

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
МИНИСТЕРСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ ПО АТОМНОЙ ЭНЕРГИИ  
МИНИСТЕРСТВО ПРОМЫШЛЕННОСТИ, НАУКИ И ТЕХНОЛОГИЙ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
РОССИЙСКАЯ АССОЦИАЦИЯ НЕЙРОИНФОРМАТИКИ  
МОСКОВСКИЙ ИНЖЕНЕРНО-ФИЗИЧЕСКИЙ ИНСТИТУТ  
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

---

**НАУЧНАЯ СЕССИЯ МИФИ–2003**

**НЕЙРОИНФОРМАТИКА–2003**

**V ВСЕРОССИЙСКАЯ  
НАУЧНО-ТЕХНИЧЕСКАЯ  
КОНФЕРЕНЦИЯ**

**ЛЕКЦИИ  
ПО НЕЙРОИНФОРМАТИКЕ**

**Часть 1**

По материалам Школы-семинара  
«Современные проблемы нейроинформатики»

Москва 2003

УДК 004.032.26 (06)

ББК 32.818я5

М82

**НАУЧНАЯ СЕССИЯ МИФИ–2003. V ВСЕРОССИЙСКАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ «НЕЙРОИНФОРМАТИКА–2003»: ЛЕКЦИИ ПО НЕЙРОИНФОРМАТИКЕ. Часть 1.** – М.: МИФИ, 2003. – 188 с.

В книге публикуются тексты лекций, прочитанных на Школе-семинаре «Современные проблемы нейроинформатики», проходившей 29–31 января 2003 года в МИФИ в рамках V Всероссийской конференции «Нейроинформатика–2003».

Материалы лекций связаны с рядом проблем, актуальных для современного этапа развития нейроинформатики, включая ее взаимодействие с другими научно-техническими областями.

Ответственный редактор

*Ю. В. Тюменцев*, кандидат технических наук

ISBN 5–7262–0471–9

© *Московский инженерно-физический институт  
(государственный университет), 2003*

## Содержание

<b>Б. В. Крыжановский, Л. Б. Литинский. Векторные модели ассоциативной памяти</b>	<b>72</b>
Введение . . . . .	73
Поттс-стекольная нейросеть и параметрическая нейронная сеть .	75
Статистическая техника Чебышева–Чернова . . . . .	79
Литература . . . . .	84

**Б. В. КРЫЖАНОВСКИЙ, Л. Б. ЛИТИНСКИЙ**

Институт оптико-нейронных технологий РАН,  
Москва

**E-mail: kryzhanov@mail.ru,  
litin@mail.ru**

## **ВЕКТОРНЫЕ МОДЕЛИ АССОЦИАТИВНОЙ ПАМЯТИ**

### **Аннотация**

Дается обзор векторных моделей ассоциативной памяти. Модель Хопфилда позволяет эффективно запомнить сравнительно небольшое число паттернов – порядка 15% от размера нейронной сети. Существенно превзойти этот показатель удается только в Поттс-стекольной модели ассоциативной памяти, где нейроны могут находиться в большем чем два числе состояний. Показано, что еще большей емкостью памяти обладает параметрическая нейронная сеть (ПНС). Для оценки емкости памяти используется статистическая техника Чебышева–Чернова.

**B. V. KRYZHANOVSKY, L. B. LITINSKII**

Institute of Optical Neural Technologies, RAS,  
Moscow

**E-mail: kryzhanov@mail.ru,  
litin@mail.ru**

## **THE VECTOR MODELS OF ASSOCIATIVE MEMORY**

### **Abstract**

We present a short review of vector models for associative memory. The storage capacity of the Hopfield model is about 15% of this network size. It can be increased significantly in the Potts-glass model of the associative memory only. In this model neurons can be in more than two different states. We show that even greater storage capacity can be achieved in the parametrical neural network (PNN). We use the Chebyshev–Chernov statistical technique to estimate the PNN storage capacity.

## Введение

Последние 15 лет отмечены интересом к моделям ассоциативной памяти с  $q$ -нарными нейронами. Всё в этих моделях подобно модели Хопфилда, но число состояний  $q$ , в которых могут находиться нейроны, больше 2.

В работах [1–3] состояния нейронов изображаются *спиновой переменной*  $S_i$ , принимающей  $q$  различных значений:

$$S_i = -1 + \frac{2(k-1)}{q-1}, \quad i = 1, \dots, N; \quad k = 1, \dots, q. \quad (1)$$

В другом подходе, различные состояния нейронов изображаются точками на единичной окружности. Один из вариантов такой сети был рассмотрен в [4, 5] — *фазорная модель*. Другой вариант — в [6] (*циклическая модель*).

