Векторно-нейронные модели ассоциативной памяти

Б.В.Крыжановский, Л.Б.Литинский

Институт оптико-нейронных технологий РАН, Москва

Аннотация

Рассматриваются две модели ассоциативной памяти Хопфилдова типа с q-нарными нейронами: *Поттс-стекольная нейросеть* и *параметрическая нейронная сеть*. Нейроны могут находиться в $q \geq 2$ различных состояниях. Модели имеют рекордные показатели по объему памяти и помехоустойчивости, и значительно превосходят модель Хопфилда. Развивается векторный формализм, позволяющий описать обе модели в едином подходе. Объясняются механизмы, приводящие к большому объему нейросетевой памяти и высокой помехоустойчивости обеих нейросетей.

1 Введение

Число образов, которые способна хранить модель Хопфилда (HM) ассоциативной памяти сравнительно невелико. Если воспользоваться статфизической оценкой емкости памяти [1],[2], то $p_{\mathrm{HM}} \sim 0.14 \cdot N$, где N – число бинарных нейронов. Иначе говоря, число образов, которые сеть способна эффективно запомнить, составляет порядка 15% от размера сети. В начале 90-х несколькими авторами ([3]-[10]) были предложены q-нарные модели ассоциативной памяти, в которых число различных состояний нейронов q могло превосходить 2. Все эти модели связаны с моделью Поттса магнетика, которая обобщает модель Изинга на случай спиновой переменной, имеющей больше двух различных состояний [11],[12]. Во всех этих работах авторы использовали один и тот же прием, с помощью которого модель Изинга была в свое время связана с моделью Хопфилда (см., например, [2]). А именно, коротко-действующее взаимодействие между двумя соседними спинами заменялось межсвязями Хеббовского типа между всеми нейронами. В результате возникающего в системе дальнодействия оказывалось возможным применить среднеполевое приближение и вычислить статистическую сумму. Различные области фазовой диаграммы интерперетировались затем в терминах способности или неспособности системы распознавать зашумленные образы.

Для всех этих моделей, кроме одной, емкость памяти оказалась меньше, чем у модели Хопфилда. Исключение составила *Поттс-стекольная нейросеть* (*Potts-glass neural network*) [3]. Численное решение системы уравнений, получающихся в результате применения к Поттс-стекольной модели статфизических методов, дает следующую оценку емкости памяти:

$$p_{\mathrm{PG}} \sim \frac{q(q-1)}{2} \cdot p_{\mathrm{HM}}.$$

Если q-нарные сети использовать для обработки цветных изображений, число q можно понимать как число различных цветов, в которые может быть окрашен элементарный пискел экрана. Следовательно, при $q \sim 10$ емкость памяти Поттс-стекольной нейросети превосходит тот же показатель для модели Хопфилда в 50 раз. Если же $q = 256$, что является стандартом при компьютерной обработке изображений, то $p_{\mathrm{PG}} \sim 10^4 \cdot p_{\mathrm{HM}}$.

Это очень хороший результат. Однако причины, по которым Поттс-стекольная модель обладает таким большим объемом памяти (а другие q-нарные модели уступают даже модели Хопфилда), долгое время оставались неясны. Статфизический подход не дает ответа на этот вопрос.

С другой стороны, в [13],[14] мы разработали модель ассоциативной памяти, ориентированную на реализацию в виде оптического устройства. Сигналы в сети распространяются по оптическим межсвязям в виде квазимонохроматических импульсов на q различных частотах. Иначе говоря, сеть способна хранить и обрабатывать информацию, закодированную в виде частотно-фазовой модуляции.

Имеется целый ряд аргументов в пользу такого подхода. 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, фактически, является аналогом уплотнения канала, и позволяет в q^2 раз уменьшить число межсвязей. Заметим, что межсвязи и так занимают до 98% площади нейрокристалла.

Сердцевину нашей модели составляют процессы параметрического четырех-волнового смешения, хорошо известные в нелинейной оптике [15]. Однако, для того, чтобы система работала как память, необходимо дополнительное условие, способствующее подавлению в системе внутрених шумов. Это дополнительное условие – *принцип несоизмеримости частот*, сформулированный на языке нелинейной оптики в [13],[14]. Анализ соотношения сигнал/шум в сочетании со статистической техникой Чебышева-Чернова [16],[17] показал, что для рандомизированного набора образов емкость памяти нашей модели в q^2 раз превосходит емкость памяти модели Хопфилда. Мы назвали нашу модель *параметрический нейронной сетью.* И решили более внимательно исследовать ее связи с Поттс-стекольной нейросетью.

Был разработан *векторный формализм* – универсальный язык описания параметрической нейросети, оторванный от оптической постановки задачи [18]-[20]. Оказалось, что на этом языке возможно и очень ясное описание Поттс-стекольной модели (которая первоначально была сформулирована совершенно в других терминах). В рамках векторного формализма легко прослеживаются связи между обеими моделями и становятся понятны механизмы, обеспечивающие их замечательные распознающие свойства. Причина этого кроется в локальной архитектуре обеих сетей, которая способствует подавлению в системе внутренних шумов. В других q-нарных моделях такого подавления не происходит.

В настоящей работе мы изложим Поттс-стекольную модель на языке векторного формализма и объясним механизмы, обеспечивающие для этой сети большой объем памяти. Затем мы опишем параметрическую нейросеть – и в оптической формулировке, и на языке векторного формализма. Кроме того, для параметрической нейросети мы рассмотрим несколько различных архитектур.

Замечание. Наш векторный формализм во многом идентичен векторно-нейронному подходу к ассоциативной памяти, который за несколько лет до нас был предложен в [21]. Мы познакомились с этой работой уже после того, как разработали векторный формализм. Авторы [21] неудачно выбрали динамическое правило, однако, по-видимому, именно они впервые сформулировали весьма плодотворные идеи относительно представления матриц межсязей в виде тензорного произведения векторов-нейронов.

