

ДВОЙНОЙ ВЭЙВЛЕТ–НЕЙРОН: ТРЕУГОЛЬНЫЕ АКТИВАЦИОННЫЕ ФУНКЦИИ, АРХИТЕКТУРА, ОБУЧЕНИЕ

В статье предложена архитектура двойного вэйвлет–нейрона на основе треугольных вэйвлетов переменной формы и его алгоритм обучения. Предложенная архитектура позволяет улучшить аппроксимирующие свойства вэйвлет–нейрона. Архитектура двойного вэйвлет–нейрона может являться строительным блоком вэйвлет–нейронной сети. Проведено имитационное моделирование путем прогнозирования нестационарного временного ряда.

Введение

Задачи прогнозирования и эмуляции нестационарных временных процессов произвольной природы в условиях неопределенности, поступающих на обработку в реальном времени, достаточно часто встречаются во многих приложениях, связанных с интеллектуальным анализом данных. Интеллектуальный анализ данных подразумевает использование методов вычислительного интеллекта среди которых гибридные нейронные сети нашли широкое применение.

Для решения такого рода задач используется большое количество архитектур нейронных сетей, в том числе и гибридных структур, однако эти системы либо громоздки по своей архитектуре (например, многослойный персептрон), либо плохо приспособлены для обучения в реальном времени. В большинстве случаев активационными функциями таких сетей являются сплайны, полиномы и радиально–базисные функции, которые не обладают способностями выявлять локальные особенности сигналов, а также сложны для аппаратной реализации.

На сегодня широкое распространение получила теория вэйвлет–анализа [1–3], которая позволяет с высокой точностью выявлять локальные особенности нестационарных сигналов. На стыке этих двух подходов и возникли, так называемые, гибридные вэйвлет–нейронные сети [4–18] благодаря своим высоким аппроксимирующим свойствам и чувствительностью к изменениям характеристик анализируемых процессов.

Весьма привлекательной с точки зрения технической реализации, обеспечиваемой точности и простоты обучения является, так называемый, вэйвлет–нейрон, предложенный и описанный в [19–21]. При этом вэйвлет–функции реализованы либо на уровне синаптических весов, либо на выходе нейрона, а для обучения используется градиентный алгоритм обучения с постоянным шагом. Для улучшения аппроксимирующих свойств и ускорения процесса обучения в настоящей работе пре-

длагается использовать конструкцию двойного вэйвлет–нейрона [22] с треугольными функциями активации [8] и алгоритм обучения, обладающий как сглаживающими, так и следящими свойствами.

Используемые аналитические вэйвлет–активационные функции

В качестве активационной функции двойного вэйвлет–нейрона можно использовать различные виды аналитических вэйвлетов. Мы предлагаем использовать треугольный вэйвлет [8], который обладает большим спектром полезных свойств.

Для того чтобы выполнялось стандартное условие (основное условие существования вэйвлетов)

$$\int_{-\infty}^{+\infty} \varphi(x) dx = 0,$$

необходимо учесть ограничение вида

$$h_2 = \frac{(b - d) + h_1(c - a)}{(c - e)}.$$

Тогда треугольный вэйвлет, представленный на рис. 1, может быть описан выражением

$$\varphi(x, [a, b, c, d, e, h_1]) = \begin{cases} 0, & \text{если } x < a \text{ и } x > e, \\ -h_1 \frac{(x - a)}{(b - a)}, & \text{если } a \leq x \leq b, \\ \frac{(h_1 + 1)(x - b)}{c - b} - h_1, & \text{если } b \leq x \leq c, \\ -\frac{(h_2 + 1)(x - c)}{(d - c)} + 1, & \text{если } c \leq x \leq d, \\ h_2 \frac{(x - d)}{(e - d)} - h_2, & \text{если } d \leq x \leq e. \end{cases} \quad (1)$$

Отличительной особенностью предлагаемого треугольного вэйвлета является то, что он может быть как четной так и нечетной функцией в зависимости от значения параметров. Такая конструкция достаточно легко может быть реализована аппаратно.

Для использования таких активационных функций в двойном вэйвлет–нейроне, необходимо построить решетку на основе треугольного вэйвлета для каждого вэйвлет–синапса. На рис. 2 представлены возможные варианты задания решеток вэйвлет–активационных функций.

Синтез архитектуры двойного вэйвлет–нейрона

Введем в рассмотрение структуру двойного вэйвлет–нейрона, приведенную на рис. 3. Как видно, двойной вэйвлет–нейрон достаточно близок по конструкции к n -входовому вэйвлет нейрону [19–21], однако содержит нелинейные вэйвлет–функции как на уровне синаптических весов, так и на выходе структуры.

