

УДК 004.032.26

А.А. Дейнеко, Ж.В. Дейнеко, А.П. Турута, Е.В. Бодянский
**АДАПТИВНЫЙ МЕТОД КОМБИНИРОВАННОГО
ОБУЧЕНИЯ-САМООБУЧЕНИЯ НЕЙРО-ФАЗЗИ СИСТЕМ**

Аннотация. Предложен подход к комбинированному обучению синаптических весов и центров функций принадлежности нейро-фаззи систем, основанный на использовании как правил обучения с учителем, так и самообучения.

Полученные результаты могут быть использованы при решении широкого класса задач, возникающих в Dynamic Data Mining.

Введение. В настоящее время искусственные нейронные сети (ANN) получили широкое распространение для решения различных задач Data Mining, интеллектуального управления, прогнозирования, идентификации и эмуляции и т. п. в условиях неопределенности, нелинейности, стохастичности и хаотичности, различного рода возмущений и помех, благодаря своим универсальным аппроксимирующим возможностям и способности обучаться на основании данных, характеризующих функционирование исследуемого явления или объекта. В случае необходимости обработки информации в последовательном on-line режиме на первый план выходит вопрос скорости сходимости процесса обучения, существенно ограничивающий класс ANN, пригодных для работы в этих условиях. С точки зрения оптимизации по скорости процесса обучения весьма эффективными являются ANN, использующие ядерные (радиально-базисные, колоколообразные, потенциальные) функции активации. Здесь наибольшее распространение получили радиально-базисные нейронные сети (RBFN), выходной сигнал которых линейно зависит от синаптических весов, что позволяет использовать для их обучения алгоритмы адаптивной идентификации такие, как рекуррентный метод наименьших квадратов, алгоритм Качмажа (Уидроу-Хоффа) и т.п.. Вместе с тем RBFN подвержены, так называемому, «проклятию размерности», порождаемому экспоненциальным ростом количества настраиваемых параметров (весов) с увеличением размерности пространства входов.

© Дейнеко А.А., Дейнеко Ж.В., Турута А.П., Бодянский Е.В., 2014

В сравнении с нейронными сетями большими возможностями обладают нейро-фаззи системы (NFS) [1-4], объединяющие в себе возможности обучения, аппроксимации и лингвистической интерпретации получаемых результатов. Наибольшее распространение среди NFS получили ANFIS и TSK-системы, выходной сигнал которых также зависит линейно от синаптических весов, что позволит использовать для их обучения множество оптимальных адаптивных алгоритмов линейной идентификации. Вместе с тем, для избежания возникновения «дыр» в пространстве входов, порождаемого рассеянным разбиением (scatter partitioning) [5], используемым в ANFIS и TSK-системах, производится настройка параметров функций принадлежности в первом скрытом слое NFS. Для этого используется алгоритм обратного распространения ошибок, реализуемый на множестве эпох [6] так, что говорить об обучении в реальном времени в данном случае не приходится.

Для управления параметрами активационных функций (центров и матриц рецепторных полей) RBFN в режиме реального времени в [7-9] было предложено использовать самоорганизующуюся карту Т. Кохонена [10], обеспечивающую в процессе своего самообучения в on-line режиме настройку этих параметров.

Представляется целесообразным распространить этот подход для адаптивной настройки параметров функций принадлежности нейро-фаззи систем.

1. Архитектура нейро-фаззи системы и обучение параметров выходного слоя. Архитектура рассматриваемой системы, приведенная на рис. 1, состоит из пяти последовательно соединенных слоев.

На входной (нулевой) слой NFS подается $(n \times 1)$ -мерный вектор входных сигналов-образов $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$ (здесь $k = 1, 2, \dots$ – текущее дискретное время), подлежащих обработке. Первый скрытый слой содержит nh (по h на каждый вход) функций принадлежности $\mu_{il}(x)$, $i = 1, 2, \dots, n$; $l = 1, 2, \dots, h$ и выполняет фаззификацию входного пространства. Второй скрытый слой обеспечивает агрегирование уровней принадлежности, рассчитанных в первом слое, и состоит из h блоков умножения. Третий скрытый слой – это слой синаптических весов, подлежащих определению в процессе обучения. Четвертый слой образован двумя сумматорами и вычисляет суммы выходных сигналов

второго и третьего слоев. И, наконец, в пятом (выходном) слое производится нормализация, в результате которой вычисляется выходной сигнал системы \hat{y} .

