# НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

# НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

# NEUROINFORMATICS AND INTELLIGENT SYSTEMS

*УДК 519.7:004.8*

Дейнеко А. О.[1], Плісс І. П.[2], Бодянський Є. В.[3]

*[1]Аспірантка Харківського національного університету радіоелектроніки*
*[2]Канд. техн. наук, провідний науковий співробітник Харківського національного університету радіоелектроніки*
*[3]Д-р техн. наук, професор Харківського національного університету радіоелектроніки*

## КОМБІНОВАНЕ НАВЧАННЯ ЕВОЛЮЦІЙНОЇ НЕЙРО-ФАЗЗІ СИСТЕМИ

Запропоновано архітектуру еволюційної нейро-фаззі системи, що не схильна до прокльону розмірності, здатна обробляти інформацію в режимі реального часу, адаптуючи при цьому свої параметри і структуру до умов задачі. У якості активаційних функцій було використано ядерні функції.

**Ключові слова:** еволюційна нейро-фаззі система, нормалізована радіально-базисна нейронна мережа, узагальнена регресійна нейро-фаззі мережа, нечітка машина опорних векторів, ядерна функція активації.

### ВСТУП

В цей час штучні нейронні мережі отримали широке поширення для вирішення різноманітних задач Data Mining, інтелектуального управління, прогнозування, розпізнавання образів тощо, в умовах невизначеності, нелінійності, стохастичності, хаотичності, різного роду збурень і завад, завдяки своїм універсальним апроксимуючим властивостям і можливостями навчання за даними, що характеризують функціонування явища або об'єкта, що досліджуються.

Процес навчання, як правило, ґрунтується на використанні тієї або іншої процедури оптимізації прийнятого критерію, при цьому швидкість збіжності такої процедури може бути досить низькою, особливо при навчанні багатошарових мереж, що створює істотні проблеми в ситуаціях, коли навчальна вибірка задана не пакетом, а у вигляді послідовності спостережень, що надходять у online режимі.

Прискорити цей процес можливо в нейронних мережах, чий вихідний сигнал лінійно залежить від синаптичних ваг, наприклад, радіально-базисних (RBFN) [1, 2] і нормалізованих радіально-базисних (NRBFN) [3, 4] нейронних мереж, однак, їх використання часто ускладнюється, так званим, прокльоном розмірності. І справа тут зовсім не у виникаючих обчислювальних складнощах, а в тому, що наявних даних просто може не вистачити для визначення великої кількості синаптичних ваг.

Альтернативою навчанню, яке базується на оптимізації, є навчання, що базується на пам'яті [3] і пов'язане з концепцією «нейрони в точках даних» [5]. Найбільш характерним представником систем, навчання яких засновано на цьому принципі, є узагальнені регресійні нейронні мережі [6], проте, по-перше, вони вирішують задачу інтерполяції, а не апроксимації, а, по-друге, кількість нейронів в цих мережах визначається числом спостережень у навчальній вибірці.

Є доцільною розробка системи, яка в процесі настроювання використовувала б обидва принципи навчання, налаштовуючи при цьому не тільки синаптичні ваги, але і свою архітектуру, еволюціонуючи у часі і пристосовуючись до розв'язуваної задачі [7]. Спроба синтезу такої мережі була зроблена в [8, 9], де запропонована система складалася з кількох паралельно працюючих нейронних мереж з однаковою архітектурою, але навчених на основі різних принципів, і блоку оптимізації, що об'єднує виходи цих мереж та синтезує оптимальний вихідний сигнал системи в цілому.

У порівнянні з нейронними мережами більші можливості мають нейро-фаззі системи [2, 10–12], що поєднують в собі можливості до навчання, апроксімації та лінгвістичної інтерпретації отриманих результатів. Найбільшого поширення серед цих систем отримала ANFIS [10], вихідний шар якої може бути налаштований за допомогою традиційних процедур ідентифікації. Загалом же абсолютна більшість нейро-фаззі систем навчається на основі тих чи інших процедур оптимізації.

У зв'язку з цим, доцільним є поширення підходу, що використаний у [8, 9], на навчання нейро-фаззі систем.

## 1. АРХІТЕКТУРА НЕЙРО-ФАЗЗІ СИСТЕМИ ТА ЇЇ НАВЧАННЯ НА ОСНОВІ ОПТИМІЗАЦІЇ

Архітектура еволюційної нейро-фаззі системи, що розглядається, наведена, на рис. 1 і складається з п'ятьох послідовно з'єднаних шарів. На вхідний (нульовий) шар подається $(n \times 1)$ – вимірний вектор вхідних сигналів $x(k) = (x_1(k), x_2(k), ..., x_n(k))^T, k = 1, 2, ..., N$, що підлягає обробці. Перший прихований шар містить $nh$ функцій належності $\mu_{il}(x), i = 1, 2, ..., n; l = 1, 2, ..., h$ і виконує фаззіфікацію вхідних змінних, при цьому значення $h$ може змінюватися в процесі навчання (еволюції) системи. Другий прихований шар здійснює агрегування рівнів належностей, обчислених у першому шарі, і складається з $h$ блоків множення. Третій прихований шар – це шар синаптичних ваг, що підлягають визначенню в процесі навчання системи. Четвертий шар утворено двома суматорами й обчислює суми вихідних сигналів другого й третього шарів. І, нарешті, у п'ятому (вихідному) шарі проводиться нормалізація, в результаті якої обчислюється вихідний сигнал системи $\hat{y}^{NF}$.

