

В.А. Дяченко, О.Ф. Михаль, О.Г. Руденко

## Сеть Кохонена с параллельным обучением

Предложена модифицированная сеть Кохонена с алгоритмом поиска нейронов-победителей при параллельном предъявлении нескольких эталонных образов. Достижимое повышение производительности перспективно в развитии многоядерных многопроцессорных вычислительных систем. Объем обучающей выборки определяется с использованием элементов статистического моделирования.

A modified Kohonen network is suggested with the algorithm of searching for the neuron-winners under the parallel presentation of several master images. The reached increasing of the capacity has the prospects as to the development of multikernel multiprocessor computer systems. The number of training samples is defined by using the elements of statistical modeling.

Запропоновано модифіковану мережу Кохонена з алгоритмом пошуку нейрона-переможця за наявності кількох еталонних образів паралельно. Підвищення продуктивності, що досягається, є перспективним у розвитку багатоядерних багатопроцесорних обчислювальних систем. Обсяг навчальної виборки відшукується з використанням елементів статистичного моделювання.

**Введение.** Многие задачи распознавания изображений, речевых и других сигналов связаны с необходимостью хранения, обработки и передачи массивов данных, что требует значительных вычислительных ресурсов и временных затрат. Если массивы данных отличаются незначительно, то существенной экономии как вычислительных средств, так и необходимого для решения задачи времени можно достичь, используя сжатие данных или специальные методы их кодирования. При этом данные могут быть разбиты на кластеры – попарно непересекающиеся области – с назначением для каждого из них своего типичного представителя (опорного представителя, центра кластера). Эффективность кластерного подхода обусловлена возможностью интерпретировать и визуализировать получаемые результаты. Кроме того, кластеризация, как метод исследования, наиболее эффективна на начальных этапах обработки данных. Группирование данных позволяет, в частности, локализовать малодостоверные данные, с тем, чтобы на последующих этапах, возможно, выявить дополнительные влияющие факторы. Достаточно хорошо разработанные методы сжатия данных требуют для своей реализации наличия статистической информации об исследуемых процессах, в частности, плотности распределения. В условиях

отсутствия такой информации наиболее целесообразным представляется использование *искусственных нейронных сетей* (ИНС) [1]. Широко распространенным нейросетевым аппаратом кластеризации являются *самоорганизующиеся карты Кохонена* (СК) [2]. Обладая рядом несомненных достоинств, СК не лишены и недостатков, один из которых – использование в них последовательной процедуры обучения.

Возможность повышения эффективности данной ИНС обусловило появление сосредоточенных вычислительных систем на однокристалльных многоядерных процессорах, реализующих парадигму распределенной обработки данных [3], соответствующие прикладные направления которой связаны с интенсивным использованием параллельности обработки [4].

В [5, 6] упрощенно рассмотрены отдельные аспекты *адаптивной параллельной* процедуры обучения СК, обеспечивающей ряд преимуществ в производительности и эффективности обучения в сравнении с классическим вариантом. В настоящей статье представлены принципы построения СК с распараллеленным обучением, удобной к реализации на вычислительной системе с элементами параллельной обработки информации, в частности, на персональном компьютере с многоядерным однокристалльным процессором.

## Классический вариант самоорганизующейся карты Кохонена

Карты Кохонена – представители ИНС, обучаемых без учителя, т.е. не получающих информацию о желаемом выходном сигнале. Самоорганизующаяся сеть в процессе обучения разбивает предъявляемые входные сигналы на классы, строя соответствующие *топологические карты*. Таким образом, СК потенциально реализуют автоматизированные методы выявления новых структур в массивах данных: разведочный анализ данных (распознавание и установление близости кластеров) и обнаружение новых явлений (распознавание и отнесение данных к конкретным кластерам; если встречается наблюдение, не похожее ни на один из известных образцов, оно не классифицируется, т.е. выявляется его новизна).

