. Проведено імітаційне моделювання розроблених архітектур та методів навчання ЕШНМ. Показано їх переваги перед відомими нейромере-

жевими системами обробки інформації та методами їх навчання по точності, стійкості та швидкодії в задачах ідентифікації одно- і багатовимірних нелінійних нестационарних динамічних об'єктів при наявності різного роду завад в умовах апріорної і поточної невизначеності.

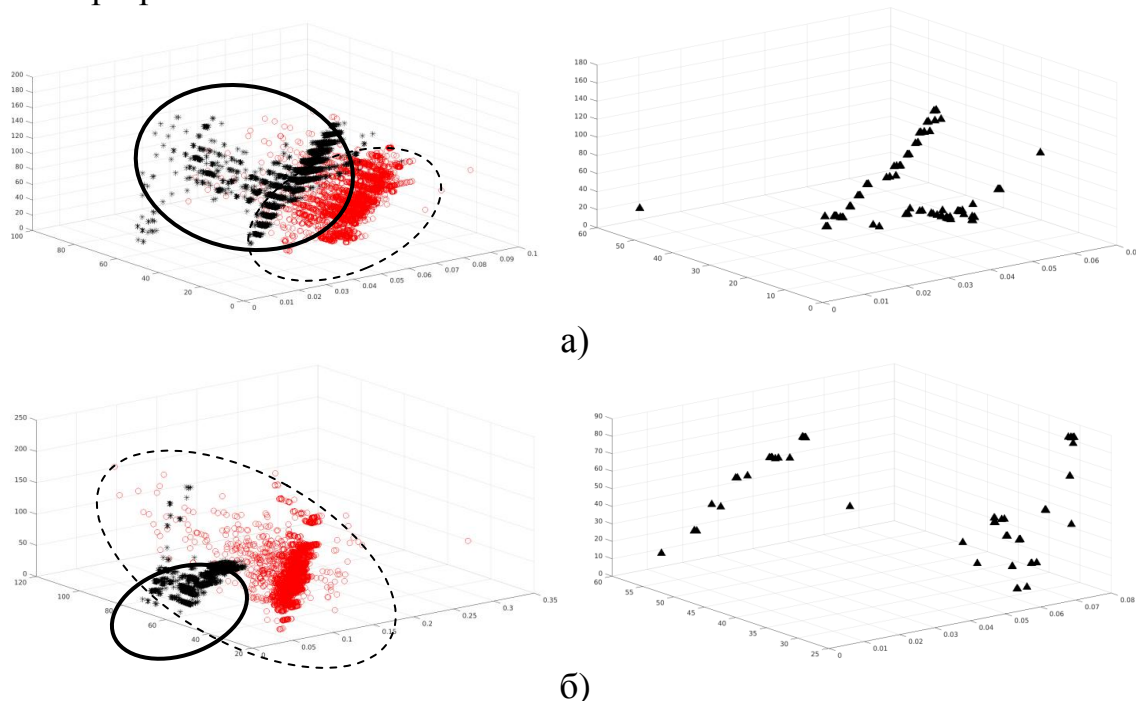


Рисунок 11 – Всі можливі рішення та фронти Парето: а) для функції (37); б) для функції (38)

Вирішувалося завдання ідентифікації багатовимірного об'єкта (МІМО), який описувався такими рівняннями:

$$y_1(k) = \frac{15u_1(k-1)y_2(k-1)}{2 + 50[u_1(k-1)]^2} + 0.5u_1(k-1) - 0.25y_2(k-1) + 0.1 + \xi_1(k);$$

$$y_2(k) = \frac{\sin(\pi u_2(k-1)y_1(k-1)) + 2u_2(k-1)}{3} + \xi_2(k),$$
(39)

де  $u_1, u_2$  – вхідні сигнали;  $y_1$  та  $y_2$  – вихідні сигнали;  $\xi_1$  та  $\xi_2$  – завади вимірювань.

Завада  $\xi_1$  описувалася моделлю (17). Завада  $\xi_2$  мала розподіл Релея (Ray(1,6)). Таким чином, вихідні сигнали об'єкта було зашумлено різними завадами. Багатозв'язаність даного об'єкта призводить до того, що остаточний вигляд розподілів встановити досить складно. Після 2000 епох навчання було отримано такі оцінки параметрів завад (незалежні по кожному виходу об'єкта):  $\sigma_1^1 = 1.2042$ ;  $\sigma_2^1 = 7.4612$ ;  $m_1^1 = 0.0$ ;  $m_2^1 = 0,6765$  – для першого виходу;  $\sigma_1^2 = 2.5593$ ;  $\sigma_2^2 = 1.5492$ ;  $m_1^2 = 0.0$ ;  $m_2^2 = 2,0294$  – для другого виходу.

Мережа містила 17 нейронів (9 нейронів с гаусівською БФ та 8 – з БФ «мексиканський капелюх»). Результати моделювання наведено на рис. 12. На рис. 12а, б представлено графіки зміни значення фітнес-функції особини-переможця та кількості активних нейронів особини-переможця.

