УДК 004.032.26

Тищенко О. К.[1], Плісс І. П.[2], Копаліані Д. С.[3]

[1]Канд. техн. наук, старший науковий співробітник Проблемної НДЛ АСУ Харківського національного університету радіоелектроніки, Україна, E-mail: lehatish@gmail.com

[2]Канд. техн. наук, старший науковий співробітник, провідний науковий співробітник Проблемної НДЛ АСУ Харківського національного університету радіоелектроніки, Україна

[3]Аспірантка, Харківський національний університет радіоелектроніки, Україна

ГІБРИДНА КАСКАДНА ОПТИМІЗОВАНА НЕЙРОННА МЕРЕЖА

Запропоновано нову архітектуру та алгоритми навчання для гібридної каскадної нейронної мережі з оптимізацією пулу нейронів у кожному каскаді. Запропонована гібридна каскадна нейронна мережа забезпечує обчислювальну простоту та характеризується як слідкуючими, так і фільтруючими властивостями.

Ключові слова: нейронна мережа, оптимальне навчання, обчислювальний інтелект, еволюціонуюча гібридна система.

ВСТУП

У цей час штучні нейронні мережі (ANNs) отримали широке поширення для розв'язання широкого класу проблем, пов'язаних з обробкою інформації, заданої або у формі таблиць «об'єкт – властивість», або часових рядів, що часто породжуються нестаціонарними нелінійними стохастичними або хаотичними системами. Переваги ANNs перед іншими підходами пояснюються, перш за все, їх універсальними апроксимуючими можливостями і здатністю до навчання.

Традиційно під навчанням розуміють процес налаштування синаптичних ваг мережі за допомогою тієї чи іншої процедури оптимізації, що відшукує екстремум заздалегідь заданого критерію навчання [1, 2]. Якість навчання може бути покращена шляхом настроювання не тільки синаптичних ваг, але й власне архітектури мережі. Ця ідея лежить в основі так званих еволюційних систем обчислювального інтелекту [3, 4], які отримують у теперішній час усе більш широке поширення. У рамках цього підходу можна виділити каскадні нейронні мережі [5–8] завдяки їх високій ефективності та простоті налаштування як синаптичних ваг, так і власне архітектури. Ця мережа стартує з найпростішої архітектури (перший каскад), утвореної пулом [5] нейронів, які навчаються незалежно. Кожен з нейронів пулу може відрізнятися від інших або активаційною функцією, або методом навчання, при цьому нейрони пулу в процесі навчання між собою не взаємодіють. Після того, як усі нейрони пулу першого каскаду налаштовані, з них обирається один найкращий у сенсі прийнятого критерію, всі ж інші видаляються, в результаті чого і формується перший каскад, утворений єдиним нейроном, синаптичні ваги якого надалі не налаштовуються – «заморожуються».

Після цього формується другий каскад, який, як правило, утворено пулом тих же нейронів з тією лише різницею, що ці нейрони мають додатковий вхід (а, отже, і додаткову синаптичну вагу), утворений виходом першого каскаду. Надалі все відбувається аналогічно до попереднього каскаду, в результаті чого другий каскад також складається з єдиного найкращого нейрону із замороженими вагами. Нейрони третього каскаду мають вже по два додаткових входи: виходи першого і другого каскадів, надалі все відбувається аналогічно до попереднього каскаду. Процес нарощування каскадів еволюційної архітектури продовжується доти, доки не буде досягнуто необхідної якості розв'язання задачі на навчальній вибірці.

Автори найбільш популярної каскадної нейронної мережі CasCorLa Фальман та Леб'єр у якості нейронів мережі використовували елементарні персептрони Ф.Розенблатта із традиційними сигмоїдальними активаційними функціями, синаптичні ваги яких налаштовуються за допомогою Quickprop-алгоритму, що є модифікацією δ-правила навчання.

У [9–16] в якості вузлів каскадної мережі були використані різні типи нейронів. Тут, однак, слід зазначити, що при роботі з різнотипними вузлами неможливо виділити в пулі єдиний найкращий нейрон. При роботі з нестаціонарними об'єктами може виникнути ситуація, коли на одній частині навчальної вибірки найкращим виявиться один нейрон, а на іншій – зовсім інший. У зв'язку з цим цілком природно в пулі зберігати всі нейрони (без визначення найкращого нейрона-переможця), а вихідний сигнал каскаду формувати шляхом об'єднання виходів усіх вузлів пулу на основі деякої оптимізаційної процедури, що породжується загальним критерієм якості роботи нейронної мережі.

Синтезу такої гібридної оптимізованої нейронної мережі і присвячено цю статтю.