Таким чином, якщо на вхід системи поданий векторний сигнал $x$, елементи першого шару обчислюють рівні належності $0 < \mu_{li} \le 1$, при цьому в якості функцій належності використовуються дзвонуваті конструкції з не строго локальним рецепторним полем, що дозволяє уникнути виникнення «дір» у фаззіфікованому просторі [4]. Найчастіше це ґавсіани

$$\mu_{li}(x) = \exp\left(-\frac{(x_i - c_{li})^2}{2\sigma_i^2}\right),$$

де $c_{li}$ – параметр центру, $\sigma_i$ – параметр ширини, що обирається емпірично або настроюється в процесі навчання за допомогою процедури зворотного поширення помилок. Помітимо також, що попереднє нормування даних на обмежений інтервал, наприклад, $-1 \le x_i \le 1$, у ряді випадків спрощує обчислення, оскільки параметр ширини $\sigma_i$ може бути прийнятий однаковим для всіх компонентів вхідного вектора-образа.

На виходах другого прихованого шару обчислюються аґреговані значення $\prod_{i=1}^{n}\mu_{li}(x)$, при цьому для ґавсіанів з однаковими параметрами ширини $\sigma$

$$\prod_{i=1}^{n}\mu_{li}(x) = \prod_{i=1}^{n}\exp\left(-\frac{(x_i - c_{li})^2}{2\sigma^2}\right) = \exp\left(-\frac{\|x - c_l\|^2}{2\sigma^2}\right)$$

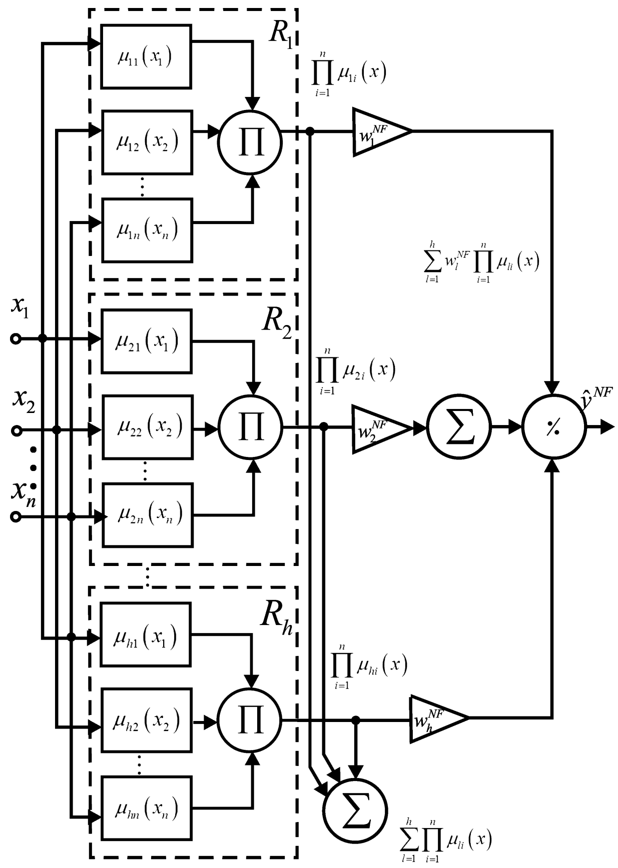


**Рис. 1.** Еволюційна нейро-фаззі система

(тут $c_l = (c_{l1}, c_{l2}, ..., c_{ln})^T$), тобто блоки системи обведені пунктиром, фактично обробляють інформацію як радіально-базисні нейрони ($R$-нейрони) нейронних мереж.

Виходами третього прихованого шару є значення $w_l^{NF}\prod_{i=1}^{n}\mu_{li}(x)$ (тут $w_l^{NF}, l = 1, 2, ..., h$ – синаптичні ваги), четвертого – $\sum_{l=1}^{h} w_l^{NF}\prod_{i=1}^{n}\mu_{li}(x)$ і $\sum_{l=1}^{h}\prod_{i=1}^{n}\mu_{li}(x)$ й, нарешті, на виході системи в цілому з'являється сигнал

$$\hat{y}^{NF}(x) = \frac{\sum_{l=1}^{h} w_l^{NF}\prod_{i=1}^{n}\mu_{li}(x)}{\sum_{l=1}^{h}\prod_{i=1}^{n}\mu_{li}(x)} = \sum_{l=1}^{h} w_l^{NF}\frac{\prod_{i=1}^{n}\mu_{li}(x)}{\sum_{l=1}^{h}\prod_{i=1}^{n}\mu_{li}(x)} = \sum_{l=1}^{h} w_l^{NF}\varphi_l^{NF}(x) = \left(w^{NF}\right)^T \varphi^{NF}(x), \tag{1}$$

