

ІНТЕЛЕКТУАЛЬНА ОБРОБКА ДАНИХ НА ОСНОВІ ГІБРИДНОЇ ВЕЙВЛЕТ-НЕЙРО-ФАЗЗИ-СИСТЕМИ НА АДАПТИВНИХ W-НЕЙРОНАХ

У статті запропоновано архітектуру гібридної вейвлет-нейро-фаззи-системи на W-нейронах, що є модифікацією відомої ANFIS-структури та має переваги кожного з підходів. Запропонована архітектура має покращені апроксимуючі властивості за рахунок введення адаптивного вейвлону в консеквент. Також синтезовано алгоритм навчання усіх її параметрів. Проведено експериментальне моделювання на різноманітних нестационарних сигналах, результати яких підтверджують актуальність підходу, що розвивається.

Ключові слова: архітектура гібридної вейвлет-нейро-фаззи-системи, W-нейрони, ANFIS-структура, адаптивний вейвлон, консеквент.

В статье предложена архитектура гибридной вейвлет-нейро-фаззи-системы на W-нейронах, которая является модификацией известной ANFIS-структуры и обладает преимуществами каждого из подходов. Предложенная архитектура обладает улучшенными аппроксимирующими способностями за счет введения адаптивного вейвлона в консеквент. Также синтезирован алгоритм обучения всех ее параметров. Проведено экспериментальное моделирование на различных нестационарных процессах, результаты, которых подтверждают актуальность развиваемого подхода.

Ключевые слова: архитектура гибридной вейвлет-нейро-фаззи-системи, W-нейроны, ANFIS-структура, адаптивный вейвлон, консеквент.

The hybrid wavelet neuro-fuzzy system based on W-neurons architecture which is the modified well-known ANFIS structure and has the advantages of each approaches is proposed. The proposed architecture has improved approximating ability by introducing of the adaptive W-neuron to the consequent. Also the all-parameters learning algorithm is proposed. The simulation experiments based on the different non-stationary process are performed. The results of them are satisfied the relevance of the developed approach.

Key words: The hybrid wavelet neuro-fuzzy system, W-neurons, ANFIS structure, adaptive W-neuron, the consequent.

ВСТУП

У цей час все більшого поширення одержують інтелектуальні методи обробки нестационарних сигналів за умов невизначеності, що засновані на технологіях обчислювального інтелекту. Такі методи дозволяють вирішувати великий клас задач обробки інформації й, насамперед, задачі ідентифікації, емуляції, інтелектуального керування, прогнозування часових рядів довільної природи за умов структурної та параметричної невизначеності.

Найбільш відомими та популярними є багатошарові мережі із прямою передачею інформації типу тришарового перцептрон, елементарними вузлами яких є так звані P-нейрони з монотонними функціями активації. Ефективність багатошарових мереж пояснюється їх універсальними апроксимуючими властивостями в сполученні з відносно компактним поданням нелінійної системи, що моделюється. Це означає, що вони з успіхом можуть бути

використані у задачах моделювання (емуляції) нелінійних систем, що описуються багатовимірними різницевиими рівняннями виду

$$y(k) = F(x(k)) + \xi(k), \quad (1)$$

де $y(k)$ – вихідний сигнал системи в k -й момент дискретного часу $k = 0, 1, 2, \dots$, $x(k) \in X$ – $(n \times 1)$ – вектор сигналів на вході, що включає в себе як екзогенні змінні, так і минулі значення вихідного сигналу, $F(\bullet)$ – довільна функція в загальному випадку невідомого виду, $\xi(k)$ – збурювання з невідомими характеристиками, що не спостерігається. Передбачається також, що функція $F(\bullet)$ задана або на одиничному гіперкубі, або на ортотопі

$$x_i(k) \in [x_i^{\min}, x_i^{\max}], \quad i = 1, 2, \dots, n, \quad (2)$$

де x_i^{\min} , x_i^{\max} відомі нижня та верхня границі варіювання i -го вхідного впливу.

Основним недоліком багатошарових мереж є досить низька швидкість їхнього навчання, заснованого на зворотному поширенні похибок, що робить неможливим їхнє використання у задачах реального часу.

Таким чином, актуальним є синтез гібридних архітектур та їх алгоритмів навчання на основі методів обчислювального інтелекту, які надалі будемо йменувати вейвлет-нейро-фаззі-системами. Такі системи поєднують у собі лінгвістичну інтерпретаваність і апроксимуючі властивості систем нечіткого виведення [1; 2] з можливостями штучних нейронних мереж навчатися та апроксимувати [3; 4] і обчислювальними перевагами вейвлет-функцій [5-8]. Це означає, що вони з успіхом можуть бути використані у задачах прогнозування стохастичних та хаотичних сигналів і послідовностей зі складним нелінійним трендом і нестационарними параметрами.

У статті пропонується модифікована архітектура гібридної вейвлет-нейро-фаззі-системи, що заснована на адаптивних W -нейронах та алгоритм навчання всіх її параметрів.

1. АРХІТЕКТУРА ГІБРИДНОЇ ВЕЙВЛЕТ-НЕЙРО-ФАЗЗІ-СИСТЕМИ НА АДАПТИВНИХ W -НЕЙРОНАХ

Введемо до розгляду п'ятишарову архітектуру, що наведено на рис. 1, і близьку, по суті, до відомої ANFIS [9; 10] та структури, що представлено в роботах [11; 12], та є у загальному випадку системою нечіткого виведення Такагі-Сугено-Канга, що навчається.

Перший прихований шар, на відміну від нейро-фаззі-систем, утворено не традиційними невід'ємними функціями належності, а набором з $h_\psi n$ вейвлетів (по h_ψ одновимірних вейвлетів на кожний вхід) $\psi_{ji}(k) = \psi_{ji}(x_i(k), c_{ji}^\psi(k), \sigma_{ji}(k))$ з $2h_\psi n$ параметрами центру c_{ji}^ψ та ширини σ_{ji} , що настроюються. Тут можна відзначити, що коливальний характер вейвлет-функцій не суперечить уніполярності функцій належності, оскільки від'ємні значення $\psi_{ji}(k)$ можуть трактуватися в сенсі малих рівнів належності або неналежності [13; 14].

