УДК 621.391.1

©2001 г. **Б.В.Крыжановский, В.Н.Кошелев, А.Б.Фонарев**

# ОЦЕНКА ЭФФЕКТИВНОСТИ РАНДОМИЗИРОВАННОЙ ПАМЯТИ ХОПФИЛДА[1]

Изучается распознающая способность сети Хопфилда, т.е. нейронной сети, взаимодействие элементов в которой определяется преобразованием Хебба. Приводится формально строгий вывод верхней границы вероятности ошибки распознавания рандомизированных объектов, основанный на стандартной технике оценивания вероятности больших уклонений по методу Чебышева–Чернова.

## §1. Введение

В этой статье изучается рандомизированная модель ассоциативной памяти Хопфилда и выводятся некоторые оценки информационной емкости этой памяти и ее разрешающей способности. Основным стимулом написания статьи послужило ознакомление авторов с работами [1-3], в которых феномен ассоциативной памяти исследовался как свойство формальной нейронной сети переходить из возбужденного состояния в ближайшее устойчивое состояние, соответствующее некоторому локальному минимуму ее энергии. Для оценки числа устойчивых состояний в [1] был использован метод рандомизации, т.е. рассматривался некоторый вероятностный ансамбль нейронных сетей и вычислялась средняя по ансамблю "вероятность ошибки", т.е. вероятность того, что возбужденный нейрон перейдет в состояние, не совпадающее с тем, какое он должен был бы иметь в ближайшем устойчивом состоянии сети. Такая модель затухания возбуждения в нейронной сети представляется очень близкой к модели декодирования сигнала на выходе шумящего канала связи. При этом ошибка распознавания как вероятностное событие имеет вид большого уклонения суммы независимых одинаково распределенных случайных величин от математического ожидания этой суммы. В схемах вероятностного кодирования метод рандомизации в конечном счете сводится к оценке вероятности больших уклонений, для чего успешно используется техника экспоненциальных оценок типа Чебышева–Чернова [4]. Было бы естественно проследить этот подход к оценке вероятности ошибки также и в нейросетевых вероятностных моделях. Метод рандомизации описан в работах [1,2], однако вместо прямого применения упомянутых экспоненциальных оценок там проводится цепочка косвенных рассуждений, основанных на асимптотической нормальности суммы н.о.р.с.в., а в работе [3] для оценки мощности ассоциативной памяти строится даже некоторый специально разработанный теоретико-числовой аппарат. Упомянутые работы относятся к периоду 1987-1989 гг., и если судить по более позднему обзору [6] (1998), метод случайного кодирования в применении к нейронным сетям остался, по-видимому, недоработанным. Как следует из списка литературы работы [6], вопрос об эффективности

[1]Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (номера проектов 99-01-00325, 00-01-00213).

рандомизированной сети Хопфилда более не затрагивался, поэтому техника вывода вероятности ошибки, предлагаемая в этой статье, может представить интерес как формально строгое завершение вопроса, поставленного в работе [1].

Выводимая нами оценка вероятности ошибки распознавания в сети Хопфилда асимптотически совпадает с указанной в [1]; однако в отличие от последней и в доасимптотическом виде, т.е. при малых значениях параметров сети, наша оценка остается нетривиальной и может оказаться полезной при изучении сетевых макроконструкций, организуемых на базе малоразмерных нейросетей.

Ассоциативная память реализуется в виде сети, составленной из взаимодействующих между собой элементов ("нейронов"), и по мере роста их числа приобретает свойство статистической устойчивости. Последнее проявляется не только в том, что сеть надежно запоминает поступающие извне сигналы, но также и в том, что обнаруживает способность к их распознаванию, относя близкие друг другу сигналы в один и тот же раздел памяти. В этом смысле сеть действует как декодер на выходе канала связи: зная a priori список всех допустимых сигналов, декодер ассоциирует принятый сигнал с одним из них. Сходство ассоциативной памяти со схемой помехоустойчивого кодирования подсказывает, что для ее изучения может быть использована стандартная техника, применяемая в теории вероятностного кодирования.

