

УДК 519.71

А.А. Бессонов, С.О. Руденко

## РОБАСТНАЯ МНОГОКРИТЕРИАЛЬНАЯ ИДЕНТИФИКАЦИЯ НЕЛИНЕЙНЫХ ОБЪЕКТОВ С ПОМОЩЬЮ СЕТЕЙ ПРЯМОГО РАСПРОСТРАНЕНИЯ

**Введение.** Задача идентификации нелинейного нестационарного динамического объекта, описываемого моделью

$$y(k) = f(\mathbf{x}(k), k) + \xi(k), \tag{1}$$

где  $\mathbf{x}(k) = [y(k-1), \dots, y(k-l), u(k-1), \dots, u(k-m)]^T$  -  $N \times 1$  вектор обобщенного входного сигнала ( $N = l + m$ );  $y(i)$ ,  $u(i)$  - выходной и входной сигналы объекта в момент времени  $i$  соответственно;  $l$  и  $m$  - порядки запаздывания по выходному и входному каналам соответственно;  $f(\bullet)$  - неизвестная нелинейная функция;  $\xi(k)$  - помеха, состоит в получении оценки функции  $f(\bullet)$  по измерениям входных и выходных переменных.

При нейросетевом подходе задача идентификации заключается в обучении сети, состоящем в определении вектора ее параметров  $\theta$  (весов, параметров активационных и базисных функций и т.д.), обеспечивающего минимум функционала

$$F(e) = \frac{1}{k} \sum_{i=1}^k \rho(e(i, \theta)), \tag{2}$$

где  $\rho(e(i, \theta))$  - некоторая функция потерь, зависящая от вида закона распределения помехи  $\xi$ ;  $e(i) = y(i) - \hat{y}(i)$ ;  $\hat{y}(i)$  - выходной сигнал модели.

Среди ИНС наибольшее распространение при решении задачи идентификации нелинейных динамических объектов (1) в настоящее время получили многослойный персептрон (МП) и радиально-базисные сети (РБС), использующие соответственно аппроксимации нелинейного оператора  $f(\bullet)$  вида

$$\hat{y}(k) = \hat{f}(k) = f^q \left[ (W^q)^T f^{q-1} \left[ (W^{q-1})^T f^{q-2} \left[ \dots f^1 \left[ (W^1 x(k) + b_1 \right] \dots \right] \right] \right] + b_q, \tag{3}$$

и

$$\hat{y}(k) = \hat{f}(k) = w_0 + \sum_{i=1}^N w_i \Phi_i(x) = w_0 + W^T \Phi(k), \tag{4}$$

где  $W^i$  - вектор весовых параметров нейронов  $i$ -го слоя сети;  $f^i[\bullet]$  - активационная функция (АФ)  $i$ -го слоя,  $b_i$  - смещение  $i$ -го нейрона;  $\Phi(k)$  - вектор выбранных базисных функций (БФ).

Так как получаемая математическая модель, с одной стороны, должна быть достаточно простой и удобной для использования ее в задачах прогнозирования, управления и т.д., а с другой - наиболее полно отражать свойства исследуемого объекта, ее качество определяется некоторым набором критериев, т.е. задача идентификации является многокритериальной.

**Задача многокритериальной оптимизации.** Задача многокритериальной оптимизации (МО), которую также часто называют мультикритериальной или векторной оптимизацией, заключается в нахождении такого вектора решений  $\mathbf{x}^* = [x_1^*, x_2^*, \dots, x_n^*]^T$ , который бы оптимизировал вектор целевых функций

$$F(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x})]^T \tag{5}$$

при наличии  $m$  ограничений в виде неравенств

$$g_i(\mathbf{x}) \leq 0, i = \overline{1, m}, \tag{6}$$

и  $p$  ограничений в виде равенств

$$h_j(\mathbf{x}) = 0, j = \overline{1, p}, \tag{7}$$

где  $\mathbf{x}^* \in \mathfrak{R}^n$  - вектор решений,  $F(\mathbf{x}) \in \mathfrak{R}^k$  - вектор целевых функций, каждая из которых должна быть оптимизирована (обычно полагают, что все целевые функции должны быть минимизированы). При многокритериальной минимизации на основе подхода Парето используется, как отмечалось выше, понятие доминирования.