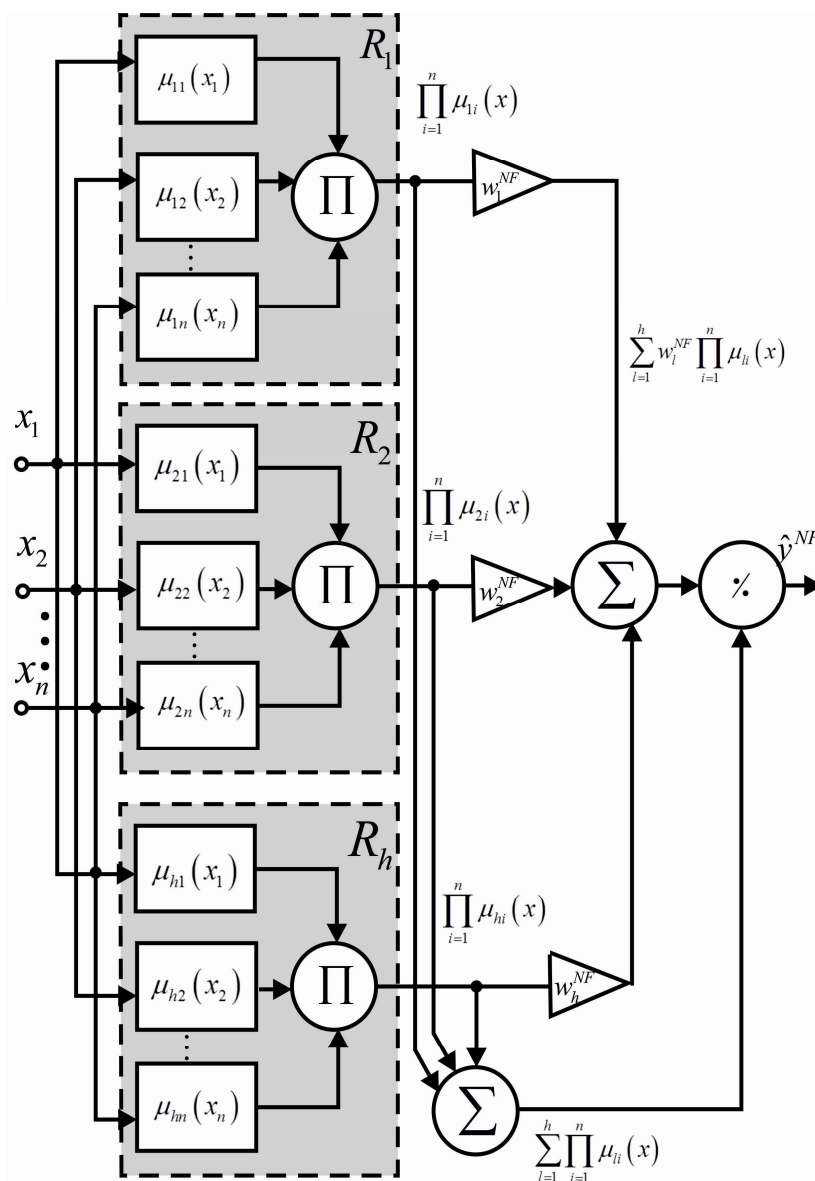


Рисунок 1 – Архитектура нейро-фаззи системы

Таким образом, если на вход NFS подан векторный сигнал $x(k)$, элементы первого скрытого слоя вычисляют уровни принадлежности $0 \leq \mu_{ii}(x(k)) \leq 1$, при этом в качестве функций принадлежности чаще всего используются традиционные гауссианы

$$\mu_{ii}(x_i(k)) = \exp\left(-\frac{(x_i(k) - c_{li})^2}{2\sigma_i^2}\right),$$

где c_{li} , σ_i – параметры центров и ширины соответственно.

Во втором скрытом слое вычисляются агрегированные значения $\prod_{i=1}^n \mu_{li}(x_i(k))$, при этом для гауссианов с одинаковыми параметрами ширины σ

$$\prod_{i=1}^n \mu_{li}(x_i(k)) = \prod_{i=1}^n \exp\left(-\frac{(x_i(k) - c_{li})^2}{2\sigma^2}\right) = \exp\left(-\frac{\|x(k) - c_l\|^2}{2\sigma^2}\right),$$

(здесь $c_l = (c_{l1}, c_{l2}, \dots, c_{ln})^T$), т.е. сигналы на выходах блоков умножения второго скрытого слоя аналогичны сигналам на выходах нейронов первого скрытого слоя радиально-базисных нейронных сетей.