Впервые модель ассоциативной памяти была описана в работе Хопфилда [7] как формальная нейронная сеть [8], которая (при достаточно большом числе нейронов) способна отождествлять наблюдаемые сигналы с "эталонами", возникающими в памяти по мере роста числа наблюдений. Тот факт, что способность к отождествлению возникает в системе довольно простой математической природы, вызвало большой интерес к модели Хопфилда. Действительно, проявление в той или иной системе признаков ассоциативного поведения, т.е. способности идентифицировать наблюдаемые объекты по мере их схожести, всегда воспринимается как элемент целенаправленной деятельности – свойство, заслуживающее самого тщательного изучения. В этом смысле можно было бы сказать, что декодер тоже действует целенаправленно, однако его действия заранее обусловлены целым рядом хорошо подобранных факторов: выбором хорошей кодовой книги, оптимального алгоритма декодирования и т.д. Однако нейронная сеть *изначально проста* и не предполагает никакой предварительной оптимизации ее устройства. Именно поэтому ее способность к принятию ассоциативных решений следует рассматривать как нейробиологический феномен, требующий формального математического объяснения.

## §2. Нейронная сеть как распознающее устройство

Введем основные определения.

Нейронная сеть состоит из $N$ узлов, называемых нейронами и принимающих значения

$$x_n \in \{-1, +1\}, \qquad n \in \overline{1, N}.$$

Нейроны взаимодействуют друг с другом через ребра сети, называемые синапсами, при этом $i$-й нейрон взаимодействует с $j$-м нейроном через $(i, j)$-й синапс, имеющий проводимость $T_{ij}$. Таким образом, формально *нейронная сеть определяется заданием пары*

$$(N, \boldsymbol{T}),$$

*где* $N \geq 1$ *– размер сети, а*

$$\boldsymbol{T} = [T_{ij}] \tag{1}$$

*–* $N \times N$*-матрица ее проводимостей.* В работе сети различаются три момента времени: начальный, переходный и конечный. Пусть вектор-строка

$$x^N = (x_1, x_2, \ldots, x_N) \in \{-1, +1\}^N \tag{2}$$

означает начальное состояние сети, т.е. численное значение всех ее нейронов в начальный момент времени. В следующий, т.е. переходный, момент времени начальное состояние, т.е. вектор (2), умножением на матрицу (1) преобразуется в $N$-мерный евклидов вектор

$$y^N \triangleq x^N \boldsymbol{T} \in \mathbb{R}^N, \tag{3}$$

который затем переводится в конечное состояние

$$z^N \triangleq \operatorname{sign} y^N \in \{-1, +1\}^N, \tag{4}$$

где sign означает покомпонентную операцию

$$z_n \triangleq \operatorname{sign} y_n = \begin{cases} -1, & y_n \leq 0, \\ +1, & y_n > 0. \end{cases}$$

Подчеркнем, что мы рассматриваем синхронную модель; это значит, что переход из начального состояния в конечное

$$x^N \to y^N \to z^N \tag{5}$$

происходит в синхронном режиме, т.е. все элементы сети срабатывают одновременно. (В этом смысле размер сети $N$ есть аналог длины блока в блоковой схеме кодирования.) Переход (5) есть результат двух последовательных операций: линейной операции (3), осуществляюшей отображение

$$\{-1, +1\}^N \to \mathbb{R}^N,$$

и нелинейной операции (4), осуществляющей отображение в обратном направлении

$$\mathbb{R}^N \to \{-1, +1\}^N. \tag{6}$$

Их последовательное объединение будем обозначать как оператор

$$\boldsymbol{G} : \{-1, +1\}^N \to \{-1, +1\}^N. \tag{7}$$

Формально оператор (7) нелинейный, однако операция (6), делающая его таковым, достаточно проста и позволяет применять к нему некоторые стандартные рассуждения из теории линейных операторов. Подчеркивая, что оператор (7) существенно зависит от выбора матрицы проводимостей $\boldsymbol{T}$, обозначим его

$$\boldsymbol{G} = \boldsymbol{G}(\boldsymbol{T}). \tag{8}$$

Существуют ли у оператора (8) неподвижные (или, в некотором смысле, “почти неподвижные”) точки и если существуют, то как их отыскать? Назовем (условно) такую постановку вопроса – задан оператор, требуется найти его неподвижные точки, – традиционной. Однако с точки зрения тех задач, на решение которых в основном и направлена разработка теории нейронных сетей, более осмысленной представляется обратная к традиционной постановка вопроса: пусть в пространстве состояний сети задан набор из $M \geq 2$ точек:

$$\boldsymbol{B} = \left\{ b_m^N \in \{-1,+1\}^N, \quad m \in \overline{1,M} \right\}; \tag{9}$$

существует ли (отличная от единичной!) $(N \times N)$-матрица

$$\boldsymbol{T} = \boldsymbol{T}(\boldsymbol{B})$$

такая, при которой точки (9) были бы неподвижными (или “почти неподвижными”) точками оператора

$$\boldsymbol{G}(\boldsymbol{T}(\boldsymbol{B}))\,?$$

Ясно, что матрицу проводимостей сети надо выбирать зависящей от набора точек (9). Но какова форма этой зависимости? Именно на этот вопрос отвечает способ построения сети, предложенный Хопфилдом.

