Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

ТИЩЕНКО ОЛЕКСІЙ КОСТЯНТИНОВИЧ

УДК 004.032.26

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ БАГАТОВИМІРНИХ НЕЛІНІЙНИХ ЧАСОВИХ РЯДІВ НА ОСНОВІ ГЕТЕРОГЕННИХ НЕЙРОННИХ МЕРЕЖ

05.13.23 – системи та засоби штучного інтелекту

Автореферат дисертації на здобуття наукового ступеня
кандидата технічних наук

Харків – 2013

Дисертацією є рукопис.

Роботу виконано в Харківському національному університеті радіоелектроніки Міністерства освіти і науки України.

Науковий керівник — доктор технічних наук, професор
Бодянський Євгеній Володимирович,
Харківський національний університет радіоелектроніки, професор кафедри штучного інтелекту

Офіційні опоненти: — доктор технічних наук, доцент
Литвиненко Володимир Іванович,
Херсонський національний технічний університет, професор кафедри інформатики та комп'ютерних технологій

доктор технічних наук, доцент
Пелешко Дмитро Дмитрович,
Національний університет «Львівська політехніка», професор кафедри інформаційних технологій видавничої справи

Захист відбудеться «*29*» *травня* 2013 р. о *15*^{00} годині на засіданні спеціалізованої вченої ради Д 64.052.01 Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, пр. Леніна, 14.

З дисертацією можна ознайомитися у бібліотеці Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, пр. Леніна, 14.

Автореферат розісланий «*24*» *квітня* 2013 р.

Учений секретар
спеціалізованої вченої ради — Є.І. Литвинова

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Задача обробки багатовимірних часових рядів досить часто виникає у багатьох технічних, медико-біологічних, соціально-економічних дослідженнях, де від якості отримуваних рішень істотно залежить точність синтезованих результатів. Треба відзначити, що у багатьох реальних завданнях часові ряди характеризуються високим рівнем нелінійності (можливо навіть хаотичності) та нестаціонарності своїх параметрів, наявністю нерегулярних трендів, стрибків та аномальних викидів. Зрозуміло, що традиційні методи аналізу часових послідовностей, що базуються на регресійному, кореляційному, спектральному та інших подібних підходах, що мають на увазі апріорну наявність достатньої вибірки спостережень, виявляються неефективними. Альтернативою традиційним статистичним методам може слугувати математичний апарат обчислювального інтелекту й, перш за все, штучні нейронні мережі та нейро-фаззі-системи завдяки своїм універсальним апроксимуючим властивостям. Разом з тим з апроксимуючих властивостей зовсім не випливають екстраполюючі, оскільки урахування давньої передісторії для побудови, наприклад, прогнозувальної моделі може погіршити якість прогнозу. У зв’язку з цим, обробляючи істотно нестаціонарні процеси, треба відмовитись від процедур навчання, заснованих на зворотному поширенні похибок або методі найменших квадратів, щоб скористатися процедурами, що ґрунтуються на локальних критеріях та «короткій» пам’яті на зразок алгоритму Качмажа-Уідроу-Хоффа. При цьому застосовані методи навчання мають забезпечувати не тільки високу швидкодію, але й фільтрувальні властивості для придушення стохастичної «шумової» компоненти в оброблюваному сигналі. У зв’язку з цим синтез спеціалізованих гетерогенних систем обчислювального інтелекту, що призначаються для вирішення задач обробки істотно нестаціонарних багатовимірних часових рядів за умов невизначеності та забезпечують разом з високою швидкістю навчання фільтрацію завад, є досить цікавою та перспективною задачею.

У розвиток зазначеного напрямку інтелектуального аналізу багатовимірних нелінійних часових рядів на основі гетерогенних нейронних мереж найбільш істотний внесок зробили J.C. Bezdek, E. Keogh, J. Abonyi, R. Babuska, H. Jaeger, W. Maass, D. Verstraeten, F. Wyffels, T. Yamakawa, T. Sugeno, Руденко О.Г., Литвиненко В.І., Пелешко Д.Д., Михальов О.І., Кондратенко Ю.П., Зайченко Ю.П.

Методи, що широко використовуються зараз для обробки істотно нестаціонарних багатовимірних часових рядів, мають ряд недоліків, а саме: радіально-базисна нейронна мережа схильна до «прокльону розмірності», а адаптивна нейро-фаззі система нечіткого виведення має громіздку архітектуру та досить повільно навчається за рахунок використання алгоритму зворотного поширення похибок. А це веде до збільшення кількості нейронів і кількості нечітких правил у базі знань, звідки випливає необхідність збільшення обсягу навчальної вибірки для налаштування подібної системи. Як було зазначено вище, нейро-фаззі системи, що використовують для навчання алгоритм зворотного поширення похибок, характеризуються низькою швидкістю, що робить їх застосування при обробці даних у послідовному режимі не дуже вдалим. Уникнути цих недоліків можна за рахунок використання гетерогенних нейронних мереж, що поєднують як теорію штучних нейронних мереж, яка дозволяє отримати універсальні апроксимувальні властивості і здатність навчатися, так і теорію нечіткої логіки, яка дає

можливість наділяти систему лінгвістичною інтерпретованістю.

Незважаючи на велику кількість наукових робіт, все ще існує задача обробки багатовимірних часових рядів, викликана потребою у покращенні якості отриманих результатів та збільшенні швидкості обробки даних. У зв'язку з цим актуальною є задача розробки методів обробки багатовимірних часових рядів на основі гетерогенних нейронних мереж для послідовної обробки нелінійних багатовимірних часових рядів, що здатні функціонувати за умов дефіциту апріорної інформації та забезпечують можливість обробки часових рядів із короткою та довгою навчальними вибірками, а також характеризуються підвищеною швидкістю навчання. Згідно з усім сказаним вище, робота є актуальною, що і визначає перспективність отримання як теоретичних, так і практичних результатів.

Зв'язок теми дисертації з планами наукових робіт. Роботу виконано на кафедрі штучного інтелекту Харківського національного університету радіоелектроніки відповідно до плану науково-дослідних робіт у рамках держбюджетних тем: «Синтез методів обробки інформації за умов невизначеності на основі самонавчання та м'яких обчислень» (№ДР 0107U003028), «Еволюційні гібридні системи обчислювального інтелекту зі змінною структурою для інтелектуального аналізу даних» (№ДР 0110U000458), які виконувались згідно наказів Міністерства освіти і науки України за результатами конкурсного відбору проектів наукових досліджень. У межах наведених тем здобувач як виконавець запропонував архітектуру прогнозувальної нейро-фаззі мережі, що базується на теорії резервуарних обчислень, прогнозувальну архітектуру, що складається з шару нечіткої кластеризації та нейро-нео-фаззі мережі зустрічного поширення, а також методи навчання таких систем.

Мета і задачі дослідження. Метою дисертаційної роботи є розробка архітектур гетерогенних нейронних мереж та методів їх навчання, що відзначаються високою швидкістю навчання параметрів, лінійною залежністю параметрів вихідного шару мережі від синаптичних ваг прихованого шару, а також здатні працювати за умов дефіциту апріорної інформації (короткої навчальної вибірки) для розв'язання завдань обробки багатовимірних нелінійних часових рядів.

Відповідно до поставленої мети у дисертаційній роботі розв'язуються такі задачі:

– розробка гетерогенних нейронних архітектур для прогнозування нелінійних багатовимірних часових рядів, що ґрунтуються на нейромережах зустрічного поширення для підвищення швидкості прогнозування часових рядів;

– розробка гетерогенних нейронних архітектур для прогнозування нелінійних багатовимірних часових рядів, що ґрунтуються на парадигмі резервуарних обчислень;

– розробка методів навчання запропонованих гетерогенних нейронних архітектур;

– розробка нейро-фаззі методів сегментації багатовимірних часових рядів;

– імітаційне моделювання розроблених методів та архітектур і розв'язання з їх допомогою реальних задач.

Об'єктом дослідження є процес інтелектуального аналізу й обробки даних, що представлено в числовій формі у вигляді часових рядів.

Предметом дослідження є методи та архітектури інтелектуального аналізу багатовимірних нелінійних часових рядів на основі гетерогенних нейронних мереж за умов невизначеності.

Методи дослідження. Теорія штучних нейронних мереж – для аналізування та

синтезу нових архітектур гетерогенних нейронних мереж, удосконалення методів їх навчання; теорія обчислювального інтелекту та м'яких обчислень – для визначення контексту дослідження; теорія нечіткої логіки – для побудови гібридних нечітких систем; інтелектуальний аналіз даних – для знаходження прихованих залежностей в інформації; імітаційне моделювання – для визначення ефективності застосування розроблених систем.

Наукова новизна отриманих результатів.

1. Вперше запропоновано архітектуру гетерогенної нейронної мережі для прогнозування даних, що складається з шару нечіткої кластеризації та нейро-нео-фаззі мережі зустрічного поширення, яка характеризується підвищеною швидкодією та покращеними апроксимувальними властивостями, що дає можливість обробляти нелінійні часові ряди довільної природи за умов невизначеності, а також запропоновано метод її навчання.

2. Вперше запропоновано багатовимірну модифікацію нео-фаззі-нейрону, що має покращені апроксимувальні властивості, та методи його навчання, які характеризуються простотою реалізації, що дозволяє забезпечити меншу обчислювальну складність за рахунок суттєвого скорочення кількості функцій належності.

3. Вперше запропоновано архітектуру нейро-фаззі мережі на основі теорії резервуарних обчислень, яка характеризується підвищеною швидкістю налаштування параметрів навчання й точністю прогнозу часового ряду, та метод її навчання, який характеризується простотою реалізації, що дозволяє істотно покращити апроксимувальні та екстраполювальні властивості системи для обробки нелінійних часових рядів за умов невизначеності.

4. Набула подальшого розвитку архітектура нейро-компресора даних на основі багатошарового персептрону «вузьке місце» за рахунок приєднання до мережі-компресора нейро-нео-фаззі мережі зустрічного поширення, що дозволило здійснювати прогнозування часових рядів зниженої розмірності без втрати внутрішніх зв'язків збережуваних даних.

5. Набув подальшого розвитку метод сегментації багатовимірних часових рядів на основі методу нечіткої кластеризації Гата-Геви шляхом спеціального налаштування цетроїдів, що дозволяє здійснювати сегментацію багатовимірних часових рядів у режимі online.