Выходами третьего скрытого слоя есть значения $w_l \prod_{i=1}^n \mu_{li}(x_i(k))$

(здесь w_l , $l=1, 2, \dots, h$ – синаптические веса, подлежащие определению),

четвертого – $\sum_{l=1}^h w_l \prod_{i=1}^n \mu_{li}(x_i(k))$ и $\sum_{l=1}^h \prod_{i=1}^n \mu_{li}(x_i(k))$ и, наконец, на выходе

де системы (пятого слоя) появляется сигнал

$$\begin{aligned} \hat{y}(x(k)) &= \frac{\sum_{l=1}^h w_l \prod_{i=1}^n \mu_{li}(x_i(k))}{\sum_{l=1}^h \prod_{i=1}^n \mu_{li}(x_i(k))} = \sum_{l=1}^h w_l \frac{\prod_{i=1}^n \mu_{li}(x_i(k))}{\sum_{l=1}^h \prod_{i=1}^n \mu_{li}(x_i(k))} = \\ &= \sum_{l=1}^h w_l \varphi_l(x(k)) = w^T \varphi(x(k)), \end{aligned}$$

где

$$\varphi_l(x(k)) = \frac{\prod_{i=1}^n \mu_{li}(x_i(k))}{\sum_{l=1}^h \prod_{i=1}^n \mu_{li}(x_i(k))},$$

$$w = (w_1, w_2, \dots, w_h)^T,$$

$$\varphi(x(k)) = (\varphi_1(x(k)), \varphi_2(x(k)), \dots, \varphi_h(x(k)))^T.$$

Несложно заметить, что рассматриваемая система реализует нелинейное отображение входного пространства в скалярный выходной сигнал подобно нормализованной RBFN [11], а по архитектуре

совпадает с системой Такаги-Сугено-Канга нулевого порядка, т.е. архитектурой Л. Ванга – Дж. Менделя [3].

Как уже отмечалось, для настройки синаптических весов NFS могут быть использованы известные адаптивные алгоритмы идентификации-обучения такие, как, например, экспоненциально-взвешенный рекуррентный метод наименьших квадратов

$$\begin{cases} w(k) = w(k-1) + \frac{P(k-1)(y(k) - w^T(k-1)\varphi(x(k)))}{\beta + \varphi^T(x(k))P(k-1)\varphi(x(k))} \varphi(x(k)) = \\ = w(k-1) + \frac{P(k-1)(y(k) - \hat{y}(k))}{\beta + \varphi^T(x(k))P(k-1)\varphi(x(k))} \varphi(x(k)), \\ P(k) = \frac{1}{\beta} \left(P(k-1) - \frac{P(k-1)\varphi(x(k))\varphi^T(x(k))P(k-1)}{\beta + \varphi^T(x(k))P(k-1)\varphi(x(k))} \right), \quad 0 < \beta \leq 1, \end{cases}$$

(здесь $y(k)$ – внешний обучающий сигнал, β – параметр забывания устаревшей информации), оптимальный по быстродействию градиентный одношаговый алгоритм Качмажа-Уидроу-Хоффа:

$$w(k) = w(k-1) + \frac{y(k) - w^T(k-1)\varphi(x(k))}{\|\varphi(x(k))\|^2} \varphi(x(k)), \quad (2)$$

алгоритм обучения, обладающий как следящими, так и сглаживающими свойствами [12,13]

$$\begin{cases} w(k) = w(k-1) + p^{-1}(k)(y(k) - w^T(k-1)\varphi(x(k)))\varphi(x(k)), \\ p(k) = \beta p(k-1) + \|\varphi(x(k))\|^2, \quad 0 \leq \beta \leq 1, \end{cases} \quad (3)$$

и тому подобные процедуры.

Интересно отметить, что процедура (3) связана с алгоритмом (1) соотношением

$$p(k) = \text{Tr } P(k),$$

а при $\beta = 0$ – приобретает форму алгоритма (2).

2. Самообучение функций принадлежности первого скрытого слоя. Процесс настройки функций принадлежности первого скрытого слоя проиллюстрируем на примере двумерного вектора входов $x(k) = (x_1(k), x_2(k))^T$ и пяти функций принадлежности $\mu_i(x_i(k))$, $l = 1, 2, 3, 4, 5$; $i = 1, 2$ на каждом из входов. В этом случае NFS содержит $nh = 10$ функций принадлежности. Начальные положения

центров $c_{li}(0)$ равномерно распределены по координатным осям x_1, x_2 , а расстояние между ними определяется соотношением

$$\Delta(0) = \frac{x_{i\max} - x_{i\min}}{h-1} = \frac{2}{h-1} = 0,5$$

для $-1 \leq x_i \leq 1$.

Данная ситуация иллюстрируется рис. 2.

