

УДК 681.518.25

А.И. Михалев, Н.В. Лысая, Д.А. Лысый

## СИНЕРГЕТИЧЕСКИЙ СИНТЕЗ МНОГОМОДЕЛЬНОГО АЛГОРИТМА ОБУЧЕНИЯ НЕЧЕТКИХ НЕЙРОСЕТЕЙ

Нечеткие нейросетевые (гибридные) методы могут рассматриваться как основа для проведения синергетического синтеза эффективных алгоритмов моделирования поведения сложных динамических систем [2,4]. В свою очередь, они, зачастую, оказываются и более эффективными в задачах прогнозирования поведения существенно нелинейных объектов, в то время как классические методы не обеспечивают приемлемого качества прогноза, а нейронные сети требуют слишком много времени на обучение. Это свойство нечетких систем может быть использовано для выявления закономерностей в стохастических процессах и предсказания их поведения в будущем.

Ставится задача синергетического синтеза алгоритма обучения гибридных (fuzzy-neuro) моделей производственных процессов при отсутствии априорной информации о входах и выходах исследуемого объекта.

Пусть динамика производственного процесса описывается уравнением вида:

$$y(k) = F(x(k)) + \eta(k), \quad (1)$$

где  $x(k) = (u(k-1), \dots, u(k-N_b), y(k-1), \dots, y(k-N_a))^T = (x_1(k), \dots, x_{N_v}(k))^T$  - обобщенный вектор входов,  $u(k)$ ,  $y(k)$ ,  $\eta(k)$  - сигналы входов, выходов и возмущений соответственно в дискретные моменты времени  $k = 0, 1, 2, \dots$ ;  $N_a, N_b$  - порядок модели по выходу и входу соответственно.

В соответствие прогнозируемому поведению процесса предлагается ставить комбинированную модель, состоящую из  $N$  нечетких нейросетей одинаковой структуры - модели Такаги-Сугено первого порядка в виде 5-слойной нечеткой нейросети. При этом в 1-м слое сети выполняется фаззификация входов; во 2-м слое вычисляются степени выполнения правил; 3-й слой выполняет нормализацию степеней выполнения правил; в 4-м слое вычисляются

значения консеквентов и значения выходов правил; в 5-м слое выполняется дефаззификация (суммирование значений выходов правил) и вычисляется значение выхода сети.

В качестве функций принадлежности входных переменных нейросети выбирались функции Гаусса:

$$\mu = \exp\left(-\frac{(x - c)^2}{2\sigma^2}\right), \quad (2)$$

где  $c$ ,  $\sigma$  - центр и ширина функции принадлежности входа соответственно.

Настройка параметров сети осуществлялась с помощью алгоритма Марквардта, модифицированного следующим образом [1]:

$$\begin{cases} \Omega(k) = \Omega(k-1) + q^{-1}(k)J(k)e(k), \\ q(k) = \alpha q(k-1) + \|J(k)\|^2, \quad 0 \leq \alpha \leq 1. \end{cases} \quad (3)$$

где  $\Omega(k)$  - вектор настраиваемых параметров функций принадлежности входов нейросети,  $J(k)$  - вектор, содержащий градиенты сети по настраиваемым параметрам antecedентов всех правил в момент времени  $k$ ,  $e(k) = y(k) - \hat{y}(k)$  - ошибка идентификации на итерации  $k$ .

Процедуру настройки параметров консеквентов можно представить как линейную модификацию процедуры (3):

$$\begin{cases} \Theta(k) = \Theta(k-1) + r^{-1}(k)e(k)\Psi(k), \\ r(k) = \beta r(k-1) + \|\Psi(k)\|^2, \quad 0 \leq \beta \leq 1. \end{cases} \quad (4)$$

здесь  $\Theta(k)$  - вектор настраиваемых параметров консеквентов,  $\Psi(k) = (\bar{w}_1(x(k))\tilde{x}(k), \dots, \bar{w}_{n_R}(x(k))\tilde{x}(k))^T$  - дополнительный вектор входов,  $\bar{w}_i(x)$  - нормализованные степени выполнения правил,  $\tilde{x} = (x^T, 1)^T$  - расширенный вектор входов.

Для повышения скорости обучения и качества прогноза в работе предлагается осуществлять обучение гибридных моделей с применением синергированного подхода [4, 5].