Вектор  $\mathbf{u} = [u_1, u_2, \dots, u_n]^T \in \mathfrak{R}^k$  доминирует над вектором  $\mathbf{v} = [v_1, v_2, \dots, v_n]^T \in \mathfrak{R}^k$  (обозначается  $\mathbf{u} \prec \mathbf{v}$ ) тогда и только тогда, когда  $\forall i \in \{1, 2, \dots, k\}, u_i \leq v_i \wedge \exists j \in \{1, 2, \dots, k\} : u_j < v_j$ . Другими словами, существует как минимум одна компонента вектора  $\mathbf{u}$  ( $u_j$ ), которая меньше, чем  $v_j$ , в то время как остальные компоненты вектора  $\mathbf{u}$  меньше либо равны соответствующим компонентам вектора  $\mathbf{v}$ .

Точка  $\mathbf{x}^* \in \Omega$  ( $\Omega$  некоторая область пространства  $\mathfrak{R}^n$ , удовлетворяющая условиям (6)-(7)) является Парето-оптимальной по отношению ко всем  $\mathbf{x} \in \Omega$  тогда и только тогда, когда  $F(\mathbf{x}^*) \prec F(\mathbf{x})$ , т.е. решение  $\mathbf{x}^*$  является Парето-оптимальным если не может быть найдено никакое другое решение, которое бы доминировало над  $\mathbf{x}^*$  с учетом определения доминирования по Парето.

В общем виде алгоритм поиска фронта Парето с помощью ЭА может быть записан следующим образом.

1. Сгенерировать начальную обучающую выборку, состоящую из векторов входных переменных  $\mathbf{x} \in \Omega$ . Вычислить векторы значений целевых функций  $F(\mathbf{x})$  для всех  $\mathbf{x}$ .
2. На основе обучающей выборки  $\mathbf{x} \in \Omega$  и соответствующих значений  $F(\mathbf{x})$  построить модели всех целевых функций  $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x})$ .
3. На основе полученных моделей  $f'_1(\mathbf{x}), f'_2(\mathbf{x}), \dots, f'_k(\mathbf{x})$  с помощью выбранного алгоритма определить Парето-оптимальное множество решений  $PF^*$ .
4. В точках полученного множества решений  $PF^*$  вычислить точные значения функций  $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x})$ . Если критерий останова (получена требуемая точность моделей и построен фронт Парето либо осуществлено максимально допустимое число итераций) не выполняется, то все полученные значения добавляются в обучающую выборку и осуществляется возврат к шагу 2, на котором уточняются модели целевых функций.

**Робастные эволюционирующие ИНС.** При переходе от ИНС к ЭИНС для всех типов сетей используются общие эволюционные процедуры (инициализация популяции, оценка популяции, селекция, скрещивание, мутации), а различия заключаются лишь в способе кодирования структуры и параметров той или иной ИНС в виде хромосомы.

В начале работы НА случайным образом инициализируется популяция  $P_0$ , состоящая из  $N$  особей (ИНС):  $P_0 = \{H_1, H_2, \dots, H_N\}$ . Каждая особь в популяции при этом получает свое уникальное описание, закодированное в хромосоме  $H_j = \{h_{1j}, h_{2j}, \dots, h_{Lj}\}$ , которая состоит из  $L$  генов, где  $h_{ij} \in [w_{\min}, w_{\max}]$  - значение  $i$ -го гена  $j$ -ой хромосомы ( $w_{\min}$  - минимальное, и  $w_{\max}$  - максимальное допустимые значения соответственно). На рис.1 показаны примеры МП и РБС, форматы хромосом и соответствие между параметрами сетей и хромосомами (представление их параметров в хромосоме). Следует отметить, что длина хромосомы зависит от размерности идентифицируемого объекта и максимально допустимого количества нейронов.

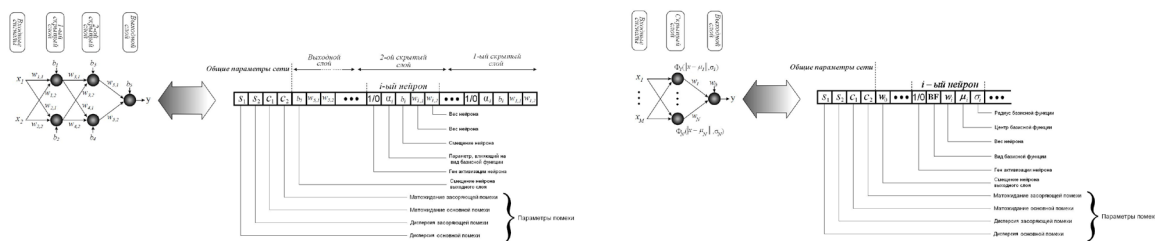


Рис. 1 – Форматы хромосом МП и РБС

Как видно из рисунков, каждая хромосома состоит из генов, в которых хранится информация о соответствующих параметрах сети. Подробное описание формата хромосом для МП и РБС приводится в работе [1].

