

УДК 004.032.26

ФУНКЦИОНАЛЬНО СВЯЗАННАЯ МГУА-ВЭЙВЛЕТ-НЕЙРО-ФАЗЗИ СИСТЕМА И РОБАСТНЫЙ АЛГОРИТМ ЕЕ ОБУЧЕНИЯ

Е.В. Бодянский, Е.А. Винокурова

*Харьковский национальный университет радиоэлектроники
bodya@kture.kharkov.ua, vinokurova@kture.kharkov.ua*

В статті запропоновано функціонально зв'язану МГУА-вейвлет-нейро-фаззі-систему, що є гібридною архітектурою обчислювального інтелекту. Синтезовано робастний алгоритм навчання на основі робастного критерію оптимізації. Така гібридна система дозволяє оброблювати сигнали за умов значного рівня завад з аномальними викидами.

Ключові слова: МГУА, функціонально зв'язані мережі, вейвлет-нейро-фаззі-система, робастний алгоритм навчання

In the paper the functional linked GMDH-wavelet-neuro-fuzzy system that is hybrid architecture of computational intelligence is proposed. The robust learning algorithm based on logistic optimization criterion is synthesized. Such hybrid system allows to process signals under significant level of a noise with abnormal outliers.

Keywords: GMDH, functional linked networks, wavelet-neuro-fuzzy system, robust learning algorithm

В статье предложена функционально связанная МГУА-вейвлет-нейро-фаззи-система, являющаяся гибридной архитектурой вычислительного интеллекта. Синтезирован робастный алгоритм обучения на основе логистического критерия оптимизации. Такая гибридная система позволяет обрабатывать сигналы в условиях большого уровня помех с аномальными выбросами.

Ключевые слова: МГУА, функционально связанные сети, вейвлет-нейро-фаззи система, робастный алгоритм обучения

Введение

В настоящее время для решения задач интеллектуального анализа данных широкое распространение получили системы, основанные на многослойных сетях, а именно персептронные архитектуры, которые характеризуются громоздкостью и низкой скоростью обучения. Другой ветвью развития многослойных архитектур являются, так называемые, функционально связанные нейронные сети, обладающие рядом преимуществ перед многослойными персептронами [1].

С другой стороны, на данный момент большое внимание уделяется гибридным архитектурам нейро-фаззи систем, которые позволяют объединить преимущества разных подходов, что, в свою очередь, позволяет повысить точность решения задач интеллектуального анализа данных.

Актуальным является объединение различных подходов, что позволяет взять от нейронных сетей – способность обучаться и хорошие аппроксимирующие свойства [2, 3], от фаззи-систем – их интерпретируемость [4], от вейвлет-теории – способность вейвлет-функций выявлять локальные особенности сигналов [5-7], а от индуктивного моделирования (МГУА-

системы) – во-первых, способность селекции входных данных (решение задачи выбора важнейших входных данных), во-вторых, способность оптимизации архитектуры системы в процессе обучения [8-10].

В статье предлагается архитектура функционально связанной вэйвлет-нейро-фаззи системы (ФС МГУА-ВНФС) и робастный алгоритм ее обучения.

1. Функционально связанная МГУА вэйвлет-нейро-фаззи система

В общем случае функционально связанная нейронная сеть [1] обладает нелинейными аппроксимирующими свойствами и обучается с помощью линейных алгоритмов. Скрытый слой такой сети расширяет входное пространство с помощью некоторого нелинейного преобразования, в данном случае с помощью многомерных адаптивных вэйвлет-функций так, что обеспечивается отображение входных сигналов в новое пространство повышенной размерности.

Основным недостатком такой сети является большое число настраиваемых параметров, которые растут как полиномиальная функция, что может привести к «проклятию размерности» или снижению скорости обучения системы.

В этом случае для сокращения количества настраиваемых параметров с успехом может быть использован метод группового учета аргументов (МГУА), предложенный А.Г. Ивахненко, развитый его учениками [8-10] и положенный в основу так называемой МГУА-нейронной сети. В отличие от стандартных нейронных сетей с полносвязной фиксированной архитектурой, МГУА-сеть имеет переменную структуру, которая может разрастаться в процессе обучения. В известном подходе [12] в качестве узла МГУА-сети используется нелинейная адалина, но в условиях высокого уровня помех с выбросами негауссовой природы такая система не будет справляться с решением задачи прогнозирования или идентификации нелинейного динамического нестационарного процесса.

