

© 2003 г. **Б.В. КРЫЖАНОВСКИЙ**, д-р физ.-мат. наук,
Л.Б. ЛИТИНСКИЙ, к-т физ.-мат. наук

(Институт оптико-нейронных технологий РАН, Москва)

ВЕКТОРНЫЕ МОДЕЛИ АССОЦИАТИВНОЙ ПАМЯТИ ¹

Модель Хопфилда позволяет эффективно запомнить сравнительно небольшое число исходных образов – порядка 15% от размера нейронной сети. Существенно превзойти этот показатель удастся 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), реализующая нелинейно-оптические принципы передачи и обработки сигналов. Развит формализм, позволяющий единым образом описывать как Поттс-стекольную ассоциативную память, так и ПНС. Для оценки емкости памяти используется статистическая техника Чебышева-Чернова.

1. Введение: параметрическая нейронная сеть

В [1,2] было показано, что искусственная нейронная сеть, устроенная на оптических принципах обработки и хранения информации, обладает существенными преимуществами по сравнению с другими моделями ассоциативной памяти – такими, как модель Хопфилда и проекционная матрица связей [3], сеть Хэмминга [4], фазорные сети [5] и др. За основу сети в [1,2] был принят *параметрический нейрон* – обладающий кубической нелинейностью элемент, способный к преобразованию и генерации частот в процессах параметрического четырехволнового смешения [6]. Соответствующую сеть было предложено назвать *параметрической нейронной сетью* (ПНС). Схематически принцип ее работы можно представить следующим образом.

Сеть составлена из N нейронов. Нейрон будем представлять себе как сложную структуру, состоящую из: 1) сумматора входных сигналов; 2) набора q идеальных частотных фильтров с собственными частотами

$$(1) \quad \omega_1, \omega_2, \dots, \omega_q;$$

3) блока сравнения отфильтрованных сигналов по амплитуде и 4) q генераторов квазимонохроматических импульсов на частотах из набора (1).

¹Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты 02-01-00457, 03-01-00355 и 01-01-00090) и программы "Интеллектуальные компьютерные системы" (проект 2.45).

Будем использовать в дальнейшем переменную l (возможно с индексами) только для нумерации частот $\{\omega_l\}_1^q$.

Состояние каждого нейрона в данный момент времени характеризуется частотой из набора (1) и фазой ψ . Например, можно отождествить каждый нейрон с элементарным пикселем экрана, а то или иное состояние нейрона - с цветом, в который окрашен данный пиксель. Цвет характеризуется частотой ω квазимонохроматического импульса $\exp(i\omega t + i\psi)$ и фазой ψ . Условимся, что фазы могут принимать только одно из двух значений: 0 или π . Тогда у квазимонохроматических импульсов появятся амплитуды ± 1 . Поскольку частота ω отвечает тому или иному цвету (красный, оранжевый, желтый и т.д.), амплитуду можно трактовать как интенсивность цвета: $+1$ отвечает интенсивности "темный", а -1 - интенсивности "светлый". Таким образом, состояние нейрона отвечает окраске соответствующего пикселя в "темно-красный" цвет или в "светло-красный", в "темно-оранжевый" или "светло-оранжевый" и т.д.:

$$\kappa_i = \pm \exp(i\omega_{l_i} t), \quad i = 1, \dots, N; \quad \omega_{l_i} \in \{\omega_l\}_1^q.$$

Здесь i - номер нейрона, κ_i - его состояние, характеризующееся частотой ω_{l_i} сигнала.

Набор одновременных состояний всех N нейронов отвечает какому-то цветному изображению на экране: $K = (\kappa_1, \kappa_2, \dots, \kappa_N)$. Введем в рассмотрение p наперед заданных исходных образов (*паттернов*) - p фиксированных цветных изображений, информация о которых хранится в памяти сети:

$$(2) \quad K^{(\mu)} = (\kappa_1^{(\mu)}, \kappa_2^{(\mu)}, \dots, \kappa_N^{(\mu)}), \quad \mu = 1, \dots, p.$$

Символ μ , заключенный всегда в скобочки и используемый только как верхний индекс, будет служить для нумерации паттернов из набора (2). Состоянию i -го нейрона в μ -м паттерне отвечает частота $\omega_{l_i^{(\mu)}}$, где $1 \leq l_i^{(\mu)} \leq q$, $i = 1, \dots, N$, $\mu = 1, \dots, p$.

Пусть состояние j -го нейрона в данный момент времени характеризуется частотой ω_{l_j} - нейрон испускает квазимонохроматический импульс $\exp(i\omega_{l_j} t)$ длительностью τ . Примем для простоты, что τ - величина, кратная всем периодам собственных колебаний нейрона

$$(3) \quad \frac{\tau}{\tau_l} - \text{целое, где } \tau_l = \frac{2\pi}{\omega_l} \quad \forall l = 1, \dots, q.$$

Память сети локализована в межсвязях T_{ij} , $i, j = 1, \dots, N$, где аккумулирована информация о состояниях i -го и j -го нейронов во всех p паттернах. Будем полагать, что межсвязи устроены динамически и организованы по правилу Хейбба:

$$(4) \quad \begin{cases} T_{ij} = \sum_{\mu=1}^p \kappa_i^{(\mu)} \kappa_j^{(\mu)*}, & i \neq j; \\ T_{ii} = 0; \end{cases} \quad i, j = 1, \dots, N.$$

