

УДК 681.5.015

НЕЙРОСЕТЕВОЙ ПОДХОД К МОНИТОРИНГУ СОСТОЯНИЯ ТЕХНИЧЕСКИХ СИСТЕМ

А.П. ПАШКЕВИЧ, М.М. КОЖЕВНИКОВ, Ф.В. ФУРМАН

*Белорусский государственный университет информатики и радиоэлектроники
П. Бровка, 6, Минск, 220013, Беларусь*

Поступила в редакцию 5 мая 2003

Предложен новый метод мониторинга состояния электромеханических манипуляторов, основанный на анализе результатов идентификации параметров динамической модели при помощи нейронных сетей. Получены аналитические выражения, позволяющие минимизировать влияние погрешностей квантования на точность идентификации. Эффективность предложенного метода подтверждается примерами практического применения в промышленной системе мониторинга

Ключевые слова: нейронные сети, мониторинг состояния, идентификация

Введение

Возрастание сложности технических систем ставит задачу их мониторинга и обнаружения неисправностей в процессе эксплуатации. Несмотря на то что эта задача рассматривалась рядом авторов [2, 3, 5, 9, 14], в настоящее время нет единого метода, позволяющего решить проблему самодиагностики в полном объеме. Поэтому в данной работе предлагается гибридный подход, объединяющий традиционные статистические модели, а также методы, основанные на принципах искусственного интеллекта.

В настоящее время наибольшее распространение получили методы мониторинга состояния технических объектов по контрольным картам [2], в основе которых лежит предположение, что при отсутствии отказов наблюдаемая величина представляет собой некоррелированный статистический процесс [11]. Однако для рассматриваемого класса технических объектов это предположение нарушается. Кроме того, требуется разработка специальных методов идентификации, робастных к шуму измерений и квантованию входных сигналов.

При анализе контрольных карт проводится статистическая проверка ряда гипотез, позволяющая обнаружить необычные вариации контролируемого параметра, обусловленные особыми (неслучайными) причинами. Наиболее часто используются тестовые правила, основанные на выявлении маловероятных "опасных" ситуаций, таких, как выход за контрольные границы, тренд, приближение к контрольным границам и т.д. [7]. Однако, как отмечается рядом авторов [6, 12], такой подход не гарантирует от ошибок. Поэтому проводятся активные исследования по разработке других методов анализа контрольных карт, использующие аппарат экспертных систем, нейронных сетей и эвристические алгоритмы [1, 4, 8, 10, 13], но предложенные в этих работах алгоритмы не учитывают специфику рассматриваемой проблемы.

В данной статье предложен новый нейросетевой метод, который позволяет распознать "опасные" ситуации в контрольных картах промышленных роботов-манипуляторов с инкрементальными датчиками перемещения в цепи обратной связи. В отличие от известных подхо-

дов, основанных на анализе данных, полученных в результате периодических измерений некоторого контролируемого параметра, предложенный метод анализирует характер изменений параметров динамической модели, которые оцениваются путем идентификации в режиме реального времени.

Идентификация параметров модели

Динамика исполнительных приводов электромеханических роботов-манипуляторов может быть описана следующими уравнениями:

$$\dot{p}(t) = v(t); \quad T \cdot \dot{v}(t) + v(t) + m \cdot \text{sgn}(v(t)) = k \cdot u(t), \quad (1)$$

где $p(t)$ — угол поворота вала двигателя (выходная переменная); $v(t)$ — скорость вращения вала; $u(t)$ — управляющее воздействие; k — коэффициент передачи; T — электромеханическая постоянная времени; m — амплитуда возмущающего воздействия. Следует отметить, что в этой модели параметры T и m рассматриваются как основные характеристики состояния объекта, которые далее используются при мониторинге. Дискретизация исходной модели (1) по времени приводит к следующим выражениям:

$$v_{k+1} = \mu \cdot v_k + (1 - \mu) \cdot (k u_k - m); \quad v_k = (p_{k+1} - 2p_k + p_{k-1}) / 2T_0, \quad (2)$$