В данной статье предлагается гибридная функционально связанная МГУА-вэйвлет-нейро-фаззи система, в качестве узла которой используется W -нейрон с робастным алгоритмом обучения [11].

Процесс обучения гибридной МГУА-сети состоит в настройке, начиная со входного слоя, всех параметров каждого узла и увеличении числа слоев для достижения требуемой точности [12]. Заметим сразу, что количество нейронов первого слоя определяется размерностью входного вектора и не превышает значения C_n^k , где n - размерность входного вектора в функционально связанной МГУА-вэйвлет-нейро-фаззи системе, k - размерность входного вектора узла (в стандартном случае $k = 2$, но в данном случае количество входов узла может быть расширено без усложнения структуры узла). Зачастую количество узлов второго слоя выбирается так же, как и в первом. После предъявления сети всей обучающей выборки оценивается точность каждого узла и формируется группа нейронов, дающих ошибку ниже минимальной ошибки предыдущего слоя. Именно выходы этой группы узлов являются входами следующего скрытого слоя. Процесс наращивания слоев и настройки параметров узлов продолжается

до тех пор, пока не будет получена желаемая точность, либо до момента возрастания ошибки на проверочной выборке, что является индикатором переобучения. На рис. 1 представлена архитектура функционально связанной МГУА-вэйвлет-нейро-фаззи системы.