У якості функцій активації-належності в адаптивній вейвлет-нейро-фаззі-мережі можуть використовуватися різні аналітичні вейвлети: вейвлет Морле, вейвлет «Mexican Hat», POLYWOG-вейвлети, RASP-вейвлети [15], генератор аналітичних вейвлетів [16], трикутний вейвлет [17] тощо.

Грунтуючись на вейвлет-функції «Mexican Hat», будемо використовувати запропоновану нами в [18] нову активаційну функцію, що настроюється та має вигляд

$$\varphi_j(x(k)) = (1 - \alpha_j \tau_j^2) \exp\left(-\frac{\tau_j^2}{2}\right), \quad (3)$$

де $\tau_j(x(k)) = \left((x(k) - c_j^\varphi(k))^T Q_j^{-1}(k)(x(k) - c_j^\varphi(k))\right)$, α_j – параметр, що настроюється ($0 \leq \alpha \leq 1$).

Параметр α_j , що уточнюється, дозволяє настроювати форму активаційної функції в процесі навчання складеного адаптивного W -нейрона (вейвлону), при цьому при $\alpha = 0$ одержуємо Гаусову функцію активації, при $\alpha = 1$ одержуємо вейвлет-функцію «Mexican Hat», а при $0 < \alpha < 1$ – гібридну функцію активації.

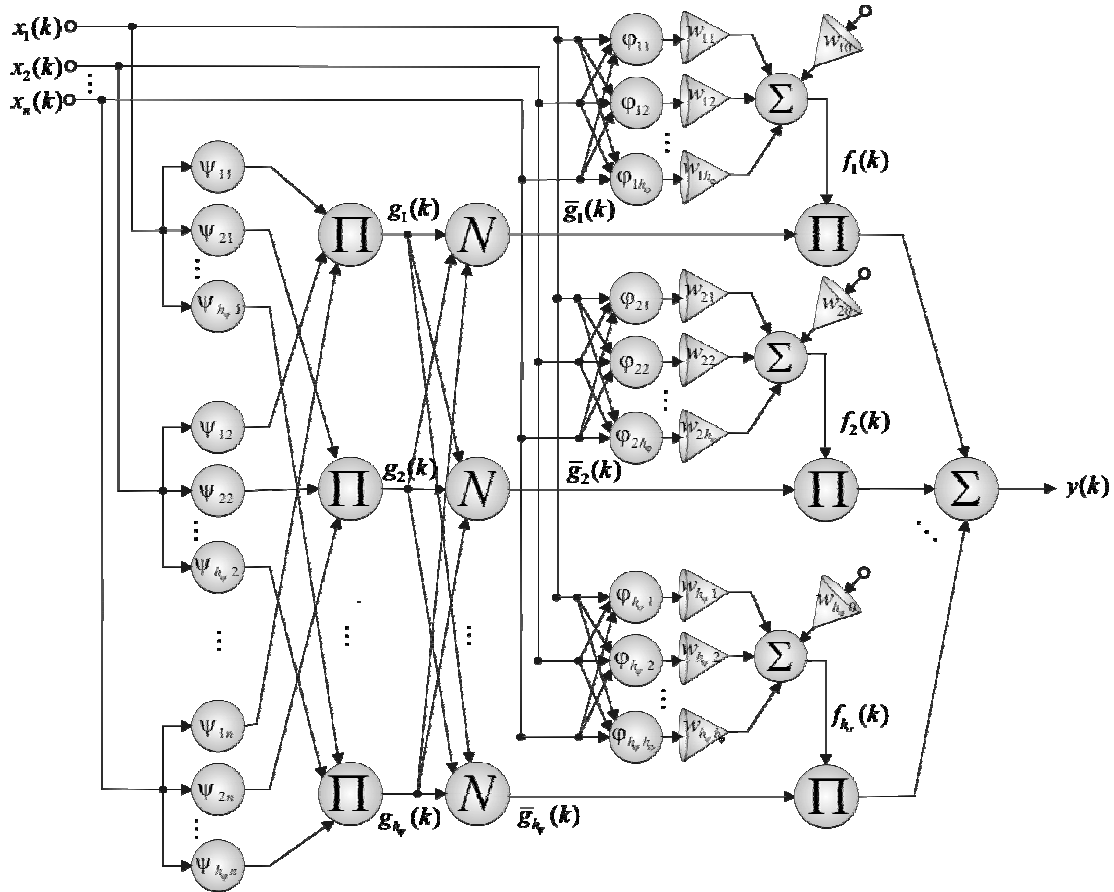


Рис. 1. Гібридна вейвлет-нейро-фаззі-система, що заснована на адаптивних W -нейронах

Другий прихований шар реалізує операцію, аналогічну обчисленню нечіткої T -норми

$$g_j(k) = \prod_{i=1}^n \psi_{ji}(k), j = 1, 2, \dots, h_\psi, \quad (4)$$

після чого в третьому прихованому шарі проводиться нормалізація

$$\bar{g}_j(k) = \frac{g_j(k)}{\sum_{j=1}^{h_\psi} g_j(k)} = \frac{\prod_{i=1}^n \psi_{ji}(k)}{\sum_{j=1}^{h_\psi} \prod_{i=1}^n \psi_{ji}(k)}, \quad (5)$$

що забезпечує виконання умови Руспіні

$$\sum_{j=1}^{h_\psi} \bar{g}_j(k) = 1. \quad (6)$$

Четвертий прихований шар реалізує операцію, яка аналогічна обчисленню консеквента в системах нечіткого виведення, при цьому замість лінійних вихідних функцій $f_j(X(k))$ використовується структура адаптивного W -нейрона [19; 20], що дозволяє покращити апроксимуючі властивості системи

$$f_j(X(k)) = w_{j0}(k) + \sum_{m=1}^{h_\varphi} w_{jm}(k) \varphi_{jm}(\tau_{jm}(k)) = w_j^T(k) \varphi_j(\tau_j(k)), \quad (7)$$