АРХІТЕКТУРА КАСКАДНОЇ ОПТИМІЗОВАНОЇ НЕЙРОННОЇ МЕРЕЖІ

На вхід мережі (рецепторний шар) надходить векторний сигнал $x(k)=\left(x_1(k),x_2(k),\ldots,x_n(k)\right)^T$, де $k=1,2,\ldots$ – або номер образу в таблиці «об'єкт – властивість», або поточний дискретний час. Цей сигнал подається на входи всіх нейронів мережі $N_j^{[m]}$ ($j=1,2,\ldots,q$ – число нейронів у пулі, $m=1,2,\ldots$ – номер каскаду), на

DOI 10.15588/1607-3274-2014-1-18

виходах яких з’являються сигнали $\hat{y}_j^{[m]}(k)$. Надалі ці сигнали об’єднуються за допомогою узагальнюючого нейрону $GN^{[m]}$, котрий формує оптимальний вихід m-го *каскаду* $\hat{y}^{*[m]}(k)$. При цьому, якщо на нейрони першого каскаду подається тільки вектор $x(k)$, котрий у загальному випадку може містити і сигнал зміщення $x_0(k) \equiv 1$, то нейрони другого каскаду мають додатковий вхід для сигналу $\hat{y}^{*[1]}(k)$, третього каскаду – два додаткових входи $\hat{y}^{*[1]}(k)$, $\hat{y}^{*[2]}(k)$, m-го каскаду – (m–1) додаткових входів $\hat{y}^{*[1]}(k)$, $\hat{y}^{*[2]}(k), \dots, \hat{y}^{*[m-1]}(k)$. Каскади формуються в процесі навчання мережі, коли стає зрозуміло, що всі попередні каскади не забезпечують необхідну якість навчання.

НАВЧАННЯ ОКРЕМИХ НЕЙРОНІВ У КАСКАДНІЙ НЕЙРОННІЙ МЕРЕЖІ

Розглянемо ситуацію, коли j-ий нейрон m-го каскаду мережі є традиційним елементарним персептроном Розенблатта з активаційною функцією

$$0 < \sigma_j^{[m]}\left(\gamma_j^{[m]} u_j^{[m]}\right) = \frac{1}{1 + e^{-\gamma_j^{[m]} u_j^{[m]}}} < 1,$$

де $u_j^{[m]}$ – сигнал внутрішньої активації j-го нейрону m-го каскаду, $\gamma_j^{[m]}$ – параметр крутизни. Тоді вихідні сигнали нейронів пулу першого каскаду можуть бути представлені у формі

$$\hat{y}_j^{[1]} = \sigma_j^{[1]}\left(\gamma_j^{[1]} \sum_{i=0}^{n} w_{ji}^{[1]} x_i\right) = \sigma_j^{[1]}\left(\gamma_j^{[1]} w_j^{[1]T} x\right),$$

(тут $w_{ji}^{[1]}$ – i-а синаптична вага j-ого нейрону першого каскаду, $x = (1, x_1, x_2, \dots, x_n)^T$, при цьому звичайно вхідні сигнали за допомогою елементарного перетворення кодуються так, що $0 \le x_i \le 1$, виходи нейронів другого каскаду:

$$\hat{y}_j^{[2]} = \sigma_j^{[2]}\left(\gamma_j^{[2]}\left(\sum_{i=0}^{n} w_{ji}^{[2]} x_i + w_{j,n+1}^{[2]} \hat{y}^{*[1]}\right)\right),$$

@

виходи m-го каскаду:

$$\hat{y}_j^{[m]} = \sigma_j^{[m]}\left(\gamma_j^{[m]}\left(\sum_{i=0}^{n} w_{ji}^{[m]} x_i + w_{j,n+1}^{[m]} \hat{y}^{*[1]} + w_{j,n+2}^{[m]} \hat{y}^{*[2]} + \dots + w_{j,n+m-1}^{[m]} \hat{y}^{*[m-1]}\right)\right) = \sigma_j^{[m]}\left(\gamma_j^{[m]} \sum_{i=0}^{n+m-1} w_{ji}^{[m]} x_j^{[m]}\right) =$$
$$= \sigma_j^{[m]}\left(w_j^{[m]T} x^{[m]}\right),$$

де $x^{[m]} = \left(x^T, \hat{y}^{*[1]}, \dots, \hat{y}^{*[m-1]}\right)^T$.

Таким чином, каскадна мережа, що утворена персептронами Розенблатта та складається з m каскадів, містить $m(n+2) + \sum_{p=1}^{m-1} p$ параметрів включно з параметрами крутизни $\gamma_j^{[p]}$, $p = 1, 2, \dots, m$.

