

ВІДНОВЛЕННЯ ФУНКЦІОНАЛЬНИХ ЗАЛЕЖНОСТЕЙ НА ОСНОВІ НЕЙРО-ФАЗЗИ- ПІДХОДУ В ПРОЦЕСІ ПОСЛІДОВНОЇ ОБРОБКИ ІНФОРМАЦІЇ

У статті запропоновано узагальнену регресійну нейро-фаззи-мережу, що дозволила забезпечити підвищення точності й швидкості прогнозування та ідентифікації нелінійних нестационарних об'єктів за рахунок використання в першому прихованому шарі одновимірних функцій належності замість багатовимірних функцій активації. Результати чисельних експериментів, що наведені в статті, підтверджують переваги запропонованої нейронної мережі.

Ключові слова: мережі, засновані на пам'яті; нейрони в точках даних; фаззи-базисні функції; нестационарність; адаптивність; системи нечіткого виведення.

В статье предложена обобщенная регрессионная нейро-фаззи-сеть, которая позволила обеспечить повышение точности и скорости прогнозирования и идентификации нелинейных нестационарных объектов за счет использования в первом скрытом слое одномерных функций принадлежности вместо многомерных функций активации. Результаты численных экспериментов, приведенные в статье, подтверждают преимущества предложенной нейронной сети.

Ключевые слова: сети, основанные на памяти; нейроны в точках данных; фаззи-базисные функции; нестационарность; адаптивность; системы нечеткого вывода.

General regression neuro-fuzzy network, that allowed to provide accuracy and rate increasing of nonlinear nonstationary plants prediction and identification by using in the first hidden layer one-dimensional membership functions instead of multi-dimensional activation functions is proposed in this article. The results of numeric experiments corroborate the advantages of proposed neural network.

Key words: memory-based networks; neurons at data points; fuzzy-basis functions; nonstationarity, adaptivity; fuzzy inference systems.

ВСТУП

У теперішній час більшість реальних задач обробки інформації характеризуються тим, що дані надходять послідовно в реальному часі, їх аналіз повинен проводитися в темпі функціонування об'єкта, а сам об'єкт є нестационарним. Для вирішення подібних задач широкого поширення набули штучні нейронні мережі. Проте класичний багатошаровий перцептрон, який є універсальним апроксиматором, у цьому випадку не підходить, а в якості альтернативи йому в ряді випадків можуть бути використані радіально-базисні нейронні мережі (РБНМ) [1; 2]. Разом з тим практичне застосування РБНМ обмежується так званим прокляттям розмірності, а також виникненням пропусків у просторі радіально-базисних функцій, що призводять до виникнення областей, де всі нейрони мережі не є активними. Крім того, ці мережі навчаються за допомогою рекурентних градієнтних алгоритмів, загальним недоліком

яких є недостатня швидкість збіжності та можливість попадання в локальні мінімуми прийнятого критерію навчання.

Таким чином, традиційні нейромережі, що використовують рекурентні процедури навчання та об'єднані загальною назвою «мережі, засновані на оптимізації», можуть виявитися неефективними в задачах адаптивної ідентифікації, прогнозування та управління реального часу, коли інформація на обробку надходить з високою частотою. Мережа в цьому випадку не встигає навчатися та не має можливості слідкувати за параметрами об'єкта, що змінюються.

Альтернативою мережам, що засновані на оптимізації, є «мережі, засновані на пам'яті», найбільш яскравим представником яких є узагальнена регресійна нейронна мережа (УРНМ), що була запропонована Д. Шпехтом [3]. Навчання цієї мережі зводиться до одноразової установки багатовимірних радіально-базисних функцій (РБФ) у точках одиничного центрованого гіперкуба, які однозначно задаються навчальною вибіркою. Таким чином, ці мережі можуть бути віднесені до так званих моделей, що миттєво навчаються [4] та настроюються за допомогою одного проходу алгоритму навчання. УРНМ навчається дуже швидко, встановлюючи центри РБФ в точках з координатами, що визначаються вхідними сигналами об'єкта за принципом «нейрони в точках даних», та з висотою РБФ, що співпадає з відповідними значеннями вихідного сигналу об'єкта. Саме висока швидкість навчання УРНМ забезпечила їх ефективне використання в задачах реального часу [5; 6].

