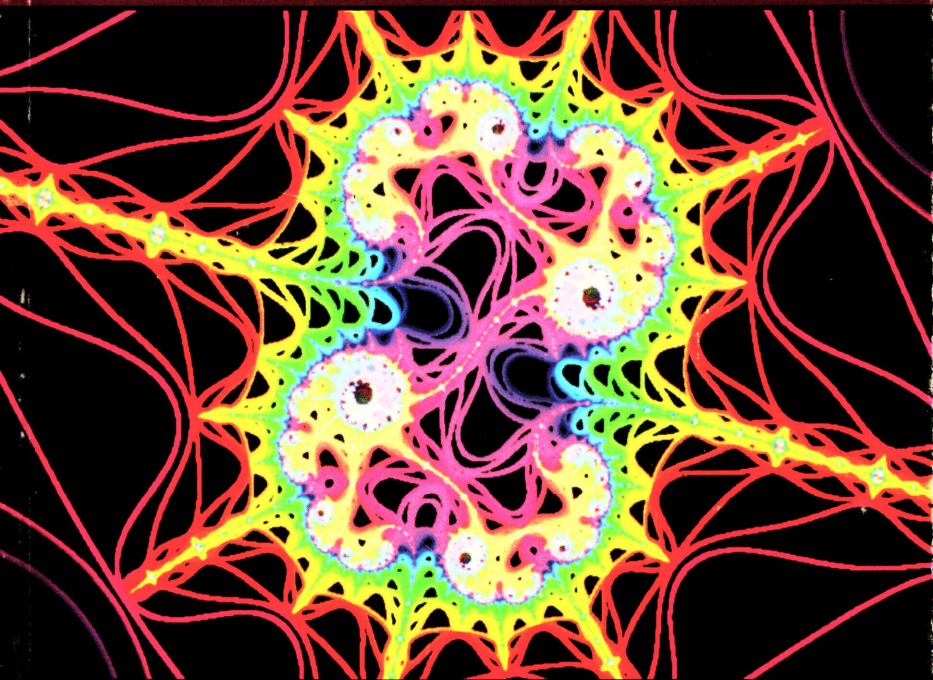


НЕЙРОННЫЕ СЕТИ И ИХ ПРИМЕНЕНИЕ В СИСТЕМАХ УПРАВЛЕНИЯ И СВЯЗИ



В. И. КОМАШИНСКИЙ
Д. А. СМЕРНОВ



**В. И. КОМАШИНСКИЙ
Д. А. СМЕРНОВ**

**НЕЙРОННЫЕ СЕТИ
И ИХ ПРИМЕНЕНИЕ
В СИСТЕМАХ УПРАВЛЕНИЯ
И СВЯЗИ**



**Москва
Горячая линия-Телеком
2003**

ББК 30.17
К 63



Комашинский В. И., Смирнов Д. А.

К63 Нейронные сети и их применение в системах управления и связи.–
М.: Горячая линия–Телеком, 2003. – 94 с.

ISBN 5-93517-094-9

Изложены основы теории искусственных нейронных сетей. Показано место нейронных сетей в эволюции интеллектуальных систем управления. Рассмотрены общие вопросы применения искусственных нейронных сетей в системах управления и связи. Отмечены преимущества, которые дает применение нейроинформационных технологий при решении многих как нетрадиционных, так и традиционных задач управления и связи.

Для научных работников и инженеров занимающихся разработкой систем управления и связи. Некоторые разделы могут быть интересны для специалистов других отраслей, а также широкому кругу читателей.

ББК 30.17

Адрес издательства в Интернет radios_hl@mtu-net.ru

Комашинский Владимир Ильич, Смирнов Дмитрий Анатольевич
Нейронные сети и их применение в системах управления и связи

ЛР № 071825 от 16 марта 1999 г.
Подписано в печать 28.11.2002. Печать офсетная
Формат 60х84/16. Тираж 1000 экз. Уч. изд. л. 5,9

ISBN 5-93517-094-9

© Комашинский В. И., Смирнов Д. А., 2002, 2003
© Оформление издательства «Горячая линия–Телеком», 2003

Предисловие

Ядром нейроинформационных технологий является представление о том, что естественные биологические нейроны можно моделировать довольно простыми искусственными автоматами, а вся сложность мозга, его гибкость в обработке различного рода информации и другие его важнейшие качества, определяются связями между нейронами. Каждая связь представляется как совсем простой элемент, служащий для обмена сигналами. Предельным выражением этой точки зрения является лозунг: "Структура связей – все, свойства элементов – ничто".

Совокупность идей, определяющая описанное представление о мозге носит название коннекционизма (по-английски connection – связь). Коннекционизм оперирует рядом несложных идей, включающих понятия однородности системы (элементы одинаковы и чрезвычайно просты, все определяется структурой связей), надежности системы из ненадежных элементов, "голографичности" системы – при разрушении случайно выбранной части система сохраняет свои полезные свойства.

Для описания устройств и алгоритмов функционирования искусственных нейронных систем разработана специальная системотехника (сумматоры, синапсы, нейроны, дендриты), предусматривающая объединение простейших устройств в функционально ориентированные сети, предназначенные для решения конкретных задач.

С середины 1980-х гг. непрерывно растет интерес к созданию специализированных устройств, получивших название нейрокомпьютеров.

Существует большое разнообразие нейрокомпьютеров – от специализированных интегральных схем, в которые вводится заранее определенная структура нейронной сети, до универсальных программируемых сопроцессоров к вычислительным машинам, на которых можно реализовать модель любой нейронной сети. Существует также и целый ряд промежуточных типов нейрокомпьютеров с той или иной степенью специализации. Особенностью нейрокомпьютеров является возможность сформировать стандартный способ решения многих нестандартных задач. Вместо программирования в нейрокомпьютерах применяются различные процедуры обучения. При этом труд программиста заменяется трудом – "учителя" (в некоторых случаях в качестве учителя может выступать имитационная или аналитическая модель).

Анализ зарубежных разработок нейрокомпьютеров позволил выделить основные перспективные направления современного развития нейроинформационных технологий: нейросетевые экспертные системы, СУБД с включением нейросетевых алгоритмов, обработка изображений и

сигналов, управление динамическими системами и в том числе сетями связи, управление финансовой деятельностью и др. Сегодня разработками в этой области занимается более 300 зарубежных компаний, причем число их постоянно увеличивается. Среди них такие гиганты, как Intel, DEC, IBM и Motorola. Вместе с тем главным в развитии нейроинформационных технологий является интеллектуализация вычислительных систем, придание им свойств человеческого мышления и восприятия. Многие отечественные и зарубежные специалисты предполагают, что нейрокompьютеры станут основной платформой для развития вычислительной техники 21 века. В этой связи необходимо констатировать, что в Японии с 1993 года принята программа "Real world computing program". Ее основная цель – создание адаптивной, эволюционирующей ЭВМ. Проект рассчитан на 10 лет. Основой разработки является нейротехнология, используемая для распознавания образов, обработки семантической информации, управления информационными потоками и роботами, которые способны адаптироваться к окружающей обстановке.

Монография является второй большой работой авторов, посвященной нейроинформатике. В отличие от носившей обобщающий характер первой, данная работа подробно рассматривает проблемы, связанные с применением искусственных нейронных сетей в системах управления и связи.

Академик МАИ,
доктор технических наук,
профессор Н. И. Буренин



Введение

Теория нейронных сетей (НС) является активно развивающимся направлением науки. Основные перспективы использования этой теории связаны с решением сложных практических задач. Нейронные сети – один из основных архитектурных принципов построения ЭВМ шестого поколения.

Интерес к нейронным сетям как к моделям биологической обработки информации не угасает более 30 лет. Начало современным моделям НС было положено в работе У. Маккаллока и У. Питтса [26]. Эти авторы сделали первую попытку эмулировать человеческие способности, классифицировать и распознавать образы. Дальнейшее развитие связано с работой Ф. Розенблатта [29]. Его модель была названа перцептроном. После некоторого затишья, с начала 1980-х годов начался и продолжается до настоящего времени новый виток развития моделей НС. Он связан с работами С. Гроссберга, Т. Кохонена, Д. Хопфилда и др. [63, 77, 78, 82], а также влиянием успехов развития оптических технологий и СБИС [7, 10, 12, 19, 22, 32], способных обеспечить реализацию новых архитектур.

Широкий круг задач, решаемый НС, не позволяет в настоящее время создавать универсальные, мощные сети, вынуждая разрабатывать специализированные НС, функционирующие по различным алгоритмам. В настоящее время существует большое количество разнообразных типов НС, предназначенных для решения различных задач. Эти модели отличаются структурой связей, правилами определения весов или правилами обучения, программирования.

Зачем использовать НС вместо старых методов? Всегда ли нейросетевые методы эффективнее (быстрее, точнее, экономичнее ...) старых методов? В чем же истинные преимущества применения НС в системах управления и связи, и существуют ли они? Данная книга – попытка ответить на эти и другие вопросы.

Основу книги составляют результаты анализа авторами большого количества отечественной и зарубежной литературы. Основное внимание уделено анализу применения нейроинформационных технологий в инфотелекоммуникационных системах. Высокая эффективность нейроинформационных технологий при решении задач адаптивного управления динамичными системами может уже в ближайшее время сделать их незаменимыми при создании новых поколений сетей мобильной связи и других беспроводных сетей.

Ряд результатов представляют собой предмет исследования авторов и публикуются впервые.

Книга предназначена для преподавателей, научных работников и инженеров, интересующихся развитием прикладных вычислительных систем на базе нейрокомпьютеров.

Для более детального изучения теории нейронных сетей и применения нейроинформационных технологий авторы рекомендуют ознакомиться с литературой [32, 41, 70, 93].

Глава 1. Основы теории искусственных нейронных сетей

1.1. Нейронные сети: общие положения

Актуальность исследований искусственных нейронных сетей подтверждается многообразием их возможных применений. Теория НС является алгоритмическим базисом развития нейрокомпьютеров, подобно тому, как булева алгебра более 50 лет являлась алгоритмическим базисом однопроцессорных и многопроцессорных ЭВМ.

Современные цифровые вычислительные машины превосходят человека по способности производить числовые и символьные вычисления. Однако человек может без усилий решать сложные задачи восприятия внешних данных (например, узнавать в толпе знакомого только по его промелькнувшему лицу) с такой скоростью и точностью, что самый мощный в мире компьютер по сравнению с ним окажется безнадёжным тупоумом. Причина столь значительного различия в их производительности заключена в том, что архитектура биологической нейронной системы совершенно не похожа на архитектуру машины фон Неймана (табл. 1.1) [14], а это существенно влияет на типы функций, которые более эффективно исполняются каждой из моделей.

Т а б л и ц а 1.1

Машина фон Неймана по сравнению с биологической
нейронной системой

Параметры сравнения	Машина фон Неймана	Биологическая нейронная система
	1	2
Процессор	Сложный	Простой
	Высокоскоростной	Низкоскоростной
	Один или несколько	Большое количество
Память	Отделена от процессора	Интегрирована в процессор
	Локализована	Распределенная
	Адресация не по содержанию	Адресация по содержанию
Надежность	Высокая уязвимость	Живучесть

Параметры сравнения	Машина фон Неймана	Биологическая нейронная система
	1	2
Вычисления	Централизованные	Распределенные
	Последовательные	Параллельные
	Хранимые программы	Самообучение
Специализация	Численные и символьные операции	Проблемы восприятия
Среда функционирования	Строго определенная	Плохо определенная
	Строго ограниченная	Без ограничений

Нейронная сеть – это сеть с конечным числом слоев из однотипных элементов – аналогов нейронов с различными типами связей между слоями [32]. При этом число нейронов в слоях выбирается исходя из необходимости обеспечения заданного качества решения задачи, а число слоев нейронов – как можно меньшее для сокращения времени решения задачи.

Развитие искусственных нейронных сетей вдохновлялось и вдохновляется биологией. Однако наши знания о работе мозга ограничены, и разработчикам искусственных НС приходится выходить за пределы современных биологических знаний в поисках структур, позволяющих выполнять полезные функции. Так часто создаются сети невозможные в живой материи или требующие неправдоподобно больших допущений об анатомии и функционировании мозга.

Нервная система человека, построенная из элементов, называемых нейронами, очень сложна. Около 10^{11} нейронов участвуют в примерно 10^{15} передающих связях. Уникальными способностями нейронов являются прием, обработка и передача электрохимических сигналов по нервным путям, которые образуют коммуникационную систему мозга.

Несмотря на огромное количество нейронов, их тела занимают всего несколько процентов общего объема мозга. Почти все остальное пространство занято межнейронными связями. Число связей каждого нейрона не имеет аналогов в современной технике. Понимание того, что межней-

ронные связи относятся к основным структурным компонентам мозга, в первую очередь, определяющим его функциональные характеристики, является одним из наиболее существенных выводов, сделанных нейрофизиологами. В подтверждение можно привести высказывание известного нейрофизиолога Э. Кэндела: "По убеждению многих нейробиологов в конце концов будет доказано, что уникальные свойства каждого человека – способность чувствовать, думать, обучаться и помнить – заключены в строго организованных сетях синаптических взаимосвязей между нейронами головного мозга" [32].

На рис. 1.1 приведена структура типичного биологического нейрона. Дендриты идут от тела нервной клетки к другим нейронам, где они принимают сигналы в точках соединения, называемых синапсами. Принятые синапсом входные сигналы подводятся к телу нейрона. Здесь они суммируются, причем одни входы стремятся возбудить нейрон, а другие воспрепятствовать его возбуждению. В соответствии с этим различают процессы синаптического возбуждения и торможения нейрона.



Рис. 1.1. Биологический нейрон

Когда суммарное возбуждение в теле нейрона превышает некоторый порог, нейрон возбуждается, посылая по аксону, который на конце сильно ветвится, сигнал другим нейронам. Нервный импульс обладает способностью распространяться по нервным волокнам без затухания, вплоть до синаптических окончаний со скоростью распространения около 10 м/с.

Синаптическое торможение снижает возбуждение нейрона и тем самым затрудняет переход возбуждения нейрона в нервный импульс.

После генерации импульса наступает так называемый рефрактерный период, во время которого нейрон восстанавливает способность к генерации следующего импульса. Длительность рефрактерного периода менее 1 мс. Поэтому максимальная частота генерации импульсов нейронов около 1000 с^{-1} .

Искусственный нейрон имитирует в первом приближении свойства биологического нейрона. Он обладает группой синапсов – однонаправленных входных связей, соединенных с выходами других нейронов, а также имеет аксон – выходную связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов. Каждый вход умножается на соответствующий вес, аналогичный синаптической силе, и все произведения суммируются, определяя уровень активации нейрона. Общий вид нейрона, реализующего эти свойства, приведен на рис. 1.2. Здесь множество входных сигналов обозначены вектором X . Каждый вес w_i соответствует "силе" одной биологической синаптической связи. Множество весов в совокупности обозначается вектором W .

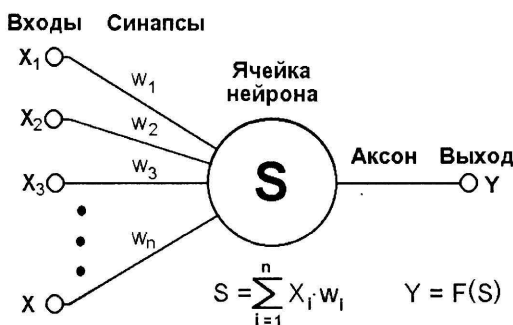


Рис. 1.2. Искусственный нейрон

Суммирующий блок, соответствующий телу биологического элемента, складывает взвешенные входы алгебраически

$$S = \sum_{i=1}^n x_i \cdot w_i . \quad (1.1)$$

Выход нейрона является функцией его состояния:

$$Y = F(x) . \quad (1.2)$$

Нелинейная функция F называется активационной и может иметь различный вид, что, в частности, показано на рис. 1.3.

В случае, когда функция активации одна и та же для всех нейронов сети, сеть называют однородной (гомогенной). Если же активационная функция зависит еще от одного или нескольких параметров, значения которых меняются от нейрона к нейрону, то сеть называют неоднородной (гетерогенной).

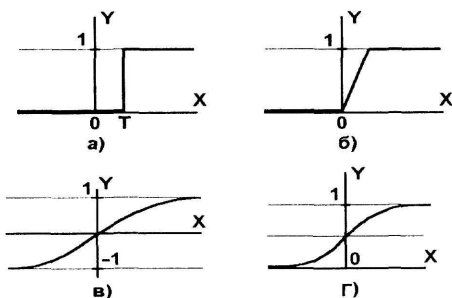


Рис. 1.3. Типы активационных функций:

- а – функция единичного скачка; б – линейный порог (гистерезис);
в – сигмоид – гиперболический тангенс; г – сигмоид – формула (1.3)

Одной из наиболее распространенных является нелинейная функция с насыщением, так называемая логистическая функция или сигмоид (т.е. функция S-образного вида) [32]:

$$f(x) = \frac{1}{1 + e^{-\alpha x}}. \quad (1.3)$$

При уменьшении α сигмоид становится более пологим, в пределе при $\alpha = 0$ вырождаясь в горизонтальную линию на уровне 0.5. При увеличении α сигмоид приближается по внешнему виду к функции единичного скачка с порогом T в точке $x = 0$. Из выражения для сигмоида очевидно, что выходное значение нейрона лежит в диапазоне $[0, 1]$. Одно из ценных свойств сигмоидной функции – простое выражение для ее производной, применение которого будет рассмотрено в дальнейшем,

$$f'(x) = \alpha f(x)(1 - f(x)). \quad (1.4)$$

Следует отметить, что сигмоидная функция дифференцируема на всей оси абсцисс, что широко используется во многих алгоритмах обучения. Кроме того, она обладает свойством усиливать слабые сигналы лучше, чем сильные, и предотвращает насыщение от сильных сигналов, так как они соответствуют областям аргументов, где сигмоид имеет пологий наклон. Другой широко используемой активационной функцией является гиперболический тангенс. В отличие от логистической функции гиперболический тангенс принимает значения различных знаков, что для ряда сетей оказывается выгодным.

Говоря о возможной классификации НС, важно отметить существование бинарных и аналоговых сетей. Первые из них оперируют с двоичными сигналами, и выход каждого нейрона может принимать только два значения: логический ноль ("заторможенное" состояние) и логическая единица ("возбужденное" состояние). В аналоговых сетях выходные зна-

чения нейронов способны принимать непрерывные значения. Еще одна классификация делит НС на синхронные и асинхронные [13]. В первом случае в каждый момент времени свое состояние меняет лишь один нейрон. Во втором – состояние меняется сразу у целой группы нейронов, как правило, у всего слоя. Для программных имитаторов нейронных сетей на цифровых ЭВМ, вопросы связанные с синхронизацией решаются компьютером, на котором реализуются НС.

Рассмотренная простая модель искусственного нейрона существенно упрощает ряд свойств своего биологического двойника. Например, она не принимает во внимание задержки во времени, которые воздействуют на динамику системы. Входные сигналы сразу же порождают выходной сигнал. И, что более важно, она не учитывает воздействие синхронизирующей функции биологического нейрона, которую ряд исследователей [32] считают решающей. Несмотря на эти ограничения, сети, построенные из этих нейронов, обнаруживают свойства, сильно напоминающие биологическую систему. Только время и дальнейшие исследования могут дать ответ на вопрос, являются ли подобные совпадения случайными или это следствие того, что в модели верно схвачены основные черты биологического нейрона.

Итак, дано описание основных элементов, из которых составляются нейронные сети. Перейдем теперь к вопросу: как составлять эти сети, как их конструировать? Строгих ограничений здесь нет, как угодно, лишь бы входы получали какие-нибудь сигналы. Возможности безграничны, но обычно используют несколько стандартных архитектур, из которых при некоторых небольших модификациях строят большинство используемых сетей.

1.2. Архитектура искусственных нейронных сетей

1.2.1. Искусственные нейронные сети прямого распространения

Хотя один нейрон и способен выполнять простейшие процедуры распознавания, сила нейронных вычислений проистекает от соединений нейронов в сетях.

Широкий круг задач, решаемый НС, не позволяет в настоящее время создавать универсальные, мощные сети, вынуждая разрабатывать специализированные НС, функционирующие по различным алгоритмам.

Выбор структуры НС осуществляется в соответствии с особенностями и сложностью задачи. Для решения некоторых отдельных типов задач уже существуют оптимальные, на сегодняшний день, конфигурации, описан-

ные, например, в [13, 14, 32, 70, 77, 82] и других изданиях. Основные из них приведены на рис. 1.4 [14].

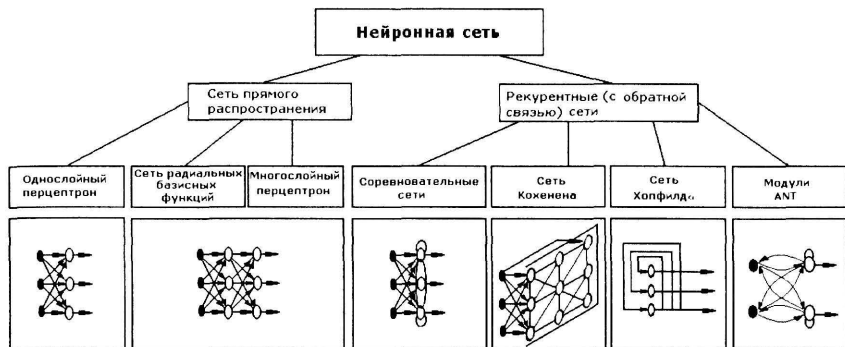


Рис. 1.4. Варианты наиболее распространенных архитектур искусственных нейронных сетей

Если же задача не может быть сведена ни к одному из известных типов, разработчику приходится решать сложную проблему синтеза новой конфигурации. При этом он руководствуется несколькими основополагающими принципами: возможности сети возрастают с увеличением числа ячеек сети, плотности связей между ними и числом выделенных слоев; введение обратных связей наряду с увеличением возможностей сети поднимает вопрос о динамической устойчивости сети; сложность алгоритмов функционирования сети (в том числе, например, введение нескольких типов синапсов – возбуждающих, тормозящих и др.) также способствует усилению мощи НС. Вопрос о необходимых и достаточных свойствах сети для решения того или иного рода задач представляет собой целое направление нейрокомпьютерной науки. Так как проблема синтеза НС сильно зависит от решаемой задачи, дать общие подробные рекомендации затруднительно. В большинстве случаев оптимальный вариант получается на основе интуитивного подбора. Единственное жесткое требование, предъявляемое архитектурой к элементам сети, это соответствие размерности вектора входных сигналов сети числу ее входов.

Простейшая однослойная сеть состоящая из группы нейронов, показана на рис. 1.5.

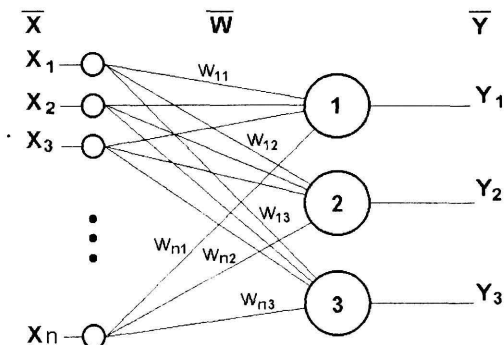


Рис. 1.5. Однослойная НС

На n входов поступают некие сигналы, проходящие по синапсам на 3 нейрона, образующие единственный слой этой НС и выдающие три выходных сигнала:

$$y_j = f \left[\sum_{i=1}^n x_i w_{ij} \right], \quad (1.5)$$

где $j = 1, 2, 3$.

В искусственных и естественных (биологических) сетях многие соединения могут отсутствовать, все соединения показаны в целях общности.

Очевидно, что все весовые коэффициенты синапсов одного слоя нейронов можно свести в матрицу $\mathbf{W}$, в которой каждый элемент w_{ij} задает величину i -й синаптической связи j -го нейрона. Таким образом процесс, происходящий в НС, может быть записан в матричной форме:

$$\mathbf{Y} = F(\mathbf{XW}), \quad (1.6)$$

где $\mathbf{X}$ и $\mathbf{Y}$ – соответственно входной и выходной сигнальные векторы; $F(\mathbf{V})$ – активационная функция, применяемая поэлементно к компонентам вектора $\mathbf{V}$.

На рис. 1.6 представлена двухслойная НС, полученная из однослойной (см. рис. 1.5) путем добавления второго слоя, состоящего из двух нейронов. Здесь уместно отметить важность роли рассмотренной ранее нелинейности активационной функции, так как, если бы она не обладала данным свойством или не входила в алгоритм работы каждого нейрона, результат функционирования любой p -слойной НС с весовыми матрицами $\mathbf{W}^{(i)}$, где $i = 1, 2, \dots, p$ для каждого слоя i сводился бы к перемножению входного вектора сигналов $\mathbf{X}$ на матрицу

$$\mathbf{W}^{(\Sigma)} = \mathbf{W}^{(1)} \mathbf{W}^{(2)} \dots \mathbf{W}^{(p)}, \quad (1.7)$$

то есть фактически такая р-слойная НС эквивалентна однослойной НС с весовой матрицей единственного слоя $\mathbf{W}^{(\Sigma)}$;

$$\mathbf{Y} = \mathbf{X} \mathbf{W}^{(\Sigma)}. \quad (1.8)$$

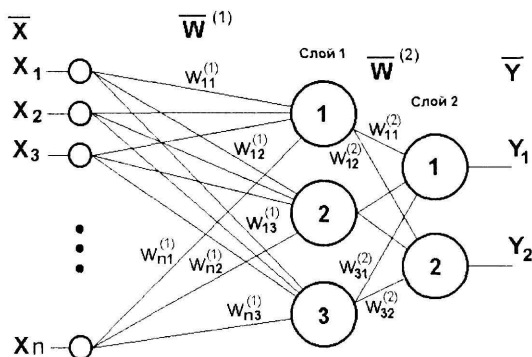


Рис. 1.6. Двухслойная НС

Таким образом, для расширения вычислительных возможностей многослойных НС по сравнению с однослойными НС необходимо использование нелинейных активационных функций.

Продолжая разговор о нелинейности, необходимо отметить, что она иногда может вводиться и в синаптические связи. В большинстве известных на сегодняшний день НС для нахождения взвешенной суммы входов нейрона используют формулу (1.1), однако в некоторых приложениях НС полезно ввести другую запись, например:

$$S = \sum_{i=1}^n x_i^2 w_i, \quad (1.9)$$

или

$$S = \sum_{i=1}^n x_i x_{((i+1) \bmod n)} w_i. \quad (1.10)$$

Вопрос в том, чтобы разработчик НС четко понимал, для чего он это делает, какими ценными свойствами он тем самым дополнительно наделяет нейрон, и каких лишает. Введение такого рода нелинейности, вообще говоря, увеличивает вычислительную мощь сети, то есть позволяет из меньшего числа нейронов с "нелинейными" синапсами сконструировать

НС, выполняющую работу обычной НС с большим числом стандартных нейронов и более сложной конфигурацией [32].