де, $\varphi_{j0} = 1$, $w_j(k) = (w_{j0}(k), w_{j1}(k), w_{j2}(k), \dots, w_{jh_\varphi}(k))^T$, $\varphi_j(\tau_j(k)) = (1, \varphi_{j1}(\tau_{j1}(k)), \varphi_{j2}(\tau_{j2}(k)),$

$$\dots, \varphi_{jh_\varphi}(\tau_{jh_\varphi}(k)))$$
, $\varphi_{jm}(\tau_{jm}(k)) = \varphi_{jm}((X(k) - c_{jm}^\varphi(k))^T Q_{jm}^{-1}(k)(X(k) - c_{jm}^\varphi(k)), \alpha_{jm}(k))$,

$$\tau_{jm}(k) = ((X(k) - c_{jm}^\varphi(k))^T Q_{jm}^{-1}(k)(X(k) - c_{jm}^\varphi(k)), \alpha_{jm}(k)).$$

У цьому випадку в четвертому шарі обчислюються сигнали такого вигляду

$$\bar{g}_j(k) \left(w_{j0}(k) + \sum_{m=1}^{h_\varphi} w_{jm}(k) \varphi_{jm}(\tau_{jm}(k)) \right) = \bar{g}_j w_j^T(k) \varphi_j(\tau_j(k)), \quad (8)$$

де $h_\varphi(n+1)$ параметрів w_{jm} , $j = 1, 2, \dots, h_\varphi$, $m = 0, 1, 2, \dots, h_\varphi$ підлягають визначенню.

І, нарешті, у п'ятому вихідному шарі обчислюється вихідний сигнал системи

$$y(k) = \sum_{j=1}^{h_\psi} \bar{g}_j(k) f_j(X(k)), \quad (9)$$

або у векторній формі

$$y(k) = \bar{g}^T(k) f(X(k)), \quad (10)$$

де $f(X(k)) = (f_1(X(k)), f_2(X(k)), \dots, f_{h_\psi}(X(k)))^T = (w_1^T(k) \varphi_1(\tau_1(k)), \dots, w_{h_\psi}^T(k) \varphi_{h_\psi}(\tau_{h_\psi}(k)))$.

Вводячи заміну змінних $w = (w_1, w_1, \dots, w_{h_\psi})^T$ і $F(X(k)) = (\bar{g}_1(k) \varphi_1(\tau_1(k)), \bar{g}_2(k) \varphi_2(\tau_2(k)), \dots, \bar{g}_{h_\psi}(k) \varphi_{h_\psi}(\tau_{h_\psi}(k)))^T$, вихід такої архітектури можна записати в компактній формі

$$y(k) = w^T(k) F(X(k)). \quad (11)$$

Таким чином, можна помітити, що параметри цієї системи, які настроюються, перебувають у першому і четвертому прихованих шарах. Це $2h_\psi n$ параметрів одновимірних вейвлетів c_{ji}^ψ і σ_{ji} ; $h_\psi(h_\varphi + 1)$ синаптичних ваг w_{jm} ; $h_\psi h_\varphi$ параметрів центрів c_{jm}^φ ; матриць розтягання Q_{jm}^{-1} та $h_\psi h_\varphi$ параметрів форми α_{jm} багатовимірних вейвлет-функцій адаптивного W -нейрона.

2. АЛГОРИТМ НАВЧАННЯ ГІБРИДНОЇ ВЕЙВЛЕТ-НЕЙРО-ФАЗЗИ-СИСТЕМИ

Для настроювання параметрів першого прихованого шару в ANFIS використовується алгоритм зворотного поширення похибки, що заснований на ланцюговому правилі диференціювання та градієнтної оптимізації локального критерію

$$E(k) = \frac{1}{2} e^2(k) = \frac{1}{2} (d(k) - y(k))^2 = \frac{1}{2} (d(k) - w^T(k) F(X(k)))^2. \quad (12)$$

У загальному випадку процедура навчання параметрів цього шару має вигляд

$$\begin{cases} c_{ji}^\psi(k+1) = c_{ji}^\psi(k) - \eta_{c^\psi}(k) \frac{\partial E(k)}{\partial c_{ji}^\psi(k)}, \\ \sigma_{ji}^{-1}(k+1) = \sigma_{ji}^{-1}(k) - \eta_\sigma(k) \frac{\partial E(k)}{\partial \sigma_{ji}^{-1}(k)}, \end{cases} \quad (13)$$

а її властивості повністю визначено параметрами кроку $\eta_{c^\psi}(k)$, $\eta_\sigma(k)$, що обираються, як правило, з емпіричних міркувань. Слід зазначити, що, якщо параметри четвертого шару можуть бути настроєні максимально швидко, то саме в першому шарі ця швидкодія губиться.

Підвищення швидкості збіжності може бути досягнуто шляхом використання більш складних, ніж градієнтні, процедур типу алгоритму Хартлі [21] або Марквардта [22], які можуть бути записані в узагальненій формі [23]

$$\Phi(k+1) = \Phi(k) + \lambda(J(k)J^T(k) + \eta I)^{-1}J(k)e(k), \quad (14)$$

де $\Phi(k) = (c_{11}^{\psi}(k), \sigma_{11}^{-1}(k), c_{21}^{\psi}(k), \sigma_{21}^{-1}(k), \dots, c_{ji}^{\psi}(k), \sigma_{ji}^{-1}(k), \dots, c_{hn}^{\psi}(k), \sigma_{hn}^{-1}(k))^T - (2h_{\psi}n \times 1)$ – вектор параметрів, що настроюється, причому для спрощення чисельної реалізації в нього входить не самий параметр σ_{ji} ширини, а зворотна йому величина σ_{ji}^{-1} , $J(k) - (2hn \times 1)$ – вектор-градієнт вихідного сигналу $y(k)$ за параметрами, що настроюються, $I - (2hn \times 2hn)$, – одинична матриця, η – скалярний регуляризуючий параметр, λ – додатний скалярний коефіцієнт підсилення (демпфірування).