где $\mu = \exp(-T_0/T)$; T_0 — период квантования по времени. Для идентификации параметров модели первоначально предположим, что коэффициент передачи k и величина скорости вала v_k могут быть оценены с достаточной точностью. Тогда основное уравнение для идентификации может быть записано в виде

$$A_k + m = (B_k + m) \cdot \mu; \quad A_k = v_{k+1} - k u_k; \quad B_k = v_k - k u_k. \quad (3)$$

Определим критерий качества следующим образом

$$F(\mu, m) = \sum_{k=1}^n [(A_k + m) - (B_k + m) \cdot \mu]^2. \quad (4)$$

Приравнявая производные функции $F(\mu, m)$ к нулю и решая соответствующую систему уравнений, получаем выражения для расчета параметров μ, m

$$\mu = \frac{n \sum_{k=1}^n A_k B_k - \sum_{k=1}^n A_k \sum_{k=1}^n B_k}{n \sum_{k=1}^n (B_k)^2 - \left(\sum_{k=1}^n B_k \right)^2}; \quad m = \frac{\mu \sum_{k=1}^n B_k - \sum_{k=1}^n A_k}{n(1 - \mu)}, \quad (5)$$

где n — размер выборки. Если коэффициент передачи k неизвестен, то динамическая модель (2) может быть представлена как линейная регрессия вида $v_{k+1} = a \cdot v_k + b \cdot u_k + c$, где $a = \mu$; $b = (1 - \mu)k$; $c = -(1 - \mu)m$ и соответствующая задача идентификации может быть решена с использованием псевдоинверсии Мура–Пенроуза:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_k v_k^2 & \sum_k v_k u_k & \sum_k v_k \\ \sum_k v_k u_k & \sum_k u_k^2 & \sum_k u_k \\ \sum_k v_k & \sum_k u_k & N \end{pmatrix}^{-1} \cdot \begin{pmatrix} \sum_k v_{k+1} v_k \\ \sum_k v_{k+1} u_k \\ \sum_k v_{k+1} \end{pmatrix}. \quad (6)$$

Тогда искомые параметры модели вычисляются на основе следующих выражений:

$$\mu = a; \quad m = -c/(1-a); \quad k = b/(1-a). \quad (7)$$

Однако проведенные исследования показали, что это выражение менее робастно к шуму измерений по сравнению с выражением (5). Поэтому целесообразно разбить идентификационную процедуру на два шага. На первом шаге производится оценка коэффициента передачи (в установившемся режиме), а на втором осуществляется идентификация параметров T и m , которые необходимы для мониторинга состояния и диагностики отказов робота.

Чтобы использовать выражения (5) и (7) для идентификации, требуется массив скоростей и соответствующих управляющих воздействий $\{v_k, u_k \mid k = 1 : n\}$. Однако рассматриваемые манипуляторы оснащены инкрементальными датчиками, которые позволяют получить только код положения. Поэтому скорость должна оцениваться численно. Очевидно, что простейшая оценка скорости $v_i^{(1)} = (p_{i+1} - p_i)/T_0$ является очень чувствительной к шуму квантования. Чтобы показать это, предположим, что ошибка измерений описывается белым шумом с нулевым математическим ожиданием и дисперсией σ^2 . Тогда среднеквадратическое отклонение оценки скорости имеет вид $\delta[v^{(1)}] = \sigma\sqrt{2}/T_0$, причем $\sigma^2 = \sigma_1^2 + \sigma_2^2$, где $\sigma_1 = \Delta/\sqrt{12}$ обусловлено погрешностями квантования по уровню с шагом Δ , а $\sigma_2 = v_{max} \cdot T_0/\sqrt{12}$ учитывает погрешности квантования по времени при максимальной скорости v_{max} . Соответствующие вычисления при типичных значениях $\Delta = 2\pi/1024$ и $v_{max} = 10 \text{ inc/tic}$ дают $\delta[v^{(1)}]/v_{max} \approx 40\%$ (рис. 1), поэтому для идентификации параметров динамической модели должны использоваться более точные оценки скорости.