У сетей, рассмотренных до сих пор, не было обратных связей, то есть соединений, идущих от выходов некоторого слоя к входам этого же слоя или предшествующих слоев. Этот специальный класс сетей, называемых сетями без обратных связей или сетями прямого распространения, представляет интерес и широко используется.

Несколько слов необходимо сказать о требуемой мощности выходного слоя сети, выполняющего окончательную классификацию пространства состояний. Дело в том, что для разделения множества входных образов, например, по двум классам достаточно всего одного выхода. При этом каждый логический уровень – "1" и "0" – будет обозначать отдельный класс. На двух выходах можно закодировать уже четыре класса и т. д. Однако результаты работы сети, организованной таким образом недостаточны надежны. Для повышения достоверности классификации желательно ввести избыточность путем выделения каждому классу одного нейрона в выходном слое или, что еще лучше, нескольких, каждый из которых обучается определять принадлежность конкретных состояний входов к определенному классу со своей степенью достоверности, например высокой, средней и низкой. Такие НС. позволяют проводить классификацию входных неявно выраженных состояний, объединенных в нечеткие (размытые или пересекающиеся) множества. Это свойство позволяет широко использовать НС в практических приложениях.

1.2.2. Искусственные нейронные сети с обратными связями

Ранее были рассмотрены сети, которые не имели обратных связей. Отсутствие обратных связей гарантирует устойчивость сетей. Однако такое желательное свойство ограничивает возможности сетей прямого распространения по сравнению с сетями с обратными связями. К тому же в работе [78] была получена теорема, описавшая подмножество сетей с обратными связями, выходы которых в конце концов достигают устойчивого состояния. Важный вклад как в теорию, так и в применение систем с обратными связями сделал Дж. Хопфилд. Поэтому некоторые из конфигураций носят его имя.

Структурная схема сети Хопфилда приведена на рис.1.7. Она состоит из единственного слоя нейронов, число которых является одновременно числом входов и выходов сети.

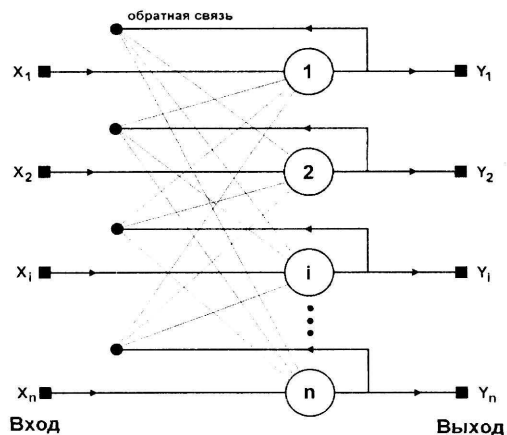


Рис. 1.7. Структурная схема сети Хопфилда

Каждый нейрон связан синапсами со всеми остальными нейронами, а также имеет один входной синапс, через который осуществляется ввод сигнала. Выходные сигналы, как обычно, образуются на аксонах.

Задача, решаемая данной сетью в качестве ассоциативной памяти, как правило, формулируется следующим образом. Известен некоторый набор двоичных сигналов (изображений, звуковых оцифровок, прочих данных, описывающих некие объекты или характеристики процессов), которые считаются образцовыми. Сеть должна уметь из произвольного неидеального сигнала, поданного на ее вход, выделить ("вспомнить" по частичной информации) соответствующий образец (если такой есть) или "дать заключение" о том, что входные данные не соответствуют ни одному из образцов. В общем случае, любой сигнал может быть описан вектором $\mathbf{X} = \{x_i : i = 0 \dots (n-1)\}$, n – число нейронов в сети и размерность входных и выходных векторов. Каждый элемент x_i равен либо $+1$, либо -1 . Обозначим вектор, описывающий k -й образец, через $\mathbf{X}^k$, а его компоненты, соответственно, $-x_i^k$, $k = 0 \dots (m-1)$, m – число образцов. Когда сеть распознает (или "вспомнит") какой-либо образец на основе предъявленных ей данных, ее выходы будут содержать именно его, то есть $\mathbf{Y} = \mathbf{X}^k$, где $\mathbf{Y}$ – вектор выходных значений сети: $\mathbf{Y} = \{y_i : i = 0, \dots, (n-1)\}$. В противном случае, выходной вектор не совпадет ни с одним образцовым.

На стадии инициализации сети весовые коэффициенты синапсов устанавливаются следующим образом [32]:

$$w_{ij} = \begin{cases} \sum_{k=0}^{m-1} x_i^k x_j^k, & i \neq j \\ 0, & i = j \end{cases} \quad (1.11)$$

Здесь i и j – индексы, соответственно, предсинаптического и постсинаптического нейронов; x_i^k, x_j^k – i -й и j -й элементы вектора k -го образца.

Алгоритм функционирования сети следующий (p – номер итерации):

1. На входы сети подается неизвестный сигнал. Фактически его ввод осуществляется непосредственной установкой значений аксонов:

$$y_i(0) = x_i, i = 0 \dots (n-1), \quad (1.12)$$

поэтому обозначение на схеме сети входных синапсов в явном виде носит чисто условный характер. Ноль в скобке справа от y_i означает нулевую итерацию в цикле работы сети.

2. Рассчитывается новое состояние нейронов

$$s_j(p+1) = \sum_{i=0}^{n-1} w_{ij} y_i(p), \quad j = 0 \dots (n-1) \quad (1.13)$$

и новые значения аксонов

$$y_j(p+1) = f[s_j(p+1)], \quad (1.14)$$

где f – активационная функция в виде скачка, приведенная на рис. 1.8а.

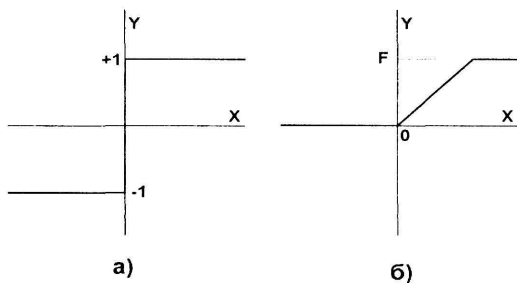


Рис. 1.8 Активационные функции

3. Проверка, изменились ли выходные значения аксонов за последнюю итерацию. Если да – переход к пункту 2, иначе (если выходы застabilizировались) – конец.

При этом выходной вектор представляет собой образец, наилучшим образом сочетающийся с входными данными.

Как говорилось выше, иногда сеть не может провести распознавание и выдает на выходе несуществующий образ. Это связано с проблемой ограниченности возможностей сети. Для сети Хопфилда число запоминаемых образов m не должно превышать значения равного $0.15 \cdot n$. Кроме того, если два образа А и Б сильно похожи, они, возможно, будут вызывать у сети перекрестные ассоциации, то есть предъявление на входы сети вектора А приведет к появлению на ее выходах вектора Б и наоборот. Еще одним недостатком сетей Хопфилда является их тенденция стабилизироваться в локальном, а не в глобальном минимуме. Эта трудность преодолевается в основном с помощью класса сетей, известных под названием машин Больцмана, в которых изменения состояний нейронов обусловлены статистическими, а не детерминированными закономерностями [32]. Принцип машины Больцмана может быть перенесен на сети практически любой конфигурации, хотя устойчивость не гарантируется.

Когда нет необходимости, чтобы сеть в явном виде выдавала образец, то есть достаточно, скажем, получать номер образца, ассоциативную память успешно реализует сеть Хэмминга. Данная сеть характеризуется, по сравнению с сетью Хопфилда, меньшими затратами на память и объемом вычислений, что становится очевидным из ее структуры (рис. 1.9).

Сеть состоит из двух слоев. Первый и второй слои имеют по m нейронов, где m – число образцов. Нейроны первого слоя имеют по n синапсов, соединенных с входами сети (образующими фиктивный нулевой слой). Нейроны второго слоя связаны между собой ингибиторными (отрицательными обратными) синаптическими связями. Единственный синапс с положительной обратной связью для каждого нейрона соединен с его же аксоном.

Идея работы сети состоит в нахождении расстояния Хэмминга от тестируемого образа до всех образцов. Расстоянием Хэмминга называется число отличающихся битов в двух бинарных векторах. Сеть должна выбрать образец с минимальным расстоянием Хэмминга до неизвестного входного сигнала, в результате чего будет активизирован только один выход сети, соответствующий этому образцу.

На стадии инициализации весовым коэффициентам первого слоя и порогу активационной функции присваиваются следующие значения:

$$w_{ik} = \frac{x_i^k}{2}, \quad i = 0 \dots (n-1), k = 0 \dots (m-1), \quad (1.15)$$

$$T_k = \frac{n}{2}, \quad k = 0 \dots (m-1). \quad (1.16)$$

Здесь x_i^k – i -й элемент k -го образца.

Весовые коэффициенты тормозящих синапсов во втором слое берут равными некоторой величине $0 < \varepsilon < 1/m$. Синапс нейрона, связанный с его же аксоном имеет вес +1.

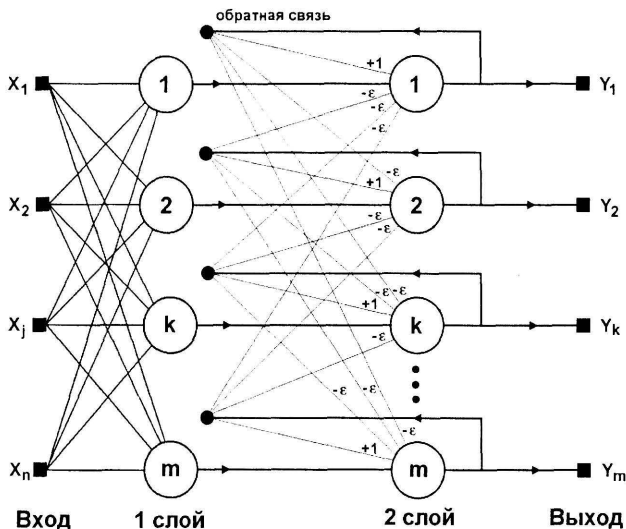


Рис. 1.9. Структурная схема сети Хэмминга

Алгоритм функционирования сети Хэмминга следующий:

1. На входы сети подается неизвестный вектор $\mathbf{X} = \{x_i : i = 0 \dots (n-1)\}$, исходя из которого рассчитываются состояния нейронов первого слоя (верхний индекс в скобках указывает номер слоя):

$$y_j^{(1)} = s_j^{(1)} = \sum_{i=0}^{n-1} w_{ij} x_i + T_j, \quad j = 0 \dots (m-1). \quad (1.17)$$

После этого полученными значениями инициализируются значения аксонов второго слоя:

$$y_j^{(2)} = y_j^{(1)}, \quad j = 0 \dots (m-1). \quad (1.18)$$

2. Вычисляются новые состояния нейронов второго слоя:

$$s_j^{(2)}(p+1) = y_j(p) - \varepsilon \sum_{k=0}^{m-1} y_k^{(2)}(p), \quad k \neq j, \quad j = 0 \dots (m-1) \quad (1.19)$$

и значения их аксонов:

$$y_j^{(2)}(p+1) = f[s_j^{(2)}(p+1)], \quad j = 0 \dots (m-1). \quad (1.20)$$

Активационная функция f имеет вид порога (рис. 1.8б), причем величина F должна быть достаточно большой, чтобы любые возможные значения аргумента не приводили к насыщению.

3. Проверка, изменились ли выходы нейронов второго слоя за последнюю итерацию. Если да – перейти к шагу 2. Иначе – конец.

Из алгоритма видно, что роль первого слоя весьма условна: воспользовавшись один раз на шаге 1 значениями его весовых коэффициентов, сеть больше не обращается к нему, поэтому первый слой может быть вообще исключен из сети (заменен на матрицу весовых коэффициентов).

Обсуждение НС с обратными связями, реализующих ассоциативную память, было бы неполным без хотя бы краткого упоминания о двунаправленной ассоциативной памяти (ДАП). Она является логичным развитием парадигмы сети Хопфилда, к которой для этого достаточно добавить второй слой. Структура ДАП представлена на рис. 1.10.

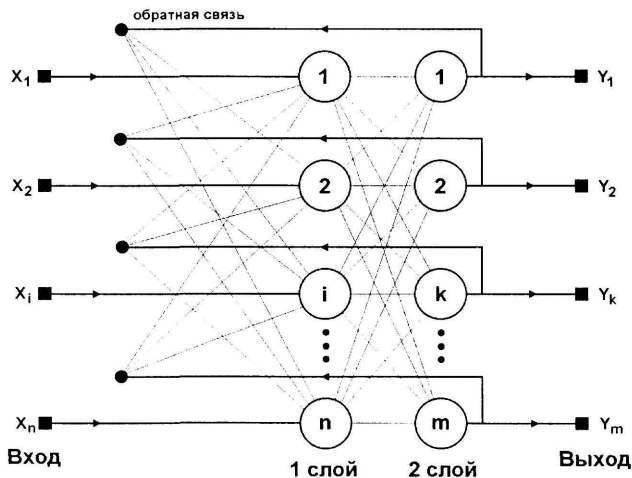


Рис. 1.10. Структурная схема ДАП

Как и сеть Хопфилда, ДАП способна к обобщению, вырабатывая правильные реакции, несмотря на искаженные входы. Кроме того, могут быть реализованы адаптивные версии ДАП, выделяющие эталонный образ из зашумленных экземпляров. Эти возможности сильно напоминают процесс мышления человека и позволяют искусственным НС сделать шаг в направлении моделирования мозга.

Сеть способна запоминать пары ассоциированных друг с другом образов. Пусть пары образов записываются в виде векторов $\mathbf{X}^k = \{x_i^k : i = 0 \dots (n-1)\}$ и $\mathbf{Y}^k = \{y_j^k : j = 0 \dots (m-1)\}$, $k = 0 \dots (r-1)$, где r – число пар. Подача на вход первого слоя некоторого вектора $\mathbf{P} = \{p_i : i = 0 \dots (n-1)\}$ вызывает образование на входе второго слоя некоего другого вектора $\mathbf{Q} = \{q_j : j = 0 \dots (m-1)\}$, который затем снова поступает на вход первого слоя. При каждом таком цикле, вектора на выходах обоих слоев приближаются к паре образцовых векторов, первый из которых – $\mathbf{X}$ – наиболее походит на $\mathbf{P}$, который был подан на вход сети в самом начале, а второй – $\mathbf{Y}$ – ассоциирован с ним. Ассоциации между векторами кодируются в весовой матрице $\mathbf{W}^{(1)}$ первого слоя. Весовая матрица второго слоя $\mathbf{W}^{(2)}$ равна транспонированной первой $(\mathbf{W}^{(1)})^T$. Процесс обучения, так же как и в случае сети Хопфилда, заключается в предварительном расчете элементов матрицы $\mathbf{W}$ (и соответственно $\mathbf{W}^T$) по формуле

$$w_{ij} = \sum_k x_i y_j, \quad i = 0 \dots (n-1), \quad j = 0 \dots (m-1). \quad (1.21)$$

Эта формула является развернутой записью матричного уравнения

$$\mathbf{W} = \sum_k \mathbf{X}^T \mathbf{Y} \quad (1.22)$$

для частного случая, когда образы записаны в виде векторов, при этом произведение двух матриц размером соответственно $[n \times 1]$ и $[1 \times m]$ приводит к (1.21).

Как и сети Хопфилда, ДАП имеет ограничения на максимальное количество ассоциаций, которые она может точно воспроизвести. Если этот лимит превышен, сеть может выработать неверный выходной сигнал, воспроизводя ассоциации, которым не обучена. В работе [70] приведены оценки, в соответствии с которыми количество запоминаемых ассоциаций не может превышать количества нейронов в меньшем слое. При этом предполагается, что емкость памяти максимизирована посредством специального кодирования, при котором количество компонент со значениями +1 равно числу компонент со значениями –1 в каждом биполярном векторе. Более точные оценки могут быть получены в результате обобщения работы [70] по оценке емкости сетей Хопфилда.

Несмотря на эти проблемы, ДАП является объектом интенсивных исследований. Основная привлекательность ДАП, как и сетей Хопфилда и Хэмминга, заключается в их простоте. Легкость построения программных и аппаратных моделей делает эти сети привлекательными для многих применений.

Еще одной разновидностью искусственных НС с обратными связями являются самоорганизующиеся карты Кохонена (SOM) [82]. Такая сеть является специальным случаем сети обучающейся методом соревнования и представляет собой векторный квантователь, задачей которого является определение принадлежности входного вектора $\mathbf{X} = \{x_1, x_2, \dots, x_N\}^T$ к одному из M возможных кластеров, представленных векторными центрами $\mathbf{W} = \{w_{j1}, w_{j2}, \dots, w_{jN}\}^T$ ($j = 1, 2, \dots, M$). Считается, что вектор $\mathbf{X}$ принадлежит к j -му кластеру, если расстояние d_j , определяемое как

$$d_j = \sum_{i=1}^N (x_i - w_{ji})^2, \quad (1.23)$$

для j -го центра кластера $\mathbf{W}_j$ минимально, то есть если $d_j \leq d_k$ для каждого $k \neq j$ [6]. Если узлы квантователя являются линейными, а вес i -го входа j -го узла равен w_{ji} для каждого i и j , то очевидно, что при соответствующих значениях порогов каждый i -й выход сети с точностью до несущественных постоянных будет равен евклидову расстоянию d_j между предъявленным входным вектором $\mathbf{X}$ и j -м центром кластера.

При обучении НС квантователя предъявляются входные векторы без указания желаемых выходов и корректируются веса согласно алгоритму, предложенному Т. Кохоненым (будет рассмотрен в п.1.3.2).

Самоорганизующиеся карты Кохонена могут быть использованы для проектирования многомерных данных, аппроксимации плотности и кластеризации. Эти сети успешно применяются для распознавания речи, обработки изображений, в робототехнике и в задачах управления [6, 14, 19].

1.3. Обучение искусственных нейронных сетей

1.3.1. Обучение с учителем

Когда идет разговор об использовании НС и нейросетевых алгоритмов, почти всегда подразумеваются определенные процедуры их обучения. НС представляет собой адаптивную систему, жизненный цикл которой состоит из двух независимых фаз – обучения и работы сети. Обучение считается законченным, когда сеть правильно выполняет преобразование на тестовых примерах и дальнейшее обучение не вызывает значительного изменения настраиваемых весовых коэффициентов. Далее сеть выполняет преобразование ранее неизвестных ей данных на основе сформированной ею

в процессе обучения нелинейной модели процесса. Сеть успешно работает до тех пор, пока существенно не изменится реальная модель отображаемого явления (например, в случае возникновения ситуации, информация о которой никогда ни предъявлялась сети при обучении). После этого сеть может быть дообучена с учетом новой информации, причем при дообучении предыдущая информация не теряется, а обобщается с вновь поступившей. При "повреждении" части весовых коэффициентов НС ее свойства могут быть полностью восстановлены в процессе дообучения.

От того, насколько качественно будет выполнен этап обучения НС, зависит способность сети решать поставленные перед ней проблемы во время эксплуатации. Теория обучения рассматривает три фундаментальных свойства, связанных с обучением по примерам: емкость, сложность образцов и вычислительная сложность. Под емкостью понимается, сколько образцов может запомнить сеть и какие функции и границы принятия решений могут быть на ней сформированы. Сложность образцов определяет число обучающих примеров, необходимых для достижения способности сети к обобщению. Важной характеристикой является время затрачиваемое на обучение. Как правило, время обучения и качество обучения связаны обратной зависимостью и выбирать эти параметры приходится на основе компромиса.

Существуют три парадигмы обучения: "с учителем", "без учителя" (самообучение) и смешанная.

В свою очередь множество различных алгоритмов обучения делятся на два больших класса: детерминистских и стохастических алгоритмов. В первом из них подстройка весов представляет собой жесткую последовательность действий, во втором – она производится на основе действий, подчиняющихся некоторому случайному процессу.

Обучение с учителем предполагает, что для каждого входного вектора существует целевой вектор, представляющий собой требуемый выход. Вместе они называются обучающей парой. Обычно сеть обучается на некотором числе таких обучающих пар.

Когда в сети только один слой, алгоритм ее обучения с учителем довольно очевиден, так как правильные выходные состояния нейронов единственного слоя заведомо известны, и подстройка синаптических связей идет в направлении, минимизирующем ошибку на выходе сети.

В многослойных же сетях оптимальные выходные значения нейронов всех слоев, кроме последнего, как правило, не известны, и двух или более слойную НС уже невозможно обучить, руководствуясь только величинами ошибок на выходах этой сети. Один из вариантов решения этой проблемы – разработка наборов выходных сигналов, соответствующих входным, для каждого слоя НС, что, конечно, является очень трудоемкой операцией и не всегда осуществимо. Второй вариант – динамическая под-

стройка весовых коэффициентов синапсов, в ходе которой выбираются, как правило, наиболее слабые связи, которые изменяются на малую величину в ту или иную сторону. Сохраняются же только те изменения, которые повлекли уменьшение ошибки на выходе всей сети. Очевидно, что данный метод "проб", несмотря на свою кажущуюся простоту, требует громоздких рутинных вычислений. И, наконец, третий, более приемлемый вариант – распространение сигналов ошибки от выходов НС к ее входам, в направлении обратном прямому распространению сигналов в обычном режиме работы. Этот алгоритм обучения НС получил название процедуры обратного распространения (back-propagation algorithm) и является наиболее широко используемым. Именно он будет более подробно рассмотрен в дальнейшем.

Согласно методу наименьших квадратов, минимизируемой целевой функцией ошибки НС является величина:

$$E(w) = \frac{1}{2} \sum_{j,p} (y_{j,p}^{(N)} - d_{j,p})^2, \quad (1.24)$$

где $y_{j,p}^{(N)}$ – реальное выходное состояние нейрона j выходного слоя N нейронной сети при подаче на ее входы p -го образа; $d_{j,p}$ – идеальное (желаемое) выходное состояние этого нейрона.

Суммирование ведется по всем нейронам выходного слоя и по всем обрабатываемым сетью образам. Минимизация ведется методом градиентного спуска, что означает подстройку весовых коэффициентов следующим образом:

$$\Delta w_{ij}^{(n)} = -\eta \frac{\partial E}{\partial w_{ij}}. \quad (1.25)$$

Здесь w_{ij} – весовой коэффициент синаптической связи, соединяющей i -й нейрон слоя $(n-1)$ с j -м нейроном слоя n , η – коэффициент скорости обучения, $0 < \eta < 1$.

Как показано в [100]:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial s_j} \frac{\partial s_j}{\partial w_{ij}}. \quad (1.26)$$

Здесь под y_j , как и раньше, подразумевается выход нейрона j , а под s_j – взвешенная сумма его входных сигналов, то есть аргумент активационной функции. Множитель $\frac{\partial y_j}{\partial s_j}$ является производной этой функции по ее аргументу. Из этого следует, что производная активационной функции должна быть определена на всей оси абсцисс. В связи с этим функция

единичного скачка и прочие активационные функции с неоднородностями не подходят для рассматриваемых НС. В них применяются такие гладкие функции, как гиперболический тангенс или классический сигмоид с экспонентой. В случае гиперболического тангенса

$$\frac{\partial y}{\partial s} = 1 - s^2. \quad (1.27)$$

Третий множитель $\frac{\partial s_j}{\partial w_{ij}}$, очевидно, равен выходу нейрона предыдущего слоя $y_i^{(n-1)}$.

Что касается первого множителя в (1.26), он легко раскладывается следующим образом [100]:

$$\frac{\partial E}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial s_k} \frac{\partial s_k}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial s_k} w_{jk}^{(n+1)}. \quad (1.28)$$

Здесь суммирование по k выполняется среди нейронов слоя $(n+1)$. Введя новую переменную

$$\delta_j^{(n)} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial s_j}, \quad (1.29)$$

мы получим рекурсивную формулу для расчетов величин $\delta_j^{(n)}$ слоя n из величин $\delta_k^{(n+1)}$ более старшего слоя $(n+1)$.

$$\delta_j^{(n)} = \left[\sum_k \delta_k^{(n+1)} w_{jk}^{(n+1)} \right] \frac{\partial y_j}{\partial s_j}. \quad (1.30)$$

Для выходного же слоя

$$\delta_j^{(N)} = (y_j^{(N)} - d_j) \frac{\partial y_j}{\partial s_j}. \quad (1.31)$$

Теперь можно записать (1.25) в раскрытом виде:

$$\Delta w_{ij}^{(n)} = -\eta \cdot \delta_j^{(n)} y_i^{(n-1)}. \quad (1.32)$$

Иногда для придания процессу коррекции весов некоторой инерционности, сглаживающей резкие скачки при перемещении по поверхности целевой функции, (1.32) дополняется значением изменения веса на предыдущей итерации

$$\Delta w_{ij}^{(n)}(t) = -\eta(\mu \Delta w_{ij}^{(n)}(t-1) + (1-\mu) \delta_j^{(n)} y_i^{(n-1)}), \quad (1.33)$$

где μ – коэффициент инерционности, t – номер текущей итерации.

Таким образом, полный алгоритм обучения НС с помощью процедуры обратного распространения строится в следующем порядке [100]:

1. Подать на входы сети один из возможных образов и в режиме обычного функционирования НС, когда сигналы распространяются от входов к выходам, рассчитать значения последних. Как было отмечено ранее

$$s_j^{(n)} = \sum_{i=0}^M y_i^{(n-1)} w_{ij}^{(n)}, \quad (1.34)$$

где M – число нейронов в слое $(n-1)$ с учетом нейрона с постоянным выходным состоянием $+1$, задающего смещение; $y_i^{(n-1)} = x_{ij}^{(n)} - i$ -й вход нейрона j слоя n .

$$y_j^{(n)} = f(s_j^{(n)}), \quad (1.35)$$

где $f()$ – сигмоид.

$$y_q^{(0)} = I_q, \quad (1.36)$$

где I_q – q -я компонента вектора входного образа.

2. Рассчитать $\delta^{(N)}$ для выходного слоя по формуле (1.31).

Рассчитать по формуле (1.32) или (1.33) изменения весов $\Delta w^{(N)}$ слоя N .

3. Рассчитать по формулам (1.30) и (1.32) (или (1.30) и (1.33)) соответственно $\delta^{(n)}$ и $\Delta w^{(n)}$ для всех остальных слоев, $n = N-1, \dots, 1$.

4. Скорректировать все веса в НС

$$w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + \Delta w_{ij}^{(n)}(t). \quad (1.37)$$

5. Если ошибка сети существенна, перейти на шаг 1. В противном случае – конец.

На шаге 1, на входы сети попеременно в случайном порядке предъявляются все тренировочные образы, чтобы сеть, образно говоря, не забывала одни по мере запоминания других. Рассмотренный алгоритм иллюстрируется рис. 1.11.

