

# АЛГОРИТМ МАКСИМИЗАЦИИ ЭНТРОПИИ ОБУЧАЮЩЕЙ ВЫБОРКИ И ЕГО ИСПОЛЬЗОВАНИЕ ПРИ СИНТЕЗЕ МОДЕЛЕЙ ПРОГНОЗА ДИСКРЕТНЫХ СОСТОЯНИЙ НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

**А. В. Назаров<sup>а</sup>**, канд. техн. наук, доцент

**В. Л. Якимов<sup>а</sup>**, канд. техн. наук, заместитель начальника кафедры

**В. А. Авдеев<sup>а</sup>**, канд. техн. наук, заместитель начальника кафедры

<sup>а</sup>Военно-космическая академия им. А. Ф. Можайского, Санкт-Петербург, РФ

**Постановка проблемы:** для описания процессов функционирования и прогнозирования состояния неисправных, но работоспособных сложных технических объектов, находящихся в условиях нестационарного воздействия факторов внешней среды, часто используется математический аппарат, учитывающий нелинейный и вероятностный характер процессов функционирования его подсистем, например автоматы с памятью, вероятностные автоматы или логико-вероятностные модели. При этом низкая достоверность прогнозных моделей дискретных состояний обусловлена статистическими свойствами обучающей выборки. Цель работы – повысить достоверность класса прогнозных моделей дискретных состояний нелинейных динамических систем, получаемых на основе последовательного статистического обучения при заданном времени их обучения. **Результаты:** сформулирована общая постановка задачи синтеза нейросетевой прогнозной модели, использующей кодированную информацию о дискретных состояниях нелинейной динамической системы, с учетом ограничений на время ее обучения. Разработан алгоритм обучения прогнозной модели, учитывающий требования к ее применению. Предложен алгоритм максимизации энтропии обучающей выборки, позволяющий значительно повысить достоверность нейросетевой прогнозной модели дискретных состояний с последовательным обучением и улучшить данный показатель по сравнению с наилучшей авторегрессионной полиномиальной моделью. Приведен пример прогнозной модели дискретных состояний бортовой аппаратуры космического аппарата в условиях нештатного функционирования. **Практическая значимость:** разработанные алгоритмы позволяют сравнительно быстро синтезировать достоверную прогнозную модель на основе кодированной информации о состояниях дискретной динамической системы при низком значении энтропии исходной обучающей выборки. Рекомендовано использовать разработанные алгоритмы в бортовых комплексах управления летательных аппаратов для синтеза автоматов, описывающих сложные процессы функционирования бортовых систем, в том числе в условиях нештатного функционирования.

**Ключевые слова** – нелинейная динамическая система, прогнозная модель, энтропия обучающей выборки, телеметрическая информация.

## Введение

При проведении автоматизированных испытаний и в процессе эксплуатации сложных технических объектов широко используются модели в виде детерминированных конечных автоматов [1–3]. Синтез таких моделей осуществляется на основе априорной информации о процессах функционирования объектов испытаний и результатах анализа телеметрической информации об их состоянии. Одним из перспективных подходов в моделировании неисправных, но работоспособных объектов, находящихся в условиях воздействия внешних факторов среды, является применение искусственных нейронных сетей. Для нейросетевых моделей, реализующих методы статистического обучения с последовательным предъявлением обучающих примеров, доказана теоретически и показана на практике возможность аппроксимации автоматов различной сложности [4–8]. Достоверность нейросетевых прогнозных моделей сильно зависит от свойств выборки обучающих примеров. В данной работе рассматривается подход к формированию об-

учающей выборки на основе максимизации энтропии, позволяющий повысить достоверность прогнозной модели дискретных состояний нелинейной динамической системы при заданных ограничениях на время обучения. Приведены результаты его применения для моделирования процессов функционирования бортовой аппаратуры космического аппарата в условиях нештатного функционирования.

## Постановка задачи прогнозирования дискретных состояний нелинейной динамической системы

Пусть объект испытаний представляет собой систему с конечным числом состояний. В каждый  $i$ -й момент времени система находится в одном из состояний, которое кодируется двоичным словом  $S_i$ . Разряды двоичного слова  $s_{g,i}$  несут информацию о состоянии  $g$ -х функциональных узлов системы, где  $g = 1..m$ ,  $m$  — количество разрядов в слове состояния. Для ликвидации неопределенности относительно процессов управления в каждый момент времени в кодую

последовательность добавлена информация о текущей команде управления  $K_i$ . Команда управления представляет собой кодовое слово с основанием кода, равным 3: «включить — 1», «выключить — 2», «оставить неизменным — 0». Разряды слова команды управления несут информацию об изменении состояния отдельных функциональных узлов системы.

На основе анализа телеметрической информации, получаемой в ходе испытаний и эксплуатации системы, можно получить большое количество кодовых последовательностей различной длины, характеризующих процессы функционирования системы:  $S_0K_0S_0K_0S_0K_1S_1K_0S_1K_2S_2...K_1S_3K_3S_4K_0S_4K_0S_5; S_1K_0S_1K_0S_0K_4S_6K_4S_0...K_5S_3K_1S_4K_0S_4K_0S_5$  и т. д. Для их описания могут быть использованы детерминированные конечные автоматы [4–6].

Неисправности могут являться причиной воздействия внешних факторов на процессы функционирования системы, вследствие чего происходят несанкционированные изменения ее состояния (отказы). В динамике таких отказов, несмотря на их вероятностный характер, может наблюдаться закономерность, обусловленная квазициклическостью воздействия некоторых внешних факторов. Рассмотренный случай интересен тем, что процессы функционирования системы становятся нелинейными, более сложными и менее предсказуемыми, а ее состояния в последующий момент времени зависят не только от текущего состояния и команды управления, но и от состояний системы в предыдущие моменты времени.

Предсказание дискретных состояний такой системы сводится к задаче экстраполяции функции многих переменных по заданному набору примеров с помощью процедуры погружения полученной кодовой последовательности слов состояния в многомерное пространство прогнозной модели [7]. Решение задачи прогнозирования состоит в нахождении функции  $f$ :

$$S_{i+1} = f(S_{i-L+1}, S_{i-L+2}, \dots, S_{i-1}, S_i, K_i), \quad (1)$$

где  $L$  — размерность «окна прогнозирования» (глубина погружения), удовлетворяющая заданной достоверности прогноза на проверочной и тестовой выборке.

Используем в качестве прогнозной модели многослойную нейронную сеть с одним входным, одним выходным и двумя скрытыми слоями.