Квазимонохроматический импульс с частотой ω_{l_j} , распространяющийся по (ij) -й межсвязи от j -го нейрона к i -му, участвует в процессах параметрического четырехволнового смешения с аккумулярованными межсвязью состояниями i -го и j -го нейронов в паттернах (2):

$$(5) \quad \omega_{l_i^{(\mu)}} - \omega_{l_j^{(\mu)}} + \omega_{l_j} \rightarrow \{\omega_l\}_1^q.$$

Амплитуды ± 1 , которые, возможно, имеются у квазимонохроматических импульсов, при этом перемножаются. Просуммировав результаты этих парциальных преобразований по всем μ , получим целый пакет квазимонохроматических импульсов, вообще говоря – на всех частотах из набора (1), с определяющимися межсвязью амплитудами. Этот пакет – результат преобразования (ij) -й межсвязью импульса ω_{l_j} – и приходит к i -му нейрону.

Пакеты квазимонохроматических импульсов, пришедшие к i -му нейрону по всем межсвязям, суммируются. Затем суммарный сигнал проходит через q параллельно соединенных частотных фильтров $\{\omega_l\}_1^q$. Выходные сигналы с фильтров сравниваются по амплитуде. И, наконец, сигнал с максимальной по модулю амплитудой инициирует генерацию i -м нейроном выходного импульса, частота и фаза которого совпадают с частотой и фазой инициирующего сигнала (*победитель получает все*).

Важным элементом излагаемой схемы является выдвинутый в [1,2] *принцип несоизмеримости частот* $\{\omega_l\}_1^q$: *никакая комбинация частот вида $\omega_l - \omega_{l'} + \omega_{l''}$ не может принадлежать набору частот (1), когда все три частоты различны*. На деле принцип несоизмеримости означает, что сигнал с частотой ω_{l_j} , участвующий в параметрическом четырехволновом смешении с хранящимися в межсвязи частотами $\omega_{l_i^{(\mu)}}$ и $\omega_{l_j^{(\mu)}}$ (см. соотношение (5)), по межсвязи не проходит, если все три частоты различны, – в этом случае сигнал ω_{l_j} межсвязью фильтруется. Описание принципа работы ПНС на этом закончено.

Несмотря на то, что описанная конструкция гораздо сложнее бинарных моделей ассоциативной памяти, имеется целый ряд аргументов в ее пользу. Во-первых, язык частотно-фазовой модуляции является естественным при оптической обработке сигналов и позволяет отказаться от искусственной адаптации оптической нейросети к амплитудно модулированным сигналам. Во-вторых, передача по межсвязям сигналов на q различных частотах (аналог уплотнения канала), фактически, позволяет в q^2 раз уменьшить число межсвязей, которые в нейросетевых архитектурах и так занимают до 98% площади нейрокристалла. И, наконец, в-третьих: проведенный в [1,2] анализ соотношения сигнал/шум в сочетании со статистической техникой Чебышева-Чернова [7,8] показывает, что для рандомизированного набора паттернов емкость па-

мноти ПНС приблизительно в q^2 раз превосходит емкость памяти модели Хопфилда. Даже при умеренных значениях $q \sim 10$ это дает выигрыш на два порядка. Заметим, что для компьютерной обработки цветных изображений стандартом является значение $q = 256$; в этом случае выигрыш по сравнению с моделью Хопфилда составит 4-5 порядков.

Вместе с увеличением емкости памяти растет и помехоустойчивость сети. В компьютерном моделировании было, например, установлено, что ПНС из 100 параметрических нейронов ($N = 100$), передающих сигналы на 22 различных частотах ($q = 22$), в память которой записаны 200 рандомизированных паттернов ($p = 200$), восстанавливает зашумленный на 80% паттерн за 100 тактов работы – за один проход по всем 100 нейронам. Та же сеть с параметрами $N = 100$, $q = 25$, $p = 1000(!)$ восстанавливала зашумленный на 65% паттерн за 4-5 проходов по всем 100 нейронам. Напомним, что в свое время демонстрацией распознающей способности модели Хопфилда считалось восстановление ею зашумленного на 30% паттерна при $N = 400$, $p = 30$ [9].

Изложенные аргументы заставляют внимательно исследовать принципы, на которых основана ПНС. Этому и посвящена настоящая работа. Прежде, чем переходить к изложению полученных результатов, сделаем одно замечание.

Возможны различные варианты параметрического четырехволнового смешения, удовлетворяющие принципу несоизмеримости частот. Например, в [1,2] изучались процессы параметрического четырехволнового смешения вида

$$(6) \quad \omega_l - \omega_{l'} + \omega_{l''} = \begin{cases} \omega_{l''}, & \text{когда } l' = l; \\ \omega_l, & \text{когда } l' = l''; \\ \rightarrow 0 & \text{в остальных случаях.} \end{cases}$$

Соответствующую сеть назовем ПНС-I. Однако большая часть излагаемых в настоящей статье результатов получена для параметрического четырехволнового смешения

$$(7) \quad \omega_l - \omega_{l'} + \omega_{l''} = \begin{cases} \omega_l, & \text{когда } l' = l''; \\ \rightarrow 0 & \text{в остальных случаях.} \end{cases}$$

Эту сеть будем называть ПНС-II.

Статья устроена следующим образом. Во втором разделе вводится векторный формализм, позволяющий сформулировать задачу в общем виде. В третьем разделе излагаются результаты, полученные для ПНС-II. В четвертом разделе на языке векторного формализма излагается Поттс-стекольная модель ассоциативной памяти и проводится ее сравнение с ПНС-II. Несколько общих замечаний отнесены в заключение, а технически сложные доказательства вынесены в Приложение.