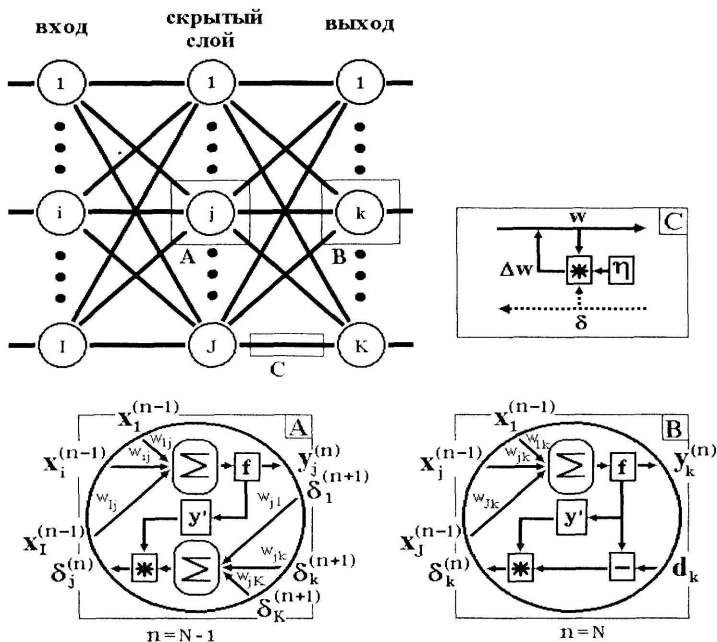


Рис. 1.11. Диаграмма сигналов в сети при обучении по алгоритму обратного распространения

Из выражения (1.32) следует, что, когда выходное значение $y_i^{(n-1)}$ стремится к нулю, эффективность обучения заметно снижается. При двоичных входных векторах в среднем половина весовых коэффициентов не будет корректироваться [110], поэтому область возможных значений выходов нейронов $[0,1]$ желательно сдвинуть в пределы $[-0.5, +0.5]$, что достигается простыми модификациями логистических функций. Например, сигмоид с экспонентой преобразуется к виду

$$f(x) = -0.5 + \frac{1}{1 + e^{-\alpha x}}. \quad (1.38)$$

Рассматриваемый алгоритм обучения НС имеет несколько "узких мест". Во-первых, в процессе обучения может возникнуть ситуация, когда большие положительные или отрицательные значения весовых коэффициентов сместят рабочую точку на сигмоидах многих нейронов в область насыщения. Малые величины производной от логистической функции приведут в соответствии с (1.30) и (1.31) к остановке обучения, что парализует НС. Во-вторых, применение метода градиентного спуска не гаран-

тирует, что будет найден глобальный, а не локальный минимум целевой функции. Эта проблема связана еще с одной, а именно – с выбором величины скорости обучения.

Доказательство сходимости обучения в процессе обратного распространения основано на вычислении производных, то есть приращении весов и, следовательно, скорость обучения должна быть бесконечно малой, однако в этом случае обучение будет происходить неприемлемо медленно. С другой стороны, слишком большие коррекции весов могут привести к постоянной неустойчивости процесса обучения. Поэтому в качестве η обычно выбирается число меньше 1, но не очень маленькое, например 0.1, и оно, в общем случае, может постепенно уменьшаться в процессе обучения. Кроме того, для исключения случайных попаданий в локальные минимумы иногда, после того как значения весовых коэффициентов стабилизируются, η кратковременно сильно увеличивают, чтобы начать градиентный спуск из новой точки. Если повторение этой процедуры несколько раз приведет алгоритм в одно и то же состояние НС, можно более или менее уверенно сказать, что найден глобальный минимум.

Алгоритмы модифицирующие процедуру обратного распространения, изменяющие процедуру выбора оптимизируемой функции критерия, выбор оптимального шага и правила коррекции параметров обучаемой сети подробно описаны в [53, 75].

Отметим, что все эти алгоритмы не согласуются с данными нейробиологии, где показано, что сигналы в биологических нейронных сетях могут распространяться только в одном, прямом направлении.

В работе [1] предложены алгоритмы настройки многослойной НС без использования процедуры обратного распространения ошибки, которые получили название “автономные алгоритмы”. Эти алгоритмы основаны на том факте, что для многослойной НС существует некоторое множество значений настраиваемых параметров, на котором вектор коррекции настраиваемых параметров вычисляется достаточно просто. Эти алгоритмы при прочих равных условиях хуже решают задачу обучения, чем алгоритмы обратного распространения ошибки, однако они являются структурно независимыми и очень простыми с вычислительной точки зрения.

1.3.2. Обучение без учителя

Обучение без учителя является более правдоподобной моделью обучения в биологической системе. Процесс обучения, как и в случае обучения с учителем, заключается в подстраивании весов синапсов. Некоторые ал-

горитмы предусматривают изменение и структуры сети, то есть количество нейронов и их взаимосвязи, но такие преобразования правильнее называть более широким термином – самоорганизацией. Очевидно, что подстройка синапсов может проводиться только на основании информации, доступной нейрону, то есть его состояния и уже имеющихся весовых коэффициентов. Исходя из этого соображения и, что более важно, по аналогии с известными принципами самоорганизации нервных клеток, построены алгоритмы обучения Хебба [32].

Сигнальный метод обучения Хебба заключается в изменении весов по следующему правилу:

$$w_{ij}(t) = w_{ij}(t-1) + \alpha y_i^{(n-1)} y_j^{(n)}, \quad (1.39)$$

где $y_i^{(n-1)}$ – выходное значение i -го нейрона $(n-1)$ -го слоя, $y_j^{(n)}$ – выходное значение j -го нейрона слоя n ; $w_{ij}(t)$ и $w_{ij}(t-1)$ – весовой коэффициент синапса, соединяющего эти нейроны, на итерациях t и $t-1$ соответственно; α – коэффициент скорости обучения. Здесь и далее для общности под n подразумевается произвольный слой сети. При обучении по данному методу усиливаются связи между возбужденными нейронами.

Существует также и дифференциальный метод обучения Хебба.

$$w_{ij}(t) = w_{ij}(t-1) + \alpha \cdot [y_i^{(n-1)}(t) - y_i^{(n-1)}(t-1)] \cdot [y_j^{(n)}(t) - y_j^{(n)}(t-1)]. \quad (1.40)$$

Здесь $y_i^{(n-1)}(t)$ и $y_i^{(n-1)}(t-1)$ – выходное значение i -го нейрона $(n-1)$ -го слоя соответственно на итерациях t и $t-1$; $y_j^{(n)}(t)$ и $y_j^{(n)}(t-1)$ – то же самое для j -го нейрона слоя n . Как видно из формулы (1.40), сильнее всего обучаются синапсы, соединяющие те нейроны, выходы которых наиболее динамично изменились в сторону увеличения.

Полный алгоритм обучения в соответствии с методом Хебба может быть представлен в следующем виде:

Шаг 1. На стадии инициализации всем весовым коэффициентам присваиваются небольшие случайные значения.

Шаг 2. На входы сети подается входной образ, и сигналы возбуждения распространяются по всем слоям согласно принципам классических прямопоточных (feedforward) сетей, то есть для каждого нейрона рассчитывается взвешенная сумма его входов, к которой затем применяется активационная (передаточная) функция нейрона, в результате чего получается его выходное значение $y_i^{(n)}$, $i = 0 \dots (M_i - 1)$, где M_i – число нейронов в слое i ; $n = 0 \dots (N - 1)$, а N – число слоев в сети.

Шаг 3. На основании полученных выходных значений нейронов по формуле (1.39) или (1.40) производится изменение весовых коэффициентов.

Шаг 4. Если выходные значения сети не стабилизируются с заданной точностью, то осуществляется повторение цикла, начиная с шага 2. Применение этого нового способа определения завершения обучения, отличного от использовавшегося для сети обратного распространения, обусловлено тем, что подстраиваемые значения синапсов фактически не ограничены.

На втором шаге цикла попеременно предъявляются все образы из входного набора.

Следует отметить, что вид откликов на каждый класс входных образов не известен заранее и будет представлять собой произвольное сочетание состояний нейронов выходного слоя, обусловленное случайным распределением весов на стадии инициализации. Вместе с тем сеть способна обобщать схожие образы, относя их к одному классу. Тестирование обученной сети позволяет определить топологию классов в выходном слое. Для приведения откликов обученной сети к удобному представлению сеть может быть дополнена еще одним слоем, который в процессе обучения по алгоритму обучения однослойного перцептрона [32] позволяет отображать выходные реакции сети в требуемом представлении.

Другой алгоритм обучения без учителя – алгоритм Кохонена – предусматривает подстройку синапсов на основании их значений от предыдущей итерации.

$$w_{ij}(t) = w_{ij}(t-1) + \alpha [y_i^{(n-1)} - w_{ij}(t-1)]. \quad (1.41)$$

Из вышеприведенной формулы видно, что обучение сводится к минимизации разницы между входными сигналами нейрона, поступающими с выходов нейронов предыдущего слоя $y_i^{(n-1)}$, и весовыми коэффициентами его синапсов.

Полный алгоритм обучения имеет примерно такую же структуру, как в методах Хейбба, но на третьем шаге из всего слоя выбирается нейрон, значения синапсов которого максимально похожи на входной образ, и подстройка весов по формуле (1.41) проводится только для него. Эта так называемая аккредитация может сопровождаться затормаживанием всех остальных нейронов слоя и введением выбранного нейрона в насыщение. Выбор такого нейрона может осуществляться, например, расчетом скалярного произведения вектора весовых коэффициентов с вектором входных значений. Максимальное произведение дает выигравший нейрон.

Другой вариант – расчет расстояния между этими векторами в p -мерном пространстве

$$D_j = \sqrt{\sum_{i=0}^{p-1} (y_i^{(n-1)} - w_{ij})^2}, \quad (1.42)$$

где p – размерность векторов, j – индекс нейрона в слое n , i – индекс суммирования по нейронам $(n-1)$ -го слоя, w_{ij} – вес синапса, соединяющего нейроны; выходы нейронов $(n-1)$ -го слоя являются входными значениями для слоя n . Корень в формуле (1.42) брать не обязательно, так как важна лишь относительная оценка различных D_j .

В данном случае, "побеждает" нейрон с наименьшим расстоянием. Иногда слишком часто получающие аккредитацию нейроны принудительно исключаются из рассмотрения, чтобы "уравнять права" всех нейронов слоя. Простейший вариант такого алгоритма заключается в торможении только что выигравшего нейрона.

При практическом использовании для обучения алгоритма Кохонена, как правило, осуществляется нормализация входных образов. При этом на стадии инициализации осуществляется нормализация начальных значений весовых коэффициентов

$$x_i = \frac{x_i}{\sqrt{\sum_{j=0}^{n-1} x_j^2}}, \quad (1.43)$$

где x_i – i -я компонента вектора входного образа или вектора весовых коэффициентов, а n – его размерность. Такой подход позволяет сократить длительность процесса обучения.

Инициализация весовых коэффициентов случайными значениями может привести к тому, что различные классы, которым соответствуют плотно распределенные входные образы, сольются или, наоборот, раздробятся на дополнительные подклассы в случае близких образов одного и того же класса. Для избежания такой ситуации используется метод выпуклой комбинации [81]. Суть его сводится к тому, что входные нормализованные образы подвергаются преобразованию:

$$x_i = \alpha(t)x_i + (1 - \alpha(t))\frac{1}{\sqrt{n}}, \quad (1.44)$$

где x_i – i -я компонента входного образа, а n – общее число его компонент, $\alpha(t)$ – коэффициент, изменяющийся в процессе обучения от нуля до единицы, в результате чего вначале на входы сети подаются практически одинаковые образы, а с течением времени они все больше сходятся к исходным. Весовые коэффициенты устанавливаются на шаге инициализации равными величине

$$w_o = \frac{1}{\sqrt{n}}, \quad (1.45)$$

где n – размерность вектора весов для нейронов инициализируемого слоя.

На основе рассмотренного выше метода строятся нейронные сети особого типа – так называемые самоорганизующиеся структуры – self-organizing feature maps. Однако речь идет не об изменении структуры сети, а только о подстройке синапсов. Для них после выбора из слоя n j -го нейрона с минимальным расстоянием D_j обучается по формуле (1.41) не только этот нейрон, но и его соседи, расположенные в окрестности R . Величина R на первых итерациях очень большая, так что обучаются все нейроны, но с течением времени она уменьшается до нуля. Таким образом, чем ближе конец обучения, тем точнее определяется группа нейронов, отвечающих каждому классу образов.

Другим распространенным методом обучения "без учителя" является обучение методом соревнования. В отличие от правила обучения Хебба, в котором множество выходных нейронов могут возбуждаться одновременно, при соревновательном обучении выходные нейроны соревнуются между собой за активизацию. Это явление известно как правило "победитель берет все". Подобное обучение имеет место в биологических нейронных сетях. Обучение посредством соревнования позволяет кластеризовать входные данные: подобные примеры группируются сетью в соответствии с корреляцией и представляются одним элементом.

При обучении модифицируются только веса "победившего" нейрона. Эффект этого правила достигается за счет такого изменения сохраненного в сети образца (вектора весов связей победившего нейрона), при котором он становится чуть ближе к входному примеру. Необходимо заметить, что сеть никогда не перестанет обучаться, если параметр скорости обучения не равен нулю. Некоторый входной образец может активизировать другой выходной нейрон на последующих итерациях в процессе обучения. Это ставит вопрос об устойчивости обучающей системы. Система считается устойчивой, если ни один из примеров обучающей выборки не изменяет своей принадлежности к категории после конечного числа итераций обучающего процесса. Один из способов достижения стабильности состоит в постепенном уменьшении до нуля параметра скорости обучения. Однако это искусственное торможение обучения вызывает другую проблему, называемую пластичностью и связанную со способностью к адаптации к новым данным. Эти особенности обучения методом соревнования известны под названием дилеммы стабильности-пластичности Гроссберга [6].

Кроме рассмотренных, следует упомянуть некоторые другие алгоритмы: Adaline и Madaline, линейный дискриминантный анализ, проекции Саммона [14], анализ главных компонентов [74].

Необходимо отметить, что обучение "без учителя" гораздо более чувствительно к выбору оптимальных параметров, нежели обучение "с учителем". Во-первых, его качество сильно зависит от начальных весовых величин синапсов. Во-вторых, обучение критично к выбору радиуса обучения и скорости его изменения. И, наконец, разумеется, очень важен характер изменения собственно коэффициента обучения. В связи с этим потребуется провести предварительную работу по подбору оптимальных параметров обучения сети.

Несмотря на некоторые сложности реализации, алгоритмы обучения "без учителя" находят обширное и успешное применение. Например, в [6] описана многослойная нейронная сеть, которая по алгоритму самоорганизующейся структуры обучается распознавать рукописные символы. Возникающее после обучения разбисие на классы может в случае необходимости уточняться с помощью обучения "с учителем". По сути дела, по алгоритму обучения "без учителя" функционируют и наиболее сложные из известных на сегодняшний день искусственные нейронные сети – когнитрон и неокогнитрон, максимально приблизившиеся в своем воплощении к структуре мозга. Однако они, конечно, существенно отличаются от рассмотренных выше сетей и намного более сложны. Тем не менее на основе вышеизложенного материала можно создать реально действующие системы для распознавания образов, сжатия информации, автоматизированного управления, экспертных оценок и других приложений.

1.4. Нейрочипы и нейрокомпьютеры

Быстрое развитие и постоянно расширяющееся внедрение нейросетевых методов и алгоритмов как в коммерческие так и в военные приложения способствует росту устойчивого интереса к созданию средств аппаратной реализации нейрокомпьютерной (НК) технологии. Прежде всего это относится к разработке функционально сложных сверхбольших интегральных схем (СБИС) со специализированной структурой, которая ориентирована на аппаратную реализацию одной или нескольких наиболее часто используемых парадигм нейронных сетей.

Общее определение нейрокомпьютера может быть сформулировано следующим образом [12]. Нейрокомпьютер – это вычислительная система с архитектурой аппаратного и программного обеспечения, адекватной выполнению алгоритмов, представленных в нейросетевом логическом базисе.

Системы, созданные на стандартной микропроцессорной элементной базе могут являться эффективными программными эмуляторами нейросетевых алгоритмов. Однако программная эмуляция нейросетевых алгоритмов на вычислительных средствах, реализованных на элементной базе, не имеющей отношения к нейросетевому логическому базису, скорее всего, является временным явлением. Для аппаратной реализации таких алгоритмов эффективнее создавать архитектуры, выполняющие операции в нейросетевом базисе – на основе логической сети нейронов при полном отказе от булевских элементов типа И, ИЛИ, НЕ.

Ограничиваясь проблемами использования в современных разрабатываемых устройствах НС и НК полупроводниковых технологий, можно выделить следующие трудности возникающие при их создании:

- высокая структурная сложность и низкая надежность систем;
- практическая трудность реализации архитектур обучения, самообучения, самоорганизации НС для непрерывно изменяемых весов взаимодействия нейронов;

- технологические сложности изготовления сверхинтегрированных НС и НК на нейрочипах и подсистемах кремний на пластине;

- увеличение потребляемой мощности и потеря системного быстродействия в подсистемах НС по мере увеличения степени их интеграции.

Выбор эффективного типа нейрокомпьютера для решения конкретной задачи или конкретного класса задач рекомендуется производить в определенной последовательности [10].

На первом этапе в общепринятых терминах формулируется математическая постановка задачи.

На втором этапе осуществляется переформулировка задачи в терминах системной постановки, требуемой для ее решения на нейрокомпьютере. При этом должно быть выделено “нейронное ядро” задачи, т. е. часть задачи адекватная нейронному базису. Для более эффективного использования нейрокомпьютеров это ядро должно составлять основную, как можно большую, долю множества операций в алгоритме, представляющем собой общую математическую формулировку задачи.

На третьем этапе для части алгоритма задачи, представляющего нейронное ядро, формулируются основные характеристики системной постановки задачи:

- входной сигнал;
- выходной сигнал;
- желаемый выходной сигнал;
- структура разомкнутой системы;
- ошибка системы;
- функционал оптимизации системы;
- градиент функционала оптимизации;

алгоритм настройки системы в процессе решения задачи; замкнутая система.

На четвертом этапе для системной постановки конкретной задачи заказчика производится подсчет необходимых ресурсов ее решения в структурных терминах НС: число эмулируемых нейронов; число эмулируемых связей; число переключений связей в процессе решения задачи нейронным ядром алгоритма.

Пятый этап требует проведения моделирования процесса решения задачи на ПЭВМ.

На шестом этапе производится оценка производительности НК на конкретной задаче с выбором конкретной технологии реализации.

Примеры существующих нейрокомпьютеров в порядке увеличения их производительности следующие [9]:

- программная эмуляция на однопроцессорной ЭВМ;

- транспьютерная реализация различной конфигурации транспьютерной сети;

- реализация нейронной архитектуры на стандартных микропроцессорных СБИС;

- реализация на аналоговых и аналого-цифровых СБИС;

- реализация на оптоэлектронных СБИС;

- реализация на пластине.

Несмотря на всплеск исследований и разработок полупроводниковых, оптических, оптоэлектронных, магнитоэлектронных, биомолекулярных НС и НК, вопрос выбора приоритетных и наиболее конкурентоспособных направлений до сих пор открыт.

Разработка функционально сложных СБИС со специализированной структурой, ориентированной на аппаратную реализацию одной или нескольких наиболее часто используемых парадигм нейронных сетей привела к созданию параллельных нейросетевых ускорительных плат для различных базовых вычислительных платформ. Примером такой реализации является разработка фирмой Telmat (Франция) нейросетевой ускорительной платы на базе специализированной СБИС L-Neuro [22, 52, 89].

СБИС L-Neuro — цифровая каскадируемая схема для моделирования дискретных нейросетевых алгоритмов, выполненная по 1,5 мк КМОП-технологии. Она совместима с интерфейсом памяти транспьютера фирмы Inmos и аппаратно реализует параллельное выполнение базовых нейросетевых операций в режимах распознавания и обучения для нейрона, имеющего 16 входов и один выход. Наличие локальной памяти весовых коэффициентов объемом 1 Кбайт с произвольной выборкой, и 32-байтного регистра состояния нейронов, обеспечивает возможность моделирования нейросетевых структур от фрагментов из 32 нейронов с 32 бай-

товыми входами до фрагментов из четырех нейронов с 256 битовыми входами. Эффективную реализацию различных нейросетевых парадигм и правил обучения обеспечивает набор из 47 команд СБИС L-Neuro.

В составе СБИС L-Neuro можно выделить следующие структурные элементы:

массив локальной памяти весовых коэффициентов (Synaptic RAM).
Время доступа к памяти со стороны внешних устройств (по адресной мультиплексной шине транспьютера) составляет 150 нс;

два блока регистров состояния нейрона (Neural State Register (NSR)).
В каждом блоке содержится 16 универсальных байтовых регистров, обеспечивающих хранение и модификацию в синхронном и асинхронном режимах значений выходов нейронов;

блок процессоров и регистров обучения (Learning Register), содержащий 16 регистров, каждый из которых обеспечивает хранение, а также выполнение операций сложения и вычитания 16-битовых операндов, представляющих собой значения весовых коэффициентов.

блок умножителей, состоящий из 16 схем умножения байтового операнда на бит (AND). С помощью данного блока вычисляются элементарные произведения для нахождения взвешенной суммы.

блок накопительного сумматора, обеспечивающий сложение частных сумм и накопление результата в накопительном регистре, в случае вычисления выходного значения нейрона, имеющего более 16 входов, или значения выхода более 1 бита;

блок дешифрации команд и микропрограммного управления, обеспечивающий выполнение 19-разрядных команд и микропрограмм.

Структурная схема СБИС L-Neuro представлена на рис. 1.12.

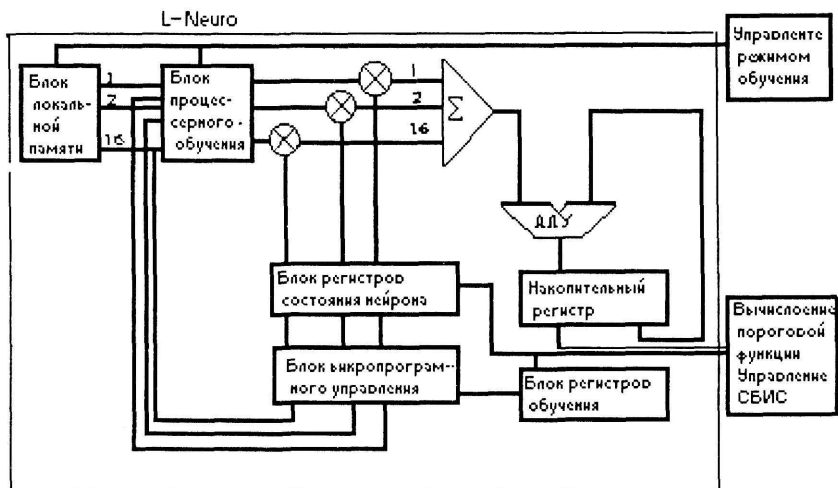


Рис. 1.12. Структурная схема СБИС L-Neuro

СБИС L-Neuro имеет два основных режима функционирования: обучение и распознавание (раскодирование). Выполнение вычислительных операций в данных режимах осуществляется с помощью микропрограммного управления. Кроме того, предусмотрено выполнение отдельных функций вычислительного процесса с помощью задания управляющих параметров, размещаемых на специально выделенных регистрах управления.

СБИС L-Neuro является основным блоком, на базе которого могут создаваться различные параллельные нейровычислительные ускорительные платы и супернейрокомпьютеры. Высокая степень параллелизма, естественно присущая нейросетевым алгоритмам, и каскадируемость СБИС L-Neuro обеспечивают возможность произвольного наращивания числа процессоров для достижения требуемой вычислительной мощности. Межпроцессорные взаимосвязи могут быть организованы в зависимости от функционального назначения и моделируемых нейросетевых парадигм. Структура взаимосвязи четырех СБИС L-Neuro для моделирования полностью связанной сети приведена на рис. 1.13.

На базе СБИС L-Neuro в настоящее время разработано два типа плат:

ускорительная плата для IBM PC совместимых компьютеров, содержащая 1 управляющий транспьютер, 4 СБИС L-Neuro, 128 Кбайт RAM;

ускорительная плата для вычислительной системы T-Node Computer фирмы Telmat Multinode, содержащая 1 управляющий транспьютер, 4 СБИС L-Neuro и 4 Мбайт RAM.

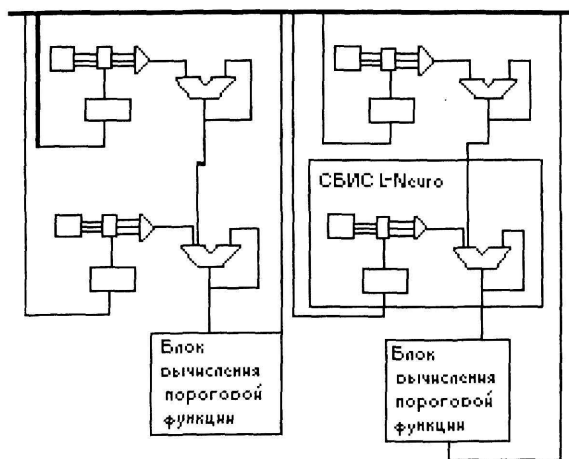


Рис. 1.13. Структура взаимосвязи четырех СБИС L-Neuro для моделирования полносвязной сети

Одним из способов реализации НС в виде СБИС является использование приборов с зарядовой связью (ПЗС). И хотя данный способ имеет существенный недостаток – низкое быстродействие серийных ПЗС, в настоящее время разработано несколько схем, свидетельствующих об успешной конкуренции реализованных на ПЗС нейронных сетей при решении некоторого класса задач [113]. Так, например, решение задач распознавания образов не требует значительного быстродействия, но предполагает увеличение размерности НС и объема обрабатываемой информации. ПЗС могут удовлетворить эти требования, так как, кроме отмеченного выше недостатка, они имеют и целый ряд преимуществ перед другими способами реализации.

Во-первых, элементарная ячейка ПЗС (МОП/МНОП конденсатор) имеет минимальные топологические размеры, и следовательно возможно создание СБИС с наивысшей степенью интеграции.

Во-вторых, ПЗС могут обрабатывать как аналоговую, так и цифровую информацию с одинаковым быстродействием.

В-третьих, при использовании ПЗС со скрытым каналом, передачу информации возможно осуществлять параллельно по нескольким слоям. Таким образом, по сравнению с другими способами реализации ПЗС имеют повышенное быстродействие.

В-четвертых, использование ПЗС детекторов во многом снимает ограничения связанные с конечным числом выводов микросхем, что снижает время необходимое на инициирование системы.

Кроме того, использование ПЗС позволяет создавать приборы, совместимые с любым другим способом реализации НС. Однако, как уже было отмечено, применение ПЗС для реализации НС возможно только в том случае, если данная элементная база удовлетворяет условиям конкретной решаемой задачи.

Использование субмикронных технологий и минипроизводств, новых схмотехнических и системотехнических принципов построения маломощных сверхбыстродействующих ультраБИС (УБИС) позволило японской фирме "Hitachi" в 1990 г. разработать и изготовить НС на кремниевой пластине диаметром 5" по 0,8-микрометровой КМОП-технологии [7]. Пластина была смонтирована на керамической подложке 15×15 см. На этой кремниевой пластине было размещено 19 млн. МОП-транзисторов в 64 нейрончиках. Такие БИС, использованные в качестве базовой ячейки на пластине, позволяли создать НС из 576 формальных нейронов (ФН) и 36 864 синапсов. В применяемом базовом матричном кристалле (БМК) размером 12×12 мм было размещено 12 ФН. Каждый нейрон представлялся 1000 КМОП-элементами с памятью 64 слова по 18 бит.