2 Поттс-стекольная нейросеть

Опишем Поттс-стекольную модель на языке векторного формализма, чтобы в дальнейшем сравнить ее с параметрической нейросетью.

2.1 Векторный формализм

Сеть состоит из N нейронов, каждый из которых может находиться в одном из q состояний. Для описания различных состояний нейронов используется набор специальных векторов – *Поттсовские векторы*. А именно, l-е состояние нейрона изображается q-мерным вектором-столбцом $\mathbf{d}_l \in \mathrm{R}^{\mathrm{q}}$, у которого l-я координата пропорциональна $q-1$, а остальные координаты пропорциональны -1:

$$\mathbf{d}_l = \frac{1}{q}\begin{pmatrix} -1 \\ \vdots \\ q-1 \\ \vdots \\ -1 \end{pmatrix},\ l = 1, \ldots, q.$$

Состояние i-го нейрона описывается вектором

$$\mathbf{x}_i = \mathbf{d}_{l_i}, 1 \le l_i \le q.$$

Состояние сети как целого определяется набором из N q-мерных векторов-столбцов $\mathbf{x}_i$: $X = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, а p наперед заданных образов – это p подобных наборов $X^{(\mu)}$:

$$\begin{aligned} &X^{(\mu)} = (\mathbf{x}_1^{(\mu)}, \ldots, \mathbf{x}_N^{(\mu)}),\ \mathbf{x}_i^{(\mu)} = \mathbf{d}_{l_i^{(\mu)}}, \\ &1 \le l_i^{(\mu)} \le q,\ \mu = 1, 2, \ldots, p. \end{aligned}$$

Поскольку нейроны являются векторами, локальное поле $\mathbf{h}_i$, действующее на i-й нейрон, тоже будет вектором:

$$\mathbf{h}_i = \frac{1}{N}\sum_{j=1}^{N} \mathbf{T}_{ij}\mathbf{x}_j.$$

Межсвязь между i-м и j-м нейронами описывается $(q \times q)$-матрицей $\mathbf{T}_{ij}$. По аналогии с моделью Хопфилда эта матрица выбирается в обобщенном Хеббовском виде:

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

где $\mathbf{x}^+$ – обозначение для q-мерной вектор-строки, а δ_{ij} – символ Кронекера. Матричные элементы матриц межсвязей могут быть вычислены по формуле:

$$T_{ij}^{(kl)} = (1 - \delta_{ij}) \sum_{\mu=1}^{p} (\mathbf{e}_k \mathbf{x}_i^{(\mu)})(\mathbf{x}_j^{(\mu)} \mathbf{e}_l) \quad k, l = 1, \ldots, q.$$

Матрица $\mathbf{T}_{ij}$, действуя на вектор $\mathbf{x}_j \in \mathrm{R}^{\mathrm{q}}$, переводит его в линейную комбинацию векторов-столбцов $\mathbf{d}_l$. После суммирования по всем j получим локальное поле $\mathbf{h}_i$ как линейную комбинацию векторов $\mathbf{d}_l$ с суммарными коэффициентами

$$\mathbf{h}_i = \sum_{l=1}^{q} A_l^{(i)} \mathbf{d}_l.$$

Пусть k – индекс, отвечающий максимальному коэффициенту в этом разложении: $A_k^{(i)} > A_l^{(i)} \; \forall \; l$. Тогда, по определению, в следующий момент временени $t+1$, i-й нейрон ориентируется вдоль направления, максимально близкого к направлению локального поля $\mathbf{h_i}$:

$$\mathbf{x}_i(t+1) = \mathbf{d}_k. \tag{2}$$

Эволюция системы состоит в последовательной переориентации векторов-нейронов согласно правилу (2). Условимся, что если несколько коэффициентов $A_l^{(i)}$ являются максимальными одновременно, и нейрон находится в одном из этих, неулучшаемых состояний, его состояние не меняется. Тогда нетрудно показать, что в процессе эволюции сети ее *энергия* $H(t) \sim -\sum_{i=1}^{N} \mathbf{x}_i^+(t) \mathbf{h}_i(t)$ монотонно убывает. В конце концов система свалится в локальный минимум по энергии. В этом состоянии все векторы-нейроны $\mathbf{x}_i$ будут ориентированы неулучшаемым образом и эволюция системы прекратится. Такие состояния являются неподвижными точками системы. Необходимым и достаточным условием того, чтобы состояние X было неподвижной точкой, является выполнение нервенств:

$$(\mathbf{x}_i \mathbf{h}_i) \geq (\mathbf{d}_l \mathbf{h}_i), \quad \forall \; l = 1, \ldots, q; \; \forall \; i = 1, \ldots, N. \tag{3}$$

Когда $q = 2$, эта схема превращается в стандартную модель Хопфилда.

2.2 Емкость памяти Поттс-стекольной модели

Будем считать набор образов $\{X^{(\mu)}\}_1^p$ рандомизированным. Пусть на вход сети подается искаженный m-й образ

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

Независимые случайные операторы $\hat{b}_j$ задают мультпликативный шум: с вероятностью b оператор $\hat{b}_j$ меняет состояние вектора-нейрона $\mathbf{x}_j^{(m)}$ на какое-нибудь другое, а с вероятностью $1-b$ оставляет вектор-нейрон неизменным. Таким образом, степень искажения образа характеризуется вероятностью b искажения каждого нейрона. Оценим вероятность того, что m-й образ не будет восстановлен сетью.

Элементарные вычисления показывают, что вероятность выполнения неравенства $(\mathbf{x}_i^{(m)}\mathbf{h}_i) < (\mathbf{d}_l\mathbf{h}_i)$ при $\mathbf{d}_l \neq \mathbf{x}_i^{(m)}$ можно записать в виде

$$P_{err}^{(i,l)} = \mathrm{Prob}\left\{\xi < \eta\right\} = \mathrm{Prob}\left\{\frac{1}{N}\sum_{j\neq i}^{N}\xi_j < \frac{1}{N}\sum_{j\neq i}^{N}\sum_{\mu\neq m}^{p}\eta_j^{(\mu)}\right\}, \qquad (4)$$

где $\xi_j = (\mathbf{x}_j^{(m)}\hat{b}_j\mathbf{x}_j^{(m)})$, а $\eta_j^{(\mu)} = (\mathbf{d}_l - \mathbf{x}_i^{(m)}, \mathbf{x}_i^{(\mu)})(\mathbf{x}_j^{(\mu)}\hat{b}_j\mathbf{x}_j^{(m)})$.