Такая структура показала свою эффективность при решении сложных задач обучения «с учителем» [8]. На выходе модели имеется пороговое устройство, позволяющее преобразовать непрерывные значения выходных нейронов в дискретные величины. Параметрами модели являются матрицы весовых коэффициентов сети  $W^1, W^2, W^3, C^1, C^2, C^3$ ; количество нейронов в скрытых слоях  $H^1$  и  $H^2$ ; количество нейронов во входном слое  $L$ . Векторы входных данных  $X_i$  и выходных данных  $Y_i$  образуют обучающую пару (таблица). В состав обучающей пары входят задержанные значения тех параметров состояния, которые оказывают влияние на состояние системы в последующий момент времени. Множество таких параметров состояния определяется экспертом.

Устройство прогнозирования дискретных состояний представляет собой нейросетевой автомат, который должен по команде от бортового комплекса управления гарантированно обучиться на имеющихся данных о функционировании контролируемой системы в течение заданного времени  $t_3$  не зависимо от начальных значений весовых коэффициентов и, также по команде, выдать сведения о достигнутой в ходе обучения достоверности прогнозной модели и прогноз дискретных состояний контролируемой системы с заданным временем упреждения. При этом достоверность прогнозной модели должна быть как можно выше. Введение ограничения на достоверность прогнозной модели снизу ведет к появлению вероятности невыполнения задачи за указанное время  $t_3$ , что недопустимо для обеспечивающей системы, так как может помешать выполнению целевой задачи бортового комплекса управления.

Для уменьшения чувствительности прогнозной модели к начальным значениям весовых коэффициентов целесообразно на интервале времени  $[0, t_3]$  провести несколько экспериментов по обучению нейронной сети. Статистическую оценку математического ожидания показателя достоверности прогнозной модели  $\bar{G}$  определим в соответствии с выражением

$$\bar{G} = -\bar{P}_{\text{ош}} = -\frac{1}{h_3} \sum_{h=1}^{h_3} P_{\text{ош } h}, \quad (2)$$

где  $\bar{P}_{\text{ош}}$  — статистическая оценка математического ожидания вероятности ошибок первого и второго рода по результатам экспериментов;  $h_3$  — количество экспериментов;  $h$  — номер экспери-

■ Структура обучающей пары векторов

Входной вектор $X_i$			Целевой вектор $Y_i$
Состояние $S_i$	$K_i$ , команда перехода в $S_{i+1}$	Задержанные значения параметров состояния $(s_{g,i-1}, \dots, s_{g,i-L+1})$	Состояние $S_{i+1}$
$m$ бит	$d$ бит	$L$ бит	$m$ бит

мента,  $h = 0...h_3$ ;  $P_{\text{ош } h}$  — статистическая оценка вероятности ошибки многошагового прогноза дискретных состояний системы, полученная в эксперименте:

$$P_{\text{ош } h} = \frac{\sum_{l=1}^{N-Tl+T} \sum_{i=l}^{N-Tl+T} (S_i^* - S_i)}{(N-T)T}. \quad (3)$$

Здесь  $N$  — интервал дискретных отсчетов времени, на котором вычисляется  $P_{\text{ош } h}$ ;  $T$  — интервал упреждения прогноза в дискретных отсчетах;  $l, i$  — номера дискретных отсчетов времени;  $S_i^*$  — прогнозное значение слова-состояния в  $i$ -й момент времени;  $S_i$  — истинное значение слова-состояния в  $i$ -й момент времени:

$$S_i^* - S_i = \begin{cases} 0, & S_i = S_i^* \\ 1, & S_i \neq S_i^* \end{cases}$$

Определим доверительный интервал  $\mu$  математического ожидания достоверности прогнозной модели  $\bar{G}$  по результатам экспериментов с доверительной вероятностью  $P_3 = 0,99$  [9]. Зададимся целью максимизации математического ожидания достоверности прогнозной модели  $\bar{G}$  дискретных состояний системы. Тогда задача оптимизации прогнозной модели дискретных состояний примет следующий вид:

$$\begin{aligned} & \bar{G}(W^1, W^2, W^3, C^1, C^2, C^3, H^1, H^2, t_0) \rightarrow \max; \\ & \mu = \frac{3,25\sigma(\bar{G})}{\sqrt{h_3}}, \bar{G} \in (0,1), \mu \in (0,1), t_0 \leq t_3, \end{aligned} \quad (4)$$

где  $\sigma(\bar{G})$  — среднеквадратическое отклонение величины  $\bar{G}$  по результатам  $h_3$  экспериментов;  $t_0$  — время, выделяемое для обучения моделей.

Достоверность прогнозной модели дискретных состояний системы  $G = \bar{G} \pm \mu$  сильно зависит от свойств обучающей выборки. В результате значительной неравномерности распределения во времени режимов работы системы и управляющих воздействий, наличия в обучающей выборке информации не о всех переходах между состояниями системы, а также необходимости учета задержанных значений параметров состояния отдельных функциональных узлов возникает неравномерное распределение символов в обучающих примерах и примеров в обучающей выборке. При этом нейронная сеть с последовательным предъявлением обучающих примеров реагирует только на примеры обучающей выборки, характеризующие наиболее частые переходы между состояниями системы, в результате чего часть весовых коэффициентов сети не участвует в процессе обучения и эффект обобщения символьной информации не проявляется.

Для решения данной проблемы осуществляют поиск оптимального основания кода путем пере-

бора для обучающей пары векторов нейронной сети либо реализуют перекодирование имеющейся информации с использованием слоя Кохонена, что требует значительных временных затрат. Оба пути основаны на общем принципе преобработки исходных данных, который состоит в максимизации энтропии входов и выходов нейросетевой модели [10]. С другой стороны, энтропия будет максимальна, если дискретные величины, поступающие на каждый входной и выходной нейрон сети, будут распределены во времени равномерно. Преодолеть неравномерность возбуждения входных и выходных нейронов символами различных классов можно, изменяя частоту повторения тех или иных обучающих примеров на входе нейронной сети.

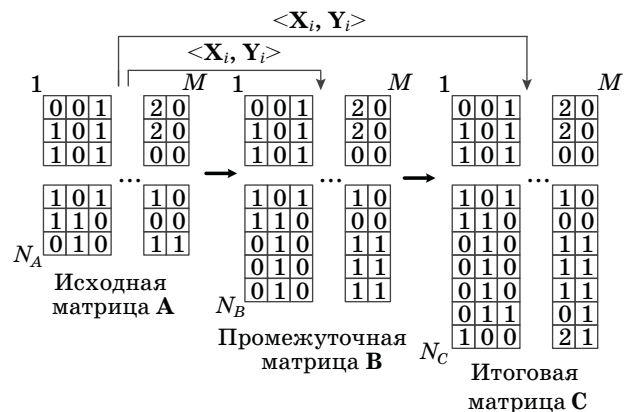
Рассмотрим матрицу исходных данных  $A_{N_A \times M}$ , где  $M = (2m + d + L - 1)$  — размер обучающей пары векторов (рис. 1).

