

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ГОСУДАРСТВЕННАЯ КОРПОРАЦИЯ ПО АТОМНОЙ ЭНЕРГИИ «РОСАТОМ»
РОССИЙСКАЯ АКАДЕМИЯ НАУК
РОССИЙСКАЯ АССОЦИАЦИЯ НЕЙРОИНФОРМАТИКИ
МОСКОВСКИЙ ИНЖЕНЕРНО-ФИЗИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ
СИСТЕМНЫХ ИССЛЕДОВАНИЙ РАН

НАУЧНАЯ СЕССИЯ МИФИ–2009

НЕЙРОИНФОРМАТИКА–2009

**XI ВСЕРОССИЙСКАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ
КОНФЕРЕНЦИЯ**

**ЛЕКЦИИ
ПО НЕЙРОИНФОРМАТИКЕ**

По материалам Школы-семинара
«Современные проблемы нейроинформатики»

Москва 2009

УДК 001(06)+004.032.26 (06) Нейронные сети
ББК 72я5+32.818я5
М82

НАУЧНАЯ СЕССИЯ МИФИ–2009. XI ВСЕРОССИЙСКАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ «НЕЙРОИНФОРМАТИКА–2009»: ЛЕКЦИИ ПО НЕЙРОИНФОРМАТИКЕ. – М.: МИФИ, 2009. – 280 с.

В книге публикуются тексты лекций, прочитанных на Школе-семинаре «Современные проблемы нейроинформатики», проходившей 27–30 января 2009 года в МИФИ в рамках XI Всероссийской конференции «Нейроинформатика–2009».

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

Ответственный редактор
Ю. В. Тюменцев, кандидат технических наук

ISBN 978–5–7262–1053–7 © *Московский инженерно-физический институт (государственный университет), 2009*

Содержание

<i>Г. А. Ососков. Практические аспекты нейровычислений в экспериментальной физике</i>	163
Введение	164
Краткое резюме	114



Г. А. ОСОСКОВ

Объединенный институт ядерных исследований,

г. Дубна

E-mail: ososkov@jinr.ru

**ПРАКТИЧЕСКИЕ АСПЕКТЫ НЕЙРОВЫЧИСЛЕНИЙ
В ЭКСПЕРИМЕНТАЛЬНОЙ ФИЗИКЕ**

Аннотация

Обсуждаются методы обработки экспериментальных данных с помощью искусственных нейронных сетей (ИНС), которые интенсивно применялись в последнее десятилетие в Объединенном институте ядерных исследований. После краткого введения в теорию нейросетей автор приводит ряд наиболее полезных и перспективных примеров применений прямооточных и полносвязных ИНС в экспериментальной физике, генетике и разработке экспертных систем. Помимо интересных деталей различных применений ИНС, автор уделяет внимание также таким практическим аспектам, как выбор наиболее подходящих типов и структуры ИНС, методов сжатия данных для ввода их в ИНС и критериям качества при оптимизации эффективности работы ИНС.

G. A. OSOSKOV

Joint Institute for Nuclear Research,

Dubna

E-mail: ososkov@jinr.ru

**PRACTICAL ASPECTS OF NEUROCOMPUTING APPLICATIONS
IN EXPERIMENTAL PHYSICS**

Abstract

Experimental data processing methods by means of artificial neural networks (ANN) are discussed. These methods are used intensively during the last decade at the Joint Institute for Nuclear Research. After a brief introduction to the ANN field some useful and promising application examples for feedforward and recurrent neural nets are considered in regard to experimental physics, genetics and expert systems. Considerable attention is focused on appropriate ANN applications as well as on various practical aspects including a choice of the most suitable ANN class and topology, compression techniques for ANN input data, performance measure to optimize ANN.

Введение

Растущий интерес к использованию искусственных нейронных сетей (ИНС) вполне оправдан необычайно широкими возможностями их применения в самых разнообразных областях научных и прикладных исследований. Это подтверждается лавинообразным потоком журнальных статей, трудов конференций и, особенно, ссылок в интернете на темы многочисленных приложений ИНС для решения обширного круга проблем классификации [1, 2], распознавания образов [3], аппроксимации многомерных функций [4], разработки систем ассоциативной памяти и экспертных систем [5], принятия оптимальных решений в сложных многокритериальных задачах [6], а в последние годы и для несеточных методов приближенных решений дифференциальных уравнений математической физики [7, 8].

Восполнен, наконец, дефицит монографий и учебников, посвященных нейросетям и нейровычислениям: в дополнение к вышедшим еще в начале 90-х годов двум монографиям [9, 10], переведены книги [11, 12], а в издательстве «Радиотехника» появилась целая серия книг «Нейрокомпьютеры и их применение» под ред. *А. И. Галушкина*.

На рынке появились пакеты программ-конструкторов ИНС для выполнения компьютерных реализаций нейросетевых алгоритмов [14].

Тем не менее, исследователь, который собирается использовать ИНС для решения какой-то своей конкретной задачи, немедленно сталкивается с целым рядом проблем их практического применения, таких как:

- выбор наиболее подходящего типа ИНС (прямоточная, полносвязная и т. д.);
- выбор структуры и других параметров ИНС (число слоев, число нейронов в скрытых слоях, вид активационной функции, метод обучения и т. д.);
- способ преобразования исходных данных для их ввода в ИНС;
- как обеспечить сходимость процесса обучения ИНС;
- по каким критериям сравнивать различные ИНС и результаты их работы.

Хотя к настоящему времени уже появились в печати весьма содержательные рекомендации по решению указанных проблем [13, 14], существует определенная специфика в подходах к подготовке и организации нейровычислений, характерная для разных областей их применения. В настоящей статье делается попытка показать, как подобные проблемы преодолевались

на примере ряда задач, связанных с нейровычислениями в Объединенном институте ядерных исследований в Дубне.

Краткое напоминание об искусственных нейронных сетях

Напомним вкратце об основных понятиях ИНС и типах ИНС, используемых в физике:

- система ИНС состоит из очень большого числа одинаковых нейронов, причем результат ее работы малочувствителен к характеристикам отдельного нейрона;
- каждый нейрон представляет собой простое логическое устройство, характеризуемое:
 - уровнем активации,
 - топологией связей нейронов друг с другом,
 - мерой взаимодействия с другими нейронами, называемой синаптической силой связи или весом. Веса этих связей различны и могут определяться в зависимости от решаемой задачи,
 - выходным уровнем, который связан с уровнем активации посредством некоторой функции обычно сигмоидального типа;
- в системе существенна возможность параллельной обработки информации.

Нейроны можно классифицировать по их положению в топологии сети следующим образом:

- *входные нейроны* — по сути они не являются нейронами, а только принимают исходный вектор, кодирующий входной сигнал;
- *промежуточные или скрытые нейроны* — выполняют основные вычислительные операции;
- *выходные нейроны* — обеспечивают выходы сети, в них также могут производиться какие-либо вычислительные операции.

Как правило, для унификации определения нейрона скрытые и выходные нейроны задаются одинаково, так что математический формализм работы искусственного нейрона можно выразить следующим образом. Обозначим сигнал, исходящий от j -го нейрона сети как x_j , вес синаптической связи i -го и j -го нейронов как w_{ij} , что в векторной форме задает входной вектор \vec{X} и матрицу весов \mathbf{W} , в которой j -й столбец может рассматриваться

как вектор весовых коэффициентов \vec{W}_j . Для входного сигнала, поступающего на i -й нейрон от всех остальных нейронов, используется линейное выражение

$$h_i = \sum_j w_{ij}x_j + w_0x_0, \quad (1)$$

являющееся записью скалярного произведения векторов \vec{X} и \vec{W}_j , где дополнительный вход $x_0 = 1$ и соответствующий ему вес необходимы для того, чтобы линейная функция (1), аппроксимирующая входной вектор, могла иметь смещение относительно начала координат, чем формируется порог чувствительности нейрона. Выходной сигнал i -го нейрона получается путем воздействия на h_i активационной функцией $g(u)$, которая определяет зависимость сигнала на выходе нейрона от взвешенной суммы сигналов на его входах, т. е.

$$y_i = g(h_i), \quad (2)$$

Функция $g(u)$ выбирается в зависимости от решаемой задачи. В большинстве случаев она является монотонно возрастающей и имеет область значений $[-1, 1]$ или $[0, 1]$.

Так, пороговая функция $g(u)$ обеспечивает переход нейрона в возбужденное состояние ($y_i = +1$), если сумма (1) превысит заданный порог.

Однако чаще всего $g(u)$ — это сжимающая функция сигмоидального вида

$$g(u) = \frac{1}{1 + e^{-\lambda u}}, \quad (3)$$

что обеспечивает уменьшение влияния выбросов, т. е. больших входных значений, а также необходимо для упрощения алгоритмов обучения сети, требующих, чтобы она была непрерывно дифференцируемой на всей числовой оси и имела просто вычисляемую производную. Для $g(u)$ вида (3) имеем

$$g'(u) = g(u)(1 - g(u)).$$

Параметр λ в (3) меняет наклон сигмоиды, увеличивая крутизну кривой при росте λ .

Существуют также и радиально-базисные функции (РБФ) активации. Этот тип функций принимает в качестве аргумента расстояние $\rho(\vec{X}, C)$, между входным вектором \vec{X} и некоторым наперед заданным центром активационной функции C . Значение этой функции тем выше, чем ближе

входной вектор к центру. В качестве РБФ можно, например, использовать функцию Гаусса:

$$y = e^{-\frac{\rho(\vec{X}, C)^2}{2\sigma^2}}, \quad (4)$$

Скалярный параметр σ определяет скорость спадания функции при удалении вектора от центра и называется *шириной окна*. Сети, с нейронами, использующими такие функции, называются РБФ-сетями. В качестве расстояния между векторами могут быть использованы различные метрики, обычно используется евклидово расстояние:

$$\rho(\vec{X}, C) = \sqrt{\sum_{j=1}^n (x_j - c_j)^2}. \quad (5)$$

Здесь x_j это j -я компонента вектора, поданного на вход нейрона, а c_j — j -я компонента вектора, определяющего положение центра передаточной функции, $j = 1, 2, \dots, n$, где n — размерность входных данных (например, для точек на плоскости $n = 2$).

Однако в зависимости от задачи в качестве метрики можно использовать сумму модулей разностей компонент $\sum_j |x_j - c_j|$ (так называемое манхэттенское расстояние), или максимум этих модулей $\max_i |x_i - c_i|$ или так называемое расстояние Махаланобиса

$$d^2(\vec{X}, C) = (\vec{X} - C)^{-T} \Sigma^{-1} (\vec{X} - C)^T,$$

где Σ — это оценка функции ковариации этих векторов.

Сами радиальные нейроны могут задаваться своим центром C и «радиусом», т. е. величиной порога для функции активации (веса в таком случае выбирают равными 1). Однако формулу (4) можно трактовать и по-иному, как в [14, 15], где предлагается рассматривать точку центра и отклонение как «вес» входящего сигнала, так что в отличие от скалярного произведения (1) радиальный нейрон обрабатывает входы как

$$h_i = \sum_j (x_j - w_{ij})^2. \quad (6)$$

По этой же схеме функционируют нейроны в самоорганизующейся сети Кохонена [16].