Величина ξ является *полезным сигналом*, связанным с влиянием на i-й нейрон именно m-го образа (у нас этот сигнал искажен мультипликативным шумом). Напротив, величина η описывает *внутренний шум* в системе, связанный с мешающим воздействием всех остальных образов. Случайные величины ξ_j независимы и одинаково распределены. Парциальные шумовые компоненты $\eta_j^{(\mu)}$ также независимы и одинаково распределены. Нетрудно получить распределения для ξ_j и $\eta_j^{(\mu)}$:

$$\xi_j = \left\{\begin{array}{rl} (q-1)/q, & 1-b \\ -1/q, & b \end{array}\right., \quad \eta_j^{(\mu)} = \left\{\begin{array}{rl} (q-1)/q, & 1/q^2 \\ 1/q, & (q-1)/q^2 \\ 0, & (q-2)/q \\ -1/q, & (q-1)/q^2 \\ -(q-1)/q, & 1/q^2 \end{array}\right. . \qquad (5)$$

Обратим внимание на то, что при $q >> 1$ шумовая компонента $\eta_j^{(\mu)}$ сосредоточена, в основном, в нуле: $\mathrm{Prob}\left\{\eta_j^{(\mu)} = 0\right\} = (q-2)/q \sim 1$.

Суммарные случайные величины ξ и η асимптотически нормальны с параметрами

$$\begin{array}{ll} E(\xi) = \frac{q-1}{q} - b, & E(\eta) = 0, \\ D(\xi) \to 0; & D(\eta) = \frac{2(q-1)}{q^3}\cdot\alpha, \end{array} \qquad (6)$$

где, как обычно, *параметр загрузки* $\alpha = \frac{p}{N}$. Теперь можно вычислить вероятность $P_{err}^{(i,l)}$ (4) как интеграл под тем хвостом нормально распределенной величины η, где $\eta > E(\xi)$:

$$P_{err}^{(i,l)} = \frac{1 - erf(c)}{2}, \ c = \frac{1 - \bar{b}}{\sqrt{2\alpha}} \sqrt{\frac{q(q-1)}{2}}, \ \bar{b} = \frac{q}{q-1} b;$$

здесь $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du$. Важнее другое – данное место является ключевым для понимания того, почему Поттс-стекольная нейросеть существенно превосходит модель Хопфилда.

В самом деле, те же самые рассуждения можно проделать и для модели Хопфилда – это сделано, например, в [2]. Опять мы получим разделение на полезный сигнал ξ и внутренний шум η, и неравенство (4) для вероятности ошибки распознавания. Опять эти случайные величины будут асимптотически нормальны. Распределения для парциальных случайных компонент ξ_j и $\eta_j^{(\mu)}$ можно получить непосредственно из (5), положив $q = 2$ – ведь в этом случае Поттс-стекольная нейросеть переходит в модель Хопфилда. А средние и дисперсии для ξ и η аналогично получаются из (6). В результате имеем:

$$\begin{cases} \xi_j = \begin{cases} \ \ 1/2, & 1-b \\ -1/2, & b \end{cases}, & \eta_j^{(\mu)} = \begin{cases} \ \ 1/2, & 1/2 \\ -1/2, & 1/2 \end{cases}, \\ \\ E(\xi) = \frac{1}{2} - b, & E(\eta) = 0, \\ D(\xi) \to 0; & D(\eta) = \frac{\alpha}{4}. \end{cases} \tag{7}$$

Сопоставляя (7) с (5) и (6) можно видеть, что дисперсия внутреннего шума для Поттс-стекольной нейросети во много раз меньше аналогичной характеристики для модели Хопфилда:

$$D_{\mathrm{PG}}(\eta)/D_{\mathrm{HM}}(\eta) = \frac{8(q-1)}{q^3} << 1, \text{ когда } q >> 1.$$

Уже при $q \sim 10$ дисперсия внутреннего шума для Поттс-стекольной нейросети на порядок меньше дисперсии внутреннего шума для модели Хопфилда. А при $q \sim 10^2$ дисперсия уменьшается на 4 порядка! Этим и определяется превосходство Поттс-стекольной нейросети над моделью Хопфилда.

Ответ на вопрос: какими механизмами обеспечивается сжатие внутреннего шума в Поттс-стекольной нейросети? - мы отложим до следующего раздела.

Переходя от анализа ситуации с одной вектор-координатой $\mathbf{x}_i^{(m)}$ к анализу ситуации с целым образом, и используя при этом стандартные приближения ([19],[20]), можно получить верхнюю оценку для вероятности неправильного распознавания образа $X^{(m)}$,

$$\Pr_{err} < \sqrt{Np}\exp\left(-\frac{N}{2p}\frac{q(q-1)}{2}(1-\bar{b})^2\right), \quad \bar{b} = \frac{q}{q-1}b. \tag{8}$$

Выражение в правой части стремится к нулю, когда число паттренов p растет медленнее, чем

$$p_{\rm PG} = \frac{N}{2\ln N}\frac{q(q-1)}{2}(1-\bar{b})^2. \tag{9}$$

Величину $p_{\rm PG}$ можно считать асимптотически достижимой емкостью памяти Поттс-стекольной модели.

Когда $q = 2$ эти выражения дают известные оценки для модели Хопфилда. При $q > 2$ емкость памяти Поттс-стекольной модели в $q(q-1)/2$ раз превосходит емкость памяти модели Хопфилда. В работе [3] этот множитель был получен подгонкой результатов численного расчета. Мы получили этот результат строго.