Обучение карты предусматривает следующие циклически повторяющиеся действия:

- предъявление векторов *эталонных образцов* (ЭО) параллельно каждому из нейронов рецепторного поля;
- формирование по каждому (очередному) ЭО реакций каждого из нейронов;
- выявление нейрона-победителя  $x \in A$ , вектор весов которого покомпонентно минимально удален от вектора предъявляемого ЭО;
- корректировку векторов весов нейронов ближнего окружения  $x$  с учетом их удаленности и номера шага ( $k$ ) процесса обучения в соответствии с алгоритмом:

$$a_{ij}(k+1) = a_{ij}(k) + \alpha(k) f_{ij}(k) (x(k) - a_{ij}(k)), \quad (1)$$

где  $\alpha(k) \in (0, 1]$  – коэффициент усиления;  $f_{ij}(k)$  – монотонно убывающая функция

$$f_{ij}(k) = f(\|r_i - r_j\|, k) = f(d, k) = f(d, \sigma);$$

$r_i, r_j$  – векторы, определяющие положение нейронов  $i$  и  $j$  в решетке;  $d = r_i - r_j$  – принятая метрика;  $\sigma$  – параметр расстояния, задающий величину области «соседства» и уменьшающийся с течением времени до нуля.

### Адаптивная параллельная процедура

Описанная последовательная процедура обучения СК отличается громоздкостью. Выявлен-

ный на  $k$ -м шаге нейрон-победитель «актуален» только для его ближайшего окружения. Если для  $(k+1)$ -го ЭО, нейрон-победитель будет находиться далеко от предыдущего, процедура корректировки его окружения (1) не повлияет на результаты предыдущего шага обучения. Аналогично,  $(k+2)$ -й шаг обучения может быть таким, что его нейрон-победитель окажется далеко от предыдущих двух. Следовательно, соответствующие ЭО, с точки зрения их реальной *одинаковой* значимости для процесса обучения данной СК, могли бы предъявляться в произвольном порядке. Таким образом, процедура (1) с  $\alpha(k) \in (0, 1]$  и монотонно убывающей функцией  $f_{ij}(k)$  искажает реальную значимость ЭО для процесса обучения, вводя дополнительную («паразитную», не мотивированную) зависимость от порядка предъявления ЭО.

Следует также отметить, что параллельность ввода и начальной обработки информации рецепторным полем СК фундаментально противоречат последовательному характеру предъявления образов и обучения СК. Снятие данного противоречия посредством модифицирования структуры связей и процедуры обучения СК потенциально обеспечивает *положительный эффект*: повышает скорость обучения, а следовательно, эффективность процесса обучения и последующей эксплуатации СК. Соответственно изменяется структура организации информации, приобретаемой и хранимой СК. Следовательно, изменяются и эксплуатационные характеристики СК после обучения.

Вариант сокращения продолжительности процедуры обучения СК, включающий распараллеливание обработки предложен в [5, 6]. Предъявление ЭО, формирование реакций нейронов рецепторного поля и выявление нейрона-победителя – действия сугубо последовательные, поскольку в каждом шаге процедуры обучения по определению может быть единственный нейрон-победитель. Несколько иной смысл имеет коррекция векторов весов нейронов. После предъявления  $k$ -го ЭО победителем является нейрон  $x(k) \in \{a_{11}, a_{12}, \dots, a_{ij}, \dots, a_{mn}\}$ ; после предъявления  $(k+1)$ -го образца – ней-

рон  $x(k+1)$ . Допустима ситуация непересекаемости их областей корректировки:

$$F(k) \cap F(k+1) = \emptyset; F(k), \\ F(k+1) \subset A; f(k), f(k+1) > 0. \quad (2)$$

При этом при больших  $n$  и  $m$  вероятность непустого пересечения мала. Пусть последовательно предъявляется  $g$  образцов, по каждому из которых определяется свой нейрон-победитель, а с учетом (2) их расстановка на множестве  $A$  такова, что области корректировки  $F(1), F(2), \dots, F(g)$  попарно не пересекаются:

$$F(p) \cap F(q) = \emptyset; p, q \in (1, 2, \dots, g); p \neq q. \quad (3)$$

Подобная ситуация достижима последовательным предъявлением произвольных ЭО с изъятием, если очередной предъявленный ЭО не удовлетворяет (3). При этом в каждый цикл из  $g$  ЭО подбираются образы, потенциально относящиеся к разным кластерам, т.е. процедура разбивки на кластеры дополнительно поддерживается указанным изъятием при предъявлении ЭО, реализуемым аналогично «временному откату».