В фазорной сети для изображения различных состояний нейронов используются двумерные векторы единичной длины (*фазоры*). Сеть состоит из  $N$  фазоров  $S_i$ ,  $i = 1, \dots, N$  — комплексных чисел, по модулю равных единице. В первоначальной версии  $S_i$  могли принимать непрерывный ряд значений (непрерывная модель), а в более поздней — только  $q$  корней уравнения  $S_i^q = 1$  (дискретная модель). Пусть  $p$  паттернов — это  $p$  наперед заданных  $N$ -мерных наборов  $(\xi_1^\mu, \xi_2^\mu, \dots, \xi_N^\mu)$ ,  $\mu = 1, \dots, p$ , где  $\xi_j^\mu = \exp(i\theta_j^\mu)$  — состояние  $i$ -го нейрона в  $\mu$ -м паттерне. Тогда синаптическая связь задается выражением

$$C_{ij} = (1 - \delta_{ij}) \sum_{\mu=1}^p \xi_i^\mu \bar{\xi}_j^\mu; \quad (2)$$

черта сверху означает здесь комплексное сопряжение. Динамическое решающее правило определяется с помощью (комплексного) локального поля:

$$h_i = \sum_j C_{ij} S_j, \quad S_i(t+1) = \frac{h_i}{|h_i|}. \quad (3)$$

В случае последовательной динамики энергия состояния

$$E = -\frac{1}{2} \sum_{i \neq j} C_{ij} S_j \bar{S}_i$$

монотонно убывает. Эта сеть изучалась также в [7].

Близкая по духу циклическая модель нейронной сети рассматривалась в [6]. Гамильтониан этой системы имеет вид

$$H = - \sum_{i \neq j}^N \sum_{\mu} \cos \frac{2\pi}{q} \left\{ (n_i - \xi_i^{\mu}) - (n_j - \xi_j^{\mu}) \right\},$$

где  $N$  — число нейронов,  $n_i = 0, 1, \dots, (q-1)$  — состояние  $i$ -го нейрона, а  $\xi_i^{\mu}$  — состояние  $i$ -го нейрона в  $\mu$ -м паттерне, записанном в память сети ( $\mu = 1, \dots, p$ ). При  $q = 2$  эта схема сводится к модели Хопфилда.

Очевидной особенностью гамильтониана этой модели является его инвариантность относительно преобразования  $\{n_i\} \rightarrow \{n_i + k \pmod{q}\}$ , где  $k$  — целое число. Связи между фазорной сетью и циклической сетью исследовались в [8], где установлено, что эти модели имеют много общего. Емкость памяти всех описанных выше моделей достигает своего максимума при  $q = 2$ , когда они превращаются в модель Хопфилда, а затем, с ростом  $q$ , емкость памяти падает.

Альтернативный подход к проблеме состоит в том, что  $q$  различных состояний нейронов изображаются  $q$ -мерными векторами. Исторически первой работой такого рода была статья [9], где рассматривалась модель, получившая название *Поттс-стекольной нейросети* (термин заимствован из физики). Здесь использовался так называемый *анизотропный* гамильтониан (см. ниже; некоторые подробности см. также в [10]). Согласно оценкам автора статьи [9], емкость памяти такой сети должна превосходить емкость памяти модели Хопфилда приблизительно в  $q(q-1)/2$  раз. При больших значениях  $q$  это является весомым фактором. Данная модель исследовалась также в работах [11–13].

Поттс-стекольная модель с *изотропным* гамильтонианом изучалась в [10], а  $q$ -мерное обобщение непрерывной фазорной модели (когда состояния нейронов изображаются  $q$ -мерными векторами с вещественными координатами) — в работе [8]. Емкость памяти этих моделей оказалась меньше емкости памяти модели Хопфилда.

По-видимому, последней по времени создания моделью ассоциативной памяти на  $q$ -мерных нейронах является *параметрическая нейронная сеть* (ПНС), предложенная в [14, 15]. Сформулированная как вариант сети, ориентированной на нелинейно-оптические принципы обработки информации, эта модель была затем формализована в рамках векторного подхода к описанию нейронов [16].

Использование для оценки емкости памяти *статистической техники Чебышева–Чернова* [17,18] позволило установить, что среди всех  $q$ -нарных сетей именно ПНС обладает наилучшими показателями по объему памяти и помехоустойчивости. К изложению этого круга вопросов мы и переходим. Вначале будут описаны Поттс-стекольная и параметрическая нейросети, а затем, на примере анализа последней, будет изложена статистическая техника Чебышева–Чернова.