У якості критерію навчання використовуватимемо традиційну квадратичну функцію

$$\begin{aligned} E_j^{[m]}(k) &= \frac{1}{2}\left(e_j^{[m]}(k)\right)^2 = \frac{1}{2}\left(y(k) - \hat{y}_j^{[m]}(k)\right)^2 = \\ &= \frac{1}{2}\left(y(k) - \sigma_j^{[m]}\left(\gamma_j^{[m]} u_j^{[m]}(k)\right)\right)^2 = \\ &= \frac{1}{2}\left(y(k) - \sigma_j^{[m]}\left(\gamma_j^{[m]} \sum_{i=0}^{n+m-1} w_{ji}^{[m]} x_i^{[m]}(k)\right)\right)^2 = \\ &= \frac{1}{2}\left(y(k) - \sigma_j^{[m]}\left(\gamma_j^{[m]} w_j^{[m]T} x^{[m]}(k)\right)\right)^2, \end{aligned} \tag{1}$$

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

Процедура градієнтної оптимізації критерію (1) за $w_j^{[m]}$ може бути записана у рекурентній формі

$$\begin{aligned} w_j^{[m]}(k+1) &= w_j^{[m]}(k) - \eta_j^{[m]}(k+1) \nabla_{w_j^{[m]}} E_j^{[m]}(k+1) = \\ &= w_j^{[m]}(k) + \eta_j^{[m]}(k+1) e_j^{[m]}(k+1) \gamma_j^{[m]} \hat{y}_j^{[m]}(k+1) \times \\ &\times \left(1 - \hat{y}_j^{[m]}(k+1)\right) x^{[m]}(k+1) = \\ &= w_j^{[m]}(k) + \eta_j^{[m]}(k+1) e_j^{[m]}(k+1) \gamma_j^{[m]} J_j^{[m]}(k+1) \end{aligned} \tag{2}$$

(тут $\eta_j^{[m]}(k+1)$ – параметр ***кроку*** настроювання), а мінімізація (1) за параметром $\gamma_j^{[m]}$ може бути забезпечена за допомогою методу Крушке-Мовеллана) [17]:

$$\begin{aligned} \gamma_j^{[m]}(k+1) &= \gamma_j^{[m]}(k) - \eta_\gamma^{[m]}(k+1) \partial E_j^{[m]}(k+1) / \partial \gamma_j^{[m]} = \\ &= \gamma_j^{[m]}(k) + \eta_\gamma^{[m]}(k+1) e_j^{[m]}(k+1) \hat{y}_j^{[m]}(k+1) \times \\ &\times \left(1 - \hat{y}_j^{[m]}(k+1)\right) u_j^{[m]}(k+1). \end{aligned} \tag{3}$$

Об’єднуючи процедури (2) та (3), приходимо до спільного методу навчання j-го нейрону m-го каскаду:

$$\begin{pmatrix} w_j^{[m]}(k+1) \\ \gamma_j^{[m]}(k+1) \end{pmatrix} = \begin{pmatrix} w_j^{[m]}(k) \\ \gamma_j^{[m]}(k) \end{pmatrix} + \eta_j^{[m]}(k+1) e_j^{[m]}(k+1) \times$$
$$\times \hat{y}_j^{[m]}(k+1)\left(1 - \hat{y}_j^{[m]}(k+1)\right) \begin{pmatrix} \gamma_j^{[m]} x^{[m]}(k+1) \\ u_j^{[m]}(k+1) \end{pmatrix},$$

або, вводячи нові позначення, у більш компактній формі:

$$\tilde{w}_j^{[m]}(k+1) = \tilde{w}_j^{[m]}(k) + \eta_j^{[m]}(k+1)e_j^{[m]}(k+1)\hat{y}_j^{[m]} \times$$
$$\times(k+1)\left(1-\hat{y}_j^{[m]}(k+1)\right)\tilde{x}^{[m]}(k+1) =$$
$$= \tilde{w}_j^{[m]}(k) + \eta_j^{[m]}(k+1)e_j^{[m]}(k+1)\tilde{J}_j^{[m]}(k+1).$$

Поліпшити характер процесу настроювання можна, вводячи до алгоритму навчання регуляризуючий член [18–20], при цьому замість критерію навчання (1) використовується функція

$$E_j^{[m]}(k) = \frac{\eta}{2}\left(e_j^{[m]}(k)\right)^2 + \frac{1-\eta}{2}\left\|\tilde{w}_j^{[m]}(k) - \tilde{w}_j^{[m]}(k-1)\right\|^2, 0 < \eta \le 1, \quad (4)$$

а сам метод набуває вигляду:

$$\tilde{w}_j^{[m]}(k+1) = \tilde{w}_j^{[m]}(k) +$$
$$+\eta_j^{[m]}(k+1)\left(\eta e_j^{[m]}(k+1)\tilde{J}_j^{[m]}(k+1) + (1-\eta)\times\right.$$
$$\left.\times\left(\tilde{w}_j^{[m]}(k) - \tilde{w}_j^{[m]}(k-1)\right)\right), \quad (5)$$

що є модифікацією відомої процедури Сільви-Альмейди [19].