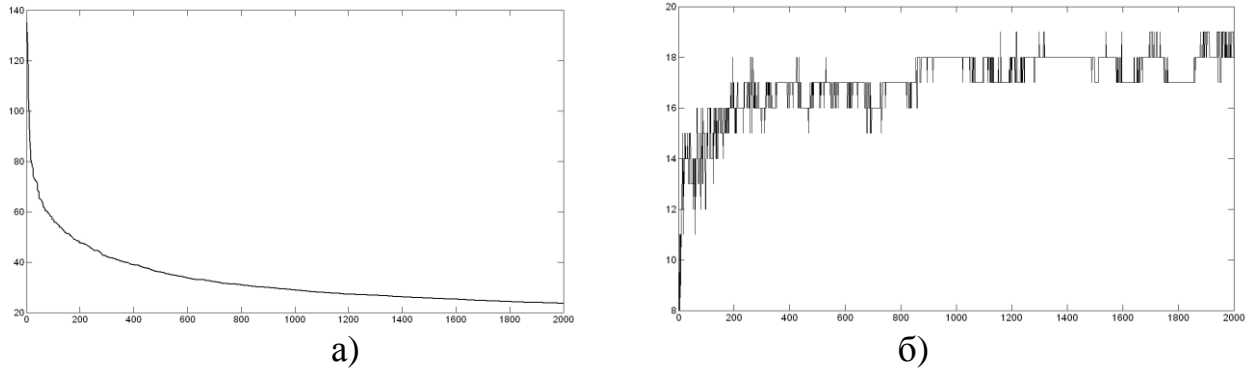


Рисунок 12 – Результати ідентифікації об'єкта (39): а) графік зміни значення фітнес-функції; б) графік зміни кількості активних нейронів

При управлінні для знаходження мінімуму функціоналу помилки використовується градієнтна процедура

$$\mathbf{u}(k) = \mathbf{u}(k-1) + \nabla_{\mathbf{u}} \varepsilon(k) \left[ \delta \mathbf{I} + \nabla_{\mathbf{u}}^T \varepsilon(k) \nabla_{\mathbf{u}} \varepsilon(k) \right]^{-1} \varepsilon(k), \quad (40)$$

де  $\nabla_{\mathbf{u}} \varepsilon(k) = \frac{\partial \varepsilon(k)}{\partial \mathbf{u}(k)}$ ;  $\gamma(k) > 0$ ;  $\delta > 0$ .

Результати роботи нейрорегулятора, який управляє нелінійним динамічним об'єктом, що описується рівняннями (39), та реалізує процедуру (40) з  $\delta = 0.05$ , наведено на рис. 13. На всіх рисунках пунктирною лінією показано необхідний вихідний сигнал  $y_i^*(k)$ , суцільною – реальний  $\hat{y}_i(k)$ , а лінією з кружками – відповідну зміну керуючого сигналу  $u_i(k)$  ( $i=1,2$ ). Необхідні значення вихідних сигналів задавалися такі:

$$y_1^*(k) = 0.3 \sin(\pi k / 100);$$

$$y_2^*(k) = \begin{cases} 0.1 & \text{при } k = \overline{1,500}; \\ -0.1 & \text{при } k = \overline{501,1000}. \end{cases}$$

Рис. 13а, б відображають роботу регулятора при наявності рівномірно розподіленої в інтервалі  $[-0.3, 0.3]$  випадкової завади  $\zeta(k)$ . Поява даної завади призвела до того, що при побудові нейромережевої моделі збільшилася кількість нейронів і стала рівною 69, а при управлінні моделлю – змінився вигляд управляючого сигналу (особливо це помітно для  $u_2(k)$ ). Крім цього, збільшилася помилка  $\zeta_2(k)$ .

Розроблено та досліджено нові закони адаптивного прогнозуючого нейроуправління багатовимірними нелінійними динамічними об'єктами, що функціонують в умовах невизначеності. Для корекції еталонної траєкторії запропоновано використовувати еволюційний підхід.

Результати роботи нейрорегулятора, при керуванні об'єктом (39), показано на рис. 14. На всіх рисунках пунктирною лінією показано необхідний вихідний сигнал  $r_i(k)$ , суцільною – реальний  $\hat{y}_i(k)$ , а лінією з кружками – зміну керуючого сигналу  $u_i(k)$ ,  $i=1,2$ . Рисунки 14а та 14б відображають результати моделювання при наявності завод вимірювання по обох каналах управління,

при цьому  $\xi_1(k)$  являла собою нормально розподілену заваду з  $\sigma^2 = 0.6$ , а  $\xi_2(k)$  – з  $\sigma^2 = 0.8$ .

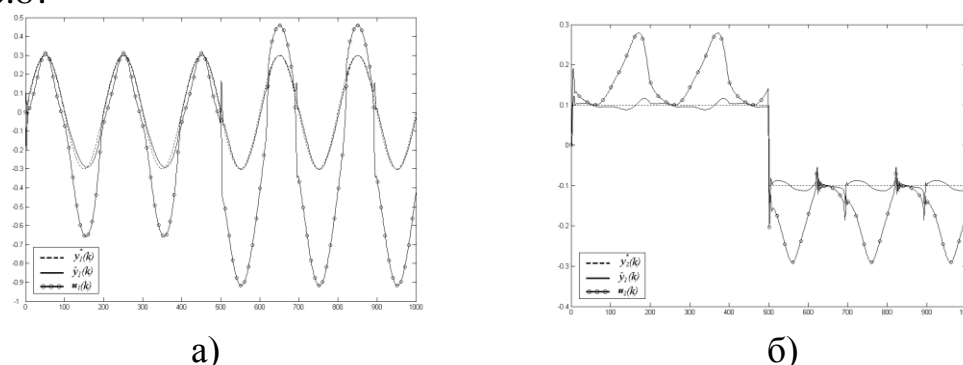


Рисунок 13 – Управління об'єктом (39) при наявності завади  $\zeta(k)$ : а) по першому каналу; б) по другому каналу