де

$$\varphi_l^{NF}(x) = \prod_{i=1}^{n}\mu_{li}(x)\left(\sum_{l=1}^{h}\prod_{i=1}^{n}\mu_{li}(x)\right)^{-1}, w^{NF} = \left(w_1^{NF}, w_2^{NF}, ..., w_h^{NF}\right),$$

$$\varphi^{NF}(x) = \left(\varphi_1^{NF}(x), \varphi_2^{NF}(x), ..., \varphi_h^{NF}(x)\right)^T.$$

Нескладно бачити, що розглянута система реалізує нелінійне відображення входів у вихідний сигнал подібно нормалізованій радіально-базисній мережі (NRBFN), а по архітектурі збігається із системами Такаґі-Суґено-Канґа нульового порядку [10, 13], Ванґа-Менделя [1, 11, 14] і структурою Ларсена [2].

Навчання синаптичних ваг цих систем (надалі будемо позначати їх $w^{TS}$) здійснюється шляхом оптимізації (мінімізації) критерію навчання

$$E^{TS}=\frac{1}{2}\sum_{k=1}^{N}\left(y(k)-(w^{TS})^{T}\varphi^{NF}(x(k))\right)^{2} \qquad (2)$$

що веде до стандартної оцінки найменших квадратів

$$w^{TS}(N)=\left(\sum_{k=1}^{N}\varphi^{NF}(x(k))(\varphi^{NF}(x(k)))^{T}\right)^{-1}\left(\sum_{k=1}^{N}\varphi^{NF}(x(k))y(k)\right), \qquad (3)$$

де $y(k)$ – зовнішній навчальний сигнал.

Вводячи у розгляд $(h\times N)$-матрицю значень функцій $\Phi^{NF}=(\varphi^{NF}(x(1)),\ldots,\varphi^{NF}(x(k)),\ldots,\varphi^{NF}(x(N)))$ і $(N\times 1)$-вектори $Y=(y(1),\ldots,y(k),\ldots,y(N))^{T}$, $\hat{Y}^{TS}=(\hat{y}^{TS}(1),\ldots,\hat{y}^{TS}(k),\ldots,\hat{y}^{TS}(N))^{T}$, можна переписати співвідношення (1)–(3) у вигляді

$$\hat{Y}^{TS}=(\Phi^{NF})^{T}w^{TS},$$

$$E^{TS}=\frac{1}{2}\left\|Y-(\Phi^{NF})^{T}w^{TS}\right\|^{2}, \qquad (4)$$

$$w^{TS}(N)=\left(\Phi^{NF}\left(\Phi^{NF}\right)^{T}\right)^{-1}\Phi^{NF}Y.$$

Якість апроксимації, що забезпечується розв'язком задачі оптимізації, у ряді випадків можна підвищити, використовуючи замість критерію навчання (4), його регуляризовану модифікацію

$$E_{R}^{TS}=\frac{1}{2}\left\|Y-(\Phi^{NF})^{T}w_{R}^{TS}\right\|^{2}+\frac{\delta}{2}\left\|w_{R}^{TS}\right\|^{2},$$

що веде до ридж-оцінок синаптичних ваг

$$w_{R}^{TS}(N)=(\Phi^{NF}(\Phi^{NF})^{T}+\delta I_{h,h})^{-1}\Phi^{NF}Y, \qquad (5)$$

де $\delta>0$ – параметр регуляризації, $I_{h,h}-(h\times h)$ – одинична матриця.

Якщо дані надходять на обробку послідовно в on-line режимі, для отримання оцінок (3)–(5) може бути використаний рекурентний метод найменших квадратів:

$$w^{TS}(k+1)=w^{TS}(k)+\frac{P(k)(y(k+1)-(w^{TS}(k))^{T}\varphi^{NF}(x(k+1)))}{1+(\varphi^{NF}(x(k+1)))^{T}P(k)\varphi^{NF}(x(k+1))}\times$$

$$\times\varphi^{NF}(x(k+1)),$$

$$P(k+1)=P(k)-\frac{P(k)\varphi^{NF}(x(k+1))(\varphi^{NF}(x(k+1)))^{T}P(k)}{1+(\varphi^{NF}(x(k+1)))^{T}P(k)\varphi^{NF}(x(k+1))},$$

причому для ридж-оцінки (5) у якості початкових умов приймається

$$P(0)=\delta^{-1}I_{h,h}.$$

## 2. УЗАГАЛЬНЕНА РЕГРЕСІЙНА НЕЙРО-ФАЗЗІ СИСТЕМА

Як відзначалося вище, альтернативою навчанню, заснованому на оптимізації, є навчання, яке базується на пам'яті, що лежить в основі узагальненої регресійної нейронної мережі (GRNN). В [15–17] була запропонована узагальнена регресійна нейро-фаззі мережа (GRNFN), що є поширенням мережі Шпехта на нейро-фаззі архітектуру, наведену на рис. 1, та яка має низку переваг у порівнянні зі своїм нейромережевим прототипом.