Чтобы повысить точность вычисления скорости, применим метод скользящего окна для интервала времени $[t - LT_0; t + LT_0]$. Простейшее выражение такого типа $v_i^{(2)} = (p_{i+L} - p_{i-L})/2LT_0$ использует граничные значения выходной координаты $p(t)$ и дает ошибку $\delta[v^{(2)}] = \sigma\sqrt{2}/2LT_0$. Дальнейшее снижение ошибки может быть достигнуто при использовании линейного сплайна, обеспечивающего сглаживание $p(t)$:

$$v_i^{(3)} = \arg \min_{v,a} \sum_i [a + vt_i - p_i]^2. \quad (8)$$

Дифференцируя это выражение по параметрам a и v и решая соответствующие уравнения, получим

$$v^{(3)} = \bar{S}_{pt} / \bar{S}_{tt}, \quad (9)$$

где

$$\bar{S}_{pt} = \sum_k \left(p_k - n^{-1} \sum_k p_k \right) \cdot \left(t_k - n^{-1} \sum_k t_k \right); \quad \bar{S}_{tt} = \sum_k \left(t_k - n^{-1} \sum_k t_k \right)^2.$$

Далее предположим, что $p(t) = vt + \xi$, где ξ — белый шум с параметрами $0, \sigma$, и рассмотрим скользящее окно $t \in [-LT_0; LT_0]$. Тогда эти суммы равны:

$$S_{pt} = vT_0^2 \sum_{i=-L}^L i^2 + T_0 \sum_{i=-L}^L i \xi_i, \quad S_{tt} = T_0^2 \sum_{i=-L}^L i^2, \quad (10)$$

и искомая оценка скорости может быть представлена как

$$v^{(3)} = v + 3 \sum_{i=-L}^L i \xi_i / T_0 (2L^3 + 3L^2 + L). \quad (11)$$

Очевидно, что среднеквадратическое отклонение $v^{(3)}$ рассчитывается следующим образом:

$$\delta[v^{(3)}] = \sigma / v T_0 \sqrt{2L^3 + 3L^2 + L}. \quad (12)$$

и существенно ниже, чем для оценок $v^{(1)}, v^{(2)}$. Как показали экспериментальные исследования (рис. 2), необходимая для рассматриваемых приложений точность $\delta[v^{(3)}] \leq 0,004$ достигается для скользящего окна с параметром $L \approx 20$.