Практичне значення отриманих результатів. Розроблені методи та архітектури гетерогенних нейронних мереж для обробки багатовимірних часових рядів за умов апріорної невизначеності, що базуються на нейромережевих та нейро-фаззі методах і архітектурах, доведено до програмної реалізації, що дозволило автоматизувати процес обробки багатовимірних часових рядів для розв'язання задач прогнозування, компресії та сегментації. Запропоновані методи та архітектури, покладені в основу розроблених програмних модулів, дали змогу підвищили швидкість обробки даних (5-6%) у задачах прогнозування у режимі їх послідовного надходження та високу точність визначення моментів змін властивостей у задачах сегментації, що підтвердило доцільність дослідження для розв'язання практичних задач.

Результати дисертаційної роботи впроваджено у ТОВ «НВП «Енергобуд», м. Харків, при розв'язанні задач прогнозування споживання електроенергії в автоматизованих системах диспетчерського керування (акт впровадження від

19.03.2012), що дозволило підвищити точність прогнозування у порівнянні із широковживаними методами прогнозування та знизити втрати, що виникають внаслідок похибок прогнозування. Також було розв'язано задачу сегментування даних електрокардіограм серед пацієнтів ТОВ «Небозвід», м. Харків, з метою прогнозування виникнення порушень серцевого ритму у хворих на гіпертонічну та ішемічну хвороби серця та відповідне корегування дій медичного персоналу при наданні невідкладної медичної допомоги відповідним групам хворих (акт впровадження від 04.07.2012).

Результати дисертаційної роботи також впроваджено у навчальному процесі на кафедрі штучного інтелекту Харківського національного університету радіоелектроніки у дисциплінах «Нейромережеві методи обчислювального інтелекту», «Інтелектуальний аналіз даних» (акт впровадження від 10.09.2012), а також у науково-дослідну роботу Харківського національного університету радіоелектроніки (акт впровадження від 02.07.2012).

Особистий внесок здобувача. Всі результати дисертації автор отримав особисто. У роботах, опублікованих із співавторами, здобувачу належать: [1] - метод прогнозування стиснутих у результаті компресії даних; [2] - метод сегментування, що базується на методі нечіткої кластеризації; [3] - архітектура шару нео-фаззі-нейронів; [4] - архітектура резервуарної нейро-фаззі мережі; [5] - архітектура багатовимірного нео-фаззі-нейрону; [8] - метод навчання нейро-нео-фаззі мережі; [9] - використання різних типів функцій належності в нео-фаззі-нейронах; [10] - використання різних типів нечітких кластеризаторів у нейро-фаззі мережі; [11] - архітектура нейро-фаззі мережі для прогнозування на основі резервуарних обчислень; [12] - використання рекурентного методу нечіткої кластеризації у нейро-нео-фаззі мережі; [14] - метод навчання багатовимірного нео-фаззі-нейрону; [15] - метод навчання нейро-фаззі предиктору стислих даних; [16] - збільшення швидкості обробки інформації у предикторі стислих даних; [18] - метод навчання модифікованого нео-фаззі-нейрону.

Апробація результатів дисертації. Результати дисертаційної роботи доповідалися й обговорювалися на 13-му Міжнародному молодіжному форумі «Радіоелектроніка та молодь у XXI столітті» (м. Харків, Україна, 2009); на 16-й Міжнародній конференції з автоматичного управління «Автоматика» (м. Чернівці, Україна, 2009); на Міжнародній науково-практичній конференції «Інформаційні технології та інформаційна безпека в науці, техніці та освіті «ІНФОТЕХ» (м. Севастополь, Україна, 2009, 2011); на Міжнародній науковій конференції «Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту» (м. Євпаторія, Україна, 2009, 2010, 2011); на 5-й Міжнародній школі-семінарі «Теорія прийняття рішень» (м. Ужгород, Україна, 2010); на 28-й, 29-й, 30-й, 31-й Науково-технічних конференціях «Моделювання» (м. Київ, 2009, 2010, 2011, 2012); на 4-й, 5-й Міжнародних науково-технічних конференціях «Сучасні комп'ютерні системи і мережі: розробка та використання» (м. Львів, 2009, 2011).

Публікації. За результатами досліджень опубліковано 19 наукових праць, з яких 5 статей (серед яких 2 статті у журналах та 3 у збірках наукових статей) у наукових фахових виданнях України з технічних наук та 14 публікацій (5 одноосібно) у збірниках праць міжнародних наукових конференцій і семінарів.

Структура й обсяг дисертаційної роботи. Дисертація складається із вступу, п'яти розділів, висновків, переліку використаних джерел та додатків. Загальний обсяг дисертації становить 130 сторінок. Робота містить 22 рисунки, 5 таблиць, 1 додаток на 4

сторінках та перелік використаних джерел із 140 найменувань на 14 сторінках.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету і задачі дослідження, наведено відомості щодо наукової новизни отриманих у дисертації результатів, визначено їх практичну цінність, наведено відомості про апробацію та впровадження результатів.

Перший розділ містить огляд проблемної області і постановку задач дисертаційного дослідження. Проведено огляд мереж зустрічного поширення, їх методів навчання. Докладно описано резервуарні обчислення, способи побудови резервуарів та методи їх навчання.

У результаті проведеного аналізу сформульовано основні завдання наукового дослідження, викладеного у дисертаційній роботі.

У **другому розділі** наведено архітектуру гетерогенної нейронної мережі для прогнозування даних, що складається з шару нечіткої кластеризації та нейро-нео-фаззі мережі зустрічного поширення, яка характеризується підвищеною швидкодією та покращеними апроксимувальними властивостями, що дає можливість обробляти нелінійні часові ряди довільної природи за умов невизначеності, а також запропоновано метод її навчання. Докладно описано архітектуру нео-фаззі-нейрону (NFN). NFN містить нелінійні синапси NS_i, в яких реалізуються елементарні правила нечіткого висновування вигляду

$$\text{IF } x_i(k) \text{ IS } X_i \text{ THEN } f_i(x_i(k)) = \sum_{j=1}^{h} \mu_{ij}(x_i(k)) w_{ij}, \ i = 1 \ldots n, \tag{1}$$

де X_i – лінгвістичне значення (нечітка множина) на i-му вході, $x_i(k)$ – вхідний сигнал нео-фаззі-нейрону, $\mu_{ij}(x_i(k))$ – функція належності, w_{ij} – синглтон, h – кількість функцій належності в кожному нелінійному синапсі, k – поточний дискретний час.

При подачі на вхід NFN сигналу $x_i(k)$ на його виході з'явиться скалярне значення

$$\hat{x}_i(k) = \sum_{i=1}^{n} f_i(x_i(k)) = \sum_{i=1}^{n} \sum_{j=1}^{h} \mu_{ij}(x_i(k)) w_{ij}(k-1), \tag{2}$$

що визначається як функціями належності μ_{ij}, так і поточними значеннями настроюваних синаптичних ваг $w_{ij}(k-1)$. У якості функцій належності у NFN зазвичай використовуються трикутні конструкції, при цьому вважається, що всі вхідні дані попередньо закодовано на інтервал $x_i \in [0,1]$.

Важливо відмітити, що така конструкція функцій належності автоматично забезпечує розбиття Руспіні (одиничне розбиття):

$$\sum_{j=1}^{h} \mu_{ij}(x_i) = 1. \tag{3}$$

Запропоновано архітектуру нової прогнозувальної гібридної нейронної мережі (рис.

1), що є поєднанням нечіткої кластеризуючої мережі Кохонена та шару NFN (нейро-нео-фаззі мережа). Тут $x(k)=(x_1(k),x_2(k),\ldots,x_n(k))^T$ – спостереження, що надходять до входів мережі, $y(k)$ – вихід мережі, $k=1,2,\ldots,N,\ldots$ – поточний дискретний час.

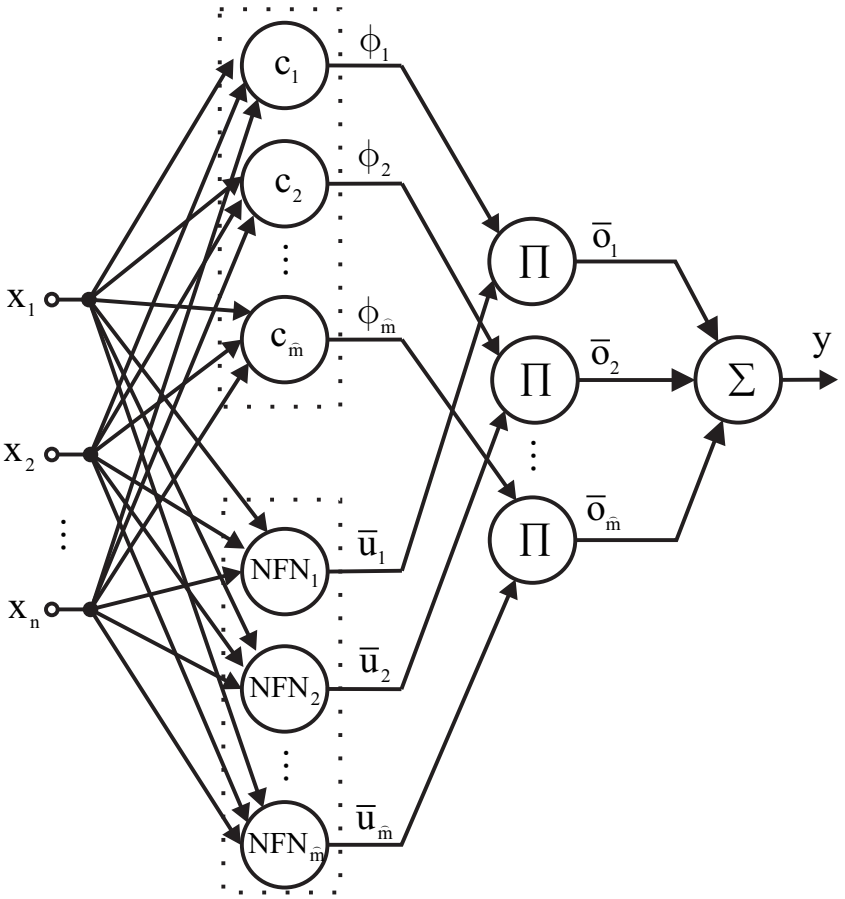

Рисунок 1 - Архітектура нейро-нео-фаззі мережі

Робота мережі полягає в обчисленні рівней належності поточного спостереження до кластерів, які визначаються вагами нейронів шару Кохонена, обчисленні значень вихідних сигналів NFN та ранжуванні цих значень у відповідності до ступенів належності на виходах шару Кохонена. Виходом мережі є зважена сума виходів NFN.