## §3. Сеть Хопфилда

Будем записывать точки подмножества (9) как вектор-строки, а само подмножество – как $(M \times N)$-матрицу

$$\boldsymbol{B} = \begin{pmatrix} b_1^N \\ b_2^N \\ \dots \\ b_M^N \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1N} \\ b_{21} & b_{22} & \dots & b_{2N} \\ \dots & \dots & \dots & \dots \\ b_{M1} & b_{M2} & \dots & b_{MN} \end{pmatrix}. \tag{10}$$

**Определение.** *Пусть заданы целые $N \geq 1$, $M \geq 1$, $(M \times N)$-матрица $\boldsymbol{B}$ и некоторое $a \geq 0$. Тогда сеть*

$$\left(N, \boldsymbol{T}^{(H)}(\boldsymbol{B})\right),$$

*где*

$$\boldsymbol{T}^{(H)}(\boldsymbol{B}) \stackrel{\triangle}{=} \boldsymbol{B'}\boldsymbol{B} - a\boldsymbol{I}, \tag{11}$$

*будем называть сетью Хопфилда.*

В (11) штрих означает транспонирование.

Очевидно, матрица $\boldsymbol{B'}\boldsymbol{B}$ симметрична, а ее $(i,j)$-й элемент представляет собой скалярное произведение $i$-го и $j$-го столбцов матрицы $\boldsymbol{B}$:

$$T_{ij}^{(H)}(\boldsymbol{B}) = (b_{1i} b_{2i} \dots b_{Mi}) \begin{pmatrix} b_{1j} \\ b_{2j} \\ \dots \\ b_{Mj} \end{pmatrix}. \tag{12}$$

Поскольку элементы матрицы $\boldsymbol{B}$ суть $\pm 1$, то все диагональные элементы в $\boldsymbol{B'B}$ равны $M$ и, таким образом, не зависят от выбора матрицы $B$; далее мы убедимся в том, что выбирая в (11) диагональное смещение

$$a = M,$$

мы поступаем, в определенном смысле, наилучшим образом. Скалярное произведение (12) выражает корреляцию между $i$-м и $j$-м столбцами матрицы (10). На корреляционную природу взаимодействия нейронов, т.е. на усиление проводимости $(i, j)$-го синапса пропорционально величине (12), было впервые указано в работе Хебба [9]. Хопфилд [7] обратил внимание на то, что для значений $M = M(N)$, достаточно медленно растущих с ростом $N$, нейронная сеть, проводимость которой определяется по Хеббу, обладает следующими свойствами: оператор этой сети

$$\boldsymbol{G}\left(\boldsymbol{T}^{(H)}(\boldsymbol{B})\right),$$

будучи применен к пространству $\{-1, +1\}^N$, оставляет строки матрицы (10) “почти неподвижными”, а всякую другую точку этого пространства переводит в точку, расположенную в окрестности ближайшей из строк этой же матрицы.

Попытаемся, вслед за [1], придать точный смысл этому утверждению. Однако прежде чем переходить к формальной стороне дела, представляется полезным описать физическую ситуацию, адекватную модели Хопфилда.

Имеется “внешний мир”, в котором возможны $M$ явлений, и имеется “память”, состоящая из $N$ нейронов. Наступление того или иного внешнего явления приводит к “возбуждению” памяти, в результате чего каждый нейрон принимает значение $-1$ либо $+1$. Связь между явлением и возбуждением, вообще говоря, не однозначная, но “статистически устойчивая”, т.е. все возможные реакции памяти в ответ на одно и то же явление, как правило, “близки” друг другу, так что для каждого

$$m \in \overline{1, M} \tag{13}$$

можно указать некоторую точку

$$b_m^N \in \{-1, +1\}^N, \tag{14}$$

играющую роль “центра тяжести” всех возбуждений, возможных в ответ на $m$-е явление. Будем понимать (14) как “эталонное возбуждение”, или “эталон”, а всякое реальное возбуждение как “отклонение” от эталона. Соответственно, матрицу $\boldsymbol{B}$, введенную в (10), будем называть матрицей эталонов. Следуя [1], будем рассматривать простейшую статистическую модель отклонений – мультипликативную модель без памяти, согласно которой реальное возбуждение $x^N$ получается из эталона $b_m^N$ путем независимых покомпонентных искажений, так что

$$x^N \triangleq b_m^N \Theta^N \triangleq (b_{m1}\Theta_1, b_{m2}\Theta_2, \ldots, b_{mN}\Theta_N), \tag{15}$$

