

УДК 004.032.26

А.К. Тищенко, И.П. Плисс

Харьковский национальный университет радиоэлектроники, Харьков

РЕКУРРЕНТНЫЙ АЛГОРИТМ НЕЧЕТКОЙ КЛАСТЕРИЗАЦИИ В НЕЙРО-НЕО-ФАЗЗИ СЕТИ ВСТРЕЧНОГО РАСПРОСТРАНЕНИЯ

Введена архитектура, состоящая из нео-фаззи-нейронов и слоя рекуррентной нечеткой кластеризации, а также алгоритмы обучения, демонстрирующие высокое быстродействие. Предлагаемая архитектура предназначена для решения задач прогнозирования нестационарных временных рядов.

Ключевые слова: нечеткая кластеризация, нео-фаззи-нейрон, алгоритм обучения, нестационарные временные ряды.

Введение

Задача прогнозирования нестационарных временных рядов занимает одно из основных мест среди множества задач, которые можно решить с помощью аппарата искусственных нейронных сетей. В задачах прогнозирования зачастую от точности получаемых прогнозов зависит качество принимаемых решений. Поскольку часто в реальных задачах приходится сталкиваться с временными рядами, которые характеризуются высоким уровнем нестационарности и наличием аномальных выбросов, то традиционные методы анализа данных оказываются неэффективными, поэтому в последнее время получила широкое распространение идея гибридных нейро-фаззи сетей, объединяющих в себе искусственные нейронные сети и методы нечеткой кластеризации. Также в последнее время возрос интерес к задаче анализа нестационарных временных рядов, изменяющих свои свойства в неизвестные моменты времени. Для решения подобных задач необходимо произвести разделение временной последовательности на внутренне однородные участки. В реальных задачах, как правило, внутренние изменения в исследуемом объекте происходят плавно, без резких скачков, что затрудняет использование традиционных методов. Предлагается алгоритм обучения нейронной сети, отличающийся высоким быстродействием. Так как состояния исследуемого объекта часто изменяются сразу по нескольким характеристикам, то целесообразно использовать методы нечеткой кластеризации.

Архитектура гибридной нейро-нео-фаззи-сети и нео-фаззи-нейрона

Предлагается архитектура новой прогнозирующей гибридной нейронной сети, основанной на идее нейронных сетей встречного распространения [1, 2] и являющейся объединением слоя рекуррентной кластеризации [3], и слоя нео-фаззи нейронов [4], называемой в дальнейшем гибридной нейро-нео-фаззи сетью. Поступающие на входы сети наблюдения обозначим $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$, выход сети – $y(k)$, $k = 1, 2, \dots, N$ – текущее дискретное время.

Работа сети заключается в вычислении уровней принадлежности текущего наблюдения к кластерам, определяемым весами нейронов кластеризационного слоя, вычислении значений выходных сигналов нео-фаззи-нейронов и ранжировании этих значений по соответствующим степеням принадлежности на выходах кластеризационного слоя. Выходом сети является прогноз, представленный в виде взвешенной суммы выходов нео-фаззи-нейронов.

Между нейронами кластеризационного слоя и слоем нео-фаззи-нейронов устанавливается взаимнооднозначное соответствие: каждому нейрону слоя кластеризации ϕ_1 соответствует нео-фаззи-нейрон NFN_1 , $1 = 1, 2, \dots, h$, h – количество нейронов в скрытых слоях.

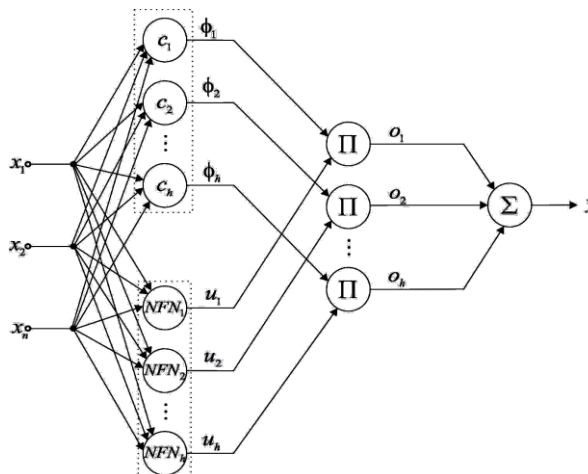


Рис. 1. Архитектура нейро-нео-фаззи сети.