2. Векторный формализм

Фактически, ПНС реализует ассоциативную память Хопфилдова типа на нейронах, которые могут находиться в большем чем 2 числе состояний. Подобные модели рассматривались и раньше - см., например, [10-15]. Стандартный прием здесь состоит в том, что состояния нейронов моделируются не скалярными величинами (со значениями ± 1 или $0/1$), а векторными, причем число изображающих векторов берут равным числу различных состояний нейронов.

В ПНС нейроны могут находиться в одном из q частотно различных состояний – тех, что определяются различными частотами квазимонохроматических импульсов (1). Кроме того, из-за наличия фазы передаваемые сигналы имеют амплитуды ± 1 . С учетом (3) имеем:

$$\frac{1}{\tau} \int_0^\tau e^{i(\omega_l - \omega_{l'})t} dt = \delta_{ll'}, \text{ где } \delta_{ll'} - \text{символ Кронекера.}$$

Следовательно, моделируя параметрические нейроны векторами, необходимо, во-первых, сопоставить частотно различным состояниям нейрона взаимно ортогональные векторы. И, во-вторых, двум состояниям, отличающимся друг от друга только фазой, следует сопоставить два противоположных по знаку вектора единичной длины. Этим требованиям легко удовлетворить, сопоставив l -му из частотно различных состояний нейрона l -й вектор-орт вещественного q -мерного пространства, а двум возможным значениям фазы $\psi = \{0, \pi\}$ – знаки $+$ и $-$ перед таким ортом:

$$(8) \quad \left\{ \begin{array}{l} \exp(i\omega_l t) \longrightarrow \mathbf{e}_l = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in \mathbf{R}^q, \quad l = 1, \dots, q; \\ \kappa_i = \pm \exp(i\omega_{l_i} t) \longrightarrow \mathbf{x}_i = x_i \mathbf{e}_{l_i}, \quad x_i = \pm 1, \quad i = 1, \dots, N; \quad 1 \leq l_i \leq q. \end{array} \right.$$

Состоянию всей сети отвечает набор из N q -мерных векторов: $X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$, а исходным образам (2) – совокупность фиксированных паттернов

$$(9) \quad X^{(\mu)} = (\mathbf{x}_1^{(\mu)}, \mathbf{x}_2^{(\mu)}, \dots, \mathbf{x}_N^{(\mu)}), \quad \mathbf{x}_i^{(\mu)} = x_i^{(\mu)} \mathbf{e}_{l_i^{(\mu)}}, \quad x_i^{(\mu)} = \pm 1, \quad \left\{ \begin{array}{l} 1 \leq l_i^{(\mu)} \leq q; \\ \mu = 1, \dots, p. \end{array} \right.$$

Поскольку нейроны теперь суть q -мерные векторы, межсвязь между i -м и j -м нейронами следует задавать не скаляром, а $(q \times q)$ -матрицей межсвязи \mathbf{T}_{ij} . Матрицы \mathbf{T}_{ij} действуют на векторы $\mathbf{x}_j \in \mathbf{R}^q$

$$\mathbf{T}_{ij} \mathbf{x}_j = \sum_{l=1}^q (\mathbf{T}_{ij} \mathbf{x}_j, \mathbf{e}_l) \mathbf{e}_l,$$

превращая их в линейную комбинацию ортов \mathbf{e}_l – аналог пакета квазимонохроматических импульсов, который приходит к i -му нейрону от j -го после преобразования в межсвязи (см. предыдущий раздел). При этом необходимо удовлетворить двум требованиям: 1) матрицы \mathbf{T}_{ij} должны определяться состояниями i -го и j -го нейронов во всех p паттернах в духе Хеббовского правила (4); 2) матрицы \mathbf{T}_{ij} должны преобразовывать векторы \mathbf{x}_j таким образом, чтобы воспроизводились все особенности процессов параметрического четырехволнового смешения, на которых основана ПНС, – соотношения (6) или (7). Матрицы \mathbf{T}_{ij} будут определены ниже, а пока, считая их заданными, определим динамику сети из векторов-нейронов.

Пусть $X(t) = (\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_N(t))$ – состояние сети в момент времени t . Локальное поле, действующее на i -й нейрон, дается выражением

$$(10) \quad \mathbf{h}_i(t) = \sum_{j=1}^N \mathbf{T}_{ij} \mathbf{x}_j(t) = \sum_{l=1}^q A_{il}(t) \mathbf{e}_l, \text{ где } A_{il}(t) = \sum_{j=1}^N (\mathbf{T}_{ij} \mathbf{x}_j(t), \mathbf{e}_l).$$

(Обычно употребляемый множитель $1/N$ здесь опущен.) Пусть $A_{ik}(t)$ – амплитуда, наибольшая по модулю среди всех амплитуд $A_{il}(t)$ в выражении (10):

$$(11) \quad |A_{ik}(t)| = \max_{1 \leq l \leq q} |A_{il}(t)|.$$

Положим, что в следующий момент времени, $t + 1$, i -й нейрон принимает значение

$$(12) \quad \mathbf{x}_i(t + 1) = \text{sgn}(A_{ik}(t)) \mathbf{e}_k.$$

Иначе говоря, i -й вектор-нейрон ориентируется в направлении, ближайшем к направлению локального поля $\mathbf{h}_i(t)$. Это и есть эквивалент правила "победитель получает все".