На рис. 1.14 представлена цифровая НС на пластине, эквивалентная сети Хопфилда, на основе архитектуры шины с разделением времени [7].

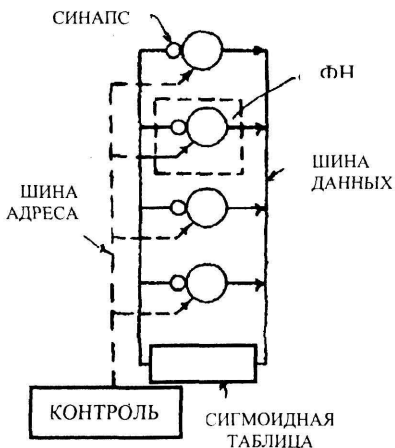


Рис 1.14. Структурная схема НС, эквивалентной сети Хопфилда, с использованием шины данных с разделением времени

На рис. 1.15 дана схема ФН, функционирование которого описывается стандартными выражениями и реализуется с помощью стандартных булевых элементов и ячеек памяти [7].

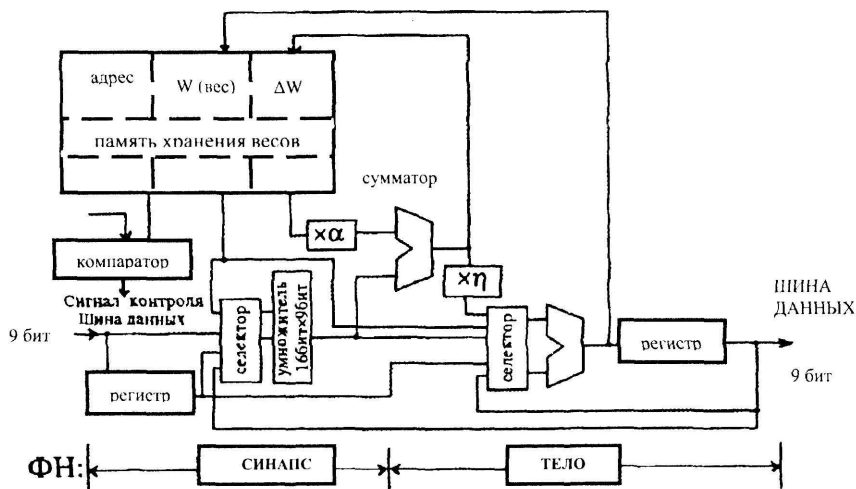


Рис. 1.15. Структурная схема формального нейрона, реализованного с помощью стандартных булевых элементов и ячеек памяти

Предложенная архитектура шины данных позволяет резко сократить число соединений между ФН внутри и вне нейрочипа. Весовые коэффициенты ФН изменяются в диапазоне 8 бит, выходные сигналы кодируются 9-битовыми числами (см. рис. 1.14), а шина с разделением времени обеспечивает асинхронный обмен данными между ячейками (см. рис. 1.15).

НС на пластине может быть достаточно просто соединена с управляющим внешним компьютером и использована для широкого спектра интеллектуальных задач в рамках высокоорганизованных систем искусственного интеллекта. Например, задача коммивояжера для 16 городов была решена с помощью созданных НС и НК, работающего с временным циклом 267 мкс, за 0,1 с, причем скорость вычисления и решения подобных задач превышала на порядок возможности суперкомпьютера фирмы "Hitachi" производительностью свыше миллиарда операций в секунду [7].

Затем в 1992–1993 гг. в рамках использования эффективной архитектуры шины данных с разделением времени (см. рис. 1.14) разработчики

фирмы заложили в проектируемые НС возможности обучения и самообучения. Такая возможность появилась в измененной схеме всех используемых в НС новых ФН за счет дополнительной подстройки весовых коэффициентов (рис. 1.16).

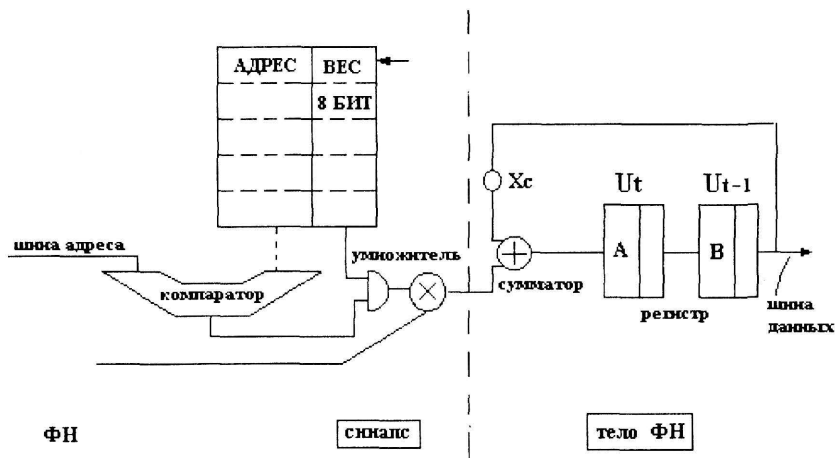


Рис. 1.16. Модифицированная структурная схема формального нейрона с возможностью обучения и самообучения за счет дополнительной подстройки весовых коэффициентов

На рис. 1.17 представлена структура полной синтезированной НС с управляющей высокопроизводительной рабочей станцией, которая контролирует функционирование и необходимые операции всей системы. Плата контроля накапливает и сохраняет данные входных и выходных слоев и базовой микропрограммы для последовательности обучающих шагов.

Сегодня уже имеется впечатляющая демонстрация возможностей трехмерных систем искусственных многослойных обучаемых перцептронных НС, реализованных на пластинах, с оценкой общей производительности системы в классе распознавания образов в 24 биллиона переключении/с [50].

Наиболее изученный мощный адаптивный алгоритм для обучения многослойных перцептронных НС – метод обратного распространения, использованный в 3D-НС, имел производительность 1,1 биллиона переключении/с для матричного воплощения элементарных процессов 128×128 и давал оценку более впечатляющего качества трехмерных сис-

тем ближайшего будущего для матричного воплощения 512×512 около 54 биллионов переключении/с [50].

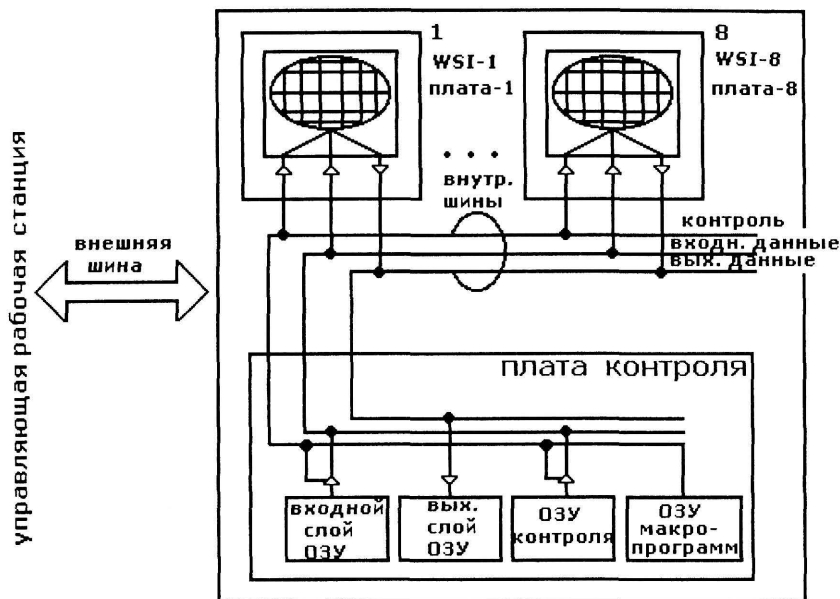


Рис. 1.17. Структурная схема нейросети на кремниевых пластинах с управляющей высокопроизводительной рабочей станцией, контролирующей функционирование и необходимые операции всей системы

Потенциальная возможность быстрого обучения, переобучения, самообучения НС вместе с высокими скоростными характеристиками обработки информации (зависящими от "плотности" представления собственно нейрона перспективного типа, его быстродействия и т.д., а также числа нейронов в НС, т. е. от уровня интеграции 3D-НС) является основным залогом возможности быстрого и качественного решения задач распознавания образов и сцен в реальном масштабе времени с большим числом имеющихся параметров. Более того, плотно упакованные системы из 3D-НС, построенные из одинаковых элементов с определенной избыточностью, позволяют весьма эффективно выходить из ситуации локальной неработоспособности отдельных элементов или фрагментов НС за счет

дополнительной перестройки соответствующих весов еще работающих элементов так, что, хотя и с потерей в точности, надежности и системной скорости, НС будет продолжать функционировать.

При уменьшении линейных размеров линий связи (ЛС) до 1 мкм и в субмикронной области для увеличения плотности компоновки систем, резко ухудшается их функционирование из-за возрастания сопротивлений ЛС, связанного с этим резкого возрастания времени распространения в них сигнала, а также из-за уменьшения срока службы ЛС вследствие электромиграции и других нежелательных эффектов (например, термических), связанных с высокими плотностями тока в ЛС. Поэтому оптимальным и наиболее кардинальным выходом из ситуации, обусловленной "отказом" длинных межсоединений на пластине и в 3D-системах, является переход к чисто оптическим связям между чипами (в отдаленной перспективе и внутри чипа), несмотря на казалось бы избыточную необходимость преобразования электрических сигналов в оптические, и, наоборот, в 3D-НС.

Использование оптических связей в симбиозных 3D-оптоэлектронных НС позволяет:

- увеличить уровень интеграции и системного быстродействия систем;
- получить максимальную быстроту передачи информации, ограниченную скоростью света;
- повысить помехоустойчивость ЛС и системы в целом;
- увеличить срок службы за счет увеличения надежности системы;
- повысить эффективность возможности параллельного ввода, передачи и обработки информации.

В то же время обработка больших массивов информации в реальном масштабе времени в будущих оптоэлектронных 3D-НС обостряет проблему создания сверхбыстродействующих микроминиатюрных оптоэлектронных компонент, совместимых с субмикронными КМОП, КБИ/КБИКМОП-элементами на кремниевых пластинах. К середине 90-х годов все более широкое внимание стало уделяться гибридной Si-GaAs-технологии применительно к 2D- и 3D-НС [7] с усовершенствованием методов и средств интеграции на кремниевой подложке оптоэлектронных приборов (излучателей, модуляторов, фотоприемников) на основе слоев GaAs и традиционных микроэлектронных компонентов: *n*-МОП, *p*-МОП, *n-p-n*, *p-n-p* транзисторов, резисторов, емкостей, созданных в рамках субмикронных кремниевых технологий.

Соответствующие фотоприемники (фотодиоды, фототранзисторы) в каждом чипе в 3D-НС преобразуют оптические сигналы от чипа к чипу в электрические и служат конечными элементами оптических соединений и ЛС. В 3D-НС оптические параметры миниатюрных фотоприемников (спектральная характеристика чувствительности, апертура, площадь све-

точувствительной площадки и др.) согласуются с оптическим каналом лазерных диодов, а электрические характеристики – с характеристиками последующих логических элементов КМОП, а также БИКМОП, КБИ/КБИКМОП схем.

Наиболее перспективной с точки зрения снижения потребляемой мощности и увеличения интеграции 3D-НС представляется организация оптических связей с внешней модуляцией периферийных источников света – лазерных диодов, что приводит к практическому исключению рассеиваемой мощности лазеров на периферии чипов (эквивалентов выходных усилителей мощности в стандартных электронных БИС) и соответственно к существенному ослаблению проблем, связанных с нагрузочной способностью выходов оптомикроэлектронных УБИС.

Если в архитектурах 3D-НС и НК удастся не только уменьшить площади периферийных лазеров, но и их число, то это приведет к значительному улучшению технико-экономических характеристик системы, в первую очередь, по снижению потребляемой мощности, уровню интеграции, а также системному быстродействию. По мере дальнейшего технологического прогресса масштабирования микроэлектронных и оптомикроэлектронных компонент и улучшения их характеристик в 3D-НС можно ожидать наиболее полной реализации преимуществ оптоэлектроники для передачи информации между нейрочипами и нейропластинами в НК.

Перспективным материалом для оптоэлектронных технологий НК и НС является наноструктурированный пористый кремний. Использование которого, благодаря его принципиальной совместимости со стандартной кремниевой технологией, позволяет повысить надежность производимых нейрочипов.

Еще одним направлением реализации нейросетевых структур является создание устройств биокомпьютинга. Качество и производительность устройств и систем биокомпьютинга, использующих химические реакции типа Белоусова–Жаботинского, ферментативные реакции с нелинейной кинетикой, сред сопряженных химических и биохимических реакций обеспечиваются не повышенным быстродействием исходных элементов, а динамикой и самоорганизацией биосистем и определяющимися ими информационными процессами.

Перспективным также является симбиоз методов и средств биокомпьютинга и микро- и оптоэлектронных принципов обработки информации на основе нейрочипов, 2D- и 3D-НС. Элементно-технологическим базисом таких симбиозных систем могут быть двумерные и трехмерные кремниевые пластины максимального диаметра, на которых будут размещаться как традиционные кремниевые чипы и нейрочипы, так и интеллектуальные биочипы предельной вычислительной сложности.

От успеха разработки, производства и эффективного внедрения систем нейрокомпьютинга, а также биокомпьютинга и симбиозных систем био-нейрокомпьютинга с поистине, неограниченными потенциальными вычислительными и логическими возможностями зависит решение жизненно важных для цивилизации и невероятно сложных для современной вычислительной техники задач.

Глава 2. Применение искусственных нейронных сетей в системах управления

2.1. Особенности использования искусственных нейронных сетей в системах управления

В историческом плане можно утверждать, что разработка систем управления (СУ) всегда происходит поэтапно. В качестве таких этапов можно выделить:

- этап разработки концепции построения СУ;
- этап моделирования СУ, в соответствии с предлагаемой концепцией построения;
- этап анализа получаемых результатов;
- этап доработки (модернизации) концепции построения СУ.

На протяжении всех этих этапов не прекращаются теоретические исследования, которые позволяют выбирать основные направления совершенствования первоначально сформулированной концепции построения СУ и распространять ее основные идеи на ряд смежных областей.

По аналогичной схеме происходит развитие СУ, использующих искусственные НС. Однако следует заметить, что большое количество разработанных аппаратных и программных моделей нейросетевых СУ часто опережают теоретическое понимание происходящих при этом процессов и имеющихся проблем [93].

Необходимым этапом решения задач управления нелинейными динамическими системами является получение их адекватных математических моделей, базирующееся, как правило, на теоретическом и экспериментальном анализе свойств этих систем. Теоретический анализ процессов, происходящих в системе, позволяет получить математическое описание в виде, например, дифференциальных уравнений. При экспериментальном анализе на основе наблюдений входных и выходных сигналов системы получают либо ее параметрическую, либо непараметрическую модель. Наиболее широкое распространение получили параметрические модели, требующие решения задач структурной и параметрической идентификации и использующие ограниченное число параметров. Несмотря на огромное количество работ, многообразие видов нелинейностей не позволяет создать единую теорию идентификации нелинейных систем. Применяемый чаще всего классический подход основан на аппроксимации нелинейностей, например рядами Вольтера, Гаммерштейна, Винера, полиномами Колмогорова-Габора и др. Однако область применения таких моделей ограничена. Кроме того, дополнительные трудности получения адекватного математического описания обусловлено наличием в реальных сигналах помех.

Одной из классических моделей СУ является модель с обратной связью с регулируемыми в реальном масштабе времени коэффициентами, например самонастраивающийся регулятор Астрома [43]. Коэффициенты такого контроллера регулируются в течении каждого цикла управления в соответствии с оценкой параметров системы. Блок-схема управления с обратной связью и регулируемыми в реальном масштабе времени коэффициентами приведена на рис. 2.1.

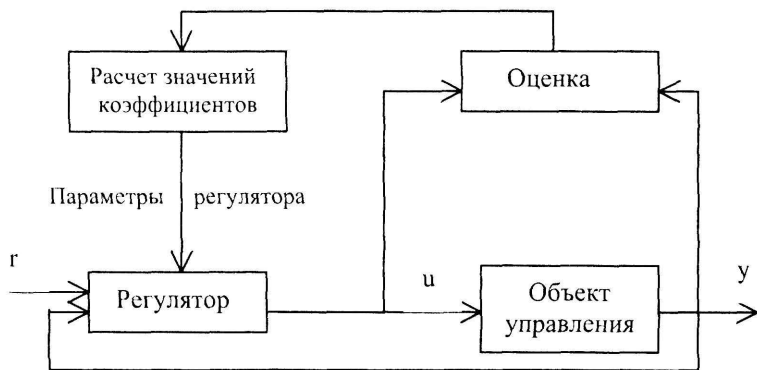


Рис. 2.1. Блок-схема управления с обратной связью и регулируемыми коэффициентами

Другой хорошо известной моделью СУ является модель Ляпунова. Системы адаптивного управления, использующие эталонную модель Ляпунова проектируются таким образом, чтобы выходной сигнал управляемой модели в конце концов соответствовал выходному сигналу предварительно определенной модели, которая имеет желаемые характеристики [96]. Такая система должна быть асимптотически устойчивой, то есть управляемая система в итоге отслеживает эталонную модель с нулевой ошибкой. Более того, переходные процессы на этапе адаптивного управления или обучающего управления имеют гарантированные пределы. Блок-схема адаптивного управления с эталонной моделью представлена на рис. 2.2.

Системы управления, так или иначе использующие искусственные НС являются одной из возможных альтернатив классическим методам управления. Возможность использования НС для решения задач управления во многом основывается на том, что НС, состоящая из двух слоев и имеющая в скрытом слое произвольное большое количество узлов, может аппроксимировать любую функцию действительных чисел с заданной степенью точности. Доказательство этого положения, основанное на известной тео-

реме Вейерштрасса, приведено в [79]. Таким образом, для решения задач идентификации и управления могут быть использованы НС даже с одним скрытым слоем.

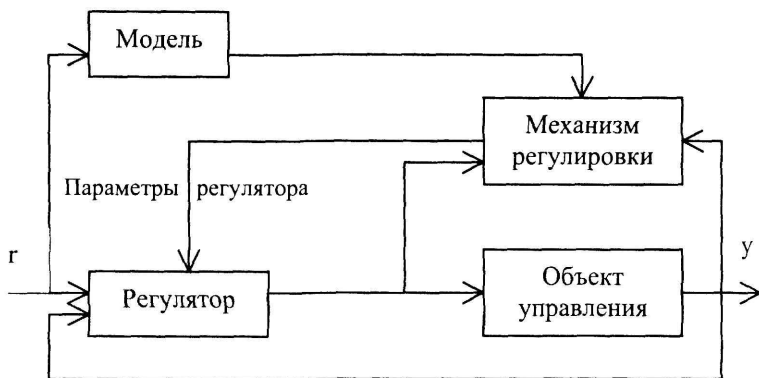


Рис. 2.2. Блок-схема адаптивного управления с эталонной моделью

Одним из первых используемых методов построения нейросетевых СУ был метод, основанный на “копировании” существующего контроллера. Применив этот метод в 1964 Уидроу назвал его методом построения экспертной системы за счет получения знаний от уже существующего эксперта. Архитектура такой СУ представлена на рис. 2.3.

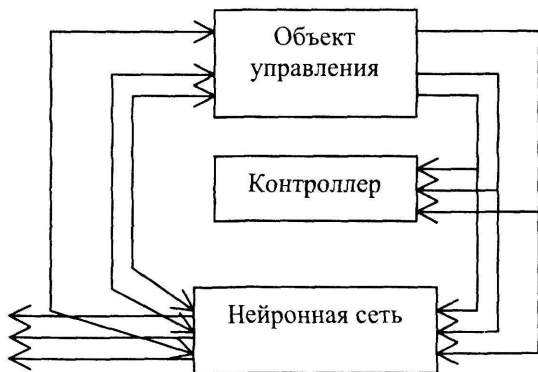


Рис. 2.3. Нейросетевая СУ, основанная на “копировании” существующего контроллера

Глядя на этот рисунок можно усомниться в полезности этого метода. Зачем нужно использовать еще один управляющий контроллер (в виде НС), если один уже существует? Однако, во-первых, существующий контроллер может быть неудобен при использовании (например, в роли такого контроллера может выступать человек), а во-вторых, для выработки эффективного управления НС может использовать отличную, от существующего контроллера, по способу представления (легче измерить, формализовать и т. д.) информацию о состоянии объекта управления.

В настоящее время достаточно хорошо разработан и широко используется целый ряд других возможных архитектур построения нейросетевых СУ [73, 93]. Во всех из них, назначением нейросетевого контроллера является выработка адекватного управляющего сигнала для управления динамикой состояний объекта управления от начального состояния до желаемого итогового состояния. Причем смена состояний должна происходить по оптимальной траектории. Организация контроля за состоянием объекта управления и реализация нейросетевого контроллера в значительной степени зависят от выбранного алгоритма обучения и используемой структуры управления. Наиболее широко используемыми являются схема прямого (непосредственного) управления и схема косвенного управления. При этом чаще всего в качестве алгоритма обучения используется алгоритм обратного распространения ошибки [73].

В схеме косвенного управления параметры объекта управления определяются в каждый момент времени и затем используются для оценки параметров контроллера (см. рис. 2.4). Таким образом, имеет место явно выраженный процесс идентификации.

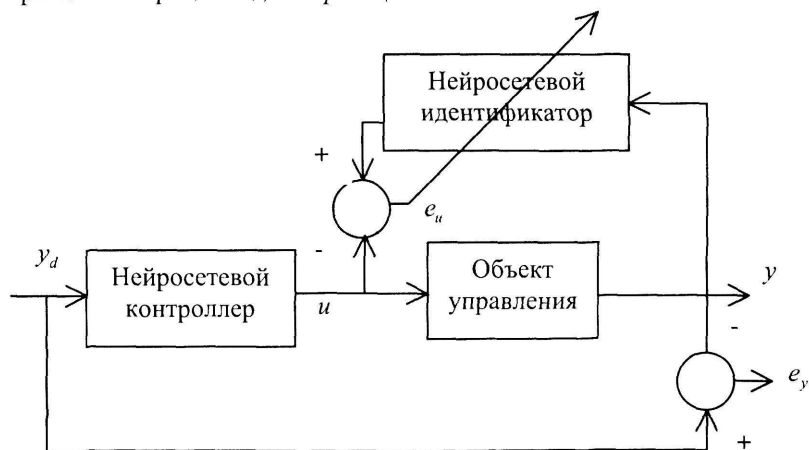


Рис. 2.4. Схема косвенного управления

Недостатком такой схемы является то, что идентификация и управление основываются исключительно на ошибке e_u , и, следовательно, минимизацию ошибки на выходе системы e_y гарантировать нельзя.

В схеме прямого управления параметры нейросетевого контроллера регулируются таким образом, чтобы уменьшить непосредственно ошибку выхода e_y (см. рис. 2.5).

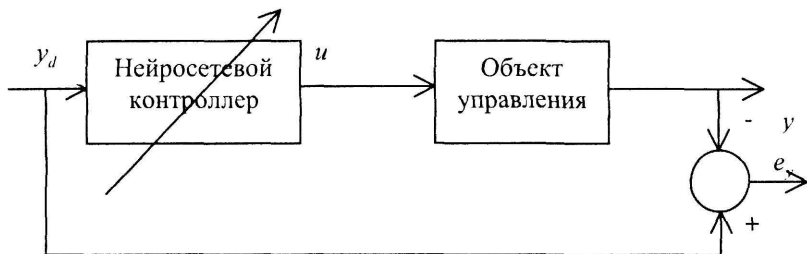


Рис. 2.5. Схема прямого управления

В качестве целевой функции, которая должна быть минимизирована контроллером используется среднеквадратическая ошибка на выходе объекта управления

$$E_y = \frac{1}{2}(y_d - y)^2. \quad (2.1)$$

Вопросы устойчивости и управляемости таких СУ подробно рассматриваются в работах [32, 70, 73, 78, 92, 93, 108].

Одной из областей теоретических исследований, рассматривающих проблемы использования НС в системах управления, является сравнение таких методов управления с другими известными типами СУ, выявление присущих нейросетевым методам особенностей и их анализ. В работе [93] дан краткий обзор характеристик упомянутых выше СУ. Сравнивались устойчивость, скорость сходимости, работа в условиях шумов, требуемый объем памяти и др. Хотя каждый из рассмотренных методов имеет как хорошие, так и плохие характеристики следует заметить, что метод нейросетевого управления имеет такие полезные свойства, которые плохо реализуются в двух других методах.

Основные результаты, полученные при сравнении, приведены в табл. 2.1.

Обзор характеристик методов управления

Критерий	Управление с обратной связью и регулируемые коэффициентами	Адаптивное управление с эталонной моделью Ляпунова	Нейросетевое управление
Устойчивость обратной связи	Наихудшая	Наилучшая	Средняя
Скорость сходимости	Наилучшая	Средняя	Наихудшая
Работа в реальном времени	Средняя	Средняя	Наилучшая
Сложность программы управления	Наихудшая	Средняя	Средняя
Ошибка слежения	Средняя	Наилучшая	Средняя
Подавление помех	Наилучшая	Наихудшая	Средняя
Робастность Рассогласования модели	Наихудшая	Средняя	Наилучшая

В методе с использованием НС отсутствуют ограничения на линейность системы, он эффективен в условиях шумов и после окончания обучения обеспечивает управление в реальном масштабе времени. Нейросетевые СУ более гибко настраиваются на реальные условия, образуя модели полностью адекватные поставленной задаче, не содержащие ограничений, связанных с построением формальных систем. Кроме того, нейросетевые СУ не только реализуют стандартные адаптивные методы управления, но и предлагают свои алгоритмические подходы к ряду задач, решение которых вызывает затруднение вследствие неформализованности. Так, появляется возможность обработки в рамках одной модели данных одной природы – для НС важна лишь их коррелированность.

Таким образом, напрашивается следующий вывод. Будущее интеллектуального управления лежит в сочетании традиционного управления с потенциальными возможностями и перспективами использования систем, основанных на использовании искусственных НС.

2.2. Использование иеронных сетей в системах управления р-настойчим множественным доступом

Эволюция сетей связи с подвижными объектами осуществляется в направлении обеспечения интегральных услуг предоставляемых пользователям.