Використовуючи надалі підхід, запропонований у [21, 22], можна ввести до (5) згладжувальні і фільтруючі властивості. При цьому приходимо до кінцевої форми:

$$\tilde{w}_j^{[m]}(k+1) = \tilde{w}_j^{[m]}(k) +$$
$$+\frac{\eta e_j^{[m]}(k+1)\tilde{J}_j^{[m]}(k+1) + (1-\eta)\left(\tilde{w}_j^{[m]}(k) - \tilde{w}_j^{[m]}(k-1)\right)}{r_j^{[m]}(k+1)},$$
$$r_j^{[m]}(k+1) = r_j^{[m]}(k) + \left\|\tilde{J}_j^{[m]}(k+1)\right\|^2 - \left\|\tilde{J}_j^{[m]}(k-s)\right\|^2, \quad (6)$$

де s – розмір ковзкого вікна.

Цікаво, що при $s = 1$, $\eta = 1$ приходимо до нелінійного варіанту оптимального за швидкодією методу Качмажа-Уідроу-Хоффа [23–25]

$$\tilde{w}_j^{[m]}(k+1) = \tilde{w}_j^{[m]}(k) + \frac{e_j^{[m]}(k+1)\tilde{J}_j^{[m]}(k+1)}{\left\|\tilde{J}_j^{[m]}(k+1)\right\|^2},$$

що широко використовується у практиці навчання штучних нейронних мереж.

ОПТИМІЗАЦІЯ ВИХІДНОГО СИГНАЛУ ПУЛУ НЕЙРОНІВ

Вихідні сигнали всіх нейронів пулу кожного каскаду об'єднуються нейроном $GN^{[m]}$, вихід якого $\hat{y}^{*[m]}(k)$ за точністю повинен перевершувати будь-який з сигналів $\hat{y}_j^{[m]}(k)$. Ця задача може бути розглянута з позицій нелінійного програмування з використанням адаптивного узагальненого прогнозування [26–31].

Вводячи до розгляду вектор вихідних сигналів пулу m-го каскаду $\hat{y}^{[m]}(k) = \left(\hat{y}_1^{[m]}(k), \hat{y}_2^{[m]}(k), \ldots, \hat{y}_q^{[m]}(k)\right)^T$, формуватимемо оптимальний вихідний сигнал нейрону $GN^{[m]}$, що є по суті адаптивним лінійним асоціатором [1, 2], у вигляді

$$\hat{y}^{*[m]}(k) = \sum_{j=1}^{q} c_j^{[m]}\hat{y}_j^{[m]}(k) = c^{[m]T}\hat{y}^{[m]}(k)$$

за додаткових обмежень на незміщеність

$$\sum_{j=1}^{q} c_j^{[m]} = E^T c^{[m]} = 1, \quad (7)$$

де $c^{[m]} = \left(c_1^{[m]}, c_2^{[m]}, \ldots, c_q^{[m]}\right)^T$, $E = (1, 1, \ldots, 1)^T - (q \times 1)$ – вектори.

Вводячи надалі критерій навчання на ковзному вікні

$$E^{[m]}(k) = \frac{1}{2}\sum_{\tau=k-s+1}^{k}\left(y(\tau) - \hat{y}^{*[m]}(\tau)\right)^2 =$$
$$= \frac{1}{2}\sum_{\tau=k-s+1}^{k}\left(y(\tau) - c^{[m]T}\hat{y}^{[m]}(\tau)\right)^2,$$

з урахуванням обмеження (7), запишемо функцію Лагранжа вигляду

$$L^{[m]}(k) = E^{[m]}(k) + \lambda\left(1 - E^T c^{[m]}\right), \quad (8)$$

де λ – невизначений множник Лагранжа.

Пряма мінімізація (8) за $c^{[m]}$ веде до співвідношення

$$\hat{y}^{*[m]}(k+1) = \frac{\hat{y}^{[m]T}(k+1)P^{[m]}(k+1)E}{E^T P^{[m]}(k+1)E},$$
$$P^{[m]}(k+1) = \left(\sum_{\tau=k-s+2}^{k+1}\hat{y}^{[m]}(\tau)\hat{y}^{[m]T}(\tau)\right)^{-1} \quad (9)$$