Дополняя матрицу  $A_{N_A \times M}$  содержащимися в ней векторами-строками  $\langle X_i, Y_i \rangle, i = 1...N_A$ , необходимо сформировать матрицу  $C_{N_C \times M}$ , для которой характерно максимально равномерное распределение символов различного класса в столбцах. При этом необходимо ввести промежуточную матрицу  $B_{J \times M}$ , количество векторов  $J$  в которой ( $N_A \leq J \leq N_C$ ) изменяется в процессе решения следующей задачи оптимизации:

$$\begin{aligned} & V(r) \rightarrow \min, r \rightarrow \infty, \\ & V(r) = \sqrt{\sum_{j=1}^M \sum_{s=1}^M \left( \frac{n_{js}(r)}{J(r)} - \frac{1}{\rho_{js}(r)} \right)^2}; \end{aligned} \quad (5)$$

$$\begin{cases} \bar{G}(r+1) > \bar{G}(r), \mu(r+1) < \mu(r) \\ V(r+1) < V(r) \end{cases} \begin{cases} \varepsilon_{d1 \text{ min}} \leq |\bar{G}(r+1) - \bar{G}(r)| \leq \varepsilon_{d1 \text{ max}}, \\ \varepsilon_{d2 \text{ min}} \leq |\mu(r+1) - \mu(r)| \leq \varepsilon_{d2 \text{ max}} \end{cases} \quad (6)$$

где  $r$  — количество итераций процедуры оптимизации;  $n_{js}$  — число элементов  $x_{ij}$  матрицы  $B_{J \times M}$



■ Рис. 1. Схема формирования обучающей выборки

в  $j$ -м столбце, принадлежащих классу  $s$ ;  $i, j$  — номер строки и столбца промежуточной матрицы  $\mathbf{V}_{J \times M}$  соответственно;  $n_{js}/J$  — показатель плотности вероятности распределения элементов  $s$ -го класса на входе  $j$ -го нейрона;  $\rho_{js}$  — количество классов элементов в  $j$ -м столбце;  $1/\rho_{js}$  — показатель плотности вероятности распределения элементов на входе  $j$ -го нейрона в случае равномерного закона распределения;  $\varepsilon_{д1 \min}, \varepsilon_{д2 \min}, \varepsilon_{д1 \max}, \varepsilon_{д2 \max}$  — границы допусков на изменение  $\bar{G}$  и  $\mu$ .

Минимизация функции (5) сопровождается увеличением количества обучающих пар векторов в выборке. Условия (6) означают, что с уменьшением значения целевой функции математическое ожидание достоверности прогнозной модели  $\bar{G}$  должно расти, а доверительный интервал  $\mu$  — уменьшаться, при этом допускаются случайные изменения указанных характеристик в пределах заданных допусков. Критерием останова процедуры минимизации целевой функции (5) будем считать достижение такого значения  $r$ , при котором характеристики модели прогнозирования ( $\bar{G}$ ,  $\mu$ ) не соответствуют условиям (6). Эффективность решения задачи (5) для повышения достоверности прогнозной модели может быть низкой в следующих случаях: исходная обучающая выборка обладает высокой энтропией; свойства обучающей выборки таковы, что при решении задачи (5) изменения целевой функции будут крайне малы.

### Алгоритм обучения нейросетевой прогнозной модели

Алгоритм обучения и применения нейросетевой прогнозной модели, позволяющий решить задачу (4), использует принцип обратного распространения ошибки [10]. Схема алгоритма приведена на рис. 2.

**Шаг 1.** Задать структуру модели, сформировать обучающую, проверочную и тестовую выборки; присвоить номер эксперименту обучения:  $h = 1, h = 1..h_3$ ; задать матрицы  $\mathbf{U}^k$  для хранения значений весовых коэффициентов каждого слоя наилучшей прогнозной модели, полученной по результатам  $h$  экспериментов,  $k$  — номер слоя,  $k = 1..3$ ; задать значение промежуточной переменной для хранения минимального значения ошибки многошагового прогноза  $P_{\text{ош } h, \theta}$ , полученной по результатам  $h$  экспериментов,  $P = 1$ .

**Шаг 2.** Задать начальные значения весовых коэффициентов матриц каждого слоя  $\mathbf{W}^k$  и  $\mathbf{C}^k$  на интервале  $[-0,5; 0,5]$ ; присвоить элементам матриц изменений весовых коэффициентов  $d\mathbf{C}_i^k$  и  $d\mathbf{W}_i^k$  нулевые значения; задать количество циклов обучения в каждом эксперименте; задать начальное значение статистической оценки вероятности ошибки на проверочной выборке  $P_{\text{ош } h, \theta} = 1$ ;

значение периода упреждения прогноза  $T$ ; номер цикла обучения  $\theta = 1, \theta = 1.. \theta_3$ , где  $\theta_3$  — заданное количество циклов обучения; матрицы  $\mathbf{U}_h^k$  для хранения значений весовых коэффициентов каждого слоя наилучшей прогнозной модели, полученной в эксперименте; задать значение промежуточной переменной для хранения минимального значения ошибки многошагового прогноза  $P_{\text{ош } h, \theta}$ , полученной в эксперименте,  $P_h = 1$ .

**Шаг 3.** Присвоить значение счетчику векторов обучающей выборки  $i = 1$ , где  $i = 1..N$ ,  $N$  — количество векторов в обучающей выборке; выбрать соответствующую пару векторов из обучающей выборки  $\langle \mathbf{X}_i, \mathbf{Y}_i \rangle$ .

**Шаг 4.** Подать на вход/выход модели обучающую пару  $\langle \mathbf{X}_i, \mathbf{Y}_i \rangle$ ; вычислить компоненты вектора выходных значений каждого слоя  $\mathbf{Y}_i^k$ ; вычислить взвешенные суммы  $\mathbf{Q}_i^k$  каждого слоя нейронной сети; осуществить нелинейное преобразование с использованием сигмоидной функции активации и получить векторы выходных значений каждого слоя  $\mathbf{Y}_i^k$ :

$$\mathbf{Q}_i^k \equiv [q_v^k] = \mathbf{X}_i \mathbf{W}_i^k + \mathbf{C}_i^k, \\ \mathbf{Y}_i^k \equiv [y_v^k] = \left( y_v = \frac{1}{1 + \exp(-q_v^k)} \right), \quad (7)$$

где  $\mathbf{X}_i$  — входной вектор-строка параметров модели;  $\mathbf{W}_i^k$  — текущая матрица весовых коэффициентов  $k$ -го слоя;  $\mathbf{C}_i^k$  — текущий вектор постоянно-го смещения;  $v$  — номер элемента вектора-строки. В результате прямого прохода сигналов на выходе нейросетевой прогнозной модели получаем вектор

$$\mathbf{Y}_i \equiv [y_v] = f \left( f \left( \mathbf{X}_i \mathbf{W}_i^1 + \mathbf{C}_i^1 \right) \mathbf{W}_i^2 + \mathbf{C}_i^2 \right) \mathbf{W}_i^3 + \mathbf{C}_i^3, \quad (8)$$