3 Параметрическая нейронная сеть

Здесь мы опишем нашу модель ассоциативной памяти – как на языке нелинейной оптики, так и в рамках векторного формализма. И изложим полученные для нее результаты.

3.1 Нелинейно-оптическая формулировка

Положим, что информативные сигналы в сети распространяются по оптическим межсвязям в виде квазимонохроматических импульсов на q различных частотах:

$$\{\omega_l\}_1^q \equiv \{\omega_1, \omega_2, ..., \omega_q\}. \tag{10}$$

За основу сети принимается *параметрический нейрон* – обладающий кубической нелинейностью элемент, способный к преобразованию и генерации частот в процессах параметрического четырех-волнового смешения: $\omega_i - \omega_j + \omega_k \to \omega_r$ (то же, что и четырехфотонное взаимодействие [15]). Нейрон будем представлять себе как сложную структуру, состоящую из: 1) сумматора входных сигналов; 2) набора q идеальных частотных фильтров с собственными частотами (10); 3) блока сравнения отфильтрованных сигналов по амплитуде и 4) q генераторов квазимонохроматических импульсов на частотах из набора (10).

Состояние нейрона в каждый момент времени характеризуется частотой ω и фазой ψ. Условимся, что фазы могут принимать только одно из двух значений: 0 или π. Тогда у квазимонохроматических импульсов появятся амплитуды ± 1:

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

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 $K = (\kappa_1, \kappa_2, \ldots, \kappa_N)$. Введем в рассмотрение p наперед заданных исходных образов – p фиксированных цветных изображений, информация о которых хранится в памяти сети:

$$\begin{aligned} &K^{(\mu)} = (\kappa_1^{(\mu)}, \ldots, \kappa_N^{(\mu)}),\ \text{где}\ \kappa_i^{(\mu)} = \pm \exp(\imath \omega_{l_i^{(\mu)}} t), \\ &\mu = 1, \ldots, p;\ i = 1, \ldots, N;\ 1 \le l_i^{(\mu)} \le q. \end{aligned} \tag{11}$$

Память сети локализована в межсвязях $T_{ij},\ i, j = 1, \ldots, N$, которые аккумулируют информацию о состояниях i-го и j-го нейронов во всех p образах. Будем полагать, что межсвязи организованы по Хеббовскому правилу:

$$T_{ij} = (1 - \delta_{ij}) \sum_{\mu=1}^{p} \kappa_i^{(\mu)} \kappa_j^{(\mu)*}, \quad i, j = 1, \ldots, N. \tag{12}$$

Схематически, принцип работы сети состоит в следующем. 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 аккумулированными межсвязью состояниями i-го и j-го нейронов во всех образах: $\omega_{l_i^{(\mu)}} - \omega_{l_j^{(\mu)}} + \omega_{l_j} \to \{\omega_l\}_1^q$.

Амплитуды сигналов при этом перемножаются. Просуммировав результаты этих парциальных преобразований по всем μ, получим целый пакет квазимонохроматических импульсов, вообще говоря – на всех частотах из набора (10), с определяющимися межсвязью амплитудами. Этот пакет – результат преобразования (ij)-й межсвязью импульса ω_{l_j} – и приходит к i-му нейрону.

Пакеты 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*).

Вообще говоря, в результате взаимодействия трех квазимонохроматических импульсов, всегда возникает четвертый импульс. Его частота определяется только законами сохранения энергии и импульса. Однако, для того, чтобы описанная схема работала как память, ее необходимо дополнить важным условием. Речь идет о предложенном в [13], [14] *принципе несоизмеримости частот* $\{\omega_l\}_1^q$: *никакая комбинация частот вида* $\omega_l - \omega_{l'} + \omega_{l''}$ *не может принадлежать набору частот (10), 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 $\omega_{l_j^{(\mu)}}$, по межсвязи не проходит, если все три частоты различны – в этом случае сигнал ω_{l_j} межсвязью фильтруется.

Описанная схема получила название *параметрической нейронной сети* (ПНС). Прежде, чем переходить к изложению результатов, сделаем одно замечание. Вообще говоря, возможны различные варианты параметрического четырехволнового смешения, удовлетворяющие принципу несоизмеримости частот. Например, в [13], [14] изучались процессы параметрического четырехволнового смешения вида

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

Соответствующую сеть назовем ПНС-1. Однако лучшие результаты получаются для параметрического четырехволнового смешения

$$\omega_l - \omega_{l'} + \omega_{l''} = \begin{cases} \omega_l, & \text{когда } l' = l''; \\ 0, & \text{в остальных случаях.} \end{cases} \tag{13}$$

Эту сеть будем называть ПНС-2. В настоящей работе излагаются результаты, полученные для ПНС-2 и других архитектур на ее основе.

Остаток статьи устроен следующим образом. В следующих пунктах ПНС-2 описывается на языке векторного формализма и для нее излагаются полученные результаты. В следующем разделе рассказывается о других нейро-архитектурах, выполненных на основе ПНС-2. Несколько общих замечаний вынесены в Заключение.

3.2 ПНС-2 на языке векторного формализма

Для описания q различных состояний нейронов (10) будем использовать набор векторов-ортов $\mathbf{e}_l$ пространства R^{q}, $\mathrm{q} \geq 1$:

$$\mathbf{e}_l = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ l = 1, \ldots, q.$$

Состояние i-го нейрона описывается вектором $\mathbf{x}_i$,

$$\mathbf{x}_i = x_i \mathbf{e}_{l_i}, \ x_i = \pm 1, \ \mathbf{e}_{l_i} \in \mathrm{R}^{\mathrm{q}}, \begin{cases} 1 \leq l_i \leq q; \\ i = 1, \ldots, N. \end{cases} \tag{14}$$

Множитель x_i моделирует наличие у сигналов фазы.

