

УДК 681.513.7

Е.В. БОДЯНСКИЙ

ОБНАРУЖЕНИЕ РАЗЛАДОВ В НЕЛИНЕЙНЫХ СТОХАСТИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЯХ С ПОМОЩЬЮ РЕКУРРЕНТНЫХ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Задача обнаружения изменения свойств стохастических последовательностей тесно связана с проблемой технической и медицинской диагностики. Эта задача широко обсуждалась в научной литературе [1–5], а для ее решения предложено множество подходов, связанных большей частью с идеями математической статистики, теории случайных процессов, распознавания образов, кластер-анализа и т.п. Не вдаваясь в критику полученных результатов, заметим, что жесткие предположения о статистических свойствах рядов ограничивают возможности данных подходов.

Более универсальным представляется многомодельный подход [1; 5—7]. Его суть состоит в том, что диагностируемый сигнал проходит через множество моделей, каждая из которых базируется на определенной гипотезе относительно характера возможных изменений. Если какая-либо из гипотез действительно подтверждается, то сигналы обновлений на выходе соответствующей модели должны быть малыми. Таким образом, решающий механизм, по сути, основан на отыскании той модели, на выходе которой обновления минимальны, а вероятность соответствующей гипотезы максимальна. Достоинства изложенного подхода несомненны, однако реальная последовательность обычно настолько многообразна, что никакая из моделей (как правило, линейных) полностью не отражает ее изменяющихся свойств.

Последние годы характеризуются резким увеличением количества исследований в области теории и практики искусственных нейронных сетей, в том числе и для решения задач диагностики [8–14]. Предложенные диагностирующие нейронные сети реализуют в основном идеи теории классификации при наличии обучающей выборки, при этом появление не предусмотренных априори состояний процесса может быть не обнаружено сетью.

Предлагаются архитектура многослойной рекуррентной искусственной нейронной сети и алгоритмы настройки ее параметров, сочетающие в себе достоинства многомодельного подхода и аппроксимирующих свойств прогнозирующих нейронных сетей [15–18] с нелинейными функциями активации. Изменение свойств стохастической последовательности фиксируется с помощью диагностирующего вектора, элементы которого являются синаптическими весами выходного нейрона.

Предлагаемая нами архитектура диагностирующей рекуррентной нейронной сети показана на рис. 1 и представляет собой сеть элементарных нейронов. Последние отличаются друг от друга видом функций активации и алгоритмами обучения, являющимися в общем случае градиентными процедурами безусловной или условной оптимизации.