Выходы нейронов слоя кластеризации вычисляются в виде

$$\phi_1(k) = \Phi_1(k) / \sum_{p=1}^h \Phi_p(k), \quad (1)$$

где $\Phi_1(k) = \Phi(\|x(k) - c_1\|_{R_1^{-1}}^2)$ – функции принадлежности нейронов; $\|\cdot\|_{R_1^{-1}}$ – норма вектора по матрице R_1^{-1} .

Следует отметить, что выходы нейронов слоя кластеризации удовлетворяют условию единичного разбиения

$$\sum_{l=1}^h \phi_l(k) = 1, \quad (2)$$

что следует из (1).

Архитектура нео-фаззи-нейрона представлена на рис. 2.

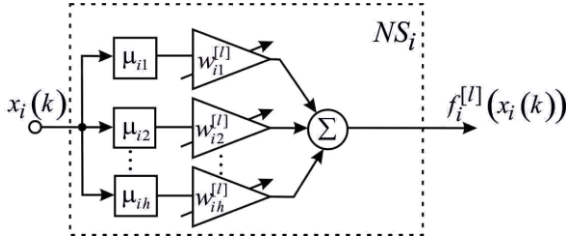


Рис. 2. Архитектура нелинейного синапса

Выходы нео-фаззи-нейрона рассчитываются в виде

$$u_i(k) = \sum_{j=1}^m f_i^{[j]}(x_i(k)) = \sum_{j=1}^m \sum_{i=1}^n \mu_{ij}(x_i(k)) w_{ij}^{[j]}(k), \quad (3)$$

где m – количество функций принадлежности в нелинейном синапсе NS_i .

Функции принадлежности μ_{ij} нелинейных синапсов обеспечивают разложение единицы на некотором интервале $[a, b]$. В дальнейшем в качестве μ_{ij} будем рассматривать треугольные функции, которые вычисляются в форме

$$\mu_{ij}(a) = \begin{cases} \frac{v_{ij} - a}{v_{ij} - v_{i,j-1}}, & a \in (v_{i,j-1}; v_{ij}], \\ \frac{a - v_{ij}}{v_{i,j+1} - v_{ij}}, & a \in (v_{ij}; v_{i,j+1}], \\ 0, & \text{иначе} \end{cases} \quad (4)$$

для $i = 1 \dots 3$; $j = 1, 2, \dots, h$.

Выход всей сети вычисляется в соответствии с выражением:

$$y(k) = \sum_{l=1}^h o_l(k) = \sum_{l=1}^h u_l(k) \phi_l(k) = \phi^T(k) u(k), \quad (5)$$

где $\phi(k) = (\phi_1(k), \phi_2(k), \dots, \phi_h(k))^T$,
 $u(k) = (u_1(k), u_2(k), \dots, u_h(k))^T$.

Таким образом, предложенную нейро-нео-фаззи сеть можно рассматривать как ансамбль нео-фаззи нейронов, вклад которых в выход сети определяется сигналами выхода слоя нечеткой кластеризации.

Предобработка данных для слоя нечеткой рекуррентной кластеризации

Для сегментации временной последовательности

$$X = \{x(1), x(2), \dots, x(k), \dots, x(N)\} \quad (6)$$

используется подход, базирующийся на непрямой кластеризации последовательности. В соответствии с этим подходом выделяются некоторые характеристики для дальнейшего отображения в преобразованное пространство признаков. В качестве характеристик начальной временной последовательности могут быть выбраны корреляционные, регрессионные, спектральные и другие характеристики.

В частности могут быть использованы оценки среднего значения, дисперсии, коэффициентов автокорреляции. Для придания адаптивных свойств, эти оценки должны быть вычислены с помощью процедуры экспоненциального сглаживания [5]. Среднее значение может быть вычислено в виде

$$s(k) = \alpha x(k) + (1 - \alpha)s(k-1), \quad 0 < \alpha < 1, \quad (7)$$

где $\alpha = 2 / (T + 1)$ – коэффициент, который определяет сглаживание на окне ширины T .