После того как начальная популяция сформирована, производится оценка приспособленности каждой входящей в нее особи, определяемой некоторой функцией приспособленности (фитнес-функцией). При идентификации параметров нелинейных систем с помощью ГА в них традиционно используется в качестве фитнес-функции квадратичный функционал

$$\rho(e(i, \theta)) = 0.5e^2(i, \theta). \tag{8}$$

Следует отметить, что фитнес-функция служит двум основным целям: во-первых, она оценивает, насколько нейросетевая модель соответствует реальной системе, и, во-вторых, она должна быть способна

устранить влияние незначительных (или зашумленных) измерений на систему идентификации, так как в противном случае получение адекватной модели нелинейного объекта будет весьма проблематичным.

Если информация о помехе недоступна или помеха засорена негауссовским шумом, использование квадратичной фитнес-функции вследствие ее неробастности не представляется возможным. Однако существует большое количество таких функций, обеспечивающих получение робастных оценок при наличии помех, распределения которых отличны от нормального, например, следующие

$$F(e) = \begin{cases} \frac{e^2}{2}, & |e| \leq c; \\ c|e| - \frac{c^2}{2}, & c < |e|; \end{cases} \quad (9)$$

$$F(e) = \arctg |e|^\lambda. \quad (10)$$

Классические робастные методы, минимизирующие неквадратичные функционалы, ориентированы на симметричность засорения, когда выбросы одинаково часто появляются как в области отрицательных, так и в области положительных значений.

Указанные методы позволяют эффективно бороться с помехами, описываемыми моделью Тьюки-Хьюбера [2]

$$p(x) = (1 - \varepsilon)p_0(x) + \varepsilon q(x), \quad (11)$$

где  $p_0(x)$  – плотность соответствующего основного распределения  $N(0, \sigma_1^2)$ ;  $q(x)$  – плотность засоряющего (произвольного) распределения  $N(0, \sigma_2^2)$ ,  $\sigma_1^2 \ll \sigma_2^2$ ;  $\varepsilon \in [0, 1]$  – параметр, характеризующий степень засорения основного распределения.

**Выбор модели.** Выбор оптимального решения из фронта Парето является наиболее важным этапом всей процедуры идентификации нелинейного объекта с помощью ЭРБС. Так как фронтом Парето обычно представлен широкий набор возможных оптимальных решений, окончательный выбор модели должен быть достаточно точным и робастным.

В последнее время в литературе по статистическому выбору модели значительное внимание уделяется принципу минимальной длины описания (МДО, Minimum Description Length, MDL), являющемуся весьма эффективным подходом при выборе модели и решении других проблем статистической обработки информации, а также обобщающему методы максимального правдоподобия и информационные критерии. Ключевым компонентом в принципе МДО является стохастическая сложность, введенная и использованная в работах [3-6], оценивающая насколько хорошо вероятностная модель соответствует статистическим закономерностям данных. Из этих работ следует, что стохастическая сложность наблюдаемых данных по отношению к данному классу параметрических моделей может быть выражена следующим образом:

$$SC(\theta) = -\log \rho(\theta) + \frac{1}{2} \log |I(\theta)| + \sum_{i=1}^N \log (|\hat{\theta}_i| + K^{-1/4}), \quad (12)$$

где  $I(\theta) = -\frac{\partial^2 \log \rho}{\partial \theta \partial \theta^T}$  – информационная матрица  $N \times N$ ;  $\hat{\theta}$  – М-оценка вектора  $\theta$ , полученная путем

минимизации (2);  $K$  – количество математических моделей.

Таким образом, штрафная излишнюю сложность модели, МДО обосновывает идею максимизации регуляризованного правдоподобия, т.е. позволяет обосновать корректность регуляризации правдоподобия.

**Моделирование.** Проводилась многокритериальная идентификация сильнозашумленного нелинейного стационарного объекта

$$y(k) = 0.725 \sin \left( \frac{16u(k-1) + 8y(k-1)}{(3 + 4u^2(k-1) + 4y^2(k-1))} \right) + 0.2u(k-1) + 0.2y(k-1), \quad (13)$$

при наличии помехи  $\xi(k)$ , описываемой моделью (16)

$$\xi(k) = (1 - \varepsilon)q_1(k) + \varepsilon q_2(k),$$

где  $\varepsilon = 0,1$ ;  $q_1(k)$ ,  $q_2(k)$  – нормально распределенные помехи с математическими ожиданиями  $m_1 = m_2 = 0$  и дисперсиями  $\sigma_1 = 0,6$ ;  $\sigma_2 = 12$  соответственно.