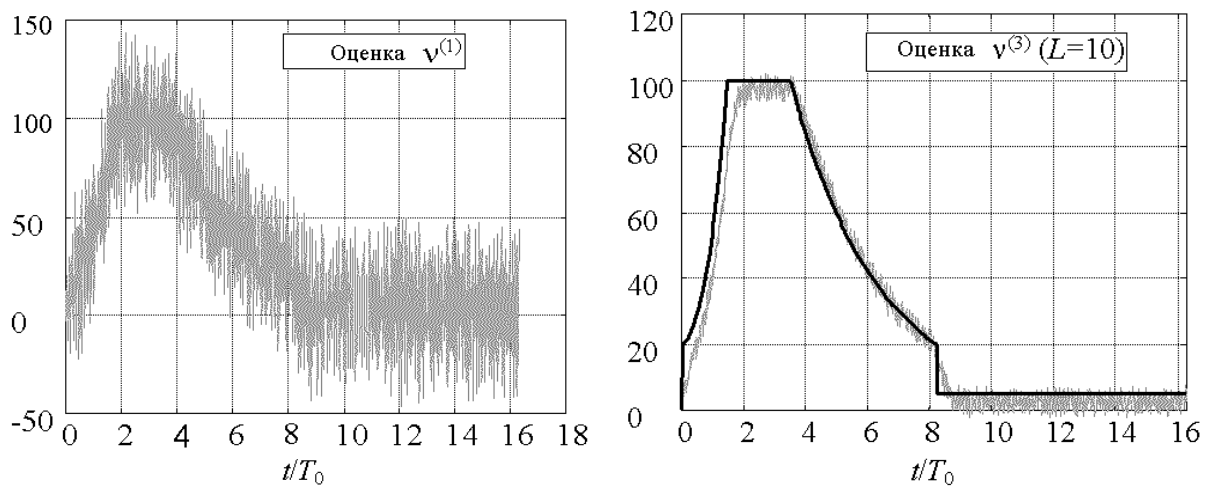


Рис. 1. Сравнение методов оценки скорости

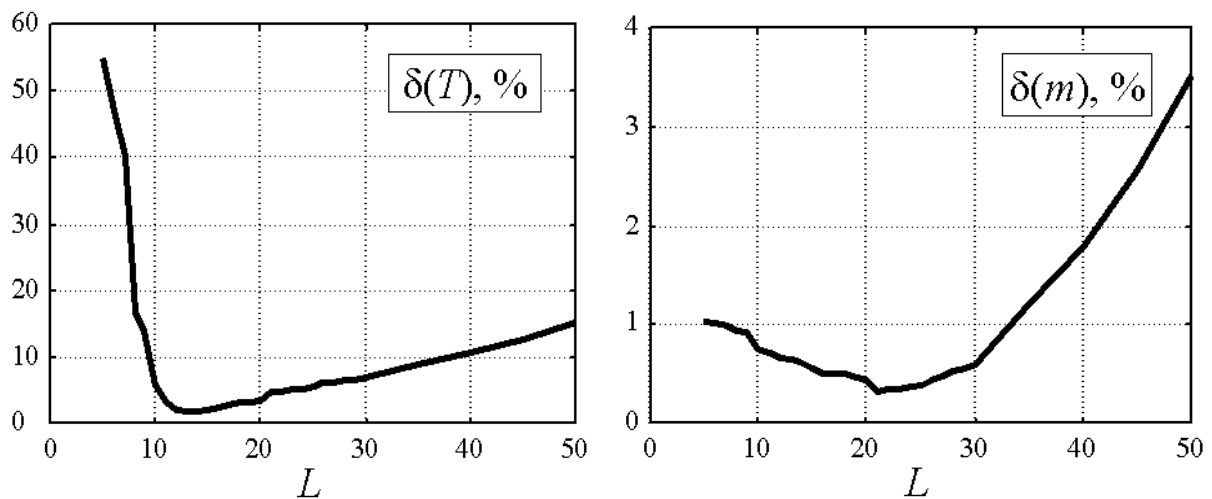


Рис. 2. Ошибки оценок $\delta(T)$ и $\delta(m)$ для разных значений L

Анализ результатов идентификации

Для обнаружения отклонений в динамическом поведении объекта используем контрольные карты типа X, R, S, которые, как показал проведенный анализ, позволяют достичь наилучшей достоверности при мониторинге состояния робота. При этом разделим диапазон изменения анализируемых параметров T, m на три равные зоны (A, B, C) и применим 10 правил,

нарушение которых в режиме нормальной эксплуатации маловероятно. К числу этих правил относятся: "Выход за контрольные границы", "Выход ниже/выше центральной линии" (7 из 7, 10 из 11, 12 из 14, и 14 из 16 точек), "Попадание в зону С" (15 из 15 точек), "Избегание зоны С" (4 из 5 точек), "Точки в зоне А или вне зоны А" (2 из 3 точек), "Линейный тренд" (7 точек), "Колебательный тренд". При этом для снижения вероятности ошибок диагностики эти правила применяются совместно.

Нейросетевой анализатор

Для дальнейшего повышения достоверности мониторинга используем нейронную сеть обратного распространения (FFN), которая позволяет описать существенно нелинейные соотношения между входным и выходным сигналами. Предложенная модификация FFN-сети (рис. 4) представляет собой четырехслойную структуру, содержащую 21 нейрон во входном слое. Этот слой используется для ввода данных от контрольных карт (20 последовательных точек и математическое ожидание M). Выходной слой сети состоит из одного нейрона, а его выход принимает значения 1 или 0 (соответственно для нормального и аварийного состояния робота-манипулятора). Два скрытых слоя сети содержат 40 и 60 нейронов соответственно. Для активации нейронов в скрытых и выходном слоях используется функция гиперболического тангенса. Для генерации требуемого для обучения сети множества контрольных карт использован метод Монте – Карло.