где $\Theta^N \in \{-1, +1\}^N$ означает последовательность независимых одинаково распределенных случайных величин

$$\Theta_n = \begin{cases} -1, & p, \\ +1, & 1-p, \end{cases} \qquad n \in \overline{1, N}, \quad 0 \le p \le \frac{1}{2}. \tag{16}$$

Например, если внешний мир есть мир животных, а число $m$ в (13) означает *кошку как вид*, то вектор (14) означает набор признаков, присутствующих (+1) или отсутствующих (−1) у *виртуальной* кошки-эталона; тогда (15) означает *реальную* случайно наблюдаемую кошку, а (16) – вектор отклонений этой кошки от эталонной. Способна ли сеть Хопфилда опознать предъявленную ей реальную кошку как животное, относящееся к $m$-му виду? В данной работе мы попытаемся ответить на этот вопрос, проследив типичную схему рассуждений, применяемых при доказательстве прямой теоремы кодирования с использованием стандартных оценок вероятности больших уклонений, известных как экспоненциальные оценки Чебышева–Чернова [4,5].

## §4. Вероятность ошибки

Из вышесказанного следует, что нейросетевая система, определяемая пятеркой

$$(N, M, \boldsymbol{T}, \boldsymbol{B}, p)$$

и предназначенная для распознавания $M$ равновероятных явлений, совершает ошибку с вероятностью

$$\begin{aligned} P_{\text{ош}}(\boldsymbol{B}) = P_{\text{ош}}(N, M, \boldsymbol{T}, \boldsymbol{B}, p) = \\ &= \sum_{m=1}^{M} M^{-1}\mathbf{P}\{\operatorname{sign}[(b_m^N\Theta^N)\boldsymbol{T}] \neq b_m^N\} = && \text{(17а)} \\ &= \sum_{m=1}^{M} M^{-1}\mathbf{P}\{\bigcup_{n=1}^{N}(\operatorname{sign}[(b_m^N\Theta^N)\boldsymbol{T}]_n \neq b_{mn})\} \leq && \text{(17б)} \\ &\leq \sum_{m=1}^{M} M^{-1}\sum_{n=1}^{N}\mathbf{P}\{\operatorname{sign}[(b_m^N\Theta^N)\boldsymbol{T}]_n \neq b_{mn}\}. && \text{(17в)} \end{aligned}$$

В (17а)–(17в) вероятность $\mathbf{P}\{\cdot\}$ вычисляется в соответствии с р.в.

$$\mathbf{P}\{\Theta^N = t^N\} = \prod_{n=1}^{N}\mathbf{P}\{\Theta_n = t_n\} = (1-p)^{N-|t^N|}p^{|t^N|}, \tag{18}$$

где

$$|t^N| \triangleq \sum_{n=1}^{N}\chi\{t_n = -1\}$$

означает вес мультипликативной помехи $\Theta^N = t^N$; в правой части равенства (17а) операция sign, примененная к вектору $x^N\boldsymbol{T}$, означает, как сказано выше, покомпонентное преобразование

$$\operatorname{sign}[x^N\boldsymbol{T}]_n \qquad \forall n \in \overline{1, N},$$

где $[x^N\boldsymbol{T}]_n$ означает $n$-ую компоненту вектора $x^N\boldsymbol{T}$; в (17б) вероятность ошибки при распознавании $m$-го явления записана как вероятность суммы покомпонентных ошибок, а неравенство (17в) следует из оценки вероятности суммы событий суммой их вероятностей.