где  $y_v$  — компоненты выходного вектора-строки;  $f$  — сигмоидная функция активации (7). Для сети с двумя скрытыми слоями целевая функция будет иметь вид

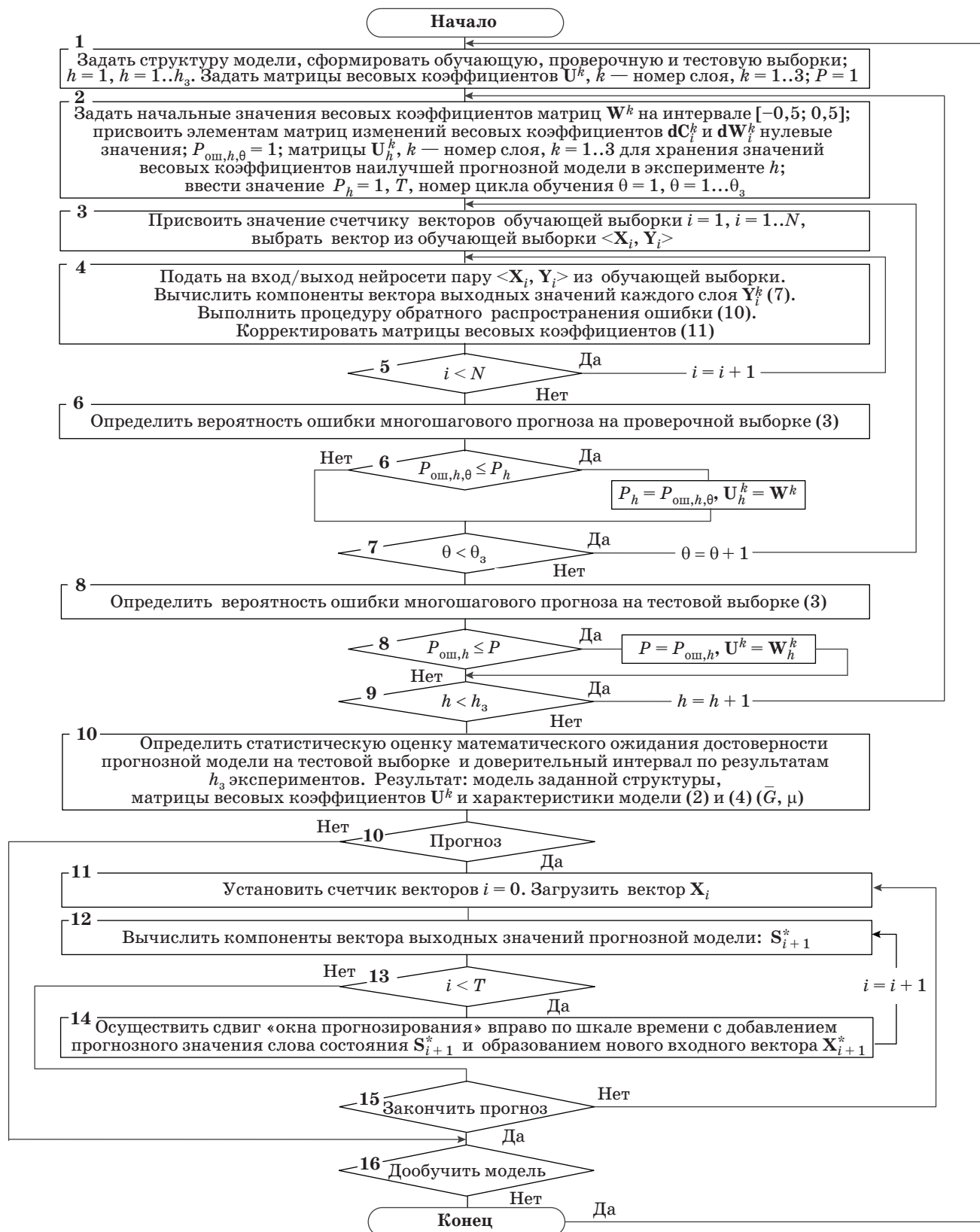
$$E(\mathbf{W}^k, \mathbf{C}^k) = \sum_{i=1}^N \sum_{v=1}^m \left( y_v - y_v^* \right)^2, \quad (9)$$

где  $\mathbf{Y}_i^* \equiv [y_v^*]$  — целевой вектор;  $y_v^*$  — компоненты целевого вектора;  $m$  — размер выходного вектора.

Далее реализуется процедура обратного распространения ошибки и вычисляется вектор ошибок каждого слоя нейронной сети  $\Delta_i^k$ :

$$\Delta_i \equiv [\delta_v] = \mathbf{Y}_i - \mathbf{Y}_i^*; \\ \Delta_i^k \equiv [\delta_v^k] = \left( \delta_v = y_v^k (1 - y_v^k) \delta_v \right); \\ \mathbf{D} \equiv [d_v^k] = \Delta_i^k \mathbf{W}_i^{kT}; \\ \Delta_i^{k-1} \equiv [\delta_v^{k-1}] = \left( \delta_v^{k-1} = y_v^{k-1} (1 - y_v^{k-1}) d_v^k \right). \quad (10)$$





■ Рис. 2. Алгоритм обучения нейросетевой прогнозной модели

Затем осуществляется коррекция значений элементов матриц  $dC_i^k$ ,  $dW_i^k$ ,  $C_i^k$ ,  $W_i^k$ :

$$\begin{aligned} dW_i^k &= \eta Y_i^{k-1T} \Delta_i^k + \alpha dW_{i-1}^k; \\ dC_i^k &= \eta \Delta_i^k + \alpha dC_{i-1}^k; \\ W_{i+1}^k &= W_i^k + dW_i^k; C_{i+1}^k = C_i^k + dC_i^k, \end{aligned} \quad (11)$$

где  $\eta$  — скорость обучения;  $\alpha$  — момент обучения.

**Шаг 5.** Проверить выполнение условия использования всех примеров обучающей выборки: если  $i < N$ , то увеличить значение счетчика векторов  $i$  на единицу и повторить шаг 4, иначе перейти к шагу 6.

**Шаг 6.** Определить статистическую оценку вероятности ошибки многошагового прогноза  $P_{\text{ош } h, \theta}$  (3) на проверочной выборке после каждого цикла обучения. Если  $P_{\text{ош } h, \theta} \leq P_h$ , то присвоить переменной  $P_h$  значение  $P_{\text{ош } h, \theta}$ , а элементам матриц  $U_h^k$  — текущие значения матриц весовых коэффициентов каждого слоя сети  $W^k$  и  $C^k$ .