Значение дисперсии временной последовательности может быть оценено в форме

$$\sigma^2(k) = \alpha(x(k) - s(k))^2 + (1 - \alpha)\sigma^2(k-1), \quad (8)$$

и коэффициенты автокорреляции:

$$\rho(k, \tau) = \alpha(x(k) - s(k))(x(k - \tau) - s(k - \tau)) + (1 - \alpha)\rho(k - 1, \tau), \quad (9)$$

где $\tau = 1, 2, \dots, \tau_{\max}$ – коэффициенты запаздывания.

Векторы признаков

$$x(k) = (s(k), \sigma^2(k), \rho(k, 1), \dots, \rho(k, \tau_{\max}))^T, \quad (10)$$

содержат $2 + \tau_{\max}$ элементов, вычисляются на каждом шаге дискретного времени k и образуют множество, состоящее из N n -мерных векторов признаков

$$X' = \{x'(1), x'(2), \dots, x'(N)\}, \quad (11)$$

где $x'(k) \in \mathbb{R}^n$, $k = 1, 2, \dots, N$.

Результатом применения процедуры нечеткой кластеризации должно быть распределение исходных данных на m кластеров с некоторой степенью принадлежности $w_j(k)$ k -го вектора признаков $x'(k)$ j -му кластеру.

Участки временных последовательностей, идущие подряд во времени и принадлежащие одинаковым кластерам, будут образовывать сегменты временной последовательности на выходе.

Алгоритм обучения гибридной сети

Процедура обучения нейро-нео-фаззи сети состоит из двух этапов: на первом этапе происходит настройка параметров слоя нечеткой класте-

ризации, на втором этапе выполняется обучение нео-фаззи-нейронов с учетом сигналов на выходах нейронов слоя кластеризации. Такая процедура обучения согласуется с подходом к обучению сети встречного распространения. В режиме последовательного обучения можно выполнять обучение слоев поочередно.

Рекуррентный алгоритм обучения слоя кластеризации

Используемый нами подход принадлежит к классу алгоритмов, основанных на целевых функциях, разработанных для решения задач кластеризации путем оптимизации определенного критерия, заданного предварительно.

Введем в рассмотрение целевую функцию

$$E v_j k , c_j = \sum_{k=1}^T \sum_{j=1}^m v_j^\beta D x k , c_j , \quad (12)$$

и ограничения

$$\sum_{j=1}^m v_j k = 1, k = 1...T; \quad (13)$$

$$0 < \sum_{k=1}^T v_j k \leq T, j = 1...m, \quad (14)$$

где T – количество элементов в выборке; m – количество центроидов; $v_j k \in [0,1]$ – степень принадлежности вектора $x k$ j-му кластеру; c_j – прототип (центроид) j-ого кластера; β – неотрицательный параметр (фаззификатор), $D x k , c_j$ – расстояние между $x k$ и c_j . Результатом кластеризации является матрица $W^{fp} = w_j^{fp} k$ размерности $T \times m$, называемая «матрицей нечеткого разбиения».

В используемом алгоритме фаззи-кластеризации используется робастная конструкция для измерения функции расстояния [3]

$$D x k , c_j = \sum_{i=1}^n \frac{1 - \text{sech}^2 x_i k - c_{ji}}{2} \times \times x_i k - c_{ji}^{2/5}, \quad (15)$$

где $x_i k$ – i-ая компонента вектора входов $x k$; c_{ji} – i-ая компонента вектора центроидов кластеров c_j .