Для обчислення компонент вектора-градієнта

$$J^{\psi}(k) = \left(\frac{\partial y(k)}{\partial c_{11}^{\psi}}, \frac{\partial y(k)}{\partial \sigma_{11}^{-1}}, \frac{\partial y(k)}{\partial c_{21}^{\psi}}, \frac{\partial y(k)}{\partial \sigma_{21}^{-1}}, \dots, \frac{\partial y(k)}{\partial c_{ji}^{\psi}}, \frac{\partial y(k)}{\partial \sigma_{ji}^{-1}}, \dots, \frac{\partial y(k)}{\partial c_{hn}^{\psi}}, \frac{\partial y(k)}{\partial \sigma_{hn}^{-1}} \right)^T \quad (15)$$

використовуються вирази

$$\begin{cases} \frac{\partial y(k)}{\partial c_{ji}^{\psi}} = f_j(X(k))\bar{g}_j(k)(1 - \bar{g}_j(k)) \frac{1}{\psi_{ji}(x_i(k), c_{ji}^{\psi}, \sigma_{ji}^{-1})} \cdot \frac{\partial \psi_{ji}}{\partial c_{ji}^{\psi}}, \\ \frac{\partial y(k)}{\partial \sigma_{ji}^{-1}} = f_j(X(k))\bar{g}_j(k)(1 - \bar{g}_j(k)) \frac{1}{\psi_{ji}(x_i(k), c_{ji}^{\psi}, \sigma_{ji}^{-1})} \cdot \frac{\partial \psi_{ji}}{\partial \sigma_{ji}^{-1}}. \end{cases} \quad (16)$$

З метою спрощення обчислювальної складності алгоритму навчання можна скористатися лемою обернення матриць у вигляді

$$(JJ^T + \eta I)^{-1} = \eta^{-1}I - \frac{\eta^{-1}IJJ^T\eta^{-1}I}{1 + J^T\eta^{-1}IJ}, \quad (17)$$

за допомогою якої можна одержати співвідношення

$$\lambda(JJ^T + \eta I)^{-1}J = \lambda \frac{J}{\eta + \|J\|^2}, \quad (18)$$

підставляючи яке в (14), отримуємо алгоритм навчання параметрів першого прихованого шару у вигляді

$$\Phi(k+1) = \Phi(k) + \lambda \frac{J^{\psi}(k)e(k)}{\eta + \|J^{\psi}(k)\|^2}. \quad (19)$$

Нескладно бачити, що алгоритм (19) є нелінійною адитивно-мультиплікативною модифікацією алгоритму Качмажа, а при $\lambda = 1, \eta = 0$ структурно з ним збігається.

З метою додання алгоритму (19) фільтруючих властивостей введемо додаткову процедуру настроювання регуляризуючого параметра η вигляду [24]

$$\begin{cases} \Phi(k+1) = \Phi(k) + \lambda \left(\frac{J^{\psi}(k)e(k)}{\eta(k)} \right), \\ \eta(k+1) = \alpha\eta(k) + \|J^{\psi}(k+1)\|^2, \end{cases} \quad (20)$$

що співпадає при $\alpha = 0$ с (19) і має при цьому максимальну швидкість збіжності, а при $\alpha = 1$ набуває властивостей стохастичної апроксимації. Тут слід зазначити, що алгоритм (20) є стійким при будь-яких значеннях параметра α , що вигідно відрізняє його від експоненційно-зваженого рекурентного методу найменших квадратів.

Далі розглянемо алгоритм навчання параметрів четвертого шару, а саме алгоритм навчання адаптивного W -нейрона.

Оскільки $(h_\varphi + 1) \times 1$ вектор синаптичних ваг w входить в опис мережі лінійно, для його уточнення може використовуватися будь-який з алгоритмів адаптивної ідентифікації [25] і, насамперед, традиційний рекурентний метод найменших квадратів:

$$\begin{cases} w(k+1) = w(k) + \frac{P(k)(y(k) - w^T(k)\varphi(x(k)))\varphi(x(k))}{1 + \varphi^T(x(k))P(k)\varphi(x(k))}, \\ P(k+1) = P(k) - \frac{P(k)\varphi(x(k+1))\varphi^T(x(k+1))P(k)}{1 + \varphi^T(x(k+1))P(k)\varphi(x(k+1))}, \end{cases} \quad (21)$$

що має фільтруючі властивості.

Для настроювання параметрів вейвлонів (векторів c_j , матриць Q_j^{-1} , параметра α_j) будемо використовувати градієнтну мінімізацію локального критерію (12), при цьому, на відміну від покомпонентного навчання, розглянутого в [26], будемо робити уточнення у векторно-матричній формі, що, по-перше, простіше з обчислювальної точки зору, а, по-друге, дозволить оптимізувати процес навчання за швидкодією.

У загальному випадку алгоритм навчання може бути записаний у вигляді

$$\begin{cases} c_j^\varphi(k+1) = c_j^\varphi(k) - \eta_{c_j^\varphi} \nabla_{c_j^\varphi} E(k), \quad j = 1, 2, \dots, h, \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) - \eta_{Q_j^{-1}} \left\{ \partial E(k) / \partial Q_j^{-1} \right\}, \quad j = 1, 2, \dots, h, \\ \alpha_j(k+1) = \alpha_j(k) - \eta_\alpha \nabla_{\alpha_j} E(k), \quad j = 1, 2, \dots, h, \end{cases} \quad (22)$$

де $\nabla_{c_j^\varphi} E$, $\nabla_{\alpha_j} E$ – $(n \times 1)$ -вектори-градієнти критерію (12) по c_j^φ та α_j відповідно; $\left\{ \partial E(k) / \partial Q_j^{-1} \right\}$ – $(n \times n)$ – матриця, що утворена частинними похідними $E(k)$ по компонентах Q_j^{-1} ; $\eta_{c_j^\varphi}$, $\eta_{Q_j^{-1}}$ і η_α – параметри кроку алгоритму навчання.