### Поттс-стекольная нейросеть и параметрическая нейронная сеть

1°. Все варианты Поттс-стекольной ассоциативной памяти являются производными от известной модели магнетика, построенной Поттсом, которая обобщает *модель Изинга* на случай спиновой переменной, принимающей не два значения  $\{-1, +1\}$ , а произвольное число  $q$  различных значений [19–21]. Во всех случаях переход от модели магнетика к модели нейронной сети происходил по одной и той же схеме, с помощью которой модель Изинга была в свое время связана с моделью Хопфилда [22].

А именно, характерное для физических рассмотрений короткодействующее взаимодействие между двумя соседними спинами заменялось межсвязями хеббовского типа между векторами-нейронами. В результате возникающего в системе дальнего действия оказывалось возможным применить для вычисления статистической суммы приближение среднего поля и построить фазовую диаграмму системы. Различные области фазовой диаграммы интерпретировались затем в терминах способности или неспособности сети к восстановлению зашумленных паттернов.

Поттс-стекольную сеть с анизотропным гамильтонианом [9] можно описать следующим образом.

Сеть составлена из  $N$  нейронов, каждый из которых может находиться в одном из  $q$  различных состояний. При этом  $l$ -ое состояние нейрона изображается  $q$ -мерным вектором-столбцом  $\vec{\xi}_l$ , у которого  $l$ -я координата пропорциональна  $(q - 1)$ , а все остальные координаты пропорциональны  $-1$ :

$$l \longrightarrow \vec{\xi}_l = \frac{1}{q} \begin{pmatrix} -1 \\ \vdots \\ q - 1 \\ \vdots \\ -1 \end{pmatrix}, \quad l = 1, \dots, q.$$

Состояние всей сети задается набором из  $N$   $q$ -мерных векторов:  $X = (\vec{x}_1, \dots, \vec{x}_N)$ , где  $\vec{x}_i = \vec{\xi}_{l_i}$ ,  $1 \leq l_i \leq q$ . Паттерны, информация о которых хранится в межсвязях сети — это  $p$  наперед заданных наборов по  $N$   $q$ -мерных векторов каждый:

$$X^\mu = (\vec{x}_1^\mu, \vec{x}_2^\mu, \dots, \vec{x}_N^\mu), \quad \mu = 1, \dots, p,$$

где

$$\vec{x}_i^\mu = \vec{\xi}_{l_i^\mu}, \quad 1 \leq l_i^\mu \leq q.$$

Межсвязь между  $i$ -м и  $j$ -м нейронами задается  $(q \times q)$ -матрицей  $\mathbf{J}_{ij}$ , которая в духе хеббовского правила аккумулирует состояния  $i$ -го и  $j$ -го нейронов во всех  $p$  паттернах

$$\mathbf{J}_{ij} = (1 - \delta_{ij}) \sum_{\mu=1}^p (., \vec{x}_j^\mu) \vec{x}_i^\mu. \quad (4)$$

Здесь через  $(., \vec{x})\vec{y}$  обозначена  $(q \times q)$ -матрица<sup>1</sup>, которая произвольный вектор  $\vec{z} \in \mathbb{R}^q$  переводит в вектор  $\vec{y}$ , домножая его на величину скалярного произведения  $(\vec{z}, \vec{x})$ :  $(., \vec{x})\vec{y}\vec{z} = (\vec{z}, \vec{x})\vec{y}$ . (Заметим, что физическая традиция использует другое обозначение для этой матрицы:  $\vec{y}\vec{x}^\Gamma$ , где  $\vec{x}^\Gamma = (x_1, \dots, x_q)$  —  $q$ -мерная вектор-строка. Формальное матричное умножение вектор-столбца  $\vec{y}$ , расположенного слева, на вектор-строку  $\vec{x}^\Gamma$ , расположенную справа, дает  $(q \times q)$ -матрицу  $(y_k x_l)_{k,l=1}^q$ , которая у нас обозначена через  $(., \vec{x})\vec{y}$ . В то же время, скалярное произведение векторов  $\vec{x}$  и  $\vec{y}$ , в физических обозначениях, имеет вид:  $\vec{x}^\Gamma \vec{y} = \vec{y}^\Gamma \vec{x}$ .)