Основываясь на формулах (12) и (15), перепишем целевую функцию для робастной кластеризации

$$E v_j k , c_j = \sum_{k=1}^T \sum_{j=1}^m v_j^\beta k D x k , c_j =$$

$$= \sum_{k=1}^T \sum_{j=1}^m v_j^\beta k \sum_{i=1}^n \frac{1 - \text{sech}^2 x_i k - c_{ji}}{2} \times \times x_i k - c_{ji}^{2/5} \quad (16)$$

и введем соответствующую ей функцию Лагранжа

$$L v_j k , c_j, \lambda k = \sum_{k=1}^T \sum_{j=1}^m v_j^\beta k \times \left(\frac{1 - \text{sech}^2 x_i k - c_{ji}}{2} \times x_i k - c_{ji}^{2/5} + \sum_{k=1}^T \lambda k \left(\sum_{j=1}^m v_j k - 1 \right) \right), \quad (17)$$

где λk – неопределенный множитель Лагранжа, обеспечивающий выполнение ограничений (13), (14). Седловая точка функции Лагранжа может быть найдена из системы уравнений Куна-Такера:

$$\begin{cases} \frac{\partial L v_j k , c_j, \lambda k}{\partial v_j k} = 0, \\ \frac{\partial L v_j k , c_j, \lambda k}{\partial \lambda k} = 0, \\ \nabla_{c_j} L v_j k , c_j, \lambda k = 0. \end{cases} \quad (18)$$

Решение первых двух уравнений приводит к

$$\begin{cases} v_j k = \frac{D x k , c_j \frac{1}{1-\beta}}{\sum_{l=1}^m D x k , c_l \frac{1}{1-\beta}}, \\ \lambda k = - \left(\sum_{l=1}^m \beta D x k , c_l \frac{1}{1-\beta} \right)^{1-\beta}, \end{cases} \quad (19)$$

а третье уравнение

$$\nabla_{c_j} L v_j k , c_j, \lambda k = \sum_{j=1}^T v_j^\beta k \nabla_{c_j} D x k , c_j = 0 \quad (20)$$

не имеет аналитического решения. Решение (20) может быть вычислено с помощью локальной модификации функции Лагранжа и рекуррентного алгоритма фаззи-кластеризации. Поиск седловой точки локальной функции Лагранжа

$$L_k v_j k , c_j, \lambda k = \sum_{j=1}^m v_j^\beta k D x k , c_j + \lambda k \left(\sum_{j=1}^m v_j k - 1 \right) \quad (21)$$

с помощью процедуры Эрроу-Гурвица-Удзавы приводит к алгоритму (22), где ηk – параметр уров-

ня обучения, c_{ji}^k – i -ая компонента j -го вектора прототипов, вычисленная на k -м шаге:

$$\left\{ \begin{aligned} v_j^k &= \frac{D_{xk}, c_j \frac{1}{1-\beta}}{\sum_{l=1}^m D_{xk}, c_l \frac{1}{1-\beta}}, \\ c_{ji}^{k+1} &= c_{ji}^k - \eta^k \frac{\partial}{\partial c_{ji}} L_k(v_j^k, c_j, \lambda^k) = \\ &= c_{ji}^k + \eta^k v_j^{\beta} \left[2 \operatorname{sech}^2(x_i^k - c_{ji}^k) \times \right. \\ &\times \left. \tanh(x_i^k - c_{ji}^k) |x_i^k - c_{ji}^k|^{2/5} + \right. \\ &+ 0.4 \left. 1 - \operatorname{sech}^2(x_i^k - c_{ji}^k) \times \right. \\ &\left. \times |x_i^k - c_{ji}^k|^{-3/5} \operatorname{sign}(x_i^k - c_{ji}^k) \right]. \end{aligned} \right. \quad (22)$$

Вычислительная сложность данного алгоритма находится на одном уровне с алгоритмами нечетких s -средних и Густафсона-Кесселя и зависит линейно от количества точек в исходной выборке.

Алгоритм обучения нео-фаззи-нейрона

Пусть обучающая выборка задана наборами векторов

$$X = (x(1), x(2), \dots, x(N)),$$

где X – значения входных сигналов для обучения сети.

Для обучения параметров нео-фаззи-нейронов используется процедура наименьших квадратов. Настройка параметров выполняется с учетом выхода соответствующего нейрона слоя кластеризации в качестве весового коэффициента.