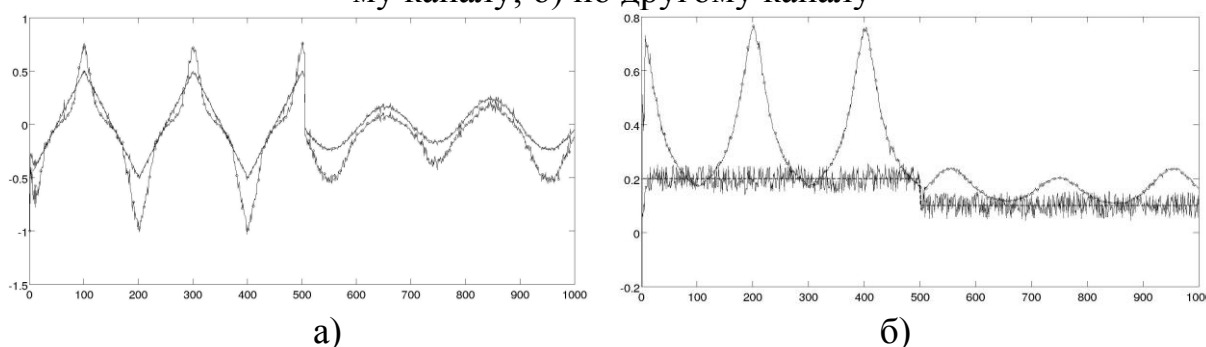


Рисунок 14 – Результати прогнозуючого управління об'єктом (39): а) по першому каналу; б) по другому каналу

Результати моделювання свідчать про те, що підхід, який розвивається, дозволяє синтезувати нейромережевий контролер для управління багатовимірними нестационарними об'єктами при наявності завад вимірювання.

Отримані в роботі результати були використані при вирішенні низки практичних задач, зокрема при побудові системи вимірювання температурних полів трубчастих печей в ТОВ "Побузький феронікелевий комбінат", управлінні процесом травлення смугової сталі на металургійних підприємствах України (АТ «Співдружність-Т») та при розробці АСУ ТП дифузії, дефекосатурації, випарювання і кристалізації в ТОВ "Кириківський цукровий завод".

У **висновках** сформульовано основні наукові та практичні результати дисертаційної роботи.

**Додатки** містять опис основних генетичних операторів, методи отримання рекурентних форм процедур навчання та аналіз їх збіжності, а також копії документів про впровадження результатів дисертаційної роботи.

## ОСНОВНІ РЕЗУЛЬТАТИ ТА ВИСНОВКИ

Дисертаційна робота присвячена вирішенню важливої науково-практичної проблеми розвитку теоретичних основ інтелектуального аналізу даних та розробці нових еволюційних ШНМ прямого розповсюдження з метою підвищення ефективності обробки інформації в умовах апріорної та поточної



невизначеності, які мають важливе наукове і практичне значення для створення ефективних інтелектуальних систем обробки інформації, емуляції, управління, прогнозування, ідентифікації як функціонуючих об'єктів, так і новостворених.

Внаслідок виконання роботи отримано такі результати.

1. Проведений аналіз сучасного стану проблеми інтелектуальної обробки інформації показав, що в якості базису для синтезу подібних систем обробки доцільно використовувати статичні ШНМ прямого розповсюдження, а також обґрунтовано необхідність застосування для поліпшення їх властивостей методів теорії адаптації та еволюції.

2. Розроблено новий метод робастної багатокритеріальної оптимізації (Парето-оптимізації) на основі робастних фітнес-функцій та інформаційних критеріїв оцінки складності моделі, який дозволяє визначити структуру оптимальної нейромережевої моделі при наявності негаусовських завад.

3. Розроблено узагальнений ЕА Парето-оптимізації на основі коеволюційного підходу, який дозволяє змінювати архітектуру мережі, адаптуючись до мінливого зовнішнього середовища.

4. Розроблено метод спрощення структури ШНМ прямого розповсюдження з метою прискорення процесів обробки інформації при допустимій неточності та нові методи навчання ШНМ, що забезпечують необхідну точність навчання при наявності обмежених завад. Ці методи характеризуються обчислювальною простотою і підвищеною швидкістю за рахунок використання додаткових процедур адаптивної корекції їх параметрів.

5. Розроблено нові робастні методи навчання ШНМ, що дають можливість ефективно обробляти інформацію при наявності завад з негаусовськими, зокрема асиметричними, розподілами, здійснювати корекцію одержуваних оцінок, усувати їх зміщення.

6. Розроблено нові процедури корекції параметрів використовуваних при навчанні функціоналів і оцінюванні параметрів завад, що описуються моделлю Тьюки-Хьюбера. Це дозволяє при відсутності апріорної інформації про статистичні властивості завад адаптивно коригувати параметри ШНМ.

7. Запропоновано нові методи апроксимації гаусівських базисних функцій в РБС нульового і першого порядків. Використання кусково-лінійної апроксимації дозволяє істотно спростити обчислення, супутні процесам побудови нейромережевої моделі досліджуваних об'єктів і управління ними.