Эволюция сети состоит в последовательной переориентации векторов-нейронов \mathbf{x}_i по правилам (11),(12). Условимся, что когда максимальное по модулю значение принимают несколько амплитуд и нейрон находится в одном из этих *неулучшаемых* состояний, его состояние не меняется. Тогда нетрудно показать, что если для матриц \mathbf{T}_{ij} выполняется равенство $\mathbf{T}_{ji} = \mathbf{T}_{ij}^*$, каждый шаг эволюции сети будет сопровождаться понижением *энергии*

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

Рано или поздно система свалится в локальный минимум по энергии. Все нейроны \mathbf{x}_i будут при этом ориентированы неулучшаемым образом и эволюция сети прекратится. Такие состояния суть неподвижные точки системы. Необходимым и достаточным

условием того, чтобы конфигурация X была неподвижной точкой, является выполнение системы неравенств:

$$(13) \quad (\mathbf{x}_i, \mathbf{h}_i) \geq |(\mathbf{e}_l, \mathbf{h}_i)|, \quad \forall l = 1, \dots, q; \quad \forall i = 1, \dots, N.$$

Определим, наконец, матрицы межсвязи \mathbf{T}_{ij} . Делается это с помощью набора из q^2 элементарных матриц $\{\mathbf{A}_{ll'}\}_{l,l'=1}^q$ размерности $(q \times q)$, в совокупности моделирующих процесс параметрического четырехволнового смещения в ПНС. Достигается это тем, что для каждой матрицы $\mathbf{A}_{ll'}$ явным образом указывается правило преобразования ею ортов $\mathbf{e}_{l''}$ в соответствии с существом избранного типа параметрического четырехволнового смещения. Например, рассмотрения [1,2] относились к четырехволновому смещению (6). На векторно-матричном языке соотношениям (6) отвечают следующие правила преобразования ортов $\mathbf{e}_{l''}$:

$$\mathbf{A}_{ll'}\mathbf{e}_{l''} = \begin{cases} \mathbf{e}_{l''}, & \text{когда } l' = l; \\ \mathbf{e}_l, & \text{когда } l' = l''; \\ \mathbf{0} & \text{когда } l \neq l' \neq l''. \end{cases} \implies \begin{cases} \mathbf{A}_{ll}\mathbf{e}_{l''} = \mathbf{e}_{l''}; \\ \mathbf{A}_{ll'}\mathbf{e}_{l''} = \delta_{l',l''}\mathbf{e}_l, \quad l \neq l', \end{cases} \quad \forall l, l', l'' = 1, \dots, q.$$

Этими правилами однозначно определяется набор элементарных матриц ²:

$$\begin{cases} \mathbf{A}_{ll} = \mathbf{I} - \text{единичная матрица в } \mathbf{R}^q; \\ \mathbf{A}_{ll'} = \mathbf{e}_l\mathbf{e}_{l'}^+, \quad \text{когда } l \neq l'. \end{cases}$$

Параметрическому четырехволновому смещению (7) отвечают другие правила преобразования ортов $\mathbf{e}_{l''}$ и, соответственно, несколько отличающийся набор элементарных матриц:

$$(14) \quad \mathbf{A}_{ll'}\mathbf{e}_{l''} = \delta_{l',l''}\mathbf{e}_l \implies \mathbf{A}_{ll'} = \mathbf{e}_l\mathbf{e}_{l'}^+, \quad \forall l, l', l'' = 1, \dots, q.$$

Фактически, с помощью элементарных матриц $\mathbf{A}_{ll'}$ моделируется парциальный акт параметрического четырехволнового смещения (5):

$$\mathbf{A}_{l_i^{(\mu)}l_j^{(\mu)}}\mathbf{x}_j \iff \omega_{l_i^{(\mu)}} - \omega_{l_j^{(\mu)}} + \omega_{l_j} \rightarrow \{\omega_l\}_1^q.$$

²Согласно физической традиции запись $\mathbf{A}_{ll'} = \mathbf{e}_l\mathbf{e}_{l'}^+$ означает $(q \times q)$ -матрицу, полученную формальным умножением q -мерного вектор-столбца \mathbf{e}_l на q -мерную вектор-строку $\mathbf{e}_{l'}^+$. Верхний индекс "+" здесь означает операцию транспонирования. В этой системе обозначений умножение q -мерной вектор-строки \mathbf{y}^+ на q -мерный вектор-столбец \mathbf{z} есть просто скалярное произведение двух q -мерных векторов: $\mathbf{y}^+\mathbf{z} = (\mathbf{y}, \mathbf{z})$. Пользуясь ассоциативностью матричного умножения, действие матрицы $\mathbf{x}\mathbf{y}^+$ на вектор \mathbf{z} можно записать в виде: $(\mathbf{x}\mathbf{y}^+)\mathbf{z} = \mathbf{x}(\mathbf{y}^+\mathbf{z}) = \mathbf{x}(\mathbf{y}, \mathbf{z})$.

Тогда, чтобы удовлетворить требованию 1), предъявляющемуся к матрицам \mathbf{T}_{ij} , необходимо взять их в виде

$$(15) \quad \begin{cases} \mathbf{T}_{ij} = \sum_{\mu=1}^p \mathbf{A}_{l_i^{(\mu)} l_j^{(\mu)}}, & i \neq j; \\ \mathbf{T}_{ii} = \mathbf{0}; \end{cases} \quad i, j = 1, \dots, N.$$

Действие матрицы \mathbf{T}_{ij} (15) на вектор \mathbf{x}_j моделирует процесс преобразования (ij) -й межсвязью импульса ω_{lj} в духе Хеббовского правила (4).