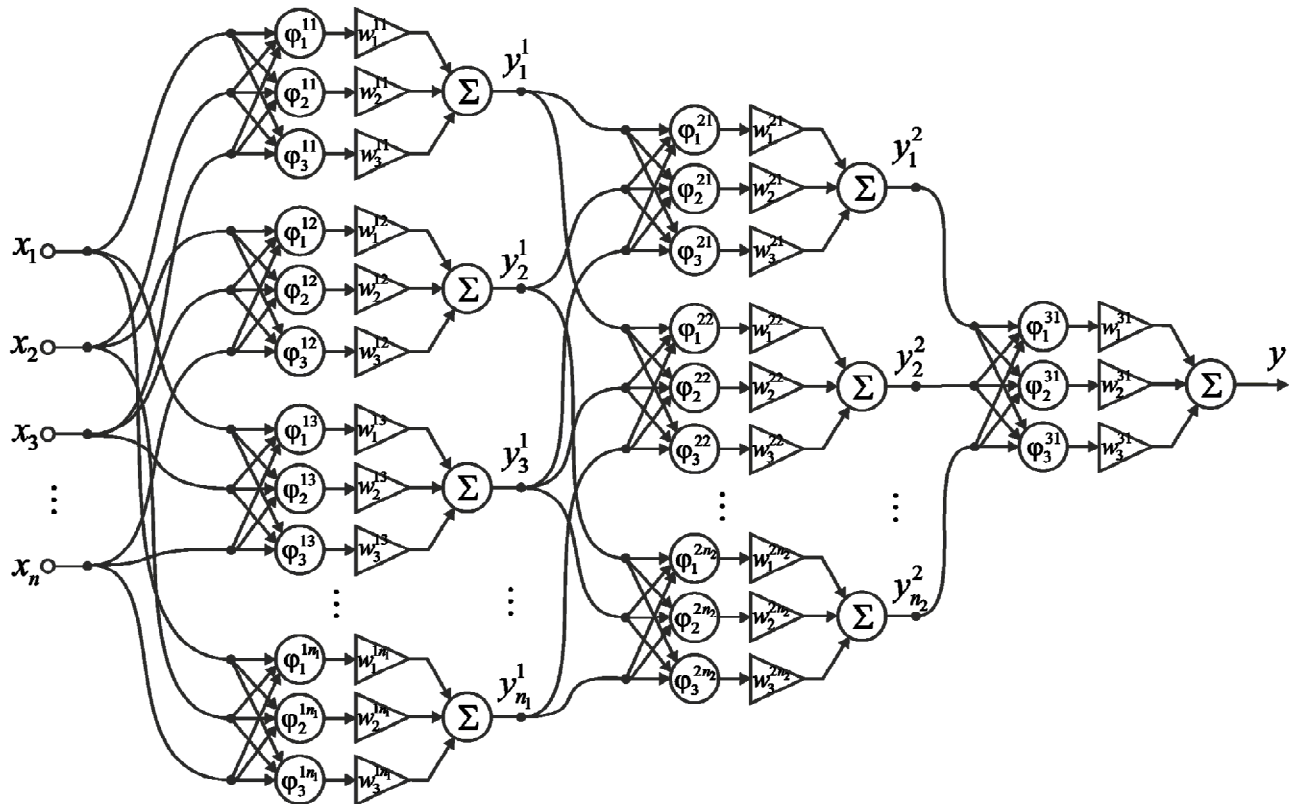


Рис. 1. Архитектура ФС МГУА-ВНФС

Далее рассмотрим архитектуру узла сети и алгоритм обучения всех его параметров. Рассмотрим структуру W -нейрона с робастным блоком обучения, приведенную на рис. 2 и совпадающую, по сути, с традиционной радиально-базисной нейронной сетью.

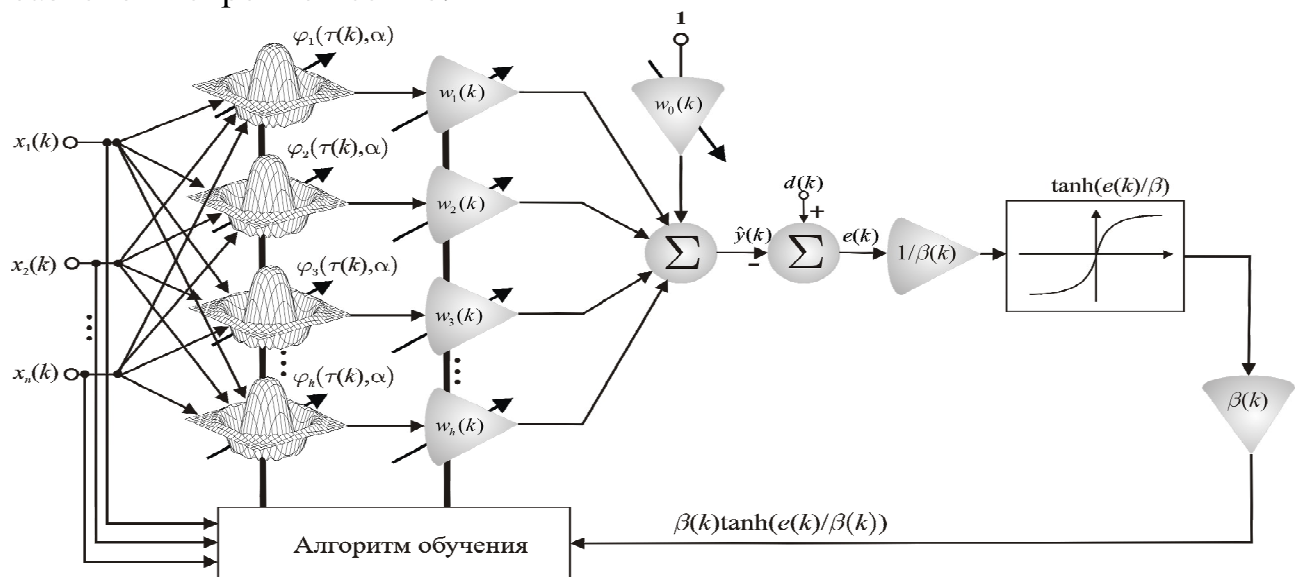


Рис. 2. Архитектура W -нейрона с робастным алгоритмом обучения

Нулевой слой архитектуры является рецепторным и в текущий момент времени k на него подается входной сигнал в форме вектора $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$. Скрытый слой в отличие от радиально-базисных сетей образован не R -нейронами, а многомерными вэйвлет-активационными функциями вида

$$\varphi_j(x(k)) = \varphi_j\left(\left((x(k) - c_j)^T Q_j^{-1}(x(k) - c_j)\right), \alpha_j\right), \quad j = 1, 2, \dots, h, \quad (1)$$

где Q_j - некоторая положительно определенная матрица (метрика Итакуры-Саито [13]).

Это приводит к тому, что рецепторные поля – гиперэллипсоиды могут иметь произвольную ориентацию относительно координатных осей пространства X , что расширяет функциональные возможности W -нейронов.