Основні проблеми, пов'язані з використанням УРНМ, визначаються можливим прокляттям розмірності, що породжується зростанням об'єму навчальної вибірки та труднощами з обґрунтованим завданням порогового параметра, який достатньо важко вибирати та інтерпретувати в багатовимірному просторі.

УЗАГАЛЬНЕНА РЕГРЕСІЙНА НЕЙРО-ФАЗЗИ-МЕРЕЖА

У статті запропоновано узагальнену регресійну нейро-фаззі-мережу (УРНФМ), яка є нейро-фаззі-системою (НФС), що навчається за типом узагальненої регресійної нейронної мережі (УРНМ), яка забезпечує їй апроксимуючі можливості відомої ANFIS [7] зі швидкістю навчання мереж, заснованих на пам'яті.

Так, для вирішення задачі ідентифікації нелінійного об'єкта, $(y(k), x(k))$ – скаляр та $(n \times 1)$ – вектор вихідних і вхідних сигналів відповідно в дискретний момент часу $k = 1, 2, \dots, F(\bullet)$ – невідомий нелінійний оператор об'єкта). За допомогою УРНМ необхідно сформулювати навчальну вибірку $\{x^*(k), y^*(k)\}$, $k = 1, 2, \dots, l$, після чого можна отримати оцінку $\hat{y}(k)$ реакції об'єкта $y(k)$ на довільний вхідний сигнал x у вигляді

$$\hat{y}(x) = \frac{\sum_{k=1}^l y^*(k) \varphi(D(k))}{\sum_{k=1}^l \varphi(D(k))}, \quad (1)$$

де $D(k)$ – міра відстані в прийнятій метриці між $x(k)$ та $x^*(k)$; $\varphi(\bullet)$ – деяка ядерна функція, частіше за все гауссіан.

В якості відстані зазвичай використовується евклідова метрика

$$D^2(k) = \sum_{i=1}^n \left(\frac{x_i(k) - x_i^*(k)}{\sigma(k)} \right)^2, \quad (2)$$

(тут $\sigma(k)$ – скалярний параметр, що визначає радіус рецепторного поля ядерної функції $\varphi(\bullet)$), хоча в більш загальному випадку можливе використання метрики Мінковського

$$D^p(k) = \sum_{i=1}^n \left| \frac{x_i(k) - x_i^*(k)}{\sigma(k)} \right|^p, \quad p \geq 1. \quad (3)$$

Із збільшенням об'єму навчальної вибірки l УРНМ асимптотично сходиться до оптимальної нелінійної регресійної поверхні [8].

Процес навчання УРНМ досить просто організувати в реальному часі. При цьому навчальні пари $x^*(k), y(k)$ надходять в мережу послідовно, формуючи нові радіально-базисні функції-нейрони. При цьому оцінюється відстань між знов сформованою і вже існуючими функціями. Якщо ця відстань менша деякого наперед заданого порогового значення r , новий нейрон виключається з мережі.

Архітектура запропонованої узагальненої регресійної нейро-фаззі-мережі наведена на рис. 1 та складається з п'яти послідовно з'єднаних шарів. Перший прихований шар складається з l блоків по n фаззі-базисних функцій (ФБФ) у кожному та виконує фаззіфікацію вектора вхідних змінних. Другий прихований шар здійснює агрегування рівнів належності, що були обчислені в першому шарі, і складається з l блоків множення. Третій прихований шар – це шар синаптичних ваг, які визначаються спеціальним способом. Четвертий шар утворений двома суматорами та обчислює суми вихідних сигналів другого і третього шарів. І, нарешті, у п'ятому (вихідному) шарі проводиться нормалізація, у результаті якої обчислюється вихідний сигнал мережі.

Як видно, архітектура УРНФМ близька до архітектури нейро-фаззі системи Ванга-Менделя, що є, у свою чергу, модифікацією системи нечіткого виведення Такагі-Сугено нульового порядку [7; 9; 10]. Проте, якщо нейро-фаззі-системи навчаються за допомогою тих або інших процедур оптимізації, УРНФМ настраюється за допомогою одного проходу алгоритму навчання, що істотно підвищує швидкість обробки даних.