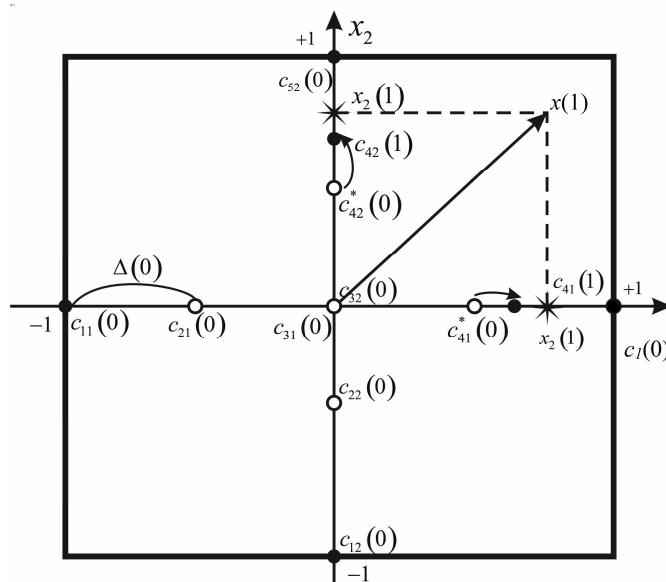


Рисунок 2 – Самообучение функций принадлежности на первом шаге

В случае многомерного вектора входов $x(k) \in R^n$, центры $c_{li}(0)$ распределяются равномерно по осям гиперкуба $[-1, 1]^n$.

Пусть на вход системы подан первый вектор-образ x_1 (на рис. 2 – $x(1) = (x_1(1), x_2(2))^T$). Далее на каждой из осей находятся центры «победители» $c_{li}^*(0)$ ближайшие к $x_i(1)$ в смысле расстояния

$$d_{li} = |x_i(1) - c_{li}(0)|,$$

т.е.

$$c_{li}^*(0) = \arg \min \{d_{1i}, d_{2i}, \dots, d_{hi}\}.$$

Заметим, что данная процедура по сути является реализацией процесса конкуренции по Т. Кохонену [10] с той лишь разницей, что «победители» по каждой из осей могут принадлежать функциям принадлежности с различными индексами l . На рис. 2 такими «победителями» оказались $c_{41}^*(0)$ и $c_{42}^*(0)$.

Далее эти «победители» подтягиваются к компонентам входного сигнала $x_i(1)$ согласно правилу самообучения Т. Кохонена «Победитель получает все» (WTA), которое для ситуации на рис. 2 может быть записано в виде

$$c_{li}(1) = \begin{cases} c_{li}^*(0) + \eta_{li}(1)(x_i(1) - c_{li}^*(0)) - \text{для победителя } l=4, \\ c_{li}(0) - \text{для всех остальных, } l=1,2,3,5; \end{cases}$$

а для общего случая:

$$c_{li}(k) = \begin{cases} c_{li}^*(k-1) + \eta_{li}(k)(x_i(k) - c_{li}^*(k-1)) - \text{для победителя } l=1,2,\dots,h; i=1,2,\dots,n; \\ c_{li}(k-1) - \text{для всех остальных.} \end{cases}$$

При этом в качестве значения параметра шага обучения в простейшем случае может быть принято значение

$$\eta_{li}(k) = \frac{1}{k_{li}},$$

где k_{li} – число раз, когда $c_{li}(k)$ был «победителем», что соответствует популярному методу кластеризации К-средних.

В более общем случае может быть использована оценка, предложенная для стандартной карты Т. Кохонена, вида [14]:

$$\begin{cases} \eta_{li}(k) = p_{li}^{-1}, \\ p_{li}(k) = \beta p_{li}(k-1) + x_i^2, \quad 0 \leq \beta \leq 1. \end{cases}$$

Как можно заметить, предлагаемый подход является модификацией правила самообучения Т. Кохонена с той разницей, что стандартное самообучение реализуется на гипершаре $\|x(k)\|=1$, а в нашем случае – на гиперкубе $[-1,1]^n$.

На рис. 3 приведена структурная схема комбинированного обучения-самообучения нейро-фаззи системы.