Перейдем теперь к вопросу об организации нейронных сетей, т. е. о топологии связей нейронов в сети. Простейшая сеть имеет *прямоточную*

структуру, когда сигналы передаются от входа через скрытые слои к выходу без обратных связей. Такая структура имеет устойчивое поведение. Гораздо сложнее поведение *полносвязных* или рекуррентных сетей, в которых все нейроны связаны друг с другом. Такие ИНС могут быть неустойчивыми и иметь очень сложную динамику поведения.

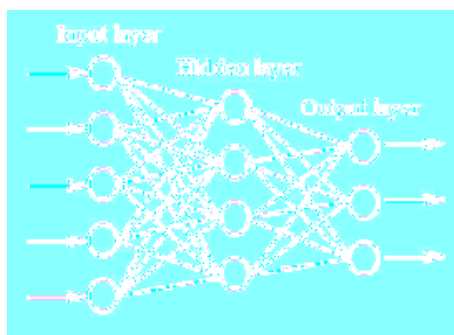


Рис. 1. Прямоточная ИНС (МСП или РБФ-сети)

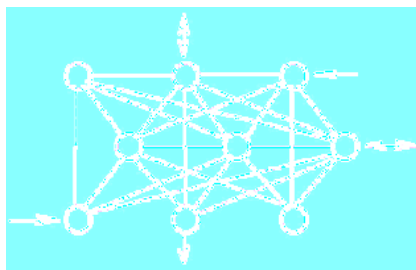


Рис. 2. Полносвязная ИНС (сеть Хопфилда)

Оба типа ИНС активно применяются в физике. Типичный пример прямоточной сети, к которым относятся *многослойные перцептроны* (МСП) или радиально-базисные (РБФ) сети, показан на рис. 1. Структура полносвязной сети такой, как нейронная *сеть Хопфилда* (ХНС) показан на рис. 2. Мы можем также рассматривать и *клеточные автоматы* как специальный тип нейронных сетей с локальными связями.

Напомним кратко основные особенности этих типов ИНС.

Прямоточные сети

Существенная особенность нейронных сетей состоит в том, что зависимость между входом и выходом находится в процессе обучения сети.

Для обучения нейронных сетей применяются алгоритмы двух типов (разные типы сетей используют разные типы обучения): управляемое («обучение с учителем») и не управляемое («обучение без учителя»). Для многослойных перцептронов применяется обучение с учителем, при котором на вход МСП подается предварительно подготовленная *обучающая выборка* МСП, которая должна обучать, подавая множество образов по одному на его вход и подстраивая веса до тех пор, пока для всех образов не будет достигнут требуемый выход. При обучении предполагается, что для каждого входного вектора существует парный ему целевой вектор, задающий требуемый выход. Вместе они называются *обучающей парой*. Сеть обучается на многих парах, которые и составляют обучающую выборку. Наиболее распространенным алгоритмом обучения МСП является *алгоритм обратного распространения ошибки*, состоящий в минимизации суммарной ошибки сети:

$$E = \frac{1}{2} \sum_{k=1}^m \sum_j (y_j^{(k)} - z_j)^2, \quad (7)$$

где y_j — значение j -го выхода сети, z_j — идеальное (желаемое) значение j -го выхода сети, $k = 1, \dots, m$ — индекс по числу обучающих выборок (эпох обучения).

Для поиска минимума E обычно применяется метод наискорейшего спуска в пространстве весов. При этом ошибки (точнее величины коррекции весов) распространяется в обратном направлении от выходов к входам, сквозь веса, соединяющие нейроны. Дифференцируя (7) по всем весам, входящим в y_j и, соответственно, в h_j , мы получим систему уравнений для определения весов. Размерность этой системы равна числу межнейронных связей в сети, т. е. уже для простой сети, изображенной на рис. 1, будет равна 32. Для сетей с большим числом входных нейронов число уравнений может достигать сотен и даже тысяч, что образно называют «проклятием размерности» многослойных нейросетей. Рекуррентные правила метода наискорейшего спуска приводят к итерационным формулам

- для весов выходного слоя

$$\Delta w_{kj}(t+1) = -\eta(y_j^t - z_j^t)g'(y_j^t)h_k^t,$$

- для весов скрытого слоя

$$\Delta w_{ik}(t+1) = -\eta \sum_j w_{kj}g'(h_k^t)x_k^t,$$

где η — параметр скорости обучения, $g(x)$ — активационная функция, z — известный результат классификации.

Каждая модификация всех весов в сети называется эпохой обучения. В общем случае количество эпох неограниченно. Сеть считается обученной, если в какой-либо эпохе максимальная ошибка обучения $\max_{t,j} |y_{t,j} - z_{t,j}|$ уменьшится до заданной точности. Выбор структуры сети весьма важен для повышения ее эффективности. Существуют рекомендации по поводу выбора числа нейронов в скрытом слое позволяющие оценить необходимое число синаптических весов для успешного обучения многослойного персептрона [13]:

$$\frac{N_y Q}{1 + \log_2 Q} \leq N_w \leq N_y \left(\frac{Q}{N_x} + 1 \right) (N_x + N_y + 1) + N_y, \quad (8)$$

где N_y — размерность выходного сигнала; Q — число элементов обучающей выборки; N_w — необходимое число синаптических весов; N_x — размерность входного сигнала.

Полносвязная сеть Хопфилда

Как упоминалось, в общую схему ИНС помимо широко применяемых персептронов, укладывается также сеть Хопфилда (НСХ). Это полносвязная сеть, являющаяся в простейшем случае системой простых бинарных нейронов s_i , которые могут принимать одно из двух различных значений, например, $\{+1, -1\}$. Эволюция НСХ приводит ее в некоторое состояние стабильности, устойчивого равновесия. Рассматривая ИНС как динамическую систему бинарных нейронов, Хопфилд использовал билинейную функцию Ляпунова как функционал энергии сети [17]

$$E(s) = -\frac{1}{2} \sum_{ij} s_i w_{ij} s_j \quad (9)$$

и показал, что для симметричной матрицы весов $w_{ij} = w_{ji}$ с нулевой диагональю $w_{ii} = 0$ и асинхронной динамикой сети ее энергия в результате эволюции убывает в локальные минимумы, соответствующие точкам стабильности сети.

В соответствии с вышесказанным, для нахождения стационарного состояния сети требуется найти точку минимума энергетического функционала по значениям состояний нейронов. Применение с этой целью метода градиентного спуска к функционалу энергии (9) приводит к системе уравнений, определяющей динамику сети:

$$s_i = -\frac{1}{2} \left(1 + \operatorname{sign} \left(-\frac{\partial E}{\partial s_i} \right) \right). \quad (10)$$

Однако процедура итерационного решения системы (10) для случая бинарных нейронов часто приводит в какой-нибудь локальный минимум функционала энергии. Кроме того, во многих практических приложениях НСХ бинарные нейроны оказываются нереалистичной идеализацией. В этой связи Хопфилд в [18] предложил обобщение своей модели ИНС на случай нейронов с непрерывным множеством состояний.

Стандартным путем перехода к нейронам с непрерывными состояниями является введение статистического шума в систему с последующим применением теории среднего поля, что приводит к усреднению значений состояний нейронов и замене ступенчатой функции (3) на специальную функцию сигмоидального вида:

$$v_i = \frac{1}{2} \left(1 + \tanh \left(-\frac{\partial E}{\partial v_i} \cdot \frac{1}{T} \right) \right) = \frac{1}{2} \left(1 + \tanh \left(-\frac{H_i}{T} \right) \right). \quad (11)$$

Здесь температура T соответствует уровню статистического шума, а $H_i = \langle \sum_j w_{ij} s_j \rangle_T$ — локальное среднее поле нейрона. Значения v_i , представившие быть целочисленными, определяют уровень активности нейрона, т. е. в случае $v_i > v_{min}$ нейрон считается активным. Температура убывает по схеме «имитационного отжига» (simulated annealing) [19]. С помощью системы уравнений (11) состояния нейронов сети итеративно обновляются до достижения ею стабильной точки. Однако то, что нейроны v_i теперь не являются бинарными, позволяет оперативно следить за теми из них, которые соединяют точки треков и под стимулирующим воздействием весов постепенно увеличивают уровень своей активности в процессе эволюции сети. В качестве порога уровня активности обычно выбирается $v_{min} = 0.5$.

Работы Хопфилда породили чрезвычайный интерес к таким сетям, в частности, потому, что устанавливалась явная связь процессов их эволюции с обширным кругом проблем оптимизации, формализуемых обычно как задача поиска экстремума функционала при наличии ограничений на его параметры.

Кроме того, сети Хопфилда можно использовать как ассоциативную память, т. е. восстанавливать предъявленный образ даже при наличии некоторых изменений. Должна быть задана фиксированная конфигурация сети. Требуется так подобрать синаптические веса w_{ij} , чтобы для любой начальной конфигурации, расположенной достаточно близко к требуемому образу, получалась требуемая конфигурация нейронов. Для матрицы синаптических весов следующее выражение решает задачу запоминания одного образа:

$$w_{ij} = \frac{1}{N} x_i x_j, \quad (12)$$

где N — количество нейронов, x_i — i -й нейрон, x_j — j -й нейрон. Множитель $1/N$ введен для того, чтобы для больших N энергия имела необходимую пропорциональность числу нейронов. Для нескольких образов $p = 1, \dots, M$ при не очень большом M решением будет следующее выражение:

$$w_{ij} = \frac{1}{N} \sum_{p=1}^M x_i^p x_j^p. \quad (13)$$

В НСХ реализуется алгоритм обучения без учителя, т. е. используется только входные векторы и на их основании сеть учится давать наилучшие значения выходов.

Выражение (13) имеет глубокую биологическую основу. Если при запоминании образа i -й нейрон посылает импульс на j -й нейрон, то соответствующая этой паре синаптическая связь усиливается и сигналы между этими нейронами проходят с большим весом. Образы должны быть *некоррелированными*, т. е. не иметь общих характерных черт.

Возможно также увеличение числа образов, записанных в нейронную сеть, так что их число будет пропорционально N , а именно, $M = \beta N$, где β — коэффициент пропорциональности. Его максимально возможное значение, при котором еще происходит восстановление большей части образов, было найдено Хопфилдом. Получено, что при $0.14 < \beta < 0.16$ нейронная сеть быстро насыщается, при этом возникают ложные минимумы от перекрытия разных образов, а часть минимумов исчезает. Восстановление

нужного образа происходит с вероятностью 0.85, в 10% случаев сеть приходит в минимум, не отвечающий ни одному образцу. Это позволяет на основе нейронной сети создавать базы данных, поиск в которых осуществляется не по каталогу, а в процессе вычислений на основе предъявленного фрагмента требуемого образа.

Если образы имеют общие характерные черты, то сеть может выдавать неверный результат: только совпадающие фрагменты изображений. Выход оказался следующим. Информация в нейронной сети хранится не в нейронах, а в матрице синаптических весов. Общее число различных элементов в симметричной матрице равно $N(N - 1)/2$. Даже если $M = 0.15N$, то в нейронной сети содержится много избыточной информации об образах, что повышает надежность НСХ даже при исключении большого числа связей между нейронами. В случае коррелированных образов достаточно исключить связи, вычисляемые следующим образом. Если хотя бы для одного образа знак соответствующего слагаемого не совпадает со знаком всей суммы, то эта связь между нейронами исключается, $w_{ij} = 0$. Если до исключения нежелательных элементов сеть отказывалась распознавать даже некоторые чистые символы, то после исключения нежелательных связей, когда их число составило 16% от первоначального числа, нейронная сеть восстанавливает символы с зашумлением до 20%.