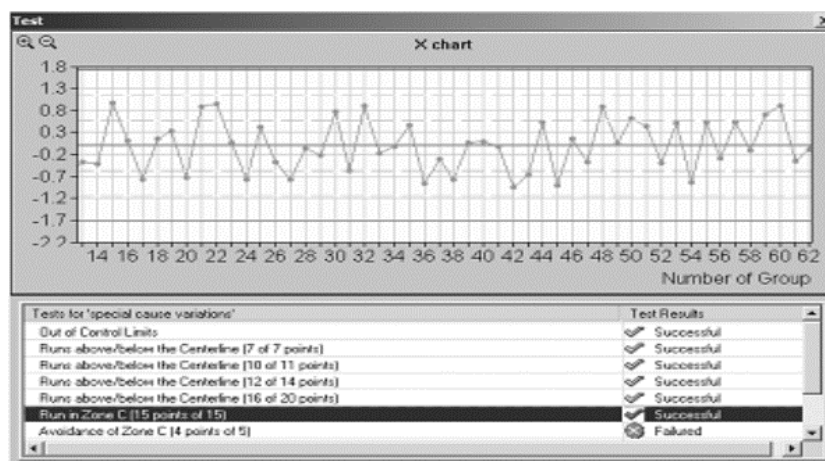


Рис. 3. Пример анализа данных идентификации

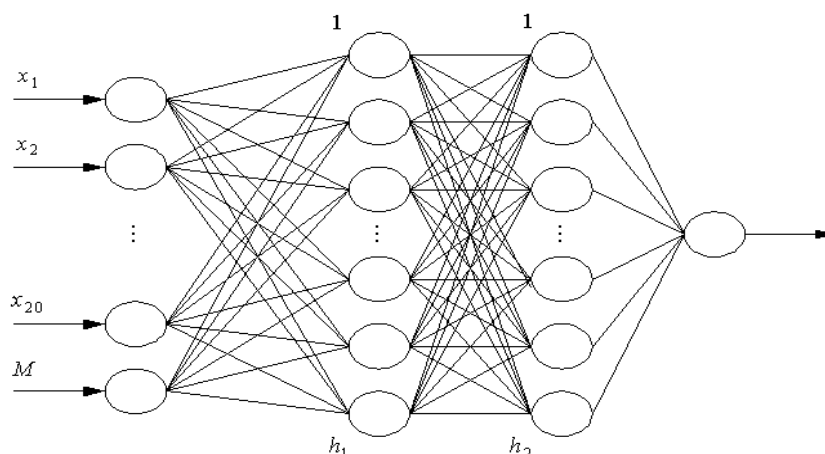


Рис. 4. Нейронная сеть для анализа контрольных карт

Для обучения сети применялся алгоритм обратного распространения с параметром скорости обучения 0,5. В процессе обучения весовые коэффициенты настраивались таким образом,

чтобы минимизировать среднеквадратичную ошибку (MSE) между вектором состояния сети и обучающим множеством. Для обучения было сгенерировано 3800 наборов данных (1000 для нормального функционирования робота и 2800 для различных типов аварийных ситуаций). Веса нейронной сети модифицируются до тех пор, пока не выполняются условия $\varepsilon \leq \varepsilon_{max}$ или $i \leq i_{max}$, где ε — величина MSE , i — количество итераций.