Виходи нейронів шару Кохонена обчислюються у вигляді

$$\phi_l(k)=\frac{\Phi_l(k)}{\sum_{p=1}^{\tilde{m}}\Phi_p(k)}, \qquad (4)$$

де $\phi_l(k)=\Phi(\|x(k)-c_l\|^2_{R_l^{-1}})$ – функції належності нейронів (функції сусідства), $\|\cdot\|_{R_l^{-1}}$ – норма вектору за матрицею R_l^{-1}. Таким чином, виходи нейронів шару Кохонена задовольняють умові

$$\sum_{l=1}^{\tilde{m}}\phi_l(k)=1. \qquad (5)$$

Треба відмітити, що функції належності NFN рівномірно розподілені на інтервалі $[0,1]$. Позначимо

$$h_\Sigma=\sum_{i=1}^{n}h_i, \qquad (6)$$

$$\mu(k)=(\mu_{11},\mu_{12},\ldots,\mu_{1h_1},\mu_{21},\mu_{22},\ldots,\mu_{2h_2},\ldots,\mu_{n1},\mu_{n2},\ldots,\mu_{nh_n})^T \qquad (7)$$

– вектор функцій належності NFN розмірності $(h_\Sigma\times 1)$ (тут параметр $x_i(k)$ у μ_{ij} пропущено для стислості запису) та

$$W^T = \begin{pmatrix} w_{11}^{[1]} & w_{21}^{[1]} & \cdots & w_{nh_n}^{[1]} \\ w_{11}^{[2]} & w_{21}^{[2]} & \cdots & w_{nh_n}^{[2]} \\ \cdots & \cdots & \ddots & \cdots \\ w_{11}^{[\tilde{m}]} & w_{21}^{[\tilde{m}]} & \cdots & w_{nh_n}^{[\tilde{m}]} \end{pmatrix} = (w^{[1]}, w^{[2]}, \ldots, w^{[\tilde{m}]})^T \tag{8}$$

– матрицю ваг NFN розмірністю $(\tilde{m} \times h_\Sigma)$.

Для простоти можна прийняти $h_1 = h_2 = \cdots = h_n = h$, тоді $h_\Sigma = hn$.

Враховуючи (7) та (8), вихід шару нео-фаззі-нейронів можна записати у векторній формі як

$$\bar{u}(k) = W^T(k)\mu(k), \tag{9}$$

а вихід всієї мережі обчислюється згідно виразу:

$$y(k) = \sum_{l=1}^{\tilde{m}} \bar{o}_l(k) = \sum_{l=1}^{\tilde{m}} \bar{u}_l(k)\phi_l(k) = \phi^T(k)\bar{u}(k), \tag{10}$$

де $\phi(k) = (\phi_1(k), \phi_2(k), \ldots, \phi_{\tilde{m}}(k))^T$, $\bar{u}(k) = (\bar{u}_1(k), \bar{u}_2(k), \ldots, \bar{u}_{\tilde{m}}(k))^T$.

Нехай навчальну вибірку задано набором $X = (x(1), x(2), \ldots, x(N))$, де X – матриця вхідних значень сигналів для навчання мережі, Y – вихідні скалярні значення.

Процедура навчання запропонованої нейро-нео-фаззі мережі складається з двох етапів: на першому етапі відбувається настроювання параметрів шару Кохонена, на другому етапі виконується навчання NFN, враховуючи сигнали на виходах нейронів шару Кохонена. Така процедура навчання узгоджується з підходом до навчання мережі зустрічного поширення.

Навчання шару Кохонена відбувається в режимі самонавчання, тобто без використання Y в якості навчального сигналу. Для знаходження центрів $c_l(k)$ нейронів цього шару запропоновано використовувати рекурентний метод кластеризації Бездека нечітких с-середніх у модифікації для навчання нечіткої кластеризуючої мережі Кохонена:

$$\begin{cases} \phi_l(k+1) = \dfrac{\left(\| x(k+1) - c_l(k) \|^2_{R_l^{-1}}\right)^{\frac{1}{1-\beta}}}{\sum\limits_{p=1}^{\tilde{m}} \left(\| x(k+1) - c_p(k) \|^2_{R_p^{-1}}\right)^{\frac{1}{1-\beta}}}, \\ c_l(k+1) = c_l(k) + \eta_c(k)\phi_l^\beta(k+1)(x(k+1) - c_l(k)), \; l = 1, 2, \ldots, \tilde{m}, \end{cases} \tag{11}$$

де β – невід'ємний параметр, що називається «фаззіфікатором» та визначає нечітку межу між кластерами з центроїдами у точках $c_l(k)$, що відповідають нейронам шару Кохонена; $\eta_c(k)$ – параметр кроку, що визначає швидкість навчання; $\phi_l^\beta(k)$ – функції сусідства.

На другому етапі для навчання параметрів NFN використовується процедура рекурентного методу найменших квадратів. Навчальним сигналом для кожного NFN є скалярне значення на виході мережі, а настроювання параметрів виконується з

урахуванням виходу відповідного нейрону шару Кохонена в якості вагового коефіцієнту.

Процедуру навчання можна записати в рекурентному вигляді:

$$\begin{cases} w^{[1]}(k) = w^{[1]}(k-1) + \dfrac{P_1(k-1)(\tilde{y}(k) - \mu^T(k)w^{[1]}(k-1))}{1+\phi_1(k)\mu^T(k)P_1(k-1)\mu(k)}\mu(k), \\ P_1(k) = \phi_1(k)\left(P_1(k-1) - \dfrac{\phi_1(k)P_1(k-1)\mu(k)\mu^T(k)P_1(k-1)}{1+\phi_1(k)\mu^T(k)P_1(k-1)\mu(k)}\right), \end{cases} \tag{12}$$

де $P_1(k)$ – матриця розмірності $(hn \times hn)$.

Далі введено рекурентну процедуру нечіткої кластеризації у нейро-нео-фаззі мережі зустрічного поширення. Для сегментації часової послідовності $x = \left(x(1), x(2), \ldots, x(k), \ldots, x(N)\right)^T$ використовується підхід, що базується на непрямій кластеризації послідовності. Згідно цього підходу виділяються декотрі характеристики для подальшого відображення у перетворений простір ознак. Таким чином, отримуємо новий вектор ознак $x'(k)$, що містить вираховані вказані вище оцінки на кожному кроці дискретного часу k і містить N n-вимірних векторів ознак $x' = \left(x'(1), x'(2), \ldots, x'(N)\right)^T$, де $k = 1, 2, \ldots, N$.

Використовуваний метод належить до класу процедур, що базуються на цільових функціях, розроблених для вирішення завдань кластеризації шляхом оптимізації визначеного критерію, що його було задано заздалегідь.

Введемо до розгляду цільову функцію

$$E\left(\bar{\mu}_j(k), c_j\right) = \sum_{k=1}^{N}\sum_{j=1}^{h}\bar{\mu}_j^{\beta} D^2\left(x(k), c_j\right), \tag{13}$$

та обмеження

$$\sum_{j=1}^{h}\bar{\mu}_j(k) = 1,\ k = 1 \ldots N, \tag{14}$$

$$0 < \sum_{k=1}^{N}\bar{\mu}_j(k) \le N,\ j = 1 \ldots h, \tag{15}$$

де N – кількість елементів у вибірці, h – кількість центроїдів, $\bar{\mu}_j(k) \in [0,1]$ – рівень належності вектору $x(k)$ j–му кластеру, c_j – центроїд j–ого кластеру, $D^2\left(x(k), c_j\right)$ – відстань між $x(k)$ та c_j у прийнятій метриці. Результатом кластеризації є матриця $W^{fp} = \left\{w_j^{fp}(k)\right\}$ розмірності $N \times h$, що називається «матрицею нечіткого розбиття».

У використовуваному методі фаззі-кластеризації виконується міра подібності вигляду

$$D^2\left(x(k), c_j\right) = \sum_{i=1}^{n}\left(1 - \operatorname{sech}^2\left(x_i(k) - c_{ji}\right)\right)\left(x_i(k) - c_{ji}\right)^{2/5}, \tag{16}$$

де $x_i(k)-i-a$ компонента вектору входів $x(k)$, $c_{ji}-i-a$ компонента вектору центроїдів кластерів c_j.

Розглянемо цільову функцію

$$E\left(\overline{\mu}_j(k),c_j\right)=\sum_{k=1}^{N}\sum_{j=1}^{h}\overline{\mu}_j^{\beta}(k)D^2\left(x(k),c_j\right)=$$
$$=\sum_{k=1}^{N}\sum_{j=1}^{h}\overline{\mu}_j^{\beta}(k)\sum_{i=1}^{n}\left(1-\operatorname{sech}^2\left(x_i(k)-c_{ji}\right)\right)\left(x_i(k)-c_{ji}\right)^{2/5} \tag{17}$$

та функцію Лагранжа

$$L\left(\overline{\mu}_j(k),c_j,\lambda(k)\right)=\sum_{k=1}^{N}\sum_{j=1}^{h}\overline{\mu}_j^{\beta}(k)\sum_{i=1}^{n}\left(1-\operatorname{sech}^2\left(x_i(k)-c_{ji}\right)\right)\left(x_i(k)-c_{ji}\right)^{2/5}+$$
$$+\sum_{k=1}^{N}\lambda(k)\left(\sum_{j=1}^{h}\overline{\mu}_j(k)-1\right), \tag{18}$$

де $\lambda(k)-$невизначений множник Лагранжа, що забезпечує виконання обмежень (14), (15). Пошук сідлової точки локальної функції Лагранжа за допомогою процедури Ерроу-Гурвіца-Удзави призводить до методу:

$$\begin{cases}\overline{\mu}_j(k)=\left(D^2\left(x(k),c_j\right)\right)^{\frac{1}{1-\beta}}\Big/\sum_{l=1}^{h}\left(D^2\left(x(k),c_l\right)\right)^{\frac{1}{1-\beta}},\\ c_{ji}(k+1)=c_{ji}(k)-\hat{\eta}(k)\dfrac{\partial}{\partial c_{ji}}L_k\left(\overline{\mu}_j(k),c_j,\lambda(k)\right)=\\ =c_{ji}(k)+\hat{\eta}(k)\overline{\mu}_j^{\beta}(k)\Big(2\operatorname{sech}^2\left(x_i(k)-c_{ji}(k)\right)\cdot\tanh\left(x_i(k)-c_{ji}(k)\right)\cdot\left|x_i(k)-c_{ji}(k)\right|^{2/5}+\\ \qquad+0.4\left(1-\operatorname{sech}^2\left(x_i(k)-c_{ji}(k)\right)\right)\cdot\left|x_i(k)-c_{ji}(k)\right|^{-3/5}\cdot\operatorname{sign}\left(x_i(k)-c_{ji}(k)\right)\Big),\end{cases} \tag{19}$$

де $\hat{\eta}(k)-$ параметр кроку навчання, $c_{ji}(k)-$ $i-$та компонента $j-$го вектору центроїдів, що обчислена на $k-$ому кроці.