Оскільки УРНФМ належить до мереж, заснованих на пам'яті, її навчання проводиться за принципом «нейрони в точках даних», що робить його украй простим і швидким в порівнянні з навчанням за допомогою алгоритмів оптимізації. Вектори навчальної вибірки $x^*(1), \dots, x^*(k), \dots, x^*(l)$ нормуються на одиничний центрований гіперкуб так, що якщо

$$x_i^{*min} \leq x_i^*(k) \leq x_i^{*max}, \quad i = 1, 2, \dots, n, \text{ то}$$

$$-0,5 \leq \tilde{x}_i^*(k) \leq 0,5.$$

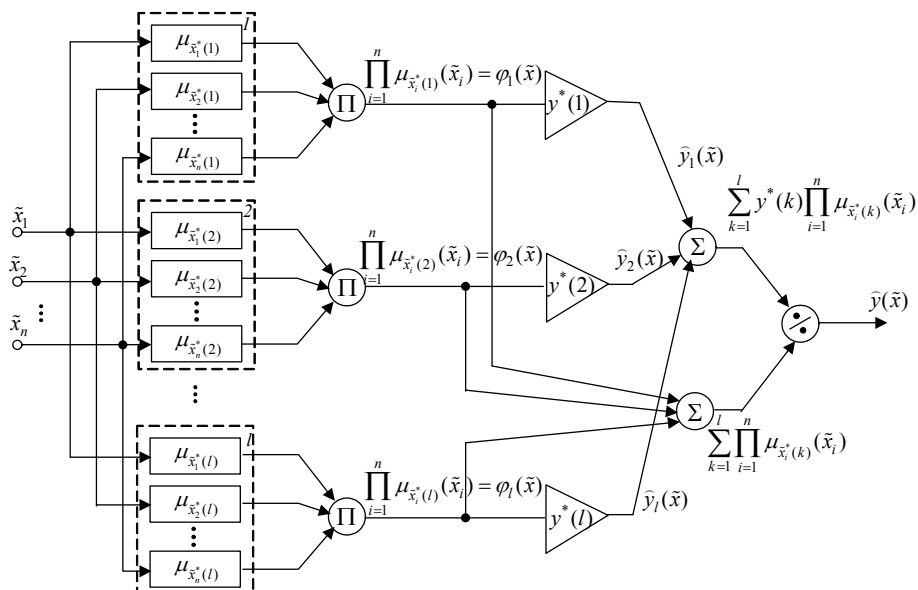


Рис. 1. Модель узагальненої регресійної нейро-фаззі-мережі

Взаємний перерахунок проводиться згідно з виразами

$$\tilde{x}_i^*(k) = \frac{x_i^*(k) - x_i^{*\min}}{x_i^{*\max} - x_i^{*\min}} - 0,5,$$

$$x_i^*(k) = (\tilde{x}_i^*(k) + 0,5)(x_i^{*\max} - x_i^{*\min}) + x_i^{*\min}.$$

Для кожного з векторів навчальної вибірки $\tilde{x}^*(k) = (\tilde{x}_1^*(k), \tilde{x}_2^*(k), \dots, \tilde{x}_n^*(k))^T$ в першому прихованому шарі формується власний набір фаззі-базисних функцій належності $\mu_{\tilde{x}_1^*(k)}, \mu_{\tilde{x}_2^*(k)}, \dots, \mu_{\tilde{x}_n^*(k)}$ так, що центри $\mu_{\tilde{x}_i^*(k)}$ співпадають з $\tilde{x}_i^*(k)$, $k = 1, 2, \dots, l$.

Процес формування фаззі-базисних функцій ілюструється на рис. 2.

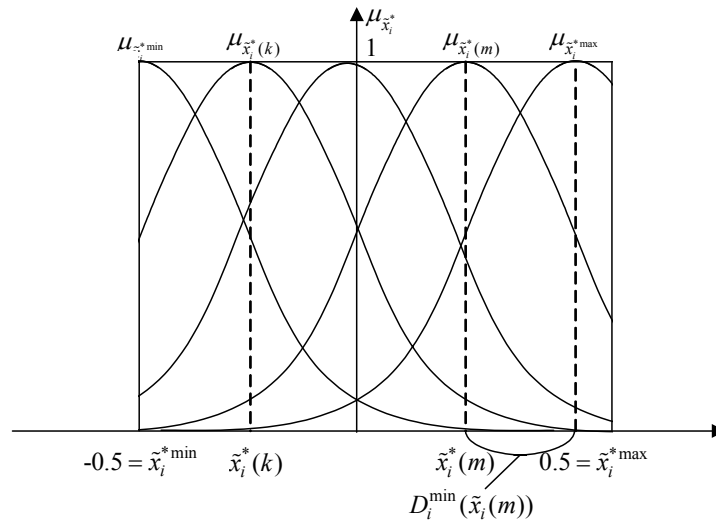


Рис. 2. Фаззі-базисні функції

Треба відмітити, що УРНФМ містить nl фаззі-базисних функцій, що ніяк не може призвести до проблеми розмірності.