Основываясь на результатах [14, 15], о том, что вэйвлет-функции могут быть использованы как функции принадлежности, введем адаптивную функцию принадлежности-активации вида [16]

$$\varphi_j(\tau_j(k)) = (1 - \alpha_j \tau_j^2(x(k))) \exp\left(-\tau_j^2(x(k))/2\right), \quad (2)$$

где $\tau_j(x(k)) = \left((x(k) - c_j(k))^T Q_j^{-1}(k)(x(k) - c_j(k))\right)$, α_j настраиваемый параметр формы вэйвлет-функции ($0 \leq \alpha \leq 1$).

Уточняемый параметр α_j позволяет перестраивать форму активационной функции в процессе обучения W -нейрона, при этом при $\alpha = 0$ получаем Гауссову функцию активации, при $\alpha = 1$ получаем вэйвлет-функцию «Mexican Hat», а при $0 < \alpha < 1$ - гибридную функцию активации.

Второй слой представляет собой обычный адаптивный линейный ассоциатор с настраиваемыми синаптическими весами w_j

$$\begin{aligned} \mathfrak{F}(k) &= w_0(k) + \sum_{j=1}^h w_j(k) \varphi\left(\left((x(k) - c_j)^T Q_j^{-1}(x(k) - c_j)\right), \alpha_j(k)\right) = \\ &= w(k)^T \varphi(\tau(k), \alpha), \end{aligned} \quad (3)$$

где $\varphi_0(\tau(k)) \equiv 1$, $\varphi(\tau(k), \alpha) = (1, \varphi_1(\tau_1(k), \alpha_1), \varphi_2(\tau_2(k), \alpha_2), \dots, \varphi_h(\tau_h(k), \alpha_h))^T$, $w(k) = (w_0(k), w_1(k), w_2(k), \dots, w_h(k))^T$, $\tau(k) = (x(k) - c_j)^T Q_j^{-1}(x(k) - c_j)$.

И, наконец, третий слой представляет собой робастный блок обучения, состоящий из функции гиперболического тангенса и усилителей.

2. Робастный алгоритм обучения узла ФС МГУА-ВНФС

Опыт показывает, что методы обучения и идентификации, основанные на критерии наименьших квадратов, оказываются чрезвычайно чувствительными

к отклонениям фактического закона распределения данных от нормального. В условиях различного рода выбросов, грубых ошибок, негауссовских возмущений с «тяжелыми хвостами» алгоритмы, связанные с критерием наименьших квадратов, теряют свою эффективность.

В этой ситуации на первый план выходят методы робастного оценивания [17], которые к настоящему времени получили известное распространение и для обучения искусственных нейронных сетей [18, 19].

Вводя в рассмотрение ошибку обучения

$$e(k) = y(k) - \hat{y}(k) = y(k) - w^T(k)\varphi(k) \quad (4)$$

и робастный критерий идентификации [18, 19]

$$E(k) = f(k) = \beta^2 \ln(\cosh(e(k)/\beta)), \quad (5)$$

где β - положительный параметр, выбираемый из эмпирических соображений и определяющий размер зоны нечувствительности к выбросам. Необходимо отметить, что робастный критерий (5) удовлетворяет всем аксиомам метрического пространства.

Далее рассмотрим процесс синтеза алгоритма обучения. Для настройки синаптических весов и параметров вэйвлонов (векторов c_j и матриц Q_j^{-1}) будем использовать градиентную процедуру минимизации критерия (5), при этом в отличие от покомпонентного обучения, будем производить уточнение в векторно-матричной форме, что, во-первых, проще с вычислительной точки зрения, а, во-вторых, позволяет оптимизировать процесс обучения по быстродействию.

В общем случае алгоритм обучения может быть записан в виде

$$\begin{cases} w(k+1) = w(k) - \eta_w \nabla_w E(k), \\ c_j(k+1) = c_j(k) - \eta_{c_j} \nabla_{c_j} E(k), \quad j = 1, 2, \dots, h, \\ Q_j^{-1}(k+1) = Q_j^{-1}(k) - \eta_{Q_j^{-1}} \left\{ \partial E(k) / \partial Q_j^{-1} \right\}, \quad j = 1, 2, \dots, h, \\ \alpha(k+1) = \alpha(k) - \eta_\alpha \nabla_\alpha E(k), \end{cases} \quad (6)$$

где $\nabla_w E$ - вектор-градиент критерия (5) по w , $\nabla_{c_j} E$ - $(n \times 1)$ -вектор-градиент критерия (5) по c_j ; $\nabla_\alpha E$ - вектор-градиент критерия (5) по α ; $\left\{ \partial E(k) / \partial Q_j^{-1} \right\}$ - $(n \times n)$ -матрица, образованная частными производными E по компонентам Q_j^{-1} ; $\eta_w, \eta_{c_j}, \eta_{Q_j^{-1}}, \eta_\alpha$ - параметры шага алгоритма обучения.