Локальное поле  $\vec{h}_i$ , действующее на  $i$ -й нейрон, определяется аналогично (3)

$$\vec{h}_i = \sum_{j=1}^N \mathbf{J}_{ij} \vec{x}_j = \sum_{\mu=1}^p \vec{x}_i^\mu \sum_{\substack{j=1 \\ j \neq i}}^N (\vec{x}_j, \vec{x}_j^\mu),$$

а динамика сети задается решающим правилом: в следующий момент времени  $i$ -й нейрон переходит в состояние, максимизирующее  $(\vec{\xi}_i, \vec{h}_i)$ ,

$$\vec{x}_i(t+1) = \vec{\xi}_i, \quad \text{где } (\vec{\xi}_i, \vec{h}_i) \geq (\vec{\xi}_{l'}, \vec{h}_i) \quad \forall l' = 1, \dots, q.$$

<sup>1</sup>О чем свидетельствует полужирный шрифт.



При  $q = 2$  такая сеть эквивалентна стандартной модели Хопфилда [9]. Оценка емкости памяти этой сети при  $N \gg 1$  будет приведена в следующем разделе.

2°. Параметрическая нейронная сеть опирается на *параметрический нейрон* — обладающий кубической нелинейностью элемент, способный к преобразованию и генерации частот в процессах параметрического четырехволнового смешения [23].

Нейроны в ПНС обмениваются по межсвязям квазимонохроматическими импульсами на  $q$  различных частотах  $\{\omega_l\}_{l=1}^q$ . Кроме того, импульсы имеют амплитуды, равные  $\pm 1$ . Таким образом, нейроны в ПНС могут находиться в  $2q$  различных состояниях. Будем описывать состояния нейронов с помощью снабженных амплитудами  $\pm 1$  единичных ортов  $\vec{e}_l$  пространства  $\mathbf{R}^q$ :

$$\vec{x}_i = x_i \vec{e}_{l_i}, \text{ где } x_i = \pm 1, \vec{e}_l = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in \mathbf{R}^q, \begin{cases} i = 1, \dots, N; \\ l = 1, \dots, q; \\ 1 \leq l_i \leq q. \end{cases} \quad (5)$$

Состояние сети как целого задается набором  $N$  таких  $q$ -мерных векторов  $X = (\vec{x}_1, \dots, \vec{x}_N)$ , а паттерны, записанные в межсвязях сети — это  $p$  наперед заданных наборов, состоящих из  $N$   $q$ -мерных векторов каждый:

$$X^\mu = (\vec{x}_1^\mu, \vec{x}_2^\mu, \dots, \vec{x}_N^\mu), \quad \mu = 1, \dots, p,$$

где

$$\vec{x}_i^\mu = x_i^\mu \vec{e}_{l_i^\mu}, \quad x_i^\mu = \pm 1, \quad 1 \leq l_i^\mu \leq q.$$

Межсвязь между  $i$ -м и  $j$ -м нейронами задается  $(q \times q)$ -матрицей  $\mathbf{J}_{ij}$  аналогично выражению (4):

$$\mathbf{J}_{ij} = (1 - \delta_{ij}) \sum_{\mu=1}^p (\cdot, \vec{x}_j^\mu) \vec{x}_i^\mu,$$

а локальное поле  $\vec{h}_i$ , действующее на  $i$ -й нейрон, имеет вид:

$$\vec{h}_i = \sum_{j=1}^N \mathbf{J}_{ij} \vec{x}_j = \sum_{\mu=1}^p \vec{x}_i^\mu \sum_{\substack{j=1 \\ j \neq i}}^N (\vec{x}_j, \vec{x}_j^\mu) = \sum_{l=1}^q A_{il} \vec{e}_l,$$

где амплитуды  $A_{il}$  равны:

$$A_{il} = \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{\mu=1}^p (\vec{x}_i^\mu, \vec{e}_l) (\vec{x}_j, \vec{x}_j^\mu). \quad (6)$$

Динамика сети задается более сложным решающим правилом, чем в случае Поттс-стекольной модели. Пусть  $A_{ik}$  — амплитуда, в данный момент времени наибольшая по модулю среди всех амплитуд (6):

$$|A_{ik}| = \max_{1 \leq l \leq q} |A_{il}|. \quad (7)$$

Тогда  $i$ -й нейрон в следующий момент времени  $(t+1)$  принимает значение:

$$\vec{x}_i(t+1) = \text{sign}(A_{ik}) \vec{e}_k. \quad (8)$$

Иначе говоря,  $i$ -й вектор-нейрон ориентируется в направлении, ближайшем к направлению локального поля  $\vec{h}_i$ .