У якості ФБФ в принципі може використовуватися будь-яка ядерна функція з нестрого локальним рецепторним полем, що дозволяє уникнути виникнення пропусків у просторі ознак [8]. У якості такої функції можна рекомендувати узагальнений гауссіан

$$\mu_{\tilde{x}_i^*(k)}(\tilde{x}_i) = \left(1 + \left| \frac{\tilde{x}_i^*(k) - \tilde{x}_i}{\sigma_i(k)} \right|^{2b} \right)^{-1}, \quad b \geq 0,5, \quad (4)$$

що є дзвонуватою функцією, форма якої визначається скалярним параметром b [11]. Відмітимо, що b визначає також і метрику $D^{2b}(k)$. Що стосується вибору параметра ширини $\sigma_i(k)$, то стандартна рекомендація зводиться до того, що він повинен забезпечувати мале перекриття сусідніх ФБФ. Нескладно бачити, що для гауссіана ця рекомендація веде до оцінки

$$\sigma_i(k) < \frac{l-1}{2 \div 3}.$$

Одночасно з формуванням ФБФ першого прихованого шару в третьому шарі відбувається установка синаптичних ваг, які приймаються рівними сигналам навчальної вибірки $y^*(k)$.

Таким чином, при подачі на вхід УРНФМ довільного сигналу \tilde{x} в першому прихованому шарі обчислюються рівні належності $\mu_{\tilde{x}_i^*(k)}(\tilde{x}_i)$, $i = 1, 2, \dots, n$, $k = 1, 2, \dots, l$, у другому шарі проводиться їх агрегування шляхом формування багатовимірних ФБФ

$$\varphi_k(\tilde{x}) = \prod_{i=1}^n \left(1 + \left| \frac{\tilde{x}_i^*(k) - \tilde{x}_i}{\sigma_i(k)} \right|^{2b} \right)^{-1}, \quad k = 1, 2, \dots, l, \quad (5)$$

у третьому шарі обчислюються добутки $\hat{y}(\tilde{x}) = y^*(k)\varphi_k(\tilde{x})$, четвертий шар обчислює значення сигналів $\sum_{k=1}^l y^*(k)\varphi_k(\tilde{x})$ та $\sum_{k=1}^l \varphi_k(\tilde{x})$, і, нарешті, у вихідному шарі формується оцінка

$$\hat{y}(\tilde{x}) = \frac{\sum_{k=1}^l y^*(k)\varphi_k(\tilde{x})}{\sum_{k=1}^l \varphi_k(\tilde{x})} = \frac{\sum_{k=1}^l y^*(k) \prod_{i=1}^n \mu_{\tilde{x}_i^*(k)}(\tilde{x}_i)}{\sum_{k=1}^l \prod_{i=1}^n \mu_{\tilde{x}_i^*(k)}(\tilde{x}_i)}, \quad (6)$$

що співпадає з (1) з тією лише різницею, що замість радіально-базисних використовуються багатовимірні фаззи-базисні функції, сформовані з одновимірних ФБФ.

Схему нечіткого виводу, що реалізовується УРНФМ, можна представити у вигляді системи логічних рівнянь

$$IF(\tilde{x}_1.IS.A_1(1)).AND.(\tilde{x}_2.IS.A_2(1)).AND.....AND.(\tilde{x}_n.IS.A_n(1)), \quad THEN \quad \hat{y}_1(\tilde{x}) = y^*(1)$$

⋮

$$IF(\tilde{x}_1.IS.A_1(k)).AND.(\tilde{x}_2.IS.A_2(k)).AND.....AND.(\tilde{x}_n.IS.A_n(k)), \quad THEN \quad \hat{y}_k(\tilde{x}) = y^*(k)$$

⋮

$$IF(\tilde{x}_1.IS.A_1(l)).AND.(\tilde{x}_2.IS.A_2(l)).AND.....AND.(\tilde{x}_n.IS.A_n(l)), \quad THEN \quad \hat{y}_l(\tilde{x}) = y^*(l),$$

де оператор $A_i(k)$ представляється функцією належності (4).