К настоящему времени, однако, школой акад. *Микаэляна* предложена модель «параметрической» НСХ, реализующей новый тип ассоциативной памяти, который существенно превосходит известные нейросетевые модели по объему памяти и обладает способностью распознавать образы даже при исключительно больших искажениях и наличии корреляции [20].

Применение ИНС в экспериментальной физике

В табл. 1 приведены некоторые из конкретных задач, где замечательные свойства искусственных нейронных сетей проявляются особенно плодотворно и позволяют получать приемлемое решение в случаях, когда общепринятые аналитические и вычислительные методы оказываются малопродуктивными.

Далее будут приведены примеры таких применений из практики Объединенного института ядерных исследований в Дубне.

Таблица 1. Возможности применения искусственных нейронных сетей в экспериментальной физике

С		
,		РБФ-аппроксимация многомерных функций
		Решение нелинейных и дифференциальных уравнений
		Реализация метода гл. элементов с помощью рециркулярных ИНС
		Решение статистических гипотез
		Разбиение на классы, экспертные системы
		Распознавание образов в ФВЭ и биологии
()		Сегментная модель треков
		Роторная модель треков (ионизационная задача)
		Эластичные ИНС (деформируемые шаблоны)
		Распознавание букв и цифр
		NP-задачи: матчинг треков (расстановка ладей)
		Имитационное моделирование физических процессов
		Генерация случайных векторов
		Устранение шумовых отсчетов и заполнение пропусков на изображениях, распознавание треков

Применения прямоточных ИНС

Сравнительная простота структуры ИНС типа многослойного персептрона и простота их применения проявились особенно при решении физических задач, где обычно квалификации дипломированного физика-экспериментатора вполне хватает для использования программной или схемной реализации ИНС. Кроме того, благодаря распространению монте-карловских пакетов программ типа GEANT [21], позволяющих моделировать физические процессы в детекторах частиц в условиях, максимально приближенных к реальным, не возникает проблем с моделированием необходимой последовательности данных, требуемых для обучения сети. Возможно именно поэтому разработка первых универсальных пакетов для генерации МСП по заданному числу слоев и нейронов в них, позволяющих также реализовать один из методов обучения сети была осуществлена еще в начале 90-х в физических институтах, как например, фортранный нейропакет JETNET, разработанный физиками Лундского университета [22]. По этим же причинам нескольким электронным фирмам еще в то время удалось осуществить аппаратные реализации ИНС широкого применения в виде интегральных чипов, работающих параллельно и допускающих настройку сети на заранее промоделированную конфигурацию (см. например, [23] или обзор в [24]). Значительное количество статей по применению различных нейрочипов (ETANN, TOTEM, ZISC036) и даже специального алгоритмического языка SNAPs для управления ими можно найти в специальном выпуске журнала Nuclear Instruments and Methods in Physics Research [25].

В настоящее время в связи со всеобщим переходом в экспериментальной физике к объектно-ориентированному программированию на языке C++ повсеместно используется программная система ROOT, где есть соответствующий нейропакет, реализующий многослойные персептроны [26].

ИНС для идентификации частиц по данным детектора RICH. Детектор черенковского излучения RICH [27] фиксирует прохождение частиц на матрице из десятков тысяч фотоэлементов, сработавшие ячейки которой (отсчеты) образуют характерные кольца (см. рис. 3). Радиусы этих колец зависят от импульсов частиц, что при известной их массе позволяет провести их идентификацию. Таким образом, при обработке данных RICH, состоящих из набора в несколько сот координат отсчетов, зарегистрированных в плоскости фотодетектора, требуется выполнить ряд последовательных шагов:

- распознать кольца, т. е. расклассифицировать отсчеты по их принадлежности к кольцам (или к шумовому фону);
- учесть оптические искажения детектора, приводящие к эллиптической (а не круговой) форме колец;
- привязать найденные кольца к трекам тех частиц, которые интересуют экспериментаторов, удалив при этом ложно найденные кольца, образованные случайными комбинациями отсчетов;
- выбрать наиболее значимые характеристики полученных колец и по ним провести идентификацию частиц, расклассифицировав их по величине приведенного радиуса.

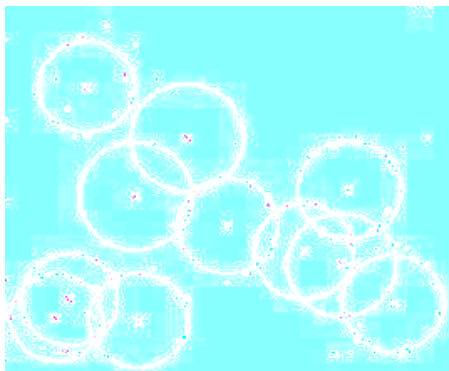


Рис. 3. Вид отсчетов на части фотодетектора. Среди распознанных видны и неправильно найденные кольца

Методы решения всех этих задач выходят далеко за рамки данной статьи, мы же остановимся на двух из них, относящихся к различению гипотез, где эффективное решение было найдено, благодаря применению ИНС.

Удаление ложных колец. Главным здесь оказалось найти наиболее значимые характеристики колец, которые можно было бы использовать для отсева ложных колец. Такие параметры должны как можно больше отличаться для ложных и правильно найденных окружностей.

После тщательного статистического анализа были отобраны 10 таких параметров:

- 1) количество отчетов в найденном кольце (правильно найденные электронные кольца имеют в среднем больше отчетов, чем ложные кольца);
- 2) расстояние от центра кольца до ближайшей проекции трека (большинство ложных колец расположено относительно далеко от треков);
- 3) сумма трех наибольших углов между соседними отчетами (типичной особенностью ложных колец является неравномерное распределение отчетов на кольце, что измеряется вычислением углов между соседними отчетами и выбором трех максимальных из них);
- 4) сумма числа отчетов, которые лежат в узком коридоре вокруг радиуса;
- 5) радиальная позиция на плоскости фотодетектора (большое количество колец находится на внутренней части фотодетектора из-за большой плотности отчетов и колец в этой области);
- 6) критерий χ^2 эллиптической подгонки кольца;
- 7) большая полуось эллипса (A);
- 8) малая полуось эллипса (B);
- 9) угол поворота эллипса φ относительно оси абсцисс;
- 10) азимутальный угол, который определяет позицию эллипса на плоскости фотодетектора (статистическое исследование показало наличие значимой зависимости этих углов для правильных колец и ее отсутствие для ложных колец).

Как показало предыдущее исследование [27] по применению пороговых процедур для удаления ложных колец, как для одномерных, так и попарных двумерных распределений выбранных параметров, не позволяет достичь приемлемых результатов при сохранении эффективности распознавания. Поэтому для задачи классификации колец была применена ИНС типа многослойного персептрона (МСП). Использовался трехслойный МСП из программного пакета ROOT [26]. Количество входных нейронов в первом слое соответствует количеству описанных выше параметров. Количество нейронов в скрытом слое на основе проведенного исследования выбирается, как удвоенное число входных нейронов, а выходной нейрон — один.

Для обучения нейронной сети использовалась обучающая выборка из 3000 ложных и 3000 правильных колец, сгенерированная методом Монте-Карло с помощью программы [21], в которой полагалось, что для ложных колец выходной сигнал равен -1 , а для правильно найденных $+1$. При

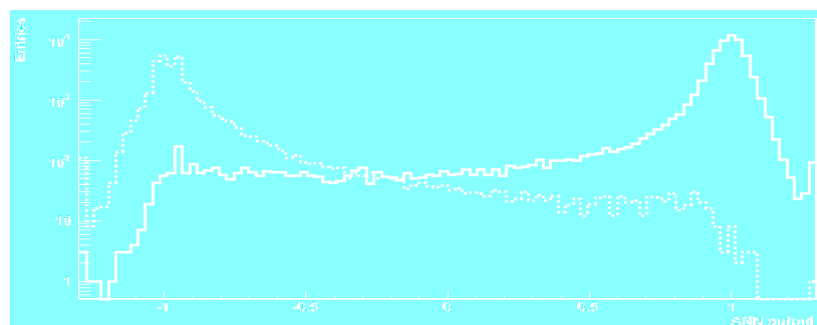


Рис. 4. Выходные значения ИНС для ложных (пунктирная линия) и правильно найденных (сплошная линия) колец (логарифмическая шкала)

применении нейросети после обучения в качестве активационной функции брался гиперболический тангенс, так что значения ее выходного нейрона уже не были бинарными, а принимали различные значения (см. рис. 4): для ложных колец эти значения лежат в районе -1 , для правильных — в районе $+1$. Порог для принятия решение по выходу МСП был выбран так, чтобы он обеспечивал максимальный процент подавления ложных колец, но при важном ограничении: выбранный порог должен обеспечивать заданную величину общей эффективности обнаружения колец. Соответствующее значение порога -0.5 для 97.5% эффективности распознавания было получено по сравнению эмпирических функций распределения ложных и правильно найденных колец.

Идентификация первичных электронов. Также применялся трехслойный МСП из программного пакета ROOT [26]. Поскольку основной задачей в детекторе RICH является идентификация электронов, и подавление конкурирующих фоновых частиц-пионов, одним из входных параметров МСП был выбран импульс частицы, так как он более всего отличается для этих двух типов частиц (радиус пионных колец зависит от импульса, а для электронных колец он почти постоянный). После изучения распределений различных параметров для пионов и первичных электронов, мы добавили к импульсу еще последние 6 параметров из списка предыдущего раздела.

При обучении нейронной сети использовалась монте-карловская выборка из 3000 пионов и 3000 первичных электронов. Порог для получения значений выходного нейрона ИНС был выбран по результатам вычисления эффективности классификации и обеспечил вполне приемлемые вероятности ошибок первого и второго рода (0.018 и $0.4 \cdot 10^{-4}$, соответственно).

Классификация белковых структур. Еще более изощренной предварительной обработке пришлось подвергнуть данные, полученные электрофорезом белков пшеницы в актуальной задаче автоматического определения ее сортности. Решение данной задачи включает в себя получение электрофоретических спектров белков посевного материала с последующим распознаванием по спектрам сортовой принадлежности данного образца. Такая задача является многокритериальной и ее решение экспертом довольно затруднительно в силу большого количества существующих сортов культур и перекрытия спектров различных экземпляров культуры в пределах одного сорта, а также между сортами. Поэтому ее решение целесообразно автоматизировать. В настоящее время показана высокая эффективность ИНС как экспертной системы при расшифровке генетических формул глиадинов (спирторастворимых белков) твердых сортов пшеницы и определения сорта по его электрофоретическому спектру [28].

Пример такого спектра из 17 дорожек, на которых получены спектры разных сортов показан на рис. 4. Сканирование одной из дорожек такого спектра (рис. 5) давало около 4000 отсчетов, показывающих изменение интенсивности, которое отражает генетическую структуру данного белка и является чем-то вроде его «отпечатка пальца». Задача состояла в том, чтобы сжать эту оцифровку на пару порядков без потери значимой информации, чтобы можно было ввести ее в ИНС. При этом были разработаны специальные алгоритмы предварительной обработки для сглаживания полученных денситограмм и вычитания фоновой подложки, а также устранения таких дефектов оцифровки, как зашумленность сигнала, неравномерная интенсивность цвета дорожек, отсутствие четких границ начала и конца спектров и нелинейное смещение дорожек относительно друг друга из-за нестационарности электрофоретических процессов. Для устранения последнего дефекта использовалась нейронная сеть Хемминга.