Для функції активації (3) можна записати

$$\begin{cases} \nabla_{c_j^\varphi} E(k) = 2e(k)w_j(k)Q_j^{-1}(k)(x(k) - c_j^\varphi(k)) \cdot \\ \cdot \left(\alpha_j \tau_j^3(x(k)) - (2\alpha_j + 1)\tau_j(x(k)) \right) \exp(-\tau_j^2(x(k))/2) = e(k)J_{c_j}(k), \\ \left\{ \partial E(k) / \partial Q_j^{-1} \right\} = e(k)w_j(k)(x(k) - c_j^\varphi(k))(x(k) - c_j^\varphi(k))^T \cdot \\ \cdot \left(\alpha_j \tau_j^3(x(k)) - (2\alpha_j + 1)\tau_j(x(k)) \right) \exp(-\tau_j^2(x(k))/2) = -e(k)J_{Q_j^{-1}}(k), \\ \nabla_{\alpha_j} E(k) = -e(k)w_j(k)\tau_j^2(x(k)) \exp(-\tau_j^2(x(k))/2) = -e(k)J_{\alpha_j}(k), \end{cases} \quad (23)$$

де $\varphi_j'(\bullet)$ – похідна j -го вейвлета по аргументу, заданому метрикою Ітакури-Сайто [27] $(x(k) - c_j^\varphi(k))^T Q_j^{-1}(k)(x(k) - c_j^\varphi(k))$.

Тоді алгоритм навчання вейвлонів прихованого шару з урахуванням (23) приймає форму

$$\left\{ \begin{array}{l} c_j^\varphi(k+1) = c_j^\varphi(k) - \eta_{c_j^\varphi} e(k) 2w_j(k) \varphi_j'((x(k) - c_j^\varphi(k))^T Q_j^{-1}(k)(x(k) - c_j^\varphi(k))) \cdot \\ \cdot Q_j^{-1}(k)(x(k) - c_j^\varphi(k)) = c_j^\varphi(k) - \eta_{c_j^\varphi} e(k) J_{c_j^\varphi}^\varphi(k), \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) + \eta_{Q_j^{-1}} e(k) w_j(k) \varphi_j'((x(k) - c_j(k))^T Q_j^{-1}(k)(x(k) - c_j(k))) \cdot \\ \cdot (x(k) - c_j(k))(x(k) - c_j(k))^T = Q_j^{-1}(k) + \eta_{Q_j^{-1}} e(k) J_{Q_j^{-1}}^\varphi(k), \\ \alpha_j(k+1) = \alpha_j(k) + \eta_{\alpha_j} e(k) w_j(k) \tau_j^2(x(k)) \exp(-\tau_j^2(x(k))/2) = \alpha_j(k) + \eta_{\alpha_j} e(k) J_{\alpha_j}^\varphi(k), \end{array} \right. \quad (24)$$

при цьому швидкість збіжності до оптимальних значень c_j^φ і Q_j^{-1} повністю визначається параметрами кроку $\eta_{c_j^\varphi}$ й $\eta_{Q_j^{-1}}$.

Як вже було сказано вище, підвищення швидкості збіжності може бути досягнуто шляхом використання більш складних, ніж градієнтні, процедури типу Хартлі або Марквардта, які для настроювання параметрів c_j^φ і α_j можуть бути записані в узагальненій формі таким чином

$$\left\{ \begin{array}{l} c_j^\varphi(k+1) = c_j^\varphi(k) - \lambda_{c_j^\varphi} (J_{c_j^\varphi}^\varphi(k) J_{c_j^\varphi}^{\varphi T}(k) + \eta_{c_j^\varphi} I)^{-1} J_{c_j^\varphi}^\varphi(k) e(k), \\ \alpha_j(k+1) = \alpha_j(k) + \lambda_{\alpha_j} (J_{\alpha_j}^\varphi(k) J_{\alpha_j}^{\varphi T}(k) + \eta_{\alpha_j} I)^{-1} J_{\alpha_j}^\varphi(k) e(k), \end{array} \right. \quad (25)$$

де I – $(n \times n)$ – одинична матриця, $\lambda_{c_j^\varphi}, \lambda_{\alpha_j}$ – додатні параметри демпфування, $\eta_{c_j^\varphi}, \eta_{\alpha_j}$ – параметри регуляризації.

Використовуючи лему обернення матриць, після досить простих, але громіздких перетворень можна отримати простий і ефективний алгоритм навчання центрів вейвлонів і параметра a_j у вигляді

$$\left\{ \begin{array}{l} c_j^\varphi(k+1) = c_j^\varphi(k) - \lambda_{c_j^\varphi} \frac{e(k) J_{c_j^\varphi}^\varphi(k)}{\eta_{c_j^\varphi} + \|J_{c_j^\varphi}^\varphi(k)\|^2}, \\ \alpha_j(k+1) = \alpha_j(k) + \lambda_{\alpha_j} \frac{e(k) J_{\alpha_j}^\varphi(k)}{\eta_{\alpha_j} + \|J_{\alpha_j}^\varphi(k)\|^2} \end{array} \right. \quad (26)$$

з точністю до позначень, що збігається з оптимальним (при $\lambda_{c_j^\varphi} = 1, \eta_{c_j^\varphi} = 0, \lambda_{\alpha_j} = 1, \eta_{\alpha_j} = 0$) за швидкодією алгоритмом Качмажа-Уїдрю-Гоффа.

Для настроювання матриць Q_j^{-1} можна скористатися матричною модифікацією алгоритму (26) у вигляді [19]

$$Q_j^{-1}(k+1) = Q_j^{-1}(k) + \lambda_{Q_j^{-1}} \frac{e(k) J_{Q_j^{-1}}^\varphi(k)}{\eta_{Q_j^{-1}} + \text{Tr}(J_{Q_j^{-1}}^{\varphi T}(k) J_{Q_j^{-1}}^\varphi(k))}, \quad (27)$$

де $\lambda_{Q_j^{-1}}, \eta_{Q_j^{-1}}$ мають такі ж самі значення, що й відповідні параметри в (25).