Таким чином, використання нейро-фаззи-підходу дозволяє забезпечити інтерпретуємість результатів, що було отримано [12].

Процес навчання УРНФМ може протікати як в пакетному режимі, коли задано всю навчальну вибірку $\{x^*(k), y^*(k)\}$, так і в реальному часі, коли пари $x^*(k), y^*(k)$ поступають послідовно, формуючи багатовимірні ФБФ φ_k . При цьому досить просто організувати процес виключення малоінформативних пар. Припустимо, що для деякого спостереження $\tilde{x}^*(m)$ виконується умова

$$\max_i D_i^{\min}(\tilde{x}_i(m)) < r < l-1, \quad (7)$$

де $D_i^{\min}(\tilde{x}_i(m))$ – найменша відстань між $\tilde{x}_i(m)$ і раніше сформованими сусідніми центрами ФБФ.

У такому разі $\tilde{x}^*(m)$ не утворює функцію φ_m і виключається з розгляду. Відмітимо, що для одновимірної ситуації пороговий параметр r і відстань D_i^{\min} визначити значно простіше, ніж в багатовимірному випадку УРНМ.

Роботу УРНФМ нескладно організувати в режимі безперервної адаптації [13], і що є важливим для ідентифікації і управління нестационарними об'єктами. Тут можливе використання двох підходів: на ковзному вікні із l спостережень, коли під час надходження на вхід мережі навчальної пари $\tilde{x}^*(l+1), y^*(l+1)$ у першому і третьому шарах виключається

набір з $\mu_{\tilde{x}_i}^*(1)$ та $y^*(1)$, а замість нього формується $\mu_{\tilde{x}_i}^*(l+1)$ та вага $y^*(l+1)$ з використанням нерівності (8).

Оскільки процес навчання відбувається практично миттєво, питання про слідкуючі властивості алгоритму навчання взагалі не виникає. Ідея, що розвивається, «нейрони в точках даних» у поєднанні з нечіткою логікою дозволяє вирішувати більш широкий клас задач в режимі послідовної обробки даних за умов нестационарності.

ЧИСЕЛЬНІ ЕКСПЕРИМЕНТИ

Для моделювання роботи узагальненої регресійної нейро-фаззи-мережі були використані репозиторні дані Вира [14], що описують порушення роботи печінки та містять 345 спостережень (6 ознак). Результати роботи УРНФМ порівнюються з результатами роботи стандартної узагальненої регресійної нейронної мережі. У якості оцінки одержаних результатів використовувалася середня абсолютна процентна похибка (Mean average percentage error, MAPE).

У таблиці 1 наведені результати роботи нейромереж УРНФМ та УРНМ для різних об'ємів навчальних вибірок та при різних значеннях порогового параметра r .

Таблиця 1

Результати роботи УРНФМ та УРНМ

Об'єм навчальної вибірки	Похибка УРНФМ	Похибка УРНМ
150	0,74	0,81
200	0,60	0,69
250	0,41	0,53
300	0,22	0,30
330	0,10	0,16
Пороговий параметр r (об'єм навчальної вибірки)		
0,9 (181)	0,69	0,73
0,7 (205)	0,55	0,65
0,5 (248)	0,40	0,48
0,3 (270)	0,33	0,40
0,1 (281)	0,21	0,29

Також був проведений експеримент щодо ідентифікації стандартного тестового об'єкта [15]

$$y(k+1) = f(y(k), y(k-1), y(k-2), u(k), u(k-1)), \quad (8)$$

де невідома функція f має вигляд

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

Вхідний сигнал об'єкта обчислюється згідно з виразом $u(k) = \sin(2\pi k / 250)$ для $k \leq 500$ та $u(k) = 0,8 \sin(2\pi k / 250) + 0,2 \sin(2\pi k / 25)$ для $k > 500$, всього 1000 сигналів.

Для різної кількості вхідних сигналів, які утворили навчальну вибірку, була побудована та навчена УРНФМ. Спостереження, що залишилися, були подані на входи навченої мережі для перевірки роботи. Також була навчена і стандартна УРНМ для порівняння результатів роботи. У якості оцінки одержаних результатів використовувалася середня абсолютна процентна похибка (Mean average percentage error, MAPE).

У таблиці 2 наведені результати роботи нейромереж.