Применим принцип многомодельного подхода [6] к синтезу алгоритма прогнозирования характеристик производственного процесса с помощью нечетких нейросетей. Для этого в соответствие прогнозируемому поведению процесса ставится комбинированная модель, состоящая из  $N$  нечетких нейросетей (моделей Такаги-Сугено первого порядка), параметры antecedентов которых настраиваются с

помощью методов нелинейной оптимизации, а параметры консеквентов – в контурах параметрической настройки идентификаторов. Кроме того, в процессе настройки параметров сетей применяется подход, сочетающий в себе адаптивную и синергичную стратегии настройки. Адаптивный подход позволяет осуществить настройку параметров каждой из  $N$  нечетких нейросетей на основе текущей информации с целью достижения оптимального критерия качества прогноза, а синергичный подход дает возможность реализовать обмен информацией между сетями с целью построения наиболее эффективной стратегии оптимизации этого критерия.

Вектор параметров antecedентов моделей  $\Omega(k)$  преобразуется в матрицу  $\Omega_i(k)$ ,  $i = \overline{1, N}$ . При этом выражение для выхода каждой из  $N$  моделей примет вид:

$$\mathfrak{f}_i = \Theta_i^T \Psi_i,$$

где  $\Theta_i$  - матрица параметров консеквентов правил,  $\Psi_i$  - дополнительный вектор входов  $i$ -й модели.

Для настройки параметров antecedентов и консеквентов используются процедуры (3), (4), модифицированные для настройки параметров  $N$  моделей:

$$\begin{cases} \Omega(k) = \Omega(k-1) + q_i^{-1}(k) J_i(k) e_i(k), \\ q_i(k) = \alpha_i q_i(k-1) + \|J_i(k)\|^2, \quad 0 \leq \alpha_i \leq 1. \end{cases} \quad i = \overline{1, N}, \quad (5)$$

$$\begin{cases} \Theta(k) = \Theta(k-1) + r_i^{-1}(k) e_i(k) \Psi_i(k), \\ r_i(k) = \beta_i r_i(k-1) + \|\Psi_i(k)\|^2, \quad 0 \leq \beta_i \leq 1. \end{cases} \quad i = \overline{1, N}. \quad (6)$$

Здесь  $e_i(k) = y(k) - \mathfrak{f}_i(k)$  - ошибка  $i$ -й модели.

Как следует из (5) и (6), каждая модель обладает своими скалярными коэффициентами «забывания»  $\alpha_i, \beta_i$  и своими параметрами  $q_i(k), r_i(k)$ , которые обеспечивают независимость поисковых движений в процессе настройки параметров  $i$ -й модели.

Критериями процесса синергирования моделей выступают четные функции ошибок в виде функции качества:

$$\begin{aligned} f_i(k) &= \exp(-\gamma e_i^2(k)) \text{ или} \\ \mathfrak{f}_i(k) &= \exp[f_i(k)], \quad \gamma > 0, \quad i = \overline{1, N}, \end{aligned} \quad (7)$$

где  $\gamma$  - весовой коэффициент,  $\gamma = const$ .

Выбор оптимальных на шаге  $k$  параметров каждой из  $N$  моделей осуществляется по синергичной стратегии в соответствии с алгоритмом [3]:

$$\Xi_i(k) = \begin{cases} \Xi_r(k) & \text{при } \forall \Delta f_i(k-1) \leq 0 \text{ и } |f_r(k-1) - f_{\ominus}| = \min |f_i(k-1) - f_{\ominus}|, \\ \Xi_i(k) & \text{при } \Delta f_i(k-1) > 0 \text{ и } \forall \Delta F_p(k-1) \leq 0, (p \neq i) \text{ или } \Delta^2 f_i(k-1) > 0, \\ \Xi_s(k) & \text{при } \forall \Delta f_i(k-1) \leq 0 \text{ и } \Delta f_s(k-1) > 0, (s \neq i). \end{cases}$$