8. Розроблено нові закони адаптивного прогноуючого нейро-управління нелінійними нестационарними динамічними об'єктами, що функціонують в умовах невизначеності. Для корекції еталонної траєкторії запропоновано використовувати еволюційний підхід. Отримані закони управління не використовують градієнтні алгоритми і характеризуються підвищеною швидкістю. Додаткова перевага забезпечується використанням результатів роботи контуру, в якому оцінюються параметри завади.

9. Отримав подальший розвиток метод гібридного навчання ШНМ. Модифікація методу полягає у використанні еволюційного алгоритму для грубої настройки параметрів мережі та розробленої робастної рекурентної процедури

Левенберга-Марквардта для остаточного тонкого налаштування. Така модифікація дозволяє підвищити якість одержуваної нейромережевої моделі і стійкість процесу навчання в умовах апріорної і поточної невизначеності.

10. Отримав подальший розвиток еволюційний метод усунення впливу завод при визначенні структур і параметрів нейромережевих моделей за рахунок використання оцінок параметрів моделі завод Тьюки-Хьюбера в процедурі М-навчання. Такий підхід дозволяє поліпшити фільтруючі властивості використовуваних процедур, спростити структуру хромосоми і скоротити тим самим процес отримання моделі ШНМ.

11. Отримав подальший розвиток еволюційний метод багатокритеріальної оптимізації структури і параметрів ШНМ шляхом виділення загальних для еволюційного і імунного підходу операторів. Така модифікація дозволяє найбільш ефективно використовувати переваги обох підходів.

12. Отримала подальший розвиток архітектура еволюційної ШНМ прямого розповсюдження, яка крім еволюції структури мережі та її параметрів враховує також еволюцію алгоритму навчання та завади. Така ШНМ має істотно поліпшені апроксимуючі і екстраполюючі властивості, що дає можливість за рахунок адаптації структури, параметрів і процедури навчання ефективно обробляти інформацію в нестаціонарних умовах і за наявності невизначеності.

13. Отримали подальший розвиток коеволюційні ШНМ прямого розповсюдження, в яких для вирішення задачі багатокритеріальної оптимізації використовується алгоритм кластеризації. Такий підхід дозволяє спростити архітектуру одержуваної нейромережевої моделі і підвищити її робастність.

14. Проведено імітаційне моделювання розроблених архітектур і методів навчання еволюційних ШНМ. Показані їх переваги перед відомими нейромережевими системами обробки інформації та методами їх навчання за точністю, стійкістю та швидкодією в задачах апроксимації (емуляції), ідентифікації нелінійних нестаціонарних процесів та інтелектуального управління нелінійними нестаціонарними динамічними об'єктами в умовах апріорної і поточної невизначеності.

15. Вирішено практичні задачі вимірювання температурних полів трубчастих печей в ТОВ "Побузький феронікелевий комбінат"; управління процесом травлення смугової сталі на металургійних підприємствах України (АТ «Співдружність-Т»); розробки АСУ ТП дифузії, дефекосатурації, випарювання і кристалізації в ТОВ "Кириківський цукровий завод".

## **СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ**

1. Руденко О.Г. Нейросетевые методы сжатия изображений / О.Г. Руденко, С.В. Мирошниченко, А.А. Бессонов // Коллективная монография «Наукоемкие технологии в инфокоммуникациях: обработка и защита информации», Харьков: СМІТ. – 2013. – 398 с.

2. Руденко О.Г. Прогнозирование экономических процессов с помощью эволюционирующих искусственных нейронных сетей / О.Г. Руденко,

А.А. Бессонов // Коллективная монография «Информационные технологии в управлении, образовании, науке и промышленности», Харьков: Издатель Рожко С.Г., 2016. – 566 с.

3. Руденко О.Г. Робастное обучение вейвлет-нейросетей / О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики. – 2010. – №5. – С. 66–79. (Входит до міжнародних наукометричних баз SCOPUS, ISI та ін.)

4. Бессонов А.А. Оценивание параметра масштаба при робастном обучении искусственных нейронных сетей / А.А. Бессонов, О.Г. Руденко, С.О. Руденко // Проблеми інформаційних технологій. – 2010. – №2 (008). – С. 81–84.

5. Бессонов А.А. Устойчивое обучение радиально-базисных сетей с ограниченной точностью / А.А. Бессонов, О.Г. Руденко, С.О. Руденко // Вестник Херсонского национального технического университета. – 2010. – №2(38). – С. 130–134.

6. Rudenko O. Function Approximation Using Robust Radial Basis Function Networks / O. Rudenko, O. Bezsonov // Journal of Intelligent Learning Systems and Applications. – 2011. – №3. – P. 17–25.

7. Руденко О.Г. Аппроксимация гауссовских базисных функций в задаче адаптивного управления нелинейными объектами / О.Г. Руденко, А.А. Бессонов, А.С. Ляшенко, Р.А. Сунна // Кибернетика и системный анализ. – 2011. – №1. – С. 3–13. (Входит до міжнародних наукометричних баз SCOPUS, INSPEC, EBSCO, ProQuest, Computer Science Index та ін.)

8. Руденко О.Г. Робастное обучение радиально-базисных сетей / О.Г. Руденко, А.А. Бессонов // Кибернетика и системный анализ. – 2011. – №6. – С. 38–46. (Входит до міжнародних наукометричних баз SCOPUS, INSPEC, EBSCO, ProQuest, Computer Science Index та ін.)