**Шаг 7.** Проверить выполнение условия: если количество циклов обучения  $\theta$  не превысило заданное значение  $\theta_3$ , то увеличить значение счетчика циклов обучения  $\theta$  на единицу и выполнить шаги 3–6, иначе перейти к шагу 8. Величина  $\theta_3$  подбирается экспериментальным путем до начала серии экспериментов, исходя из приемлемого значения времени обучения:  $t_0 < t_3$ .

**Шаг 8.** Определить статистическую оценку вероятности ошибки многошагового прогноза  $P_{\text{ош } h}$  на тестовой выборке, получаемую для прогнозной модели с матрицами весовых коэффициентов  $U_h^k$ . Если  $P_{\text{ош } h} \leq P$ , то присвоить переменной  $P$  значение  $P_{\text{ош } h}$ , а элементам матриц  $U^k$  — текущие значения матриц весовых коэффициентов каждого слоя  $U_h^k$ , при невыполнении условия перейти к шагу 9.

**Шаг 9.** Проверить выполнение условия: если достигнуто заданное количество экспериментов обучения модели  $h_3$ , то перейти к шагу 10, иначе увеличить значение счетчика экспериментов  $h$  на единицу и выполнить шаги 2–8. Величина  $h_3$  выбирается экспериментальным путем до начала серии экспериментов, исходя из приемлемого значения времени обучения:  $t_0 < t_3$ .

**Шаг 10.** Определить статистическую оценку математического ожидания достоверности прогнозной модели  $\bar{G}$  на тестовой выборке (2) и доверительный интервал  $\mu$  (4). Результатом работы шагов 1–10 алгоритма является прогнозная модель (структура нейронной сети и набор из  $k$  матриц весовых коэффициентов каждого ее слоя  $U^k$ ) и характеристики ( $\bar{G}$ ,  $\mu$ ), позволяющие оценить, насколько можно доверять ее прогнозам.

**Шаг 11.** По команде на применение прогнозной модели значению счетчика векторов присвоить

значение  $i = 0$ , на вход модели подать новый входной вектор  $X_i$ .

**Шаг 12.** Вычислить компоненты вектора выходных прогнозных значений нейронной сети  $S_{i+1}^*$ , соответствующие входному вектору  $X_i$ .

**Шаг 13.** Проверить выполнение условия реализации многошагового прогноза с периодом упреждения  $T$ : если  $i \geq T$ , то выполнить шаг 15, иначе шаг 14.

**Шаг 14.** Осуществить сдвиг «окна прогнозирования» вправо с добавлением прогнозного значения слова состояния  $S_{i+1}^*$  во входной вектор  $X_{i+1}^*$ , увеличить значение счетчика  $i$  на единицу, выполнить шаги 12–13.

**Шаг 15.** Проверить выполнение условия: если поступила команда закончить прогноз, то выполнить шаг 16, иначе повторить шаги 11–14 для очередного входного вектора.

**Шаг 16.** Проверить выполнение условия: если поступила команда дообучить модель на новых данных, то, при необходимости, верифицировать структуру прогнозной модели и выполнить шаги 1–15, иначе закончить работу.

### Алгоритм максимизации энтропии обучающей выборки

Для повышения достоверности прогнозной модели, синтезируемой в результате работы описанного алгоритма, решим задачу максимизации энтропии обучающей выборки (5). Для ее решения разработан алгоритм, представленный на рис. 3.

**Шаг 1.** Задаются исходная обучающая выборка в виде матрицы  $A$ ; структура пробной прогнозной модели и начальные значения ее весовых коэффициентов; промежуточная матрица  $B$ , элементам которой присваиваются значения элементов исходной обучающей выборки  $A$ ; матрица  $Z$  для хранения промежуточных вариантов матрицы  $B$ ; граничные значения допусков на характеристики модели (4)  $\varepsilon_{d1 \min}$ ,  $\varepsilon_{d2 \min}$ ,  $\varepsilon_{d1 \max}$ ,  $\varepsilon_{d2 \max}$ ; параметр  $r$ , показывающий количество циклов обновления выборки,  $r = 0$ ; оценивается значение целевой функции  $V_B$  для матрицы  $B$  в соответствии с (6).

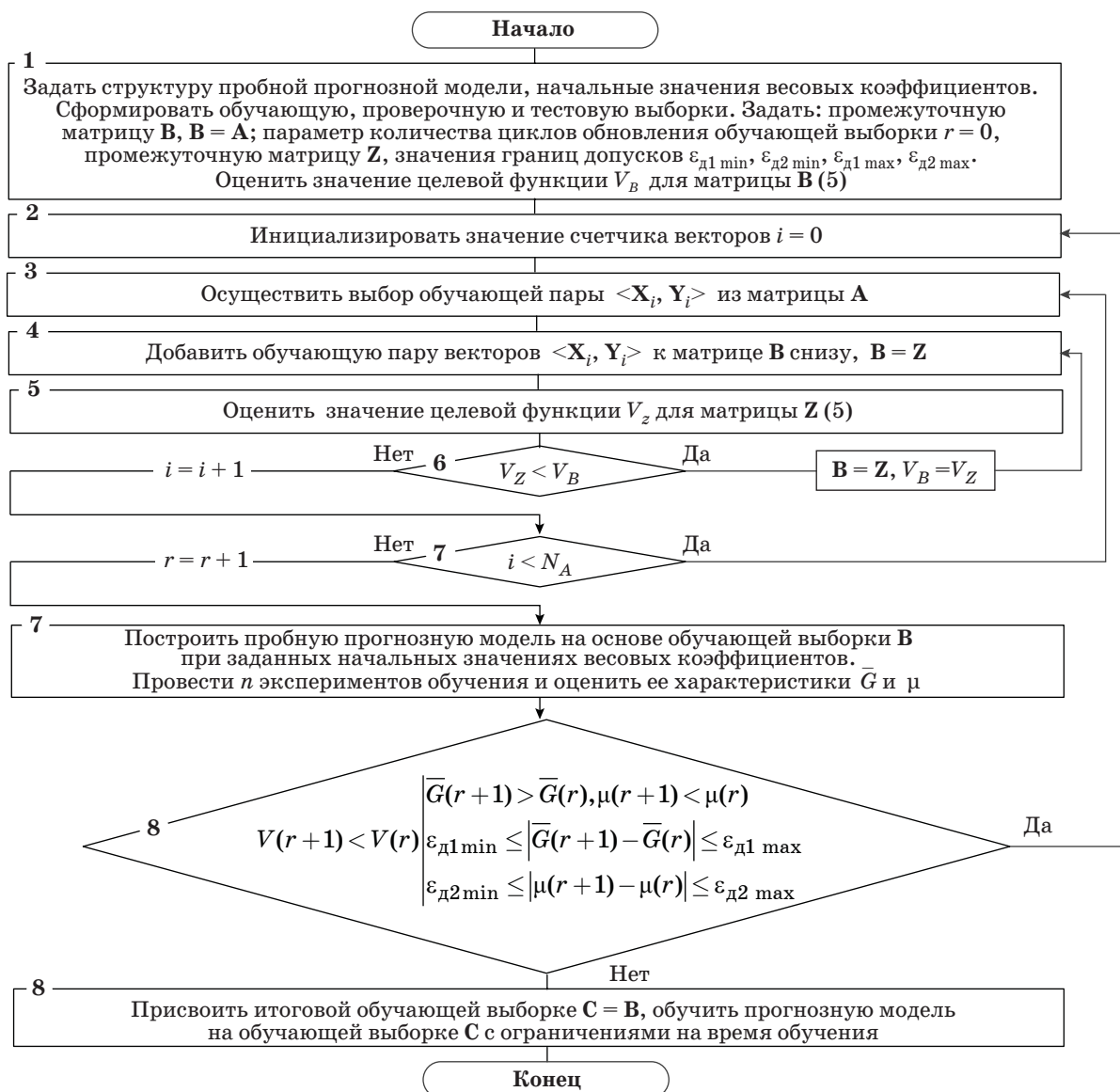
**Шаг 2.** Осуществляется инициализация значения счетчика векторов  $i = 0$ .