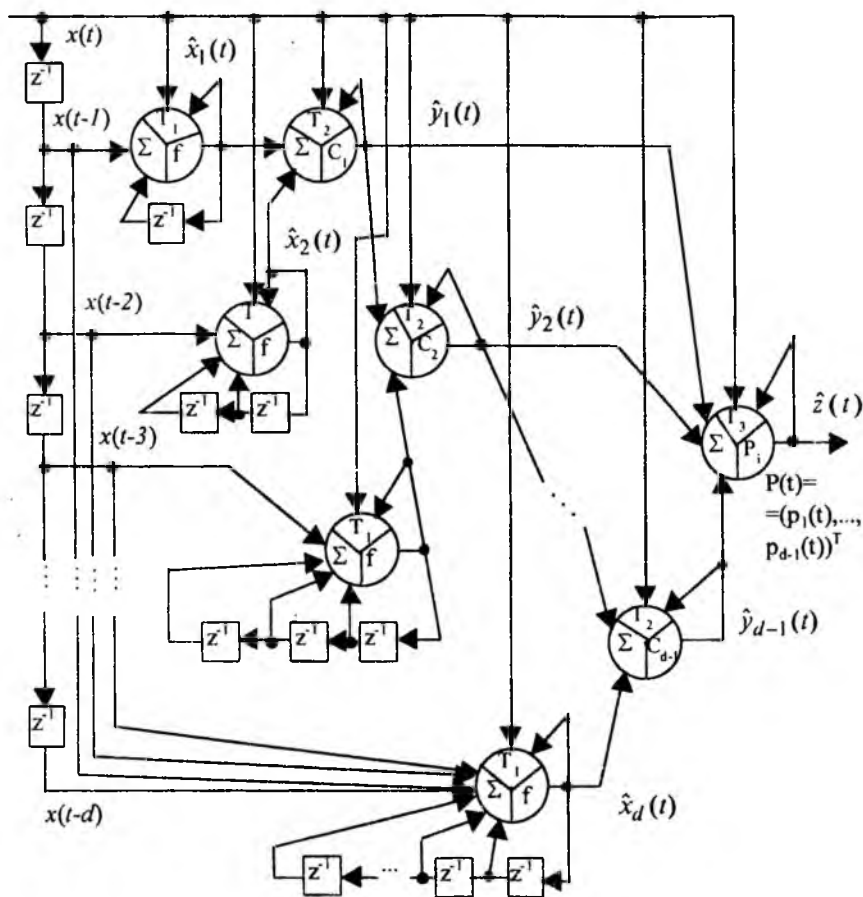


Рис. 1

На рис. 2 изображен элементарный искусственный нейрон, при этом сектор Σ соответствует операции линейного суммирования входов $\sum w_j(t)x_j(t)$ с настраиваемыми синаптическими весами $w_j(t)$, сектор f соответствует нелинейному преобразованию с помощью функции активации $\hat{y}(t) = f(\sum w_j(t)x_j(t))$, сектор T соответствует алгоритму настройки нейрона на основе рекуррентных процедур типа

$$w(t+1) = w(t) + \psi(y(t), \hat{y}(t)),$$

где $y(t)$ – обучающий сигнал.

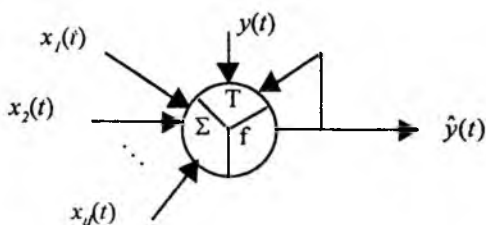


Рис. 2

Контролируемая стохастическая последовательность $\{x(t)\}$, $t=1, 2, \dots$ подается на входной слой сети, представляющий собой последовательную цепочку элементов чистого запаздывания z^{-1} ($z^{-1}x(t) = x(t-1)$), в результате чего на выходе этого слоя формируется набор задержанных значений временного ряда $x(t-1)$, $x(t-2)$, ..., $x(t-d)$. Чем больше значение d , тем более широкими диагностирующими возможностями обладает нейронная сеть.

Первый скрытый слой образован нейронами типа Мак-Каллоха — Питтса, на суммирующие входы которых подаются задержанные значения контролируемой последовательности $x(t)$ и по цепи обратной связи — задержанные значения прогноза $\hat{x}_j(t)$, $j=1, 2, \dots, d$. Входы нейронов T_1 соответствуют входам алгоритма настройки синаптических весов, а f описывает нелинейную функцию активации нейронов.

В результате обработки сигнала $x(t)$ нейронами первого слоя на их выходах появляются прогнозные оценки

$$\left\{ \begin{array}{l} \hat{x}_1(t) = f(x(t-1), \hat{x}_1(t-1)); \\ \hat{x}_2(t) = f(x(t-1), x(t-2), \hat{x}_2(t-1), \hat{x}_2(t-2)); \\ \dots\dots\dots \\ \hat{x}_d(t) = f(x(t-1), \dots, x(t-d), \hat{x}_d(t-1), \dots, \hat{x}_d(t-d)), \end{array} \right. \quad (1)$$

соответствующие нелинейному процессу авторегрессии – скользящего среднего (NARMA-процесс) [16; 17] порядка от 1 до d . Задачей, решаемой сетью, является определение текущего значения порядка NARMA-процесса и моментов возможного его изменения в реальном времени.

$$j = 1, 2, \dots, d-1:$$

Нейронами второго скрытого слоя T_2 производится попарное объединение выходов нейронов T_1 с целью получить оценки $\hat{y}_j(t)$, $j = 1, 2, \dots, d-1$:

$$\left\{ \begin{array}{l} \hat{y}_1(t) = \varphi(\hat{x}_1(t), \hat{x}_2(t), c_1), \quad \hat{x}_1(t) = \hat{y}_0(t); \\ \hat{y}_2(t) = \varphi(\hat{y}_1(t), \hat{x}_3(t), c_2, c_1); \\ \dots\dots\dots \\ \hat{y}_{d-1}(t) = \varphi(\hat{y}_{d-2}(t), \hat{x}_d(t), c_{d-1}, c_{d-2}, \dots, c_1) \end{array} \right. \quad (2)$$