Очевидно, что для адаптивного многомерного вэйвлета (2) можно записать

$$\left\{ \begin{aligned}
 \nabla_w E(k) &= -\beta \tanh(e(k)/\beta)(1 - \alpha_j \tau_j^2(x(k))) \exp(-\tau_j^2(x(k))/2) = \\
 &= -\beta \tanh(e(k)/\beta) J_w(k), \\
 \nabla_{c_j} E(k) &= \beta \tanh(e(k)/\beta) w_j(k) (\alpha \tau_j^3(x(k)) - (2\alpha + 1) \tau_j(x(k))) \cdot \\
 &\cdot \exp(-\tau_j^2(x(k))/2) Q_j^{-1}(k) (x(k) - c_j(k)) = \beta \tanh(e(k)/\beta) J_{c_j}(k), \\
 \{\partial E(k)/\partial Q_j^{-1}\} &= -\beta \tanh(e(k)/\beta) w_j(k) (\alpha \tau_j^3(x(k)) - (2\alpha + 1) \tau_j(x(k))) \cdot \\
 &\cdot \exp(-\tau_j^2(x(k))/2) (x(k) - c_j(k)) (x(k) - c_j(k))^T = -\beta \tanh(e(k)/\beta) J_{Q_j^{-1}}(k), \\
 \nabla_\alpha E(k) &= -\beta \tanh(e(k)/\beta) w(k) \tau^2(x(k)) \exp(-\tau_j^2(x(k))/2) = \\
 &= \beta \tanh(e(k)/\beta) J_\alpha(k).
 \end{aligned} \right. \quad (7)$$

Заметим также, что скорость сходимости к оптимальным значениям w , c_j , Q_j^{-1} и α полностью определяется параметрами шага η_w , η_{c_j} , $\eta_{Q_j^{-1}}$ и η_α . Повышение скорости сходимости может быть достигнуто путем использования более сложных, нежели градиентные, процедур типа Хартли или Марквардта. Воспользовавшись леммой обращения матриц и используя подход, предложенный в [20], можно ввести следующую модифицированную процедуру обучения:

$$\left\{ \begin{aligned}
 w(k+1) &= w(k) + \lambda_w (\beta \tanh(e(k)/\beta) J_w(k) / \eta_w(k)), \\
 \eta_w(k+1) &= \gamma \eta_w(k) + \|J_w(k+1)\|^2, \\
 c_j(k+1) &= c_j(k) - \lambda_{c_j} (\beta \tanh(e(k)/\beta) J_{c_j}(k) / \eta_{c_j}(k)), \\
 \eta_{c_j}(k+1) &= \alpha_c \eta_{c_j}(k) + \|J_{c_j}(k+1)\|^2, \\
 Q_j^{-1}(k+1) &= Q_j^{-1}(k) + \lambda_{Q_j^{-1}} (\beta \tanh(e(k)/\beta) J_{Q_j^{-1}}(k) / \eta_{Q_j^{-1}}(k)), \\
 \eta_{Q_j^{-1}}(k+1) &= \gamma \eta_{Q_j^{-1}}(k) + Tr(J_{Q_j^{-1}}^T(k+1) J_{Q_j^{-1}}(k+1)), \\
 \alpha(k+1) &= \alpha(k) + \lambda_\alpha (\beta \tanh(e(k)/\beta) J_\alpha(k) / \eta_\alpha(k)), \\
 \eta_\alpha(k+1) &= \gamma \eta_\alpha(k) + \|J_\alpha(k+1)\|^2,
 \end{aligned} \right. \quad (8)$$

(здесь $0 \leq \alpha_w \leq 1, 0 \leq \alpha_{c_j} \leq 1, 0 \leq \alpha_{Q_j^{-1}} \leq 1, 0 \leq \alpha_\alpha \leq 1$ - параметры взвешивания устаревшей информации, $\lambda_{\{\bullet\}}$ - положительный демпфирующий параметр), являющуюся нелинейным гибридом алгоритмов Качмажа-Уидроу-Хоффа и Гудвина-Рэмеджа-Кэйнеса и обладающую как следящими, так и фильтрующими свойствами.