Таким чином, остаточно алгоритм навчання параметрів вейвлонів прихованого шару в оптимальному за швидкодією варіанті може бути записано як

$$\begin{cases} c_j^\varphi(k+1) = c_j^\varphi(k) - \left\| J_{c_j^\varphi}^\varphi(k) \right\|^{-2} e(k) J_{c_j^\varphi}^\varphi(k), \quad j = 1, 2, \dots, h, \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) + \left(\text{Tr}(J_{Q_j^{-1}}^{\varphi T}(k) J_{Q_j^{-1}}^\varphi(k)) \right)^{-1} e(k) J_{Q_j^{-1}}^\varphi(k), \quad j = 1, 2, \dots, h, \\ \alpha_j(k+1) = \alpha_j(k) + \left\| J_{\alpha_j}^\varphi(k) \right\|^{-2} e(k) J_{\alpha_j}^\varphi(k), \quad j = 1, 2, \dots, h. \end{cases} \quad (28)$$

Відомо, що однокрокові алгоритми типу Качмажа, маючи високу швидкодію, не мають фільтруючих властивостей, тобто погано працюють за умов інтенсивних збурень і завад. Для додання процесу навчання згладжуючих властивостей, що згладжують, використовуючи підхід, який запропоновано в [24], можна ввести таку процедуру навчання:

$$\begin{cases} c_j^\varphi(k+1) = c_j^\varphi(k) - \lambda_{c_j^\varphi} \frac{e(k) J_{c_j^\varphi}^\varphi(k)}{\eta_{c_j^\varphi}(k)}, \quad \eta_{c_j^\varphi}(k+1) = \gamma_{c_j^\varphi} \eta_{c_j^\varphi}(k) + \left\| J_{c_j^\varphi}^\varphi(k+1) \right\|^2, \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) + \lambda_{Q_j^{-1}} \frac{e(k) J_{Q_j^{-1}}^\varphi(k)}{\eta_{Q_j^{-1}}(k)}, \quad \eta_{Q_j^{-1}}(k+1) = \gamma_{Q_j^{-1}} \eta_{Q_j^{-1}}(k) + \text{Tr} \left(J_{Q_j^{-1}}^{\varphi T}(k+1) J_{Q_j^{-1}}^\varphi(k+1) \right), \\ \alpha_j(k+1) = \alpha_j(k) + \lambda_{\alpha_j} \frac{e(k) J_{\alpha_j}^\varphi(k)}{\eta_{\alpha_j}(k)}, \quad \eta_{\alpha_j}(k+1) = \gamma_{\alpha_j} \eta_{\alpha_j}(k) + \left\| J_{\alpha_j}^\varphi(k+1) \right\|^2, \end{cases} \quad (29)$$

(тут $0 \leq \gamma_{c_j^\varphi} \leq 1, 0 \leq \gamma_{Q_j^{-1}} \leq 1, 0 \leq \gamma_{\alpha_j} \leq 1$ – параметри зважування застарілої інформації), що є нелінійним гібридом алгоритму Качмажа-Уїдрой-Гоффа та Гудвіна-Ремеджа-Кейнеса, що має як слідкуючі, так і фільтруючі властивості.

3. ЕКСПЕРИМЕНТАЛЬНЕ МОДЕЛЮВАННЯ

Ефективність запропонованого підходу було продемонстровано при вирішенні задачі емуляції. Типову схему задачі емуляції представлено на рис. 2. Входами вейвлет-нейро-фаззі-системи є зовнішні керуючі сигнали об'єкта та їхні затримані значення, а також затримані значення сигналів об'єкта емуляції на виході.

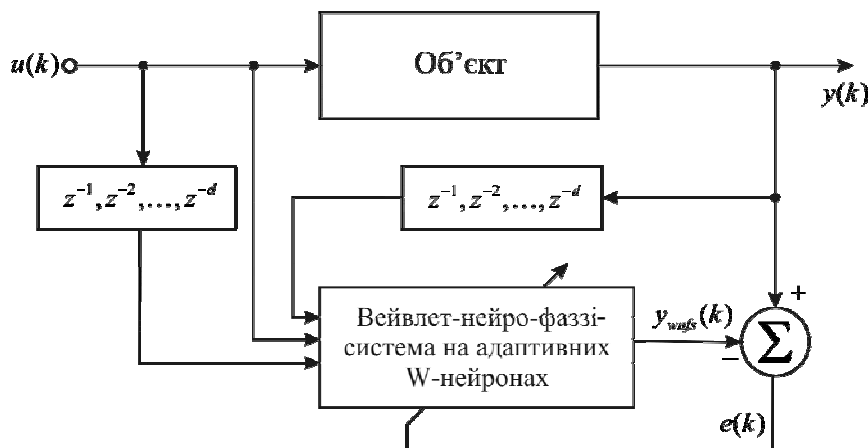


Рис. 2. Типова схема емуляції

Як об'єкт емуляції, взято динамічний нелінійний об'єкт другого порядку [28], що описується різницеvim рівнянням

$$y(k) = \theta(y(k-1), y(k-2), y(k-3), u(k), u(k-1)), \quad (30)$$

де

$$\theta(x_1, x_2, x_3, x_4, x_5) = \frac{x_1 x_2 x_3 x_5 (x_3 - 1) + x_4}{1 + x_3^2 + x_2^2}. \quad (31)$$

Як керуючий сигнал, на вході використовується послідовність

$$u(k) = \begin{cases} \sin(\pi k / 25), & \text{для } k < 250, \\ 1.0, & \text{для } 250 \leq k < 500, \\ -1.0, & \text{для } 500 \leq k < 750, \\ 0.3 \sin(\pi k / 25) + 0.1 \sin(\pi k / 32) + 0.6 \sin(\pi k / 10), & \text{для } 750 \leq k \leq 1000. \end{cases} \quad (32)$$

Для емуляції динамічного об'єкта застосовувалася запропонована гібридна вейвлет-нейро-фаззі-система на W -нейронах з кількістю входів $n = 5$. Кількість вейвлет-активаційних функцій у прихованому шарі та їх початкові параметри були отримані за допомогою методу субтрактивної кластеризації [29]. Застосування цього методу дозволяє отримати не тільки вихідну матрицю координат центрів кластерів, але також вектор, компоненти якого визначають діапазон впливу центра кластера. Початкові значення синаптичних ваг були прийняті нульовими.

У якості критерію прогнозу було використано середньоквадратичну похибку (RMSE)

$$RMSE = \frac{1}{N} \sum_{k=1}^N (y(k) - y_{wnfs}(k))^2. \quad (33)$$

Результати емуляції динамічного об'єкта (30-32) наведено на рис. 3. Як видно, дві криві, що представляють реальні (пунктирна лінія) і модельні (суцільна лінія) значення, практично ідентичні.