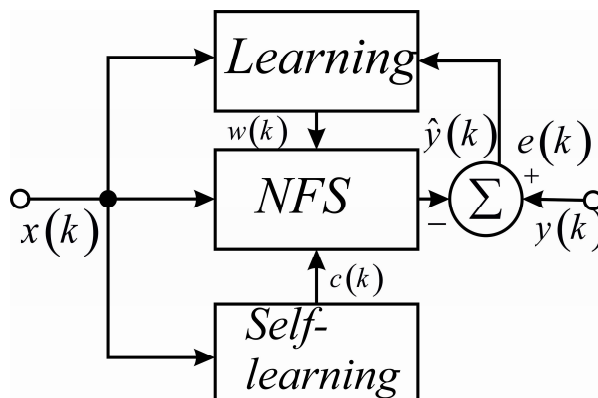


Рисунок 3 – Комбинированное обучение – самообучение

Процесс функционирования этой системы происходит следующим образом. При поступлении на вход вектора-образа $x(k)$ сначала в блоке самообучения происходит коррекция функций принадлежности $\mu_i(x_i(k))$, т.е. рассчитываются центры $c_i(k)$. Далее на основании уточненных функций принадлежности

$$\mu_i(x_i(k)) = \exp\left(-\frac{(x_i(k) - c_i(k))^2}{2\sigma^2}\right)$$

и ранее рассчитанного вектора синаптических весов $w(k-1)$ с помощью алгоритмов обучения с учителем ((1) или (2), или (3)) происходит расчет весов выходного слоя NFS $w(k)$.

Выводы. Предложен подход к комбинированному обучению синаптических весов и центров функций принадлежности нейрофаззи систем, основанный на использовании как правил обучения с учителем, так и самообучения. Основным преимуществом развиваемого подхода является возможность его использования в on-line режиме, когда обучающая выборка на вход системы подается последовательно, а ее объем заранее не фиксирован. Полученные результаты могут быть использованы при решении широкого класса задач, возникающих в Dynamic Data Mining.

ЛИТЕРАТУРА

1. Rutkowski L. Computational Intelligence. Methods and Tehniques. / Rutkowski L. // Berlin-Heidelberg: Springer-Verlag, – 2008. – 514 p.
2. Jang J.-S. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Maching Intelligence. / Jang J.-S., Sun C.-T., Mizutani E. // Upper Saddle River: Prentice Hall – 1997. – 640 p.
3. Wang L.-X. Fuzzy basis functions, universal approximation and orthogonal least squares learning. / Wang L.-X., Mendel J. M. // IEEE Trans. on Neural Networks – 1993. – 3. – P. 807-814.
4. Cios K. J. Neuro-fuzzy algorithms. / Cios K. J., Pedrycz W. – Oxford: IOP Publishing Ltd and Oxford University Press. // Handbook of Neural Computation, – 1997. – D1. 3:1 – D1. 3:7.
5. Friedman J. The Elements of Statistical Learning. Data Mining, Inference and Prediction. / Friedman J., Hastie T., Tibshirani R. // Berlin: Springer, – 2003.–552 p.
6. Осовский С. Нейронные сети для обработки информации. / Осовский С. // М.: Финансы и статистика. – 2002. – 344 с.

7. Бодянский Е.В. Адаптивное обучение архитектуры и параметров радиально-базисной нейронной сети / Бодянский Е.В., Дейнеко А.А. // Системные технологии. – Днепропетровск, 2013. – 4. – 87. – С. 166-173.
8. Бодянский Е.В. Эволюционирующая радиально-базисная нейронная сеть и ее обучение с помощью карты Кохонена / Бодянский Е.В., Дейнеко А.А. // Научно-техническая конференция: «Информационные технологии в металлургии и машиностроении.» – Днепропетровск, 2013. – С. 75-77.
9. Torres J.A. The problem of organizing and partition large data sets in learning algorithms for SOM-RBF mixed structure / Torres J.A., Martinez S., Martinez F.J., Peralta M. // Proc. 5th Int. Joint Conf. on Computational Intelligence. – Vilamoura, Algarve, Portugal – 2013 – P 497-501.
10. Kohonen T. Self-Organizing Maps / Kohonen T. – Berlin: Springer-Verlag. – 1995. – 362 p.
11. Nelles O. Nonlinear System Identification. / Nelles O. // Berlin: Springer, – 2001. – 785p.
12. Bodyanskiy Ye. An adaptive learning algorithm for a neuro-fuzzy network / Bodyanskiy Ye., Kolodyazhniy V., Stephan A. / Ed. by B. Reusch "Computational Intelligence. Theory and Applications." – Berlin – Heidelberg – New York: Springer, 2001. – P. 68-75.
13. Otto P. / A new learning algorithm for a forecasting neuro-fuzzy network. / Otto P., Bodyanskiy Ye., Kolodyazhniy V. // Integrated Computer-Aided Engineering. – 2003. – 10. – № 4. – P. 399-409.
14. Бодянский Е.В. Искусственные нейронные сети: архитектуры, обучение, применение. / Бодянский Е.В., Руденко О.Г. – Харьков: ТЕЛТЕХ. – 2004. – 372 с.