Для дальнейшего изучения оценки (17) применим метод рандомизации, т.е. будем осреднять ее по множеству всех матриц (9). Следуя [1], рассмотрим случайную $(M \times N)$-матрицу

$$\mathcal{B} = \left( \begin{array}{cccc} \beta_{11} & \beta_{12} & \ldots & \beta_{1N} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2N} \\ \multicolumn{4}{c}{\dotfill} \\ \beta_{M1} & \beta_{M2} & \ldots & \beta_{MN} \end{array} \right), \tag{19}$$

состоящую из $MN$ независимых, одинаково распределенных случайных величин вида

$$\beta_{mn} = \left\{ \begin{array}{ll} +1, & \frac{1}{2}, \\ -1, & \frac{1}{2}, \end{array} \right. \qquad \forall (m,n) \in \overline{1,M} \times \overline{1,N}. \tag{20}$$

Осредняя неравенство (17), видим, что средняя по ансамблю (19),(20) вероятность ошибки

$$\begin{gathered} \overline{P_{\text{ош}}(\mathcal{B})} = \overline{P_{\text{ош}}(N, M, \boldsymbol{T}, \mathcal{B}, p)} \leq \\ \leq \sum_{m=1}^{N} M^{-1} \sum_{n=1}^{N} \mathbf{P}\{\operatorname{sign}[(\beta_m^N \Theta^N)\boldsymbol{T}]_n \neq \beta_{mn}\}, \end{gathered} \tag{21}$$

где вероятность $\mathbf{P}\{\cdot\}$ вычисляется теперь в соответствии с совместным р.в.

$$\mathbf{P}\{\mathcal{B} = \boldsymbol{B}; \Theta^N = t^N\} = \mathbf{P}\{\mathcal{B} = \boldsymbol{B}\}\mathbf{P}\{\Theta^N = t^N\}. \tag{22}$$

Заметим, что в (22) первый сомножитель

$$\mathbf{P}\{\mathcal{B} = \boldsymbol{B}\} = 2^{-NM},$$

а второй сомножитель указан в (18). Все слагаемые в правой части (21) одинаковы, поэтому, полагая $m = n = 1$, находим, что

$$\overline{P_{\text{ош}}(\mathcal{B})} \leq N\mathbf{P}\{\operatorname{sign}[(\beta_1^N \Theta^N)\boldsymbol{T}]_1 \neq \beta_{11}\}. \tag{23}$$

Так как

$$\mathbf{P}\{\beta_{11} = -1\} = \mathbf{P}\{\beta_{11} = +1\} = \frac{1}{2},$$

то вероятность в правой части (23) достаточно оценивать для случая $\beta_{11} = -1$. Таким образом,

$$\begin{gathered} \overline{P_{\text{ош}}(\mathcal{B})} \leq N\mathbf{P}\left\{\operatorname{sign}[(\beta_1^N \Theta^N)\boldsymbol{T}]_1 = +1 \Big/ \beta_{11} = -1\right\} = \\ = N\mathbf{P}\left\{\sum_{n=1}^{N} \beta_{1n}\Theta_n T_{n1} \geq 0 \Big/ \beta_{11} = -1\right\}, \end{gathered} \tag{24}$$

где $T_{n1}, \quad n \in \overline{1,N},$ суть компоненты первого столбца матрицы $\boldsymbol{T}$. Оценка (24) справедлива для любой матрицы $\boldsymbol{T}$, и дальнейшее изучение вероятности ошибки требует конкретного выбора матрицы проводимостей сети.

### §5. Вероятность ошибки в сети Хопфилда

Применим оценку (24) для вычисления вероятности ошибки в рандомизированной сети Хопфилда. Подставим в (11) вместо $\boldsymbol{B}$ случайную матрицу $\mathcal{B}$, и в получившейся матрице

$$\boldsymbol{T}^{(H)}(\mathcal{B}) \overset{\triangle}{=} \mathcal{B}'\mathcal{B} - a\boldsymbol{I}$$

рассмотрим элементы первого столбца:

$$\mathcal{T}_{n1} \overset{\triangle}{=} \sum_{m=1}^{M} \beta_{mn}\beta_{m1} - a\chi\{n=1\}, \quad n \in \overline{1,N}, \tag{25}$$

где

$$\chi\{n=1\} = \begin{cases} 1, & n=1, \\ 0, & n \in \overline{2,N}. \end{cases}$$

Подставляя в последнее из выражений (24) вместо величин $T_{n1}$ величины (25), находим, что

$$N^{-1}\overline{P_{\text{ош}}(N,M,\mathcal{B},p)} \le \mathbf{P}\left\{\sum_{n=1}^{N} \beta_{1n}\Theta_n\mathcal{T}_{n1} \ge 0 \Big/ \beta_{11} = -1\right\} =$$

$$= \mathbf{P}\left\{\sum_{n=1}^{N} \beta_{1n}\Theta_n \left(\sum_{m=1}^{M} \beta_{m1}\beta_{mn} - a\chi\{n=1\}\right) \ge 0 \Big/ \beta_{11} = -1\right\}. \tag{26}$$