Состояние сети как целого задается набором N q-мерных векторов $\mathbf{x}_i$: $X = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, а p образов – это p наперед заданных подобных наборов:

$$X^{(\mu)} = (\mathbf{x}_1^{(\mu)}, \mathbf{x}_2^{(\mu)}, \ldots, \mathbf{x}_N^{(\mu)}), \ \mathbf{x}_i^{(\mu)} = x_i^{(\mu)} \mathbf{e}_{l_i^{(\mu)}},$$
$$x_i^{(\mu)} = \pm 1, \ 1 \leq l_i^{(\mu)} \leq q, \ \mu = 1, \ldots, p.$$

Локальное поле на i-м нейроне имеет тот же вид, что и в случае Поттс-стекольной модели:

$$\mathbf{h}_i = \frac{1}{N} \sum_{j=1}^{N} \mathbf{T}_{ij} \mathbf{x}_j, \tag{15}$$

где $(q \times q)$-матрица $\mathbf{T}_{ij}$ определяет межсвязь между i-м и j-м нейронами. Матрица $\mathbf{T}_{ij}$ действует на вектор $\mathbf{x}_j \in \mathrm{R}^{\mathrm{q}}$, превращая его в линейную комбинацию ортов $\mathbf{e}_l$. Эта комбинация – аналог пакета квазимонохроматических импульсов, приходящих к i-му нейрону от j-го после транформации в (ij)-й межсвязи (см. предыдущий пункт). Чтобы удовлетворить условиям (12) и (13), матрицы $\mathbf{T}_{ij}$ должны быть выбраны в виде

$$\mathbf{T}_{ij} = (1 - \delta_{ij}) \sum_{\mu=1}^{p} \mathbf{x}_i^{(\mu)} \mathbf{x}_j^{(\mu)^+}, \quad i, j = 1, \ldots, N. \tag{16}$$

Заметим, что конструкция этого выражения повторяет (1).

Формулируя на векторном языке правило функционирования ПНС, придем к тому же динамическому правилу, что и раньше: в момент времени $t+1$ i-й нейрон ориентируется в направлении, ближайшем к направлению локального поля $\mathbf{h}_i$ в момент времени t. Однако формулы теперь будут отличаться от (2). Действительно, с помощью (16) перепишем выражение для локального поля в более удобном для анализа виде:

$$\mathbf{h}_i(t) = \sum_{l=1}^{q} A_l^{(i)} \mathbf{e}_l, \text{ где } A_l^{(i)} \sim \sum_{j(\neq i)}^{N} \sum_{\mu=1}^{p} (\mathbf{e}_l \mathbf{x}_i^{(\mu)})(\mathbf{x}_j^{(\mu)} \mathbf{x}_j(t)). \tag{17}$$

Пусть $A_k^{(i)}$ – амплитуда, наибольшая *по модулю* среди всех амплитуд в разложении (17):

$$\mid A_k^{(i)} \mid = \max_{1 \le l \le q} \mid A_l^{(i)} \mid .$$

Тогда, согласно нашему определению,

$$\mathbf{x}_i(t+1) = \mathrm{sgn}(A_k^{(i)}) \mathbf{e}_k. \tag{18}$$

Эволюция системы состоит в последовательной переориентации векторов-нейронов по правилу (18). Необходимым и достаточным условием того, чтобы конфигурация X была неподвижной точкой, является выполнение неравенств

$$(\mathbf{x}_i \mathbf{h}_i) \ge \mid (\mathbf{e}_l \mathbf{h}_i) \mid, \quad \forall\ l = 1, \ldots, q;\ \forall\ i = 1, \ldots, N,$$

(ср. с выражениями (3)).

3.3 Емкость памяти ПНС-2

Все рассуждения здесь проводятся совершенно аналогично тому, как это делалось для Поттс-стекольной нейросети. Отличия связаны только с тем, что теперь состояния нейронов задаются как векторными, так и скалярными величинами. Искаженный m-й образ имеет вид:

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

Независимые случайные величины $\{a_j\}_1^N$ определяют *шум по фазе*: с вероятностью a случайная величина a_j принимает значение -1, а с вероятностью $1-a$ – значение $+1$. Независимые случайные операторы $\hat{b}_j$ определяют *шум по частоте*: с вероятностью b оператор $\hat{b}_j$ переориентирует вектор-нейрон $\mathbf{x}_j^{(m)}$ в каком-нибудь другом направлении, а с вероятностью $1-b$ оставляет вектор-нейрон неизменным. Оценим вероятность того, что m-й образ не будет восстановлен сетью.

Амплитуды $A_l^{(i)}$ (17) имеют вид:

$$A_l^{(i)} \sim \begin{cases} x_i^{(m)}\sum_{j\neq i}^N \xi_j \;+\; \sum_{j\neq i}^N\sum_{\mu\neq m}^p \eta_j^{(\mu)}, & \text{для } l = l_i^{(m)}; \\ \sum_{j\neq i}^N\sum_{\mu\neq m}^p \eta_j^{(\mu)}, & \text{для } l \neq l_i^{(m)}, \end{cases}$$

где $\xi_j = a_j(\mathbf{x}_j^{(m)}\hat{b}_j\mathbf{x}_j^{(m)})$, $\eta_j^{(\mu)} = a_j(\mathbf{e}_l\mathbf{x}_i^{(\mu)})(\mathbf{x}_j^{(\mu)}\hat{b}_j\mathbf{x}_j^{(m)})$, $j = 1,\ldots,N$, $\mu = 1,\ldots,p$. Для рандомизированного набора образов, ξ_j и $\eta_j^{(\mu)}$ являются независимыми случайными величинами с распределениями

$$\xi_j = \begin{cases} +1, & (1-b)(1-a) \\ 0, & b \\ -1 & (1-b)a \end{cases}, \quad \eta_j^{(\mu)} = \begin{cases} +1, & 1/2q^2 \\ 0, & 1-1/q^2 \\ -1 & 1/2q^2 \end{cases}$$

(ср. с (5)). Обратим внимание на то, что (как и в случае Поттс-стекольной нейросети) шумовая компонента $\eta_j^{(\mu)}$ при $q >> 1$ сосредоточена, в основном, в нуле: $\mathrm{Prob}\left\{\eta_j^{(\mu)} = 0\right\} = 1 - 1/q^2 \sim 1$.