**Шаг 3.** Осуществляется выборка обучающей пары  $\langle X_i, Y_i \rangle$  из матрицы  $A$ .

**Шаг 4.** Обучающая пара векторов  $\langle X_i, Y_i \rangle$  добавляется к матрице  $B$  снизу, а значения элементов промежуточной матрицы  $B$  заносятся в матрицу  $Z$ .

**Шаг 5.** Оценивается значение целевой функции  $V_Z$  для матрицы  $Z$ .

**Шаг 6.** Проверяется следующее условие: если значение целевой функции уменьшается:  $V_Z < V_B$ , то  $B = Z$ ,  $V_B = V_Z$ , осуществляется повтор шагов 4–5 для обучающей пары векторов  $\langle X_i, Y_i \rangle$ ; иначе,



■ Рис. 3. Алгоритм максимизации энтропии обучающей выборки

если  $V_Z \geq V_B$ , то значение счетчика векторов  $i$  увеличивается на единицу.

**Шаг 7.** Проверяются следующие условия: если  $i < N_A$ , то необходимо повторить шаги 3–6 для очередного вектора  $\langle X_i, Y_i \rangle$ . Если же  $i > N_A$ , то  $r = r + 1$ , осуществляется загрузка исходных данных для обучения прогнозной модели, строится пробная прогнозная модель на основе обучающей выборки  $B$  при заданных начальных значениях весовых коэффициентов, оцениваются характеристики  $\bar{G}$ ,  $\mu$  пробной прогнозной модели.

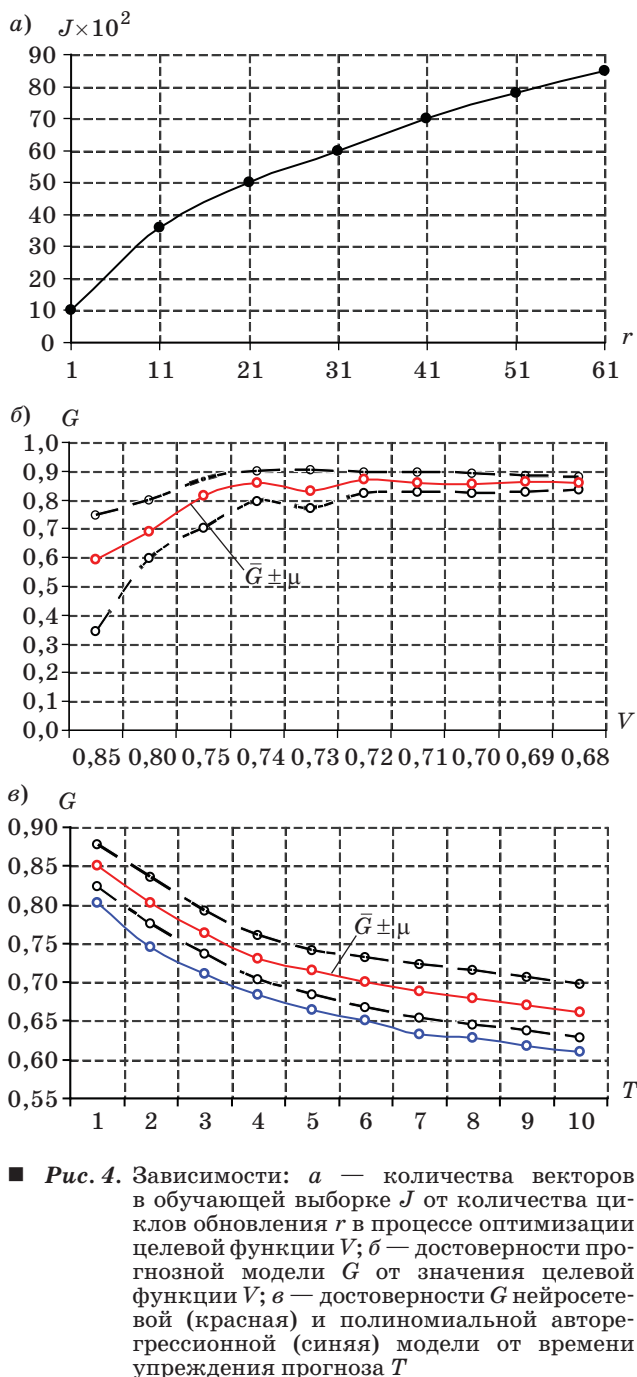
**Шаг 8.** Проверяется следующее условие: если характеристики прогнозной модели удовлетворяют условиям (6) при увеличении  $r$ , то выполняются шаги 2–7, иначе останавливается процесс оптимизации целевой функции, переписываются значения элементов матрицы  $B$  в матрицу  $C$ , обу-

чается прогнозная модель на обучающей выборке  $C$  с учетом ограничений на время обучения (4).

Пробная прогнозная модель используется в алгоритме для определения качества формируемой обучающей выборки. При сравнении разных вариантов обучающей выборки (шаг 7) необходимо обеспечить равные условия для обучения пробных прогнозных моделей: структуру модели, начальные значения весовых коэффициентов и количество векторов, поступивших на вход модели за время ее обучения  $t_0$ . Поэтому с ростом размера обучающей выборки при решении задачи (5) необходимо пропорционально уменьшать количество циклов обучения. Полученная в результате решения задачи (5) обучающая выборка используется для решения задачи (4) уже при заданных ограничениях на время обучения.