Рассмотрим сумму $\sum\limits_{n=1}^{N} \ldots$, указанную под знаком условной вероятности. Выделяя слагаемые с $n=1$ и замечая, что $\beta_{mn}^2 = 1$, получаем

$$\sum_{n=1}^{N} \ldots = \sum_{n=1}^{N} \beta_{1n}\Theta_n \left(\beta_{11}\beta_{1n} + \sum_{m=2}^{M} \beta_{m1}\beta_{mn} - a\chi\{n=1\}\right) =$$

$$= \beta_{11}\Theta_1 \left(\beta_{11}^2 + \sum_{m=2}^{M} \beta_{m1}^2 - a\right) + \sum_{n=2}^{N} \beta_{1n}\Theta_n \left(\beta_{11}\beta_{1n} + \sum_{m=2}^{M} \beta_{m1}\beta_{mn}\right) =$$

$$= \beta_{11}\Theta_1(M-a) + \sum_{n=2}^{N} \Theta_n\beta_{11} + \sum_{n=2}^{N} \beta_{1n}\Theta_n \sum_{m=2}^{M} \beta_{m1}\beta_{mn}\,. \tag{27}$$

При $\beta_{11} = -1$ выражение (27) принимает значение

$$\left(\sum_{n=1}^{N} \cdots\right)\Bigg|_{\beta_{11}=-1} = -\Theta_1(M-a) - \sum_{n=2}^{N} \Theta_n + \sum_{n=2}^{N}\sum_{m=2}^{M} \Theta_n\beta_{1n}\beta_{m1}\beta_{mn}\,. \tag{28}$$

Подставляя его в (26), находим, что

$$N^{-1}\overline{P_{\text{ош}}(N,M,\mathcal{B},p)} \le \mathbf{P}\left\{\sum_{n=2}^{N}\sum_{m=2}^{M} \Theta_n\beta_{1n}\beta_{m1}\beta_{mn} \ge \Theta_1(M-a) + \sum_{n=2}^{N} \Theta_n\right\}.$$

Левая часть неравенства, записанного под знаком вероятности, представляет собой сумму $(M-1)(N-1)$ случайных величин вида

$$\beta_{1n}\Theta_n\beta_{m1}\beta_{mn} \overset{\triangle}{=} \eta_l \in \{-1,+1\}, \quad l \in \overline{1,(M-1)(N-1)}.$$

Все они суть независимые копии случайной величины

$$\eta = \begin{cases} -1, & \frac{1}{2}, \\ +1, & \frac{1}{2}. \end{cases} \tag{29}$$

Чтобы убедиться в этом, выпишем все участвующие в правой части неравенства (28) переменные в виде $(M \times N)$-матрицы

$$\begin{pmatrix} \beta_{11}\Theta_1 & \beta_{12}\Theta_2 & \beta_{13}\Theta_3 & \ldots & \beta_{1N}\Theta_N \\ \beta_{21} & \beta_{22} & \beta_{23} & \ldots & \beta_{2N} \\ \beta_{31} & \beta_{32} & \beta_{33} & \ldots & \beta_{3N} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \beta_{M1} & \beta_{M2} & \beta_{M3} & \ldots & \beta_{MN} \end{pmatrix}.$$

Все элементы этой матрицы (за исключением $(1,1)$-го) взаимно независимы и принимают значения $\pm 1$ с вероятностью $\frac{1}{2}$. Если элементы $n$-го столбца, $n \in \overline{2,N}$, умножить на $\beta_{1n}\Theta_n$, а элементы $m$-й строки, $m \in \overline{2,M}$, умножить на $\beta_{m1}$ , то все элементы в указанных столбцах и строках останутся н.о.р. вида (29). Используя эти обозначения, перепишем оценку (28) в виде

$$N^{-1}\overline{P_{\text{ош}}(N,M,\mathcal{B},p)} \le \mathbf{P}\left\{\Theta_1(M-a) + \sum_{n=2}^{N}\Theta_n \le \sum_{l=1}^{(M-1)(N-1)} \eta_l\right\}. \tag{30}$$