Как показал опыт экспертов, умеющих классифицировать сорта по спектрам, ключевой является информация не столько о самих пиках, сколько о порядке их чередования. Поэтому при распознавании значимых пиков

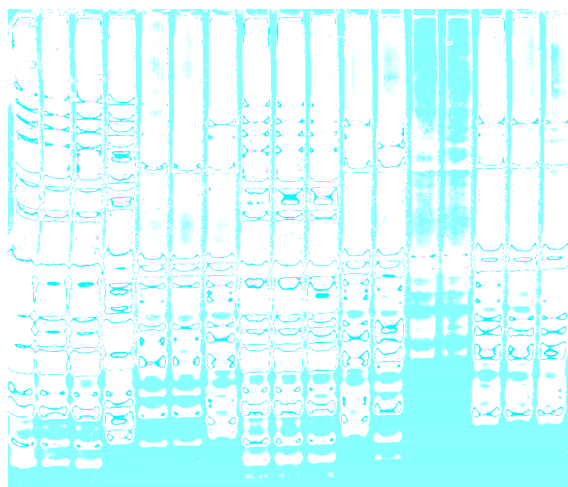


Рис. 5. 17-тидорожечная электрофореграмма

данные об их амплитудах и положении использовались для позонного ранжирования денситограмм, при котором весь спектр разбивается на определенное количество зон и для каждой зоны вычисляется свое значение — ранг по амплитуде пика, попавшего в эту зону. Проверка этой методики на обучающей выборке, сформированной на основе экспертной классификации более чем 100 электрофореграмм для каждого из пяти сортов пшеницы, показала, что после обучения трехслойный перцептрон с 45 скрытыми нейронами позволяет правильно распознавать с 95%-й эффективностью.

В настоящее время эта методика проходит апробацию на большем количестве сортов и обучающих выборках большего размера.

Диагностические экспертные системы. Ситуация, когда объем обучающей выборки ограничен небольшим экспериментальным материалом, особенно характерна для применения ИНС в медицине при разработке диагностических и прогностических экспертных систем. Трудности усугубляются тем, что количество диагностических параметров, которые должны быть введены в нейронную сеть, может достигать нескольких десятков, что увеличивает размерность ИНС и, в свою очередь, требует увеличения

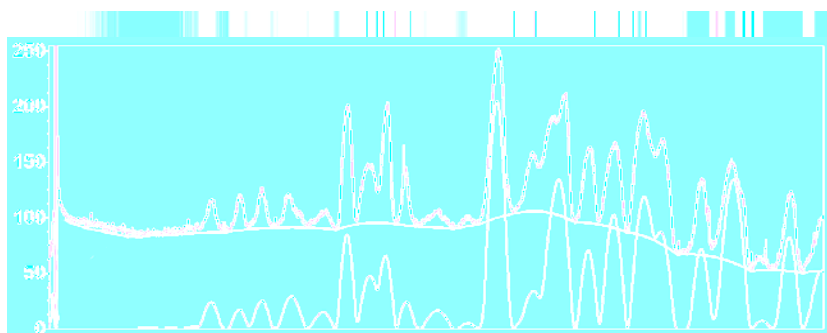


Рис. 6. Пример одной из дорожек электрофоретического спектра (вверху) и его график после оцифровки, сглаживания и устранения фоновой подложки

размеров обучающих выборок. Тем не менее, статистический анализ показывает, что эти диагностические параметры по большей части оказываются сильно скоррелированы [29]. Это дает возможность значительно сократить их число с помощью метода главных компонент (МГК) [30]. Обычно используется следующая схема:

1. Рассчитывается матрица ковариаций для исходной выборки.
2. Производится вращение матрицы, используя, скажем, метод Якоби, с целью ее диагонализации. Таким образом, мы переходим в новый базис из собственных значений матрицы, в котором данные уже не коррелируют между собой, так как мы произвели ортогонализацию входных параметров.
3. Кардинально уменьшается размерность базиса за счет отбрасывания большего числа компонент, малозначимых с точки зрения суммы их диагональных элементов, которая обычно не должна превышать 10% от общей суммы, называемой дисперсией системы.

Оставшихся главных компонент, сумма дисперсий которых составляет 90% дисперсии системы, достаточно для того, чтобы корректно обучить сеть. Благодаря тому, что преобразование к меньшему числу главных компонент является линейным, его можно «встроить» в саму сеть, добавив к ней после вычисления коэффициентов переходного преобразования дополнительный слой нейронов, осуществляющий переход от старого к новому

базису. Следует заметить, однако, что практическое применение МГК требует значительных предварительных расчетов, необходимых для реализации вышеприведенной схемы, причем эти вычисления приходится повторять всякий раз при изменении структуры сети (например, из-за введения нового диагностического признака). Поэтому в работе [29] был предложен менее затратный нейросетевой способ реализации МГК, в котором за основу был взят метод сжатия изображений с помощью *рециркуляционной* (или *автоассоциативной*) нейронной сети [31]. В такой сети, являющейся по структуре обычным трехслойным персептроном, изображение, подаваемое на вход, подвергается сжатию (выход скрытого слоя и есть сжатое изображение), а затем восстанавливается (выход сети). Сжатие изображения эквивалентно получению главных компонент, так что их количество должно быть равно числу нейронов скрытого слоя, которое должно быть установлено пользователем до обучения. Поэтому необходимо заставить сеть в ходе ее обучения самой устанавливать минимально необходимое количество нейронов скрытого слоя на основе некоего критерия оценки эффективности работы этого метода. Очевидно, что лучшим критерием будет точность распознавания предъявляемых образов.



Рис. 7. Средняя погрешность обучения как функция от числа скрытых нейронов

На графике рис. 7 видно, что при увеличении числа нейронов скрытого слоя погрешность сети вначале убывает довольно быстро, а затем измене-

ние стабилизируется на уровне, который можно считать оптимальным по числу скрытых нейронов. Сравнение с аналогичным графиком для МГК, где отображены процентные вклады дисперсий каждого из факторов в дисперсию системы показывает, что когда при увеличении количества нейронов погрешность сети изменяется мало, нет необходимости в дальнейшем увеличении количества нейронов, т. е. достигнут оптимум.

Таким образом, мы понизили размерность входных данных не алгебраическими методами, а используя саму сеть. Однако, чтобы построить зависимость погрешности от числа нейронов, мы использовали простой перебор сетей с разным числом нейронов в скрытом слое, хотя желательно делать подбор количества нейронов прямо в ходе обучения сети. В [29] для подбора оптимальной структуры сети предлагается использовать *динамическое добавление нейронов* в сеть, которое состоит во включении нейронов в действующую сеть во время ее обучения. Сеть начинает обучение с количеством нейронов, заведомо недостаточным для решения задачи. В ходе обучения мы добавляем дополнительный скрытый нейрон в том случае, когда относительная скорость изменения погрешности становится меньше заданной величины. Обучение сети прекращается, когда выполняется какое-либо из условий остановки, определяемое пользователем, например, достижение критического уровня погрешности или заданной эпохи. Дальнейшее развитие такой динамической оптимизации структуры ИНС путем еще и рационального удаления мало используемых нейронов и соответствующих им связей (pruning) позволило авторам работы [29] разработать нейросетевую медицинскую экспертную систему, сократив размеры исходной ИНС с 53/30/1 (число нейронов во входном/скрытом /выходном слоях) до 15/22/1, что, среди прочего, позволило с01973(о)-5.89017(й)-272.015(И)-0.69988(Н)-0.69988(С)-Д7883(е)3я

как выходные значения сети:

$$y_j = \sum_{k=1}^M w_k g(\rho(\vec{X}_j, C_k)), \quad (14)$$

где w_k — веса связи единственного нейрона (аппроксиматора) выходного слоя и k -го нейрона скрытого слоя с центром в точке C_k (веса связей между входным и скрытым слоем берутся все равными 1), функция активации g задана как (5), а евклидово расстояние ρ определяется формулой (6). В матричной форме уравнение (14) выглядит так: $\Phi w = Y$, где матрица Φ состоит из элементов $\phi_{kj} = g(\rho(\vec{X}_j, C_k))$, а Y и w , соответственно, есть векторы-столбцы выходов y_j и весов w_k . Это позволяет легко вычислить искомый вектор весов аппроксиматора безо всякого обучения простым обращением матрицы Φ : $w = \Phi^{-1}Y$, что обеспечивает прохождение интерполяционной поверхности через заданные M точек в пространстве входных образов, чтобы затем с помощью полученных весов вычислять значения функции F в произвольной точке, лежащей *внутри* области измерений.

Аналогичным образом можно решить задачу нейроаппроксимации и с помощью многослойного персептрона, однако для обеспечения достаточно высокой точности аппроксимации потребуется очень большая обучающая выборка и большое число эпох обучения, так что время на обучение МСП будет на порядок больше, чем для РБФ-сети.

Следует однако иметь в виду, что РБФ-сеть очень чувствительна к выбору числа скрытых нейронов, размещению их центров, а также к ширине окон активационной функции σ . Кроме того, сети РБФ не могут экстраполировать свои выводы за область известных данных, так как при удалении от обучающего множества значение активационной функции быстро спадает до нуля. Напротив, сеть MLP выдает более определенные решения при обработке сильно отклоняющихся данных.

К этому же кругу аппроксимационных задач относится и активно развивающийся нейросетевой подход к решению сложных нелинейных и дифференциальных уравнений со сложными граничными условиями [7, 8]. Поясним идею подхода на примере поиска решения $u = u(x, y)$ двумерного уравнения Лапласа $\Delta u = 0$ в области L с граничными условиями Дирихле: $u|_{\Gamma_k} = f_k$. Предлагается искать приближенное решение задачи

как выход нейронной сети:

$$u(x, y) = \sum_{i=1}^N c_i \varphi_i(x, y; a_i), \quad (15)$$

обучение которой состоит в процессе минимизации некоторого функционала ошибки $J(u)$. Преимущество этого подхода состоит в том, что радиально-базисная сеть позволяет свести вычисление производных функции $u = u(x, y)$ и ее граничных условий к значительно более простым операциям с аппроксимирующими ее РБФ φ_i . Функционал ошибки $J(u)$ берется как сумма квадратов невязок, получаемых при подстановке разложения (15) и производных $u(x, y)$ в уравнение $\Delta u = 0$ и в граничные условия $u|_{\Gamma_k} = f_k$:

$$J(u) = \sum_{j=1}^M |\Delta u(x_j, y_j)|^2 + \sum_k \delta_k \cdot \sum_{j_k=1}^{m_k} |u(x_{j_k}, y_{j_k}) - f_k(x_{j_k}, y_{j_k})|^2. \quad (16)$$

Первое слагаемое в (16) отвечает за удовлетворение уравнению, следующие слагаемые, входящие со штрафными множителями $\delta_k > 0$ (обычно их берут достаточно большими), отвечают за удовлетворение краевым условиям на соответствующих участках границы. Отличие такого подхода от классических методов типа Галеркина состоит в том, что можно оптимально подбирать не только весовые параметры c_i , но и сами базисные функции φ_i и входящие в них нелинейно параметры a_i , определяющие ширину окон РБФ. Условия проверяются на множествах контрольных точек: внутри области

$$L - \{(x_j, y_j)\}_{j=1}^M$$

на частях Γ_k границы

$$\{(x_{j_k}, y_{j_k})\}_{j_k=1}^{m_k}, \quad k = 1, \dots$$

Эти тестовые множества обновляются после завершения каждого этапа обучения (прохождения определенного числа шагов в процессе минимизации функционала). Как указано в [9], в процессе обучения сети необходимо производить регенерацию тестовых точек, чтобы избежать переобучения сети и сделать этот процесс более устойчивым.