Аналогами выражений (6) теперь будут

$$\begin{array}{ll} E(\xi) = (1-2a)(1-b), & E(\eta) = 0, \\ D(\xi) \to 0; & D(\eta) = \frac{1}{q^2}\cdot\alpha. \end{array}$$

При больших q дисперсия внутреннего шума в ПНС-2 еще меньше, чем в Поттс-стекольной нейросети:

$$D_{\text{ПНС}}(\eta)/D_{\text{PG}}(\eta) = 1/2, \text{ когда } q >> 1.$$

Этим и определяется в конечном итоге некоторое превосходство ПНС-2 над Поттс-стекольной моделью в объеме памяти и помехоустойчивости (см. ниже). Здесь, наконец, удобно сказать о механизмах, приводящих к подавлению внутренних шумов. В обеих моделях они одинаковы, но яснее всего их можно продемонстрировать на примере ПНС.

Распространяясь по межсвязи, сигнал взаимодейтсвует с хранящимися там частотами $\omega_{l_i^{(\mu)}} - \omega_{l_j^{(\mu)}} + \omega_{l_j} \to \{\omega_l\}_1^q$ при условии выполнения принципа несоизмеримости частот (13). На векторном языке это моделируется соотношением:

$$\mathbf{x}_i^{(\mu)} \mathbf{x}_j^{(\mu)+} \mathbf{x}_j = \mathbf{x}_i^{(\mu)} (\mathbf{x}_j^{(\mu)} \mathbf{x}_j) = \begin{cases} \mathbf{x}_i^{(\mu)}, & \text{если } l_j^{(\mu)} = l_j; \\ 0, & \text{в остальных случаях.} \end{cases} \tag{19}$$

Из последнего выражения ясно видно, что большая часть распространяющихся по межсвязям сигналов будет подавляться. Действительно, из массы возможных комбинаций индексов $l_j^{(\mu)}$ и l_j межсвязь выделяет только одну комбинацию – когда индексы совпадают – и только в этом случае пропускает сигнал. Все остальные комбинации индексов дают ноль – сигнал по межсвязи не проходит. Только с этим и связано то, что у ПНС-2 большая часть внутреннего шума η сосредоточена в нуле. Иначе говоря, межсвязь фильтрует сигналы.

Легко убедиться в том, что подобная фильтрация происходит и в Поттс-стекольной модели. Аналогом соотношения (19) в этом случае будет

$$\mathbf{x}_i^{(\mu)} \mathbf{x}_j^{(\mu)+} \mathbf{x}_j = \begin{cases} \frac{q-1}{q} \mathbf{x}_i^{(\mu)}, & \text{если } l_j^{(\mu)} = l_j; \\ -\frac{1}{q} \mathbf{x}_i^{(\mu)}, & \text{в остальных случаях.} \end{cases}$$

На первый взгляд все выглядит совершенно иначе: сигнал проходит по межсвязи всегда. Но при совпадении индексов $l_j^{(\mu)}$ и l_j сигнал снабжается большой положительной амплитудой порядка 1, а при несовпадении индексов – маленькой отрицательной аплитудой порядка $-1/q$. Эта фильтрация сигналов и приводит к подавлению в Поттс-стекольной модели внутренних шумов. Во всех остальных q-нарных моделях подобная фильтрация отсутствует.

Закончим с описанием свойств ПНС-2. Аналогами выражений (8) и (9) для помехоустйчивости и объема памяти будут:

$$\Pr_{err} < \sqrt{Np} \exp\left(-\frac{N(1-2a)^2}{2p} \cdot q^2 (1-b)^2 \right), \tag{20}$$

$$p_c = \frac{N(1-2a)^2}{2\ln N} \cdot q^2(1-b)^2. \tag{21}$$

При $q = 1$ выражения (20),(21) превращаются в известные результаты для модели Хопфилда (в этом случае шум по частоте отусутствует, $b = 0$). С ростом q экспоненциально спадает вероятность ошибки распознавания (20) – существенно растет помехоустойчивость сети. Одновременно, пропорционально q^2 растет и емкость памяти (21). В отличие от модели Хопфилда оказывается возможным эффективное запоминание числа образов p, большего, чем число нейронов N.

Например, зададимся уровнем ошибки распознавания $\mathrm{Pr}_{err} = 0.01$. В модели Хопфилда с такой вероятностью ошибки можно хранить только $p = N/10$ образов, причем уровень искажений не должен превышать 30%. В то же время, ПНС-2 при $q = 64$ способна восстанавливать любой из $p = 5N$ образов, искаженный не более, чем на 90%; или – любой из $p = 50N$ образов с уровнем искажений не более 65%. Компьютерные эксперименты подтверждают эти оценки.

По объему памяти ПНС-2 превосходит Поттс-стекольную модель в 2 раза. По-видимому, это связано с тем, что при одном и том же значении q число различных состояний нейронов в ПНС-2 в 2 раза больше, чем в Поттс-стекольной модели (за счет наличия у наших сигналов еще и амплитуд ± 1). В целом, обе модели очень близки по своим характеристикам.

4 Другие ПНС-архитектуры

4.1 Фазово-независимая ПНС-3

При реализации ПНС в виде устройства отдельную проблему для разработчиков составляет необходимость следить за согласованием фаз всех сигналов, приходящих к данному нейрону по межсвязям – известная проблема *набегания фазы*. Казалось бы, этой проблемы легко избежать, положив все фазы у сигналов одинаковыми. На формальном языке, мы просто должны приравнять 1 все амплитуды ± 1 в выражениях (11) и (14). Внимательный анализ, однако, показывает, что тогда парциальные шумовые компоненты $\eta_j^{(\mu)}$ перестают быть независимыми. В результате, дисперсия шума катастрофически нарастает. Выход из положения [22] состоит в том, чтобы в определении локального поля (15) использовать специально подобранные векторные пороги:

$$\mathbf{h}_i = \frac{1}{N}\left\{\sum_{j=1}^{N} \mathbf{T}_{ij}\mathbf{x}_j - \frac{1}{q}\sum_{\mu=1}^{p} \mathbf{x}_i^{(\mu)}\right\},$$

где матрицы $\mathbf{T}_{ij}$ определяются выражениями (16). Тогда парциальные шумовые компоненты $\eta_j^{(\mu)}$ становятся некоррелированными и можно применять описанный выше теоретико-вероятностный подход.