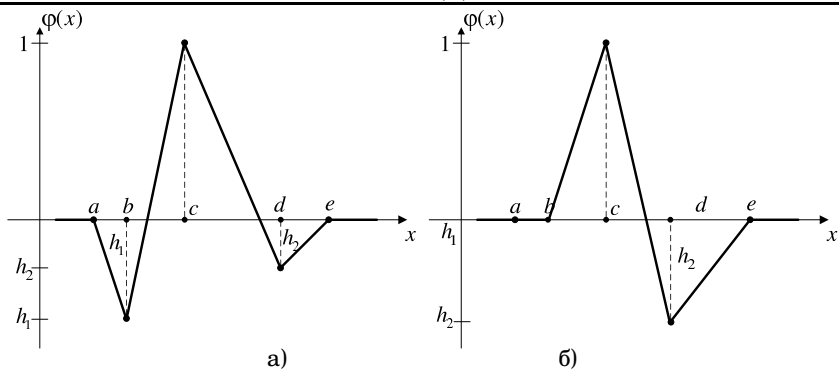


Рис. 1 – Треугольный вэйвлет: а – четный; б – нечетный;

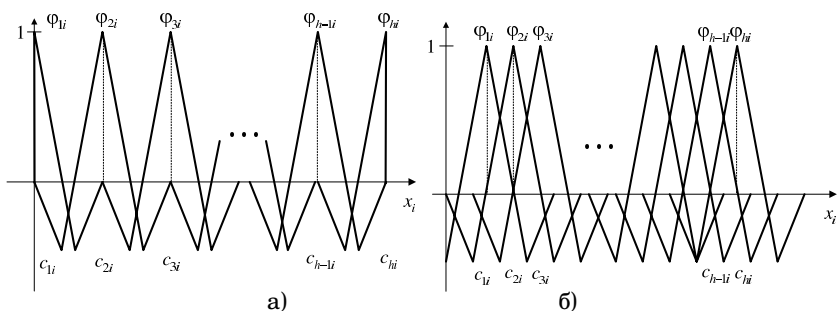


Рис. 2 – Решетки на основе треугольных вэйвлетов

При подаче на вход двойного вэйвлет–нейрона, приведенного на рис. 4, векторного сигнала $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$ (здесь $k = 0, 1, 2, \dots$ – номер наблюдения в обучающей выборке или текущее дискретное время) на его выходе появляется скалярный сигнал вида

$$y(k) = f_0 \left(\sum_{i=1}^n f_i(x_i(x)) \right) = f_0(u(k)) = \sum_{l=0}^{h_2} \left(\sum_{i=1}^n \sum_{j=0}^{h_1} \varphi_{ji}(x_i(x)) w_{ji}(k) \right) w_{j0}, \quad (2)$$

определяемый как настраиваемыми синаптическими весами $w_{ji}(k), w_{l0}(k)$, так и значениями используемых вэйвлет–функций $\varphi_{ji}(x_i(k)), \varphi_{l0}(u(k))$, при этом полагается $\varphi_{00}(\bullet) = \varphi_{0i}(\bullet) \equiv 1$.

Двойной вэйвлет–нейрон состоит из двух слоев: скрытого слоя, в котором n вэйвлет–синапсов по h_1 вэйвлет–функций в каждом и выходного слоя, состоящего из одного вэйвлет–синапса с h_2 вэйвлет–функциями.

В каждом вэйвлет–синапсе реализованы вэйвлеты, отличающиеся между собой параметрами растяжения и смещения.