Здесь  $\Xi_i(k) = (\Omega_i(k), \Theta_i(k))$  – набор параметров antecedентов и консеквентов  $i$ -й модели;  $\Delta^2 f_i(k-1)$  – вектор приращения второго порядка функции качества настройки  $i$ -й модели;  $f_{\ominus}$  – экстремальное значение функций  $f_i(t)$ ;  $|f_i(k-1) - f_{\ominus}| \leq \varepsilon$  – критерий выбора нелинейной модели, настроенной наилучшим образом, при условии неудачной настройки параметров на  $k-1$  шаге для всех  $N$  нейросетей ( $\forall \Delta f_i(k-1) \leq 0$ ). При этом, если наилучшей моделью на шаге  $k$  оказалась нейросеть с индексом  $i=m$ , т.е. норма  $\|\Xi_i(k) - \Phi(k)\| \leq \varepsilon$ , ( $\forall \varepsilon > 0$ ), где  $\Phi(k)$  – набор параметров прогнозируемого процесса, то в соответствии с синергичной стратегией в этот момент произойдет переключение всех нелинейных моделей на модель с индексом  $m$ :

$$\Omega_i(k) = \Omega_m(k), \quad \Theta_i(k) = \Theta_m(k), \quad i = \overline{1, N}. \quad (8)$$

Предложенный алгоритм (5) - (8) реализует как возможность независимого поиска критерия оптимальности, который совершает каждая из  $N$  нечетких нейросетей (выражения (5),(6)) в процессе настройки своих параметров, так и одновременное эффективное взаимодействие всех  $N$  нелинейных моделей для оптимальной организации процесса настройки параметров (выражения (7) и (8)).

На рис. 1 - 3 представлены графики процесса настройки параметров нечетких нейросетей, обучающихся независимо по алгоритму (3) - (4) и этих же нейросетей, обучавшихся с применением синергированного подхода (алгоритм (5) – (6)), при прогнозировании зависимости процесса выплавки ферросплавов «Кратность шлака =  $f(\text{Ств}/\text{Mn})$ ». При сопоставлении графиков заметно изменение процесса настройки параметров сети, обучающейся без синергии и сети, обучающейся с применением синергированного подхода.

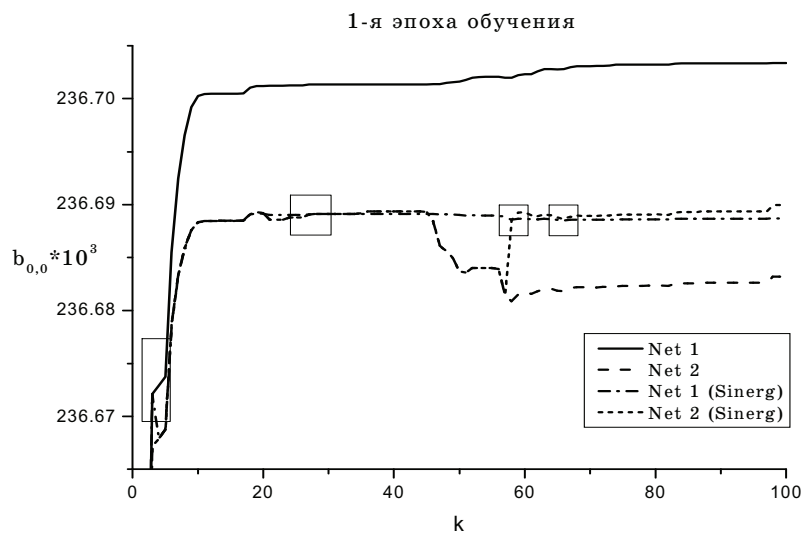


Рисунок 1 - Процесс настройки параметра  $b_{0,0}$  (центра функции принадлежности 1-го входа) нечеткой нейросетевой модели: изменение параметра на 1-й эпохе обучения.

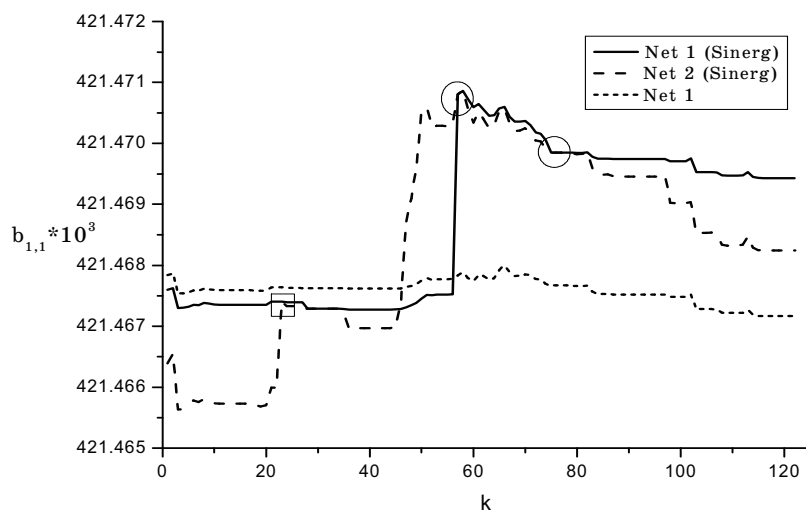


Рисунок 2 - Процесс настройки параметра  $b_{1,1}$  (центра функции принадлежности 2-го входа) нечетких нейросетей в течение 2-й эпохи обучения