Применения полносвязных ИНС Хопфилда (НСХ)

Как уже отмечалось выше, работы Хопфилда [17, 18] вызвали всплеск интереса к применению полносвязных сетей, особенно в экспериментальной физике, где способность таких сетей, эволюционируя, приходить в устойчивое состояние, удалось использовать для решения задач распознавания траекторий заряженных частиц (треков) по данным их координатных измерений в современных электронных детекторах физики высоких энергий. Проблема эта является весьма сложной в силу большой множественности событий, число треков в которых составляет многие сотни, как это изображено на рис. 8.

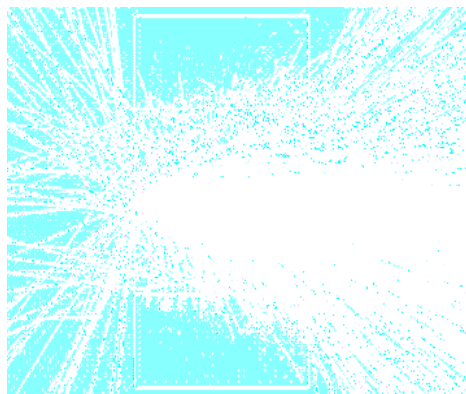


Рис. 8. Модель центрального столкновения Au+Au в CBM-эксперименте. Нейрон s_{ij} — направленный сегмент, соединяющий точки i, j

Задача распознавания треков состоит в классификации всех измеренных точек на группы по их принадлежности к трекам или к группе шумовых измерений. С математической точки зрения трек — это гладкая кривая, проходящая через точки пересечения траектории частицы с координатными плоскостями, так что правила классификации должны учитывать факт близости измерений к неким гладким кривым, число которых неизвестно заранее. Попытка просто ввести нейронную сеть, принимая за нейроны измеренные точки, приводит к чрезмерно большому числу нейронов и еще большему числу межнейронных связей, подавляющее большинство кото-

рых не относится к трекам. В итоге, любая программная реализация сети оказывалась слишком громоздкой и работала настолько медленно, что становилась бесполезной в практических приложениях. Для борьбы с этим «проклятием размерности» весовые функции должны вводиться так, чтобы поощрять связи нейронов, принадлежащих одному и тому же треку и, кроме того, в функционал энергии необходимо вводить «ограничительные» члены, запрещающие как межтрековые связи, так и чрезмерный рост числа самих треков.

Сегментная модель НСХ. Первую удачную попытку использовать нейронные сети для распознавания треков сделали Петерсон [32] и Денби [33]. Их подход, известный как *метод сегментов*, состоял в том, что на множестве экспериментальных точек на плоскости вводились нейроны s_{ij} , определяющие, соединяются точки i и j или нет, т. е. принадлежит данный направленный сегмент s_{ij} треку или нет (см. рис. 9). В дальнейших вычислениях в соответствии со схемой (4) состояния нейронов v_{ij} перестают быть целочисленными. Их значения $v_{ij} = \langle s_{ij} \rangle_T$ и определяют уровень активности нейрона. Энергетический функционал в работах [32, 33] был определен как состоящий из двух частей:

$$E = E_{cost} + E_{constraint} = -\frac{1}{2} \sum_{ijkl} \delta_{jk} \frac{\cos^m \theta_{ijl}}{r_{ij} r_{jl}} - \frac{\alpha}{2} \left[\sum_{ik} s_{ik} s_{kl} + \sum_{jl} s_{ij} s_{jl} \right] + \frac{\beta}{2} \left[\sum_{ij} s_{ij} - N \right]^2.$$

Предполагалось, что распознавание велось для гладких прямых треков без разветвлений. Поэтому первый стоимостной член выбирался так, чтобы он поощрял *короткие смежные сегменты с малым углом между ними*, а второй член (штрафной) устанавливал *запрет разветвлений* (т. е. запрет ситуаций, когда к одной точке присоединено больше одного сегмента-нейрона) и *баланс между числом активных нейронов и числом экспериментальных точек*.

К сожалению, эта схема Денби-Петерсона, принципиально запрещающая бифуркацию трека, была неприменима в случаях распадов нейтральных и рождения заряженных частиц в объеме детектора. Поэтому в работе [34], в которой исследовались данные распада нейтральных каонов и гиперонов, пришлось модифицировать штрафной член энергетической функции

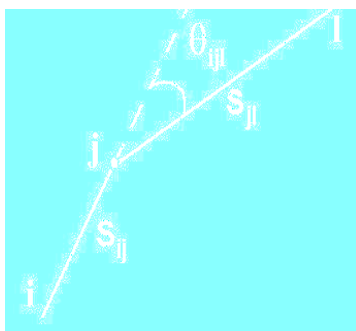


Рис. 9. Метод сегментов

так, чтобы он позволял осуществлять распознавание разветвлений. Пример работы сети дан на рис. 10, где реальные треки показаны сплошной линией, а ложные связи отмечены пунктиром.

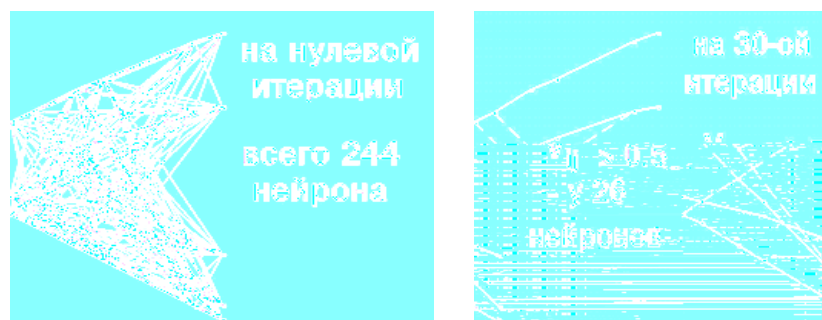


Рис. 10. Пример применения полносвязной ИНС для распознавания событий с короткоживущими частицами

В тяжелых фоновых условиях и экспериментальной неэффективности камер действующая программа обработки эксперимента EXCHARM не распознала 5.2% событий, а программа, основанная на ИНС, 3.2%. При этом из нераспознанных событий для разных программ совпали только 0.8% от общего числа событий, т. е. эти два множества плохо распознанных событий

имеют малую область пересечения. Таким образом возникла идея эффективного комбинированного алгоритма восстановления треков при незначительном увеличении временных затрат на обработку. Идея заключается в следующем: после распознавания X -проекций и Y -проекций треков более быстрым методом опорных дорожек в программе обработки производится пространственная «сшивка» этих проекций.

Если часть надежно найденных проекций не имеет сшивок, то включается нейросетевой алгоритм, который в большинстве случаев успешно решает задачу распознавания треков в таких событиях. Комбинированный алгоритм с использованием ИНС был опробован и на реальных данных эксперимента ЕХСНАРМ, где позволил достичь 99% эффективности распознавания событий.

Роторная модель НСХ. Идея роторных ИНС, приведенная *К. Петерсоном* в работе [35] определяет нейрон как единичный вектор, выходящий из i -й точки, измеренной на треке (см. рис. 11). Энергетическая функция роторной сети выглядит следующим образом:

$$E = -\frac{1}{2} \sum_{ij} \frac{1}{|r_{ij}|^m} v_i v_j - \frac{1}{2} \alpha \sum_{ij} \frac{1}{|r_{ij}|^m} (v_i r_{ij})^2.$$

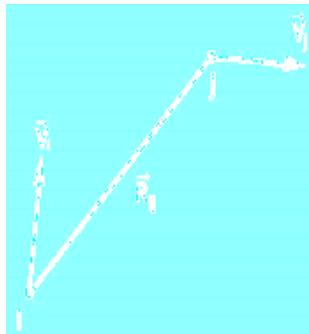


Рис. 11. Роторная модель ИНС

Первый его член сближает роторы по направлению, а второй — выстраивает их вдоль трека. В исходном состоянии (рис. 12а) направления

роторов генерировалось случайно, но после нескольких десятков итераций эволюции сети они обычно выстраивались по касательным к треку, как показано на рис. 12b.

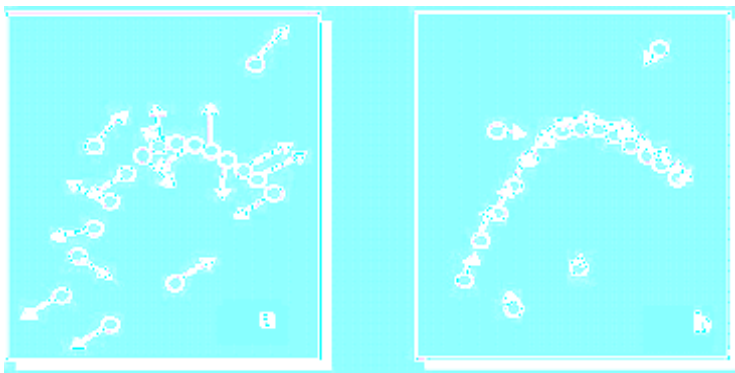


Рис. 12. Начальное (a) и конечное (b) состояния эволюции роторной ИНС

Однако в таком виде эта идея не нашла применения из-за слишком медленной сходимости процесса эволюции и явной тенденции к спрямлению треков с большой кривизной. Чтобы реально применить роторные ИНС для распознавания треков в магнитном поле, потребовалась ввести совсем другую энергетическую функцию, а также существенно повысить эффективность работы ИНС за счет следующих дополнительных усилий:

- префилтрации данных с помощью клеточного автомата;
- удачной процедуре начальной установки роторов;
- введения робастных множителей к синаптическим весам для «оттачивания» от соседних треков.