Синтез, навчання й еволюція GRNFN відбуваються в такий спосіб. Вектори навчальної вибірки $x(1),\ldots,x(k),\ldots,x(h)$, $h\le N$, попередньо нормуються так, що $-1\le x_i\le 1$, $i=1,2,\ldots,n$, після чого для кожного з $x(k)$ у першому прихованому шарі формується власний набір функцій належності $\mu_{k1}(x),\mu_{k2}(x),\ldots,\mu_{kn}(x)$ так, що їх центри співпадають з $x_i(k)$, тобто

$$\mu_{ki}(x)=\exp\left(-\frac{(x_i-x_i(k))^{2}}{2\sigma_i^{2}}\right).$$

Одночасно з формуванням функцій належності першого прихованого шару в третьому шарі відбувається установка синаптических ваг $w_k^{GF}$ (індексом GF позначено параметри й сигнали GRNFN), які покладаються рівними навчальним сигналам $y(k)$. Таким чином, при подачі на вхід GRNFN довільного сигналу $x$, в першому прихованому шарі обчислюються рівні належності $\mu_{ki}(x_i)$, $i=1,2,\ldots,n$; $k=1,2,\ldots,h$; у другому шарі відбувається їхнє агрегування шляхом формування багатовимірних функцій

$$y^{GF}(x(k))=\prod_{k=1}^{h\le N}\mu_{ki}(x_i),$$

у третьому шарі обчислюються добутки $y(k)\varphi^{GF}(x(k))$, четвертий шар обчислює значення сигналів $\sum_{k=1}^{h}y(k)\varphi^{GF}(x(k))$ і $\sum_{k=1}^{h}\varphi^{GF}(x(k))$ й, нарешті, у вихідному шарі формується оцінка вихідного сигналу

$$\hat{y}^{GF}(x)=\frac{\sum_{k=1}^{h}y(k)\varphi^{GF}(x(k))}{\sum_{k=1}^{h}\varphi^{GF}(x(k))}=\sum_{k=1}^{h}w_k^{GF}\frac{\prod_{i-1}^{n}\mu_{ki}(x_i)}{\sum_{k=1}^{h}\prod_{i=1}^{n}\mu_{ki}(x_i)}=$$

$$= \sum_{k=1}^{h} w_k^{GF} \varphi^{GF}(x(k)) = (w^{GF})^T \varphi^{GF}(x).$$

Процес навчання GRNFN може протікати як у пакетному режимі, коли задані навчальна вибірка $\{x(k), y(k)\}$ й максимальне число функцій належності для кожного входу $h$, так і в on-line режимі, коли пари $x(k)$, $y(k)$ надходять на обробку послідовно, формуючи багатовимірні функції $\varphi_k^{GF}$. При цьому досить просто організувати процес виключення спостережень, що містять не суттєву інформацію. Якщо для якогось $x(k)$ виконується нерівність

$$\max_i D_i^{\min}(x_i(k)) < r < \frac{2}{h-1} \qquad (6)$$

(тут $D_i^{\min}(x_i(k))$ – найменша відстань між $x_i(k)$ і раніше сформованими центрами функцій належності), то $x(k)$ не формує нову $\varphi^{GF}(x(k))$ й виключається з розгляду. Помітимо, що в нейро-фаззі системі з одновимірними функціями, граничний параметр $r$ і відстань $D_i^{\min}(x_i(k))$ визначити значно простіше ніж в GRNN з багатовимірними функціями активації, де загальна рекомендація до вибору $r$ визначається малим перекриттям сусідніх багатовимірних ґавсіанів [18].

Роботу GRNFN нескладно організувати в режимі безперервного еволюціонування, що важливо при обробці суттєво нестаціонарних сигналів. Тут можливе використання двох підходів: на ковзному вікні з $h$ спостережень, коли при надходжені на вхід системи навчальної пари $x(h+1)$, $y(h+1)$ у першому й третьому шарах виключаються всі $\mu_{1i}$ й $w_1^{GF}$, а замість них установлюються $\mu_{h+1,i}$ й $w_{h+1}^{GF}$, і заснований на нерівності (6). У цьому випадку нова пара замінює найближчу до неї пару «старих» даних.

Оскільки процес навчання відбувається практично миттєво, еволюція нейро-фаззі системи також відбувається дуже швидко.

## 3. НЕЧІТКА МАШИНА ОПОРНИХ ВЕКТОРІВ

Своєрідним гібридом систем, які базуються як на оптимізації, так і на пам'яті одночасно, є машини опорних векторів (SVM), які є нейронними мережами, що збігаються по архітектурі з RBFN і GRNN, їх синаптичні ваги визначаються в результаті розв'язання задачі оптимізації, а центри активаційних функцій установлюються за принципом «нейрони в точках даних». І хоча ці мережі мають цілу низку безсумнівних переваг [19, 20], їхнє навчання з обчислювальної точки зору є досить трудомістким, оскільки пов'язане з розв'язанням задачі нелінійного програмування високої розмірності.

У зв'язку із цим, у якості альтернативи SVM були запропоновані машини опорних векторів, засновані на методі найменших квадратів (LS-SVM) [21], навчання яких зводиться до розв'язання системи лінійних рівнянь.