Далее для сформированного  $g$ -элементного набора нейронов-победителей конструируется суммарная (совокупная, композиционная) функция корректировки

$$f_g = f(1) \cup f(2) \cup \dots \cup f(h) \dots \cup f(g) = \\ = \min(f(1), f(2), \dots, f(h), \dots, f(g)), \quad (4)$$

описывающая степень принадлежности окрестных нейронов к соответствующим нейронам-победителям, имеющая  $g$  локальных максимумов в точках расположения нейронов-победителей. С учетом (4) множество нейронов  $A$  (рецепторное поле) корректируется за один раз, т.е. для нейронов-победителей процедура корректировки является параллельной. Процедура повторяется до исчерпания набора ЭО (рис. 1).

Особенность данного подхода состоит в том, что число  $g$  задается изначально, по результатам предварительного вероятностного оценивания, индивидуального для конкретной предметной ситуации с учетом (2) и (3). При этом в процессе формирования очередного набора ЭО

те ЭО, которые не удовлетворяют ситуации, откладываются и предъявляются в следующих циклах обучения. Очевидно, что в каждом из циклов по  $g$  предъявлений ЭО амплитудного изменения (коэффициент  $\alpha(k)$  в (1)) и изменения размеров окрестности, охватываемой  $f_{ij}(k)$ , не требуется, поскольку внутри группы отдельные нейроны-победители не конкурируют за ближнее окружение. В результате, обучение СК становится менее продолжительным и потенциально более эффективным: требуется обучающая выборка ЭО меньшего объема либо при фиксированном объеме обучающей выборки результат тренировки будет более устойчивым при воспроизводимости конфигураций кластеров.

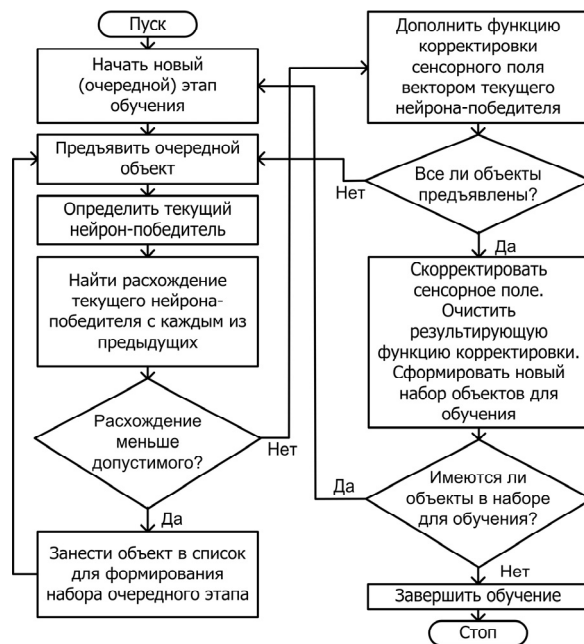


Рис. 1. Схема алгоритма обучения модифицированной сети Кохонена

### Вероятностное оценивание объемов обучающей выборки

Для вероятностной оценки объемов  $g$  наборов ЭО целесообразно проведение статистического моделирования. Рассмотрим случай двумерного рецепторного поля размером  $m \times n$ , с неизменной конфигурацией окрестности нейрона-победителя (коэффициент  $f_{ij}$  в (1)), включающей в себя только четыре ближайших (соседних) нейрона. В данном варианте допустимы три структурно различных вида нейронов:

- *угловые* (если такой нейрон становится победителем, изменяются весовые значения только у двух близлежащих нейронов), таких нейронов  $k_1 = 4$ ;

- *краевые* (изменяются значения трех близлежащих нейронов), число таких нейронов *линейно* зависит от размеров сети:  $k_2 = 2(m + n - 4)$ ;

- *центральные* (изменяются значения четырех нейронов), их число *квадратично* зависит от размеров сети:  $k_3 = mn - 2(m + n - 2)$ .

Упрощаем модельную ситуацию: пусть  $m = n$ . Тогда

$$k_1 = 4; k_2 = 4(n - 2); k_3 = n^2 - 4(n - 1).$$

Непосредственным подсчетом находим, что при  $n \sim 40$  величина  $(k_1 + k_2)$  составляет порядка 10% от  $k_3$ ; при  $n \sim 400$  – порядка 1%; при  $n \sim 4000$  – порядка 0,1%, т.е. краевые эффекты сглаживаются только при существенном расширении рецепторного поля. Поэтому возможности упрощения вероятностного оценивания ограничены.