Зашумленная таким образом поверхность (13) приведена на рис.2-б. Для решения данной задачи использовалась фитнес-функция (10), популяция состояла из 150 особей (сетей). Максимально

возможное количество нейронов в каждом скрытом слое МП сети было ограничено 10. Результирующие фронты Парето и восстановленные поверхности некоторыми входящими в него сетями после 1000 эпох обучения представлены на рис.3. Квадратиком обозначена особь, для которой критерий Акаике является минимальным.

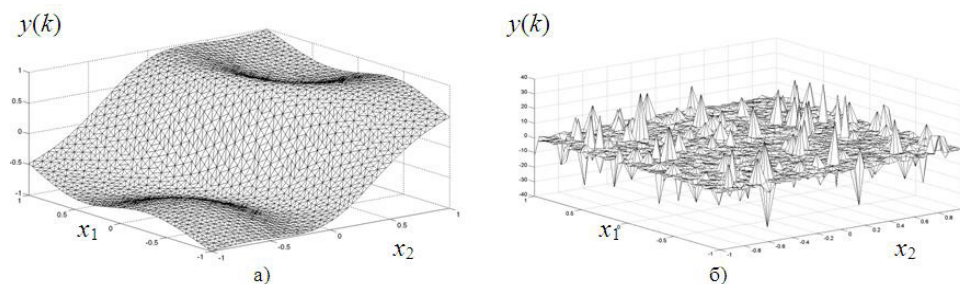


Рис. 2

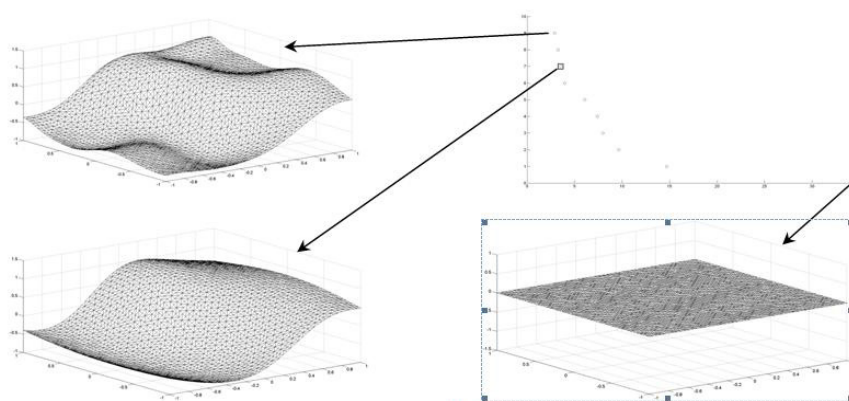


Рис. 3

**Заключение.** Эффективность использования эволюционного подхода для решения задачи многокритериальной нейросетевой идентификации нелинейных объектов при наличии негауссовских помех, т.е. получение Парето-оптимального решения, существенно возрастает при применении робастных фитнес-функций. Для выбора оптимального решения из получаемого фронта Парето следует использовать какой-либо робастный информационный критерий.

ЛИТЕРАТУРА:

1. Руденко О.Г., Бессонов А.А. Идентификация нелинейных нестационарных объектов с помощью эволюционирующей радиально-базисной сети // Проблемы управления и информатики – 2012. – №4. – С. 5-14.
2. Хьюбер П. Робастность в статистике. – М.:Мир, 1984. – 304 с.
3. Rissanen J. Order estimation by accumulated prediction errors // In Essay in Time Series and Allied Processes (Gani J. and Priestley H.B., eds) // J. Appl. Probab. – 1989. – 23 A. – P. 55-61.
4. Qian G., Künsch H.R. On model selection in robust linear regression // Research Report 80, ETH Zentrum, Zurich, 1996. – 33 p.
5. Qian G. Computing minimum description length for robust linear regression model selection // Pacific Symposium of Biocomputing, 1999: Big Island of Hawaii, USA. – 1999. – 4. – P.314-325.
6. Akaike H.A. A new look at statistical model identification // IEEE Trans. on Automatic Control. - 1974. – 19. – P.716-723.

БЕССОНОВ Александр Александрович – к.т.н., доцент, кафедра ЭВМ, Харьковский национальный университет радиоэлектроники

Научные интересы: искусственные нейронные сети, искусственные иммунные системы

РУДЕНКО Сергей Олегович – аспирант, кафедра ЭВМ, Харьковский национальный университет радиоэлектроники

Научные интересы: искусственный интеллект, интеллектуальные системы