Фаззі-аналогом SVM є нечітка машина опорних векторів (FSVM) [22], у якій багатовимірні активаційні функції замінені одновимірними функціями належності. В [23] була розглянута FSVM, заснована на методі найменших квадратів (LS-FSVM) і призначена для розв'язання задачі розпізнавання образів, що коли в процесі навчання сигнали $y(k)$ можуть приймати значення $+1$ або $-1$.

У межах досліджуваної нами задачі перетворення, що реалізується LS-FSVM (далі використовується позначення FS), може бути записане у вигляді

$$\hat{y}^{FS}(x) = \sum_{k=1}^{N} w_k^{FS} \frac{\prod_{i=1}^{n} \mu_{ki}(x_i)}{\sum_{k=1}^{N}\prod_{i=1}^{n} \mu_{ki}(x_i)} = \sum_{k=1}^{N} w_k^{FS} \varphi^{FS}(x(k)) =$$

$$= (w^{FS})^T \varphi^{FS}(x), \qquad (7)$$

а її навчання зводиться до установки центрів функцій належності $\mu_{ki}(x_i)$ в точках $x_i(k)$ і оптимізації квадратичного критерія навчання

$$E^{FS} = \frac{1}{2}\left\|w^{FS}\right\|^2 + \frac{\gamma}{2}\sum_{k=1}^{N} e^2(k) \qquad (8)$$

з урахуванням обмежень у вигляді системи з $N$ лінійних рівнянь

$$y(k) = (w^{FS})^T \varphi^{FS}(x(k)) + e(k) \qquad (9)$$

де $\gamma > 0$ – параметр регуляризації,

$$e(k) = y(k) - \hat{y}^{FS}(x(k)).$$

Оптимізація критерію (8) без урахування обмежень (9) веде до виразу

$$w^{FS}(N) = \left(\sum_{k=1}^{N} \varphi^{FS}(x(k))(\varphi^{FS}(x(k)))^T + \gamma^{-1} I_{N,N}\right)^{-1} \times$$

$$\times \left(\sum_{k=1}^{N} \varphi^{FS}(x(k)) y(k)\right),$$

що є по суті ридж-оцінкою при $\delta = \gamma^{-1}$.

Для урахування системи обмежень (9) введемо в розгляд функцію Лаґранжа

$$L(w^{FS}, e(k), \lambda(k)) = E^{FS} + \sum_{k=1}^{N} \lambda(k)(y(k) - (w^{FS})^T \varphi^{FS}(x(k)) - e(k)) =$$

$$= \frac{1}{2}(w^{FS})^T w^{FS} + \frac{\gamma}{2}\sum_{k=1}^{N} e^2(k) + \sum_{k=1}^{N}\left(y(k) - (w^{FS})^T \varphi^{FS}(x(k)) - e(k)\right)$$

(тут $\lambda(k) - N$ невизначених множників Лаґранжа) і систему рівнянь Каруша-Куна-Таккера:

$$\begin{cases} \nabla_{w^{FS}} L = w^{FS} - \sum_{k=1}^{N} \lambda(k)\varphi^{FS}(x(k)) = \vec{0}_N, \\ \dfrac{\partial L}{\partial e(k)} = \gamma e(k) - \lambda(k) = 0, \\ \dfrac{\partial L}{\partial \lambda(k)} = y(k) - (w^{FS})^T \varphi^{FS}(x(k)) - e(k) = 0, \end{cases} \tag{10}$$

де $\vec{0}_N - (N \times 1)$ – вектор, утворений нулями.

З (10) випливає

$$\begin{cases} w^{FS}(N) = \sum_{k=1}^{N} \lambda(k)\varphi^{FS}(x(k)), \\ \lambda(k) = \gamma e(k), \\ y(k) = (w^{FS})^T \varphi^{FS}(x(k)) + e(k), \end{cases} \tag{11}$$

або в матричному вигляді

$$\left(\gamma^{-1} I_{N,N} + \Omega\right) \begin{pmatrix} \lambda(1) \\ \lambda(2) \\ \vdots \\ \lambda(N) \end{pmatrix} = \begin{pmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{pmatrix},$$

або

$$\left(\gamma^{-1} I_{N,N} + \Omega\right)\vec{\lambda} = Y,$$

(тут $\Omega = \left\{\Omega_{ij} = (\varphi^{FS}(x(i)))^T \varphi^{FS}(x(j))\right\}$),

звідки

$$\vec{\lambda} = (\gamma^{-1} I_{N,N} + \Omega)^{-1} Y. \tag{12}$$

Остаточно (7) з урахуванням (11) і (12) набуває вигляд

$$\hat{y}^{FS}(x) = \left(\sum_{k=1}^{N} \lambda(k)\varphi^{FS}(x(k))\right)^T \varphi^{FS}(x).$$

## 4. ОПТИМАЛЬНИЙ АНСАМБЛЬ НЕЙРО-ФАЗЗІ СИСТЕМ

При розв'язанні конкретної задачі не можна встановити заздалегідь яка з окремих синтезованих систем (TS, GRNFN, FSVM або яка-небудь ще) покаже найкращі результати. Більше того, при обробці нестаціонарних сигналів на різних часових інтервалах, найкращими можуть виявитися різні системи. У цій ситуації доцільно скористатися ансамблем одночасно працюючих систем [24–27] з наступним об'єднанням їх вихідних сигналів з метою одержання оптимального результату так, як це показане на рис. 2.