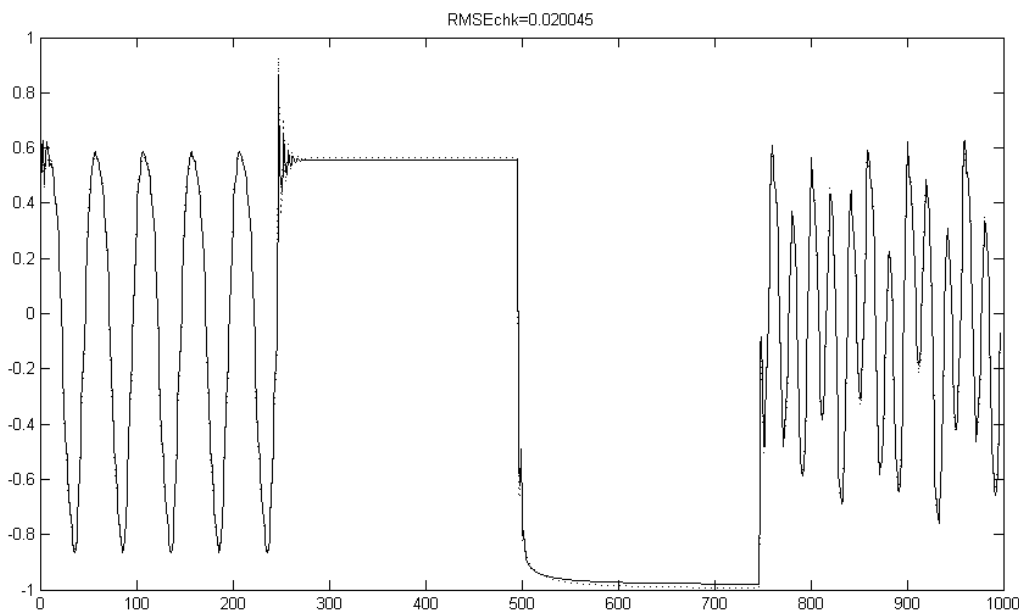


Рис. 3. Результати емуляції динамічного об'єкта (29) – (31)

У таблиці 1 наведено порівняльний аналіз процесу емуляції на основі запропонованої гібридної вейвлет-нейро-фаззі-системи на основі W -нейронів з настроюванням всіх її параметрів з іншими підходами, описаними в літературних джерелах, а саме рекурентною фаззі-нейро-мережею (Recurrent Fuzzy Neural Network, RFNN), рекурентною нейро-фаззі-мережею, що самоорганізується (recurrent self-organizing neural fuzzy inference network, RSONFIN) [30], нейро-фаззі-системою із прямою передачею інформації (feedforward neural fuzzy systems, NFS), рекурентною фаззі-мережею ТСК-типу (TSK-type recurrent fuzzy network, TSK-RFN) [31] і фаззі-вейвлет-нейронною мережею [11].

Порівняльний аналіз результатів емуляції динамічного об'єкта

Нейронна мережа / Алгоритм навчання	Кількість епох навчання	RMSE	
		Навчальна вибірка	Тестова вибірка
Запропонована вейвлет-нейро-фаззі-система на основі адаптивних W -нейронів (Wavelet-Neuro-Fuzzy Systems based on W -neurons, WNFS- W) / РМНК (21) + Запропонований алгоритм навчання всіх параметрів мережі (20), (29)	30	0,0183	0,02004
Фаззі-вейвлет-нейронна мережа [11]	200	0,0282	0,0301
Рекурентна фаззі-нейро-мережа (Recurrent Fuzzy Neural Network, RFNN)	200	0,0114	0,0575
Рекурентна нейро-фаззі-мережа, що самоорганізується (Recurrent Self-Organizing Neural Fuzzy Inference Network, RSONFIN) [30]	200	0,0248	0,0780
Нейро-фаззі-система із прямою передачею інформації (Feedforward Neural Fuzzy Systems, NFS)	200	0,0203	0,0521
Рекурентна фаззі-мережа ТСК-типу (TSK-type Recurrent Fuzzy Network, TSK-RFN) [31]	200	0,0084	0,0346

З наведених вище результатів випливає, що запропонована архітектура гібридної вейвлет-нейро-фаззі-системи та алгоритм навчання всіх її параметрів забезпечують більш високу точність емуляції в порівнянні з іншими підходами за менший час навчання.

ВИСНОВКИ

Запропонована архітектура має покращені апроксимуючі властивості за рахунок введення адаптивного вейвлону в консеквент, а також настроювання всіх параметрів мережі. Таку вейвлет-нейро-фаззі-архітектуру може бути використано для вирішення задач діагностики, прогнозування, емуляції та ідентифікації нестационарних процесів. Експериментальне моделювання було проведено на нестационарних рядах технічної та біологічної природи. Порівняння результатів експериментів з результатами відомих нейро-фаззі-систем довело переваги запропонованого підходу.

ЛІТЕРАТУРА

1. Kosko B. Fuzzy systems as universal approximators / B. Kosko // Proc. 1-st IEEE Int. Conf. on Fuzzy Systems. – San Diego, CA. – 1992. – P. 1153-1162.
2. Kreinovich V. Fuzzy rule based modeling as a universal approximation tool / V. Kreinovich, Y.C. Mouzouris, H.T. Nguen – In. «Fuzzy systems: Modeling and Control». Eds.: H.T. Nguen, M. Sugeno. – Boston: Kluwer Academic Publishers, 1998. – P. 135-195.
3. Hornik K. Multilayer feedforward networks are universal approximators / K. Hornik, M. Stinchcombe, H. White // Neural Networks. – 1982. – 2. – P. 359-366.
4. Scarcelli F. Universal approximation using feedforward neural networks: a survey of some existing methods and some new results / F. Scarcelli, A.S. Tsoi // Neural Networks. – 1998. – 11. – P. 15-37.
5. Zhang Q.H. Using wavelet network in nonparametric estimation / Q.H. Zhang // IEEE Trans. on Neural Networks. – 8 (2). – 1997. – P. 227-236.
6. Bodyanskiy Ye. Adaptive wavelet-neuro-fuzzy network in the forecasting and emulation tasks / Ye. Bodyanskiy, I. Pliss, O. Vynokurova // Int. Journal on Information Theory and Applications. – V.15. – 1. – 2008. – P. 47-55
7. Бодянский Е.В. Анализ клинических данных в медицинских исследованиях на основе методов вычислительного интеллекта / В.В. Бойко, Е.В. Бодянский, Е.А. Винокурова, С.В. Сушков, А.А. Павлов – Харьков: ТО Ексклюзив, 2008. – 120 с.