Заметим в заключение, что при $q = 1$ ПНС переходит в стандартную модель Хопфилда: векторы \mathbf{x}_i (8) превращаются в обычные бинарные переменные $x_i = \pm 1$, а все остальное в выражениях (9)-(15) просто копирует модель Хопфилда.

3. ПНС-II

В настоящем разделе рассматривается ПНС, устроенная на основе семейства элементарных матриц (14). Согласно (15), матрица межсвязи между i -м и j -м нейронами дается выражением

$$\mathbf{T}_{ij} = \sum_{\mu=1}^p \mathbf{x}_i^{(\mu)} \mathbf{x}_j^{(\mu)+}, \quad i \neq j,$$

а локальное поле, действующее на i -й вектор-нейрон \mathbf{x}_i , имеет вид:

$$(16) \quad \mathbf{h}_i = \sum_{\mu=1}^p \mathbf{x}_i^{(\mu)} \sum_{j=1(\neq i)}^N (\mathbf{x}_j, \mathbf{x}_j^{(\mu)}) = \sum_{l=1}^q A_{il} \mathbf{e}_l, \quad \text{где } A_{il} = \sum_{j=1(\neq i)}^N \sum_{\mu=1}^p (\mathbf{x}_i^{(\mu)}, \mathbf{e}_l) (\mathbf{x}_j, \mathbf{x}_j^{(\mu)}).$$

С помощью (16) можно доказать несколько общих утверждений относительно неподвижных точек ПНС-II (они справедливы и для модели Хопфилда).

1) Если X – неподвижная точка сети, то и обратная к ней конфигурация $-X = (-\mathbf{x}_{l_1}, -\mathbf{x}_{l_2}, \dots, -\mathbf{x}_{l_N})$ тоже является неподвижной точкой.

2) Множество неподвижных точек не зависит от расстановки знаков перед паттернами $\{X^{(\mu)}\}_{\mu=1}^p$.

3) Если число паттернов $p \leq 2$, то единственными неподвижными точками сети будут сами эти паттерны.

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

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

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 . Случайный оператор \hat{b}_i с вероятностью b изменяет состояние

i -го нейрона на другое, а с вероятностью $1 - b$ оставляет вектор $\mathbf{x}_i^{(m)}$ неизменным. Иными словами, a определяет вероятность искажения по знаку, а b – вероятность искажения вектор-состояния нейрона. Оценим, при каких условиях m -й паттерн будет распознаваться сетью правильно. Амплитуды A_{il} (16) имеют вид

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

где $\xi_j = a_j(\mathbf{x}_j^{(m)}, \hat{b}_j \mathbf{x}_j^{(m)})$, $\eta_r \equiv \eta_j^{(\mu)}(l) = a_j(\mathbf{e}_l, \mathbf{x}_i^{(\mu)})(\mathbf{x}_j^{(\mu)}, \hat{b}_j \mathbf{x}_j^{(m)})$, $j = 1, \dots, N$, $j \neq i$, $\mu = 1, \dots, p$, $\mu \neq m$ и для простоты величины η проиндексированы обобщенным индексом $r = (j, \mu)$, который принимает $L = (N - 1)(p - 1)$ различных значений: $r = 1, \dots, L$; зависимость величин η от индекса l в дальнейшем никакой роли не играет.

Для *рандомизированного* набора паттернов $\{X^{(\mu)}\}_{\mu=1}^p$ случайные величины ξ_j и η_r независимы и имеют распределения

$$(17) \quad \xi_j = \begin{cases} +1 & \text{с вероятн. } (1 - b)(1 - a), \\ 0 & \text{с вероятн. } b, \\ -1 & \text{с вероятн. } (1 - b)a; \end{cases} \quad \eta_r = \begin{cases} +1 & \text{с вероятн. } 1/2q^2, \\ 0 & \text{с вероятн. } 1 - 1/q^2, \\ -1 & \text{с вероятн. } 1/2q^2. \end{cases}$$

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

$$\text{sgn}(A_{il_i^{(m)}}) = (\mathbf{e}_{l_i^{(m)}}, \mathbf{x}_i^{(m)}), \quad \sum_{j=1}^{N-1} \xi_j + (\mathbf{e}_{l_i^{(m)}}, \mathbf{x}_i^{(m)}) \sum_{r=1}^L \eta_r \geq \left| \sum_{r=1}^L \eta_r \right|.$$

В противном случае произойдет ошибка распознавания вектор-координаты $\mathbf{x}_i^{(m)}$. Поскольку случайная величина $(\mathbf{e}_{l_i^{(m)}}, \mathbf{x}_i^{(m)})\eta_r$ имеет то же распределение, что и η_r , вероятность ошибки распознавания можно представить в виде

$$(18) \quad \text{Pr}_i = \text{Pr} \left\{ \sum_{j=1}^{N-1} \xi_j + \sum_{r=1}^L \eta_r < 0 \right\}.$$

Для оценки вероятности (18) воспользуемся известной техникой Чебышева-Чернова [7,8] (см. Приложение). В результате получим верхнюю оценку для вероятности ошибки распознавания паттерна $X^{(m)}$:

$$(19) \quad \text{Pr}_{err} \leq N \exp \left(-\frac{N(1 - 2a)^2}{2p} q^2 (1 - b)^2 \right).$$

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

$$(20) \quad p_c = \frac{N(1 - 2a)^2}{2 \ln N} q^2 (1 - b)^2.$$

Оценку (20) можно рассматривать как асимптотически достижимую емкость памяти ПНС-II.