Эволюция сети состоит в последовательной переориентации векторов  $\vec{x}_i$  по правилам (7), (8). Условимся, что когда максимальное по модулю значение принимают несколько амплитуд и нейрон находится в одном из этих неуправляемых состояний, его состояние не меняется. Тогда нетрудно показать, что каждый шаг эволюции сети будет сопровождаться понижением энергии

$$E = -\frac{1}{2} \sum_{i=1}^N (\vec{h}_i, \vec{x}_i) = -\sum_{i,j=1}^N (\mathbf{J}_{ij} \vec{x}_j, \vec{x}_i).$$

Рано или поздно система «свалится» в локальный минимум по энергии. Все нейроны  $\vec{x}_i$  будут при этом ориентированы неуправляемым образом и эволюция сети прекратится. Такие состояния суть *неподвижные точки* системы. Необходимым и достаточным условием того, чтобы состояние  $X$  было неподвижной точкой, является выполнение системы неравенств:

$$(\vec{x}_i, \vec{h}_i) \geq |(\vec{e}_l, \vec{h}_i)|, \quad \forall l = 1, \dots, q; \quad \forall i = 1, \dots, N. \quad (9)$$

Заметим в заключение, что при  $q = 1$  ПНС переходит в стандартную модель Хопфилда: векторы  $\vec{x}_i$  (5) превращаются в обычные бинарные переменные  $x_i = \pm 1$ , а все остальное здесь просто копирует модель Хопфилда.

### Статистическая техника Чебышева–Чернова

Дадим асимптотическую оценку емкости памяти ПНС при  $N \gg 1$ . Пусть начальным состоянием сети является искаженный  $m$ -й паттерн

$$\tilde{X}^m = (a_1 \hat{b}_1 \bar{x}_1^m, a_2 \hat{b}_2 \bar{x}_2^m, \dots, a_N \hat{b}_N \bar{x}_N^m).$$

Независимые случайные величины  $\{a_i\}_1^N$  и  $\{\hat{b}_i\}_1^N$  задают мультипликативный шум. Случайная величина  $a_i$  с вероятностью  $a$  принимает значение  $-1$ , а с вероятностью  $(1 - a)$  — значение  $1$  (в оптической терминологии — *амплитудный шум*). Случайный оператор  $\hat{b}_i$  с вероятностью  $b$  изменяет состояние  $i$ -го нейрона на другое, а с вероятностью  $(1 - b)$  оставляет вектор  $\bar{x}_i^m$  неизменным (в оптической терминологии — *частотный шум*). Оценим, при каких условиях  $m$ -й паттерн будет распознаваться сетью правильно.

Амплитуды  $A_{il}$  (6) имеют вид:

$$A_{il} = \begin{cases} x_i^m \sum_{j=1}^{N-1} \xi_j + \sum_{r=1}^L \eta_r, & \text{для } l = l_i^{(m)}; \\ \sum_{r=1}^L \eta_r, & \text{для } l \neq l_i^{(m)}, \end{cases}$$

где

$$\begin{aligned} \xi_j &= a_j(\bar{x}_j^m, \hat{b}_j \bar{x}_j^m), \\ \eta_r &\equiv \eta_j^\mu(l) = a_j(\bar{e}_l, \bar{x}_i^\mu)(\bar{x}_j^\mu, \hat{b}_j \bar{x}_j^\mu), \\ & j = 1, \dots, N, \quad j \neq i, \\ & \mu = 1, \dots, p, \quad \mu \neq m \end{aligned}$$

и, для простоты, величины  $\eta$  проиндексированы обобщенным индексом  $r = (j, \mu)$ , который принимает  $L = (N - 1)(p - 1)$  различных значений; зависимость величин  $\eta$  от  $l$  в дальнейшем никакой роли не играет, поэтому  $l$  можно опустить.

Для рандомизированного набора паттернов  $\{X^\mu\}_{\mu=1}^p$  случайные величины  $\xi_j$  и  $\eta_r$  независимы и имеют распределения

$$\begin{aligned} \xi_j &= \begin{cases} +1, & \text{с вероятностью } (1 - b)(1 - a), \\ 0, & \text{с вероятностью } b, \\ -1, & \text{с вероятностью } (1 - b)a; \end{cases} \\ \eta_r &= \begin{cases} +1, & \text{с вероятностью } 1/2q^2, \\ 0, & \text{с вероятностью } 1 - 1/q^2, \\ -1, & \text{с вероятностью } 1/2q^2. \end{cases} \end{aligned} \quad (10)$$

Для того, чтобы  $i$ -й нейрон перешел в состояние  $\vec{x}_i^m$ , согласно (9) одновременно должно выполняться

$$\text{sign}(A_{ij^{(m)}}) = x_i^m, \quad \sum_{j=1}^{N-1} \xi_j + x_i^m \sum_{r=1}^L \eta_r \geq \left| \sum_{r=1}^L \eta_r \right|.$$