або в рекурентній формі:

$$\tilde{P}^{[m]}(k+1) = P^{[m]}(k) -$$
$$-\frac{P^{[m]}(k)\hat{y}^{[m]}(k+1)\hat{y}^{[m]T}(k+1)P^{[m]}(k)}{1+\hat{y}^{[m]T}(k+1)P^{[m]}(k)\hat{y}^{[m]}(k+1)},$$
$$P^{[m]}(k+1) = \tilde{P}^{[m]}(k+1) + \quad (10)$$
$$+\frac{\tilde{P}^{[m]}(k+1)\hat{y}^{[m]}(k-s+1)\hat{y}^{[m]T}(k-s+1)\tilde{P}^{[m]}(k+1)}{1-\hat{y}^{[m]T}(k-s+1)\tilde{P}^{[m]}(k+1)\hat{y}^{[m]}(k-s+1)},$$
$$\hat{y}^{*[m]}(k+1) = \frac{\hat{y}^{[m]}(k+1)P^{[m]}(k+1)E}{E^T P^{[m]}(k+1)E}.$$

При $s = 1$ співвідношення (9), (10) набувають вкрай простого вигляду:

$$\hat{y}^{*[m]}(k+1) = \frac{\hat{y}^{[m]T}(k+1)\,\hat{y}^{[m]}(k+1)}{E^T \hat{y}^{[m]}(k+1)} =$$

$$= \frac{\left\|\hat{y}^{[m]}(k+1)\right\|^2}{E^T \hat{y}^{[m]}(k+1)} = \frac{\sum_{j=1}^{q}\left(\hat{y}_j^{[m]}(k+1)\right)^2}{\sum_{j=1}^{q}\hat{y}_j^{[m]}(k+1)}. \quad (11)$$

Тут важливо відзначити, що як навчання нейронів у каскадах, так і навчання узагальнюючих нейронів можна організувати в адаптивному режимі. При цьому ваги всіх попередніх каскадів не заморожуються, а постійно налаштовуються, число каскадів може як збільшуватися, так і зменшуватися, що вигідно відрізняє запропоновану нейронну мережу від відомих каскадних систем.

ВИСНОВОК

У статті запропоновано архітектуру та методи навчання гібридної оптимізованої каскадної нейронної мережі, що відрізняється від відомих каскадних систем обчислювального інтелекту можливістю обробки часових рядів в адаптивному режимі, що дає можливість обробляти нестаціонарні стохастичні та хаотичні сигнали нелінійних об'єктів з необхідною точністю. У порівнянні зі своїми прототипами запропонована система відрізняється обчислювальною простотою і відзначається як слідкуючими, так і фільтруючими властивостями.