Необходимо рассчитать вероятности непересечения областей, которые в первую очередь зависят от того, где расположится нейрон-победитель, так как область возбужденных нейронов может иметь различные размеры. Далее нужно рассчитать количество возможных вариантов с первым нейроном-победителем. Этим нейроном может оказаться как угловой, так и краевой и центральный нейроны. Рассмотрим вероятности для указанных случаев:

- угловой нейрон:  $P = 4/(mn)$ ;
- краевой нейрон:  $P = 2(m + n - 4) / (mn)$ ;
- центральный нейрон:

$$P = (mn - 2(m + n - 2)) / (mn) = 1 - (2(m + n - 2) / (mn)).$$

Далее вероятность непересечения может быть рассчитана исходя из того, что надо учитывать местоположение областей уже возбужденных нейронов, т.е. чтобы следующий нейрон-победитель и нейроны, находящиеся возле него, не попадали в такие области. Соответствующие аналитические выражения остаются простыми лишь в предположении малости вероятности непустого пересечения, т.е. для больших ре-

цепторных полей. Оценку предельных допустимых объемов параллельной обучаемости СК представляется более простым выполнить методами статистического моделирования, позволяющими учесть также конфигурацию рецепторного поля (случай  $m \neq n$ ). Моделирование позволяет получить наборы статистических оценок для вероятностей заданного процента использования исходного обучающего материала при равномерном двумерном статистическом распределении выпадений нейронов-победителей по рецепторному полю СК. Зная указанные вероятности и имея данные о выигрыше в производительности модифицированной процедуры обучения СК в сравнении с традиционным вариантом, можно оценить оптимальные объемы выборок ЭО. Критерием оптимальности может быть, в частности, минимальное время обучения СК.

Моделирование реализуется с использованием генераторов случайных чисел, воспроизводящих процесс нахождения нейрона-победителя. В структурном отношении модель аналогична алгоритму обучения, приведенному на рис. 1, и дополнена стандартными блоками статистического анализа и документирования результатов. Укрупненно, моделирование включает в себя следующие этапы.

- Ввод матрицы  $A$  рецепторного поля (размеров  $m, n$  и элементов  $a_{ij}$ , где  $i \in (1, 2, \dots, m)$ ,  $j \in (1, 2, \dots, n)$ ).

- Подсчет количества краевых  $k_2$  и центральных  $k_3$  нейронов в соответствии с заданной размерностью матрицы.

- На первом шаге (после предъявления первого образца) количество возможных нейронов-победителей равно  $S = m \times n$ .

- Получение нейрона-победителя  $a_{ij}$  после предъявления ЭО. Определение его типа: краевой ( $i \in (1, m), j \notin (1, n)$  или  $i \notin (1, m), j \in (1, n)$ ); центральный ( $i, j \in (1, n, m)$ ); угловой ( $i \in (1, m), j \in (1, n)$ ). В упрощенном варианте модели активационная функция изменяет значения весов только нейронов, окружающих нейрон-победитель ( $a_{i-1,j-1}, a_{i-1,j}, a_{i-1,j+1}, a_{i,j-1}, a_{i,j+1}, a_{i+1,j-1}, a_{i+1,j}, a_{i+1,j+1}$ ). При этом количество изменяемых ней-

ронов составляет три – для угловых, пять – для краевых и восемь – для центральных. Таким образом, после предъявления каждого образца и определения нейрона-победителя из  $S$  вычитается количество измененных нейронов, включая нейрон-победитель.

• После предъявления каждого образца, сравниваем  $S$  с пороговым значением  $Q$ . Если  $S > Q$ , продолжаем предъявление образцов; если  $S = Q$ , – цикл предъявления образцов окончен.

Возможен также альтернативный вариант построения модели: многократная генерация ЭО (псевдослучайное задание координат нового нейрона-победителя), проверка расхождения его с уже размещенными нейронами-победителями, статистическое оценивание вероятности успешного размещения в зависимости от числа предъявленных ЭО.