В противном случае, произойдет ошибка распознавания вектор-координаты  $\vec{x}_i^m$ . Поскольку случайная величина  $x_i^m \eta_r$  имеет то же самое распределение, что и  $\eta_r$ , вероятность ошибки распознавания можно представить в виде:

$$\text{Pr}_i = \text{Pr} \left\{ \sum_{j=1}^{N-1} \xi_j + \sum_{r=1}^L \eta_r < 0 \right\}. \quad (11)$$

Для оценки вероятности (11) при  $N \gg 1$  воспользуемся известной техникой Чебышева-Чернова [17, 18]. Для начала, запишем:

$$\text{Pr}_i \leq \text{Pr} \left\{ \sum_{j=1}^{N-1} \xi_j + \sum_{r=1}^L \eta_r \leq 0 \right\} = \text{Pr} \left\{ - \sum_{j=1}^{N-1} \xi_j - \sum_{r=1}^L \eta_r \geq 0 \right\}.$$

Далее, используя экспоненциальные оценки чебышевского типа [24], для любого положительного  $z > 0$ , получаем:

$$\text{Pr}_i \leq \exp \left[ z \left( - \sum_{j=1}^{N-1} \xi_j - \sum_{r=1}^L \eta_r \right) \right] = \left[ \overline{\exp(-z\xi_j)} \left( \overline{\exp(-z\eta_r)} \right)^{p-1} \right]^{N-1}.$$

Черта сверху означает усреднение по всем возможным реализациям, а последнее равенство следует из независимости случайных величин  $\xi_j$  и  $\eta_r$ .

С учетом (10), легко получить выражения для средних

$$\overline{\exp(-z\xi_j)} = (1-a)(1-b)e^{-z} + b + a(1-b)e^z,$$

$$\overline{\exp(-z\eta_r)} = e^{-z}/2q^2 + 1 - 1/q^2 + e^z/2q^2.$$

Делая здесь замену переменных  $e^z = y$  и вводя функции  $f_1(y)$  и  $f_2(y)$ ,

$$f_1(y) = a(1-b)y + b + (1-a)(1-b)/y,$$

$$f_2(y) = 1/2q^2 (y + 1/y) + 1 - 1/q^2,$$

получаем, что при любом положительном  $y$  для  $\text{Pr}_i$  справедливо:

$$\text{Pr}_i \leq \left[ f_1(y) f_2^{p-1}(y) \right]^{N-1}. \quad (12)$$

Для получения наименьшей оценки вероятности  $\text{Pr}_i$  необходимо отыскать значение переменной  $y$ , минимизирующее правую часть (12). Это приводит к уравнению:

$$(p-1)(y^2-1) + \frac{a(1-b)y^2 - (1-a)(1-b)}{a(1-b)y^2 + by + (1-a)(1-b)} [y^2 + 2(q^2-1)y + 1] = 0.$$

В случае  $p \gg 1$  интересующий нас корень этого уравнения, с точностью до членов более высокого порядка малости по  $1/p$ , равен

$$y_1 = 1 + \frac{q^2(1-2a)(1-b)}{(p-1)}.$$

Подставляя  $y_1$  в правую часть (11), окончательно получаем оценку для вероятности неправильного распознавания вектор-координаты  $\vec{x}_i^m$ :

$$\text{Pr}_i \leq \left[ 1 - \frac{q^2(1-2a)^2(1-b)^2}{2(p-1)} \right]^{N-1} \cong \exp \left[ -\frac{N(1-2a)^2}{2p} \cdot q^2(1-b)^2 \right].$$

В результате, для вероятности ошибки распознавания паттерна  $X^m$  имеем:

$$\text{Pr}_{err} = N \exp \left[ -\frac{N(1-2a)^2}{2p} \cdot q^2(1-b)^2 \right]. \quad (13)$$

С ростом  $N$  вероятность ошибки распознавания стремится к нулю, если  $p$  растет медленнее, чем

$$p_c = \frac{N(1-2a)^2}{2 \ln N} \cdot q^2(1-b)^2. \quad (14)$$

Оценку (14) можно рассматривать как *асимптотически достижимую емкость памяти* ПНС.