и весовых коэффициентов c_j , характеризующих точность объединяемых прогнозов $\hat{y}_{j-1}(t)$, $\hat{x}_{j+1}(t)$ и объединенного $\hat{y}_j(t)$. Вектор текущих весов $C(t) = (c_1(t), c_2(t), \dots, c_{d-1}(t))^T$ описывает качество прогнозирования, достигаемое во втором скрытом слое в каждый текущий момент времени, а изменение соотношений между его элементами уже само по себе свидетельствует об изменении структуры и параметров контролируемой последовательности $x(t)$.

Выходной слой сети образован одним нейроном T_3 . Его выходом являются прогноз $\hat{z}(t)$ и вектор диагностических признаков $P(t) = (p_1(t), p_2(t), \dots, p_{d-1}(t))^T$, элементы которого $p_j(t)$ соответствуют вероятностям того, что “истинное” состояние процесса $x(t)$ наилучшим образом описывается оценкой $\hat{y}_j(t)$. Максимальное значение $p_j(t)$ определяет порядок диагностируемой NARMA-последовательности в текущий момент времени t , а непрерывное уточнение вектора $P(t)$ с помощью соответствующих алгоритмов настройки нейронов позволяет обнаруживать момент “разладки” процесса $x(t)$.

Выходной сигнал j -го нейрона первого скрытого слоя может быть представлен в виде

$$\begin{aligned} \hat{x}_j(t) &= f_j \left(\sum_{i=1}^j w_{ji}(t)x(t-j) + \sum_{i=1}^j \omega_{ji}(t)\hat{x}_j(t-j) + w_{j0}(t) \right) = \\ &= f_j(w_j^T(t)X_j(t)) = f_j(\tilde{X}_j(t)), \end{aligned} \quad (3)$$

где $f_j(\bullet)$ – функция активации j -го нейрона; $w_j(t) = (w_{j0}(t), w_{j1}(t), \dots, w_{jj}(t), \omega_{j1}(t), \dots, \omega_{jj}(t))^T$ – $(2j+1) \times 1$ – вектор настраиваемых синаптических весов; $X_j(t) = (1, x(t-1), \dots, x(t-j), \hat{x}(t-1), \dots, \hat{x}(t-j))^T$ – вектор обобщенных входов; $\tilde{X}_j(t) = w_j^T(t)x_j(t)$; $j = 1, 2, \dots, d$; $t = 1, 2, \dots$ – текущее дискретное время. Выражение (3) представляет собой описание нелинейной стохастической последовательности авторегрессии – скользящего среднего j -го порядка, причем, как отмечалось в [16], именно определение значения j является наиболее сложной проблемой, в связи с чем для решения конкретных задач целесообразно принимать достаточно большие значения d .

Вводя в рассмотрение ошибку прогнозирования j -го нейрона

$$\varepsilon_j(t) = x(t) - \hat{x}_j(t) = x(t) - f_j(\tilde{X}_j(t)), \quad (4)$$

можно записать градиентную процедуру настройки синаптических весов в виде [19; 20]

$$\begin{aligned} w_j(t+1) &= w_j(t) + \eta_j(t)\varepsilon_j(t)\nabla_{w_j} f_j(\tilde{X}_j(t)) = \\ &= w_j(t) + \eta_j(t)\varepsilon_j(t)G_j(t). \end{aligned} \quad (5)$$

Здесь $\eta_j(t)$ – параметр шага поиска, принимаемый чаще всего постоянным; $\nabla_{w_j} f_j(\tilde{X}_j(t)) = G_j(t)$ – градиент функции активации по синаптическим весам.