Этот подход был с успехом применен для распознавания треков по данным, полученным на магнитном спектрометре ARES [36], а также для обработки данных вертикального зондирования ионосферы [37]. Мы использовали замечательное свойство окружности согласно которому касательные к ней в двух точках, соединенных хордой, становятся параллельными, если одну из этих касательных зеркально отразить относительно хорды, как это показано на рис. 13. Это свойство позволяет выразить отраженный ротор-

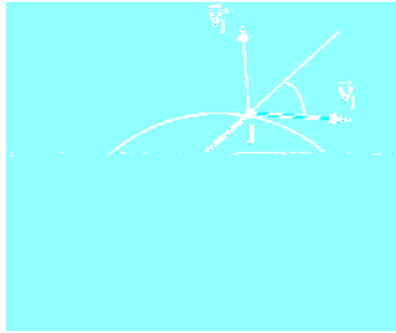


Рис. 13. Роторная модель для треков в магнитном поле

нейрон как

$$v'_j = \begin{bmatrix} \cos 2\phi_{ij} & \sin 2\phi_{ij} \\ \sin 2\phi_{ij} & -\cos 2\phi_{ij} \end{bmatrix} \cdot v_j = W_{ij}v_j,$$

что приводит к следующему простому виду энергетической функции сети безо всяких ограничивающих членов (6)

$$E = -\frac{1}{2} \sum_{ij} v_i \cdot v'_j.$$

Анализ ионограмм. Данные, полученные в результате вертикального зондирования ионосферы, представляют собой последовательность отсчетов частоты сигнала, его амплитуды и времени прихода после отражения от ионосферного слоя [38]. Типичная ионограмма вертикального зондирования показана на рис.14а, где по оси абсцисс дана измеренная частота сигнала, по оси ординат — высота слоя. Расплывчатые кривые, сформированные отсчетами, по аналогии с ФВЭ называются треками. Ионосферные треки могут пересекаться и искажены вертикальными помехами. Если учесть еще и амплитуды приходящих отсчетов, то получим трехмерную картинку, изображенную на рис. 14b, напоминающую горный пейзаж.

Задача состоит в том чтобы с минимальными временными затратами без использования априорной информации распознать ионосферные треки и аппроксимировать их по точкам наибольшей амплитуды, т.е., по альпинистской терминологии, зафиксировать траверсы всех «хребтов». Эта

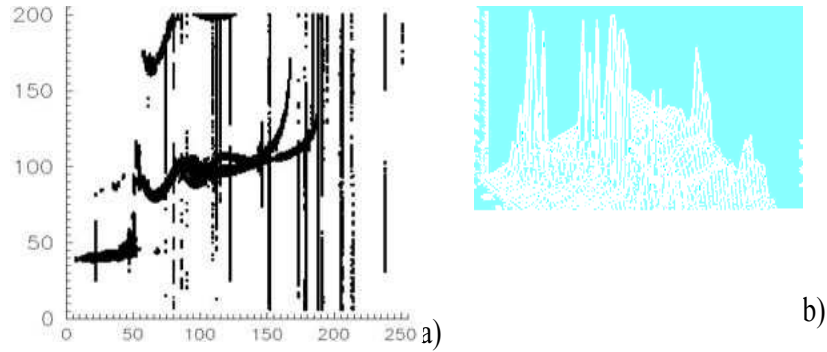


Рис. 14. Ионограмма вертикального зондирования (а) и трехмерное изображение той же ионограммы (б)

задача была успешно решена в [37] методом модифицированных роторных НСХ с предварительной обработкой данных с помощью клеточных автоматов. Основной целью этой предобработки было устранение вертикальных помех и заполнение пропущенных отсчетов. После этого для каждого частотного кластера находилась наибольшая амплитуда. После предобработки, сокращавшей объем данных в 5–10 раз, данные поступали в роторную ИНС, которую пришлось модифицировать, чтобы справиться в переменной кривизной ионотреков. Модификация состояла в отказе от глобальной аппроксимации трека окружностью и введении скользящего окна просмотра, размер которого выбирался так, чтобы в его пределах локальная аппроксимация окружностью выполнялась с заданной точностью. Идея скользящего окна упростила и ускорила также и процедуру определения начальной конфигурации сети путем углового гистограммирования (см. [37]), ограничив ее пределами текущего окна. Таким образом, инициализация по сжатым данным выполнялась следующим образом: каждая точка помещалась в середину левой стороны окна, разбитого на угловые сектора в 1° . Подсчитывалось число точек в пределах окна, попавших в каждый сектор. Центр тяжести трех смежных ячеек с максимумом этой гистограммы в середине и брался в качестве направления ротора, а за модуль принималось среднее от числа точек в этих трех ячейках.

Динамика среднего поля для эволюции ротора определялась формулой



Рис. 15. Этапы обработки ионограммы: **(а)** клеточный автомат убрал вертикальные помехи и заполнил пропущенные отсчеты; **(б)** инициализация роторов путем углового гистограммирования в скользящем окне переменного размера в зависимости от изменения кривизны трека; **(с)** результат распознавания

(11) с $T = 1.5$. Процедура сходилась в среднем за 3–5 итераций, что подтвердило ее эффективность, а также оправдало значительные временные затраты на инициализацию роторов, занимавшую иногда до 90% времени обработки. Основные этапы обработки ионограммы отображены на рис. 15.

В дальнейшем этот метод и реализующая его программы активно использовались в реальных экспериментальных исследованиях ионосферы в России и США [38, 39].

Управляемые нейронные сети. Решение широкого класса комбинаторных задач, объем вычислений которых растет как экспонента с размерностью задачи, оказалось возможным выполнять в разумное время с помощью

особого типа нейронных сетей, названных в работе [40] управляемыми ИНС. Одной из классических проблем комбинаторной оптимизации, относящаяся к классу NP-проблем, т. е. задач, не решаемых за число шагов, растущее как полином от n , является, например, задача о расстановке N ферзей на шахматной доске размером $N \times N$ так, чтобы они не били друг друга. Программы, основанные на обычных комбинаторных методах, уже при $N = 97$ не могут решить эту задачу за сколько-нибудь реальное время, в то время как управляемая ИНС, построенная на базе сети Хопфилда, позволила решить эту задачу для $N = 1024$ за 10 минут на компьютере 90-х годов. Идея управляемой ИНС (УИНС) заключается во введении в сеть дополнительных управляющих нейронов, предназначенных для управления эволюцией определенных групп нейронов с целью предотвратить «застревания» энергетической функции сети в локальном минимуме. Выбор групп управляемых нейронов зависит от специфики задачи, в частности, в задаче расстановки ферзей в качестве группы выбирается одна из сторон шахматной доски.

В ФВЭ актуальная задача сшивки треков (track-matching), распознанных на разных проекциях, является именно такой комбинаторной NP-проблемой. В терминологии теории графов задача сшивки треков сводится к построению матрицы инцидентности с выбором ненулевого элемента на каждой ее строке так, чтобы все эти элементы лежали в разных столбцах. На шахматном языке эта задача означает, что мы должны на шахматной доске размера $N \times N$ поставить более чем N ладей, не оставив ни одной пустой строки, а потом снять «лишние» ладьи так, чтобы на каждой строке оставалось ровно по одной ладье и ни одна не попадала бы под удар другой. Описание деталей конфигурации соответствующей УИНС и определения весовой матрицы, обеспечивающей функционирование нейрон-контролеров чересчур громоздки для изложения в рамках данной статьи. Их можно найти в [40]. В табл. 2 приведен результат сравнительного прогона программы, реализующей УИНС и программы, основанной на быстром стандартном алгоритме перебора (САП). Вычисления велись на старой ЭВМ 90-х годов VAX-8350.

Как видно, начиная с $N = 30$, УИНС продолжает показывать разумное время, тогда как применение стандартного алгоритма становится нецелесообразным. Алгоритм УИНС может быть адаптирован для решения и других задач, например, составления расписаний больших размерностей.

Таблица 2. Время решения проблемы шивки треков в зависимости от их числа

N	30	15	13
УИНС	105.32 с	0.26 с	0.13 с
САП	7.5 ч	23.78 с	0.135 с

Эластичные нейронные сети. В ходе многочисленных применений нейросетей в ФВЭ обнаружили такие их недостатки, отмеченные рядом исследователей (см. напр., [40, 41]), как:

- слишком медленная сходимость из-за излишнего числа степеней свободы;
- с помощью ИНС решается только проблема распознавания без учета известной модели трека;
- отмечается чрезмерная чувствительность ИНС к шумам.

В этой связи в работе [41] было предложено объединить этапы распознавания и фитирования кривых, максимально используя априорную информацию. С этой целью предлагалось с учетом известного уравнение трека, создать эластичный шаблон и изгибать его (меняя параметры уравнения) так, чтобы он прошел по «своим» измеренным точкам. Подобный подход немедленно порождает естественные вопросы:

- неизвестно число шаблонов;
- где взять начальные значения их параметров;
- как организовать подгонку сразу всех кривых (треков);
- где тут нейронная сеть?

Проблема нахождения шаблонов (templates) и грубых значений их параметров требует специального рассмотрения. Мы ограничимся здесь ссылкой на работы [41–44], в которых использовались эластичные шаблоны и отметим, что всюду для их поиска применяется тот или иной вариант преобразования Радона–Хафа [45]. При практической реализации это преобразование сводится к гистограммированию в пространстве параметров с последующим поиском максимального значения подобно тому, как это делалось в конце предыдущего раздела при поиске начальных значений роторов. В силу приближенного характера процедуры нахождения шаблонов,

их число могло превышать число реальных треков в событии. Появляющиеся из-за этого лишние, искусственные, треки должны быть удалены впоследствии в ходе специальной процедуры отбраковки.

Для поиска сразу всех треков, порожденных взаимодействием в однородном магнитном поле, в работе [41] предложено минимизировать следующий функционал:

$$E(\{S_{ia}\}; \vec{\pi}) = \sum_{i=1}^N \sum_{a=1}^M S_{ia} D_{ia}(\vec{x}, \vec{\pi}) + \lambda \sum_{i=1}^N \left(\sum_{a=1}^M S_{ia} - 1 \right)^2.$$

Здесь $\vec{\pi}$ — вектор параметров геликоиды, описывающей траекторию движения a -го трека; $D_{ia}(x, \vec{\pi})$ — квадрат расстояния от точки x до трека a ; S_{ia} — бинарный нейрон, определяющий принадлежность i -й точки к a -му шаблону трека:

$$\begin{aligned} S_{ia} &= 1, & i\text{-я точка принадлежит } a\text{-му треку,} \\ S_{ia} &= 0, & \text{в противном случае.} \end{aligned}$$

Минимизация $E(S_{ia}; \pi)$ ведется при условии, что каждая точка может принадлежать только одному треку или ни одному ($\sum_{ia} S_{ia} = 0$). В последнем случае функционал штрафует на величину λ , что определяет критическое расстояние $\sqrt{\lambda}$, до которого точки энергетически выгодно включать в трек, а после — считать такую точку шумовой ($S_{ia} = 0, \forall i$).

Дальнейшие вычисления по поиску глобального минимума ведутся в соответствии с обычной схемой сетей Хопфилда: сеть термализуется и применяется теория среднего поля. Для последовательности уменьшающихся температур $T_k > T_{k-1} > \dots > T_0$ метод наискорейшего спуска дает пошаговый итерационный алгоритм поиска минимума [41]:

$$\Delta \pi_a^{(k)} = -\eta_a^{(k)} \sum_i V_{ia} \frac{\partial D_{ia}}{\partial \pi_a^{(k)}},$$

где

$$V_{ia} = \frac{e^{-\beta D_{ia}}}{e^{-\beta \lambda} + \sum_{b=1}^M e^{-\beta D_{ib}}}$$

представляют собой факторы Поттса, $\beta = 1/T$.

Метод эластичных нейросетей был успешно применен в работах [42, 43] для обработки данных, параметризуемых уравнением окружности: колец черенковского излучения и треков в однородном магнитном поле.