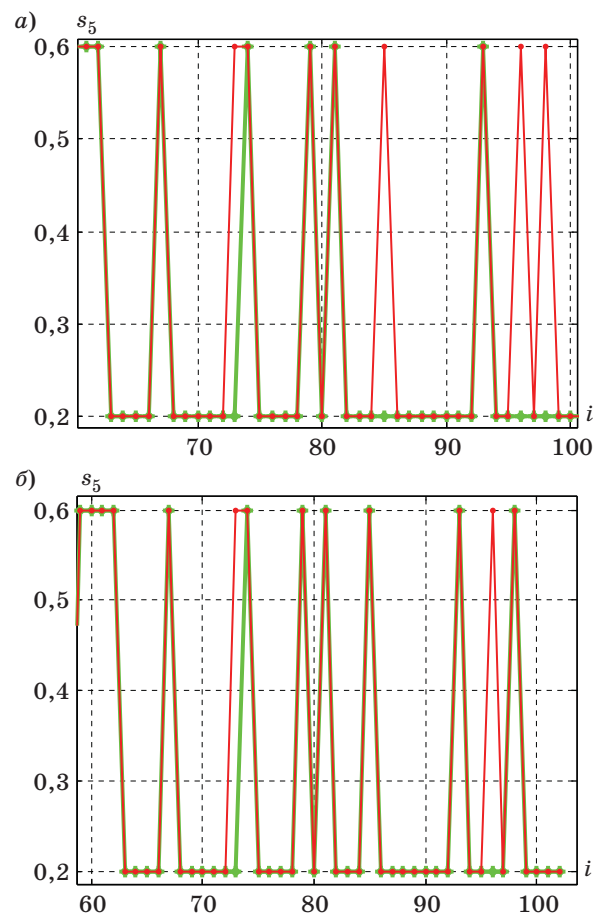
Результаты моделирования

Рассмотренные алгоритмы были использованы для синтеза прогнозной модели дискретных состояний бортовой аппаратуры космического аппарата с неисправной системой электропитания. Неисправность стала причиной срабатывания систем автоматики и нештатных отключений солнечной батареи. Результаты моделирования представлены на рис. 4. Во входной вектор прогнозной модели вошли слово состояния бортовой аппаратуры и солнечной батареи в текущий



момент времени (5 бит), команда управления (4 бит) и задержанные значения параметра состояния солнечной батареи (20 бит). Выходной вектор представлял собой слово состояния бортовой аппаратуры и солнечной батареи в следующий момент времени (5 бит). Размер обучающей выборки составил 1000 примеров на интервале времени, равном одному году. Время обучения прогнозной модели  $t_3$  было выбрано равным периоду одного витка космического аппарата. В результате работы алгоритма максимизации энтропии обучающей выборки и решения задачи (5) количество векторов в ней увеличилось до 8500 (рис. 4, а).

Полученная обучающая выборка была использована для обучения нейронной сети и решения задачи (4) при заданных ограничениях на время обучения, в результате чего достоверность пробной прогнозной модели увеличилась с 60 до 85 % (рис. 4, б). На рис. 4, в представлены результаты применения полученной прогнозной модели при различном времени упреждения прогноза  $T$



■ **Рис. 5.** Фрагмент зависимости истинных (красные) и прогнозных (зеленые) значений параметра состояния, характеризующего отключения солнечной батареи без использования (а) и с использованием (б) алгоритма максимизации энтропии обучающей выборки



в сравнении с наилучшей полиномиальной авторегрессионной моделью.

На рис. 5 представлены фрагменты зависимости истинных и прогнозных значений параметра состояния солнечной батареи (разряд  $s_5$  в слове состоянии), характеризующего ее отключения (штатные и несанкционированные), полученные нейросетевой прогнозной моделью на интервале упреждения, равном одному дискретному отсчету времени (3 мин), без использования алгоритма максимизации энтропии обучающей выборки (рис. 5, а) и с использованием разработанного алгоритма (рис. 5, б).

На рассмотренном интервале времени алгоритм максимизации энтропии обучающей выборки позволил уменьшить вероятность пропуска отказа.

### Заключение

Для оперативного решения задач технической диагностики сложных объектов испытаний используется телеметрическая информация, содер-

жащая сведения о состояниях объекта и значения некоторых наиболее важных его функциональных параметров. В работе показана возможность использования такого рода данных для синтеза нейросетевых прогнозных моделей дискретных состояний (нейросетевых автоматов) нелинейных динамических систем на примере бортовой аппаратуры космического аппарата в условиях нештатного функционирования. Использование алгоритма максимизации энтропии позволило повысить достоверность прогнозной модели дискретных состояний бортовой аппаратуры космического аппарата на основе многослойной нейронной сети на 25 % и улучшить данный показатель по сравнению с авторегрессионной полиномиальной моделью на 5–7 %. Разработанный алгоритм максимизации энтропии обучающей выборки позволяет сформировать такую последовательность обучающих примеров на входе/выходе прогнозной модели, которая улучшает ее возможности к обобщению символьной информации и повышает достоверность при введении ограничений на время обучения.

### Литература

1. Александровская Л. Н., Круглов В. И., Кузнецов А. Г. Теоретические основы испытаний и экспериментальная отработка сложных технических систем: учеб. пособие. — М.: Логос, 2003. — 736 с.
2. Горбатов В. А. Фундаментальные основы дискретной математики. — М.: Наука. Физматлит, 2000. — 544 с.
3. Luger G. F. Artificial Intelligence: Structures and Strategies for Complex Problem Solving. — Addison Wesley, 2002. — 850 p.
4. Callan R. The Essence of Neural Networks (Essence of Computing). — Prentice Hall, 1998. — 248 p.
5. Кретинин А. В., Солдатов Д. В., Шалыто А. А., Шостаков А. В. Ракеты. Автоматы. Нейронные сети // Нейрокомпьютеры: разработка и применение. 2005. № 5. С. 50–59.
6. Мальцев Г. Н., Назаров А. В., Якимов В. Л. Алгоритм реконструкции фазового пространства динамической системы и его применение для разработки прогнозных моделей // Информационно-управляющие системы. 2014. № 2(69). С. 33–40.
7. Назаров А. В., Якимов В. Л. Прогнозирование технического состояния малых космических аппаратов с использованием многослойных нейронных сетей // Изв. вузов. Приборостроение. 2006. № 1. С. 7–12.
8. Горбань А. Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей // Сибирский журнал вычислительной математики. 1998. Т. 1. № 1. С. 12–24.
9. Хартман К., Лецкий Э., Шефер В. Планирование эксперимента в исследованиях технологических процессов. — М.: Мир, 1977. — 544 с.
10. Haykin S. S. Neural Networks: A Comprehensive Foundation. — Prentice Hall, 1998. — 842 p.