Заметим, что для наиболее распространенных в теории и практике многослойных нейронных сетей функций активации

$$\begin{cases} f_j'(\tilde{X}_j(t)) = \tanh(\gamma_j \tilde{X}_j(t)) = \frac{1 - e^{-2\gamma_j \tilde{X}_j(t)}}{1 + e^{-2\gamma_j \tilde{X}_j(t)}}; \\ f_j''(\tilde{X}_j(t)) = \frac{1}{1 + e^{-\gamma_j \tilde{X}_j(t)}} \end{cases} \quad (6)$$

градиенты имеют вид

$$\left\{ \begin{aligned} \nabla_{w_j} f_j'(\tilde{X}_j(t)) &= G_j'(t) = \gamma_j(1 - (f_j'(\tilde{X}_j(t))))^2 X_j(t) = \\ &= \gamma_j(1 - \hat{x}_j^2(t)) X_j(t); \\ \nabla_{w_j} f_j''(\tilde{X}_j(t)) &= G_j''(t) = \gamma_j f_j''(\tilde{X}_j(t))(1 - f_j''(\tilde{X}_j(t))) X_j(t) = \\ &= \gamma_j \hat{x}_j(t)(1 - \hat{x}_j(t)) X_j(t). \end{aligned} \right. \quad (7)$$

Сходимость градиентных процедур типа (5) обеспечивается в достаточно широком интервале варьирования параметра шага $\eta_j(t)$: для детерминированного случая данный параметр должен удовлетворять условиям $0 < \eta_j(t) < 2/Lp$ (здесь Lp – константа Липшица для оптимизируемой функции), а для стохастического случая – условиям Дворецкого. При этом вполне естественным представляется выбор шага, обеспечивающего максимальную скорость сходимости.

Максимизация функции

$$W_j(t) = \|\tilde{w}_j(t)\|^2 - \|\tilde{w}_j(t+1)\|^2, \quad (8)$$

где $\tilde{w}_j(t) = w_j - w_j(t)$; w_j – оптимальный вектор синаптических весов, приводит к неконструктивной оценке

$$\eta_j(t) = \frac{(w_j - w_j(t))^T G_j(t)}{\varepsilon_j(t) \|G_j(t)\|^2}. \quad (9)$$

Однако, если выполняется соотношение, справедливое для выпуклых функций,

$$\begin{aligned} (w_j - w_j(t))^T G_j(t) &\leq f_j(w_j^T X_j(t)) - f_j(w_j^T(t) X_j(t)) = \\ &= x(t) - f_j(w_j^T(t) X_j(t)), \end{aligned} \quad (10)$$

из (9) следует, что

$$0 < \eta_j(t) \leq \|G_j(t)\|^2. \quad (11)$$

Рассмотрев одношаговый вариант алгоритма Маквардта [21]

$$w_j(t+1) = w_j(t) + \left(G_j(t) G_j^T(t) + \rho(t) E \right)^{-1} G_j(t) \varepsilon_j(t), \quad (12)$$

где $\rho(t) > 0$; E – единичная матрица, и используя известные соотношения теории псевдообратных матриц

$$\begin{cases} \lim_{\rho(t) \rightarrow 0} \left(G_j(t) G_j^T(t) + \rho(t) E \right)^{-1} = \left(G_j(t) G_j^T(t) \right)^+; \\ \left(G_j(t) G_j^T(t) \right)^+ G_j(t) = \left(G_j^T(t) \right)^+ = G_j(t) \left\| G_j(t) \right\|^{-2}, \end{cases} \quad (13)$$

можно записать оптимальный по быстродействию вариант (5) в виде [22]

$$w_j(t+1) = w_j(t) + \frac{x(t) - \hat{x}_j(t)}{\left\| G_j(t) \right\|^2} G_j(t), \quad (14)$$

совпадающий в линейном случае с алгоритмом настройки синаптических весов Уидроу–Хоффа.

Заметим, что для функций активации типа (6) алгоритм (14) приобретает форму

$$\begin{cases} w_j'(t+1) = w_j'(t) + \frac{x(t) - \hat{x}_j(t)}{\gamma_j (1 - \hat{x}_j(t)) \left\| X_j(t) \right\|^2} X_j(t); \\ w_j''(t+1) = w_j''(t) + \frac{x(t) - \hat{x}_j(t)}{\gamma_j \hat{x}_j(t) (1 - \hat{x}_j(t)) \left\| X_j(t) \right\|^2} X_j(t). \end{cases} \quad (15)$$

Для того чтобы обеспечить дополнительные сглаживающие свойства алгоритму (14), можно ввести следующую экспоненциально взвешенную модификацию:

$$\begin{cases} w_j(t+1) = w_j(t) + r_j^{-1}(t) (x(t) - \hat{x}_j(t)) G_j(t); \\ r_j(t) = \alpha r_j(t-1) + \left\| G_j(t) \right\|^2, \quad 0 \leq \alpha \leq 1, \quad r_j(0) = 1. \end{cases} \quad (16)$$

При $\alpha = 0$ она совпадает с (14), а в линейном случае при $\alpha = 1$ – с процедурой стохастической аппроксимации Гудвина–Рэмеджа–Кейнеса [23].