9. Руденко О.Г. М-обучение радиально-базисных сетей с использованием асимметричных функций влияния / О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики. – 2012. – №1. – С. 79–93. (Входит до міжнародних наукометричних баз SCOPUS, ISI та ін.)

10. Руденко О.Г. Идентификация нелинейных нестационарных объектов с помощью эволюционирующей радиально-базисной сети / О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики. – 2012. – №4. – С. 5–14. (Входит до міжнародних наукометричних баз SCOPUS, ISI та ін.)

11. Бессонов А.А. Обучение радиально-базисных сетей с помощью генетических алгоритмов с адаптивной мутацией / А.А. Бессонов // Системи обробки інформації. – 2012. – №3(101). – С. 177–180.

12. Бессонов А.А. Идентификация нелинейных нестационарных объектов с помощью эволюционного многослойного персептрона / А.А. Бессонов, С.О. Руденко // Вестник Херсонского национального технического университета. – 2012. – №1(44). – С. 117–121.

13. Бессонов А.А. Применение робастных фитнес-функций при оценивании параметров нелинейных объектов с помощью эволюционирующей радиально-базисной сети / А.А. Бессонов // Системи управління, навігації та зв'язку. – 2012. – №3(23). – С. 51–56.

14. Бессонов А.А. Многокритериальная нейроэволюционная оптимизация нелинейных функций / А.А. Бессонов // Системи обробки інформації. – 2012. – №9(107). – С. 5–10.

15. Руденко О.Г. Робастная идентификация нелинейных объектов с помощью эволюционирующей радиально-базисной сети / О.Г. Руденко, А.А. Бессонов, С.О. Руденко // Кибернетика и системный анализ. – 2013. – №2. – С. 15–26. (Входит до міжнародних наукометричних баз SCOPUS, INSPEC, EBSCO, ProQuest, Computer Science Index та ін.)

16. Руденко О.Г. Робастная многокритериальная идентификация нелинейных объектов на основе эволюционирующих радиально-базисных сетей / О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики. – 2013. – №5. – С. 22–32. (Входит до міжнародних наукометричних баз SCOPUS, ISI та ін.)

17. Бессонов А.А. Робастная многокритериальная идентификация нелинейных объектов с помощью сетей прямого распространения / А.А. Бессонов, С.О. Руденко // Вестник Херсонского национального технического университета. – 2013. – №1(46). – С. 142–145.

18. Руденко О.Г. Использование нормализующей компоненты при нейросетевом сжатии изображений / О.Г. Руденко, А.А. Бессонов, Р.В. Бобнев // Управляющие системы и машины. – 2013. – №5. – С. 2–31.

19. Руденко О.Г. Робастная нейроэволюционная идентификация нелинейных нестационарных объектов / О.Г. Руденко, А.А. Бессонов // Кибернетика и системный анализ. – 2014. – №1. – С. 21–36. (Входит до міжнародних наукометричних баз SCOPUS, INSPEC, EBSCO, ProQuest, Computer Science Index та ін.)

20. Руденко О.Г. Многокритериальная оптимизация эволюционирующих сетей прямого распространения / О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики. – 2014. – №6. – С. 29–41. (Входит до міжнародних наукометричних баз SCOPUS, ISI та ін.)

21. Руденко О.Г. Прогнозирующее управление нелинейными объектами на основе эволюционирующих нейронных сетей прямого распространения / О.Г. Руденко, А.А. Бессонов // Проблемы управления и информатики, №6. – 2015. – С. 46–55. (Входит до міжнародних наукометричних баз SCOPUS, ISI та ін.)

22. Руденко О.Г. Программирование с экспрессией генов: способы кодирования / О.Г. Руденко, С.В. Мирошниченко, А.А. Бессонов // Управляющие системы и машины. – 2015. – №3(257). – С. 82–92.

23. Руденко О.Г. Программирование с экспрессией генов: генетические операторы / О.Г. Руденко, С.В. Мирошниченко, А.А. Бессонов // Управляющие системы и машины. – 2015. – №4(258). – С. 72–82.

24. Руденко О.Г. Программирование с экспрессией генов: модификации эволюционного процесса / О.Г. Руденко, С.В. Мирошниченко, А.А. Бессонов // Управляющие системы и машины. – 2015. – №5(259). – С. 73–78.

25. Руденко О.Г. Робастная идентификация нелинейных объектов /

О.Г. Руденко, А.А. Бессонов, С.О. Руденко // Проблемы информационных технологий. – 2015. – №1(17). – С. 83–86.

26. Бессонов А.А. Решение задачи управления с прогнозирующей моделью на основе эволюционирующего многослойного персептрона / А.А. Бессонов // Системы обработки информации. – 2015. – №1(126). – С. 7–11.

27. Бессонов А.А. Обобщенный алгоритм обучения эволюционирующей радиально-базисной сети / А.А. Бессонов // Системы обработки информации. – №10(135). – 2015 – С. 163–166.

28. Бессонов А.А. Исследование одношаговых алгоритмов обучения искусственных нейронных сетей / А.А. Бессонов // Системы обработки информации. – №7(144). – 2016. – С. 127–129.

29. Руденко О.Г. Устойчивые алгоритмы обучения радиально-базисных сетей / О.Г. Руденко, А.А. Бессонов // Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту : міжнар. наук. конф. : тези доп. – Євпаторія, 18-22 травня 2009. – С. 421–422.