Предоставление интегральных услуг для любых пользователей тесно связано с использованием цифровых методов передачи всех видов информации, представленной в едином пакетном виде и использовании методов множественного доступа (МД). К таким методам можно отнести семейство гибридных методов основанных на временном, кодовом и частотном разделении каналов [20, 21]. Гибридные методы доступа позволяют найти компромисс между обеспечиваемым гарантированным временем доставки сообщения при использовании детерминированных методов доступа и высокой степенью использования пропускной способности канала, присущей большинству случайных методов МД. Однако высокая степень использования пропускной способности канала при случайном МД достигается при определенной (оптимальной) входной нагрузке. При увеличении поступающей нагрузки сверх допустимого порогового значения коэффициент использования пропускной способности канала снижается. Эффективность использования ресурсов пропускной способности канала и в гибридных, и в случайных методах МД может быть существенно повышена при управлении потоком поступающих в радиоканал. В качестве управляемого параметра может использоваться вероятность первичной и повторной передачи пакетов в радиоканал. Влияние выбора вероятности передачи на поддержание коэффициента использования пропускной способности канала МД можно пояснить следующим образом. Для подсистемы радиодоступа при малом значении вероятности повторной передачи у абонентов, в канале МД имеется большое количество свободных временных окон. При значительном увеличении вероятности повторной передачи возрастает число конфликтных ситуаций. И в первом и во втором случае пропускная способность канала используется не полностью. Управление вероятностью передачи и выбор ее оптимального значения позволяет поддерживать коэффициент использования пропускной способности канала МД на почти постоянном, близком к максимально возможному, уровне. В ряде работ [62, 83] доказана теорема о существовании такой вероятности повторной передачи p_r^{opt} и ее влиянии на эффективность функционирования сети в целом. Управление значением p_r^{opt} осуществляется на основе анализа пространства состояний радиоканала. Для оценки пространства состояний складывающегося в канале множественного доступа и нахождения требуемого значения

управляемого параметра предлагается использовать нейросетевые модели.

В общем виде модель управления p -настойчивым множественным доступом может быть представлена в виде:

$$x(k+1) = f(x(k), u(k)), \quad (2.2)$$

где $u(k)$ - вектор входных сигналов, $x(k+1)$ - вектор выходных сигналов, $k = 0, 1, \dots$ - дискретное время.

Частными случаями представления вида (2.2) являются четыре модели идентификации объектов управления [91] (объект управления имеет один вход и один выход). Эти модели могут быть описаны следующими нелинейными разностными уравнениями:

$$\text{Модель 1. } x(k+1) = \sum_{i=0}^{n-1} \alpha_i x(k-i) + f[u(k), u(k-1), \dots, u(k-m+1)]. \quad (2.3)$$

$$\text{Модель 2. } x(k+1) = f[x(k), x(k-1), \dots, x(k-n+1)] + \sum_{i=0}^{m-1} \beta_i u(k-i). \quad (2.4)$$

Модель 3.

$$x(k+1) = f[x(k), x(k-1), \dots, x(k-n+1)] + g[u(k), u(k-1), \dots, u(k-m+1)]. \quad (2.5)$$

Модель 4.

$$x(k+1) = f[x(k), x(k-1), \dots, x(k-n+1); u(k), u(k-1), \dots, u(k-m+1)]. \quad (2.6)$$

где $[u(k), x(k)]$ – пара значений входа-выхода в момент k . Функции f и g являются дифференцируемыми по всем параметрам.

Модель 1: выход неизвестного нелинейного объекта управления зависит линейно от его предыдущих значений и нелинейно от предыдущих значений входа. Модель 1 показана на рис. 2.6 и состоит из секционированных линий задержки на входе и в цепи обратной связи.

Модель 2: может быть реализована как показано на рис. 2.7. В этом случае выход зависит линейно от входа $u(k)$ и его предыдущих значений и нелинейно от собственных предыдущих значений.

Модель 3: из нелинейного разностного уравнения (2.5) видно, что выход нелинейного объекта управления зависит нелинейно как от предыдущих значений входа, так и от предыдущих значений выхода. Схематичное представление нелинейного разностного уравнения для модели 3 приведено на рис. 2.8.

Модель 4: эта наиболее общая из всех моделей, представленных ранее. Выход в любой момент времени является нелинейной функцией предыдущих значений и входа и выхода. Она изображена на рис. 2.9.

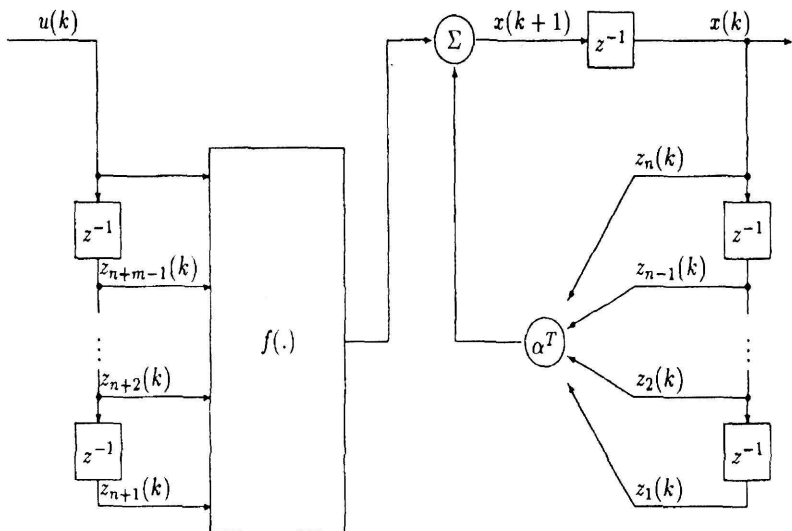


Рис.2.6. Модель 1

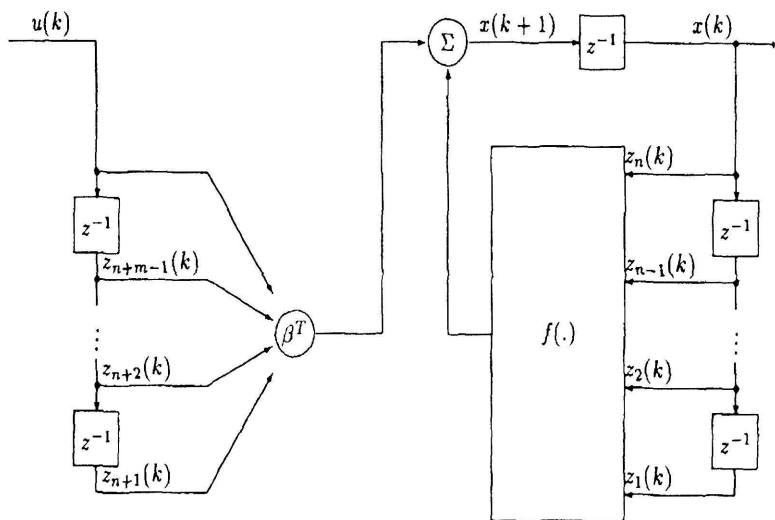


Рис. 2.7. Модель 2

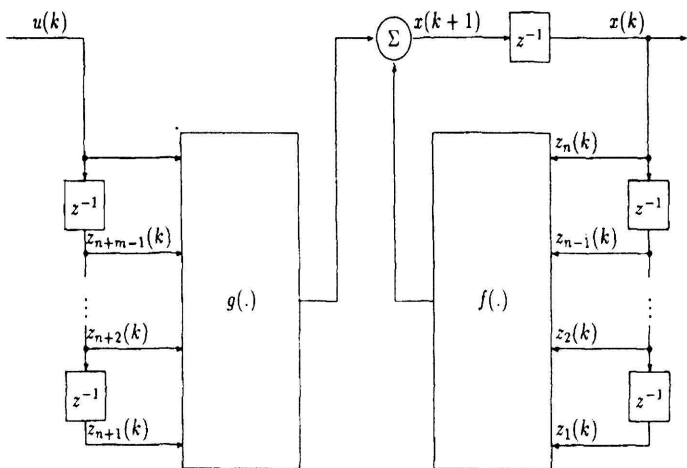


Рис. 2.8. Модель 3

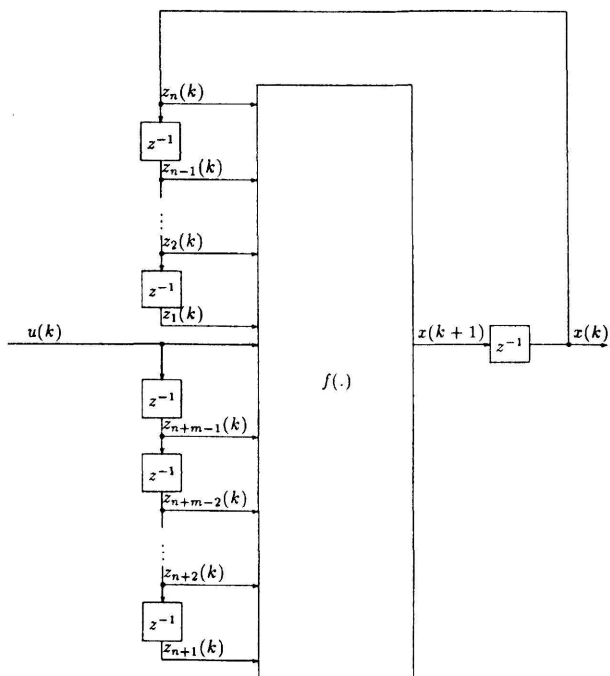


Рис. 2.9. Модель 4

Из (2.3)–(2.6) и рис. 2.6–2.9 следует, что модели идентификации объектов управления в общем случае должны состоять из нейросетевых моделей и секционированных линий задержки. При этом используемая НС должна содержать достаточное для соответствующего нелинейного отображения состояния объекта управления число слоев и узлов в каждом слое. С точки зрения математики это подразумевает, что нелинейные функции в разностных уравнениях, описывающих состояние объекта управления могут быть заменены нейросетевыми моделями с фиксированными матрицами весовых коэффициентов W_i . Чтобы правильно идентифицировать состояние объекта управления и на основе этого формировать соответствующее управляющее воздействие модель идентификации должна выбираться на основе имеющейся информации относительно класса, которому принадлежит рассматриваемый объект.

При построении контроллера для управления p -настойчивым множественным доступом могут использоваться типовые варианты реализации управляющих контроллеров, содержащих нейросетевые модели описанные в [73]. Нейронная сеть будет являться одной из основных компонент системы управления. При этом НС используется в виде либо параллельной, либо последовательно-параллельной (рис. 2.10) модели. Настройка весов НС (обучение НС), может рассматриваться как процесс ее адаптации к выбранной системе управления.

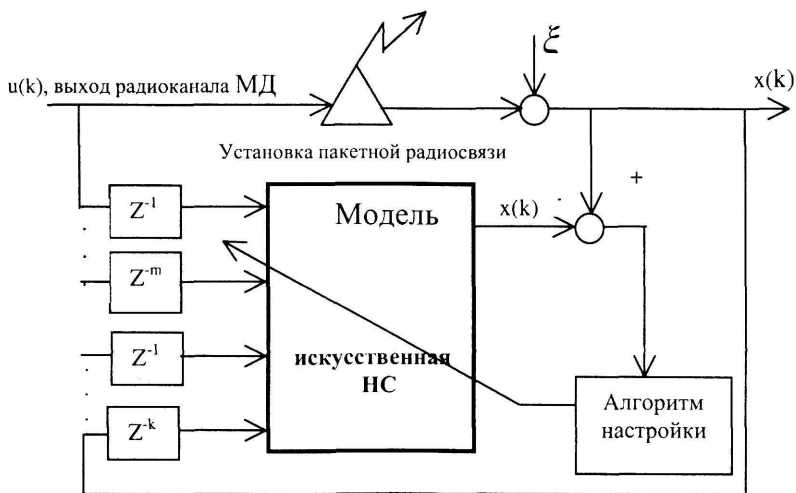


Рис. 2.10. Обобщенная модель управления p -настойчивым множественным доступом

Вариант использования многослойной НС для решения задачи децентрализованного управления р-настойчивым МД представлен на рис. 2.11.

Слева схематично изображен частотно-временной план тактированного радиоканала случайного МД. Предполагается, что возможно наличие трех событий: "Конфликт", "Успех", "Свободно" (пустое временное окно). Справа схематично изображена q -слойная НС с H_i нейронами в i -м слое.

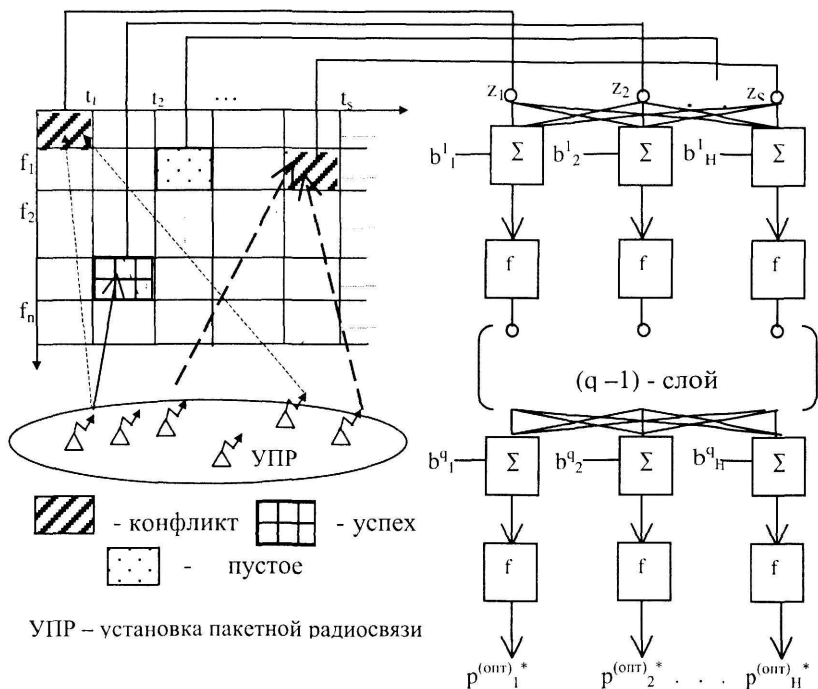


Рис.2.11. Вариант использования НС для решения задачи управления МД

Математическая модель такой сети имеет вид:

$$x(k+1) = \sum_{i=1}^{H_q} w_i^q f \left[\sum_{j=1}^{H_{q-1}} w_j^{q-1} \sigma \left[\dots \varphi \left[\sum_{p=1}^q w_p^1 x(k-p+1) + \sum_{l=1}^m w_{q+l}^1 u(k-l+1) \dots \right] \right] \right], \quad (2.7)$$

или в векторно-матричной форме

$$x(k+1) = (w^q)^T f \left[(w^{q-1})^T \sigma \left[\dots \left[\varphi \left[(w^1)^T z \right] \dots \right] \right] \right]. \quad (2.8)$$

Здесь $x(\cdot), u(\cdot)$ – соответственно выходной и входной сигналы;

$z = [y(\cdot), u(\cdot)]^T$; w^q – весовой коэффициент i -го нейрона в q -ом слое;

$f(\cdot), \sigma(\cdot), \varphi(\cdot)$ – функции активации.

Наиболее часто используется модель искусственной НС, содержащая помимо входного и выходного слоев, один скрытый слой. Кроме того, для простоты принимается, что все функции активации имеют одинаковый вид. Для того случая уравнение (2.7) принимает вид

$$x(k+1) = (w^3)^T f \left[(w^2)^T f \left[(w^1)^T z(k) \right] \right]; \quad (2.9)$$

где $z(k) = [x(k), x, \dots, x(k-m+1); u(k), \dots, u(k-m+1)]^T$.

Такая нейросетевая модель и была использована при компьютерном моделировании.

Для обучения многослойных НС может быть использовано несколько процедур, оптимизирующих некоторую функцию критерия с помощью градиентных методов первого и второго порядков, а также с помощью методов линеаризации (см. п.1.2.2). Обучающее множество должно быть представительным, что гарантирует правильную работу системы идентификации даже в случае поступления на ее вход набора, не включенного в обучающее множество.

При моделировании среднеквадратическая ошибка

$$E = \frac{1}{2} \sum_{c=1}^C \sum_{s=1}^S (p_{sc}^{(onm)*} - p_{sc}^{(onm)})^2$$

рассчитывалась при использовании про-

цедуры обратного распространения. Здесь через C обозначается число выборок образов в обучающем множестве, а через $p_c^{(opt)*} = \{p_{1c}^{(opt)*}, p_{2c}^{(opt)*}, \dots, p_{nc}^{(opt)*}\}^E$ и $p_c^{(opt)} = \{p_{1c}^{(opt)}, p_{2c}^{(opt)}, \dots, p_{nc}^{(opt)}\}^T$ – соответственно получаемый и желаемый выходные векторы сети. Для расчета желаемого выходного вектора сети, используемого для обучения может использоваться, в частности, аналитическая модель, предлагаемая в [18, 62].

Результаты моделирования показали, что использование НС для управления р-множественным доступом позволяет получать оптимальные значения вероятности повторной передачи в соответствии с выбранной стратегией при неизвестных ранее входных данных. Причем после завершения процесса обучения не требуется производить дополнительных вычислений, и процесс принятия решения ускоряется.

Глава 3. Применение нейронных сетей в телекоммуникационных системах

3.1. Основные направления применения искусственных нейронных сетей в телекоммуникационных системах

Длительное время считалось, что нейрокомпьютеры эффективны и применимы лишь для решения так называемых неформализуемых и плохо формализуемых задач, связанных с необходимостью включения в алгоритм их решения данных обучения на реальном экспериментальном материале. В первую очередь к таким задачам относятся задачи распознавания образов. В последнее время область применения нейроинформационных технологий динамично расширяется [19]. Они все шире и шире используются в задачах с ярко выраженным естественным параллелизмом: обработка сигналов, изображений и т. п.

Среди основных выделяют четыре области применения нейрокомпьютеров в системах связи [9]:

- управление коммутацией;
- маршрутизация;
- управление трафиком;
- распределение каналов в подвижных системах радиосвязи.

Решение практически любой задачи в нейросетевом логическом базисе предполагает наличие следующих этапов [12]:

- формирование входного и выходного сигналов НС;
- формирование желаемого выходного сигнала НС;
- формирование сигнала ошибки и функционала оптимизации;
- формирование структуры НС, адекватной выбранной задаче;
- разработка алгоритма настройки НС, эквивалентного процессу решения задачи в нейросетевом логическом базисе;
- проведение исследований процесса решения задачи.

Метод обучения и нейронная сеть, используемая для управления работой высокоскоростной коммутационной сети с пакетной передачей в асинхронном режиме, описаны в работе [44]. Сеть используется для управления коммутацией пакетов при передаче речи, изображений и данных. Коммутатор представлен как логическое устройство, на вход которого поступает N сигналов и которое на выходе воспроизводит эти сигналы в любом представленном порядке.

В работе [48] рассматривается пространственный коммутатор (входы и выходы коммутатора являются различными физическими линиями). Приводятся примеры формирования структуры НС для управления процессом коммутации в различных телекоммуникационных системах.

В настоящее время число связей, эмулируемых в нейрокомпьютере, может достигать нескольких сот миллионов. Поэтому становится возможным построение коммутаторов с нейросетевым управлением на несколько сот каналов [9]. В работе [49] рассматривается многоступенчатая коммутационная схема Баньяна и ее управление с помощью нейроконтроллера. Подробнее об управлении коммутацией с использованием искусственных НС можно ознакомиться в [9, 103] и других работах.

Использование НС для управления трафиком в сложных многоступенчатых системах связи предложено в работе [55]. Трудность задачи обусловлена тем, что, во-первых, заранее неизвестны параметры, характеризующие потоки информации, а во-вторых, требования к качеству могут меняться со временем. НС решает задачи оптимизации, связанные с нахождением бесконфликтных потоков при заданных входных и выходных значениях. При этом НС легко адаптируется к изменениям условий.

С нейросетевыми алгоритмами маршрутизации в системах связи можно ознакомиться в работах [9, 97], а один из возможных алгоритмов маршрутизации будет рассмотрен в п. 3.1.2.

Постановка и решение задачи распределения каналов в подвижных системах радиосвязи в нейросетевом базисе мало отличаются от постановки и решения задачи маршрутизации. Разница заключена в сотовой структуре радиосети и большом числе коммутируемых узлов [9, 56]. Один из возможных алгоритмов распределения каналов в подвижных системах радиосвязи рассматривается в п.3.1.4.

В п.3.1.3 описан нейросетевой алгоритм прогнозирования напряженности поля, который как и алгоритм распределения каналов в подвижных системах радиосвязи может быть использован на этапе проектирования таких систем.

Кроме вышеперечисленных областей применения нейронных сетей в телекоммуникационных системах, перспективным является использование нейросетевых алгоритмов в задачах кодирования и декодирования информации. В качестве первоочередных можно рассматривать задачи связанные с обработкой речевой информации и изображений [12] (см. п.2.1.1). Широко известным является метод сжатия информации, предложенный в 1987 году [13]. При этом используется трехслойный перцептрон, у которого число элементов входного и выходного слоев одинаково, а число элементов скрытого слоя значительно меньше. Если обучение произведено таким образом, что на выходе воспроизводится тот же самый вектор, который подается на входной слой, то такой перцептрон автоматически осуществляет сжатие информации (на элементах скрытого слоя возникает представление каждого вектора, которое значительно короче, чем длина вектора, подаваемого на вход). Таким образом, передача предварительно сжатой информации по линии связи позволяет уменьшить

число необходимых для этого каналов. Реализуется это следующим образом. На одном конце линии помещают входной и скрытый слои перцептрона, а результат работы элементов скрытого слоя (короткие векторы) подают в канал. Поместив на другом конце линии копию скрытого слоя и выходной слой, можно на выходе последнего воспроизвести исходный вектор.

Большое количество работ посвящено построению нейросетевых приемников систем множественного доступа [39, 71].

С другими вариантами применения НС в телекоммуникационных системах можно ознакомиться в [9, 12, 19, 44, 70, 75, 85, 93].

Таким образом, область применения искусственных НС в телекоммуникационных системах постоянно расширяется. В этих условиях появление на рынке недорогих как многофункциональных, так и ориентированных на решение конкретных задач нейрочипов и нейрокомпьютеров будет способствовать быстрому переходу к очередному этапу развития телекоммуникационных систем – появлению интеллектуальных телекоммуникационных систем. Ждать этого, по всей видимости, остается совсем недолго.

3.2. Нейронные сети в системах автоматического распознавания речи

Хорошо известно, что речь человека характеризуется высокой степенью изменчивости. Это обусловлено несколькими причинами [6]. Во-первых, даже для одного и того же говорящего, реализации одних и тех же акустических единиц будут отличаться по своему спектральному составу и длительности произношения. Это может быть связано с изменениями эмоционального состояния человека, условий, в которых он находится. Во-вторых, наличие коартикуляционных эффектов приводит к тому, что произношение слов и фонем сильно зависит от их контекста. В-третьих, к изменениям в речевом сигнале приводят помехи различного характера. Принимая во внимание все эти факторы, и учитывая ряд других ограничений, следует отметить, что для высококачественного распознавания речи в реальном времени требуются вычислительные средства с высоким быстродействием. Одним из способов снижения этого требования является распараллеливание вычислений, которое естественным образом достигается при использовании искусственных НС, реализованных на нейрокомпьютерах.

К настоящему времени разработаны высокоэффективные нейронносетевые модели для распознавания коротких речевых сегментов и наборов схожих по звучанию изолированных слов. Часть из этих моделей успешно

объединены с традиционными подходами для создания распознавателей больших словарей, изолированных слов и слитной речи [6].

Рассмотрим простейшую схему распознавания изолированных слов, представленную на рис. 3.1 [6].

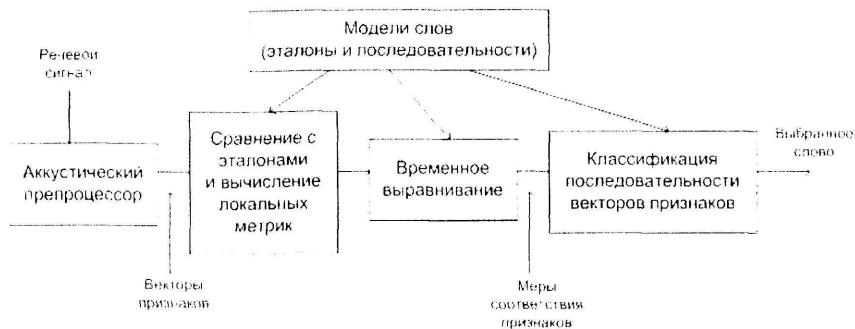


Рис. 3.1. Схема распознавателя изолированных слов

Процесс распознавания в этом случае можно разделить на три этапа. На первом этапе акустический препроцессор преобразует входной речевой сигнал в последовательность векторов признаков или акустических векторов, извлекаемых через фиксированные промежутки времени. Как правило, эти векторы содержат спектральные или кепстральные коэффициенты, характеризующие короткие отрезки речевого сигнала.

На втором этапе векторы сравниваются с эталонами, содержащимися в моделях слов, и вычисляются их локальные метрики или меры соответствия (в общем случае сравниваются речевые сегменты, представленные несколькими векторами признаков). На третьем этапе эти метрики используются для временного выравнивания последовательностей векторов признаков с последовательностями эталонов, образующими модели слов, и вычисляются меры соответствия для слов. Временное выравнивание используется для компенсации изменений в скорости произнесения.

После выполнения всех этих операций распознаватель выбирает слово, для которого мера соответствия максимальна. При распознавании слитной речи локальные метрики полученные на втором этапе вычислений, используются для временного выравнивания и определения мер соответствия для отдельных предложений или высказываний. С целью высококачественного распознавания обычно используется дополнительный этап, позволяющий учесть семантические, синтаксические и прагматические ограничения.

В схеме распознавания, изображенной на рис.3.1, НС наиболее успешно используются на второй стадии вычислений при расчете локальных метрик [87]. Для статистических распознавателей с непрерывным наблюдением данные метрики являются монотонными функциями функций правдоподобия векторов признаков.

Простейшие из этих функций, такие как логарифм функции правдоподобия для гауссовского распределения векторов независимых величин, могут быть рассчитаны с помощью однослойных сетей без их предварительного обучения (для известных параметров распределений) [6]. При вычислении более сложных метрик могут быть использованы многослойные перцептроны, способные вычислять функции любой сложности. При настройке весовых коэффициентов таких сетей используется способность многослойного перцептрона, имеющего достаточное число связей, аппроксимировать апостериорную вероятность классов после его обучения для выполнения классификации [6, 98]. Данное свойство было успешно использовано для создания высокоэффективных гибридных подходов к распознаванию слитной речи, основанных на скрытых марковских моделях (СММ), где многослойные сети служат для вычисления правдоподобий состояний СММ [47, 90]. Использование НС в таких подходах позволяет учитывать при выполнении распознавания акустический контекст векторов наблюдений СММ и снять допущения относительно формы распределения этих векторов [6].

Распознаватели речи с дискретным наблюдением сначала выполняют векторное квантование и присваивают каждому вектору признаков определенный символ из кодовой книги. Затем на основе этих символов с помощью специальных таблиц, содержащих вероятности наблюдения символов для каждого эталонного вектора, вычисляются локальные метрики. Такие вычисления могут быть выполнены однослойными перцептронами, состоящими из линейных узлов, число которых равно числу эталонов. Число входов такого перцептрона должно быть равным числу возможных символов.