Во втором слое предлагаемой нейронной сети производится попарное объединение выходных сигналов первого слоя в виде

$$\hat{y}_j(t) = c_j(t) \hat{y}_{j-1}(t) + (1 - c_j(t)) \hat{x}_{j+1}(t), \quad (17)$$

где $\hat{y}_0(t) \equiv \hat{x}_1(t)$; $j = 1, 2, \dots, d-1$, а веса $c_j(t)$ задают сравнительную точность прогнозов $\hat{y}_{j-1}(t)$ и $\hat{x}_{j+1}(t)$ и обеспечивают несмещенность прогноза $\hat{y}_j(t)$.

Чтобы найти значения $c_j(t)$, обеспечивающие оптимальность $\hat{y}_j(t)$, введем $(t \times 1)$ – векторы наблюдений и ошибок

$$\begin{cases} X(t) = (x(1), x(2), \dots, x(t))^T; \\ \hat{Y}_j(t) = (\hat{y}_j(1), \hat{y}_j(2), \dots, \hat{y}_j(t))^T; \\ \hat{X}_j(t) = (\hat{x}_j(1), \hat{x}_j(2), \dots, \hat{x}_j(t))^T; \\ V_j(t) = X(t) - \hat{Y}_j(t); \\ V_{j-1}(t) = X(t) - \hat{Y}_{j-1}(t); \\ V_{x,j-1}(t) = X - \hat{X}_{j-1}(t) \end{cases}$$

и запишем очевидные соотношения

$$V_j(t) = c_j(t)V_{j-1}(t) + (1 - c_j(t))V_{x,j+1}(t); \quad (18)$$

$$\begin{aligned} \|V_j(t)\|^2 &= c_j^2(t)\|V_{j-1}(t)\|^2 + 2c_j(t)(1 - c_j(t))V_{j-1}^T(t)V_{x,j+1}(t) + \\ &+ (1 - c_j(t))^2\|V_{x,j+1}(t)\|^2. \end{aligned} \quad (19)$$

Тогда, решив уравнение

$$\begin{aligned} \frac{\partial \|V_j(t)\|}{\partial c_j(t)} &= c_j(t)\|V_{j-1}(t)\|^2 + V_{j-1}^T(t)V_{x,j+1}(t) - 2c_j(t)V_{j-1}^T(t)V_{x,j+1}(t) - \\ &- \|V_{x,j+1}(t)\|^2 + c_j(t)\|V_{x,j+1}(t)\|^2 = 0, \end{aligned} \quad (20)$$

несложно получить

$$\begin{cases} c_j(t) = V_{x,j+1}^T(t) \frac{V_{x,j+1}(t) - V_{j-1}(t)}{\|V_{x,j+1}(t) - V_{j-1}(t)\|^2}; \\ 1 - c_j(t) = V_{j-1}^T(t) \frac{V_{j-1}(t) - V_{x,j+1}(t)}{\|V_{j-1}(t) - V_{x,j+1}(t)\|^2}. \end{cases} \quad (21)$$

Подставив (21) в (19), найдем

$$\begin{aligned} \|V_j(t)\| = & \left\| \left(\|V_{x,j+1}(t)\|^2 - V_{j-1}^T(t)V_{x,j+1}(t) \right) V_{j-1}(t) + \right. \\ & \left. + \left(\|V_{j-1}(t)\|^2 - V_{j-1}^T(t)V_{x,j+1}(t) \right) V_{x,j+1}(t) \right\|^2 \|V_{j-1}(t) - V_{x,j+1}(t)\|^{-4}. \end{aligned} \quad (22)$$

Можно показать, что

$$\begin{cases} \|V_j(t)\|^2 - \|V_{x,j+1}(t)\|^2 = - \frac{\left(\|V_{x,j+1}(t)\|^2 - V_{j-1}^T(t)V_{x,j+1}(t) \right)^2}{\|V_{j-1}(t) - V_{x,j+1}(t)\|^2} \leq 0; \\ \|V_j(t)\|^2 - \|V_{j-1}(t)\|^2 = - \frac{\left(\|V_{j-1}(t)\|^2 - V_{j-1}^T(t)V_{x,j+1}(t) \right)^2}{\|V_{j-1}(t) - V_{x,j+1}(t)\|^2} \leq 0, \end{cases} \quad (23)$$