Средние и дисперсии суммарных случайных величин ξ and η получаются здесь в точности такими же, как в выражениях (6). А вся фазово-независимая ПНС – мы назвали ее ПНС-3, – по своим характеристикам эквивалентна Поттс-стекольной модели. По сравнению с ПНС-3 Поттс-стекольная модель выглядит неоправданно усложненной – это связано с использованием Поттсовских векторов $\mathbf{d}_l$ вместо ортов $\mathbf{e}_l$. Будучи реализованной на компьютере ПНС-3 работает в q раз быстрее, чем Поттс-стекольная модель.

4.2 Декоррелирующая ПНС

Мы придумали [23],[24], как использовать ПНС-архитектуру для существенного увеличения бинарной ассоциативной памяти в том случае, когда между бинарными образами имеются корреляции. Как известно, при наличии корреляций емкость памяти модели Хопфилда катастрофичеки уменьшается и единственным выходом из положения является так называемое *редкое кодирование* [25]-[29]. Наш метод составляет альтернативу этому подходу.

Основу нашего подхода составляет взаимно-однозначное отображение бинарных образов во внутреннее, промежуточное представление, использующее векторы-нейроны большой размерности. Свойства отображения таковы, что корреляции между векторно-нейронными образами становятся пренебрежимо малыми, а размерность q векторов-нейронов растет *экспоненциально* по параметру отображения. Затем на полученных векторно-нейронных образах строится ПНС. Поскольку распознающие свойтсва ПНС тем лучше, чем больше размерность q, а последняя растет экспоненциально, в результате мы имеем *экспоненциальное* по параметру отображения увеличнение емкости памяти.

Основная идея, лежащая в основе отображения бинарных образов в векторно-нейронные, очень проста. Пусть имеется N-мерный бинарный вектор

$$Y = (y_1, y_2, \dots, y_N),\ y_i = \pm 1.$$

Разобьем его мысленно на n фрагментов, содержащих по k элементов каждый: $N = n \cdot k$. Каждый из этих k-фрагментов превратим в целое число l по правилу:

$$l = 1 + \sum_{i=1}^{k} (y_i + 1) \cdot 2^{i-2}. \tag{22}$$

(Если бы для кодирования бинарных переменных использовалось представление $\{0/1\}$, а не представление $\{-1/1\}$, можно было бы просто сказать, что к каждому k-фрагменту мы относимся как к двоичной записи целого числа. Фактически, преобразование (22) выполняет ровно эту операцию.)

Возможные значения l заключены в интервале $1 \leq l \leq 2^k$. Для величины 2^k введем специальное обозначение: $q = 2^k$. Это q есть размерность векторов-нейронов того внутреннего представления, в которое происходит отображение исходных бинарных образов.

Сопоставим теперь каждому фрагменту вектор $\mathbf{x}$ из 2^k-мерного пространства R^{q}, равный декартову орту с номером l:

$$(y_1, y_2, \dots, y_k) \rightarrow\ l \rightarrow\ \mathbf{x} = \mathbf{e}_l\ \in\ \mathrm{R}^{\mathrm{q}},\ \text{где } \mathrm{q} = 2^{\mathrm{k}}.$$

Тогда всему бинарному образу будет сопоставлен набор из n таких 2^k-мерных ортов:

$$Y = (y_1, .., y_k, y_{k+1}, .., y_{2k}, \dots, y_{N-k+1}, .., y_N) \rightarrow X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n),$$

где $\mathbf{x}_i = \mathbf{e}_{l_i},\ i = 1, .., n;\ 1 \leq l_i \leq 2^k$. Очевидно, что это отображение взаимно-однозачно: по заданному набору X 2^k-мерных ортов бинарный образ восстанавливается единственным образом.

Описание отображения бинарных образов в векторно-нейронное представление на этом закончено: векторы $\mathbf{x}_i$ в дальнейшем будут играть роль векторов-нейронов. Проделав такое отображение для всех бинарных образов $\{Y^{(\mu)}\}_1^p$, получим набор векторно-нейронных образов $\{X^{(\mu)}\}_1^p$, на которых можно построить ПНС-3. Искаженный бинарный образ, который необходимо распознать с помощью модели Хопфилда, тоже подвергается такому отображению в вектороно-нейронное представление и распознается с помощью ПНС-3. А затем результат распознавания отображается обратно в бинарное представление.

Что дает такое отображение? Во-первых, при отображении бинарных образов в векторно-нейронные устраняются корреляции между образами. Это легко понять на качественном уровне (хотя может быть строго выражено и на количественном): достаточно двум бинарным k-фрагментам отличаться хотя бы в одной координате, чтобы они отобразились в два орта с совершенно различными индексами. Иначе говоря, даже при очень высокой коррелированности двух бинарных фрагментов они отобразятся в два разных, ортогональных друг другу орта. И второе: размерность q векторов-нейронов растет экспоненциально по параметру отображения k. Поскольку емкость памяти ПНС пропорциональна q^2, в результате получаем экспоненциальный (по параметру отображения k) рост объема памяти. Компьютерные эксперименты подтверждают эффективность описанной схемы.