В процессе исследования рассмотрен ряд четырехслойных сетей обратного распространения, отличающихся значениями размерности скрытых слоев h_1 и h_2 , а обучение проводилось до достижения значений $\varepsilon_{max}=10^{-5}$ или $i_{max}=3000$. Графики изменения MSE в процессе обучения для сетей с различной размерностью скрытого слоя приведены на рис. 5 (моделирование проведено в пакете *MatLab 6.0*). Для тестирования этих сетей использовано 140 наборов данных для аварийных режимов и 50 наборов данных для режима нормального функционирования. Результаты обучения и тестирования сетей приведены в таблице (величины достигнутых в процессе обучения ошибок (MSE) и процент корректно распознанных ситуаций на тестовом множестве P). Моделирование показало, что минимальная величина $\varepsilon=0,00015$ достигается при размерности скрытых слоев $h_1=50$ и $h_2=40$. Однако, как следует из таблицы, для случая $h_1=60$ и $h_2=40$, получено большее число корректно распознанных ситуаций (94%), поэтому сеть с такими параметрами применена для диагностики состояния манипуляционного робота в промышленной робототехнической системе, которая включает 5 подвесных манипуляторов, обслуживающих 200 рабочих станций.

Эти исследования проводились в рамках проекта INTAS 2000-757.