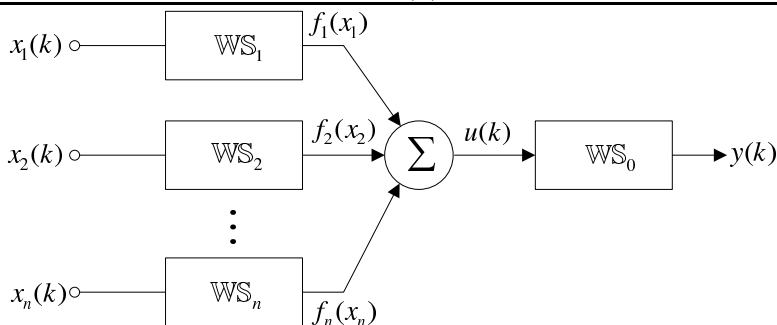


Рис. 3 – Обобщенная структура двойного вэйвлет-нейрона

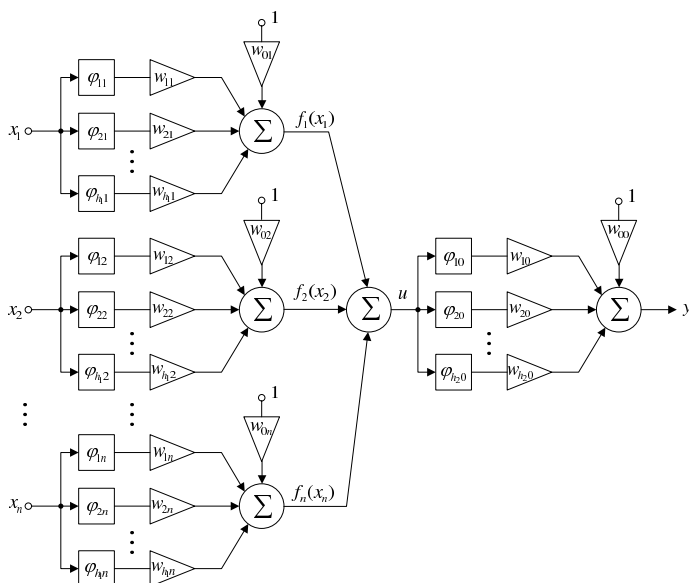


Рис. 4 – Архитектура двойного вэйвлет-нейрона с нелинейными вэйвлет-синапсами

Синтез алгоритма обучения двойного вэйвлет-нейрона

Для обучения выходного слоя двойного вэйвлет-нейрона будем использовать критерий вида

$$E(k) = \frac{1}{2}(d(k) - y(k))^2 = \frac{1}{2}e^2(k), \quad (3)$$

где $d(k)$ – внешний обучающий сигнал.

Алгоритм обучения выходного слоя двойного вэйвлет-нейрона на

основе градиентной процедуры имеет вид

$$\begin{cases} w_{j0}(k+1) = w_{j0}(k) + \eta_0(k)e(k)\varphi_{j0}(u(k)), & j - \text{активированный нейрон,} \\ w_{j0}(k+1) = w_{j0}(k) & \text{в остальных случаях,} \end{cases} \quad (4)$$

где $w_0(k) = (w_{00}(k), w_{10}(k), w_{20}(k), \dots, w_{h_20}(k))^T$ – вектор синаптических весов, $\varphi_0(u(k)) = (1, \varphi_{10}(u(k)), \varphi_{20}(u(k)), \dots, \varphi_{h_20}(u(k)))^T$ – вектор вэйвлет-активационных функций, $e(k)$ – ошибка обучения, $\eta_0(k)$ – шаг обучения, подлежащий определению.

Обучение скрытого слоя проводится аналогичным образом на основе обратного распространения ошибки с использованием того же критерия, записанного в форме

$$E(k) = \frac{1}{2}(d(k) - f_0(u(k)))^2 = \frac{1}{2}\left(d(k) - f_0\left(\sum_{i=1}^n \sum_{j=0}^{h_1} \varphi_{ji}(x_i(x))w_{ji}(k)\right)\right)^2, \quad (5)$$

Алгоритм обучения скрытого слоя двойного вэйвлет-нейрона на основе градиентной оптимизации имеет вид

$$\begin{cases} w_{ji}(k+1) = w_{ji}(k) + \eta(k)e(k)f'_0(u(k))\varphi_{ji}(x_i(k)), & j - \text{активированный нейрон,} \\ w_{ji}(k+1) = w_{ji}(k) & \text{в остальных случаях,} \end{cases} \quad (6)$$

где $w_i(k) = (w_{0i}(k), w_{1i}(k), w_{2i}(k), \dots, w_{h_{1i}}(k))^T$ – вектор синаптических весов, $\varphi_i(u(k)) = (1, \varphi_{1i}(u(k)), \varphi_{2i}(u(k)), \dots, \varphi_{h_{1i}}(u(k)))^T$ – вектор вэйвлет-активационных функций, $\eta(k)$ – шаг обучения, подлежащий определению и $f'_0(u(k))$ – производная функции $f_0(u(k))$.

Выводы

В работе предложена структура двойного вэйвлет-нейрона и алгоритм его обучения на основе треугольных вэйвлетов. Двойной вэйвлет-нейрон может использоваться как самостоятельно, так и являться строительным блоком вэйвлет-нейронной сети. Алгоритм прост в численной реализации и обладает высокой скоростью сходимости. Целесообразно использование предложенного алгоритма для решения задач прогнозирования, эмуляции, классификации временных рядов различной природы, а также в других приложениях, связанных с интеллектуальной обработкой данных в реальном времени. Также одним из преимуществ предлагаемого подхода является возможность реализации треугольной вэйвлет-активационной функции в аппаратном виде. Имитационное моделирование подтверждает эффективность развиваемого подхода.