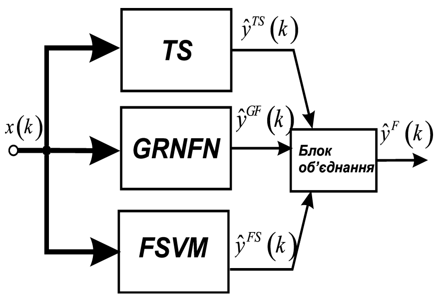


**Рис. 2.** Ансамбль нейро-фаззі систем

Вводячи в розгляд вектори вихідних сигналів $\hat{y}(k) = \left(\hat{y}^{TS}(k), \hat{y}^{GF}(k), \hat{y}^{FS}(k)\right)^T$ і параметрів об'єднання $w^F = \left(w_{TS}^F, w_{GF}^F, w_{FS}^F\right)^T$, будемо шукати оптимальний вихідний сигнал у вигляді

$$\hat{y}^F(k) = (w^F)^T \hat{y}(k)$$

при обмеженнях на незміщенність

$$I_3^T w^F = 1,$$

де $I_3 - (3 \times 1)$ – вектор, утворений одиницями.

Невідомий вектор параметрів $w^F$ може бути визначений за допомогою методу штрафних функцій, для чого вводиться $(N \times 3)$ матриця вихідних сигналів $\hat{Y} = \left(\hat{y}(1), \hat{y}(2), ..., \hat{y}(N)\right)^T$ і критерій оптимізації

$$E^F = \left(Y - \hat{Y}w^F\right)^T \left(Y - \hat{Y}w^F\right) + \rho^{-2}\left(1 - I_3^T w^F\right), \tag{13}$$

де $\rho$ – штрафний коефіцієнт.

Мінімізація (13) по $w^F$ веде до виразу

$$w^F(\rho) = \left(\hat{Y}^T\hat{Y} + \rho^{-2} I_3 I_3^T\right)^{-1}\left(\hat{Y}^T Y + \rho^{-2} I_3\right),$$

яке після нескладних перетворень [28] і прирівнювання $\rho$ до нуля, набуває вигляд

$$w^F(N) = \lim_{\rho \to 0} w^F(\rho) = w^{LS}(N) + P^F(N)\frac{1 - I_3^T w^{LS}(N)}{I_3^T P^F(N) I_3}, \tag{14}$$

де

$$w^{LS}(N) = \left(\hat{Y}^T\hat{Y}\right)^{-1}\hat{Y}^T Y = P^F(N)\hat{Y}^T Y \tag{15}$$

– звичайна оцінка найменших квадратів типу (4).

Елементам вектора $w^F(N)$ (14) можна надати зміст належностей кожної з підсистем ансамблю TS, GRNFN і FSVM до деякої гіпотетичної оптимальної системи, якщо забезпечити їх невід'ємність у процесі об'єднання.

Вводячи у розгляд лаґранжіан

$$L\left(\mu^F,\lambda^F,\rho^F\right)=\left(Y-\hat{Y}\mu^F\right)^T\left(Y-\hat{Y}\mu^F\right)+\lambda^F\left(I_3^T\mu^F-1\right)-\left(\rho^F\right)^T\mu^F,$$

(тут $\mu^F$ – $(3\times1)$ – вектор рівнів належності, $\lambda^F$ – невизначений множник Лагранжа, $\rho^F$ – $(3\times1)$ – вектор невід'ємних множників Лаґранжа, що відповідають умовам додаткової нежорсткості), і систему рівнянь

$$\begin{cases}\nabla_{\mu^F} L=\vec{0}_3,\\ \dfrac{\partial L}{\partial\lambda^F}=0,\end{cases}$$

розв'язок якої має вигляд

$$\begin{cases}\mu^F=P^F(N)\left(\hat{Y}^T\hat{Y}+0,5\lambda^F I_3+0,5\rho^F\right),\\ \lambda^F=\dfrac{I_3^T P^F(N)\hat{Y}^T Y-1+0,5I_3^T P^F(N)\rho^F}{0,5I_3^T P^F(N)I_3},\end{cases}$$

і використовуючи процедуру Ерроу-Гурвіца-Удзави [29], одержуємо алгоритм навчання вектора $\mu^F$ у вигляді

$$\begin{cases}\mu^F(k+1)=w^{LS}(k+1)-P^F(k+1)\times\\ \times\left(I_3^T w^{LS}(k+1)-1+0,5I_3^T P^F(k+1)\rho^F(k)\right)\left(I_3^T P^F(k+1)I_3\right)^{-1}I_3+0,5P^F(k+1)\rho^F(k),\\ \rho^F(k+1)=\mathrm{Pr}_+\left(\rho^F(k)-\eta_\rho(k+1)\mu^F(k+1)\right),\end{cases}\qquad(16)$$

де $w^{LS}$ – визначається виразом (15), $\mathrm{Pr}_+(\cdot)$ – проектор на додатний ортант.