3. Экспериментальное моделирование

Экспериментальное исследование предложенной функционально связанной МГУА-вэйвлет-нейро-фаззи системы и робастного алгоритма обучения проводилось на основе сигнала, зашумленного интенсивными выбросами. Сигнал был получен на основе нелинейного динамического объекта Нарендры-Партхасаратхи [21] (стандартный тест, широко распространенный для сравнения нейро- и нейро-фаззи систем при моделировании и прогнозировании временных рядов), чей выходной сигнал был искусственно загрязнен случайным шумом с распределением Коши (распределение с тяжелыми хвостами) вида

$$F_X^{-1}(x) = x_0 + \gamma_{ch} \operatorname{tg} [\pi(x - 0.5)], \quad (9)$$

где x_0 параметр локализации, γ_{ch} параметр масштаба ($\gamma_{ch} > 0$), x носитель ($x \in (-\infty, +\infty)$).

Нелинейный динамический объект был сгенерирован согласно уравнению

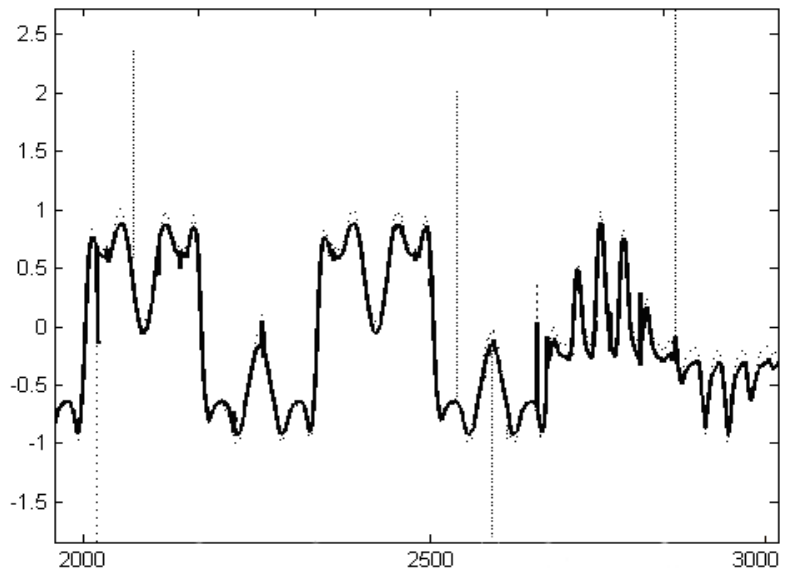
$$y(k+1) = 0.3y(k) + 0.6y(k-1) + f(u(k)), \quad (10)$$

где $f(u(k)) = 0.6\sin(u(k)) + 0.3\sin(3u(k)) + 0.1\sin(5u(k))$ и $u(k) = \sin(2k/250)$, k дискретное время. Значения $x(k-3), x(k-2), x(k-1), x(k)$ были использованы для эмуляции $x(k+1)$. Таким образом, начальное значение числа входов было равно $n = 4$, а количество узлов первого слоя функционально связанной МГУА-вэйвлет-нейро-фаззи системы было равно $C_4^2 = 6$. Параметры алгоритма обучения были приняты $\beta_w = 0.5$, $\beta_c = 0.5$, $\beta_Q = 0.5$, $\beta_\alpha = 0.5$, $\gamma_w = \gamma_c = \gamma_Q = \gamma_\alpha = 0.99$, $\lambda_w = \lambda_{c_j} = \lambda_{Q_j^{-1}} = \lambda_\alpha = 0.99$. Начальные значения равны $\eta_w(0) = \eta_{c_j}(0) = \eta_{Q_j^{-1}}(0) = \eta_\alpha(0) = 10000$.