Векторное квантование может быть выполнено с помощью сети, подобной карте признаков Кохонена. Такая сеть представляет собой двумерный массив узлов кодовой книги, содержащий по одному узлу на каждый возможный символ. Каждый узел вычисляет евклидово расстояние между входным вектором сети и соответствующим эталоном, представленным весами узла, после чего выбирается узел с наименьшим евклидовым расстоянием. Веса данной сети вычисляются с помощью алгоритма Кохонена, его модификаций [6] или с помощью любого другого традиционного алгоритма векторного квантования, использующего в качестве метрики евклидово расстояние (например, с помощью кластеризационного алгоритма k-средних [6, 87]).

Многослойные нейронные сети также могут быть использованы для снижения размерности векторов признаков, извлекаемых препроцессором на начальном этапе распознавания. Такая НС имеет столько же выходов, сколько и входов, и один или более слоев скрытых узлов. При обучении НС ее веса подбираются так, чтобы она могла воспроизводить на выходе любой входной вектор через небольшой слой скрытых узлов. Выходы этих узлов после обучения сети могут быть использованы в качестве входных векторов меньшей размерности для дальнейшей обработки речи.

В случае использования НС для классификации статических образов фонем, слогов и небольших словарей изолированных слов в качестве входного образа может быть выбран вектор признаков, характеризующий стационарный участок ее реализации. Эксперименты показывают, что в этом случае нейронносетевые классификаторы имеют примерно такую же точность распознавания, как и традиционные гауссовский классификатор и классификатор k ближайших соседей [87]. Для учета динамической природы речи, то есть характера изменения параметров речевого сигнала во времени, в качестве статического входа НС может быть выбрано окно, включающее несколько последовательных во времени векторов признаков. Данное окно обычно размещается в начале или в конце распознаваемой фонемы, так как именно там наиболее проявляется динамический характер речи и имеется возможность учитывать ее контекст. Однако для высококачественного опознавания необходима точная сегментация обучающих и контрольных выборок для их временного выравнивания, которое на практике осуществить довольно сложно.

В [87] описаны результаты экспериментов по исследованию перцептронов и некоторых типов иерархических НС, используемых для распознавания статических образов изолированных слов и цифр. Показано, что в случае небольших словарей точность распознавания примерно равна точности распознавания коммерческих распознавателей и распознавателей, основанных на СММ.

Особый интерес вызывают динамические нейросетевые классификаторы, разработанные специально для распознавания речи и включающие в свой состав короткие временные задержки и узлы, выполняющие временное интегрирование, или рекуррентные связи. Обычно такие классификаторы мало чувствительны к небольшим временным сдвигам обучающих и контрольных выборок и, следовательно, не требуют для высококачественной работы точной сегментации речевых данных. Использование динамических сетей при распознавании речи позволяет преодолеть основные недостатки, присущие статическим сетям, и, как показывают экспериментальные исследования, приводит к превосходному качеству распознавания для акустически схожих слов, согласных и гласных [6]. Частота ошибок у динамических сетей для задач с малым

словарем часто оказывалась значительно ниже, чем у лучших альтернативных распознавателей, в том числе и основанных на СММ.

Нейронная сеть с временными задержками (НСВЗ) представляет собой многослойный перцептрон, узлы которого модифицированы введением временных задержек. Узел, имеющий N задержек $\tau, 2\tau, \dots, N\tau$, показан на рис. 3.2. Он суммирует взятые в $N+1$ последовательных моментов времени J своих входов, умноженных на соответствующие весовые коэффициенты, вычитает порог и вычисляет нелинейную функцию F полученного результата.

Архитектура трехслойной НСВЗ, предложенной для распознавания трех фонем (или трех классов фонем), показана на рис. 3.3 (на нем показаны связи только для одного выходного узла).

На рис. 3.3 показано, что обработка сдвинутой входной последовательности акустических векторов эквивалентна прохождению окон временных задержек над образами узлов нижнего уровня. На самом нижнем уровне эти образы состоят из сенсорного входа, т. е. акустических векторов. Узлы скрытых слоев сети представляют собой движущиеся детекторы признаков и способны обнаруживать требуемые образы в любом месте входных последовательностей. Благодаря тому что выходные узлы имеют равные веса связей со вторым слоем, любые моменты времени для таких детекторов являются равноправными. Это делает сеть инвариантной к временным сдвигам обучающих и контрольных образцов фонем (для случая, когда эти сдвиги не столь велики, чтобы важные ключевые признаки оказывались за пределами входной последовательности сети). Простая структура делает НСВЗ подходящей для стандартизированной СБИС-реализации с загружаемыми извне весами.

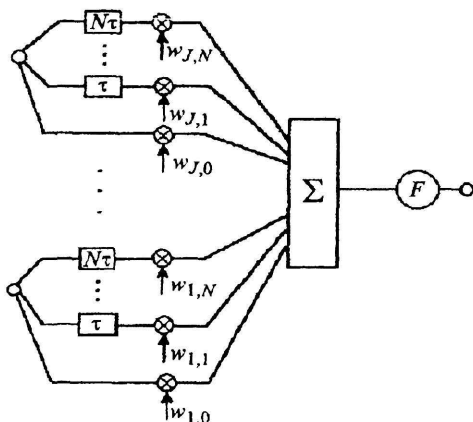


Рис. 3.2. Узел НСВЗ с задержками

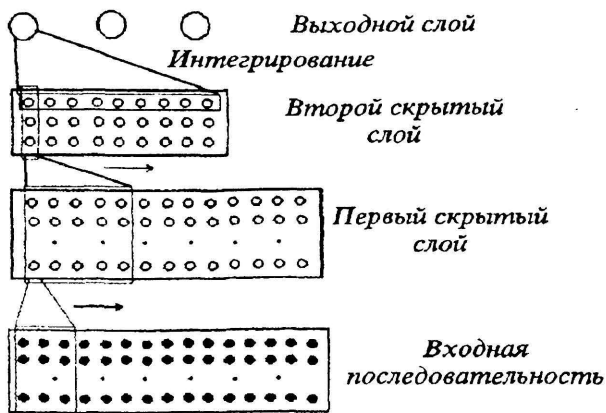


Рис. 3.3. Архитектура HCBV

Обзор нейросетевых структур, предназначенных для выполнения других функций при распознавании речи можно найти, например, в [87].

3.3. Использование нейронных сетей для решения задач маршрутизации

Маршрутизация является одной из важных задач для телекоммуникационных сетей различного назначения. Задачи, связанные с выбором маршрута, планированием работы средств связи и т. п., относятся к классу сложных комбинаторно-оптимизационных задач, как правило, не имеющих простых аналитических решений. Кроме того, сложность необходимых вычислений экспоненциально возрастает при увеличении количества узлов в сети. Поэтому в настоящее время широко применяют различные эвристические алгоритмы и процедуры, полученные путем творческого поиска, интуиции и опыта исследователя. Альтернативой существующим методам решения задач маршрутизации является использование нейросетевых моделей, которые позволяют при значительном снижении временных затрат получить хорошие субоптимальные решения. Так, для решения комбинаторно-оптимизационных задач широко используются модели построенные на основе НС Хопфилда, впервые примененные для решения задачи о коммивояжере. Эти модели явились началом развития нейронных методов решения сложных оптимизационных задач. Большинство последующих исследований так или иначе базировались именно на них

[77, 78]. Коротко остановимся на формулировке и основных принципах организации вычислений при решении задачи коммивояжера.

Для некоторой группы городов с известными расстояниями между ними требуется найти кратчайший маршрут посещения каждого города один раз с возвращением в исходную точку.

Обозначим города, которые необходимо посетить, буквами А, В, С..., а расстояния – $d_{AB}, d_{AC} \dots d_{BC} \dots$. Решением является упорядоченное множество из n городов. Последовательность, в которой обходятся города удобно представлять матрицей $n \times n$, строки которой соответствуют городам, а столбцы номерам городов в последовательности. Например, имеется пять городов А, В, С, D, E, а последовательность обхода этих городов задана матрицей

$$\begin{array}{c|ccccc}
 & 1 & 2 & 3 & 4 & 5 \\
 \hline
 \text{A} & 0 & 1 & 0 & 0 & 0 \\
 \text{B} & 0 & 0 & 0 & 1 & 0 \\
 \text{C} & 1 & 0 & 0 & 0 & 0 \\
 \text{D} & 0 & 0 & 0 & 0 & 1 \\
 \text{E} & 0 & 0 & 1 & 0 & 0
 \end{array} \quad (3.1)$$

Таким образом город С посещается первым, город А – вторым и т. д. Длина маршрута равна $d_{CA} + d_{AE} + \dots + d_{DC}$. В каждом столбце и в каждой строке этой матрицы может быть только одна единица, так как в каждый момент посещается только один город и каждый город посещается только один раз. Матрицу вида (3.1) можно воспринимать как состояние нейронной сети из $N = n^2$ нейронов. Задача состоит в том, чтобы из $n!/2n$ маршрутов выбрать один с наименьшей длиной. Состояние каждого нейрона описывается двумя индексами, которые соответствуют городу и порядковому номеру его посещения в маршруте. Например, $Y_{xj} = 1$ показывает, что город x был j -м по порядку городом маршрута.

Запишем функцию вычислительной энергии для сети, предназначенной для решения задачи коммивояжера, в которой состояние с наименьшей энергией соответствует самому короткому маршруту. В общем виде такая функция для рассматриваемой сети может иметь следующий вид [19]:

$$E = -\frac{1}{2} \sum_i \sum_j w_{ij} Y_i Y_j - \sum_j I_j Y_j + \sum_j T_j Y_j, \quad (3.2)$$

где E – искусственная энергия сети, w_{ij} – вес от выхода нейрона i к входу нейрона j , Y_j – выход нейрона j , I_j – внешний вход нейрона j , T_j – порог нейрона j .

Изменение энергии, вызванное изменением состояния j -нейрона, можно вычислить следующим образом:

$$\delta E = \left(\sum_{i \neq j} (w_{ij} Y_i) + I_j - T_j \right) \delta Y_j, \quad (3.3)$$

где δY_j – изменение выхода j -го нейрона.

Каждому состоянию системы соответствует конкретная величина вычислительной энергии. Устойчивое состояние имеет меньшую энергию, чем неустойчивое. Эволюция системы во времени – это движение в пространстве состояний в поисках минимума энергии и остановка в этой точке.

Для рассматриваемой системы функция энергии должна удовлетворять следующим требованиям [32]. Во-первых, она должна поддерживать устойчивые состояния в форме матрицы (3.1). Во-вторых, из всех возможных решений функция энергии должна поддерживать те, которые соответствуют коротким маршрутам. Этим требованиям удовлетворяет функция энергии вида (при этом, $Y_{xj} = 0, 1$)

$$E = \frac{A}{2} \sum_x \sum_i \sum_{j \neq 1} Y_{xi} Y_{xj} + \frac{B}{2} \sum_i \sum_x \sum_{k \neq x} Y_{xi} Y_{ki} + \frac{C}{2} \left(\sum_x \sum_i Y_{xi} - n \right)^2 + \frac{D}{2} \sum_x \sum_{k \neq x} \sum_i d_{xk} Y_{xi} (Y_{k,i+1} + Y_{k,i-1}). \quad (3.4)$$

Первые три члена выражения (3.4) поддерживают первое требование, четвертый член – второе; A, B, C, D – положительные множители. Первый член равен нулю, если каждая строка x содержит не больше одной единицы. Второй член равен нулю, если каждый столбец содержит не более одной единицы. Третий член равен нулю, если в матрице вида (3.1) n единиц. Таким образом, без учета четвертого члена функция энергии имеет минимумы ($E = 0$) во всех состояниях, представленных матрицей с одной единицей в каждом столбце и каждой строке. Все другие состояния имеют более высокую энергию. Короткие маршруты поддерживает четвертый член. В нем индексы i берутся по $\text{mod } n$, для того чтобы показать, что i -й город соседствует в маршруте с $(n-1)$ -м и первым, т. е. $Y_{k,n+j} = Y_{kj}$. Четвертый член численно равен длине маршрута.

Раскрывая скобки в (3.4) и приравнявая коэффициенты при квадратичных и линейных членах в полученном выражении и общей формуле энергии (2.2), определяем матрицу связей и внешние взаимодействия:

$$w_{xi,kj} = -A\delta_{xk}(1-\delta_{ik}) - B\delta_{ij}(1-\delta_{xk}) - C - Dd_{xk}(\delta_{j,i+1} + \delta_{j,i-1}), \quad (3.5)$$

где $\delta_{ij} = 1$, если $i = j$, в противном случае $\delta_{ij} = 0$. Кроме того, каждый нейрон имеет смещающий вес $I_{xi} = Cn$.

Первый член в (3.5) задает связи нейронов в каждой строке, второй – внутри каждого столбца, третий и четвертый – глобальные связи. И в (3.4) и в (3.5) три первых члена отвечают за общие ограничения для любой задачи коммивояжера и приводят сеть к финальному состоянию в виде (3.1).

Четвертый член управляет тем, какое из $n!/2n$ возможных различных финальных состояний соответствует самому короткому маршруту.

Рассмотрим вариант совместного решения задачи маршрутизации и планирования использования линий радиосвязи для сети пакетной радиосвязи с многоскачковой топологией. Важность взаимосвязи между маршрутизацией и вопросами планирования последовательности выбора направления для передачи по используемым линиям связи показана в [65]. Там же обобщен случай непрерывного трафика для определенного класса сетей. При этом выбор маршрутов максимизирующих степень узла в сети, позволяет спланировать работу так, чтобы время ее выполнения было минимальным. Степень узла для этого случая, определяется как сумма всех потоков, поступающих в узел и исходящих от узла. Например, линия, которая должна активироваться, три раза добавляет поток из трех единиц к обоим узлам, которые она соединяет. При этом, критерий качества работы, выбираемый для задачи маршрутизации, должен отражать цели, связанные с соответствующей задачей составления плана работы линий связи.

Пусть заданы: граф связности сети пакетной радиосвязи, ряд пар N_{SD} – исходная точка – пункт назначения (SD) и ряд линий связи, соединяющих каждую пару SD. Предполагается, что в системе используется тактированный множественный доступ и длительность временных окон соответствует длине пакета (все пакеты имеют фиксированную длину), а на каждом узле имеется только один приемопередатчик. Между каждой парой узлов сети SD имеет место одинаковый трафик, равный одному пакету на цикл передачи. Считаем также, что линии связи между соответствующими парами узлов активизируются (используются для передачи) по мере необходимости.

Требуется выбрать единственный маршрут между каждой парой SD с таким расчетом, чтобы минимизировать желаемый критерий качества работы.

Показатель качества работы должен согласовываться со структурой НС Хопфилда. По аналогии с рассмотренной выше задачей коммивояжера такой показатель, называемый “энергией перегрузки” задается формулой [111]

$$E_b = \frac{1}{2} \sum_{i=1}^{N_{SD}} \sum_{k=1, k \neq i}^{N_{SD}} \sum_{j=1}^{N_p(i)} \sum_{l=1}^{N_p(k)} |P_{ij} \cap P_{kl}| V_{ij} V_{kl}, \quad (3.6)$$

где P_{ij} – j -й маршрут между SD парой i ,

$|P_{ij} \cap P_{kl}|$ – число узлов, которые совместно используют маршруты P_{ij}

и P_{kl} ,

$$V_{ij} = \begin{cases} 1, & \text{если выбирается } P_{ij} \\ 0, & \text{если не выбирается } P_{ij} \end{cases}$$

$N_p(i)$ – число вариантов маршрутов, определенных между SD парой i .

Цель состоит в минимизации E_b с учетом того, что для каждой пары SD выбирается только один маршрут (т. е. $V_{ij} = 1$ для единственного значения j для каждого значения i). В этом случае энергия перегрузки соответствует сумме числа общих узлов всех выбранных маршрутов (одного для каждой SD пары), взятых попарно. Например, на рис. 3.4 показана простая сеть пакетной радиосвязи с шестью узлами и двумя маршрутами между каждой из двух SD пар.

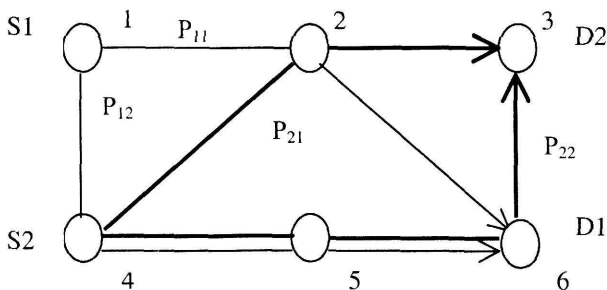


Рис. 3.4. Пример сети связи из шести узлов

Задача маршрутизации состоит в выборе либо пары P_{11} или P_{12} для соединения S1 с D1 и либо пары P_{21} или P_{22} для соединения S2 с D2. Допустимое решение, которое задается выбором трасс P_{11} и P_{22} , имеет энер-

гию перегрузки $E_b = 1$, так как эти маршруты имеют один общий узел (узел 6).

Теперь рассмотрим модель НС Хопфилда, используемую в этом случае для выбора маршрута между несколькими SD парами в сети пакетной радиосвязи. Выходные напряжения нейронов (которые и определяют их состояния) такой НС приближаются к двоичным значениям по мере перехода сети к состоянию устойчивого равновесия с минимальной “энергией”. Соединения между нейронами i и j описываются весом T_{ij} , который положителен если соединение возбуждающее и отрицателен, если соединение тормозящее (запрещающее). В рассматриваемой модели НС для каждого маршрута между каждой SD парой определяется один нейрон. Вариант модели НС для сети изображенной на рис. 3.4 представлен на рис. 3.5. В соответствии с рис. 3.5 нейрон ij отображает j маршрут между SD парой i .

НС эволюционирует от какого-то начального состояния до состояния равновесия, которое отображает минимум (не обязательно глобальный) функции энергии Ляпунова, которая по аналогии с (3.2) может быть записана через веса соединений, токи смещения и напряжения на выходах нейронов следующим образом [111]:

$$E_{total} = -\frac{1}{2} \sum_{i=1}^{N_{SD}} \sum_{k=1}^{N_{SD}} \sum_{j=1}^{N_p(i)} \sum_{l=1}^{N_p(k)} T_{ij,kl} V_{ij} V_{kl} - \sum_{i=1}^{N_{SD}} \sum_{j=1}^{N_p(i)} V_{ij} I_{ij}. \quad (3.7)$$

В выражении (3.7) $T_{ij,kl}$ – вес соединения между нейронами ij и kl , I_{ij} – ток смещения, прикладываемый к нейрону ij , $N_p(i)$ – число маршрутов между SD парой i . В рассматриваемой модели веса соединений являются симметричными (т. е. $T_{ij,kl} = T_{kl,ij}$). Эта симметрия гарантирует сходимость к устойчивому состоянию [111]. Общее число нейронов N задается как $N = \sum_{i=1}^{N_{SD}} N_p(i)$. Таким образом, веса соединений $T_{ij,kl}$ являются элементами матрицы связности размерности $N \times N$.

Таким образом, целевая функция, моделируемая с помощью НС Хопфилда, включает взвешенные суммы произведений пар выходных напряжений нейрона и выходных напряжений взятых по отдельности. При выборе E_b вида (3.6) предполагалось, что оценивается энергия перегрузки сети в допустимом состоянии, т. е. активируется только один маршрут для каждой SD пары (т. е. $V_{ij} = 1$ для единственного значения j для каждого значения i , а остальные значения $V_{ij} = 0$). Однако до достижения сходимости величины V_{ij} принимают значения в континууме $[0, 1]$ и выражение

для энергии перегрузки, определяемое с помощью выражения (3.6), применимо в полной мере только для аналоговой реализации системы.

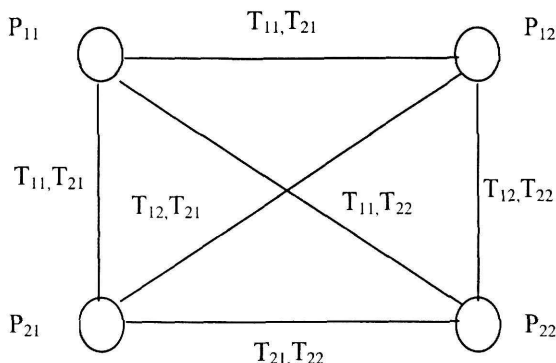


Рис. 3.5. Модель НС для сети, изображенной на рис. 2.4

Рассматриваемая задача оптимизации с целым рядом ограничений может быть сведена к задаче без ограничений за счет включения ограничений в целевую функцию посредством использования множителей Лагранжа [111]. Функция энергии перегрузки при этом приобретает следующий вид:

$$E_{total} = bE_b + \sum_{c=1}^3 \lambda_c E_c - I \sum_{i=1}^{N_{SD}} \sum_{j=1}^{N_p(i)} V_{ij}. \quad (3.8)$$

Ограничения для задачи являются соответствующими членами уравнения энергии перегрузки E_c (равны нулю, если ограничение выполняется) и формулируются так:

1) На SD пару активизируется (выбирается) не более одного маршрута:

$$E_1 = \frac{1}{2} \sum_{i=1}^{N_{SD}} \sum_{j=1}^{N_p(i)} \sum_{\substack{i=1 \\ i \neq j}}^{N_p(i)} V_{ij} V_{il} = 0.$$

2) В сети выбираются строго N_{SD} маршрутов:

$$E_2 = \frac{1}{2} \left(\sum_{i=1}^{N_{SD}} \sum_{j=1}^{N_p(i)} V_{ij} - N_{SD} \right)^2 = 0.$$

3) На SD пару выбирается строго один маршрут:

$$E_3 = \frac{1}{2} \sum_{i=1}^{N_{SD}} \left(\sum_{i=1}^{N_p(i)} V_{ij} - 1 \right)^2 = 0.$$

Хотя последнее ограничение представляется избыточным (выполнение первых двух гарантирует удовлетворение последнего), его включение в уравнение энергии полезно для достижения более быстрой сходимости.

Подстановка выражений для E_b и E_c в (3.8) дает:

$$E_b = \frac{b}{2} \sum_{i=1}^{N_{SD}} \sum_{k=1, k \neq i}^{N_{SD}} \sum_{j=1}^{N_p(i)} \sum_{l=1}^{N_p(k)} |P_{ij} \cap P_{kl}| V_{ij} V_{kl} + \frac{\lambda_1}{2} \sum_{i=1}^{N_{SD}} \sum_{j=1}^{N_p(i)} \sum_{\substack{l=1 \\ l \neq j}}^{N_p(i)} V_{ij} V_{il} + \frac{\lambda_2}{2} \left(\sum_{i=1}^{N_{SD}} \sum_{j=1}^{N_p(i)} V_{ij} - N_{SD} \right)^2 + \frac{1}{2} \sum_{i=1}^{N_{SD}} \left(\sum_{i=1}^{N_p(i)} V_{ij} - 1 \right)^2 + \frac{\lambda_3}{2} \sum_{i=1}^{N_{SD}} \left(\sum_{i=1}^{N_p(i)} V_{ij} - 1 \right)^2 - I \sum_{i=1}^{N_{SD}} \sum_{j=1}^{N_p(i)} V_{ij}. \quad (3.9)$$

Одним из самых важных вопросов при разработке модели НС Хопфилда и дальнейшем моделировании работы системы является вопрос выбора коэффициентов λ_c . Фактически, любые значения λ_c приведут к получению справедливых выражений для E_{total} . Однако при эволюции системы может быть гарантирован только локальный минимум, то есть конечное состояние зависит от начального состояния, при котором начинается эволюция системы. Таким образом, различные значения коэффициентов приводят к получению различных результатов. В большинстве исследований, посвященных использованию НС Хопфилда величины коэффициентов полагаются постоянными, лучшие значения которых обычно определяются в ходе испытаний при программном моделировании. Однако существует ряд подходов, позволяющих во всей полноте использовать метод множителей Лагранжа [111, 112]. В этом случае величины λ_c изменяются по мере изменения состояния системы [111].

Оценить качество решения задачи обычно не представляется возможным, так как число возможных решений для больших сетей очень велико. Например, для 100-узловой сети существует приблизительно $5 \cdot 10^{35}$ различных решений. Поскольку исчерпывающий поиск для такой сети исключается, то при моделировании выполнялся случайный поиск $2 \cdot 10^6$ выборок решений для получения опорного уровня качества работы для оценки работы НС [65]. Наилучшее решение, полученное с помощью случайного поиска, имело энергию перегрузки $E_b = 567$. Использование традиционных эвристических методов решения задачи маршрутизации позволило получить $E_b = 313$. Наилучшее решение найденное с помощью НС Хопфилда дало $E_b = 291$. Наибольшее значение энергии пере-

грузки в этом случае было $E_b = 303$. Моделирование выполнялось от 50 различных начальных состояний. Таким образом, результаты моделирования показывают эффективность рассмотренной модели для минимизации перегрузки в больших сетях. Тот факт, что глобальный минимум находится не всегда, скрашивается тем обстоятельством, что возможно осуществление нескольких испытаний при различных начальных условиях, так что найденное наилучшее решение может выбираться в качестве решения задачи.

Большое количество работ посвященных использованию НС при решении задачи маршрутизации и близость получаемых результатов к оптимальным свидетельствуют о робастности таких моделей.

3.4. Использование нейронных сетей при планировании сотовых сетей подвижной радиосвязи

Одним из основным этапов проектирования сотовых сетей подвижной радиосвязи (ССПР) является этап территориального планирования, в ходе которого выбираются структура (конфигурация) сети, места размещения базовых станций (БС), рассчитываются возможность обеспечения охвата (покрытия) требуемой зоны обслуживания с заданным качеством связи и емкость сети, требуемая для обслуживания абонентской нагрузки с заданной интенсивностью потерь (отказов в обслуживании).

В качестве технических основ планирования используются характеристики стандартов СППР, характеристики приемо-передающего оборудования и антенн, условия распространения радиоволн, требуемая напряженность поля полезного сигнала, средняя нагрузка на одного абонента, допустимое значение вероятности блокирования вызовов [3].

Таким образом, прогнозирование напряженности поля в зоне обслуживания БС является одним из этапов планирования современных систем подвижной связи. В настоящее время широко используются две стратегии прогнозирования напряженности электрического поля. Одна из них состоит в выводе эмпирических формул на основе измерений, а другая основывается на аналитических моделях, использующих теорию дифракции. Практика показывает, что и та и другая модели имеют ряд недостатков [24].

Использование эмпирического метода прогнозирования напряженности поля, основанного на использовании нейронных сетей, позволяет устранить некоторые недостатки [102]. В этом случае НС решает задачу аппроксимации некоторой функции по имеющимся значениям ряда точек. Для НС эти точки являются обучающим множеством. Кроме известных

значений средней напряженности электрического поля на вход НС подается приведенная к дискретной форме и нормализованная информация о покрове земной поверхности, рабочей частоте, высоте антенн и т. д. Параметры НС настраиваются таким образом, чтобы для любой входной комбинации погрешность значения формируемого на выходе НС была минимальна. Выход сети в зависимости от требований может определяться в виде нормализованной напряженности поля в заданной точке, в виде потерь на трассе и др. Динамический диапазон напряженности поля во многом определяется распространением радиоволн в свободном пространстве. Для повышения точности аппроксимации среднего значения напряженности поля, вклад последнего может вычисляться аналитически и исключаться из процесса обучения.