Далі вводиться адаптивна нео-фаззі прогнозувальна архітектура (надалі називатимемо адаптивним нео-фаззі-предиктором). Ця архітектура (рис. 2) в якості вузлів містить n паралельно з'єднаних нео-фаззі-нейронів з $nd+1$ входами кожен (тут $d\geq 1-$ глибина використаної для прогнозування передісторії часового ряду), що є істотно нелінійними елементами. Кожен з нео-фаззі-нейронів замість звичайних синаптичних ваг містить nd нелінійних синапсів NS_{ji} з h синаптичними вагами $w_{i\overline{l}j}$ кожен, де $i=1,2,\ldots,n;\ \overline{l}=1,2,\ldots nd;\ j=1,2,\ldots h$ при цьому архітектура в цілому містить $n(ndh+1)$ настроюваних ваг.

Таким чином, перетворення, реалізоване даною архітектурою, може бути записано у вигляді

$$\hat{x}(k)=f\left(x(k)\right), \tag{20}$$

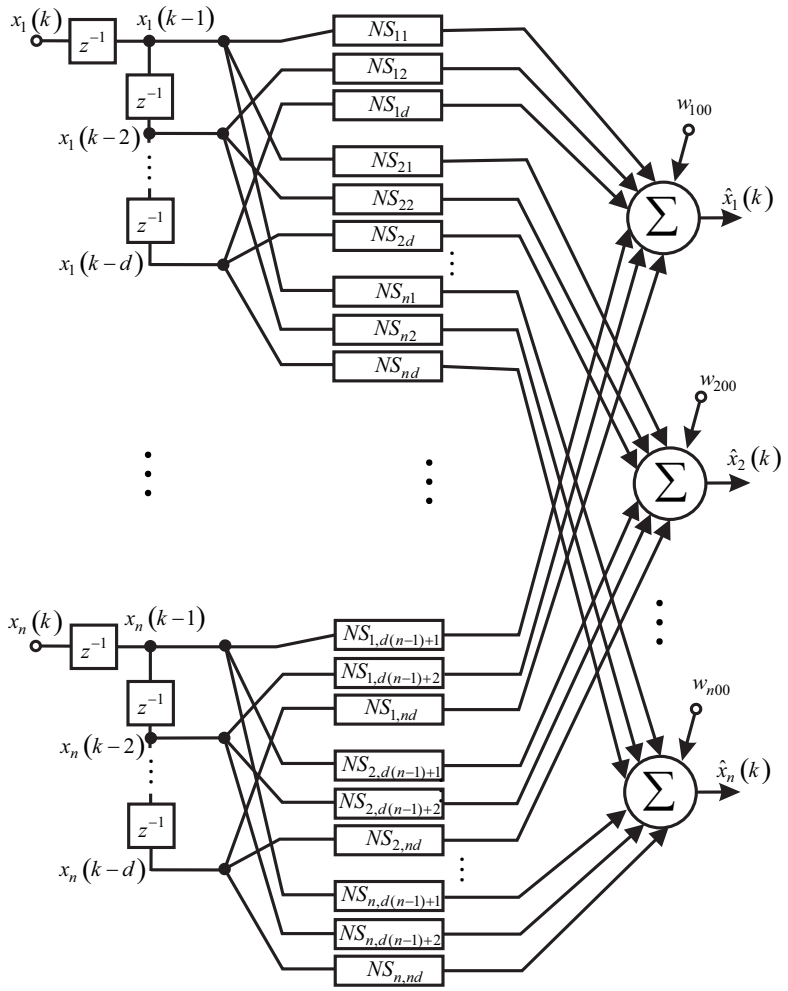

Рисунок 2 - Архітектура адаптивного нео-фаззі-предиктору

де $\hat{x}(k)=\left(\hat{x}_1(k),\hat{x}_2(k),\ldots,\hat{x}_n(k)\right)^T$ $-(n\times 1)$- вектор прогнозів, отриманих з передісторії доступної на момент часу k, $x(k)=\left(1,x_1(k-1),\ldots,x_1(k-d),\ldots,\right.$ $\left.x_n(k-1),\ldots,x_n(k-d)\right)^T-$ $\left((nd+1)\times 1\right)$-вектор цієї передісторії, $W^{pr}-\left(n\times(nd+1)\right)$ – матриця настроюваних синаптичних ваг, $f(\bullet)$-нелінійний оператор спеціального вигляду, що забезпечує оптимальне в сенсі прийнятого критерія прогнозування багатовимірної послідовності $x(k)$.

Вводячи до розгляду $(ndh+1)\times 1$-вектор поточних значень функцій належності та $\left(n\times(ndh+1)\right)$-матрицю синаптичних ваг можна записати сигнал на виході нео-фаззі-предиктору в k-й момент часу у вигляді

$$\hat{x}(k)=f\left(x(k)\right)=W^{pr}(k-1)\mu(k). \tag{21}$$

Навчання може бути реалізовано за допомогою багатовимірної модифікації адаптивного методу ідентифікації у вигляді

$$\begin{cases} W^{pr}(k)=W^{pr}(k-1)+r^{-1}(k)\left(x(k)-W^{pr}(k-1)\mu(k)\right)\mu^T(k), \\ r(k)=\alpha r(k-1)+\left\|\mu(k)\right\|^2, \end{cases} \tag{22}$$

де $0\le\alpha\le 1$ – параметр згладжування, що забезпечує компроміс між слідкуючими и фільтруючими якостями методу.

У кожний поточний момент часу k з надходженням нового вектору спостережень часового ряду $x(k)$ спочатку уточнюється матриця синаптичних ваг $W^{pr}(k)$, а на її основі будується однокроковий прогноз

$$\hat{x}(k+1)=W^{pr}(k)\mu(k+1), \tag{23}$$

де вектор $\mu(k+1)$ містить фаззіфіковані дані з часовими індексами, що не перевищують

k.

Вперше запропоновано багатовимірну модифікацію нео-фаззі-нейрону, що має покращені апроксимувальні властивості, та методи його навчання, які характеризуються простотою реалізації, що дозволяє забезпечити меншу обчислювальну складність за рахунок суттєвого скорочення кількості функцій належності. Архітектура нео-фаззі-предиктору є надлишковою, якщо містить однотипні нелінійні синапси $NS_{i\bar{1}}$. Це пояснюється тим, що $\mu_{1\bar{1}j}(\tilde{x}_i)=\mu_{2\bar{1}j}(\tilde{x}_i)=\ldots=\mu_{i\bar{1}j}(\tilde{x}_i)=\ldots=\mu_{n\bar{1}j}(\tilde{x}_i)$. Аби запобігти цієї надлишковості, було введено до розгляду багатовимірний нео-фаззі-нейрон. Вузлами цієї конструкції є складені нелінійні синапси MNS_j, кожен з котрих містить h функцій належності $\mu_{i\bar{1}j}$ (індекс i тут в принципі може бути прийнятим довільним $1\le i\le n$) та nh настроюваних синаптичних ваг $w_{i\bar{1}h}$. Таким чином, багатовимірний нео-фаззі-нейрон в цілому містить $nd\cdot nh+n=n(ndh+1)$ синаптичних ваг, однак тільки ndh функцій належності, що в n разів менше, ніж в архітектурі нео-фаззі-предиктору. Вводячи далі $(ndh+1)\times 1$-вектор поточних значень функцій належності $\mu(k)=\left(1,\mu_{i11}(\tilde{x}_1(k)),\mu_{i12}(\tilde{x}_1(k)),\ldots,\mu_{i1h}(\tilde{x}_1(k)),\mu_{i21}(\tilde{x}_2(k)),\ldots,\mu_{i,nd,h}(\tilde{x}_{nd}(k))\right)^T$ і $(n\times(ndh+1))$-матрицю синаптичних ваг W^{pr}, знов повертаємось до опису (21) та методу навчання (22). Таким чином, нео-фаззі-предиктор досить просто може бути реалізований на базі багатовимірного нео-фаззі-нейрону.

У **третьому розділі** описано архітектуру нейро-фаззі мережі на основі теорії резервуарних обчислень, яка характеризується підвищеною швидкістю налаштування параметрів навчання й точністю прогнозу часового ряду, та метод її навчання який характеризується простотою реалізації, що дозволяє істотно покращити апроксимувальні та екстраполювальні властивості системи для обробки нелінійних часових рядів за умов невизначеності.

Запропонована архітектура (рис. 3) має вхідний вектор даних $x(k)=\left(x(k-1),\ldots,x(k-d_{in}),\hat{x}(k),\ldots,\hat{x}(k-d_{out})\right)^T$, а $\hat{x}(k)$ є вихідним сигналом (прогнозом) мережі. Прогноз мережі описується в формі

$$\hat{x}(k)=w^T\mu(k), \tag{24}$$

де $\mu(k)=\left(\mu_{11}(x(k)),\mu_{12}(x(k)),\ldots,\mu_{1h}(x(k)),\mu_{21}(x(k)),\ldots,\mu_{d_{in}+d_{out},h}(x(k))\right)$.

Вектор $\mu(k)$ має розмірність $((d_{in}+d_{out})\cdot h)\times 1$, де d_{in} - кількість вхідних елементів чистого затримування, d_{out} - кількість вихідних елементів чистого затримування та h - кількість функцій належності в нелінійному синапсі.