8. Chui C.K. An Introduction to Wavelets / C.K. Chui. – New York: Academic, 1992. – 264 p.
9. Jang J.-S. R. ANFIS: Adaptive network-based fuzzy inference systems / J.-S. R. Jang // IEEE Trans. on Systems, Man, and Cybernetics. – 23. – 1993. – P. 665-685.
10. Jang J.-S. R. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence / J.-S. R. Jang, C.-T. Sun, E. Mizutani. – Upper Saddle River, N.J.: Prentice Hall, Inc., 1997. – 614 p.
11. Abiyev R.H. Identification and control of dynamic plant using fuzzy wavelet neural networks / R.H. Abiyev, O. Kaynak // Proc. of IEEE Int. Symposium on Intelligent Control. – USA: San Antonio. – 2008. – P. 1295-1301.
12. Abiyev R.H. Fuzzy wavelet neural networks for identification and control of dynamic plants – A novel structure and a comparative study / R.H. Abiyev, O. Kaynak // IEEE Trans. on Industrial Electronics. – 2008. – 55(8). – P. 3133-3140.
13. Mitaim S. What is the best shape for a fuzzy set in function approximation? / S. Mitaim, B. Kosko // Proc. 5th IEEE Int. Conf on Fuzzy Systems «Fuzz-96». – V. 2. – 1996. – P. 1237-1213.
14. Mitaim S. Adaptive joint fuzzy sets for function approximation / S. Mitaim, B. Kosko // Proc. Int. Conf. on Neural Networks. – 1997. – P. 537-542.
15. Lekutai G. Self-tuning control of nonlinear systems using neural network adaptive frame wavelets / G. Lekutai, H.F. VanLandingham // Proc. IEEE Int. Conf. on Systems, Man and Cybernetics. Piscataway, N.J. – 2. – 1997. – P. 1017-1022.
16. Винокурова Е.А.. Генератор аналитических вэйвлетов. / Е.А. Винокурова, Н.С. Ламонова, И.П. Плисс // Проблемы бионики. – Вып. 60. – 2004. – С. 104-109.
17. Бодянский Е.В. Треугольный вэйвлет и формальный нейрон на его основе / Е.В. Бодянский, Е.А. Винокурова // Сб. наук. праць 3-ої Міжнародної науково-практичної конференції «Математичне та програмне забезпечення інтелектуальних систем» (MPZIS-2005) – Дніпропетровськ: ДНУ. – 2005. – С. 14-15.
18. Бодянский Е.В., Винокурова Е.А. Адаптивный вэйвлон и алгоритм его обучения // Управляющие системы и машины. – 2009. – 1 (219). – С. 47-53.
19. Bodyanskiy Ye. Hybrid radial-basis neuro-fuzzy wavelon in the non-stationary sequences forecasting problems / Ye. Bodyanskiy, O. Vynokurova // Proc. 2nd Int. Conf. on Inductive Modelling. – Kyiv. 2008.– P.144-147.
20. Bodyanskiy Ye. Radial-basis-fuzzy-wavelet-neural network with adaptive activation-membership function / Ye. Bodyanskiy, E. Yegorova, O. Vynokurova // Int. Journal on Artificial Intelligence and Machine Learning. – 2008. – V. 8. – II. – P. 9-15.
21. Hartley H. The modified Gauss-Newton method for the fitting of nonlinear regression functions of least squares / H. Hartley // Technometrics. – 3. – 1961. – P. 269-280.
22. Marquardt D. An algorithm for least squares estimation of nonlinear parameters. / D. Marquardt // SIAM J. Appl. Math. – 11. – 1963. – P. 431-441.
23. Бодянский Е.В. Адаптивные алгоритмы идентификации нелинейных объектов управления / Е.В. Бодянский // АСУ и приборы автоматики. – Харьков: Выща шк., 1987. – Вып. 81. – С. 43-46.
24. Bodyanskiy Ye. An adaptive learning algorithm for a neuro-fuzzy network. / Ye. Bodyanskiy, V. Kolodyazhnyi, A. Stephan / Ed. by B. Reusch «Computational Intelligence. Theory and Applications.» – Berlin-Heidelberg-New York: Springer-Verlag. – 2001. – P. 68-75.
25. Льюнг Л. Идентификация систем. Теория для пользователя / Л. Льюнг. – Москва: Наука. ГРФМЛ, 1991. – 432 с.
26. Bishop C. M. Neural Networks for Pattern Recognition / C.M. Bishop. – Oxford: Clarendon Press, 1995. – 482 p.
27. Itakura F. Maximum prediction residual principle applied to speech recognition / F. Itakura // IEEE Trans. on Acoustics, Speech and Signal Processing. – 23. – 1975. – P. 67-72.
28. Narendra K. S. Identification and control of dynamic systems using neural networks / K.S. Narendra, K. Parthasarathy // IEEE Trans. on Neural Networks. – 1990 – 1. – P. 4-26.
29. Chiu S. Fuzzy model identification based on cluster estimation / S. Chiu// Journal of Intelligent and Fuzzy Systems. – 1994. – 2. – №3. – P. 267-278.
30. Juang C.-F. A recurrent self-organizing neural fuzzy inference network / C.-F. Juang, C.-T. Lin // IEEE Trans. Neural Networks. – 10 (4). – 1999 – P. 828-845.
31. Juang C.-F. A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithm / C.-F. Juang // IEEE Trans. Fuzzy Systems. – 10. – 2002. – P. 155-170.

Рецензенти: д.т.н., проф. Фісун М.Т.,
д.т.н., проф. Кондратенко Ю.П.

© Бодянский С.В., Винокурова О.А., 2009

Стаття надійшла до редакції 15.08.09