СПИСОК ЛІТЕРАТУРИ

1. *Cichocki, A.* Neural Networks for Optimization and Signal Processing / A. Cichocki, R. Unbehauen. – Stuttgart : Teubner, 1993. – 526 p.
2. *Haykin, S.* Neural Networks. A Comprehensive Foundation / S. Haykin. – Upper Saddle River : Prentice Hall, 1999. – 842 p.
3. *Kasabov, N.* Evolving Connectionist Systems / Kasabov N. – London : Springer-Verlag, 2003. – 307 p.
4. *Lughofer, E.* Evolving Fuzzy Systems – Methodologies, Advanced Concepts and Applications / Lughofer E. – Berlin-Heidelberg: Springer-Verlag, 2011. – 454 p.
5. *Fahlman, S. E.* The cascade-correlation learning architecture. Advances in Neural Information Processing Systems / S. E. Fahlman, C. Lebiere ; Ed. by D. S. Touretzky. – San Mateo, CA : Morgan Kaufman, 1990. – P. 524–532.
6. *Prechelt, L.* Investigation of the Cascor family of learning algorithms / Prechelt L. // Neural Networks. – 1997. – vol. 10. – P. 885–896.
7. *Schalkoff, R. J.* Artificial Neural Networks / Schalkoff R. J. – N. Y. : The McGraw-Hill Comp., 1997. – 528 p.
8. *Avedjan, E. D.* Cascade neural networks / E. D. Avedjan, G. V. Barkan, I. K. Levin // Avtomatika i telemekhanika. – 1999. – No. 3. – P. 38–55.
9. *Bodyanskiy, Ye.* The cascaded orthogonal neural network / Ye. Bodyanskiy, A. Dolotov, I. Pliss, Ye. Viktorov ; Eds. by K. Markov, K. Ivanova, I. Mitov // Information Science & Computing. – Sofia, Bulgaria : FOI ITHEA. – 2008. – Vol. 2. – P. 13–20.
10. *Bodyanskiy, Ye.* The cascaded neo-fuzzy architecture and its on-line learning algorithm / Ye. Bodyanskiy, Ye. Viktorov ; Eds. by K. Markov, P. Stanchev, K. Ivanova, I. Mitov // Intelligent Processing. – 9. – Sofia : FOI ITHEA, 2009. – P. 110–116.
11. *Bodyanskiy, Ye.* The cascaded neo-fuzzy architecture using cubic-spline activation functions / Ye. Bodyanskiy, Ye. Viktorov // Int. J. «Information Theories & Applications». – 2009. – vol. 16, No. 3. – P. 245–259.
12. *Bodyanskiy, Ye.* The cascade growing neural network using quadratic neurons and its learning algorithms for on-line information processing / Ye. Bodyanskiy, Ye.Viktorov, I. Pliss ; Eds. by G. Setlak, K. Markov // Intelligent Information and Engineering Systems. – 13. – Rzeszov-Sofia : FOI ITHEA, 2009. – P. 27–34.
13. *Kolodyazhniy, V.* Cascaded multi-resolution spline-based fuzzy neural / V. Kolodyazhniy, Ye. Bodyanskiy ; Eds. by P. Angelov, D. Filev, N. Kasabov // Proc. Int. Symp. on Evolving Intelligent Systems. – Leicester, UK : De Montfort University, 2010. – P. 26–29.
14. *Bodyanskiy, Ye.* Cascaded GMDH-wavelet-neuro-fuzzy network / Ye. Bodyanskiy, O. Vynokurova, N. Teslenko // Proc 4th Int. Workshop on Inductive Modelling «IWIM 2011». – Kyiv, 2011. – P. 22–30.
15. *Bodyanskiy, Ye.* Hybrid cascaded neural network based on wavelet-neuron / Ye. Bodyanskiy, O. Kharchenko, O. Vynokurova // Int. J. Information Theories & Applications. – 2011. – vol. 18, No. 4. – P. 335–343.
16. *Bodyanskiy, Ye.* Evolving cascaded neural network based on multidimensional Epanechnikov's kernels and its learning algorithm / Ye. Bodyanskiy, P. Grimm, N. Teslenko // Int. J. Information Technologies & Knowledge. – 2011. – vol. 5, No. 1. – P. 25–30.
17. *Kruschke, J. K.* Benefits of gain: speed learning and minimum layers backpropagation networks / J. K. Kruschke, J. R. Movellan // IEEE Trans. on Syst., Man. And Cybern. – 1991. – vol. 21. – P. 273–280.
18. *Chan, L. W.* An adaptive learning algorithm for backpropagation networks / L. W. Chan, F. Fallside // Computer Speech and Language. – 1987. – vol. 2. – P. 205–218.
19. *Silva, F. M.* Speeding up backpropagation / F. M. Silva, L. B. Almeida ; Ed. by R. Eckmiller // Advances of Neural Computers. – North-Holland : Elsevier Science Publishers. – B. V., 1990. – P. 151–158.
20. *Veitch, A. C.* A modified quickprop algorithm / A. C. Veitch, G. Holmes // Neural Computation. – 1991. – vol. 3. – P. 310–311.
21. *Bodyanskiy, Ye.* An adaptive learning algorithm for a neuro-fuzzy network / Ye. Bodyanskiy, V. Kolodyazhniy, A. Stephan ; Ed. by B. Reusch // Computational Intelligence: Theory and Applications. – Berlin-Heidelberg-New-York: Springer, 2001. – P. 68–75.
22. *Otto, P.* A new learning algorithm for a forecasting neuro-fuzzy network / P. Otto, Ye. Bodyanskiy, V. Kolodyazhniy // Integrated Computer-Aided Engineering. – 2003. – vol. 10, No. 4. – P. 399–409.
23. *Kaczmarz, S.* Angenaeherte Ausloesung von Systemen linearer Gleichungen / Kaczmarz S. // Bull. Int. Acad. Polon. Sci. – 1937. – Let. A. – P. 355–357.
24. *Kaczmarz, S.* Approximate solution of systems of linear equations / Kaczmarz S. // Int. J. Control. – 1993. – vol. 53. – P. 1269–1271.

25. *Widrow, B.* / Adaptive switching circuits / Widrow B., Hoff Jr. M. E. // 1960 URE WESCON Convention Record. – N. Y. : IRE, 1960. – Part 4. – P. 96–104.
26. *Бодянский, Е.* Адаптивно прогнозиране на нестационарни процеси / Е. Бодянский, И. Плисс, Н. Маджаров // Автоматика и изчислительна техника. – 1983. – № 6. – С. 5–12.
27. *Бодянский, Е. В.* Адаптивное обобщенное прогнозирование многомерных случайных последовательностей / Е. В. Бодянский, И. П. Плисс, Т. В. Соловьева // Докл. АН УССР. – 1989. – Сер.А. – №m9. – С. 73–75.
28. *Bodyanskiy, Ye.* Adaptive generalized forecasting of multivariate stochastic signals / Ye. Bodyanskiy, I. Pliss // Proc. Latvian Sign. Proc. Int. Conf. – Riga, 1990. – V. 2. – P. 80–83.
29. *Bodyanskiy, Ye.* Algorithm for adaptive identification of dynamical parametrically nonstationary objects / Ye. Bodyanskiy, S. Vorobyov, A. Stephan // J. Computer and Systems Sci. Int. – 1999. – vol. 38, No. 1. – P. 14–38.
30. *Bodyanskiy, Ye.* Recurrent neural network detecting changes in the properties of nonlinear stochastic sequences / Ye. Bodyanskiy, S. Vorobyov // Automation and Remote Control. – 2000. – vol. 61, No. 7. – Part 1. – P. 1113–1124.
31. *Vorobyov, S.* An adaptive noise cancellation for multisensory signals / S. Vorobyov, A. Cichocki, Ye. Bodyanskiy // Fluctuation and Noise Letters. – 2001. – vol. 1, No. 1. – P. 13–24.