Величина $a$ означает произвольное смещение диагональных элементов в матрице (11), поэтому правую часть оценки (30), вообще говоря, следовало бы минимизировать по этому параметру. Интуитивно, однако, ясно, что при фиксированном $p > 0$ эта оптимизация существенно не улучшит асимптотику оценки при $N \to \infty$. Действительно, если взять $a < M$, то с вероятностью

$$\mathbf{P}\{\Theta_1 = -1\} = p$$

левая часть неравенства, записанного под знаком вероятности, получает смещение $-(M-a)$; если же взять $a > M$, то это же смещение возникает с вероятностью

$$\mathbf{P}\{\Theta_1 = +1\} = 1 - p.$$

В любом случае произойдет существенное уменьшение скорости экспоненциального убывания оценки (30) по сравнению с той, которая достигается при выборе $a = M$. Если же выбрать $a = M$, то произойдет лишь уменьшение размерности схемы – она понизится на единицу, что не повлияет на асимптотику оценки. Полагая $a = M$ и заменяя для удобства $M$ и $N$ на $M+1$ и $N+1$ соответственно, перепишем неравенство (30) в виде

$$(N+1)^{-1}\overline{P_{\text{ош}}(N+1;M+1;\mathcal{B};p)} \le \mathbf{P}\left\{\sum_{n=1}^{N}\Theta_n \le \sum_{l=1}^{MN}\eta_l\right\}. \tag{31}$$

Далее воспользуемся известной техникой Чебышева–Чернова [4], согласно которой для любого $\rho \geq 0$

$$\mathbf{P}\left\{\sum_{n=1}^{N}\Theta_n \leq \sum_{l=1}^{MN}\eta_l\right\} = \overline{\chi\left\{\sum_{n=1}^{N}\Theta_n \leq \sum_{l=1}^{MN}\eta_l\right\}} =$$
$$= \overline{\chi\left\{1 \leq e^{\rho\left(\sum\limits_{l=1}^{MN}\eta_l - \sum\limits_{n=1}^{N}\Theta_n\right)}\right\}} \leq \overline{e^{\rho\left(\sum\limits_{l=1}^{MN}\eta_l - \sum\limits_{n=1}^{N}\Theta_n\right)}} =$$
$$= \left(\overline{e^{-\rho\Theta}}\right)^N \left(\overline{e^{\rho\eta}}\right)^{MN} = \left[\left((1-p)e^{-\rho} + pe^{\rho}\right)\left(\frac{e^{-\rho}+e^{\rho}}{2}\right)^M\right]^N. \tag{32}$$

Минимизируя правую часть (32) по $\rho \geq 0$, имеем

$$\mathbf{P}\left\{\sum_{n=1}^{N}\Theta_n \leq \sum_{l=1}^{MN}\eta_l\right\} \leq \left[\min_{\rho\geq 0}\left(\frac{e^{-\rho}+e^{\rho}}{2}\right)^M \left((1-p)e^{-\rho} + pe^{\rho}\right)\right]^N. \tag{33}$$

Заметим, что минимизируемое выражение как функция от $\rho$ выпукло книзу. Дифференцируя по $\rho$ и приравнивая производную нулю, находим, что его минимум достигается при $\rho$, удовлетворяющем уравнению

$$(M+1)pe^{4\rho} + (M-1)(1-2p)e^{2\rho} - (M+1)(1-p) = 0. \tag{34}$$

Последнее представляет собой квадратное уравнение, решение которого

$$e^{2\rho} = \frac{\sqrt{(M-1)^2 + 16Mp(1-p)} - (1-2p)(M-1)}{2p(M+1)} \tag{35}$$

и доставляет искомый минимум в правой части (33).

## §6. Анализ

Начнем со случая

$$p = 0, \tag{36}$$

который следует рассмотреть отдельно, чтобы показать, насколько хорошо работает сеть Хопфилда "без внешней нагрузки", т.е. когда для идентификации ей предъявляется один из $M+1$ порождающих ее эталонов в его неискаженном виде. Подставляя (36) в (31) и (33), находим, что

$$(N+1)^{-1}\overline{P_{\text{ош}}(N+1; M+1; \mathcal{B}; 0)} \leq$$
$$\leq \mathbf{P}\left\{\sum_{l=1}^{MN}\eta_l \geq N\right\} \leq \left[\min_{\rho\geq 0}\left(\frac{e^{-\rho}+e^{\rho}}{2}\right)^M e^{-\rho}\right]^N; \tag{37}$$