5 Заключение

С начала 90-х годов интенсивность исследования q-нарных нейросетей заметно снизилась. По-видимому, это объясняется отсутствием прогресса в разработке эффективных моделей ассоциативной памяти. Моделирование ПНС-архитектур на компьютере показывает, что здесь мы приближаемся к таким характеристикам по объему нейросетевой памяти и помехоустойчивости, которые могут представлять интерес для практических приложений. Использование ПНС-архитектур представляется нам весьма перспективным.

Данная работа выполнялась в рамках выполнения проекта "Интеллектуальные компьютерные системы"(программа 2.45) при финансовой поддержке грантами Президента РФ НШ-1152-2003-1 и РФФИ 03-01-00355.

Список литературы

[1] D. Amit, H. Gutfreund, H. Sompolinsky. Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks. Phys. Rev. Lett., vol. 55, pp. 1530-1533, 1985; Information storage in neural networks with low levels of activity. Phys. Rev. A, vol. 35, pp. 2293-2303, 1987.

[2] J. Hertz, A. Krogh, R. Palmer. Introduction to the Theory of Neural Computation, NY: Addison-Wesley, 1991.

[3] I. Kanter. Potts-glass models of neural networks. Phys. Rev. A, vol. 37, pp. 2739-2742, 1988.

[4] J. Noest. Discrete-state phasor neural networks. Phys. Rev. A, vol. 38, pp. 2196-2199, 1988.

[5] J. Cook. The mean-field theory of a Q-state neural network model. Journal of Physics A, vol. 22, pp. 2000-2012, 1989.

[6] H. Rieger. Storing an extensive number of grey-toned patterns in a neural network using multistate neurons. Journal of Physics A, vol. 23, pp. L1273-L1279, 1990.

[7] D. Bolle, P. Dupont, J. van Mourik. Stability properties of Potts neural networks with biased patterns and low loading. Journal of Physics A, vol. 24, pp. 1065-1081, 1991.

[8] H. Vogt, A. Zippelius. Invariant recognition in Potts glass neural networks. Journal of Physics A, vol. 25, pp. 2209-2226, 1992.

[9] D. Bolle, P. Dupont, J. Huyghebaert. Thermodynamics properties of the q-state Potts-glass neural network. Phys. Rev. A, vol. 45, pp. 4194-4197, 1992.

[10] D. Bolle, J. van Mourik. Capacity of diluted multi-state neural networks. Journal of Physics A, vol. 27, pp. 1151-1162, 1994.

[11] F.Y Wu. The Potts model. Review of Modern Physics, vol. 54, pp. 235-268, 1982.

[12] Р. Бэкстер. Точно решаемые модели в статистической механике. Москва: Мир, 1985.

[13] Б.В.Крыжановский, А.Л.Микаэлян. О распознающей способности нейронной сети на нейронах с параметрическим преобразованием частот. ДАН (мат.-физ.), т. 383(3), стр. 318-321, 2002.

[14] A. Fonarev, B.V. Kryzhanovsky et al. Parametric dynamic neural network recognition power. Optical Memory and Neural Networks, vol. 10(4), pp. 31-48, 2001.

[15] N. Bloembergen. Nonlinear optics. NY: Benjamin, 1965.

[16] N. Chernov. A mesure of asymptotic efficiency for tests of hypothesis based on the sum of observations. Ann. Math. Statistics, vol. 23, pp. 493-507, 1952.

[17] B.V. Kryzhanovsky, A.L. Mikaelian, V.N. Koshelev et al. On recognition error bound for associative Hopfield memory. Optical Memory and Neural Networks, vol. 9(4), pp. 267-276, 2000.

[18] B.V. Kryzhanovsky, L.B. Litinskii, A. Fonarev. Optical neural network based on the parametrical four-wave mixing process. In: Wang, L., Rajapakse, J.C, Fukushima, K., Lee, S.-Y., and Yao, X. (eds.): Proceedings of the 9th International Conference on Neural Information Processing (ICONIP'02), Singapore: Orchid Country Club, vol. 4, pp. 1704-1707, 2002.

[19] Б.В.Крыжановский, Л.Б.Литинский. Векторные модели ассоциативной памяти. Автоматика и телемеханика, 2003, № 11, стр. 152-165.

[20] B.V. Kryzhanovsky, L.B. Litinskii, A.L. Mikaelian. Parametrical Neural Network. Optical Memory and Neural Networks, vol. 12(3), pp. 138-156, 2003.

[21] Y. Nakamura, K. Torii, T. Munaka. Neural-network model composed of multidimensional spin neurona. Phys. Rev. B, vol. 51, pp. 1538-1546, 1995.

[22] Д.И.Алиева, Б.В.Крыжановский. Бесфазовая модель параметрической нейронной сети. Международная конференция по искусственному интеллекту IEEE AIS'2003, т. 1, стр. 511-517, Москва: Физматлит, 2003.

[23] Б.В.Крыжановский, А.Л.Микаэлян. Ассоциативная память, способная распознавать сильно коррелированные образы. ДАН (информатика), т. 390(1), стр. 27-31, 2003.

[24] B. Kryzhanovsky, L. Litinskii, and A. Fonarev. An effective associative memory for pattern recognition. In: Advances in Intelligent Data Analysis V. 5th International Symposium on Intelligent Data Analysis IDA 2003, pp. 179-186, Berlin: Springer, 2003.

[25] G. Palm. Memory capacity of local rules for synaptic modification. Concepts in Neuroscience, vol. 2, pp. 97-128, 1991.

[26] G. Palm. On the information storage capacity of local learning rules. Neural Computation, vol. 4, pp. 703-711, 1992.

[27] G. Palm, F.T. Sommer. Information capacity in recurrent McCulloch-Pitts networks with sparsely coded memory states. Network, vol. 3, pp. 1-10, 1992.

[28] C.J. Perez-Vicente, D.J. Amit. Optimized network for sparsely coded patterns. Journal of Physics A, vol. 22, pp. 559-569, 1989.

[29] A.A. Frolov, I.P. Murav'ev. Informational characteristics of neural networks capable of associative learning based on Hebbiean plasticity. Network, vol. 4, pp. 495-536, 1995.