При  $q = 1$  выражения (13) и (14) превращаются в известные результаты для модели Хопфилда (в этом случае  $b = 0$ ). С ростом  $q$  экспоненциально уменьшается вероятность ошибки распознавания (13) — существенно растет *помехоустойчивость сети*. Одновременно, пропорционально  $q^2$  растет

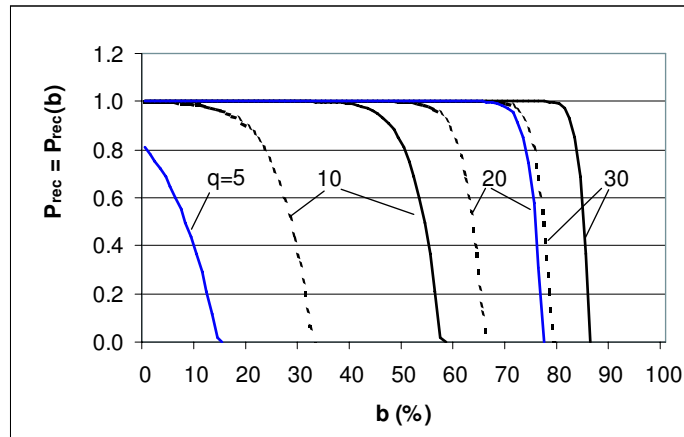


Рис. 1. Вероятность распознавания паттерна  $P_{rec} = 1 - Pr_{err}$  как функция частотного шума  $b$  (в процентах) для различных значений  $q = 5, 10, 20, 30$  при числе паттернов, вдвое большем числа нейронов,  $\alpha = p/N = 2$ . Здесь: сплошная линия — ПНС, штриховая линия — Поттс-стеклянная нейросеть.

и емкость памяти (14). В отличие от модели Хопфилда, оказывается возможным эффективным запоминание большего, чем  $N$ , числа паттернов  $p$ .

На рис. 1, для различных значений  $q$ , сплошной линией показана зависимость вероятности правильного распознавания  $P_{rec} = 1 - Pr_{err}$  от величины частотного шума  $b = b \cdot 100\%$ ,  $b \in [0, 1]$  при числе паттернов, вдвое большем числа нейронов ( $p = 2N$ ) и  $a = 0$ . Мы видим, что для  $q = 20$  практически со стопроцентной вероятностью будет правильно восстановлен любой паттерн, зашумленный не более, чем на 70%, а для  $q = 30$  — любой паттерн, зашумленный не более, чем на 85%. Вообще, если уровень шума  $b$  меньше критического значения  $b_c = 1 - 2/q\sqrt{p/N}$ , ПНС практически со стопроцентной вероятностью восстановит зашумленный паттерн. Если же  $b > b_c$ , вероятность правильного распознавания паттерна стремится к нулю. Машинные эксперименты подтверждают эти оценки.

Аналогично можно оценить емкость памяти Поттс-стекольной нейросети. Проводя вычисления, подобные тем, что делались выше, для рандомизированного набора паттернов получим:

$$\text{Pr}_{err} = N \exp \left[ -\frac{N}{2p} \frac{q(q-1)}{2} (1 - \bar{b})^2 \right], \quad \bar{b} = \frac{q}{q-1} b, \quad (15)$$

$$p_c = \frac{N}{2 \ln N} \frac{q(q-1)}{2} (1 - \bar{b})^2. \quad (16)$$

При  $q = 2$  эти выражения переходят в известные оценки для модели Хопфилда. В то же время, при  $q \gg 1$  емкость памяти Поттс-стекольной нейросети в два раза меньше емкости памяти ПНС (ср. (16) с (14) при  $a = 0$ ). Эта двойка становится существенной, когда речь заходит о помехоустойчивости сети — для вероятности неправильного распознавания эта двойка попадает в показатель экспоненты (см. (13) и (15)), что ведет к заметному снижению помехоустойчивости Поттс-стекольной модели по сравнению с ПНС (особенно, в области умеренных значений  $q \sim 10$ ). Это хорошо видно на рис. 1, где для Поттс-стекольной модели штриховой линией показана зависимость вероятности правильного распознавания  $\mathbf{P}_{rec}$  от величины частотного шума  $\mathbf{b}$  при тех же условиях, что и для ПНС (сплошная линия).

В заключение заметим, что в основе хороших емкостных характеристик обеих рассмотренных нами моделей лежит одно и то же — высокая степень фильтрации внутренних шумов как на стадии передачи сигналов по межсвязям, так и в процессе выполнения решающего правила. Это приводит к уменьшению дисперсии внутреннего шума на величину порядка  $q^2$ , что и обеспечивает успех. В других  $q$ -нарных моделях ассоциативной памяти такая фильтрация полностью отсутствует. Более того, с ростом  $q$  все больше дробится шкала, градуирующая различные состояния нейронов (см. (1)), и все меньше становится разница между двумя их различными состояниями. В то же время, дисперсия шума остается практически неизменной. Это и приводит к уменьшению емкости памяти этих моделей по сравнению с моделью Хопфилда.