Стаття надійшла до редакції 21.02.2014.

Тищенко А. К.[1], Плисс И. П.[2], Копалиани Д. С.[3]

[1]Канд. техн. наук, старший научный сотрудник Проблемной НИЛ АСУ Харьковского национального университета радиоэлектроники, Украина

[2]Канд. техн. наук, старший научный сотрудник, ведущий научный сотрудник Проблемной НИЛ АСУ Харьковского национального университета радиоэлектроники, Украина

[3]Аспирантка, Харьковский национальный университет радиоэлектроники, Украина

ГИБРИДНАЯ КАСКАДНАЯ ОПТИМИЗИРОВАННАЯ НЕЙРОННАЯ СЕТЬ

Предложена новая архитектура и алгоритмы ее обучения для гибридной каскадной нейронной сети с оптимизацией пула нейронов в каждом каскаде. Предложенная гибридная каскадная нейронная сеть обеспечивает вычислительную простоту и характеризуется следящими и фильтрующими свойствами.

Ключевые слова: нейронная сеть, оптимальное обучение, вычислительный интеллект, эволюционирующая гибридная система.

Tyshchenko O. K.[1], Pliss I. P.[2], Kopaliani D. S.[3]

[1]Ph.D, Senior Researcher at Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Ukraine

[2]Ph.D, Leading Researcher, Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Ukraine

[3]Post-graduate student, Kharkiv National University of Radio Electronics, Ukraine

A HYBRID CASCADE OPTIMIZED NEURAL NETWORK

A new architecture and learning algorithms for a hybrid cascade optimized neural network is proposed. The proposed hybrid system is different from existing cascade systems in its capability to operate in an online mode, which allows it to work with both non-stationary and stochastic nonlinear chaotic signals with the required accuracy. The proposed hybrid cascade neural network provides computational simplicity and possesses both tracking and filtering capabilities.

Keywords: neural network, optimal learning, computational intelligence, evolving hybrid system.

REFERENCES

1. Cichocki A., Unbehauen R. Neural Networks for Optimization and Signal Processing. Stuttgart, Teubner, 1993, 526 p.
2. Haykin S. Neural Networks. A Comprehensive Foundation, Upper Saddle River, Prentice Hall, 1999, 842 p.
3. Kasabov N. Evolving Connectionist Systems. London, Springer-Verlag, 2003, 307 p.
4. Lughofer E. Evolving Fuzzy Systems – Methodologies, Advanced Concepts and Applications. Berlin-Heidelberg, Springer-Verlag, 2011, 454 p.
5. Fahlman S. E., Lebiere C. The cascade-correlation learning architecture. Advances in Neural Information Processing Systems, Ed. by D. S. Touretzky, San Mateo, CA, Morgan Kaufman, 1990, pp. 524–532.
6. Prechelt L. Investigation of the Cascor family of learning algorithms, *Neural Networks*, 1997, vol. 10, pp. 885–896.
7. Schalkoff R. J. Artificial Neural Networks. N.Y, The McGraw-Hill Comp., 1997, 528 p.
8 Avedjan E. D., Barkan G. V., Levin I. K. Cascade neural networks, *Avtomatika i telemekhanika,* 1999, No. 3, pp. 38–55.
9. Bodyanskiy Ye., Dolotov A., Pliss I., Viktorov Ye. The cascaded orthogonal neural network, Eds. by K. Markov, K. Ivanova, I. Mitov, *Information Science & Computing*, Sofia, Bulgaria, FOI ITHEA, 2008, Vol. 2, pp. 13–20.
10. Bodyanskiy Ye., Viktorov Ye. The cascaded neo-fuzzy architecture and its on-line learning algorithm, Eds. by K. Markov, P. Stanchev, K. Ivanova, I. Mitov, Intelligent Processing, 9, Sofia, FOI ITHEA, 2009, pp. 110–116.
11. Bodyanskiy Ye., Viktorov Ye. The cascaded neo-fuzzy architecture using cubic-spline activation functions, *Int. J. «Information Theories & Applications»,* 2009, vol. 16, No. 3, pp. 245–259.
12. Bodyanskiy Ye., Viktorov Ye., Pliss I. The cascade growing neural network using quadratic neurons and its learning algorithms for on-line information processing, Eds. by G. Setlak, K. Markov, *Intelligent Information and Engineering Systems*, 13, Rzeszov-Sofia, FOI ITHEA, 2009, pp. 27–34.