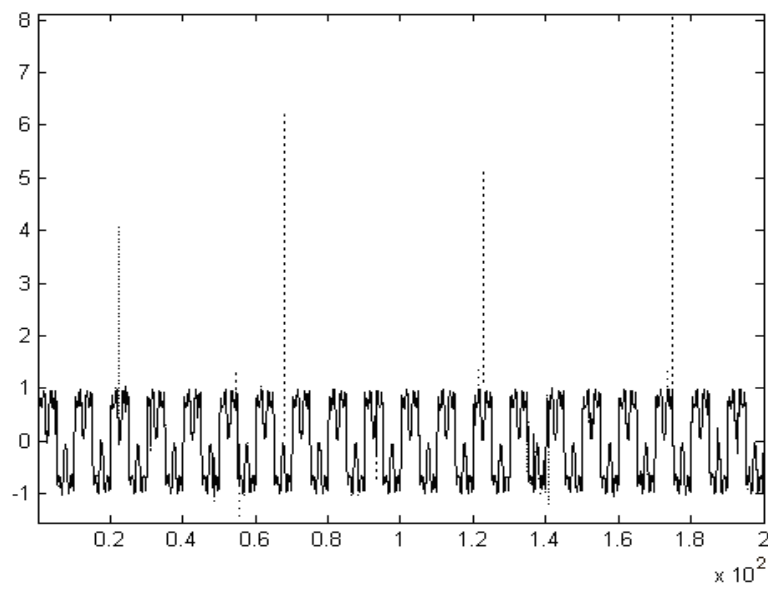
Обучающая выборка содержит 2000 точек, при этом каждый узел сети был обучен за пять эпох по 2000 итераций в каждой. Далее следующие 1000 точек были использованы как тестовая выборка. Тестовая выборка включает в себя $f(u(k)) = 0.6\sin(\pi u(k)) + 0.3\sin(3\pi u(k)) + 0.1\sin(5\pi u(k))$ и $u(k) = \sin(2\pi k/250)$ для $2001 < k < 2500$ и $f(u(k)) = u^3(k) + 0.3u^2(k) - 0.4u(k)$ и $u(k) = \sin(2\pi k/250) + \sin(2\pi k/25)$ для $2501 < k < 3000$. Начальные значения синаптических весов были случайно сгенерированы в интервале от -0.1 до $+0.1$.

На рис. 3 а представлено результаты прогнозирования зашумленного нестационарного динамического сигнала, а на рис. 3 б показано сегмент процесса обучения; можно видеть, что выброс с большой амплитудой находящийся в начале выборки, не повлиял на процесс обучения.

Сравнение результатов прогнозирования на основе робастного алгоритма обучения проводилось с результатами прогнозирования на основе такой же сети с робастным алгоритмом обучения только синаптических весов, с градиентным алгоритмом, с алгоритмом на основе рекуррентного метода наименьших квадратов, а также со стандартной МГУА-нейронной сетью с градиентным алгоритмом обучения.



а)



б)

Рис. 3. Результаты прогнозирования

Наилучший результат прогнозирования обеспечивает предложенная система с робастным алгоритмом обучения всех параметров узла. При обучении только синаптических весов с помощью робастного алгоритма ошибка прогнозирования увеличивается. При обучении градиентным алгоритмом первый же выброс в начале выборки сильно влияет на процесс обучения и как результат -

большая ошибка прогнозирования. При обучении системы рекуррентным методом наименьших квадратов при первом же выбросе происходит, так называемый, «взрыв параметров» ковариационной матрицы и как результат - невозможность прогнозирования сигналов, зашумленных аномальными выбросами. Результаты сравнения приведены в табл. 1.

Таблица 1

Результаты прогнозирования

Нейронная сеть / Алгоритм обучения	Кол. слоев сети / Кол. входов в узле	RMSE
ФС МГУА-ВНФС / Предложенный робастный алгоритм (настраиваются все параметры) (8)	3/3	0.0493
ФС МГУА-ВНФС / Предложенный робастный алгоритм (настраиваются все параметры) (8)	3/2	0.0557
ФС МГУА-ВНФС / Предложенный робастный алгоритм (настраиваются только синаптические веса) (8)	4/2	0.0899
ФС МГУА-ВНФС / Градиентный алгоритм (6)	4/2	0.1398
ФС МГУА-ВНФС / Метод наименьших квадратов	4/2	∞
МГУА-нейронная сеть / Градиентный алгоритм	4/2	1.4436

Таким образом, видно, что предложенный робастный алгоритм обучения позволяет обрабатывать сигналы в условиях существенного загрязнения аномальными выбросами.