Работа выполнялась при финансовой поддержке РФФИ (гранты 02-01-00457 и 01-01-00090) и программы «Интеллектуальные компьютерные системы» (проект 2. 45).

**Литература**

1. *Rieger H.* Storing an extensive number of grey-toned patterns in a neural network using multistate neurons // *J. Phys. A*, v. 23, pp. L1273–L1279 (1990).
2. *Bolle D., van Mourik J.* Capacity of diluted multi-state neural networks // *J. Phys. A*, v. 27, pp. 1151–1162 (1994).
3. *Bolle D., Shim G. M., Vinck B., Zagrebnov V. A.* // *J. Stat. Phys.*, v. 74, p. 565 (1987).
4. *Noest A. J.* // *Europhys. Letters*, v. 6, p. 469 (1988).
5. *Noest A. J.* Discrete-state phasor neural networks. // *Phys. Rev. A*, v. 38, pp. 2196–2199 (1988).
6. *Cook J.* The mean-field theory of a Q-state neural network model // *J. Phys. A*, v. 22, pp. 2000–2012 (1989).
7. *Gerl F., Bauer K., Krey U.* // *Z. Phys. B*, v. 88, p. 339 (1992).
8. *Nakamura Y., Torii K., Munaka T.* Neural-network model composed of multidimensional spin neurons // *Phys. Rev. B*, v. 51, pp. 1538–1546 (1995)
9. *Kanter I.* Potts-glass models of neural networks. // *Phys. Rev. A*, v. 37, pp. 2739–2742 (1988).
10. *Vogt H., Zippelius A.* Invariant recognition in Potts glass neural networks // *J. Physics A*, v. 25, pp. 2209–2226 (1992).
11. *Bolle D., Dupont P., van Mourik J.* Stability properties of Potts neural networks with biased patterns and low loading // *Journal of Physics A*, v. 24, pp. 1065–1081 (1991).
12. *Bolle D., Dupont P., Huyghebaert J.* Thermodynamics properties of the  $q$ -state Potts-glass neural network // *Phys. Rev. A*, v. 45, pp. 4194–4197 (1992).
13. *Bolle D., Jongen G., Shim G. M.* Q-Ising neural network dynamics: A comparative review of various architectures // *Cond-mat /9907390* (1999).
14. *Крыжановский Б. В., Микаэлян А. Л.* О распознающей способности нейросети на нейронах с параметрическим преобразованием частот // *ДАН (мат.-физ.)*, т. 65(2), с. 286–288 (2002).
15. *Fonarev A., Kryzhanovsky B. V. et al.* Parametric dynamic neural network recognition power // *Optical Memory and Neural Networks*, v. 10(4), pp. 31–48 (2001).
16. *Крыжановский Б. В., Литинский Л. Б.* Векторные модели ассоциативной памяти // *Теоретическая и математическая физика*, 2003 (в печати)
17. *Chernov N.* A measure of asymptotic efficiency for tests of hypothesis based on the sum of observations // *Ann. Math. Statistics*, v. 23, pp. 493–507 (1952).
18. *Kryzhanovsky B. V., Mikaelyan A. L., Koshelev V. N., Fonarev A.* On recognition error bound for associative Hopfield memory // *Optical Memory and Neural Networks*, v. 9(4), pp. 267–276 (2000).



19. Potts R.B. // *Proc. Camb. Phil. Soc.*, v. 48, p. 106 (1952).
20. Wu F. Y. The Potts model // *Review of Modern Physics*, v. 54(1), pp. 235–268 (1982).
21. Бэкстер Р. Точно решаемые модели статистической физики. – М.: Мир, 1985.
22. Amit D., Gutfreund H., Sompolinsky H. Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks // *Phys. Rev. Lett.*, v. 55, pp. 1530–1533 (1985); *See also*: Information storage in neural networks with low levels of activity // *Phys. Rev. A*, v. 35, pp. 2293–2303 (1987).
23. Bloembergen N. *Nonlinear optics*. 1966.
24. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 206 с.

**Борис Владимирович КРЫЖАНОВСКИЙ**, заместитель директора Института оптико-нейронных технологий РАН (Москва), доктор физико-математических наук. Область научных интересов: нейронные сети, квантовая механика, нелинейная оптика. Имеет более 80 научных публикаций.

**Леонид Борисович ЛИТИНСКИЙ**, старший научный сотрудник, кандидат физико-математических наук, заведующий сектором Динамических нейронных сетей в Институте оптико-нейронных технологий РАН, Москва. Область научных интересов: ассоциативные нейронные сети, распознавание образов, многоэкстремальные задачи. Имеет более 30 научных публикаций.