т.е. точность объединенного прогноза $\hat{y}_j(t)$ никогда не может быть ниже точности объединяемых прогнозов $\hat{y}_{j-1}(t)$ и $\hat{x}_{j+1}(t)$. Весовой коэффициент $c_j(t)$ задает вклад $\hat{y}_{j-1}(t)$ в $\hat{y}_j(t)$ и тем самым близость реального процесса $x(t)$ к $\hat{y}_{j-1}(t)$ или $\hat{x}_{j+1}(t)$. Изменение значения $c_j(t)$ может служить признаком изменения свойств последовательности $x(t)$, а вектор $C(t) = (c_1(t), c_2(t), \dots, c_{d-1}(t))^T$ может использоваться в качестве вектора диагностических признаков.

Для работы в режиме текущего времени целесообразно представить (21) в рекуррентной форме. Вводя обозначения

$$\begin{cases} E_j(t) = V_{x,j+1}(t) - V_{j-1}(t); \\ v_{j-1}(t+1) = x(t+1) - \hat{y}_{j-1}(t+1); \\ v_{x,j+1}(t+1) = x(t+1) - \hat{x}_{j+1}(t+1); \\ e_j(t+1) = v_{x,j+1}(t+1) - v_{j-1}(t+1), \end{cases}$$

можно записать

$$\left\{ \begin{aligned} c_j(t) &= V_{x,j+1}^T(t) \frac{E_j(t)}{\|E_j(t)\|^2}; \\ c_j(t+1) &= V_{x,j+1}^T(t+1) \frac{E_j(t+1)}{\|E_j(t+1)\|^2} = \\ &= \frac{V_{x,j+1}^T(t)E_j(t) + v_{x,j+1}(t+1)e_j(t+1)}{\|E_j(t)\|^2 + e_j^2(t+1)}, \end{aligned} \right. \quad (24)$$

или окончательно

$$\left\{ \begin{aligned} c_j(t+1) &= \frac{\Gamma_j(t)}{\Gamma_j(t+1)} c_j(t) + \frac{v_{x,j+1}(t+1)e_j(t+1)}{\Gamma_j(t+1)}; \\ \Gamma_j(t+1) &= \Gamma_j(t) + e_j^2(t+1). \end{aligned} \right. \quad (25)$$

В ряде случаев в алгоритме (25) удобнее использовать не сигналы обновлений, а непосредственно контролируруемую последовательность $x(t)$ и ее прогнозы. С учетом того что

$$\begin{aligned} E_j(t) &= X(t) - \hat{X}_{j+1}(t) - X(t) + \hat{Y}_{j-1}(t) = \hat{Y}_{j-1}(t) - \hat{X}_{j+1}(t); \\ e_j(t+1) &= x(t+1) - \hat{x}_{j+1}(t+1) - x(t+1) + \hat{y}_{j-1}(t) = \hat{y}_{j-1}(t) - \hat{x}_{j+1}(t+1), \end{aligned}$$

алгоритм настройки нейронов второго скрытого слоя может быть представлен в виде

$$\left\{ \begin{aligned} c_j(t+1) &= \frac{\Gamma_j(t)}{\Gamma_j(t+1)} c_j(t) + \frac{v_{x,j+1}(t+1)(\hat{y}_{j-1}(t+1) - \hat{x}_{j+1}(t+1))}{\Gamma_j(t+1)}; \\ \Gamma_j(t+1) &= \Gamma_j(t) + (\hat{y}_{j-1}(t+1) - \hat{x}_{j+1}(t+1))^2. \end{aligned} \right. \quad (26)$$

Таким образом уже на уровне двух слоев предлагаемой искусственной рекуррентной нейронной сети может быть обнаружено изменение свойств диагностической стохастической последовательности путем контроля за вариациями элементов $c_j(t)$. Кроме того, во втором слое решается задача оптимального одношагового прогнозирования последовательности $x(t)$.