При $q = 1$ выражения (19),(20) превращаются в известные результаты для модели Хопфилда (в этом случае $b = 0$). С ростом q экспоненциально падает вероятность ошибки распознавания (19) – существенно растет помехоустойчивость сети. Одновременно пропорционально q^2 растет и емкость памяти (20). В отличие от модели Хопфилда оказывается возможным эффективное запоминание большего чем N числа паттернов p . На рисунке для различных значений q сплошной линией показана зависимость вероятности правильного распознавания $P_{rec} = 1 - \text{Pr}_{err}$ от величины частотного шума b (в процентах) при числе паттернов вдвое большем числа нейронов: $\alpha = p/N = 2$; искажения по знаку здесь отсутствуют ($a = 0$). Мы видим, что для $q = 20$ практически со 100%ной вероятностью будет правильно восстановлен любой из паттернов, зашумленный не более чем на 70%, а для $q = 30$ – любой паттерн, зашумленный не более чем на 85%. Машинные эксперименты подтверждают эти оценки.

4. Поттс-стекольная ассоциативная память

В прошлом уже рассматривались модели ассоциативной памяти с нейронами, которые могут находиться в большем чем два числе состояний – см., например, [5,10-15]. Все эти модели являются производными от известной модели Поттса магнетика, которая обобщает модель Изинга на случай спиновой переменной, принимающей не два значения ± 1 , а произвольное число q различных значений [16,17]. 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 [18]. А именно, характерное для физических рассматриваний короткодействующее взаимодействие между двумя соседними спинами заменялось межсвязями Хейббовского типа между векторами-нейронами. В результате возникающего в системе дальнегодействия 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 искаженных паттернов.

В *планарной* модели Поттса в качестве спиновых переменных используются q симметрично направленных векторов, лежащих в одной плоскости. Планарная модель была переформулирована на язык нейронных сетей в [10] (clock neural network). Емкость памяти этой сети меньше, чем у модели Хопфилда.

В *стандартной* модели Поттса спиновыми переменными являются q векторов,

служащих ребрами $(q - 1)$ -мерного симплекса. Эта модель послужила прообразом для двух вариантов Potts-glass neural network (PGNN). Вариант PGNN с *изотропным* гамильтонианом [12] оказывается полезным для инвариантного распознавания образов. Емкость памяти этой сети максимальна при $q = 2$, когда сеть переходит в модель Хопфилда. С ростом q емкость памяти изотропной PGNN падает. Вариант PGNN с *анизотропным* гамильтонианом (APGNN), предложенный в [11] и развитый в [13-15], оказывается наиболее близким к параметрической нейронной сети как по своим характеристикам, так и по идеологии. С помощью векторного формализма APGNN можно описать следующим образом.

Сеть составлена из N нейронов, каждый из которых может находиться в одном из q различных состояний. l -е состояние нейрона изображается q -мерным вектором \mathbf{d}_l , у которого l -я координата пропорциональна $q - 1$, а все остальные координаты пропорциональны -1 :

$$(21) \quad \mathbf{d}_l = \frac{1}{q} \begin{pmatrix} -1 \\ \vdots \\ q - 1 \\ \vdots \\ -1 \end{pmatrix}, \quad l = 1, \dots, q.$$

Состояние i -го нейрона описывается вектором $\mathbf{x}_i = \mathbf{d}_{l_i}$, $1 \leq l_i \leq q$. Состояние всей сети задается набором N таких q -мерных векторов $X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$, а исходные образы – совокупностью p паттернов

$$X^{(\mu)} = (\mathbf{x}_1^{(\mu)}, \mathbf{x}_2^{(\mu)}, \dots, \mathbf{x}_N^{(\mu)}), \quad \mathbf{x}_i^{(\mu)} = \mathbf{d}_{l_i^{(\mu)}}, \quad \mu = 1, \dots, p, \quad 1 \leq l_i^{(\mu)} \leq q.$$

Как и в ПНС, межсвязь между i -м и j -м нейронами задается $(q \times q)$ -матрицей \mathbf{T}_{ij} , которая в духе Хебовского правила (15) аккумулирует состояния i -го и j -го нейронов во всех p паттернах

$$\mathbf{T}_{ij} = (1 - \delta_{ij}) \sum_{\mu=1}^p \mathbf{B}_{l_i^{(\mu)} l_j^{(\mu)}}.$$

Элементарные матрицы $\mathbf{B}_{ll'}$ определяются соотношениями

$$(22) \quad \mathbf{B}_{ll'} = \mathbf{d}_l \mathbf{d}_{l'}^+ \implies \mathbf{B}_{ll'} \mathbf{d}_{l''} = (\delta_{l', l''} - 1/q) \mathbf{d}_l, \quad \forall l, l', l'' = 1, \dots, q.$$

Локальное поле \mathbf{h}_i , действующее на i -й нейрон, имеет вид

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

а динамика сети задается правилом: в следующий момент времени i -й нейрон переходит в состояние, максимизирующее $(\mathbf{d}_l, \mathbf{h}_i(t))$,

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

Известно, что при $q = 2$ APGNN эквивалентна стандартной модели Хопфилда [11]. Фазовое пространство APGNN-динамической системы содержит q^N точек (различных конфигураций). Для APGNN отсутствует понятие о конфигурации $-X$, обратной к X : в процессе эволюции системы такие конфигурации просто не возникнут. Соответственно, говорить о свойствах 1), 2) ПНС-II здесь не приходится (см. предыдущий раздел). Однако свойство 3) имеет место и для APGNN: если число паттернов $p \leq 2$, то (при любом q) единственными неподвижными точками сети будут сами эти паттерны.