13. Kolodyazhniy V., Bodyanskiy Ye. Cascaded multi-resolution spline-based fuzzy neural network, Eds. by P. Angelov, D. Filev, N. Kasabov, *Proc. Int. Symp. on Evolving Intelligent Systems*, Leicester, UK, De Montfort University, 2010, pp. 26–29.
14. Bodyanskiy Ye., Vynokurova O., Teslenko N. Cascaded GMDH-wavelet-neuro-fuzzy network, *Proc 4th Int. Workshop on Inductive Modelling «IWIM 2011»*. Kyiv, 2011, pp. 22–30.
15. Bodyanskiy Ye., Kharchenko O., Vynokurova O. Hybrid cascaded neural network based on wavelet-neuron, *Int. J. Information Theories & Applications.* 2011, 18, No. 4, pp. 335–343.
16. Bodyanskiy Ye., Grimm P., Teslenko N. Evolving cascaded neural network based on multidimensional Epanechnikov's kernels and its learning algorithm, *Int. J. Information Technologies & Knowledge,* 2011, vol. 5, No. 1, pp. 25–30.
17. Kruschke J. K., Movellan J. R. Benefits of gain: speed learning and minimum layers backpropagation networks, *IEEE Trans. on Syst., Man. And Cybern,* 1991, 21, pp. 273–280.
18. Chan L. W., Fallside F. An adaptive learning algorithm for backpropagation networks, *Computer Speech and Language*, 1987, 2, pp. 205–218.
19. Silva F. M., Almeida L. B. Speeding up backpropagation, Ed. by R. Eckmiller, *Advances of Neural Computers.* North-Holland, Elsevier Science Publishers, B.V., 1990, pp. 151–158.
20. Veitch A. C., Holmes G. A modified quickprop algorithm, *Neural Computation*, 1991, 3, pp. 310–311.
21. Bodyanskiy Ye., Kolodyazhniy V., Stephan A. An adaptive learning algorithm for a neuro-fuzzy network, Ed. by B. Reusch, *Computational Intelligence: Theory and Applications,* Berlin-Heidelberg-New-York, Springer, 2001, pp. 68–75.
22. Otto P., Bodyanskiy Ye., Kolodyazhniy V. A new learning algorithm for a forecasting neuro-fuzzy network, *Integrated Computer-Aided Engineering,* 2003, vol. 10, No. 4, pp. 399–409.
23. Kaczmarz S. Angenaeherte Ausloesung von Systemen linearer Gleichungen, *Bull. Int. Acad. Polon. Sci*, 1937, Let. A, pp. 355–357.
24. Kaczmarz S. Approximate solution of systems of linear equations, *Int. J. Control,* 1993, vol. 53, pp. 1269–1271.
25. Widrow B., Hoff Jr. M. E. Adaptive switching circuits, *1960 URE WESCON Convention Record,* N.Y, IRE, 1960, Part 4, pp. 96–104.
26. Bodyanskiy Ye., Pliss I., Madjarov N. Adaptive forecasting of nonstationary processes, *Avtomatika I Izchislitelna Tekhnika,* 1983, No. 6, pp. 5–12.
27. Bodyanskiy Ye., Pliss I. P., Solovyova T. V. Adaptive generalized forecasting of multidimensional stochastic sequences, *Doklady AN USSR,* 1989, A, No. 9, pp. 73–75.
28. Bodyanskiy Ye., Pliss I. Adaptive generalized forecasting of multivariate stochastic signals, *Proc. Latvian Sign. Proc. Int. Conf.* Riga, 1990, V. 2, pp. 80–83.
29. Bodyanskiy Ye., Vorobyov S., Stephan A. Algorithm for adaptive identification of dynamical parametrically nonstationary objects, *J. Computer and Systems Sci. Int*, 1999, vol. 38, No. 1, pp. 14–38.
30. Bodyanskiy Ye., Vorobyov S. Recurrent neural network detecting changes in the properties of nonlinear stochastic sequences, *Automation and Remote Control,* 2000, vol. 61, No. 7, Part 1, pp. 1113–1124.
31. Vorobyov S., Cichocki A., Bodyanskiy Ye. An adaptive noise cancellation for multisensory signals, *Fluctuation and Noise Letters,* 2001, 1, No. 1, pp. 13–24.