Результати роботи УРНФМ та УРНМ

Об'єм навчальної вибірки	Похибка УРНФМ	Похибка УРНМ
300	0,85	0,88
400	0,76	0,80
500	0,62	0,67
600	0,51	0,55
700	0,42	0,46
800	0,24	0,32

Далі на вхід УРНФМ подавалася вибірка з перших 500 спостережень. Навчальну вибірку сформували 378 сигналів, відстані між якими перевищували заздалегідь задане порогове значення. Результати цього експерименту показали, що сформована таким чином навчальна вибірка дає точність апроксимації УРНФМ в середньому на 9 % вищу, ніж УРНМ.

З табл. 1-2 видно, що точність УРНФМ при заздалегідь заданій навчальній вибірці фіксованого об'єму в середньому на 7 % вища, ніж точність УРНМ. А за умови використання різних порогових значень, тобто урахування найбільш інформативних ознак для формування навчальної вибірки, точність УРНФМ в середньому на 8 % вищу, ніж точність УРНМ за тих же умов.

ВИСНОВОК

У статті запропоновано модель узагальненої регресійної нейро-фаззі-мережі, яка поєднує в собі традиційну узагальнену регресійну нейронну мережу й адаптивну систему нечіткого виведення і характеризується обчислювальною простотою, інтерпретованістю результатів, високою точністю в задачах прогнозування та ідентифікації нелінійних нестационарних об'єктів та може використовуватися в задачах відновлення функціональних залежностей у режимі послідовної обробки.

ЛІТЕРАТУРА

1. Moody J., Darken C.J. Fast learning in networks of locally-tuned processing units // *Neural Computation*. – 1989. – 1. – P. 281-294.
2. Schilling R.J., Carrol J.J., Al-Ajlouni A.F. Approximation of nonlinear systems with radial basis function neural networks // *IEEE Trans. on Neural Networks*. – 2001. – 12. – P. 1-15.
3. Specht D.E. A general regression neural network // *IEEE Trans. on Neural Networks*. – 1991. – 2. – P. 568-576.
4. Nelles O. *Nonlinear System Identification*. – Berlin: Springer, 2001. – 785 p.
5. Seng T.L., Khalid M., Yusof R., Omatu S. Adaptive neuro-fuzzy control system by RBF and GRNN neural networks // *J. of Intelligent and Robotic Systems*. – 1998. – 23. – P. 267-289.
6. Guo X.-P., Wang F.-L., Jia M.-X. A sub-stage moving window GRNN quality prediction method for injection molding process // *«Lecture Notes in Computer Science»*. – V3973. – Berlin-Heidelberg: Springer-Verlag, 2006. – P. 1138-1143.
7. Jang J.-S. R. ANFIS: Adaptive-network-based fuzzy inference systems // *IEEE Trans. Syst., Man, and Cybern.* – 1993. – 23. – № 3. – P. 665-685.
8. Friedman J., Hastie T., Tibshirani R. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. – Berlin: Springer, 2003. – 552 p.
9. Jang J.-S.R., Sun G.-T. Neuro-fuzzy modeling and control // *Proc. IEEE*. – 1995. – 83. – P. 378-406.
10. Cios K.J., Pedrycz W. Neuro-fuzzy algorithms // In: *«Handbook on Neural Computation»*. – Oxford: University Press, 1997. – D1.3:1-D1.3:7.
11. Jang J.-S. R., Sun G.-T., Mizutani E. *Neuro-Fuzzy and Soft Computing*. – Upper Saddle River, NJ: Prentice Hall, 1997. – 614 p.
12. Bodyanskiy Ye., Teslenko N. General regression neuro-fuzzy network for identification of nonstationary plants // *Information Technologies and Knowledge*. – 2008. – 2. – P. 136-142.
13. Bodyanskiy Ye., Pliss I., Otto P., Teslenko N. Nonlinear process identification and modeling using general regression neuro-fuzzy network // *Computer Science Meets Automation: Proc. 52nd Int. Sci. Colloquium*. – Ilmenau, 2007. – P. 27-33.

14. Murphy P.M., Aha D. UCI Repository of machine learning databases [Электронный ресурс] // CA: University of California, Department of Information and Computer Science: <http://www.ics.uci.edu/~mlearn/MLRepository.html>.
15. Narendra K.S., Parthasarathy K. Identification and control of dynamical systems using neural networks // IEEE Trans. on Neural Networks. – 1990. – 1. – P. 4-26.

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

© Тесленко Н.О., 2009

Стаття надійшла до редакції 14.09.09