30. Руденко О.Г. Робастное обучение радиально-базисных сетей / О.Г. Руденко, А.А. Бессонов // Автоматика-2009 : междунар. конф. по автоматическому управлению : тезисы докл. – Черновцы, 22-25 сентября 2009. – С. 368.

31. Руденко О.Г. Нейросетевое управление на основе кусочно-линейной аппроксимации базисных функций / О.Г. Руденко, А.А. Бессонов, А.С. Ляшенко // Автоматизация: проблемы, идеи решения : междунар. науч.-техн. конф. : тезисы докл. – Севастополь, 7-12 сентября 2009. – С. 26–28.

32. Бессонов А.А. Нейросетевые методы построения устойчивых моделей нелинейных объектов / А.А. Бессонов // Сучасні методи, інформаційне, програмне та технічне забезпечення систем управління організаційно-технологічними комплексами : міжнар. наук.-техн. конф. : тези доп. – Київ, 26-27 листопада 2009. – С. 24.

33. Бессонов А.А. Сравнительный анализ эффективности радиально-базисных сетей и вэйвлет-нейросетей в задаче идентификации нелинейных динамических объектов / А.А. Бессонов, С.О. Руденко // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління : міжнар. наук.-техн. конф. : тези доп. – Харків-Київ, 13-14 грудня 2010. – С. 71.

34. Бессонов А.А. Управление сложными технологическими процессами с помощью робастных вэйвлет-нейронных сетей / А.А. Бессонов, С.О. Руденко // Інформаційні технології в навігації і управлінні: стан та перспективи розвитку : міжнар. наук.-техн. конф. : тези доп. – Київ, 5-6 липня 2010. – С. 61.

35. Руденко О.Г. Об особенностях использования неквадратичных критериев в задаче обучения искусственных нейронных сетей / О.Г. Руденко, А.А. Бессонов, С.О. Руденко // Автоматизация: проблемы, идеи, решения : междунар. науч.-техн. конф. : тезисы докл. – Севастополь, 6-10 сентября 2010. – С. 9–11.

36. Руденко О.Г. Робастная нейросетевая идентификация нелинейных динамических объектов / О.Г. Руденко, А.А. Бессонов // Автоматика-2010 :

международ. конф. по автоматическому управлению : тезисы докл. – Харьков, 27-29 сентября 2010. – С. 153–154.

37. Бессонов А.А. Идентификация нелинейных объектов с помощью эволюционной радиально-базисной сети / А.А. Бессонов, С.О. Руденко // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління : міжнар. наук.-техн. конф. : тези доп. – Київ-Харків, 15-16 грудня 2011. – С. 55.

38. Бессонов А.А. Определение структуры радиально-базисной сети с помощью генетических алгоритмов / А.А. Бессонов // Інформаційні технології в навігації і управлінні: стан та перспективи розвитку : міжнар. наук.-техн. конф. : тези доп. – Київ, 16-17 липня 2011. – С. 36.

39. Руденко О.Г. М-обучение искусственных нейронных сетей / О.Г. Руденко, А.А. Бессонов // Обчислювальний інтелект (результати, проблеми, перспективи) : міжнар. наук.-техн. конф. : тези доп. – Черкаси, 10-13 травня 2011. – С. 46–47.

40. Бессонов А.А. Использование асимметричных функционалов для обучения радиально-базисных сетей / А.А. Бессонов, О.Г. Руденко, С.О. Руденко // Автоматизация: проблемы, идеи, решения : международ. науч.-техн. конф. : тезисы докл. – Севастополь, 5-9 сентября 2011. – С. 8–9.

41. Бессонов А.А. Применение механизма адаптивных мутаций при эволюционном обучении радиально-базисных сетей / А.А. Бессонов, С.О. Руденко // Автоматизация: проблемы, идеи, решения : международ. науч.-техн. конф. : тезисы докл. – Севастополь, 3-7 сентября 2012. – С. 184–185.

42. Руденко О.Г. Робастная идентификация нелинейных объектов с помощью эволюционного многослойного персептрона / О.Г. Руденко, А.А. Бессонов, С.О. Руденко // Информатика, математическое моделирование, экономика : международ. науч.-практ. конф. : тезисы докл. – Смоленск, 20 апреля 2012. – С. 7–13.

43. Руденко О.Г. Робастная многокритериальная идентификация нелинейных динамических объектов с помощью эволюционирующей РБС / О.Г. Руденко, А.А. Бессонов // Современные направления развития информационно-коммуникационных технологий и средств управления : международ. науч.-техн. конф. : тезисы докл. – Полтава-Белгород-Харьков-Киев-Кировоград, 11-12 апреля 2013. – С. 47.

44. Руденко О.Г. Адаптивное масштабирование шага дискретизации входных данных при аппроксимации функций с помощью РБФ / О.Г. Руденко, А.А. Бессонов, Р.В. Бобнев // Информатика, математическое моделирование, экономика : международ. науч.-практ. конф.: тезисы докл. – Смоленск, 24-26 апреля 2013. – Т.2. – С. 135–139.

45. Бессонов А.А. Многокритериальная оптимизация искусственных нейронных сетей прямого распространения / А.А. Бессонов, С.О. Руденко // Стратегия качества в промышленности и образовании : международ. конф. : тезисы докл. – Варна (Болгария), 31 мая-7 июня 2013. – С. 443–445.