Литература

1. Chui С.К. An Introduction to Wavelets. — New York: Academic, 1992. — 264 p.
2. Daubechies I. Ten Lectures on Wavelets. — Philadelphia, PA: SIAM, 1992. — 228 p.

3. Meyer Y. Wavelets: Algorithms and Applications. — Philadelphia, PA: SIAM, 1993. — 133 p.
4. Lekutai G., van Landingham H.F. Self-tuning control of nonlinear systems using neural network adaptive frame wavelets // Proc. IEEE Int. Conf. on Neural Networks. — Piscataway, N.J., 1997. — 2. — P 1017–1022.
5. Bodyanskiy Ye., Lamonova N., Pliss I., Vynokurova O. An adaptive learning algorithm for a wavelet neural network // Expert Systems. — 2005. — № 22. — P 235–240.
6. Bodyanskiy Ye., Kolodyazhnyi V., Pliss I., Vynokurova O. Learning wavelet neuron based on the RASP-function // Radio Electronics. Computer Science. Control. — 2004. — № 1. — P 118–122.
7. Бодянский Е.В., Винокурова Е.А., Ламонова Н.С. Адаптивная гибридная вэйвлет-нейронная сеть для решения задачи прогнозирования и эмуляции // Сб. науч. трудов 12-й международной конференции по автоматическому управлению "Автоматика 2005". — Харьков: Изд-во НТУ "ХПИ", 2005. — 3. — С. 40–41.
8. Бодянский Е.В., Винокурова Е.А. Треугольный вэйвлет и формальный нейрон на его основе // Сб. наук. праць 3-ої Міжнародної науково-практичної конференції "Математичне та програмне забезпечення інтелектуальних систем"(MPZIS- 2005). — Дніпропетровськ: ДНУ, 2005. — 3. — С. 14–15.
9. Billings S.A., Wei H.-L. A new class of wavelet networks for nonlinear system identification // IEEE Trans. on Neural networks. — 2005. — 16. — № 4. — P 862–874.
10. Szu H.H., Telfer B., Kadambe S. Neural network adaptive wavelets for signal representation and classification // Opt. Eng. — 1992. — № 31. — P 1907–1916.
11. Zhang Q. H., Benveniste A. Wavelet networks // IEEE Trans. on Neural Networks. — 1992. — 3. — № 6. — P 889–8986.
12. Dickhaus H., Heinrich H. Classifying biosignals with wavelet networks // IEEE Eng. Med. Biol. Mag. — 1996. — 15. — № 5. — P 103–111.
13. Cao L.Y., Hong Y.G., Fang H.P., He G.W. Predicting chaotic time series with wavelet networks // Phys. D. — 1995. — № 85. — P 225–238.
14. Oussar Y., Dreyfus G. Initialization by selection for wavelet network training // Neurocomputing. — 2000. — № 34. — P 131–143.
15. Zhang J., Walter G.G., Miao Y., Lee W.N. Wavelet neural networks for function learning // IEEE Trans. on Signal Processing. — 1995. — 43. — № 6. — P 1485–1497.
16. Zhang Q.H. Using wavelet network in nonparametric estimation // IEEE Trans. on Neural Networks. — 1997. — 8. — № 2. — P 227–236.
17. Casdagli M. Nonlinear prediction of chaotic time series // Phys. D. — 1989. — № 35. — P 335–356.

18. Soltani S. On the use of wavelet decomposition for time series prediction // *Neurocomputing*. – 2002. – № 48. – P. 267–277.
19. Yamakawa T., Uchino E., Samatu T. Wavelet neural networks employing over-complete number of compactly supported non-orthogonal wavelets and their applications // *IEEE Int. Conf. on Neural Networks*. – Orlando, USA, 1994. – P. 1391–1396.
20. Yamakawa T., Uchino E., Samatu T. The wavelet network using convex wavelets and its application to modeling dynamical systems // *The Trans. on the IEICE*. – 1996. – № 12. – P. 2046–2053.
21. Yamakawa T. A novel nonlinear synapse neuron model guaranteeing a global minimum - Wavelet neuron // *Proc. 28 th IEEE Int. Symp. on Multiple-Valued Logic*. – Fukuoka, Japan: IEEE Comp. Soc., 1998. – P. 335–336.
22. Bodyanskiy Ye., Lamonova N., Pliss I., Vynokurova O. Recurrent learning algorithm for double-wavelet neuron // *Proc. XII -th Int. Conf. "Knowledge - Dialogue - Solution"*. – Varna (Bulgaria), 2006. – P. 77–84.