Результаты использования НС для прогнозирования напряженности поля можно рассмотреть на следующих примерах.

В первом случае используется хорошо известная формула Хата [69], отражающая эмпирическую аппроксимацию потерь при распространении L_p в городских районах:

$$L_p(\text{дБ}) = 69,55 + 26,16 \log f(\text{МГц}) - 13,82 \log h_{\text{БС}}(\text{м}) - \\ - \{1,1 \log f(\text{МГц}) - 0,7\} h_M(\text{м}) + \{1,56 \log f(\text{МГц}) - 0,8\} + \\ + \{44,9 - 6,55 \log h_{\text{БС}}(\text{м})\} \log d(\text{км}). \quad (3.10)$$

Диапазон частот $f = 150 \dots 1500$ МГц, высота расположения базовых станций (БС) $h_{\text{БС}} = 30 \dots 200$ м., высоты расположения подвижных станций $h_M = 1 \dots 10$ м и дальность $d = 1 \dots 20$ км. Нейронная сеть, используемая для аппроксимации формулы имела четыре нейрона во входном слое и десять нейронов в скрытом слое. Выход моделируемой сети отображает потери при распространении L_p .

Применение НС для прогнозирования напряженности поля поясняется на рис. 3.6.

Нормализованная средняя ошибка моделирования M определяется следующим образом:

$$M = \frac{\frac{1}{N} \sum_{i=1}^N |L_{p,i} - L_{p,\text{net},i}|}{L_{p,\text{max}} - L_{p,\text{min}}}, \quad (3.11)$$

где $L_{p,\text{net},i}$ – потери на трассе, получаемые с выхода НС.

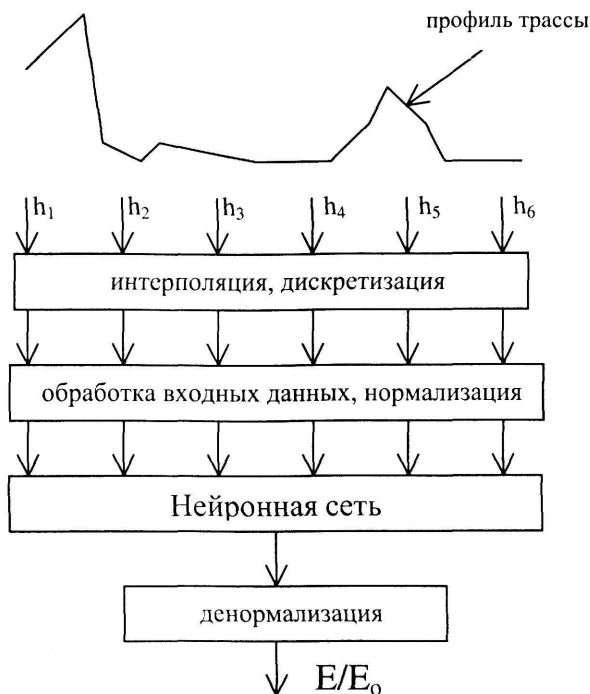


Рис. 3.6. Вариант применения НС для прогнозирования напряженности поля

После 1200 циклов обучения на примере обучающего множества из 2160 точек среднее отклонение составило приблизительно 0,18%. Для точек тестирования, выбранных между точками обучения в рассматриваемом параметрическом пространстве, отклонение увеличивается до 0,24%, при этом $M_{\max} \leq 1,23\%$. В этом эксперименте динамический диапазон составляет $L_{p,\max} - L_{p,\min} = 100 \text{ дБ}$.

Другая типичная задача связана с использованием геометрической теорией дифракции Келлера [80] для случая дифракции на клине (см. рис. 3.7а). При этом также может использоваться многослойная НС с одним скрытым слоем. Число нейронов во входном и скрытом слое выбирается исходя из задачи и предъявляемых к решению требований.

В модели для аппроксимации нормализованной напряженности поля E/E_0 на профиле трассы изображенной на рис. 3.7а использовалась НС, имеющая 11 нейронов во входном слое и 8 нейронов в скрытом слое.

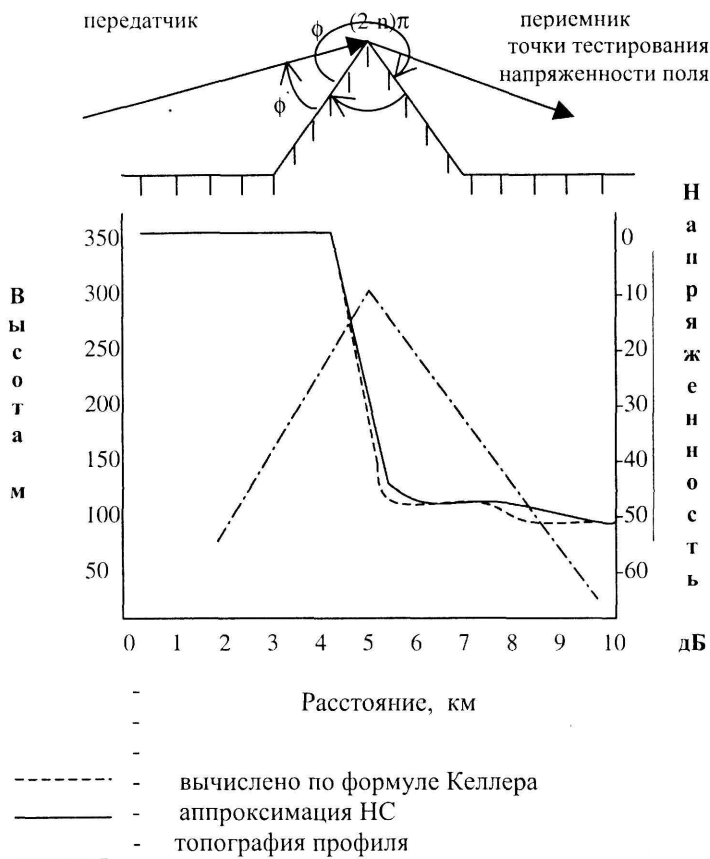


Рис. 3.7. Аппроксимация напряженности поля с помощью нейронной сети

Выборка топографических данных производилась в 10 точках, которые обеспечивают входные данные для 10 элементов. Величина выборки может изменяться. Данные о ней поступают на 11 вход. В результате последней операции выборка приспособляется к длине профиля трассы. Моделирование проводилось для различных комбинаций высот, углов участков препятствия моделируемого “клином”. Среднее отклонение вычислялось по формуле (3.11) и после 20 000 циклов обучения составило 0,3%. В обоих случаях при обучении использовался алгоритм обратного распространения. На практике для обучения НС могут использоваться как результаты аналитических расчетов, так и результаты текущих и ранее

проведенных измерений. После завершения процесса обучения для тестовых примеров максимальное отклонение $M_{\max}$ не превышало 6% (см. рис. 3.7).

Таким образом, подтверждается возможность использования НС для прогнозирования напряженности поля на этапе планирования современных систем подвижной связи и в других случаях. В целях повышения точности прогнозирования рассмотренные нейросетевые модели могут совершенствоваться. Так, может более полно учитываться вся имеющаяся картографическая информация, информация о земном покрове и др.

3.5. Использование нейронных сетей для распределения каналов в сотовых радиосетях

Решение задачи назначения частот, так же как и решение задачи прогнозирования напряженности поля, является одним из этапов проектирования ССПР. При решении этой задачи назначение частот должно производиться в рамках ограничений заданных матрицей электромагнитной совместимости (ЭМС). При этом рассматриваются такие ограничения, как:

- ограничения из-за помех по соседнему каналу;

- ограничения из-за помех, обусловленных совместным расположением (несущие используемые в одной ячейке, должны иметь необходимый разнос по частоте);

- ограничения по числу частот для каждой ячейки.

В 1982 г. Гамст и Рейв определили общий вид задачи распределения каналов в произвольной неоднородной сотовой радиосети [61]. Согласно их определению ограничения ЭМС в n -сотовой сети описываются с помощью матрицы размером $n \times n$, которую называют матрицей совместимости C . Каждый недиагональный элемент c_{ij} матрицы C представляет собой расстояние минимального разнесения в частотной области между частотой, присваиваемой i -й ячейке и частотой, присваиваемой j -й ячейке. Ограничения из-за помех по соседнему каналу могут быть записаны как $c_{ij} = 1$. Запись $c_{ij} = 2$ указывает на то, что частоты, смежные в частотной области не могут присваиваться смежным ячейкам, а значение $c_{ij} = 0$ – на то, что ячейкам i и j разрешается использовать одну и ту же частоту. Каждый диагональный элемент c_{ii} в C обозначает расстояние минимального разнесения между двумя любыми частотами, присваиваемыми

мыми i -й ячейке. При этом для удовлетворения выше сформулированных ограничений c_{ij} всегда должно удовлетворять ограничению $c_{ii} \geq 1$.

Требования, предъявляемые к частотному каналу каждой ячейки n -сотовой сети, описываются с помощью n -элементного вектора, который называют вектором спроса $\mathbf{D}$. Элемент d_i вектора $\mathbf{D}$ обозначает число частот, которые должны присваиваться i -й ячейке. Обозначая через f_{ik} k -ю частоту, присваиваемую i -й ячейке, ограничения по ЭМС можно записать следующим образом:

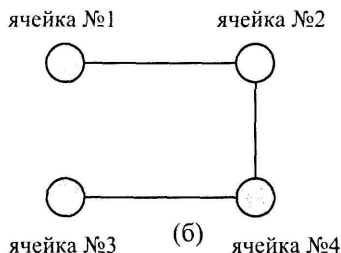
$$|f_{ik} - f_{jl}| \geq c_{ij}, \quad (3.12)$$

для $i = 1, \dots, n$, $j = 1, \dots, n$, $k = 1, \dots, d_i$ и $l = 1, \dots, d_j$, исключая $i = j, k = l$.

В качестве примера, приведем решение задачи распределения каналов (рис.3.8в) для четырехсотовой сети [101]. На рис. 3.8а приведена матрица совместимости $\mathbf{C}$ и вектор спроса $\mathbf{D}$. На рис. 3.8б показана топология сети, соответствующая матрице совместимости $\mathbf{C}$.

$$\mathbf{C} = \begin{bmatrix} 5 & 4 & 0 & 0 \\ 4 & 5 & 0 & 1 \\ 0 & 0 & 5 & 2 \\ 0 & 1 & 2 & 5 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 3 \end{bmatrix}$$

(а)



частота

		1	2	3	4	5	6	7	8	9	10	11
я	1											
ч	.											
е	.											
й	.											
к	.											
а	4											

Рис. 3.8. Распределение частотных каналов для четырехсотовой сети

Задача распределения каналов относится к математическим задачам, связанным с раскраской графов [67]. Из теории известно, что время необ-

ходимое для нахождения решения, экспоненциально увеличивается с увеличением размерности задачи (то есть, ростом числа базовых станций).

В соответствии с теорией окрашивания графов, путем поиска так называемого максимально полного подграфа, можно определить нижнюю границу для общего использованного частотного ресурса. Известные алгоритмы различной степени сложности направлены на удовлетворение всех требований с учетом заданных ограничений при минимизации использованного частотного ресурса. Если в результате решения общее число рабочих частот близко к определенной нижней границе, то оно принимается, в противном случае (если спрос превышает полный частотный ресурс, выделенный для сети подвижной связи) размеры ячеек следует уменьшить (произвести их “расщепление”), чтобы сеть хотя бы локально работала с более высоким уровнем повторного использования частот.

Рассмотрим алгоритм решения задачи распределение частотных каналов, основанный на использовании искусственной НС.

Модель НС для решения комбинаторной задачи оптимизации впервые была введена Хопфилдом и Танком [77]. Решение задачи заключается в минимизации сформированной функции энергии (см. п.1.3.2). Несмотря на то что в ряде работ [95, 109] критиковались устойчивость и качество решения, для таких моделей известно о большом количестве успешных применений НС Хопфилда [19, 57, 58, 93].

В 1991 г. Кунц впервые предложил использовать модель НС Хопфилда для решения задачи распределения каналов в сотовой радиосети [84]. Кунц не рассматривал ограничения из-за помех по соседнему каналу и установил предел для ограничения по расстоянию минимального разнесения между двумя любыми частотами, присваиваемыми i -й ячейке $c_{ii} = 2$. Однако, модель НС Кунца имеет несколько недостатков. Прежде всего, он использует “медленную” сигмоидальную активационную функцию. Кроме этого, в используемой функции энергии не учитываются особенности решаемой задачи. Эти недостатки модели Кунца препятствуют ее использованию для решения задачи распределения каналов в сотовой радиосети на практике.

Предлагаемая модель нейронной сети использует гистерезисную активационную функцию и модифицированную функцию энергии [59]. Используемая функция энергии E определяется путем рассмотрения всех ограничений, описанных выше. При этом она характеризует текущее состояние НС. Это обеспечивает более быструю сходимость решения, чем модель Кунца.

Частные производные вычислительной функции энергии $E(V_1, ..., V_n)$, которую также называют уравнением энергии, задают изменения входа i -го нейрона

$$\frac{dU_i}{dt} = -\frac{\partial E(V_1, \dots, V_n)}{\partial V_i}. \quad (3.13)$$

Здесь U_i и V_i соответственно вход и выход i -го нейрона, n – число нейронов используемых для решения конкретной задачи.

На рис. 3.9 в виде матрицы состояний представлены 4×11 выходов нейронов, используемых для решения задачи распределения каналов в четырехсотовой сети, изображенной на рис. 3.8б, и их сходимость к решению. Минимально необходимое число нейронов для решения этой задачи равно 4×11 . Это следует из матрицы совместимости **C** и вектора спроса **D**, представленных на рис.3.8а ($c_{44} = 5, d_4 = 3$, следовательно, ячейка №4 требует самое малое $11(1+5 \times 2)$ частот).

Число требуемых частот, а следовательно и необходимое для решения задачи число нейронов должно определяться перед моделированием. Обычно для задач большой размерности m приблизительно может быть определено путем перемножения c_{ii} и максимального значения в векторе спроса. Если с помощью определенного таким образом значения m задача не решается, то его необходимо увеличивать до тех пор, пока система не найдет решения. Однако можно воспользоваться и найденными в предыдущих работах, в частности [84, 101] значениями.

V_{ij}		частота j										
		1	2	3	4	5	6	7	8	9	10	11
я	1											
ч	.											
е	.											
й	і											
к	.											
а	4											

Рис. 3.9. Представление выходов 4×11 нейронов, используемых для решения задачи распределения каналов в четырехсотовой сети, изображенной на рис. 3.8б, и их сходимость к решению

Выход ij -го обрабатывающего нейрона V_{ij} указывает на то, присваивается или не присваивается j -частота i -й ячейке. Ненулевой выход ($V_{ij} = 1$) указывает на то, что j -частота присваивается i -й ячейке. Нулевой выход ($V_{ij} = 0$) указывает на то, что j -частота не присваивается i -й

ячейке. На решении задачи, приведенной на рис. 3.9, черные квадраты обозначают ненулевой выход, а белые квадраты – нулевой выход.

Уравнение движения ij -го обрабатывающего нейрона V_{ij} для n -ячеечной, m -частотной задачи задается следующим образом:

$$\frac{dU_{ij}}{dt} = -A \left(\sum_{q=1}^m V_{iq} - d_i \right) - B \left(\sum_{\substack{q=j-(c_n-1) \\ q \neq j \\ 1 \leq q \leq m}}^{j+(c_n-1)} V_{iq} + \sum_{\substack{p=1 \\ p \neq i \\ c_p > 0}}^n \sum_{\substack{q=j-(c_p-1) \\ 1 \leq q \leq m}}^{j+(c_p-1)} V_{pq} \right). \quad (3.14)$$

Первый член выражения (3.14) (А-член) вынуждает d_i нейронов обрабатывающих состояния m частот выделенных для i -й ячейки иметь ненулевой выход в случае, если соответствующие частоты присваиваются i -й ячейке. Второй член выражения (3.14) (В-член) мешает ij -у нейрону иметь ненулевой выход в случае, если присвоение j -й частоты i -й ячейке нарушает следующие ограничения:

если q -я частота в пределах расстояния c_{ii} от j -й частоты ($|j - q| < c_{ii}$) присваивается i -й ячейке, то j -я частота не должна присваиваться этой ячейке, то есть

$$\sum_{\substack{q=j-(c_n-1) \\ q \neq j \\ 1 \leq q \leq m}}^{j+(c_n-1)} V_{iq}$$

имеет ненулевой выход в случае, если присвоение j -й частоты i -й ячейке нарушает ограничение по числу частот для каждой ячейки;

при ограничении из-за помех по соседнему каналу и ограничении из-за помех, обусловленных совместным расположением, если q -я частота в пределах расстояния c_{ip} от j -й частоты ($|j - q| < c_{ip}$) присваивается p -й ячейке, для $c_{ip} > 0$ и $p \neq i$ j -я частота не должна присваиваться i -й ячейке. Таким образом

$$\sum_{\substack{p=1 \\ p \neq i \\ c_p > 0}}^n \sum_{\substack{q=j-(c_p-1) \\ 1 \leq q \leq m}}^{j+(c_p-1)} V_{pq}$$

имеет ненулевой выход, если присвоение j -й частоты i -й ячейке нарушает ограничению из-за помех по соседнему каналу и ограничению из-за помех, обусловленных совместным расположением.

A и B являются постоянными коэффициентами ($A = B = 1$).

Исходя из вышесказанного, функция энергии для задачи распределения каналов в сотовой радиосети задается следующим образом:

$$E = \frac{A}{2} \sum_{i=1}^n \left(\sum_{q=1}^m V_{iq} - d_i \right)^2 + B \sum_{i=1}^n \sum_{j=1}^m \left(\sum_{\substack{q=j-(c_n-1) \\ q \neq j \\ 1 \leq q \leq m}}^{j+(c_n-1)} V_{iq} + \sum_{\substack{p=1 \\ p \neq i \\ c_p > 0}}^n \sum_{\substack{q=j-(c_p-1) \\ 1 \leq q \leq m}}^{j+(c_p-1)} V_{pq} \right) V_{ij}. \quad (3.15)$$

Известно, что для используемой модели НС гарантируется нахождение только локального минимума. Для того чтобы обеспечить более быструю сходимость алгоритма и увеличить частоту нахождения глобального минимума введены следующие дополнения функции энергии.

1. Для того чтобы ограничить А-член между двумя значениями, он модифицируется следующим образом:

$$-A f \left(\sum_{q=1}^m V_{iq} - d_i \right), \quad (3.16)$$

где $f(x) = A_{\max}$, если $x > A_{\max}$, $f(x) = A_{\min}$, если $x < A_{\min}$ и $f(x) = x$ в других случаях. Таким образом $A_{\max}$ и $A_{\min}$ являются соответственно верхней и нижней границей А-члена.

2. Используемый в уравнении энергии В-член, может принимать две формы: если $(t \bmod T) < \varpi$

$$-B \left(\sum_{\substack{q=j-(c_n-1) \\ q \neq j \\ 1 \leq q \leq m}}^{j+(c_n-1)} V_{iq} + \sum_{\substack{p=1 \\ p \neq i \\ c_p > 0}}^n \sum_{\substack{q=j-(c_p-1) \\ 1 \leq q \leq m}}^{j+(c_p-1)} V_{pq} \right) V_{ij}, \quad (3.17)$$

в других случаях

$$-B \left(\sum_{\substack{q=j-(c_n-1) \\ q \neq j \\ 1 \leq q \leq m}}^{j+(c_n-1)} V_{iq} + \sum_{\substack{p=1 \\ p \neq i \\ c_p > 0}}^n \sum_{\substack{q=j-(c_p-1) \\ 1 \leq q \leq m}}^{j+(c_p-1)} V_{pq} \right), \quad (3.18)$$

где t – число итерационных шагов, а T и ϖ – постоянные параметры.

3. К уравнению движения добавляется следующее выражение:

$$+Ch \left(\sum_{q=1}^m V_{iq} - d_i \right) \cdot (1 - V_{ij}), \quad (3.19)$$

где $h(x) = 1$ при $x < 0$ и $h(x) = 0$ при $x \geq 0$.

Значение C выбирается на каждом итерационном шаге таким образом, чтобы устранить ситуацию, при которой два или более обрабатывающих нейрона имеют одни и те же состояния. C -член помогает ij -у обрабатывающему нейрону иметь ненулевой выход, если менее чем d_i нейронов для i -й ячейки имеют ненулевой выход и $V_{ij} = 0$.

4. Вход обрабатывающего нейрона ограничивается двумя величинами:

$$\begin{aligned} U_{ij} &= U_{\max}, \text{ если } U_{ij} > U_{\max} \\ U_{ij} &= U_{\min}, \text{ если } U_{ij} < U_{\min}, \end{aligned} \quad (3.20)$$

где $U_{\max}$ и $U_{\min}$ соответственно постоянные верхняя и нижняя границы входного значения U_{ij} .

Алгоритм распределения каналов в сотовой радиосети, основанный на использовании уравнения движения (3.14), функции энергии (3.15) и ее дополнений может быть описан следующим образом.

1 этап. Эмпирическим путем устанавливаются используемые значения коэффициентов и параметров $A, B, C, t, T, \varpi, U_{\max}, U_{\min}, T_{\max}, UTR, LTR$.

2 этап. Рандомизируем начальные значения входа $U_{ij}(t)$ для $i = 1, \dots, n$ и $j = 1, \dots, m$ между 0 и $U_{\min}$. Назначаем начальные значения выхода $V_{ij}(t)$ для $i = 1, \dots, n$ и $j = 1, \dots, m$, равными нулю.

3 этап. Используя уравнение движения (3.14), вычисляем изменение входа $\Delta U_{ij}(t)$:

если $(t \bmod T) < \varpi$

$$\begin{aligned} \Delta U_{ij}(t) = & -A f \left(\sum_{q=1}^m V_{iq}(t) - d_i \right) - B \left(\sum_{\substack{q=j-(c_q-1) \\ q \neq j \\ 1 \leq q \leq m}}^{j+(c_q-1)} V_{iq}(t) + \sum_{\substack{p=1 \\ p \neq i \\ c_p > 0}}^n \sum_{\substack{q=j-(c_p-1) \\ 1 \leq q \leq m}}^{j+(c_p-1)} V_{pq}(t) \right) V_{ij}(t) + \\ & + C h \left(\sum_{q=1}^m V_{iq}(t) - d_i \right) \cdot (1 - V_{ij}(t)), \end{aligned} \quad (3.21)$$

в других случаях

$$\Delta U_{ij}(t) = -Af \left(\sum_{q=1}^m V_{iq}(t) - d_i \right) - B \left(\sum_{\substack{q=j-(c_n-1) \\ q \neq j \\ 1 \leq q \leq m}}^{j+(c_n-1)} V_{iq}(t) + \sum_{\substack{p=1 \\ p \neq i \\ c_p > 0}}^n \sum_{\substack{q=j-(c_p-1) \\ 1 \leq q \leq m}}^{j+(c_p-1)} V_{pq}(t) \right) + \\ + Ch \left(\sum_{q=1}^m V_{iq}(t) - d_i \right) \cdot (1 - V_{ij}(t)). \quad (3.22)$$

4 этап. Обновляем значение входа

$$U_{ij}(t+1) = U_{ij}(t) + \Delta U_{ij}(t). \quad (3.23)$$

5 этап. Используя выражение (2.20) модифицируем значение входа:

$$U_{ij}(t+1) = U_{\max}, \text{ если } U_{ij}(t+1) > U_{\max}$$

$$U_{ij}(t+1) = U_{\min}, \text{ если } U_{ij}(t+1) < U_{\min}. \quad (3.24)$$

6 этап. Используя гистерезисную активационную функцию, обновляем выход обрабатывающего нейрона $V_{ij}(t+1)$:

$$V_{ij}(t+1) = 1, \text{ если } U_{ij}(t+1) > UTR,$$

$$V_{ij}(t+1) = 0, \text{ если } U_{ij}(t+1) < LTR, \quad (3.25)$$

в других случаях не изменяется.

Здесь UTR, LTR – соответственно верхняя и нижняя точки обхода.

7 этап. Проверка условия окончания работы алгоритма.

$$\text{Если } \left(\sum_{q=1}^m V_{iq} - d_i \right) = 0 \text{ и } (V_{ij}(t) = 1 \text{ и } \left(\sum_{\substack{q=j-(c_n-1) \\ q \neq j \\ 1 \leq q \leq m}}^{j+(c_n-1)} V_{iq} + \sum_{\substack{p=1 \\ p \neq i \\ c_p > 0}}^n \sum_{\substack{q=j-(c_p-1) \\ 1 \leq q \leq m}}^{j+(c_p-1)} V_{pq} \right) = 0),$$

для $i = 1, \dots, n$ и $\exists j \in \{1, \dots, m\}$ или $t = T_{\max}$ то алгоритм работу заканчивает.

В других случаях увеличиваем значение t на единицу и возвращаемся на этап 2.

Для того чтобы сократить время сходимости при решении ряда задач, можно зафиксировать назначение частот для одной из ячеек или нескольких ячеек с самым большим числом требуемых частот. Например, из рис. 3.8а видно, что четвертая ячейка имеет наибольший элемент в векторе спроса, поэтому логичным является фиксирование присвоения частот этой ячейке, что и делалось при моделировании. Кроме того, фиксированное назначение частот может быть априорно заданно, например на случай расширения сети, для того чтобы не было совпадений с частотами соседних сетей.

Результаты моделирования показали, что сходимость рассматриваемого алгоритма значительно выше, например известного алгоритма Кунца [84]. Так для 25-и ячеечной сети и 73-х распределяемых частотах для решения потребовалось 200 итерационных шагов, против 2450 при использовании алгоритма Кунца.

Модификация этого алгоритма может быть направлена на учет большого количества всевозможных ограничений, которые могут быть заданы матрицей ЭМС.

Заключение

В последние десятилетия в мире отмечается быстрое развитие нейроинформационных технологий. Актуальность исследований в этом направлении подтверждается большим количеством различных применений нейроинформационных систем. Это автоматизация процессов распознавания образов, адаптивное управление, аппроксимация функционалов, прогнозирование, создание экспертных систем и многие другие приложения.

С помощью нейроинформационных систем можно, например, управлять телекоммуникационными сетями, проводить динамичную диагностику и терапию широкого круга заболеваний, предсказывать показатели биржевого рынка, выполнять распознавание звуковых сигналов, создавать самообучающиеся системы, способные управлять оружием и оценивать ситуацию складывающуюся на поле боя.

Особый интерес вызывает применение нейроинформационных технологий в инфотелекоммуникационных системах. Высокая эффективность нейроинформационных технологий при решении задач адаптивного управления динамичными системами может уже в ближайшее время сделать их незаменимыми при создании новых поколений сетей мобильной связи и других беспроводных сетей.