46. Бессонов А.А. Построение робастных моделей технологических процессов производства сахара с помощью эволюционирующих радиально-

базисных сетей / А.А. Бессонов, С.О. Руденко // Автоматизация: проблемы, идеи, решения : междунар. науч.-техн. конф. : тезисы докл. – Севастополь, 2013. – С. 170–172.

47. Бессонов А.А. Многокритериальная оптимизация эволюционирующей радиально-базисной сети / А.А. Бессонов // Проблемы информатизации : міжнар. наук.-техн. конф. : тези доп. – Черкаси-Київ-Голыятті-Полтава, 19-20 грудня 2013. – С. 29.

48. Бессонов А.А. Применение эволюционирующей радиально-базисной сети в системах управления с прогнозирующей моделью / А.А. Бессонов // Современные направления развития информационно-коммуникационных технологий и средств управления : междунар. науч.-техн. конф. : тезисы докл. – Полтава-Баку-Белгород-Кировоград-Харьков, 4-5 декабря 2014. – С. 22.

49. Руденко О.Г. Нейроэволюционные методы обработки информации и управления / О.Г. Руденко, А.А. Бессонов, В.И. Быков // Стратегічні рішення інформаційного розвитку економіки, суспільства та бізнесу : міжнар. наук.-практ. конф. : тези доп. – Рівне, 12-14 лютого 2014. – С. 36-37.

50. Бессонов А.А. Прогнозирующее нейроэволюционное управление нелинейными динамическими объектами / А.А. Бессонов // Проблемы информатизации : междунар. науч.-техн. конф. : тезисы докл. – Киев-Полтава-Катовице-Париж-Белгород-Черкаси-Харьков, 12-13 апреля 2014. – С. 73-74.

51. Руденко О.Г. Эволюционирующие нейронные сети прямого распространения / О.Г. Руденко, А.А. Бессонов, С.О. Руденко // Обчислювальний інтелект (результати, проблеми, перспективи) : міжнар. наук.-практ. конф. : тези доп. – Київ-Черкаси, 12-15 травня 2015. – С. 106.

52. Бессонов А.А. Нейроуправление с прогнозирующей моделью / А.А. Бессонов // Информатика, математическое моделирование, экономика : междунар. науч.-практ. конф. : тезисы докл. – Смоленск, 11-15 мая 2015. – Т. 1. – С. 18–22.

53. Бессонов А.А. Коэволюция ансамблей нейронных сетей прямого распространения / А.А. Бессонов // Проблемы информатизации : міжнар. наук.-техн. конф. : тези доп. – Черкаси-Баку-Бельсько-Бяла-Полтава, 12-13 листопада 2015. – С. 23.

54. Бессонов А.А. Нечеткий регулятор скорости травления стали / А.А. Бессонов, О.О. Илюнин, А.В. Илюнин, // Современные направления развития информационно-коммуникационных технологий и средств управления : междунар. науч.-техн. конф. : тезисы докл. – Полтава-Баку-Кировоград-Харьков, 21-22 апреля 2016. – С. 25.

55. Руденко О.Г. Прогнозирование экономических временных рядов на основе коэволюционного подхода / О.Г. Руденко, А.А. Бессонов // Общество и экономическая мысль в XXI в.: пути развития и инновации : междунар. науч.-практ. конф. : тезисы докл. – Воронеж, 31 марта 2016. – С. 544–549.

56. Руденко О.Г. Прогнозирование экономических процессов с помощью эволюционирующих искусственных нейронных сетей / О.Г. Руденко, А.А. Бессонов // Проблемы та перспективи розвитку ІТ-індустрії : міжнар.

## АНОТАЦІЯ

**Безсонов О.О. Еволюційні штучні нейронні мережі прямого розповсюдження: архітектури, навчання, застосування. – На правах рукопису.**

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту. – Харківський національний університет радіоелектроніки, Міністерство освіти і науки України, Харків, 2017.

У роботі проведено аналіз проблеми розвитку теоретичних основ інтелектуального аналізу даних та створення нових еволюційних ШНМ прямого розповсюдження задля підвищення ефективності обробки інформації в умовах апріорної та поточної невизначеності; розроблено новий метод робастної багатокритеріальної оптимізації (Парето-оптимізації) на основі робастних фітнес-функцій та інформаційних критеріїв оцінки складності моделі; запропоновано нові методи навчання ШНМ, що забезпечують необхідну точність навчання при наявності обмежених симетричних та асиметричних завдань; розроблено нові процедури корекції параметрів функціоналів, що використовуються при навчанні, та оцінювання параметрів задачі, що описується моделлю Тьюки-Хьюбера; розроблено нові закони адаптивного прогнозуючого нейроуправління нелінійними нестационарними динамічними об'єктами, що функціонують в умовах невизначеності.

У середовищі *NeurophStudio* проведено імітаційне моделювання різних процедур навчання ЕШНМ; досліджено процес вирішення задач ідентифікації та прогнозуючого управління нелінійними динамічними об'єктами, розроблено програмні засоби, які реалізують запропоновані методи побудови еволюційних ШНМ прямого розповсюдження. Проведено експериментальні дослідження властивостей і характеристик розроблених методів, які підтвердили основні положення дисертації.

Достовірність результатів підтверджується експериментальними дослідженнями та впровадженнями.