На графиках (рис. 2) представлены результаты моделирования для рецепторных полей нейронов. Графики 1, 2, ..., 5 соответствуют допустимым расстояниям между элементами 1, 2, ..., 5 единиц. Вероятность  $P$  наложения элементов растет с ростом числа  $N$  предъявляемых элементов и допустимого расстояния их расположения. Вероятность существенно зависит от размеров рецепторного поля: с учетом масштаба, весь массив графиков ( $\delta$ ) может быть сжат и помещен в нижнюю строку клеток графика ( $a$ ). Зависимость вероятности от асимметрии поля (случай  $m \times n$ ,  $m \neq n$ ) на данном этапе моделирования не рассматривалась.

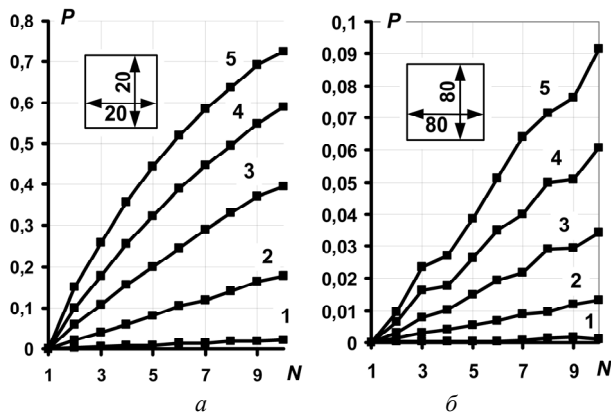


Рис. 2. Вероятность  $P$  наложения объектов от номера  $N$  предъявляемого объекта для рецепторных полей размером:  $a - 20 \times 20$  и  $b - 80 \times 80$

Результаты могут быть интерпретированы следующим образом. Согласно, например, графику 3 (рис. 2, б) при восьми последовательных предъявлениях объектов вероятность наличия наложения не превышает 3%. Для конкретной программной реализации при известном вы

игрыше в производительности процесса обучения (см. рис. 1) в сравнении с традиционным вариантом сети Кохонена, на основании результатов моделирования, представленных на рис. 2, могут быть даны конкретные рекомендации по оптимизации объема обучающей выборки. Критерий оптимальности для заданной конфигурации параметров сети Кохонена (формата рецепторного поля, допустимого расстояния между элементами и др.) может быть сформулирован, в частности, как *максимизация объема фрагмента обучающей выборки (минимизация фрагментированности обучающей выборки)* при непревышении заданной вероятности наложения элементов.

**Заключение.** Информационная емкость сети Кохонена определяется объемом рецепторного поля, с возрастанием которого существенно растет объем обучающей выборки. Перспективным направлением повышения производительности сети на этапе обучения является предложенный метод параллельного обучения, в котором нахождение нейрона-победителя осуществляется параллельно для нескольких эталонных образов – фрагментов обучающей выборки. Однако при этом возникает задача оценки оптимальных объемов фрагментов выборки, решаемая средствами статистического моделирования. Предложенный метод параллельного обучения имеет приоритет в свете развития средств распределенной обработки данных, в частности с использованием многоядерных и многопроцессорных вычислительных систем.

1. Руденко О.Г., Бодянский Е.В. Основы теории искусственных нейронных сетей. – Харьков: ТЕЛЕТЕХ, 2002. – 317 с.
2. Kohonen T. Self-organization and associative memory // Series in Information Sciences. – 1984. – 8. – Berlin: Springer Verlag.
3. Эндрюс Г.Р. Основы многопоточного, параллельного и распределённого программирования. – М: Изд. дом «Вильямс», 2003. – 512 с.
4. Михаль О.Ф., Руденко О.Г. Моделирование на сетях Петри виртуального вычислительного устройства для исследования эффективности локально-параллельных алгоритмов // УСиМ. – 2003. – № 3. – С. 18–28.
5. Дяченко В.А. Синтез карт Кохонена с адаптивным параллельным обучением // 12-й Междунар. молодежный форум «Радиоэлектроника и молодежь в XXI веке», г. Харьков, ХНУРЭ, 2008.
6. Дяченко В.А. Адаптивная параллельная обработка карт Кохонена. Области корректировки // Мат-лы междунар. науч.-практ. конф. студентов и аспирантов «Информационные технологии в экономике и образовании». – М.: Российский ун-т кооперации, 2008. – С. 54–58

Поступила 10.02.2009  
E-mail: fuzzy16@pisem.net

© В.А. Дяченко, О.Ф. Михаль, О.Г. Руденко, 2009