Нескладно показати, що алгоритм (16) елементарно поширюється на довільну кількість систем, що входять в ансамбль.

## ВИСНОВКИ

Введено систему еволюційних нейро-фаззі систем, що використовують різні принципи навчання та налаштовують не тільки синаптичні ваги, але й свою архітектуру. Запропоновано адаптивну процедуру об'єднання цих систем, що дозволяє синтезувати оптимальний вихідний сигнал і встановлювати рівні належності до деякої гіпотетичної оптимальної системи. Підхід, що розвивається, відрізняється обчислювальною простотою й дозволяє обробляти інформацію в on-line режимі в міру її надходження.

## СПИСОК ЛІТЕРАТУРИ


1. *Осовский, С.* Нейронные сети для обработки информации / С. Осовский. – М. : Финансы и статистика, 2002. – 344 с.
2. *Rutkowski, L.* Computational Intelligence. Methods and Tehniques / L. Rutkwski. – Berlin-Heidelberg : Springer-Verlag, 2008. – 514 p.
3. *Nelles, O.* Nonlinear System Identification / O. Nelles. – Berlin : Springer, 2001. – 785 p.
4. *Friedman, J.* The Elements of Statistical Learning. Data Mining, Inference and Prediction / J. Friedman, T. Hastie, R. Tibshirani. – Berlin : Springer, 2003.–552 p.
5. Pattern recognition using radial basis function network / D. Zahirniak, R. Chapman, S. Rogers [and other] // Dayton, OH : Application of AI Conf. – 1990. – P. 249–260.
6. *Specht, D. F.* A general regression neural network / D. F. Specht // IEEE Trans. on Neural Networks. – 1991. – Vol. 2. – P. 568–576.
7. *Kasabov, N.* Evolving Connectionist Systems / N. Kasabov. – London : Springer – Verlag, 2003. – 307 p.
8. *Bodyanskiy, Ye.* Hibrid evolving neural netvork using kernel activation functions / Ye. Bodyanskiy, N. Teslenko, P. Grimm // Conf. Proc. 17th Zittau East-West Fuzzy Coll. – Zittau Goerlitz : HS. – 2010. – P. 39–46.
9. *Бодянський, Є. В.* Еволюційна нейронна мережа з ядерними функціями активації та адаптивний алгоритм її навчання / Є. В. Бодянський, А. О. Дейнеко, Н. О. Тесленко. – Наукові праці – Вип. 130. – Т. 143. – Комп'ютерні технології. – Миколаїв : Вид-во ЧДУ ім. Петра Могили, 2010. – С. 71–78.
10. *Jang, J.-S.* Neuro-Fuzzy and Soft Computing : A Computational Approach to Learning and Maching Intelligence / J.-S. Jang, C.-T. Sun, E. Mizutani // Upper Saddle River : Prentice Hall. – 1997. – 640 p.
11. *Wang, L.-X.* Fuzzy basis functions, universal approximation and orthogonal least squares learning / L.-X. Wang, J. M. Mendel // IEEE Trans. on Neural Networks. – 1993. – Vol. 3. – P. 807–814.
12. *Cios, K. J.* Neuro-fuzzy algorithms / K. J. Cios, W. Pedrycz // Oxford : IOP Publishing Ltd and Oxford University Press. – Handbook of Neural Computation, 1997. – D1. 3:1 – D1. 3:7.
13. *Takagi, T.* Fuzzy identification of systems and its applications to modeling and control / Takagi T., Sugeno M. // IEEE Trans. on Systems, Man, and Cybernetics. – 1985. – Vol. 15. – P. 116–132.
14. *Wang, L.-X.* Adaptive Fuzzy Systems and Control. Design and Stability Analysis / L.-X. Wang // Upper Saddle, N. J. : Prentice Hall, 1994. – 256 p.
15. *Bodyanskiy, Ye.* Generalized regression neuro-fuzzy network / Ye. Bodyanskiy, N. Teslenko. – Proc. XIII-th Int. Conf. «Information Reaserch & Application», i. TECH 2007. – V. 1. – Varna, 2007. – P. 219–225.
16. *Bodyanskiy, Ye.* Nonliner process identification and modeling using general regression neuro-fuzzy network / Bodyanskiy Ye., Otto P., Pliss I, Teslenko N. // Proc. 52 nd Int. Sci. Coll. «Computer Science Meets Automation.» – TU Ilmenau (Thuer.) – 2007. – P. 23–27.
17. *Bodyanskiy, Ye.* General regression neuro-fuzzy network for identification of nonstationary plants / Ye. Bodyanskiy, N. Teslenko // Int. J. Informaion Technologies and Knowledge, 2008. – Vol. 2, № 2 – P. 136–142.