При подачі на вхід мережі сигналу $x(k)$ за допомогою шару елементів чистого затримування сигнал перетворюється на вектор $\left(x(k-1),\ldots,x(k-d_{in})\right)^T$, котрий потім подається на шар NFN, а система отримує прогноз $\hat{x}(k)$. Таким чином, NFN настроює ваги, використовуючи сигнал $x(k)$ в якості вчителя.

Наступна архітектура є прогнозувальною нейронною мережею (рис. 4) і також

побудована на нео-фаззі-нейронах та шарі елементів чистого затримування, що називатиметься далі резервуарною прогнозувальною нейро-нео-фаззі мережею. Тут

$$\hat{x}_l^g(k) = f^{[gl]}(x_g(k-1)) + f^{[gl]}(\hat{x}_l^g(k-2)) + \ldots + f^{[gl]}(\hat{x}_n^g(k-2)), \tag{25}$$

$$f^{[gl]}(x_g(k-1)) = \sum_{i=1}^{n+1}\sum_{j=1}^{h}\mu_{ij}^{[gl]}\left(x_g(k-1)\right)w_{ij}^{[gl]}(k), \tag{26}$$

$$f^{[gl]}(\hat{x}_l^g(k-2)) = \sum_{i=1}^{n+1}\sum_{j=1}^{h}\mu_{ij}^{[gl]}\left(\hat{x}_l^g(k-2)\right)w_{ij}^{[gl]}(k), \tag{27}$$

$$f^{[gl]}(\hat{x}_n^g(k-2)) = \sum_{i=1}^{n+1}\sum_{j=1}^{h}\mu_{ij}^{[gl]}\left(\hat{x}_n^g(k-2)\right)w_{ij}^{[gl]}(k), \tag{28}$$

де $\hat{x}_l^g$ – вихід l–го нейрону g–ої підсистеми.

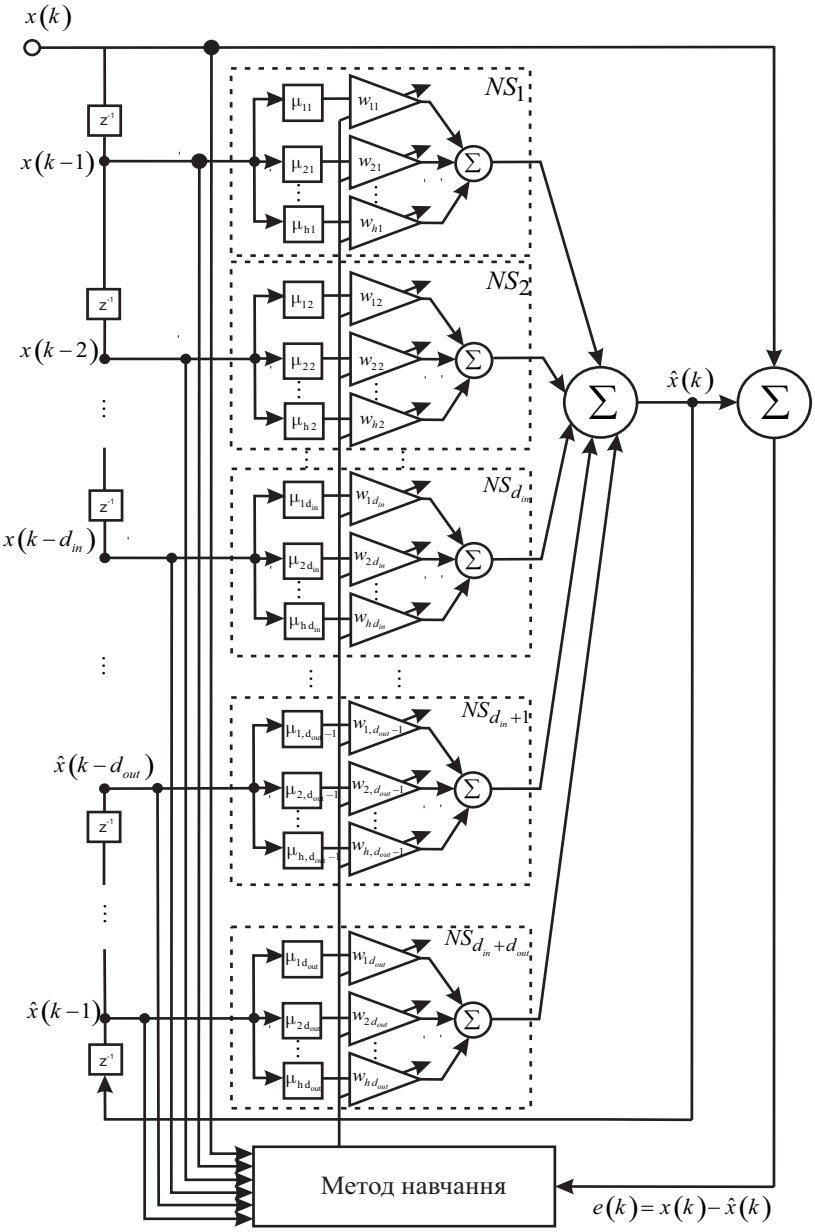

Рисунок 3 - Архітектура резервуарної прогнозувальної мережі, що складається з нео-фаззі-нейронів із зворотними зв’язками

Відзначимо, що в цьому випадку використовується модифікований NFN (рис. 5): у класичному нео-фаззі-нейроні на входи всіх синапсів подається однаковий сигнал, а в цьому випадку на вхід кожного синапсу подається власне значення (один вхідний сигнал) і n сигналів з виходів, що пройшли крізь шар елементів чистого затримування. Архітектура системи складається з g підсистем (рис. 4). Тут $x_g(k)$ – скалярне значення, що надходить на вхід підмережі, $\hat{x}_g(k)$ – вихідне значення підмережі, $k = 1, 2, \ldots, N$ – поточний дискретний час. Слід зазначити, що якщо додати нові нео-фаззі-нейрони до архітектури підмережі, то її робота принципово не зміниться.

Вхідний сигнал після проходження крізь шар затримування подається на шар нео-фаззі-нейронів.

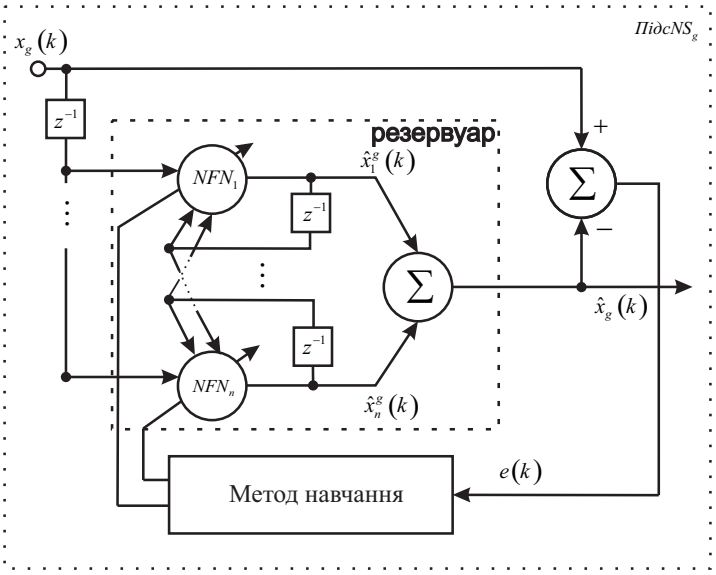

Рисунок 4 – Підсистема резервуарної прогнозуючої нейро-нео-фаззі мережі

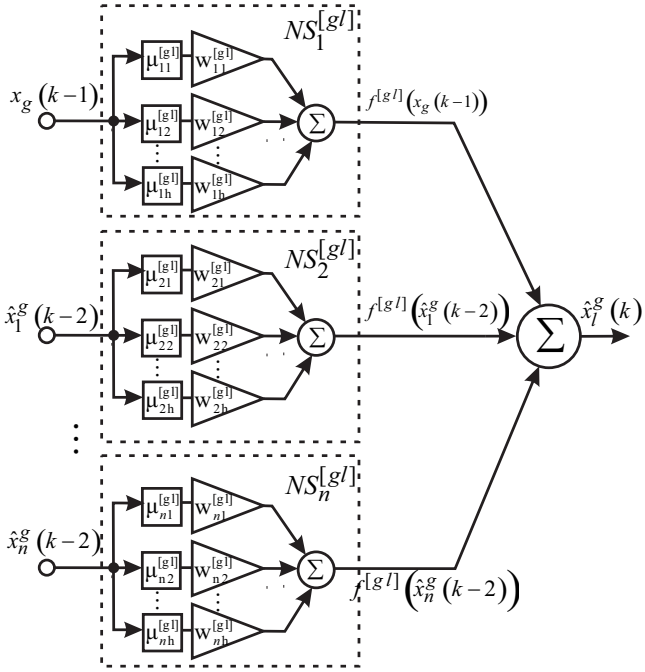

Рисунок 5 – Архітектура модифікованого нео-фаззі-нейрону

Особливістю цієї конструкції є наявність зворотних зв'язків, спрямованих крізь шар елементів затримування від виходів NFN назад до їх входів, що й дозволяє називати цю конструкцію «резервуаром». Вихід підсистеми розраховується за формулою

$$\hat{x}_g(k)=\sum_{l=1}^{n} f^{[gl]}(x_g(k-1))+f^{[gl]}(\hat{x}_1^g(k-2))+\ldots+f^{[gl]}(\hat{x}_n^g(k-2))=$$
$$=\sum_{l=1}^{n}\sum_{i=1}^{n+1}\sum_{j=1}^{h}\mu_{ij}^{[gl]}\left(x_g(k-1)\right)w_{ij}^{[gl]}(k)+\mu_{ij}^{[gl]}\left(\hat{x}_1^g(k-2)\right)w_{ij}^{[gl]}(k)+\mu_{ij}^{[gl]}\left(\hat{x}_n^g(k-2)\right)w_{ij}^{[gl]}(k), \quad (29)$$

де n – кількість входів, $\mu_{ij}^{[gl]}$ – функції належності NFN_l^g, $w_{ij}^{[gl]}$ – синаптичні ваги NFN_l^g. Відмітимо, що на вхід NFN_l^g буде подано вектор розмірністю $\left((n+1)\times 1\right)$.