**Ключові слова:** еволюція, штучна нейронна мережа, базисна функція, генетичний алгоритм, процедура навчання, хромосома, мутація, ідентифікація, нелінійний динамічний об'єкт.

## АННОТАЦИЯ

**Бессонов А.А. Эволюционирующие искусственные нейронные сети прямого распространения: архитектуры, обучение, применения. – На правах рукописи.**

Диссертация на соискание ученой степени доктора технических наук по специальности 05.13.23 – системы и средства искусственного интеллекта. – Харьковский национальный университет радиоэлектроники, Министерство об-



разования и науки Украины, Харьков, 2017.

В диссертационной работе, посвященной решению важной научно-практической проблемы развития теоретических основ интеллектуального анализа данных и созданию новых эволюционирующих ИНС прямого распространения с целью повышения эффективности обработки информации в условиях априорной и текущей неопределенности. Результаты проведенного анализа показали, что в качестве базиса для синтеза подобных систем обработки информации целесообразно использовать статические ИНС прямого распространения, а для улучшения их свойств необходимо применение методов теории адаптации и эволюции.

Разработан новый метод робастной многокритериальной оптимизации (Парето-оптимизации) на основе робастных фитнес-функций и информационных критериев оценки сложности модели.

Предложены новые робастные методы обучения ИНС, дающие возможность эффективно обрабатывать информацию при наличии помех с асимметричными распределениями. Также разработаны новые процедуры коррекции параметров используемых при обучении функционалов и оценивания параметров помехи, описываемой моделью Тьюки-Хьюбера, что позволяет при отсутствии априорной информации о статистических свойствах помех адаптивно корректировать параметры ИНС.

Разработаны новые законы адаптивного прогнозирующего нейроуправления нелинейными нестационарными динамическими объектами, функционирующими в условиях неопределенности. Для коррекции эталонной траектории предложено использовать эволюционный подход.

Предложены новые методы кусочно-линейной аппроксимации гауссовских базисных функций в РБС, которая позволяет существенно упростить вычисления, сопутствующие процессам построения нейросетевой модели исследуемых объектов и управления ими.

Получил дальнейшее развитие метод гибридного обучения ИНС. Модификация метода состоит в использовании эволюционного алгоритма для грубой настройки параметров сети и разработанной робастной рекуррентной процедуры Левенберга-Марквардта для окончательной тонкой настройки.

Получил дальнейшее развитие эволюционный метод устранения влияния помех при определении структур и параметров нейросетевых моделей за счет использования оценок параметров модели помехи Тьюки-Хьюбера в процедуре М-обучения.

Получили дальнейшее развитие коэволюционирующие ИНС прямого распространения, в которых для решения задачи многокритериальной оптимизации использован алгоритм кластеризации. Проведено имитационное моделирование разработанных архитектур и методов обучения эволюционирующих ИНС. Показаны их преимущества перед известными нейросетевыми системами обработки информации и методами их обучения по точности, устойчивости и быстродействию в задачах аппроксимации (эмуляции), идентификации нелинейных нестационарных процессов и

интеллектуального управления нелинейными нестационарными динамическими объектами в условиях априорной и текущей неопределенности.

Решены практические задачи построения системы измерения температурных полей вращающихся трубчатых печей в ООО „Побужский фероникелевый комбинат“; управления процессом травления полосовой стали на металлургических предприятиях Украины (АТ «Співдружність-Т»); разработки АСУ ТП диффузии, дефекосатурации, выпаривания и кристаллизации в ООО “Кириковский сахарный завод”. Достоверность полученных результатов диссертационной работы подтверждается экспериментальными исследованиями и результатами внедрения.

**Ключевые слова:** эволюция, искусственная нейронная сеть, базисная функция, генетический алгоритм, процедура обучения, хромосома, мутация, идентификация, нелинейный динамический объект.

## ABSTARCT

**Bezsonov O.O. Evolving artificial feedforward neural networks: architecture, training, applications. – As the Manuscript.**

A Thesis for a Doctor of Technical Sciences degree in the specialty 05.13.23 – systems and means of artificial intelligence. – Kharkiv National University of Radio Electronics, Ministry education and science of Ukraine, Kharkiv, 2017.

The thesis analyzes the problems of intellectual data mining theoretical foundations and creation of new evolutionary artificial neural networks (EANN) for improving the efficiency of information processing in conditions of priori and current uncertainty. A new method of robust multi-objective optimization (Pareto optimization) based on robust fitness functions and information criteria for assessing the complexity of the model, is developed. New methods of training EANN that ensure the required accuracy in the presence of limited symmetric and asymmetric noise are proposed. New procedures for the functional parameters correction and evaluating parameters of noise that is described by Tukey-Huber model are developed. New laws of adaptive neuro-predictive control of nonlinear non-stationary dynamic objects that operate under uncertainties are designed.

Simulation of different EANN training algorithms was performed in *NeurophStudio* environment. The process of identifying and solving problems of nonlinear dynamic objects predictive control is researched. Software tools that implement the proposed methods for constructing EANN are developed. Experimental research of developed methods properties and characteristics, that confirmed the fundamentals of the thesis, is performed.

The validity of the obtained results is confirmed by the experimental studies and real applications.

**Keywords:** evolution, artificial neural network, basis function, genetic algorithm, training algorithm, chromosome, mutation, identification, nonlinear dynamical object.