Можно оценить емкость памяти APGNN при $N \gg 1$. Проводя вычисления, аналогичные тем, что делались в случае ПНС, для рандомизированного набора паттернов получим:

$$(23) \quad \text{Pr}_{err} = N \exp \left\{ -\frac{N}{2p} \frac{q(q-1)}{2} (1 - \bar{b})^2 \right\},$$

$$(24) \quad p_c = \frac{N}{2 \ln N} \frac{q(q-1)}{2} (1 - \bar{b})^2, \text{ где } \bar{b} = \frac{q}{q-1} b.$$

При $q = 2$ эти выражения переходят в известные оценки для модели Хопфилда. В то же время при $q \gg 1$ емкость памяти APGNN в два раза меньше емкости памяти ПНС-II – ср. (24) с (20) при нулевом частотном шуме $a = 0$. Эта двойка становится существенной, когда речь заходит о помехоустойчивости сети: для вероятности ошибки распознавания эта двойка попадает в показатель экспоненты (см. (19) и (23)), что ведет к заметному снижению помехоустойчивости APGNN по сравнению с ПНС-II. Это хорошо видно на рисунке, где для APGNN штриховой линией показана зависимость вероятности правильного распознавания P_{rec} от величины частотного шума b при тех же условиях, что и для ПНС-II (сплошная линия). Превосходство ПНС-II над APGNN вполне очевидно.

Сравнивая между собой ПНС-II и APGNN, отметим следующее обстоятельство. Эти модели ассоциативной памяти отличаются друг от друга наборами базисных векторов (8) и (21) и тем, как эти векторы преобразуются под действием элементарных матриц (14) и (22) соответственно. Однако различия эти, на наш взгляд, не столь уж велики. В самом деле, в ПНС-II элементарная матрица $\mathbf{A}_{ll'}$ (14), действуя на l'' -й базисный вектор, переводит его в l -й базисный вектор *только тогда*,

когда совпадают индексы l' и l'' – во всех остальных случаях вектор $\mathbf{e}_{l''}$ переводится в ноль. Здесь работает принцип несоизмеримости частот, выдвинутый в [1,2] как средство подавления шумов: из массы возможных событий (комбинаций индексов l' , l'') отфильтровывается только одно (совпадение индексов), а все остальные игнорируются. Принцип, положенный в основу APGNN, выглядит полной противоположностью: элементарная матрица $\mathbf{B}_{ll'}$ (22) *всегда* переводит l'' -й базисный вектор в l -й. Однако противоположность эта кажущаяся. Потому что, когда индексы l' и l'' совпадают, l -й вектор снабжается большой положительной амплитудой порядка $q - 1$, а во всех остальных случаях – маленькой отрицательной амплитудой порядка -1 . Иначе говоря, в несколько иной форме здесь реализован тот же самый принцип несоизмеримости частот. Представляется, что близость характеристик обеих моделей обусловлена именно этим обстоятельством.

5. Заключение

В случае ПНС нейрон может находиться в одном из $2q$ различных состояний, поскольку изображающие векторы \mathbf{x}_i характеризуются двумя параметрами: положением в пространстве и знаком ориентации. Эти две характеристики полностью независимы, что согласуется с лежащей в основе ПНС оптической моделью: у квази-монохроматических импульсов, которыми обмениваются параметрические нейроны, сбой по фазе (инверсия знака) может быть вызван одними причинами, а сбой по частоте (неправильная ориентация вектора) – совсем другими.

В случае поттсовских нейросетей у векторов-нейронов \mathbf{x}_i знак отсутствует и нейрон может находиться в одном из q различных состояний. Это ограничение унаследовано от модели Поттса магнетика и, вообще говоря, носит случайный характер. В принципе, можно рассмотреть вариант APGNN-модели, допускающий перед векторами \mathbf{x}_i знаки \pm , скорректировав при этом и динамическое правило в духе соотношения (11). По емкости памяти и помехоустойчивости такой вариант APGNN, скорее всего, сравняется с ПНС-II. С точки зрения нейронных сетей это, однако, навряд ли имеет какой-то смысл.

Во-первых, в вычислительном отношении ПНС-схема предпочтительнее APGNN-схемы, поскольку в первом случае и базисные векторы, и элементарные матрицы в высокой степени разрежены. Во-вторых, если говорить о реализации модели в виде устройства, преимущество должно быть отдано ПНС. В сущности, ПНС является адаптацией к языку нейронных сетей тех процессов, которые реально происходят при передаче сигналов по оптическим каналам. И вопрос о реализации ПНС в виде оптического устройства сводится к вопросу о том, позволяет ли современная

технология воспроизводить эти процессы в "железе" и управлять ими. В значительной мере, это вопрос технологии. Поттсовские нейросети тоже ориентированы на обработку цветных изображений, однако их реализация в виде устройства требует, в первую очередь, выяснения того, на каких физических принципах такое устройство будет работать. Поттсовские нейросети являются адаптацией к языку нейронных сетей статфизических аспектов теории конденсированного состояния. И приходится лишь удивляться тому, насколько близким оказывается математический формализм APGNN к естественному языку проблематики, связанной с передачей и обработкой оптических сигналов.