Вводячи до розгляду $\left((n+1)\cdot nhg\right)\times 1$-вектор значень функцій належності та вектор аналогічної розмірності синаптичних ваг, можна представити вихід мережі у векторній формі у вигляді

$$\hat{x}(k)=w^T\mu(k). \quad (30)$$

Навчання розглянутих мереж здійснюється згідно стандартного методу найменших квадратів у рекурентній формі, що його було розглянуто раніше.

У **четвертому розділі** набув подальшого розвитку метод сегментації багатовимірних часових рядів на основі методу нечіткої кластеризації Гата-Геви шляхом спеціального налаштування центроїдів, що дозволяє здійснювати сегментацію часових рядів у режимі online. Розбиття часового ряду на внутрішні однорідні сегменти займає важливе місце серед задач інтелектуального аналізу даних, оскільки дозволяє по суті виділити з великих масивів інформації ключові характеристики часового ряду у більш компактній формі. Зміни властивостей багатовимірного часового ряду звичайно слабо

відчутні та не фокусуються на декотрій визначеній часовій точці.

Запропонований метод сегментації припускає послідовне оцінювання однорідності сегментів, а також їх формальне представлення у часі.

У загальному випадку задача сегментації часового ряду може бути сформульована наступним чином: знайти в декотрому розумінні найкраще розбиття часового ряду на деяку кількість внутрішньо однорідних сегментів.

Зараз існує ряд широко вживаних методів сегментації. Однак головним недоліком всіх широко використовуваних сьогодні методів сегментації є адекватні результати лише при роботі в пакетному режимі.

Перед тим, як використовувати запропонований метод сегментації, дані необхідно скомпресувати за допомогою аналізу головних компонент (PCA). Сегментація часових рядів може розглядатися як кластеризація з упорядкованою у часі структурою. Оскільки точки в кластері повинні йти друг за другом у часі, то треба прийняти до уваги під час кластеризації часову координату даних. Отже, кластеризація базується на мірі відстані, що складається з двох етапів: відстань ґрунтується на тому, як дані визначено нечіткими множинами, що їх задано у даному часовому інтервалі у вказаному сегменті; другий період вимірює те, як далеко дані знаходяться від гіперплощини моделі сегментів PCA.

Ефективність і точність запропонованого методу залежить від правильного вибору кількості головних компонент та кількості сегментів.

У загальному випадку метод сегментації може бути представлено у такому вигляді: спочатку треба задати $\overline{c}$ – кількість сегментів та $\widehat{q}$– кількість головних компонент, обрати поріг зупинки $\varepsilon > 0$, ініціалізувати величини W_i^{sg} - матриця ваг, $\overline{\mu}_{ik}$- рівень належності спостереження i-му кластеру, σ_{ix}^2 - дисперсія, c_i^x - центроїд кластеру.

На першому етапі треба обчислити центри кластерів у формі:

$$\begin{aligned} c_i^x &= \left(\sum\nolimits_{k=1}^{N}\left(\overline{\mu}_{ik}\right)^{\beta}\right)^{-1}\sum\nolimits_{k=1}^{N}\left(\overline{\mu}_{ik}\right)^{\beta}\left(x_k - W_i^{sg}y_{ik}\right) = \\ &= \left(\sum\nolimits_{k=1}^{N}\left(\overline{\mu}_{ik}\right)^{\beta}\right)^{-1}\sum\nolimits_{k=1}^{N}\left(\overline{\mu}_{ik}\right)^{\beta}\left(x_k - W_i^{sg}M_i^{-1}\left(W_i^{sg}\right)^{T}\left(x_k - c_i^x\right)\right), \end{aligned} \tag{31}$$

де $y_{ik} = M_i^{-1}\left(W_i^{sg}\right)^{T}\left(x_k - c_i^x\right)$ та $M_i = \sigma_{ix}^2 I + \left(W_i^{sg}\right)^{T} W_i^{sg}$ – матриця розмірності $\left(\widehat{q}\times\widehat{q}\right)$, I - одинична матриця.

Після цього обчислюються нові значення матриці ваг W_i^{sg}

$$W_i = F_i W_i^{sg}\left(\sigma_{ix}^2 I + M_i^{-1}\left(W_i^{sg}\right)^{T} F_i W_i^{sg}\right)^{-1}, \tag{32}$$

де F_i - коваріаційна матриця i-го кластеру.

Наступним кроком обчислюються нові значення σ_{ix}^2

$$\sigma_{ix}^2 = \frac{1}{\widehat{q}}\mathrm{Tr}\left(F_i - F_i W_i^{sg} M_i^{-1}\left(W_i^{sg}\right)^{T}\right) \tag{33}$$

та норма відстані A_i (матриця розмірності $\left(n\times n\right)$)

$$A_i = \sigma_{ix}^2 I + W_i^{sg}(l)\left(W_i^{sg}(l)\right)^T. \tag{34}$$

Після чого відбувається обчислення параметрів моделі у часі: центр моделі у часі c_i^t розраховується в загальному вигляді за допомогою рекурентної процедури

$$c_i^t(l+1) = c_i^t(l) - \breve{\eta}(l)\bar{\mu}_{it}^{\beta} D^2\left(t_k(l+1), c_i^t(l)\right)\nabla_{c_i^t} D^2\left(t_k(l+1), c_i^t(l)\right), \tag{35}$$

де $\breve{\eta}$ - параметр кроку навчання, $D^2\left(t_k(l+1), c_i^t(l)\right)$ - відстань між значенням часового ряду t_k та центром моделі у часі, а стандартна дисперсія обчислюється у вигляді

$$\sigma_{it}^2 = \left(\sum\nolimits_{k=1}^{N}\left(\bar{\mu}_{ik}\right)^{\beta}\right)^{-1} \mathrm{Tr}\left(\sum\nolimits_{k=1}^{N}\left(\bar{\mu}_{ik}\right)^{\beta}\left(t_k - c_i^t\right)\left(t_k - c_i^t\right)^T\right). \tag{36}$$

На другому етапі обчислюються міри відстаней $D^2\left(z_k, \eta_i\right)$ у формі:

$$\begin{aligned} D^2\left(z_k, \eta_j\right) = \frac{1}{\hat{\alpha}_i}\sqrt{2\pi\sigma_{it}^2}\exp\left(\frac{1}{2}\frac{\left(t_k - c_i^t\right)^2}{\sigma_{it}^2}\right)\times \\ \times(2\pi)^{\hat{r}/2}\sqrt{\det\left(A_i\right)}\exp\left(\frac{1}{2}\left(x_k - c_i^x\right)^T A_i^{-1}\left(x_k - c_i^x\right)\right), \end{aligned} \tag{37}$$

де $\hat{\alpha}_i = \dfrac{\sum\nolimits_{k=1}^{N}\mu_{ik}}{N}$, N - кількість точок у вибірці, $\hat{r}$ – ранг норми відстані A_i.

На третьому етапі змінюється матриця розбиття. В інакшому випадку треба перейти знову до першого етапу.

Якщо ж довжина вибірки досить мала або збігається з розмірністю багатовимірного часового ряду (нейромережа не встигає навчитися), дані треба деяким чином стиснути. То ж з цією метою набула подальшого розвитку архітектура нейро-компресору даних на основі багатошарового персептрону «вузьке місце» та нейро-нео-фаззі мережі зустрічного поширення, що дозволило здійснювати прогнозування часових рядів зниженої розмірності без втрати внутрішніх зв'язків збережуваних даних.

Нехай $\hat{u}_j^{[s]}$ - внутрішній сигнал j-го нейрону s-го шару багатошарового песептрону, що містить n_s нейронів, є зваженою сумою виходів попереднього шару

$$\hat{u}_j^{[s]} = \sum_{i=0}^{n_{s-1}} w_{ji}^{[s]} o_i^{[s-1]}, \tag{38}$$

де $w_{ji}^{[s]}$ - синаптичні ваги, що поєднують кожний нейрон s-го шару з усіма нейронами попереднього шару так, що $o_i^{[s-1]} = x_i^{[s]}, o_i^{[0]} = x_i, o_i^{[3]} = \tilde{x}_i$.

Тоді вихід кожного нейрону є нелінійне перетворення вхідного сигналу вигляду

$$o_j^{[s]} = \psi_j^{[s]}(\hat{u}_j^{[s]}) = \psi_j^{[s]}(\sum_{i=0}^{n_{s-1}} w_{ji}^{[s]} o_i^{[s-1]}), \tag{39}$$

де для скорочення запису величина зміщення $\theta_j^{[s]}$ позначається як синаптична вага $w_{j0}^{[s]} = \theta_j^{[s]}$, а $o_0^{[s-1]} \equiv 1$.

Використовуючи векторно-матричну форму, можна записати перетворення, виконуване кожним шаром, у компактному вигляді

$$o^{[s]} = \Psi^{[s]}(W^{[s]}x^{[s]}) = \Psi^{[s]}(W^{[s]}o^{[s-1]}), \tag{40}$$

де $o^{[s]} = (o_0^{[s]}, o_1^{[s]}, o_2^{[s]}, ..., o_{n_s}^{[s]})^T - ((n_s+1)\times 1)$ – вихідний вектор сигналів; $x^{[s]} = (x_0^{[s]}, x_1^{[s]}, ..., x_{n_{s-1}}^{[s]})^T = o^{[s-1]} - ((n_{s-1}+1)\times 1)$ – вхідний вектор; $W^{[s]} - ((n_s+1)\times(n_{s-1}+1))$ – матриця синаптичних ваг; $\Psi^{[s]}(\bullet)$- діагональний нелінійний оператор, утворений активаційними функціями всіх нейронів $\psi_j^{[s]}$.

З урахуванням введених позначень рівняння трьохшарового персептрону приймає вигляд

$$y = \Psi(x) = \Psi^{[3]}(W^{[3]}\Psi^{[2]}(W^{[2]}\Psi^{[1]}(W^{[1]}x))). \tag{41}$$

Прогнозувальна архітектура представляє собою нейро-нео-фаззі мережу зустрічного поширення, яку було докладно описано в розділі 2 та яка, в свою чергу, складається з шару нечіткої кластеризації Кохонена та шару нео-фаззі-нейронів. На вхід прогнозувальної системи подаються вихідні сигнали з другого шару персептрону (шару-компресору), таким чином, на вхід подається вектор стиснутих значень $o^{[2]}(k) = \left(o_1^{[2]}(k), o_2^{[2]}(k), \ldots, o_{\tilde{m}}^{[2]}(k)\right)^T$. Надалі позначимо цей вектор значень через $y(k) = \left(y_1(k), y_2(k), \ldots, y_{\tilde{m}}(k)\right)^T$. Робота мережі полягає в обчисленні рівнів належності поточного спостереження до кластерів, які визначаються вагами нейронів шару Кохонена, обчисленні значень вихідних сигналів NFN та ранжуванні цих значень у відповідності до ступенів належності на виходах шару Кохонена. Виходом мережі є вектор прогнозів $\tilde{y}(k)$.