18. *Bishop, C. M.* Neural Networks for Pattern Recognition / C. M. Bishop // Oxford : Clarendon Press, 1995. – 482 p.
19. *Vapnik, V. N.* The Nature of Statistical Learning Theory / V. N. Vapnik – N. Y. : Springer, 1995. – 188 p.
20. *Vapnik, V. N.* Statistical Learning Theory: Adaptive and Learning Systems / V. N. Vapnik – N. Y. : John Wiley & Sons, 1998. – 736 p.
21. Least Squares Support Vector Machines / [Suykens J.A.K., Gestel T.V., Brabanter J.D. and other]. – Singapore : World Scientific, 2002. – 294 p.
22. *Lin, Ch.-F.* Fuzzy Support Vector Machines / Ch.-F. Lin, Sh.-D. Wang. – IEEE Trans. on Neural Networks. – 2002. – T. 13., № 2. – P. 646–471.
23. *Tsujinishi, D.* Fuzzy Least Squares Support Vector Machines for multiclass problems / D. Tsujinishi, S. Abe. – Neural Networks, 2003. – Vol. 16. – P. 785–792.
24. *Hansen, L. K.* Neural networks ensembles / L. K. Hansen, P. Salamon // IEEE Trans. on Pattern Analysis and Machine Intelligence. – 1990. – Vol. 12. – P. 993–1000.
25. *Sharkey, A. J. C.* On combining artificial neural nets / A. J. C. Sharkey // Connect. Sci. – 1996. – Vol. 8. – P. 299–313.
26. *Hashem, S.* Optimal linear combination of neural networks / S. Hashem // Neural Networks.– 1997. – Vol. 10. – P. 599–614.
27. *Naftaly, U.* Optimal ensemble averaging of neural networks / U. Naftaly, N. Intrator, D. Horn // Network : Comput. Neural Syst. – 1997. – Vol. 8. – P. 283–296.
28. *Бодянський Є. В.* Адаптивне виявлення розладнань в об'єктах керування за допомогою штучних нейронних мереж / Є. В. Бодянський, О. І. Михальов , І. П. Плісс. – Дніпропетровськ : Системні технології, 2000. – 140 с.
29. *Поляк, Б. Т.* Введение в оптимизацию / Б. Т. Поляк. – М. : Мир, 1984. – 541 с.
30. *Бодянский, Е. В.* Адаптивное обобщенное прогнозирование многомерных случайных последовательностей / Е. В. Бодянский , И. П. Плисс, Т. В. Соловьева. – Доклады АН УССР. – 1989. – А. – № 9. – С. 73–75.


Стаття надійшла до редакції 20.02.2012.

Дейнеко А. А., Плисс И. П., Бодянский Е. В.

КОМБИНИРОВАННОЕ ОБУЧЕНИЕ ЭВОЛЮЦИОННОЙ НЕЙРО-ФАЗЗИ СИСТЕМЫ

Предложена архитектура эволюционной нейро-фаззи системы, которая не склонна к проклятию размерности, способна обрабатывать информацию в режиме реального времени, адаптируя при этом свои параметры и структуру к условиям задачи. В качестве активационных функций были использованы ядерные функции активации.

**Ключевые слова:** эволюционная нейро-фаззи система, нормализованная радиально-базисная нейронная сеть, обобщенная регрессионная нейро-фаззи сеть, нечеткая машина опорных векторов, ядерная функция активации.

Deineko A. A., Pliss I. P., Bodyanskiy Ye.

EVOLVING NEURO-FUZZY SYSTEM COMBINED LEARNING

In this work the evolving neuro-fuzzy system with kernel activation function that contains fuzzy support vector machine, normalized radial basis function neural network and general regression neuro-fuzzy network as subsystems is proposed. This network is tuned using both optimization and memory based approaches and does not inclined to the «curse of dimensionality», is able to real time mode information processing by adapting its parameters and structure to problem conditions.

**Key words**: evolving neuro-fuzzy system, normalized radial-basis function neural network, general regression neuro-fuzzy network, fuzzy support vector machine, kernel activation function.

*UDC 519.832.4*

Romanuke V. V.

*Candidate of Technical Science, docent of Khmelnytskyy National University*

# OPTIMAL STRATEGIES CONTINUUM FOR PROJECTING THE FOUR-MOUNT CONSTRUCTION UNDER INTERVAL UNCERTAINTIES WITH INCORRECTLY PRE-EVALUATED TWO LEFT AND ONE RIGHT ENDPOINTS

There is investigated a two-person game model of optimizing cross-section squares of the four-mount construction, where the model kernel is defined on the six-dimensional hyperparallelepiped as the product of three closed intervals of unit-normed loads and of three closed intervals of unit-normed cross-section squares. For the case of incorrectly pre-evaluated two left and one right endpoints of those interval uncertainties there has been proved that the projector may obtain an optimal strategies continuum. A criterion for singularizing that continuum has been proposed.

**Key words:** optimizing cross-section squares, two-person game model, four-mount construction, incorrect pre-evaluation.

## INVESTIGATION AREA

There are many uncertain factors in building mount constructions, one of which is interval-valued potential load on the construction pivots, pillars, bars or other mount elements [1, 2]. If the potential load on the construction with four mounts is unit-normed, then the unit-normed load on the $i$-th mount $x_i$ is enclosed within the closed interval $\left[a_i;\, b_i\right] \subset \left(0;\, 1\right)$ by $b_i > a_i$ for $i = \overline{1,\, 3}$ [3, 4]. The nonzero