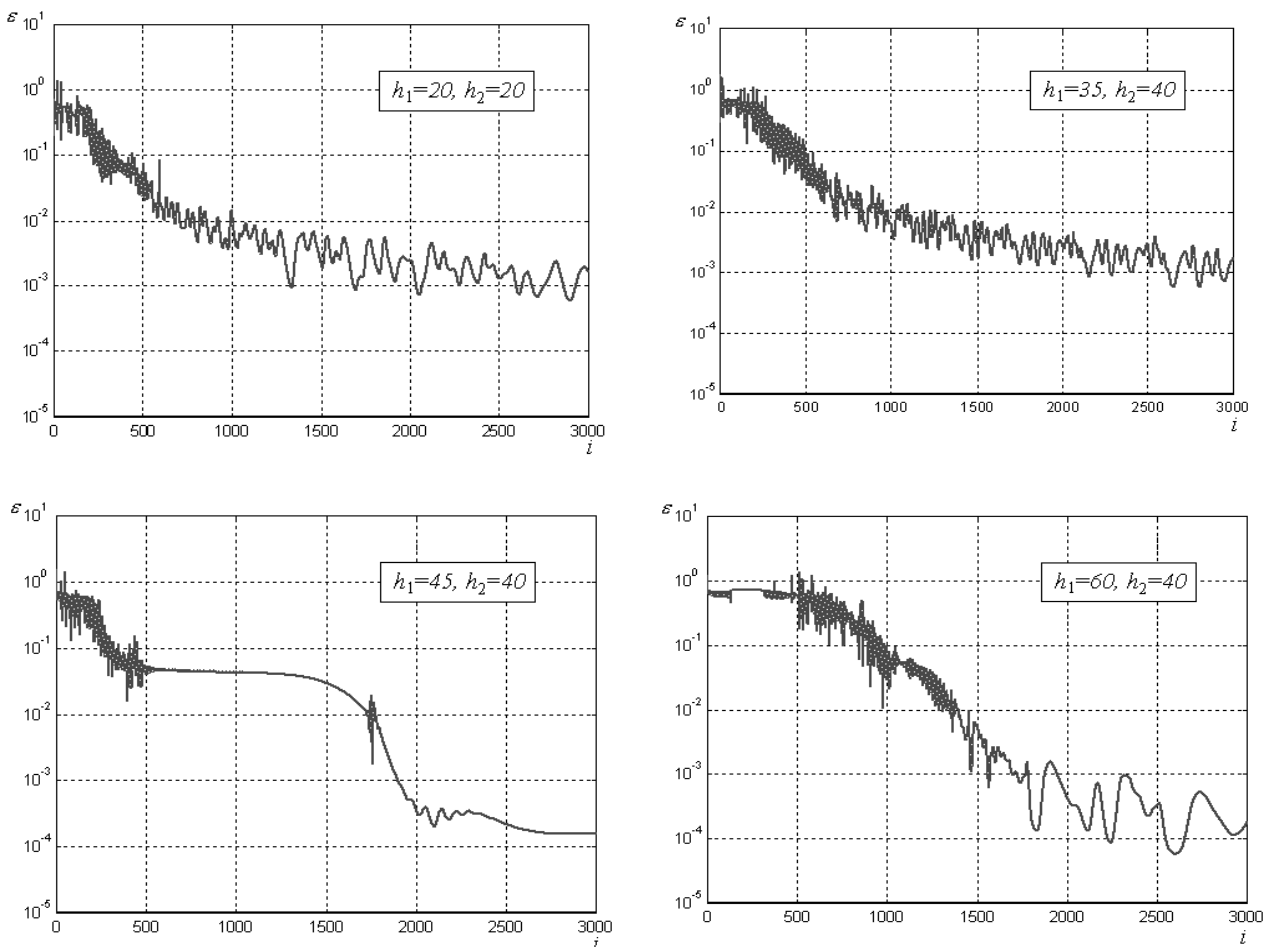


Рис. 5. Визуализация обучения сети при различных параметрах h_1 и h_2

Результаты обучения нейронных сетей

h_1	h_2	ε	Время обучения, с	$P, \%$
20	20	0,00162	300	89,5
25	25	0,000132	330	87,5
30	30	0,00035	408	91,6
35	40	0,00175	420	90,0
40	40	0,00052	480	91,1
45	40	0,00015	510	93,6
50	40	0,00015	528	93,7
60	40	0,00018	600	94,0

Заключение

Предложенный метод мониторинга состояния электромеханических роботоманипуляторов с инкрементальными датчиками обратной связи, основанный на анализе параметров динамической модели при помощи нейронной сети типа FFN, позволяет повысить достоверность обнаружения отказов. На основе предложенного метода разработан программный модуль, эффективность которого подтверждается примерами практического применения при мониторинге состояния промышленной робототехнической системы.

NEURAL NETWORK APPROACH TO CONDITION MONITORING OF TECHNICAL SYSTEMS

A.P. PASHKEVICH, M.M. KAZHEUNIKAU, F.V. FURMAN

Abstract

The paper presents a novel technique for condition monitoring of electromechanical robotic manipulators, which is based on neural network analyses of the dynamic model parameters obtained by means of on-line identification. There were derived analytical expressions that allow minimising impact of the measurement errors on the identification accuracy. Efficiency of the proposed technique has been verified by real-life case studies from industrial monitoring systems.

Литература

1. Anagun A. S. // Proceedings of the 23rd International Conference on Computers and Industrial Engineering. Chicago, IL, 1998. P. 185–188.
2. Basseville M., Nikiforov I.V. Detection of abrupt changes: theory and applications. Englewood Cliffs, NJ: Prentice-Hall. 1993.
3. Chang S. I. // Proceedings of ASQC 49th Annual Quality Congress. Cincinnati, OH, 2000. P. 725–732.
4. Cheng C.S. // International Journal of Production Research. 1997. № 35. P. 667–697.
5. Dash S., Venkatasubramanian V. // Proceedings of the conference on Process Systems Engineering. Keystone, Colorado, 2000. P. 785–791.
6. Frank P. // Automatica. 1990. № 26(3). P. 459–474.
7. Grant E.L., Leavenworth R.S. Statistical quality control. New York: McGraw-Hill. 1996.
8. Hwang H.B., Hubele N.F. // IIE Transactions. 1993. № 25. P. 27–40.
9. Kulikov G.G., Breikin, T.V., Arkov V.Y., Fleming, P.J. // Proceedings of the International Gas Turbine Congress. Kobe, Japan, 1999. P. 949–952.
10. Lucy-Bouler T.L. // Proceedings of WCNN'93—World Congress on Neural Networks. Portland, UK, 1993. P. 631–633.
11. Montgomery D.C. Introduction to Statistical Quality Control. New York: John Wiley and Sons, Inc. 1996.
12. Patton R.J., Frank P.M., Clark R.N. Fault Diagnosis in dynamic systems: Theory and Applications. Prentice Hall. 1989.
13. Pham D.T., Oztemel E. // International Journal of Production Research. 1994, № 32. P. 721–729.
14. Tsung, F. // International Journal of Production Research. 2000, № 38. P. 625.