UDC 629.76.05

doi:10.15217/issn1684-8853.2015.2.57

### Training Sample Entropy Maximization Algorithm and its Use in Synthesis of Predictive Models of Nonlinear Dynamical System Discrete States

Nazarov A. V.<sup>a</sup>, PhD, Tech., Associate Professor, naz\_av@mail.ru

Yakimov V. L.<sup>a</sup>, PhD, Tech., Deputy Head of Chair, yakim78@yandex.ru

Avdeev V. A.<sup>a</sup>, PhD, Tech., Deputy Head of Chair, apex7@rambler.ru

<sup>a</sup>A. F. Mozhayskii Military Space Academy, 13, Zhdanovskaia St., 197198, Saint-Petersburg, Russian Federation

**Purpose:** To describe the functioning processes and forecast the condition of faulty but runnable complex technical objects under non-stationary impact of environment factors, we often use a mathematical apparatus considering nonlinear and probabilistic nature of the functioning processes in its subsystems, for example, automata with memory, probabilistic automata or logical-probabilistic models. The predictive models of discrete states have low reliability because of the statistical properties of the training sample.

The goal of this work is to increase the reliability of predictive models for discrete states of nonlinear dynamical systems obtained on the base of consecutive statistical training when the training time is preset. **Results:** A general problem definition is formulated for neural network predictive model synthesis. It uses coded information on discrete conditions of a nonlinear dynamic system taking into account the restrictions on its training time. An algorithm of predictive model training is developed, considering the requirements to its application. An algorithm of maximizing the training sample entropy is offered, helping to significantly increase the reliability of discrete states neural network predictive model with consecutive training and to improve this indicator in comparison with the best autoregression polynomial model. An example is given for a predictive model of spacecraft onboard equipment discrete conditions under the conditions of emergency functioning. **Practical relevance:** The developed algorithms allow us to quickly synthesize a reliable predictive model on the base of coded information on the conditions of a discrete dynamic system at a low value of entropy of the initial training sample. It is recommended to use the developed algorithms in onboard aircraft management complexes for the synthesis of automata which describe complex processes of onboard system functioning, including emergency functioning.

**Keywords** — Nonlinear Dynamic System, Predictive Model, Training Sample Entropy, Telemetric Information.

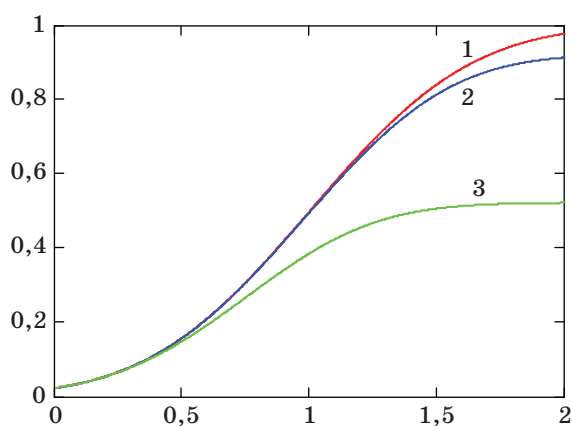
**References**

1. Aleksandrovskaia L. N., Kruglov V. I., Kuznetsov A. G. *Teoreticheskie osnovy ispytaniy i eksperimental'naya otrabotka slozhnykh tekhnicheskikh sistem: uchebnoe posobie* [Theoretical Tests Bases and Experimental Working of Difficult Technical Systems]. Moscow, Logos Publ., 2003. 736 p. (In Russian).
2. Gorbatov V. A. *Fundamental'nye osnovy diskretnoi matematiki* [Fundamental Bases of Discrete Mathematics]. Moscow, Nauka Publ., 2000. 544 p. (In Russian).
3. Luger G. F. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. AddisonWesley, 2002. 850 p.
4. Callan R. *The Essence of Neural Networks (Essence of Computing)*. Prentice Hall, 1998. 248 p.
5. Kretinin A. V., Soldatov D. V., Shalyto A. A., Shostak A. V. *Rockets. Automatos. Neural Networks. Neurokomp'iutery: razrabotka i primeneniye*, 2005, no. 5, pp. 50–59 (In Russian).
6. Maltsev G. N., Nazarov A. V., Yakimov V. L. A Reconstruction Algorithm for a Dynamic System Phase Space and its Application for Development of Predictive Models. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2014, no. 2(69), pp. 33–39 (In Russian).
7. Nazarov A. V., Yakimov V. L. Prediction of Technical State of Small Spacecrafts with Usage Multilayer Neural Networks. *Izvestiia vuzov. Priborostroeniye*, 2006, no. 1, pp. 7–12 (In Russian).
8. Gorban' A. N. Generalized Approximation Theorem and Computational Capabilities of Neural Networks. *Sibirskii zhurnal vychislitel'noi matematiki*, 1998, vol. 1, no. 1, pp. 12–24 (In Russian).
9. Khartman K., Letskii E., Shefer V. *Planirovaniye eksperimenta v issledovaniyakh tekhnologicheskikh protsessov* [Experiments Design in the Technological Processes Research]. Moscow, Mir Publ., 1977. 552 p. (In Russian).
10. Haykin S. S. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1998. 842 p.

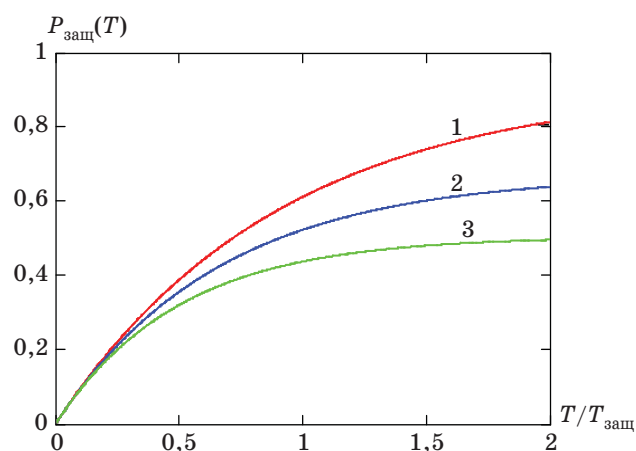
**К статье Г. Н. Мальцева, А. Н. Панкратова, Д. А. Лесняка «Исследование вероятностных характеристик изменения защищенности информационной системы от несанкционированного доступа нарушителей»**

(Информационно-управляющие системы, 2015, № 1(74), с. 50–58).

На с. 55 ошибочно помещен рисунок:



Вместо него должен быть следующий рисунок:



Редакция журнала приносит извинения за ошибку, допущенную в статье.