В первой работе [43] методом эластичных нейросетей (ЭН) осуществлялся поиск черенковских колец с одновременной оценкой их параметров по данным RICH детектора CERES/NA-45 (см. пример на рис. 16 слева). Глобальный поиск велся безо всяких априорных сведений о центрах и радиусах колец. Эта информация была получена с помощью сведения трехмерного преобразования Хафа к двумерному и одномерному. Вначале выполнялся перебор допустимых триплетов измеренных точек (т. е. таких троек точек, через которые можно провести окружность с радиусом и центром, удовлетворяющим заданным неравенствам). Идея метода основана на том, что все точки, принадлежащие некоей окружности должны отображаться в одну точку в пространстве параметров, так что суммируя, мы должны получить пик в этом месте. Из-за наличия ошибок измерений этот пик несколько размазывается, что сказывается на размерах биннинга при гистограммировании получаемых параметров. Гистограммирование разумнее выполнить в два этапа: сначала строится двумерные гистограммы центров и находятся все пики в ней, превышающие заданный порог. Затем для каждого найденного центра выбираются точки внутри области допустимых значений радиуса и также гистограммируются. Максимальный из пиков гистограммы и принимается за оценку радиуса. Процедуры выбора порогов и тестирования полученных параметров достаточно сложны. В [43] можно найти полезные рекомендации по выбору размеров биннинга гистограмм и порогов для оценки параметров колец, полученные на основе исследования о влиянии на эффективность распознавания таких факторов, как радиальный разброс точек, общая зашумленность события и среднее число фотонов на кольцо. Об эффективности полученного алгоритма реализации метода эластичных нейросетей свидетельствует то, что он с минимальной модификацией был также успешно применен для распознавания треков в магнитном поле. Примеры работы алгоритма в этих двух режимах представлены на рис. 16. Отсев ложных колец, получившихся из-за того что преобразование Хафа нашло несколько лишних шаблонов (см. пример в левом нижнем углу рис. 16), мог быть легко выполнен в последующем анализе с использованием критерия χ^2 и учетом статистики по числу и угловому распределению фотонов в кольце.

Следует, однако, сделать важное замечание, касающееся издержек глобальности метода ЭН: при попытках увеличить число колец или множественности, т. е. числа треков, в районе поиска выше 12–13, метод переставал работать (не говоря уже о быстром росте машинного времени). Частично это происходило из-за ошибок в преобразовании Хафа, но главным

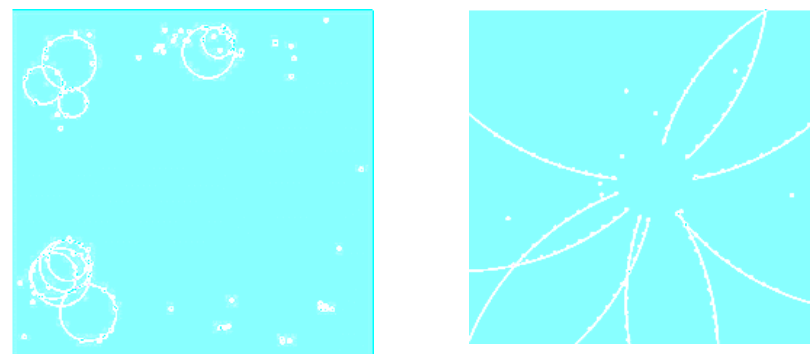


Рис. 16. Примеры применения метода ЭН из [42]: *слева* — распознавание черенковских колец; *справа* — распознавание треков в магнитном поле

образом из-за проблем с минимизацией функционала.

Поэтому во второй работе [44], касавшейся применения метода ЭН в задаче распознавания и определения параметров треков в системе дрейфовых трубок, мы отказались от глобального прослеживания сразу всех треков события и искали треки по данным Хаф-преобразования по-очереди. Каждая из газополных трубок, образующих плотные слои сотовой структуры (см. рис. 17) имеет центральную электродную проволоку, так что при прохождении частицы сквозь такую трубку регистрируется не только координата этой проволоки, но и время дрейфа до нее облака ионов, образующихся в газе при прохождении частицы. При известной скорости дрейфа это время пересчитывается в радиус дрейфа, позволяющий вычислить координату пролета частицы до микронных точностей. Таким образом данные об измеренном треке в дрейфовой камере выглядят как набор малых окружностей с центрами в зарегистрированных центральных проволоках и радиусами, равным соответствующим радиусам дрейфа. С точностью до ошибок измерения этих радиусов трек проходит по касательным ко всем этим маленьким окружностям (рис. 17). Главная неприятность здесь состоит в том, что знание радиуса дрейфа не позволяет определить, слева или справа от проволоки прошла частица, возникает известная *лево-право неопределенность*. Ситуация усугубляется как из-за ложных срабатываний некоторых из трубок, так и из-за их неэффективности, приводящей к пропускам некоторых

измерений.

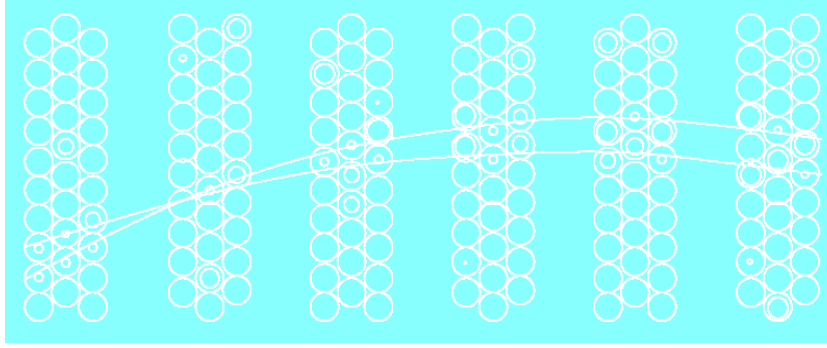


Рис. 17. Типичное событие, детектированное дрейфовой камерой в магнитном поле

Ограничимся для наглядности рассмотрением одной проекции XoZ и обозначим (x_i, z_i) — координаты центра сработавших трубок, r_i — радиус дрейфа i -го измерения. Считая трек окружностью с параметрами $(a, b; R)$, получаем для расстояния от центра i -ой трубки до трека следующее выражение:

$$D_i(a, b; R) = R - \sqrt{(x_i - a)^2 + (z_i - b)^2},$$

которое, отражая суть проблемы право-лево неопределенности, может быть как положительным, так и отрицательным.

Обозначим

$$\begin{aligned} d_i^- &= (D_i(a, b; R) - r_i)^2, & \text{если } D_i(a, b; R) > 0, \\ d_i^+ &= (D_i(a, b; R) + r_i)^2, & \text{в остальных случаях.} \end{aligned}$$

Из-за этой двойственности пришлось вместо бинарных нейронов S_{ia} , используемых для определения принадлежности точки треку, ввести двумерный вектор $S_i = (s_i^+, s_i^-)$ с допустимыми значениями $(1, 0)$, $(0, 1)$, $(0, 0)$. Функционал E зависит теперь от пяти параметров

$$E(a, b, R, s_i^+, s_i^-) = \sum_i \{d_i^- s_i^- + d_i^+ s_i^+ + \lambda[(s_i^+ + s_i^-) - 1]^2\}, \quad (17)$$

где λ — измерительная ошибка радиуса дрейфа. Выражение (17) минимизируется при ограничении $(s_i^+, s_i^-) \leq 1$. Факторов Поттса стало два

$$s_i^- = \frac{1}{1 + e^{(d_i^- - \lambda)/T} + e^{(d_i^- - d_i^+)/T}},$$

$$s_i^+ = \frac{1}{1 + e^{(d_i^+ - \lambda)/T} + e^{(d_i^+ - d_i^-)/T}}.$$

При трехступенчатой последовательности убывающих температур, используемых для реализации схемы имитационного отжига, процедура минимизации оказалась достаточно изощренной и потребовала последующей корректировки параметров (см. детали в [44]). Однако временные затраты на метод ЭН удалось компенсировать, благодаря организации гибридной структуры обработки, в которой этот метод используется только в редких случаях, когда не срабатывает более быстрая основная программа, основанная на робастном подходе.

Применение клеточных автоматов

Трактуя нейронные сети как дискретные динамические системы, можно и клеточные автоматы (КА) рассматривать как вариант такой простой нейронной сети, в которой нейроны-клетки принимают лишь конечное дискретное множество состояний (часто всего два), их связи локализованы ближайшими соседями, правила эволюции также жестко определены в зависимости от них и, наконец, сама эволюция происходит в тактовые моменты времени, синхронные для всех клеток автомата. Тем не менее эта кажущаяся простота структуры и динамики позволяет организовать самые удивительные и полезные применения КА [46].

Правила построения клеточного автомата для отсева шумовых отсчетов, заполнения пропусков и распознавания связанных групп точек в процессе эволюции КА были предложены в [47] как обобщение правил известной игры «Жизнь». В дальнейшем эти правила применялись с необходимыми модификациями для предварительного грубого распознавания сегментов треков с последующим более точным распознаванием треков или вершин событий путем применения ИНС. Эти модификации были необходимы для введения критерия близости клеток по направлению вдоль трека. Подобно тому, как это делалось в вышеописанной сегментной модели ИНС Денби-Петерсона, клетка определялась как сегмент, соединяющий два отсчета на

соседних координатных плоскостях (с учетом неэффективности камер — допускались соединения и со следующей камерой). Соседство устанавливалось по близости сегментов по направлению, т. е. запрещались большие углы (хотя возможность многократного рассеяния могла допускаться). При инициализации значения клеток с допустимыми наклонами устанавливались в 1, затем в процессе эволюции значение клетки увеличивалось на единицу, если у соседа на предыдущем слое было то же значение. Эволюция прекращалась, когда не оставалось клеток с одинаковыми значениями. После этого треки собирались, начиная с клетки с наибольшим значением, путем присоединения соседей с предыдущими значениями при движении назад вдоль дерева. Если появлялись лишние ветви, оставлялись более длинные и гладкие. На заключительной стадии проверялось качество треков для отсева ложных из них по критерию χ^2 и большему числу сегментов [48].

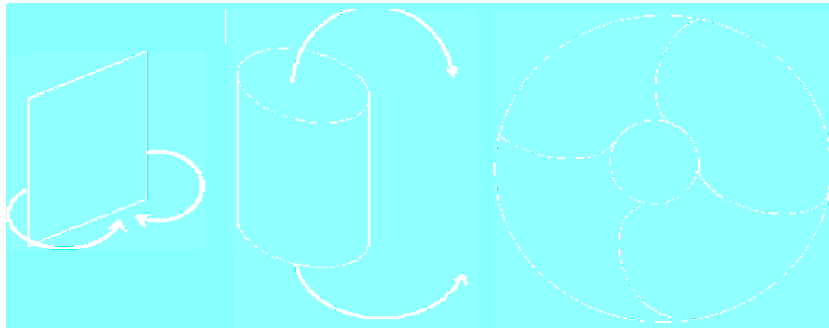


Рис. 18. Свертывание решетки в тор для выполнения граничных условий

В работе [49] удалось сконструировать двумерный клеточный автомат совершенно другой структуры для генерации случайных многомерных векторов, что необходимо при решении многих актуальных задач физического моделирования методом Монте-Карло. На решетке 32×500 строится двумерный клеточный автомат, каждая ячейка которого определяется по своим соседям по формуле:

$$\phi_{i,j} = \left(\sum_{i'=i-1}^{i+1} \sum_{j'=j-1}^{j+1} c_{i'j'} \right) \bmod 2. \quad (18)$$

Решетка сворачивается в тор для выполнения граничных условий (рис. 18).