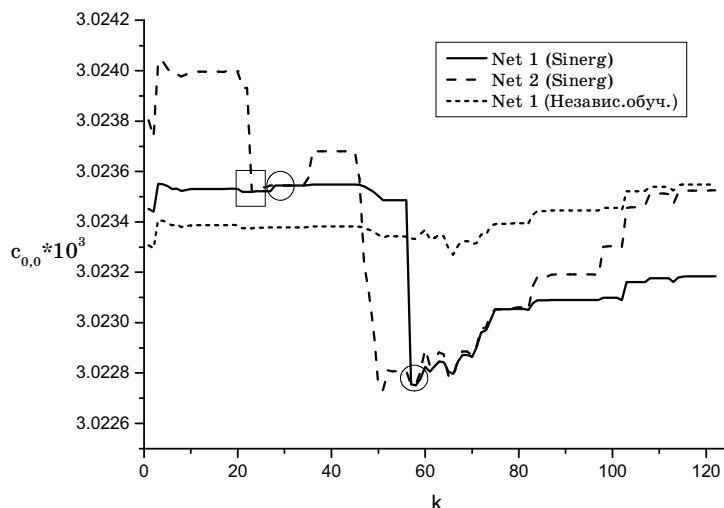


Рисунок 3 - Процесс настройки ширины функции принадлежности 1-го входа нечетких нейросетей на 2-й эпохе обучения

**Заключение.** Использование многомодельного подхода в сочетании с синергичной стратегией настройки параметров моделей позволяет существенно повысить не только скорость сходимости известных процедур настройки параметров нечетких нейросетей за счет задания различных начальных значений коэффициентов «забывания», одновременного задания нескольких начальных условий (аналог мультистарта), в соответствии с чем увеличивается вероятность удачного первого шага, но и точность аппроксимации и предсказания в целом. При этом увеличение числа синергированных моделей дает практически неограниченные возможности для изучения поведения существенно нестационарных зависимостей. В то же время при выборе числа настраиваемых моделей необходимо учитывать априорную информацию о динамике объекта, чтобы необоснованно не усложнять систему введением дополнительных моделей, снижая тем самым ее быстродействие.

#### ЛИТЕРАТУРА

1. Бодянский Е.В., Колодяжный В.В. Об одном алгоритме обучения нейро-фаззи предиктора // Адаптивные системы автоматического управления. Респ. межвед. науч.-техн. сб. - Вып. 3(23). - 2000. - С. 29-36.
2. Колесников А.А. Синергетическая теория управления. – Таганрог, ТРТУ. – М.: Энергоатомиздат, 1994. – 344 с.
3. Михалев А.И. Алгоритм поисковой идентификации с синергированными моделями существенно нестационарных динамических систем // Адаптивные системы автоматического

- управления. Респ. межвед. науч.-техн. сб. - Вып. 20. - 1992. - С. 35-40.
4. Михалёв А.И., Власова Н.В. Синергичная стратегия настройки в прогнозирующих системах //Тез. докл. 2-ой Украинской конференции по автоматическому управлению (Автоматика-95), 24-30 сентября 1995 г.-Львов.-Т.3. – С. 93-94
  5. Михалев А.И., Власова Н.В. Адаптивные методы прогнозирования оптимального состава многокомпонентной шихты для получения ферросплавов //Материалы 4-ой Укр. конф. по автомат.упр. ("Автоматика-97"), 23-28 июня 1997 г. – Черкассы. – С. 100-101.
  6. Михалев А.И., Лысяя Н.В. Многомодельный подход к прогнозированию состояний нестационарных динамических систем // Системные технологии. Компьютерная обработка экспериментальных данных: Регион. межвуз. сб. научн. тр. - Вып. 3. – Днепропетровск: "Системные технологии", 1998. – С. 99-105.

Получено 22.03.2008 г.