Навчання багатошарового персептрону може бути здійснено за допомогою будь-якої процедури зворотного поширення похибок лише з тією різницею, що в якості навчального образу використовується сам вхідний сигнал $x(k)$, що його треба стиснути. Ця архітектура побудована на формальних нейронах з поліноміальною функцією активації. Треба зауважити, що ця функція активації задовольняє всім вимогам, що їх висувають до сигмоїдальних функцій активації.

$$\psi(\gamma_j(k)\hat{u}_j(k)) = \gamma_j(k)\hat{u}_j(k) - \frac{1}{3}\gamma_j^3(k)\hat{u}_j^3(k). \tag{42}$$

Процес навчання нейронної мережі зводиться до настроювання матриць синаптичних ваг $W^{[s]}$, *s*=1,2,3, шляхом мінімізації прийнятої цільової функції, в якості котрої використовується вираз

$$E(k) = \frac{1}{2}\sum_{i=1}^{n} E_i(k) = \frac{1}{2}\sum_{i=1}^{n} e_i^2(k),$$

де $e_i(k) = x_i(k) - \hat{x}_i(k)$, $\hat{x}(k) - (n \times 1)$ – вихідний векторний сигнал мережі; $x(k) - (n \times 1)$ – вхідний вектор.

Використовуючи процедури градієнтного спуску, отримуємо метод навчання загального вигляду:

$$w_{ji}^{[s]}(k+1) = w_{ji}^{[s]}(k) - \eta(k)\frac{\partial E(k)}{\partial w_{ji}^{[s]}(k)} = w_{ji}^{[s]}(k) - \eta(k)\frac{\partial E(k)}{\partial \hat{u}_j^{[s]}(k)}\frac{\partial \hat{u}_j^{[s]}(k)}{\partial w_{ji}^{[s]}(k)} =$$

$$= w_{ji}^{[s]}(k) + \eta(k)\delta_j^{[s]}(k)o_i^{[s-1]}(k),\ \Delta w_{ji}^{[s]}(k) = -\eta(k)\frac{\partial E(k)}{\partial w_{ji}^{[s]}(k)}, \qquad (43)$$

де $\delta_j^{[s]}(k) = \dfrac{\partial E(k)}{\partial u_j^{[s]}(k)}$ - локальна похибка *j*-го нейрону *s*-го шару; $\eta(k)$ - параметр кроку, що визначає швидкість навчання мережі; $o_i^{[s-1]}$ - вихідний сигнал (*s*-1)-го (попереднього) шару.

Запишемо метод настроювання синаптичних ваг для кожного шару на основі процедури зворотного поширення похибки. Таким чином, запропонований метод навчання цієї мережі представляє собою процедуру настроювання трьох матриць синаптичних ваг, починаючи з вихідного (третього) шару та має вигляд :

$$\begin{cases} \Delta w_{ji}^{[3]}(k) = \eta(k)(1 - \hat{u}_j^{[3]2}(k))e(k)o_i^{[2]}(k), \\ \Delta w_{ji}^{[2]}(k) = \eta(k)(1 - \hat{u}_j^{[2]2}(k))o_i^{[1]}\sum_{\hat{p}=1}^{n_2}\delta_l^{[3]}(k)w_{\hat{p}i}^{[3]}(k), \\ \Delta w_{ji}^{[1]}(k) = \eta(k)(1 - \hat{u}_j^{[1]2}(k))x_i(k)\sum_{\check{p}=1}^{n_1}\delta_l^{[2]}(k)w_{\check{p}i}^{[2]}(k). \end{cases} \qquad (44)$$

Поліноміальна функція активації (44) має більш прості похідні за настроюваними параметрами порівняно з традиційниминими функціями активації, що дозволяє спростити чисельну реалізацію процедури навчання.

Процедуру навчання прогнозувальної нейро-нео-фаззі мережі описано в розділі 2.

У **п’ятому розділі** наведено розв’язки практичних задач на основі запропонованих у дисертаційній роботі гетерогенних нейронних мереж та методів їх навчання. Розроблено програмні модулі, які реалізують запропоновані методи та архітектури гетерогенних нейронних мереж, що дозволило автоматизувати процес обробки багатовимірних часових рядів для розв’язання задач прогнозування, компресії та сегментації. Результати тестування програмних модулів, що реалізують запропоновані методи та архітектури гетерогенних нейронних мереж, підтвердили теоретичні висновки щодо доцільності їх використання для розв’язання практичних задач. Розв’язано задачу прогнозування споживання електроенергії в автоматизованих системах диспетчерського керування. Це призвело до підвищення точності прогнозування у порівнянні із широковживаними методами прогнозування та до зниження матеріальних втрат, що виникають внаслідок похибок прогнозування. Розв’язано задачу сегментації даних електрокардіограм з метою прогнозування у різних груп хворих на гіпертонічну іта ішемічну хвороби серця ймовірності виникнення різних порушень серцевого ритму. Усі результати впроваджено та підтверджено актами впровадження.

У **висновках** сформульовано наукові та практичні результати, що їх одержано у дисертаційній роботі.

Додаток містить копії документів про впровадження та практичне застосування результатів, отриманих у дисертаційній роботі.

ВИСНОВКИ

У дисертаційній роботі представлено результати, що відповідно до поставленої мети є вирішенням актуальної науково-технічної задачі підвищення ефективності обробки багатовимірних нелінійних часових рядів за рахунок розробки інтелектуальних методів обробки інформації на основі гетерогенних нейронних мереж.

Автором одержано такі наукові та практичні результати:

1. Нова архітектура гетерогенної нейронної мережі для прогнозування даних, що відзначається підвищеною швидкодією та покращеними апроксимувальними властивостями, на основі шару нечіткої кластеризації та нейро-нео-фаззі мережі зустрічного поширення. Це дозволяє обробляти нелінійні часові ряди довільної природи за умов невизначеності, а також синтезовано метод її навчання.
2. Нова багатовимірна модифікація нео-фаззі-нейрону з покращеними апроксимувальними властивостями, а також синтезовано методи його навчання, які відзначаються простотою реалізації. Це дозволяє забезпечити суттєве скорочення кількості функцій належності.
3. Нова архітектура нейро-фаззі мережі на основі теорії резервуарних обчислень з підвищеною швидкістю налаштування параметрів навчання та точністю прогнозу часового ряду. Синтезовано чисельно простий метод її навчання. Це дозволило істотно покращити апроксимувальні та екстраполювальні властивості системи для обробки нелінійних часових рядів за умов невизначеності.
4. Розвинена архітектура нейро-компресора за рахунок скорочення об’єму навчальної вибірки за умов обмеженої кількості даних, що відзначається здатністю обробляти часові ряди зниженої розмірності без втрати внутрішніх зв’язків збережуваних даних.
5. Розвинений метод сегментації багатовимірних часових рядів на основі методу нечіткої кластеризації Гата-Геви. Запропонований метод відзначається процедурою налаштування цетроїдів спеціального виду та характеризується здатністю визначати моменти змін властивостей багатовимірних часових рядів у режимі online.
6. Проведено імітаційне моделювання запропонованих архітектур і методів та показано їх переваги над існуючими методами у задачах обробки багатовимірних нелінійних часових рядів. Розв’язано практичні задачі обробки багатовимірних нелінійних часових рядів. Результати досліджень впроваджено у ТОВ «НВП «Енергобуд», м. Харків, та у ТОВ «Небозвід», м. Харків, що підтверджено відповідними актами.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Тищенко А.К. Предиктор сжатых данных на основе нейро-нео-фаззи сети встречного распространения / А.К. Тищенко, И.П. Плисс, Н.А. Тесленко // Восточно-

Европейский журнал передовых технологий. – 2012. – № 3/3(57). – С. 14-17.

2. Тищенко А.К. Сегментация многомерных нестационарных временных рядов с помощью метода нечеткой кластеризации / А.К. Тищенко, И.П. Плисс // Восточно-Европейский журнал передовых технологий. – 2012. – № 4/4(58). – С. 24-27.

3. Тищенко А.К. Рекуррентный алгоритм нечеткой кластеризации в нейро-нео-фаззи сети встречного распространения / А.К. Тищенко, И.П. Плисс // Збірник наукових праць Харківського університету Повітряних Сил. – 2012. – Вип. 3(32). – С. 126-129.

4. Тищенко А.К. Прогнозирующая нейро-нео-фаззи сеть на основе резервуарных вычислений / А.К. Тищенко, И.П. Плисс // Системи управління, навігації та зв'язку. – 2012. – Вип. 1(21). – Т.2. – С. 123-126.

5. Бодянский Е.В. Адаптивный нео-фаззи-предиктор многомерных нестационарных временных рядов / Е.В. Бодянский, А.К. Тищенко // Системи обробки інформації. – 2012. – Вип. 3(101). – Т.2. – С. 8-12.

6. Тищенко А.К. Об одном алгоритме обучения нейронных сетей встречного распространения / А.К. Тищенко // Моделювання: зб. наук. праць за матеріалами XXVIII наук.-техн. конф. – Київ. – 2009. – С. 16.

7. Тищенко А.К. Нейро-фаззи сеть встречного распространения и алгоритм ее обучения / А.К. Тищенко // Радиоэлектроника и молодежь в XXI веке: сб. науч. трудов по материалам 13-го междун. мол. Форума. – Харьков: ХНУРЭ. – 2009. – Ч.2. – С. 20.

8. Бодянский Е.В. Архитектура нейро-нео-фаззи сети и алгоритм ее обучения / Е.В. Бодянский, Е.В. Горшков, А.К. Тищенко // Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту: зб. наук. праць за матеріалами міжнар. наук. конф. – Херсон: ХНТУ. – 2009. – Т.2. – С. 239-243.