Ячейки вначале случайно засеваются нулями и единицами. Потом на каждую эволюцию КА по формуле (17) получаем 500 32-значных случайных числа, образующих 500-мерный случайный вектор с равномерным распределением.

Заключение

Применения ИНС в практике ОИЯИ не исчерпывается вышеприведенными примерами. Более подробные сведения о них можно найти, например, в обзорах [50, 51]. Следует также подчеркнуть, что успех применения ИНС в прикладных задачах обеспечивался, как правило, благодаря следующей методике нейросетевых вычислений:

- 1) статистическая обработка данных (весьма ответственный этап предобработки) для отбора наиболее значимых параметров, с применением при необходимости метода главных компонент или других методов сокращения числа входных параметров ИНС;
- 2) выбор типа нейронной сети (прямоточная, с обратными связями) — в зависимости от решаемой задачи;
- 3) выбор структуры нейронной сети (число слоев, число нейронов в скрытых слоях);
- 4) выбор алгоритма обучения нейронной сети (метод обратного распространения ошибки, метод имитационного отжига);
- 5) обучение нейронной сети;
- 6) тестирование нейронной сети после обучения на специально выбранном тестовом множестве, не использованном при обучении;
- 7) оценка эффективности работы нейронной сети.

На последнем пункте стоит остановиться подробнее, так как именно на основе критериев эффективности и происходит оптимизация структуры и оценка работы ИНС. Обычно применяемый критерий эффективности в задачах распознавания как процентное отношение числа правильно распознанных образов к общему их числу в тестовой выборке, недостаточен при решении аппроксимационных задач, где не менее важна и точность нейросетевой аппроксимации. В задачах проверки гипотез вычисление вышеуказанной эффективности дает лишь вероятность ошибки первого рода и для оценки вероятности ошибки второго рода следует подсчитывать также и случаи принятия сетью альтернативного решения. Следует учитывать

и размерность применяемой ИНС, т. е. общее число нейронов или всех межнейронных связей (весов), а также временные затраты на ее обучение или эволюцию до прихода в минимум энергетической функции. Все это многообразие критериев можно объединить, например, в их взвешенной сумме с соответственно подобранными весами и использовать затем для оптимизации ИНС по совокупности этих характеристик.

Оптимизация структуры ИНС в несложных задачах может выполняться и простым перебором по числу скрытых нейронов, но в более сложных задачах применяют либо метод динамической оптимизации ИНС по ходу ее обучения, как это было описано выше, либо используют эволюционные методы оптимизации с помощью *генетических алгоритмов* [52].

Работа выполнена при частичной поддержке РФФИ (грант 06–04–08244).

Литература

1. Хехт-Нильсен Р. Нейрокомпьютинг: история, состояние, перспективы // *Открытые системы*. – 1998, № 4.
2. URL: <http://www.basegroup.ru/library/analysis/neural/classification/>
3. Галушкин А. И. Синтез многослойных систем распознавания образов. – М.: Энергия, 1974.
4. Горбань А. Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей // *Сиб. журнал выч. техн.* – 1998. – т. 1, № 1. – с. 12–24.
5. Фролов А. А., Муравьев И. П. Нейронные модели ассоциативной памяти. – М.: Наука, 1987. – 160 с.
6. Джарратано Д., Райли Г. Экспертные системы: принципы разработки и программирование. – М.: Вильямс, 2007. – 1152 с.
7. Тархов Д. А. Нейронные сети: модели и алгоритмы. – М.: Радиотехника, 2005. – 256 с.
8. Васильев А. Н., Тархов Д. А. Нейронные сети как новый универсальный подход к численному решению задач математической физики // *Нейрокомпьютеры: разработка, применение*. – 2004, № 7–8. – с. 111–118.
9. Уоссермен Ф. Нейрокомпьютерная техника. – М.: Мир, 1992.
10. Горбань А. Н. Обучение нейронных сетей. – М.: ParaGraph, 1990. – 160 с.
11. Каллан Р. Основные концепции нейронных сетей. – М.: Вильямс, 2001 – 288 с.

12. Хайкин С. Нейронные сети. Полный курс. – М.: Вильямс, 2006. – 1104 с.
13. Ясеницкий Л. Н. Введение в искусственный интеллект. – Академия, 2005. – 176 с.
14. URL: <http://www.statsoft.ru/home/textbook/default.htm>
См. также: Нейронные сети. Statistica Neural Networks. Методология и технологии современного анализа данных. – Горячая Линия–Телеком, 2008. – 392 с.
15. Ososkov G., Stadnik A. Effective training algorithms for RBF-neural networks // *Optical Memory & Neural Networks*. – v. 12, No. 3, 2003. – pp. 199–218.
16. Кохонен Т. Ассоциативные запоминающие устройства. – М.: Мир, 1982.
17. Hopfield J. Neural networks and physical systems with emergent collective computational abilities // *Proc. Nat. Acad. Sci. USA*. – 1982, Vol. 79. – pp. 2554–2558.
18. Hopfield J. Neurons with graded responses have collective computational properties like those of two-state neurons // *Proc. Nat. Acad. Sci. USA*. – Vol. 81, 1984. – p. 3088.
19. Kirkpatrick S. et al. Optimization by simulated annealing // *Science*. – **220**, 1983. – pp. 671–680.
20. Крыжановский Б. В., Микаэлян А. Л. Ассоциативная память, способная распознавать сильно скоррелированные образы // *Доклады АН. Информатика*. – 2003. – т. 390, № 1. – с. 27–31.
21. CERN Program Library GEANT Long Writeup W5013,
URL: <http://wwwasdoc.web.cern.ch/wwwasdoc/pdfdir/geant.pdf>
22. Peterson C. et al. JETNET 3.0 – a versatile artificial neural network package. – LU TP 93-29, 1993; CERN-TH 7135/94.
23. Lindsey C., Lindblad Th. Review of hardware neural networks: a user's perspective // *Intern. Journ. of Neural Systems*. – v. 6, Suppl., 1995.
24. Никитюк Н. М. Нейрочипы, нейрокомпьютеры и их применение в экспериментальной физике высоких и сверхвысоких энергий // *ЭЧАЯ*. – т. 32, вып. 2, Дубна, 2001. – с. 431–480.
25. *Nucl. Instr. Meth.* – **A389**, 1997.
26. URL: <http://root.cern.ch>
27. Lebedev S., Ososkov G., Hoehne C. Ring recognition in the CBM RICH detector // *JINR Comm.* E10-2007-88, JINR, Dubna, 2007.
28. Кудрявцев А. М. и др. Использование искусственных нейронных сетей при определении сортовых качеств семян твердой пшеницы // *Сельскохозяйственная биология*. – 2002, № 1. – с. 121–124.

29. *Ососков Г. А., Филимонов А. В.* Динамическая оптимизация структуры перцептронов // *Сообщ. ОИЯИ*, P11-2002-274, Дубна, 2002.
30. *Лоули Д., Максвелл А.* Факторный анализ как статистический метод. – М.: Мир, 1967.
31. *Брилюк Д. В., Старовойтов В. В.* Распознавание человека по изображению лица нейросетевыми методами // Препринт, Ин-т техн. кибернетики НАН Беларуси, №2, Минск, 2002. – 54 с.
32. *Peterson C.* Track finding with neural networks // *Nucl. Instr. and Meth.* – 1986, **A279**. – p. 537.
33. *Denby B.* Neural networks and cellular automata in experimental high energy physics // *Comp. Phys. Comm.* – 1988, **49**. – p. 429.
34. *Ососков Г. А. и др.* Использование нейронных сетей для улучшения интерпретации эксперимента EXCHARM // *Матем. моделир.* – 1999, т. 11, вып. 10. – с. 116–126.
35. *Peterson C.* Preprint LU TP 90-6, Lund, 1990.
36. *Baginyan S. et al.* Tracking by modified rotor model of neural network // *Comp. Phys. Comm.* **79**, 1994. – pp. 165–179.
37. *Ososkov G., Zaznobina E.* Tracking by cellular automata and neural networks // *Intern. Journ. of Neural Systems.* – v. 6. (Suppl. 1995). – pp. 269–274.
38. *Galkin I. et al.* Feedback neural networks for ARTIST ionogram processing // *Radio Science.* – v. 31, No. 5, 1996. – pp. 1119–1128.
39. *Grozov V.P., Ososkov G.A. et al.* Automatic processing of ionograms on the basis of the artificial neural network method // *Proc. Intern. Symp. on Radio Propagation (ISRP'97)*, Qingdao, China. – 1997. – pp. 514–517.
40. *Baginyan S., Ososkov G.* Controlled neural network application in track-match problem // *JINR Comm.* E-10-93-415, JINR, Dubna, 1993.
41. *Gyulassy M., Harlander M.* Elastic tracking and neural networks algorithms for complex pattern recognition // *Comp. Phys. Comm.* **66**, 1991. – pp. 31–46.
42. *Ohlsson M. et al.* Track finding with deformable templates – the elastic arms approach // *Comp. Phys. Commun.* – **71**, 1992. – p. 77.
43. *Muresan L., Panebratsev Yu.* Deformable templates for circle recognition // *JINR Rapid Comm.* 1[81]-97, Dubna, 1997. – pp. 27–44.
44. *Baginyan S., Ososkov G.* Finding tracks detected by a drift tube system // *Comp. Phys. Comm.* – v. 108, No. 1, 1998. – pp. 20–28.
45. *Hough P. V. C.* A Method and Means for Recognizing Complex Patterns. – US Patent: 3,069,654, 1962.

46. *Грабовский В. И.* Клеточные автоматы как простые модели сложных систем // *Успехи современной биологии*. – 1995, том 115, вып. 4. – с. 412–418.
47. *Glazov A. et al.* Filtering tracks in discrete detectors using a cellular automaton // *Nucl. Instr. Meth.* – **A329**, 1993. – pp. 262–268.
48. *Kisel I., Masciocchi S.* CATS – a Cellular Automaton for Tracking in Silicon for the HERA-B vertex detector // *HERA-B Note* 99-242, 1999.
49. *Осоков Г. А., Тихоненко Е. А.* Новый генератор случайных чисел на базе двумерного клеточного автомата // *Математическое моделирование*. – 1996, т. 8, вып. 12. – с. 77–84.
50. *Кисель И. В., Нескоромный В. Н., Осоков Г. А.* Применение нейронных сетей в экспериментальной физике // *ЭЧАЯ*. – 1993, т. 24, вып. 6. – с. 1551–1595.
51. *Осоков Г. А. и др.* Современные методы обработки экспериментальных данных в физике высоких энергий // *ЭЧАЯ*. – 2002, т. 33, вып. 3. – с. 676–745.
52. *Вороновский Г. К., Махотило К. В., Петрашев С. Н., Сергеев С. А.* Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности // *Основа*, 1997. – с. 112.

Геннадий Алексеевич ОСОКОВ, доктор физико-математических наук, профессор, главный научный сотрудник Лаборатории информационных технологий Объединенного института ядерных исследований, Дубна. Области научных интересов: компьютерное моделирование, вычислительная статистика, распознавание образов, искусственный интеллект.