подставляя (36) в уравнение (34) и решая последнее, находим, что минимум в правой части (37) достигается при

$$e^{2\rho} = \frac{M+1}{M-1}, \tag{38}$$

а сама оценка принимает вид

$$\begin{gathered}(N+1)^{-1}\overline{P_{\text{ош}}(N+1;M+1;\mathcal{B};0)} \leq \\ \leq \left[(1-M^{-1})^{M-1}(1+M^{-1})^{M+1}\right]^{-\frac{N}{2}} = \\ = e^{-\frac{N}{2M}\left(1+2\frac{M+6}{(M-1)^3}+\cdots\right)}. \end{gathered} \tag{39}$$

Из теоремы о среднем следует, что при любых целых $N$ и $M$ существует, по крайней мере, одна $((M+1)\times(N+1))$-матрица эталонов $\boldsymbol{B}$ такая, что вероятность ошибки, которую совершает порождаемая этой матрицей сеть Хопфилда при распознавании случайно предъявленного ей эталона, не превосходит величины, указанной в правой части неравенства (39).

Теперь рассмотрим случай, когда вероятность искажения эталона положительна, т.е. когда

$$0 < p \leq \frac{1}{2}\,.$$

Подставляя в (31) последовательно (33) и (35) и производя упрощения, аналогичные тем, которые произведены в (39), находим, что

$$\overline{P_{\text{ош}}(N+1;M+1;\mathcal{B};p)} \leq (N+1)e^{-\frac{N}{2M}(1-2p)^2(1+o(1))}\,,$$

где $o(1)\to 0,\quad M\to\infty$. Полученное неравенство устанавливает *верхнюю границу для средней вероятности ошибки в рандомизированной сети Хопфилда с параметрами* $(N+1;M+1;p)$. С ростом $N$ эта граница сходится к нулю всякий раз, когда величина $M$ как функция от $N$, растет медленнее, чем

$$M(N) \triangleq \frac{N}{2\ln N}(1-2p)^2. \tag{40}$$

Согласно [1] это дает основание рассматривать величину (40) как асимптотически достижимую мощность ассоциативной памяти Хопфилда. Следует признать, что это очень низкая мощность. Если воспользоваться рекомендациями теории информации и строить ассоциативную память на основе метода максимального правдоподобия, принимая решение в целом по всем $N$ наблюдаемым признакам, то мощность памяти была бы близка к величине $2^{NC(p)}$, где

$$C(p) = 1-\mathcal{H}(p), \quad 0 \leq p < \frac{1}{2},$$

– пропускная способность двоичного канала наблюдения, указанного в (16). К сожалению, нейробиологическая эволюция проходила в дошенноновскую эпоху, поэтому природа воспользовалась более простым принципом посимвольного принятия решений с мощностной характеристикой (40).

## СПИСОК ЛИТЕРАТУРЫ

1. *McEliece R.J., Posner E.C., Rodemich E.R., Venkatesh S.S.* The Capacity of the Hopfield Associative Memory // IEEE Trans. Inform. Theory. 1987. V.33. №4. P.461-482.

2. *Kuh A., Dickinson B.W.* Information Capacity of Associative Memories // IEEE Trans Inform. Theory. 1989. V.35. №1. P.59-68.

3. *Sussmann H.J.* On the Number of Memories that Can Be Perfectly Stored in a Neural Net with Hebb Weights // IEEE Trans. Inform. Theory. 1989. V.35. №1. P.174-178.

4. *Chernov H.* A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations // Ann. Math. Statistics. 1952. V.23. P.493-507.

5. *Shannon C.E.* Certain results in coding theory for noisy channels // Information and Control. 1957. V.1. P.6-25.

6. *Kulkarni S.R., Lugosi G., Venkatesh S.S.* Learning Pattern Classification - A Survey // IEEE Trans. Inform. Theory. 1998. V.44. №6. P.2178-2204.

7. *Hopfield J.J.* Neural Networks and Physical Systems with Emergent Collective Computational Capabilities // Proc. Nat. Acad. Sci. USA. 1982. V.79. P.2554-2558.

8. *McCulloch W.S., Pitts W.* A Logical Calculus of the Idieas Imminent in Neural Activity // Bull. Math. Biophys. 1943. V.5. P.115-133.

9. *Hebb D.O.* The Organization of Behavior //New-York: Wiley. 1949.



Авторы:
1) Кошелев Валерий Николаевич
д.ф.-м.н., НСК РАН; 135-62-11
117292 Москва ул.Профсоюзная, 8-2-510
2) Крыжановский Борис Владимирович
д.ф.-м.н., ИОНТ РАН; 135-63-31
143952 г.Реутов ул.Советская, 33, кв.32
3) Фонарев Анатолий Борисович
д.ф.-м.н.;
New-York
City University of New-York
USA