Выводы

В статье предложен численно простой и эффективный робастный алгоритм обучения всех параметров узла функционально связанной МГУА-взвешенной нейро-фаззи системы, обладающий следящими и сглаживающими свойствами и позволяющий в реальном времени обрабатывать нестационарные нелинейные сигналы в условиях различного рода выбросов и возмущений с «тяжелыми хвостами». Настройка рецепторных полей W-нейронов, включающая в себя преобразования активационных функций (сдвиг, растяжение, вращение, форма), позволяет улучшить прогнозирующие свойства сети, что подтверждается результатами экспериментального исследования.

Литература

1. Pao Y.-H. Adaptive Pattern Recognition and Neural Networks. - Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1989. – 309 p.
2. Park J., Sandberg I.W. Universal approximation using radial-basis-function networks // Neural Computation. – 3. – 1991. – P. 246-257.
3. Poggio T., Girosi F. A Theory of Networks for Approximation and Learning // A. I. Memo № 1140, C.B.I.P. Paper № 31. - Massachusetts Institute of Technology, 1994. - 63 p.

4. Bishop C. M. *Neural Networks for Pattern Recognition*. - Oxford: Clarendon Press, 1995. - 482 p.
5. Chui C. K. *An Introduction to Wavelets*. - New York: Academic, 1992. - 264 p.
6. Billings S. A., Wei H.-L. A new class of wavelet networks for nonlinear system identification // *IEEE Trans. on Neural Networks*. - 16 (4). - 2005. - P. 862-874.
7. Zhang Q. H. Using wavelet network in nonparametric estimation // *IEEE Trans. on Neural Networks*. - 8(2). - 1997. - P. 227-236.
8. Ивахненко А.Г. Долгосрочное прогнозирование и управление сложными системами. - Киев: "Техніка", 1975. - 311 с.
9. Ивахненко А.Г. Самообучающиеся системы распознавания и автоматического управления. – Киев: Техніка, 1969 – 392 с.
10. Ивахненко А.Г., Степашко В.С. Помехоустойчивость моделирования. – К.: Наукова думка, 1984. – 295 с.
11. Bodyanskiy Ye., Pavlov O., Vynokurova O. Outliers resistant learning algorithm for radial-basis-fuzzy-wavelet-neural network in stomach acute injury diagnosis tasks // Eds. by K. Markov, K. Ivanova, I. Mitov *Int. Book Series "Information Science and Computing"*, Number 2. – Sofia: Institute of Information Theories and Application, 2008. – P.55-62.
12. Бодянский Е.В., Руденко О.Г. Искусственные нейронные сети: архитектуры, обучение, применения. – Телетех: Харьков. – 2004. – 372 с.
13. Itakura F. Maximum prediction residual principle applied to speech recognition // *IEEE Trans. on Acoustics, Speech, and Signal Processing*. – 1975. – 23. – P. 67-72.
14. Mitaim S., Kosko B. What is the best shape for a fuzzy set in function approximation? // *Proc. 5th IEEE Int. Conf on Fuzzy Systems "Fuzz-96"*. - V.2. – 1996. – P. 1237-1213.
15. Mitaim S., Kosko B. Adaptive joint fuzzy sets for function approximation // *Proc. Int. Conf. on Neural Networks "ICNN-97"*. – 1997. – P. 537-542.
16. Bodyanskiy Ye., Yegorova E., Vynokurova O. Radial-basis-fuzzy-wavelet-neural network with adaptive activation-membership function // *Int. J. on Artificial Intelligence and Machine Learning*. – 2008. – V.8. – II. – P. 9-15
17. Rey W.J.J. *Robust Statistical Methods*. - Lecture Notes in Mathematics. - Berlin-Heidelberg-New York: Springer-Verlag. - V. 690. - 1978. – 128 p.
18. Cichocki A., Unbehauen R. *Neural Networks for Optimization and Signal Processing*. - Stuttgart: Teubner, 1993. – 544 p.
19. Holland P.W., Welsh R.E. Robust regression using iteratively reweighted least squares // *Commun. Statist., Theory and Methods*. – 46. – 1977. – P.813-828.
20. Bodyanskiy Ye., Kolodyazhniy V., Stephan A. An adaptive learning algorithm for a neuro-fuzzy network // *Computational Intelligence. Theory and Applications*, Ed. by B. Reusch. - Berlin-Heidelberg-New York: Springer. – 2001. - P. 68-75.
21. Narendra K.S., Parthasarathy K. Identification and control of dynamic systems using neural networks // *IEEE Trans. on Neural Networks*. – 1990. – 1. – P. 4-26.