9. Бодянский Е.В. Искусственная нейронная сеть встречного распространения на нео-фаззи-нейронах / Е.В. Бодянский, Е.В. Горшков, А.К. Тищенко // Автоматика – 2009: зб. наук. праць за матеріалами 16 міжнар. конф. з автоматичного управління. – Чернівці: Видавництво «Книги-XXI». – 2009. – С. 387-389.

10. Бодянский Е.В. Нейро-фаззи-система, основанная на нео-фаззи-нейронах и процедуре нечеткой кластеризации / Е.В. Бодянский, Е.В. Горшков, А.К. Тищенко // Информационные технологии и информационная безопасность в науке, технике и образовании: сб. науч. трудов по материалам междунар. науч.-практ. конф. – Севастополь: СевНТУ. – 2009. – С. 207-209.

11. Бодянский Е.В. Резервуарная прогнозирующая нейро-фаззи сеть / Е.В. Бодянский, А.К. Тищенко // Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту: зб. наук. праць за матеріалами міжнар. наук. конф. – Херсон: ХНТУ. – 2010. – Т.1. – С. 279-282.

12. Тищенко А.К. Гибридная нейро-нео-фаззи сеть с нечетким рекурсивным алгоритмом кластеризации / А.К. Тищенко, М.З. Стольникова // Теорія прийняття рішень: зб. наук. праць за матеріалами V міжнар. школи-семінару. – Ужгород: УжНУ. – 2010. – С. 218-219.

13. Тищенко А.К. Архитектура нейро-фаззи предиктора скомпрессированных данных / А.К. Тищенко // Моделювання: зб. наук. праць за матеріалами XXX наук.-техн. конф. – Київ. – 2011. – С. 49.

14. Бодянский Е.В. Адаптивный предиктор на основе многомерного нео-фаззи-нейрона / Е.В. Бодянский, А.К. Тищенко // Інтелектуальні системи прийняття рішень і

проблеми обчислювального інтелекту: зб. наук. праць за матеріалами міжнар. наук. конф. – Херсон: ХНТУ. – 2011. – Т.1. – С. 321-324.

15. Бодянский Е.В. Архитектура и алгоритм обучения нейро-фаззи предиктора сжатых данных / Е.В. Бодянский, А.К. Тищенко // Информационные технологии и информационная безопасность в науке, технике и образовании: сб. науч. трудов по материалам междунар. науч.-практ. конф. – Севастополь: СевНТУ. – 2011. – С. 7-8.

16. Bodyanskiy Ye. Compressed Data Predictor / Ye. Bodyanskiy, O. Tyshchenko // Сучасні комп'ютерні системи та мережі: розробка та використання: зб. наук. праць за матеріалами 5-ї міжнар. наук.-техн. конф. – Львів: НВФ «Українські технології». – 2011. – С. 217-219.

17. Тищенко А.К. Резервуарная архитектура нейро-фаззи-сети в задачах прогнозирования / А.К. Тищенко // Моделювання: зб. наук. праць за матеріалами XXIX наук.-техн. конф. – Київ. – 2012. – С. 38.

18. Bodyanskiy Ye. Neo-Fuzzy Forecasting Echo State Network / Ye. Bodyanskiy, O. Tyshchenko // Сучасні комп'ютерні системи та мережі: розробка та використання: зб. наук. праць за матеріалами 4-ї міжнар. наук.-техн. конф. – Львів: НВФ «Українські технології». – 2009. – С. 95-96.

19. Тищенко А.К. Нечеткая сегментация временных рядов / А.К. Тищенко // Моделювання: зб. наук. праць за матеріалами XXXI наук.-техн. конф. – Київ. – 2012. – С. 48.

АНОТАЦІЯ

Тищенко О.К. Інтелектуальний аналіз багатовимірних нелінійних часових рядів на основі гетерогенних нейронних мереж. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту. – Харківський національний університет радіоелектроніки, Міністерство освіти і науки України, Харків, 2013.

Дисертацію присвячено розробці методів обробки багатовимірних нелінійних часових послідовностей за допомогою гетерогенних нейронних мереж. Уперше запропоновано архітектуру нейро-фаззі мережі на основі теорії резервуарних обчислень, яка характеризується підвищеною швидкістю налаштування параметрів навчання й точністю прогнозу часового ряду, що дозволяє істотно покращити апроксимувальні та екстраполювальні властивості системи для обробки нелінійних часових рядів за умов невизначеності. Уперше запропоновано багатовимірну модифікацію нео-фаззі-нейрону, що має покращені апроксимувальні властивості, та методи його навчання, які характеризуються простотою реалізації, що дозволяє забезпечити меншу обчислювальну складність за рахунок суттєвого скорочення кількості функцій належності. Уперше запропоновано архітектуру для прогнозування даних, що складається з шару нечіткої кластеризації та нейро-нео-фаззі мережі зустрічного поширення, яка характеризується підвищеною швидкодією та покращеними апроксимувальними властивостями, що дає можливість обробляти нелінійні часові ряди довільної природи за умов невизначеності.

Набула подальшого розвитку архітектура нейро-компресора даних на основі багатошарового персептрону «вузьке місце» за рахунок приєднання до мережі-

компресора нейро-нео-фаззі мережі зустрічного поширення, що дозволило здійснювати прогнозування часових рядів зниженої розмірності без втрати внутрішніх зв'язків збережуваних даних. Набув подальшого розвитку метод сегментації багатовимірних часових рядів на основі методу нечіткої кластеризації Гата-Геви, який відрізняється процедурою налаштування цетроїдів, що дозволяє здійснювати сегментацію багатовимірних часових рядів у режимі online.

Ключові слова: гетерогенні нейронні мережі, нейро-фаззі мережа, методи навчання, функція належності, нео-фаззі-нейрон.

АННОТАЦИЯ

Тищенко А.К. Интеллектуальный анализ многомерных нелинейных временных рядов на основе гетерогенных нейронных сетей. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.23 – системы и средства искусственного интеллекта. – Харьковский национальный университет радиоэлектроники, Министерство образования и науки Украины, Харьков, 2013.

Диссертация посвящена исследованию и разработке методов обработки многомерных временных рядов с помощью гетерогенных нейронных сетей.

В первом разделе подробно рассмотрены нейронные сети встречного распространения, потому что предложенные гетерогенные нейронные сети строились, опираясь на их структуру. Также уделено внимание новому подходу вычислительного интеллекта – резервуарным вычислениям. Рассматривается механизм функционирования подобных систем, а также параметры настройки резервуаров. Выполнена постановка задачи исследования, проанализированы существующие методы обработки многомерных временных рядов, определены преимущества и недостатки рассмотренных подходов.

Во втором разделе рассматривается архитектура традиционного нео-фаззи-нейрона (NFN), так как данная вычислительная система играет важную роль в дальнейших построениях гетерогенных нейронных сетей. Впервые предложена архитектура для прогнозирования данных, состоящая из слоя нечеткой кластеризации и нейро-нео-фаззи сети встречного распространения, которые характеризуются повышенным быстродействием и улучшенными аппроксимирующими свойствами, что позволяет обрабатывать нелинейные временные ряды произвольной природы в условиях неопределенности. Впервые предложена многомерная модификация нео-фаззи-нейрона, имеющая улучшенные аппроксимирующие свойства, и методы обучения, которые характеризуются простотой реализации, позволяющие обеспечить меньшую вычислительную сложность за счет существенного сокращения количества функций принадлежности.

В третьем разделе впервые предложена архитектура нейро-фаззи сети на основе теории резервуарных вычислений, которая характеризуется повышенной скоростью настраивания параметров обучения и точностью прогноза временного ряда, что позволяет существенно улучшить аппроксимирующие и экстраполирующие свойства системы для обработки нелинейных временных рядов в условиях неопределенности.

В четвертом разделе получила дальнейшее развитие архитектура нейро-компрессора на основе многомерного персептрона «бутылочное горлышко» и нейро-

нео-фаззи сети встречного распространения, что позволило осуществлять прогнозирование временных рядов пониженной размерности без потери связей внутри данных. Также получил дальнейшее развитие метод сегментации многомерных временных рядов, основанный на методе нечеткой кластеризации и отличающийся процедурой настраивания центроидов, что позволило осуществлять сегментацию данных в режиме online.

Проведено численное моделирование ряда практических задач, на основе которых показана эффективность использования предложенных архитектур и методов обработки многомерных временных рядов на основе гетерогенных нейронных сетей. Они могут быть использованы в любой области, где данные представлены в числовой форме в виде временных последовательностей.

Ключевые слова: гетерогенные нейронные сети, нейро-фаззи сеть, методы обучения, функция принадлежности, нео-фаззи-нейрон.

ABSTRACT

Tyshchenko O.K. Multivariate nonlinear time series intelligent analysis based on heterogeneous neural networks. – Manuscript.

A thesis for the candidate degree in technical sciences in the specialty 05.13.23 – systems and tools of artificial intelligence. – Kharkiv National University of Radio Electronics, Ministry of Education and Science, Kharkiv, 2013.

The thesis is devoted to methods and architectures for multivariate time series processing with the help of heterogeneous neural networks.

A predictive neuro-fuzzy architecture based on reservoir computing techniques is proposed. It demonstrates high speed parameter setting and has linear outputs and high processing speed.

A multidimensional neo-fuzzy-neuron architecture is proposed. This architecture lessens the amount of membership functions comparing to other analogues and provides smaller computing complexity. A new learning method is proposed.

A predictive architecture is proposed that consists of a fuzzy clusterization layer and a neuro-neo-fuzzy counter propagation network. It provides simple computing realization and high processing speed in an online mode. Values from the neo-fuzzy-neuron layer outputs are ranked according to membership degrees from the fuzzy clusterization layer outputs. This procedure allows to get a high quality prediction result.

A predictive architecture of compressed data based on multilayer perceptron “bottle-neck” and a neuro-neo-fuzzy counter propagation network is improved. This system can predict time series of reduced dimension without quality loss of connections inside data.

A multivariate time series segmentation method based on fuzzy clusterization method is improved. Now it can process data in an online mode.

Lots of experiments are carried out. The effectiveness of proposed architectures and methods has been proved while solving practical problems.

Key words: heterogeneous neural networks, neuro-fuzzy networks, learning methods, membership function, neo-fuzzy-neuron.

Відповідальний випусковий В.П. Машталір