ПРИЛОЖЕНИЕ

Для асимптотической оценки емкости памяти ПНС-II оценим при $N \gg 1$ ошибку распознавания Pr_i (18) i -й вектор-координаты $\mathbf{x}_i^{(m)}$ зашумленного m -го паттерна:

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

Тогда, используя известную технику экспоненциальных оценок Чебышевского типа, для любого положительного $z > 0$ получаем:

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

Черта сверху означает усреднение по всем возможным реализациям, а последнее равенство следует из независимости случайных величин ξ_j и η_r .

С учетом (17) легко получить выражения для средних

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

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

$$f_1(x) = a(1-b)x + b + \frac{(1-a)(1-b)}{x}, \quad f_2(x) = \frac{1}{2q^2} \left(x + \frac{1}{x} \right) + 1 - \frac{1}{q^2},$$

получаем, что при любом положительном x для Pr_i справедливо:

$$(25) \quad \text{Pr}_i \leq \left(f_1(x) f_2^{p-1}(x) \right)^{N-1}.$$

Для получения наинизшей оценки вероятности Pr_i необходимо отыскать значение переменной x , минимизирующее правую часть (25). Это приводит к уравнению

$$(p-1)(x^2-1) + \frac{a(1-b)x^2 - (1-a)(1-b)}{a(1-b)x^2 + bx + (1-a)(1-b)}(x^2 + 2(q^2-1)x + 1) = 0.$$

В случае $p \gg 1$ интересующий нас корень этого уравнения, с точностью до членов более высокого порядка малости по $1/p$, равен $x_1 = 1 + q^2(1 - 2a)(1 - b)/(p - 1)$. Подставляя это значение x в правую часть (25), окончательно получаем оценку вероятности ошибки распознавания вектор-координаты $\mathbf{x}_i^{(m)}$:

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

Что дает оценку (19) для вероятности ошибки распознавания паттерна $X^{(m)}$.

СПИСОК ЛИТЕРАТУРЫ

1. *Крыжановский Б.В., Микаэлян А.Л.* О распознающей способности нейросети на нейронах с параметрическим преобразованием частот // ДАН (мат.-физ.). 2002. Т.65(2). С.286-288.
2. *A.Fonarev, B.V.Kryzhanovsky, Mikaelyan A.L.* Parametric dynamic neural network recognition power // Optical Memory & Neural Networks. 2001. V.10(4). P.31-48.
3. *Hertz J., Krogh A., Palmer R.* Introduction to the Theory of Neural Computation. Massachusetts: Addison-Wesley, 1991.
4. *Ikeda N., Watta P., Artiklar M. et al.* A two-level Hamming network for high-performance associative memory // Neural Networks. 2001. V.14, P.1189-1200.
5. *Noest J.* Discrete-state phasor neural networks // Phys. Rev. A. 1988. V.38. P.2196-2199.
6. *Bloembergen N.* Nonlinear optics. NY: Benjamin, 1965.
7. *Chernov N.* A mesure of asymptotic efficiency for tests of hypothesis based on the sum of observations // Ann. Math. Statistics. 1952. V.23. P.493-507.
8. *Kryzhanovsky B.V., Mikaelyan A.L., Koshelev V.N. et al.* On recognition error bound for associative Hopfield memory // Optical Memory & Neural Networks. 2000. V.9(4). P.267-276.
9. *Kinzel W.* Spin glasses and memory // *Physica Scripta*. 1987. V.3. P.398-401.
10. *Cook J.* The mean-field theory of a Q-state neural network model // J. Phys. A. 1989. V.22. P.2000-2012.
11. *Kanter I.* Potts-glass models of neural networks // Phys. Rev. A. 1988. V.37. P.2739-2742.
12. *Vogt H., Zippelius A.* Invariant recognition in Potts glass neural networks // J. Phys. A. 1992. V.25. P.2209-2226.
13. *Bolle D., Dupont P., van Mourik J.* Stability properties of Potts neural networks with biased patterns and low loading // J. Phys. A. 1991. V.24. P.1065-1081.
14. *Bolle D., Dupont P., Huyghebaert J.* Thermodynamics properties of the q -state Potts-glass neural network // Phys. Rev. A. 1992. V.45. P.4194-4197.
15. *Bolle D., Jongen G., Shim G.M.* Q-Ising neural network dynamics: a comparative review of various architectures // Cond-mat/9907390, 1999.
16. *Wu F.Y.* The Potts model // Review of Modern Physics. 1982. V.54(1). P.235-268.
17. *Бэкстер Р.* Точно решаемые модели статистической физики. М.: Мир. 1985.
18. *Amit D., Gutfreund H., Sompolinsky H.* Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks // Phys. Rev. Lett. 1985. V.55. P.1530-1533.

Подписи к рисункам

Вероятность правильного распознавания $P_{rec} = 1 - \text{Pr}_{err}$ как функция частотного шума b (в процентах) для различных значений $q = 5, 10, 20, 30$ при числе паттернов вдвое большем числа нейронов $\alpha = p/N = 2$: сплошная линия – параметрическая нейронная сеть (ПНС-II), штриховая линия – Поттс-стекольная ассоциативная память (APGNN).