Вводя матрицу коэффициентов

$$Y_1 = I_N \cdot (\phi_1(1), \phi_1(2), \dots, \phi_1(N))^T, \quad (23)$$

где I_N – единичная матрица размерности $(N \times N)$, процедуру обучения можно записать в форме

$$w^{[1]} = (\Psi^T Y_1^{-1} \Psi)^{-1} \Psi^T Y_1^{-1} Y, \quad (24)$$

где $\Psi = (\Psi(1), \Psi(2), \dots, \Psi(N))^T$, или в рекуррентном виде:

РЕКУРРЕНТНИЙ АЛГОРИТМ НЕЧІТКОЇ КЛАСТЕРИЗАЦІЇ У НЕЙРО-НЕО-ФАЗЗИ МЕРЕЖІ ЗУСТРІЧНОГО ПОШИРЮВАННЯ

О.К. Тищенко, І.П. Плісс

Запропоновано архітектуру, що містить нео-фаззи-нейрони та шар рекуррентної нечіткої кластеризації. Запропонована архітектура призначена для прогнозування нестационарних часових послідовностей.

Ключові слова: нечітка кластеризація, нео-фаззи-нейрон, алгоритм навчання, нестационарні тимчасові ряди.

RECURRENT FUZZY CLUSTERIZATION ALGORITHM IN THE NEURO-NEO-FUZZY COUNTERPROPAGATION NETWORK

O.K. Tyshchenko, I.P. Pliss

The architecture that consists of neo-fuzzy-neurons and a layer of recurrent fuzzy clusterization is proposed. The proposed architecture can be used for unstationary time series forecasting.

Keywords: unclear clusterization, neo-fuzzy-neurons, teaching algorithm, unstationary temporal rows.

$$\left\{ \begin{aligned} w(k) &= w(k-1) + \Psi(k) \times \\ &\times \frac{P_1(k-1)(\tilde{y}(k) - \Psi^T(k)w(k-1))}{1 + \phi_1(k)\Psi^T(k)P_1(k-1)\Psi(k)}, \\ P_1(k) &= \phi_1(k) P_1(k-1) - \\ &\times \frac{\phi_1(k)P_1(k-1)\Psi(k)\Psi^T(k)P_1(k-1)}{1 + \phi_1(k)\Psi^T(k)P_1(k-1)\Psi(k)}, \end{aligned} \right. \quad (25)$$

где $P_1(k)$ – матрица размерности $(nm \times nm)$.

Выводы

Разработана гибридная нейро-нео-фаззи-система вычислительного интеллекта, предназначенная для обработки временных последовательностей с существенно нестационарными характеристиками в условиях априорной неопределенности. Используемый в работе алгоритм кластеризации позволяет осуществлять разбиение данных с аномальными выбросами.

Список литературы

1. Hecht-Nielsen R. Counterpropagation Networks // *Applied Optics*. – 1987. – V. 26. – P. 4979-4984.
2. Zhang Z., Zheng N., Wang T. Fuzzy generalization of the counter-propagation neural network: a family of soft competitive basis function neural networks // *Soft Computing*. – 2001. – V. 5 - 6. – P. 440-450.
3. Recursive fuzzy clustering algorithms for segmentation of biological time series / Ye. Bodyanskiy, Ye. Gorshko, V. Kolodyazhnyi, O. Shylo // *Proc. 13-th East West Fuzzy Coll. 2006, Germany, 2006. – Zittau/Goerlitz: University of Applied Sciences (FH), 2006. – P. 130-139.*
4. Neo fuzzy neuron and its applications to system identification and prediction of the system behavior / T. Yamakawa, E. Uchino, T. Miki, H.A. Kusanagi // *Proc. 2nd Int. Conf. on Fuzzy Logic and Neural Networks "IIZUKA-92"*. – Iizuka, Japan, 1992. – P. 477-483.
5. L.F. Failure diagnosis and performance monitoring / L.F. Pau. – N.Y.: Marcel Dekker, 1981. – 448 p.

Поступила в редакцию 20.04.2012

Рецензент: д-р техн. наук, проф. Е.В. Бодянский, Харьковский национальный университет радиоэлектроники, Харьков.