- Список литературы: 1. *Обнаружение* изменения свойств сигналов и динамических систем: Пер. с англ. / М. Бассвилль, А. Вилски, А. Банвенист и др. М.: Мир, 1989. 278 с. 2. *Romberg T.M., Black J.L., Ledwidge T.J.* Signal Processing for Industrial Diagnostics. Chichester: John Wiley & Sons, 1996. 317 p. 3. *Basseville M., Nikiforov I.* Detection of Abrupt Changes: Theory and Application. Englewood Cliffs, New Jersey: PTR Prentice-Hall, 1993. 528 p. 4. *Kerestencioglu F.* Change Detection and Input Design in Dynamical Systems. Taunton, United Kingdom: Research Studies Press, 1993. 152 p. 5. *Pouliezos A.D., Stravrakakis G.S.* Real Time Fault Monitoring of Industrial Processes. Dordrecht: Kluwer Academic Publishers, 1994. 542 p. 6. *Бодянский Е.В., Руднева И.А.* Об одном адаптивном алгоритме обнаружения разладок в случайных последовательностях // Автоматика и телемеханика. 1995. № 10. С. 101–106. 7. *Бодянский Е.В., Плисс И.П., Соловьева Т.В.* Адаптивное обобщенное прогнозирование многомерных случайных последовательностей // Докл. АН УССР. Сер. А. 1989. № 9. С. 73–75. 8. *Venkatasubramanian V., Chan K.* A neural network methodology for process fault diagnosis // AIChE J. 1989. 35. P. 1993–2002. 9. *Naidu R.S., Zafiriou E., McAvoy T.J.* Use of neural networks for sensor failure detection in a control system // Inst. of electrical and electronic eng. (IEEE) Control Systems Mag. 1990. 10. P. 49–55. 10. *Yamashima H., Kumamoto H., Okumura S., Ikesaki T.* Failure diagnosis of a servovalve by neural networks with new learning algorithm and structure analysis // Intern. J. of Production Research. 1990. 28. N 6. P. 1009 – 1021. 11. *Sorsa T., Koivo H.N., Koivisto H.* Neural networks in process fault diagnosis // IEEE Trans. on System, Man and Cybern. 1991. 21, N 4. P. 815–825. 12. *Ray A.K.* Equipment fault diagnosis – A Neural network approach // Computers in Industry. 1991. 16. P. 169 – 177. 13. *Sorsa T., Koivo H.N.* Application of artificial neural network in process fault diagnosis // Automatica. 1993. 29, N 4. P. 843–849. 14. *Бодянский Е., Воробьев С., Ламонова Н., Штефан А.* Обнаружение изменения свойств стохастических последовательностей на основе искусственных нейронных сетей // АСУ и приборы автоматки. 1997. Вып. 106. С. 75–79. 15. *Connor J.T., Martin R.D., Atlas L.E.* Recurrent neural networks and robust time series prediction // IEEE Trans. on Neural Networks. 1994. 5, N 1. P. 240–254. 16. *Aussem A., Murtagh F., Sarazin M.* Dynamical recurrent neural networks – towards environmental time series prediction // Intern. J. Neural Systems. 1995. 6, N 2. P. 145–170. 17. *Pham D.T., Liu X.* Neural Networks for Identification, Prediction and Control. London: Springer – Verlag, 1995. 238 p. 18. *Chng E.S., Chen S., Mulgrew B.* Gradient radial basis function networks for nonlinear and nonstationary time series prediction // IEEE Trans. on Neural Networks. 1996. 7, N 1. P. 190–194. 19. *Narendra K.S., Parthasarathy K.* Identification and control of dynamical systems using neural networks // IEEE Trans. on Neural Networks. 1990. 1, N 1. P. 4–26. 20. *Narendra K.S.* Adaptive control of dynamical systems using neural networks // Handbook of Intelligent Control: Neural, Fuzzy and Adaptive Approaches / Ed. by D.A. White, D.A. Sofge. New York. 1992. P. 141–183. 21. *Marquardt D.* An algorithm for least squares estimation on nonlinear parameters // SIAM J. Appl. Math. 1963. N 11. P. 431–441. 22. *Бодянский Е.В.* Адаптивные алгоритмы идентификации нелинейных объектов управления // АСУ и приборы автоматки. 1987. Вып. 81. С. 43–46. 3. *Goodwin G.C., Ramadge P.J., Caines P.E.* A globally convergent adaptive predictor // Automatica. 1981. 17, N 1. P. 135–140.

Поступила в редколлегию 17.03.98