Необходимо отметить также, что в то время, как на Западе нейроинформационные технологии применяются уже достаточно широко, в России это все еще экзотика. Это вызывает потребность с одной стороны, в повышении усилий по их популяризации, а с другой – в расширении исследований в этом направлении. Уже сейчас достаточно очевидно, что объединение нейроинформационных технологий с другими технологиями будет способствовать существенному прорыву в решении многих актуальных проблем.



Список литературы

1. Аведьян Э. Д., Коваленко М. Л., Цитоловский Л. Е., Цыпкин Я. З. Автономные алгоритмы настройки нейронных сетей. – Тр. Международной конференции “Математика, компьютер, управление и инвестиции”. – М.: Гарант, 1993, с.2–11.
2. Аппаратная реализация ускорителя нейровычислений. / А.В. Бочаров, А.С. Грошев, М.В. Захватов и др. – Известия вузов. Приборостроение, 1995, т. 38, №1–2.
3. Бабков В.Ю., Михайлов П.А., Смирнов Д.А. Методы и алгоритмы частотно-территориального планирования сотовых сетей подвижной радиосвязи на основе геоинформационной базы данных. /VI Санкт-Петербургская международная конференция “Региональная информатика-98”. Тезисы докладов. Часть 1. – СПб., 1998. с. 57.
4. Барцев С.И., Охонин В.А. Адаптивные сети обработки информации. // Препринт ИФ СО АН СССР, Красноярск, 1986, №59Б.
5. Блум Ф., Лейзерсон А., Хофстедтер Л., Мозг, разум и поведение. – М.: Мир, 1988.
6. Бовбель Е. И., Паршин В. В. Нейронные сети в системах автоматического распознавания речи. – Зарубежная радиоэлектроника. Успехи современной радиоэлектроники, 1998, №4, с.49–65
7. Бубенников А. Н. Архитектурно-технологический облик интеллектуальных нейронных сетей на кремниевых пластинах и трехмерных нейрокомпьютеров. – Зарубежная радиоэлектроника. Успехи современной радиоэлектроники, 1998, №4, с.34–51.
8. Галушкин А. И. Нейрокомпьютеры в разработке военной техники США. – Зарубежная радиоэлектроника, 1995, №5.
9. Галушкин А. И. Нейрокомпьютеры в разработке военной техники США. – Зарубежная радиоэлектроника, 1995, №6, с.4–21.
10. Галушкин А. И., Крысанов А. И. Оценка производительности нейрокомпьютеров. – Зарубежная радиоэлектроника. Успехи современной радиоэлектроники, 1998, №4, с.22–33.
11. Галушкин А. И. Синтез многослойных систем распознавания образов. – М.: Энергия, 1974.
12. Галушкин А. И. Современные направления развития нейрокомпьютеров в России. – Зарубежная радиоэлектроника. Успехи современной радиоэлектроники, 1998, №4, с.3–17.
13. Горбань А. Н. Обучение нейронных сетей. – М.: СП ПараГраф, 1990.
14. Джейн А. К., Мао Ж., Монуддин К. М. Введение в искусственные нейронные сети. – Открытые системы, 1997, №4, с.16–24.
15. Дуда Р., Харт П. Распознавание образов и анализ сцен. – М.: Мир, 1976.
16. Ежов А., Чечеткин В. Нейронные сети в медицине. – Открытые системы, 1997, №4, с.34–37.
17. Захарченко В. М. Проблемы и перспективы создания нейроподобных устройств обработки информации. Обзоры по электронной технике. Сер. 9. Экономика и системы управления. Выпуск 1 (1006). – М.: ЦНИИ “Электроника”, 1984.

18. Комашинский В. И., Смирнов Д. А., Титов А. А., Шнуренко С. А. Алгоритм децентрализованного управления доступом для сети радиосвязи с пакетной передачей сообщений. /53 НТК. Тезисы докладов. – СПб.: НТОРЭС им. А.С. Попова, 1998, с.40.
19. Комашинский В.И., Смирнов Д.А. Введение в нейро-информационные технологии. – СПб.: Тема, 1999.
20. Комашинский В. И., Смирнов Д. А., Шнуренко С. А. Интегральное обслуживание пользователей в сотовых системах подвижной радиосвязи при множественном доступе с кодовым разделением. /51 НТК. Тезисы докладов. – СПб.: ГУТ, 1998, с.58.
21. Комашинский В. И., Смирнов Д. А. Гибридные методы множественного доступа к беспроводным АТМ сетям. /54 НТК. Апрель 1999 г. Тезисы докладов. – СПб.: НТОРЭС имени А.С. Попова, 1999, с.66–67.
22. Крысанов А. И. СБИС L-Neuro – базовый нейрочип для создания современных нейрокомпьютеров. – Зарубежная радиоэлектроника. Успехи современной радиоэлектроники, 1998, №4, с.18–21.
23. Куссуль Э. М. Ассоциативные нейроподобные структуры. – Киев: Наукова думка, 1990.
24. Ли У. Техника подвижных систем связи: пер. с англ. – М.: Радио и связь, 1985.
25. Логовский А. С. Использование нейронных сетей для решения комбинаторных задач с полным перебором. – Нейрокомпьютер, 1994, №3, 4, с.41–50.
26. Маккалок У. С., Питтс У. Логическое исследование идей, относящихся к нервной активности, сборник «Автоматы», под ред. К. Э. Шеннона и Дж. Маккарти, – ИЛ, 1956.
27. Песиков Я. С., Рыбалко С. Я. Атлас клинической аурикулотерапии. – М.: Медицина, 1990.
28. Портнов Ф. Г. Электростимуляторная рефлексотерапия. – Рига: Знание, 1988.
29. Розенблатт Ф. Принципы нейродинамики (перцептроны и теория механизмов мозга). – М.: Мир, 1965.
30. Соколов В. Н. Нейронные механизмы памяти и обучения. – М.: Наука, 1981.
31. Табеева Д. М. Руководство по иглорефлексотерапии. – М.: Медицина, 1990.
32. Уоссерман Ф. Нейрокомпьютерная техника: Теория и практика. – М.: Мир, 1990.
33. Хант Э. Искусственный интеллект. – М.: Мир, 1978.
34. Цыпкин Я. З. Адаптивные методы выбора решений в условиях неопределенности. – Автоматика и телемеханика, 1976, №4, с.78–91.
35. Цыпкин Я. З. Основы теории обучающихся систем. – М.: Наука, 1970.
36. Шаров А.Н., Степанец В.А., Комашинский В.И. Сети радиосвязи с пакетной передачей информации. – СПб.: ВАС, 1994.
37. Якушев Д. Ж. Применение нейрокомпьютеров в финансовой деятельности. – Зарубежная радиоэлектроника. Успехи современной радиоэлектроники, 1998, №4, с.61–68.
38. Яфраков М. Ф., Корчагина Л. И. Особенности комплексного подхода к нейрокомпьютерингу. – Известия вузов. Приборостроение, 1997, т.40, №3.

39. Aazhang B., Paris B.-P., Orsak G. C. Neural Networks for Multiuser Detection in CDMA Communications, *IEEE Trans. Comm.*, July 1992, vol. 40, pp. 1212–1222.
40. Anderson J. A., Pellionisz A., Rosenfeld, E. *Neurocomputing 2*. Cambridge, 1990, MA: MIT Press.
41. Anderson J. A., Rosenfeld E., *Neurocomputing: Foundation of Research*, MIT Press, Cambridge, Mass., 1988.
42. Anumolu V., Bray N. W., Reilly K. D. A neural network model of strategy selection and evolution. *World Congress on Neural Networks - '93*, pp. 1528–1535.
43. Astrom K. J., Wittenmark B. *Adaptive control.*, Reading MA: Addison Wesley, 1989.
44. Atsushi H. ATM communications network control by neural network. *IJCNN*, Washington DC, 1989, vol. 1, pp.259–266.
45. Atsushi H. ATM communications network control by neural networks. *IEEE Trans. on Neural Networks*, vol. 1, №1, 1990.
46. Bitto J., Magorla N., Hush D., Chael E. Seismic event discrimination using neural networks. 23rd Asi. Conf. Signals, Syst. And Comput., Pacific Grove, Calif., Oct.30-No, vol. 1, 1989.
47. Bourlard A., Wellekens C. J., Links between Markov models and multilayer perceptrons, *IEEE Trans. Patt. Anal. and Mach. Intell.*, 1990, vol. 12, pp.1167–1178.
48. Brown T. X. Neural network for switching. *IEEE Comm. Mag.*, Nov., 1989, pp.72–81.
49. Brown T. X., Liu Kuo-Hui. Neural network design of a Banyan network controller. *IEEE J. Selec. Areas Commun.*, 1990, vol. 8, №8, pp.1428–1438.
50. Campbell M., Toberg S., 3-D Wafer scale Architectures for neural network computing., *IEEE Trans Veh. Technol.*, 1993, vol. CHMT-16, no. 7, pp.646–655.
51. Carpenter G. A., Grossberg S., *Pattern Recognition by SelfOrganizing Neural Networks*, MIT Press, Cambridge, Mass., 1991.
52. Duranton M., Aglan F., Mauduit N., *Hardware Accelerations for Neural Networks: Simulations in Parallel machines*, L-Neuro user Guide, Telmat Multinode.
53. Fortuna L., Geaziani S., Presti M. L., Muscato G. Improving back-propagation learning using auxiliary neural networks. *Int. J. Control*. 1992. v.55, №4, pp.793–807.
54. Fujii R., Tenorio M. F., Zhu H. Use of neural nets in channelrouting. *IJCNN-89*, Washington DC, 1989, vol. 1, pp.321–325.
55. Funabiki Nobuo, Takefuji Yoshiyasu, Lee Cuo Chun. A neural network model for traffic control in multistage interconnection networks. *IJCNN-91*, Seattle, Wash., July 8–12, 1991, vol. 2, p.898.
56. Funabiki N. And oth. A neural network approach of broadcasting in multishop packed radio networks. *IJCNN-91. Theory and applications*, Singapore, 1991, pp. 2540–2545.
57. Funabiki N., Takefuji Y. A parallel algorithm for spare allocation problems. *IEEE Trans. Reliability*, vol. 40, Aug. 1991, pp.338–346.
58. Funabiki N., Takefuji Y. A parallel algorithm for channel routing problems. *IEEE Trans. Computer-Aided Design*, vol. 11, Apr. 1992, pp.464–474.
59. Funabiki N., Takefuji Y. Neural network parallel algorithm for channel assignment problems in cellular radio networks. *IEEE Transactions. Veh. Technol.*, vol. 41, №4, Nov. 1992, pp.430–437.

60. Gael de La Croix Vaubois, Catherine Moulinoux, Benolt Derot, The N Programming Language //Neurocomputing, NATO ASI series, vol.F68, pp.89-92.
61. Gamst A., Rave W. On frequency assignment in mobile automatic telephone system, in Proc. GLOBECOM'82 1982, pp.309-315.
62. Ghez S., Verdu S., Schwartz Stuart C., Optimal Decentralized Control in the Random Access Multipacket Channel, IEEE Trans. Automat. Contr., vol. 34, pp. 1153-1163, Nov. 1989.
63. Grossberg S. Adaptive pattern classification and universal recoding: I/ Parallel development and coding of neyral feature detectors. Biol. Cybern., 1976, v.23, pp.121-134.
64. Gyu-Sang Jung, Farid Dowla, Vemuri V. Application of neural networks for seismic phase identification. IJCNN-91, Singapore, 1991, pp.899-904.
65. Hajek B., Sasaki G. Link Sceduling in Polynomial Time. IEEE Trans. Inform. Theory, Sept. 1988, vol. 34, pp.910-917.
66. Hao J., Vandewalle J., Tan Sh. Predictive control of nonlinear systems based on identification by backpropagation networks. J. Neural Systems, 1994, v. 5, №4. pp. 335-344.
67. Hale W.K. Frequency assignment: Theory and applications. Proc. IEEE, vol. 68, Dec. 1980, pp.1497-1514.
68. Harris Drucker, Yann Le Cun, Improving Generalization Performance Using Backpropagation //IEEE Transactions on Neural Networks, Vol.3, N5, 1992, pp.991-997.
69. Hata M. Empirical formula for propagation loss and land mobile radio services. IEEE Transactions., 1980, VT-29(3).
70. Haykin S., Neural Networks: A comprehensive foundation, MacMillan College Publishing Co., New York, 1994.
71. He G., Tang P., Pang X. Neural Networks Approaches to Implementation of Optimum Multiuser Detection in CDMA Channals, Intl. Journ. Of Electronics, 1996, vol. 80, pp.425-431.
72. Hebb D. O., The organization of behaviour, N. Y.: Wiley, 1949.
73. Heermann P. D. Neural network techniques for stable learning control of nonlinear systems. Dissertation D. S. University of Texas at Austin, 1992.
74. Hertz J., Krogh A., and R.G. Palmer, Introduction to the Theory of Neural Computation, Addison-Wesley, Reading, Mass., 1991.
75. Hirose Y., Yamashita K., Hijiya S. Back-propagation algorithm which varies the number of hidden units. Neural Networks. 1991, v.4, pp.61-66.
76. Hopfield J., Learning algorithms and probability distributios in feed-forward and feed-back networks, Proc. Natl. Acad. Sci. USA, 1987, vol.84, pp.8429-8433.
77. Hopfield J., Tank D., Neural computation of decision in optimization problems, Biol. Cybernet, 1985, vol.52, pp.141-152.
78. Hopfield J., Neural Networks and Physical Systems with Emergent Collective Computational Abilities, in Proc. National Academy of Sciencies, USA 79, 1982, pp. 2554-2558.
79. Hornik K., Stinchcombe M., White H. Mylti-layer feed-forward networks are universal approximators. Discussion paper, Department of Economics, University of California, San Diego, La Jolla, CA, 1988.

80. Keller J. B., Geometrical theory of diffraction. Opt. Soc. Am., 1962, vol.52, pp.116-130.
81. Keun-Rong Hsieh and Wen-Tsuen Chen, A Neural Network Model which Combines Unsupervised and Supervised Learning, IEEE Trans. on Neural Networks, vol.4, No.2, march 1993.
82. Kohonen T., SelfOrganization and Associative Memory, Third Edition, Springer-Verlag, New York, 1989.
83. Kyung S. Kwak and Ramesh R. Rao, Controlled ALOHA with Geometric Capture Probabilities, IEEE Trans. Automat. Contr., vol. 35, pp. 1270-1274, Nov. 1990.
84. Kunz D. Channel assignment for cellular radio using neural networks. IEEE Transactions. Veh. Technol., vol. 40, Feb. 1991, pp.188-193.
85. Levine D. S., Leven, S. J. Motivation, emotion, and goal direction in neural networks, 1992, Hillsdale: Erlbaum.
86. Lippman R. P., An Introduction to Computing with Neural Nets, IEEE ASSP Magazine, Vol.4, No.2, Apr. 1987, pp. 4-22.
87. Lippman R. P., Review of neural networks for speech recognition, Neural Computation, 1991, vol.1, no. 1, p.1-38.
88. Malki H. A., Moghaddamjoo A. Using the Karhunen-Loe`ve Transformation in the Back-Propagation Training Algorithm //IEEE Transactions on Neural Networks, Vol.2, N1, 1991, pp.162-165.
89. Manduit N. Preliminary data sheet: the neuromimetic circuit L-Neuro 1.0, L-Neuro User Guide. Telmat Multinode.
90. Morgan N., Bourlard A. Neural networks for statistical recognition of continuous speech, Proceedings of the IEEE. Concepts and Theory, May 1995, vol.83, pp.742-769.
91. Narendra K. S., Parthasarathy K. Back propagation in dynamical systems containing neural networks. TR 8905, Center for systems science, Department of Electrical Engineering, Yale University, New Haven, CT, 1989.
92. Narendra K. S., Parthasarathy K. Neural networks and dynamical systems. Part 1: A gradient approach to Hopfield networks. TR 8820, Center for systems science, Department of Electrical Engineering, Yale University, New Haven, CT, 1988.
93. Neural networks for control / Miller W. T., Sutton R. S., Werbos P. J., Eds., The MIT Press, 1990.
94. Odom M. D., Sharda R., Oklahoma State University A Neural Network Model for Bankruptcy Prediction IJCNN, S. D., 1990, vol.2, pp. 163-168.
95. Paielli R.A. Simulation tests of the optimization method of Hopfield and Tank using neural networks. NASA Tech. Mem. 101047, Nov. 1988.
96. Parks P. C. Lyapunov redesign of model reference adaptive control systems, IEEE Journal Transactions of Automatic Control, 1966, vol. 11, pp.362-367.
97. Rauch H. E., Winarske T., Neural networks for Routing Communication Traffic, IEEE Control Syst. Mag., 1988, vol.8, pp.26-31.
98. Richard M. D., Lippman R. P., Neural network classifiers estimate Bayesian discriminant function, Neural Computation. Concepts and Theory, 1991, vol. 3, pp.461-483.
99. Rumelhart D. E., Hinton G. E., Williams R. J., Learning internal representations by error propagation, Parallel Distributed Processing. Cambridge, MA: MIT Press, 1986, vol. 1, Ch. 8, pp.318-362.

100. Sankar K. Pal, Sushmita Mitra, Multilayer Perceptron, Fuzzy Sets, and Classification //IEEE Transactions on Neural Networks, Vol.3, N5,1992, pp.683–696.
101. Sivarajan K.N., McEliece R.J., Ketchum J.W. Channel assignment in cellular radio. In Proc. 39th IEEE Veh. Technol. Soc. Conf., May 1989, pp.846–850.
102. Stocker K. E., Landstorfer F. M. Empirical prediction of radiowave propagation by neural network simulator. Electronics letters, June 1992, vol.28, pp.1177–1178.
103. Troudet Terry P., Walters Stephen M. Neural network architecture for crossbar switch control. IEEE Transactions Circuits and Syst., 1991, vol. 38, №1, pp.42–56.
104. Waibel A., Hampshire J., A novel objective function for improved phoneme recognition using time-delay neural networks, IEEE Transactions on Neural Networks, June 1990, vol. 1, no. 2, pp.216–228.
105. Waibel A., Phoneme recognition using time-delay neural networks, IEEE Transactions on Acoustics, Speech and Signal Processing, March 1989, vol. 37, pp.328–339.
106. Wen-Yuan Chen, Sin-Horng Chen, Cheng-Jung Lin, A speech recognition method based on the sequential multi-layer perceptrons, Neural Networks, 1996, vol.9, no. 4, pp.655–669.
107. Werbos P. J., Backpropagation Through Time: What It Does and How to Do It //Artificial Neural Networks: Concepts and Theory, IEEE Computer Society Press, 1992, pp.309–319.
108. Werbos P. J., Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, Phd Thesis, Dept. of Applied Mathematics, Harvard University, Cambridge, Mass., 1974.
109. Wilson G.V., Pawley G.S. On the stability of the traveling salesman problem algorithm of Hopfield and Tank. Biological Cybern., vol. 58, 1988, pp.63–70.
110. Widrow B., Lehr M. A., 30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation //Artificial Neural Networks: Concepts and Theory, IEEE Computer Society Press, 1992, pp.327–354.
111. Wieselthier J. E., Barnhart C. M., Ephremides A., A Neural Networks Approach to Routing Without Interference in Multihop Radio Networks, IEEE Transactions on Comm., 1994, vol.42, no.1, pp.166–177.
112. Wieselthier J. E., Barnhart C. M., Ephremides A. The problem of routing and scheduling in multihop radio network – a hopfield neural network approach. NRL Report 9366, Dec. 1991.
113. White M., Mack I., Borsuk J., Lampe D., Kub F., Charge-Coupled Device (CCD) Adaptive Discrete Analog Signal Processing, IEEE Journal of Solid-State circuits, Feb. 1979, vol.SC-14, no.1.





Содержание

Предисловие	3
Введение	5
Глава 1. Основы теории искусственных нейронных сетей	7
1.1. Нейронные сети: общие положения	7
1.2. Архитектура искусственных нейронных сетей	12
1.2.1. Искусственные нейронные сети прямого распространения	12
1.2.2. Искусственные нейронные сети с обратными связями	23
1.3. Обучение искусственных нейронных сетей	23
1.3.1. Обучение с учителем	24
1.3.2. Обучение без учителя	29
1.4. Нейрочипы и нейрокомпьютеры	34
Глава 2. Применение искусственных нейронных сетей в системах управления.....	47
2.1. Особенности использования искусственных нейронных сетей в системах управления	47
2.2. Использование нейронных сетей в системах управления p-настойчивым множественным доступом	53
Глава 3. Применение нейронных сетей в телекоммуникационных Системах	60
3.1. Основные направления применения искусственных нейронных сетей в телекоммуникационных системах ...	60
3.2. Нейронные сети в системах автоматического распознавания речи	62
3.3. Использование нейронных сетей для решения задач маршрутизации	67
3.4. Использование нейронных сетей при планировании сотовых сетей подвижной радиосвязи	75
3.5. Использование нейронных сетей для распределения каналов в сотовых радиосетях	79
Заключение	87
Список литературы	88



**Издательское предприятие редакции журнала “Радиотехника”
продолжает выпуск серии книг
“Нейрокомпьютеры и их применение”
под редакцией докт. тех. наук, проф. А.И.Галушкина**

- *Книга 1. А. И. Галушкин. Теория нейронных сетей.
- *Книга 2. Сигеру Омату, Марзуки Халид, Рубия Юсоф. Нейроуправление и его приложения (пер. с англ.; под ред. А.И. Галушкина, В.А. Птичкина).
- Книга 3. А. И. Галушкин. Нейрокомпьютеры.
- *Книга 4. Нейронные сети: история развития (под ред. А. И. Галушкина и акад. Я. З. Цыпкина).
- Книга 5. В. А. Терехов. Нейросетевые системы управления.
- *Книга 6. В. А. Головкин. Нейронные сети: обучение, организация и применение.
- Книга 7. Н.И. Червяков, П.А. Сахнюк, А.В. Шапошников. Модулярные нейрокомпьютеры высокой производительности и надежности с параллельно-конвейерной архитектурой.
- Книга 8. А.П. Татузов. Нейронные сети в задачах радиолокации.
- *Книга 9. Нейроматематика (Сб. статей).
- Книга 10. Нейронные сети. "STATISTICA Neural Networks" (пер. с англ. StatSoft Russia).
- Книга 11. Э. Д. Аведьян. Алгоритмы обучения нейронных сетей.
- Книга 12. А. Ю. Дорогов. Теория модульных нейронных сетей.
- Книга 13. Нейрокомпьютеры в системах обработки изображений (коллектив авторов).
- Книга 14. А. И. Галушкин. Применение нейрокомпьютеров.
- Книга 15. А. А. Фролов, И. П. Муравьев. Ассоциативная память. Нейросетевой подход.
- Книга 16. Ю. И. Зозуля. Интеллектуальные нейросистемы (под ред. А. И. Галушкина).
- Книга 17. Р. Рохас. Нейронные сети: систематизированное введение. (пер. с англ. под ред. А. И. Галушкина, Г. Г. Губайдуллина, Ю. И. Зозули).
- Книга 18. Нейрокомпьютеры в вертолетах (коллектив авторов).
- Книга 19. Н. П. Абовский. Нейроуправляемые конструкции и системы.
- Книга 20. А. А. Фролов, Е. В. Бирюкова, Р. А. Прокопенко. Нейросетевое управление антропоморфными роботами и манипуляторами.
- Книга 21. Э. Ю. Кирсанов. Системотехническое проектирование высокопроизводительных цифровых нейрокомпьютеров с параллельной структурой.
- Книга 22. А. И. Галушкин, В. А. Шахнов, А. И. Власов. Нейропроцессорные системы активного гашения промышленных вибраций.
- Книга 23. А. И. Власов, Г. Л. Яковлев. Применение нейросетевых методов в информационных и аналитических системах.
- Книга 24. А. И. Власов. Автоматизированное проектирование нейроадаптивных систем активного управления волновыми полями.
- Книга 25. Нейрокомпьютеры в космической технике (коллектив авторов).
- Книга 26. Нейрокомпьютеры в авиации (коллектив авторов).
- Книга 27. В. И. Горбаченко. Применение нейрокомпьютеров для решения краевых задач теории поля (под ред. А.И. Галушкина).
- Книга 28. Лой, Лей, Лай. Применение интеллектуальных систем в энергетике. Эволюционное программирование и нейронные сети (пер. с англ.). Y. Wiley and Sons.
- Книга 29. Л. Е. Назarov. Применение нейронных сетей в задачах обработки космических изображений земной поверхности.
- Книга 30. А. В. Тимофеев. Нейронные сети с самоорганизующейся архитектурой и мультиагентные технологии.
- Книга 31. В. В. Борисов, И. В. Огнев. Ассоциативные нейронные сети.
- Книга 32. А. И. Радченко. Механизмы нейронной памяти и модели амнезий.
- *Книга 33. Н. Г. Рамбиди, Е. П. Гребенников, А. Г. Девятков, Д. В. Яковенчук. Биомолекулярные нейронные системы.
- Книга 34. В. В. Круглов, В. В. Борисов. Нечеткие нейронные сети.
- Книга 35. А. А. Жданов. Системы автономного искусственного интеллекта.
- Книга 36. В. А. Карасев, В. В. Лучинин. Бионические микро- и наносистемы. Синтез и функционирование.

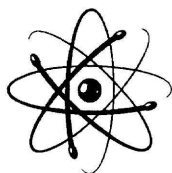
* Вышедшие книги.

Сделать предварительные заявки на книги, а также получить дискету с подробным планом-проектом каждой книги можно в ИПРЭР.

**103031, г. Москва, К-31, Кузнецкий мост, 20/6.
Тел/факс: (095) 925-92-41, 925-78-72, 921-48-37**

softline

www.softline.ru



119991 г. Москва,
ул. Губкина, д. 8
тел.: (095) 232-0023
e-mail: info@softline.ru

Все для науки

Научное ПО для исследований и расчетов

Почему студенты, преподаватели и научные работники приобретают нужные им программы в компании SoftLine?

- Низкие цены – компания работает напрямую с вендорами и является привилегированным партнером по образовательным программам для многих производителей ПО.
- Возможность получения демо-версий и обновлений для широкого спектра программ.
- Возможность удобного выбора программ по каталогу SoftLine-direct или на сайте www.softline.ru.
- Поддержка сообщества пользователей на сайте www.exponenta.ru.

Какое научное программное обеспечение поставляется SoftLine?

- Языки программирования математических задач (MathWorks)
- Моделирование электронных схем и устройств (Electronics Workbench, PCAD)
- Универсальные математические пакеты, обмен данными с AutoCAD (Wolfram Research, Waterloo Maple, MathSoft)
- ПО для химиков (CambridgeSoft)
- Статистические пакеты (MathSoft, StatSoft)
- Редакторы формул (Design Science, Multi Edit)
- Пакеты бизнес-анализа (Palisade)
- Математические надстройки к MS Office (Frontline Systems)

SoftLine – это свобода выбора

Обратившись в SoftLine, вы в кратчайшие сроки решите проблемы с программным обеспечением. Получив консультации менеджеров, вы подберете все необходимые инструменты для работы в вашей области. Компания SoftLine поможет вам также в выборе обучающих курсов и пригласит на регулярно проводимые семинары по математическим и статистическим пакетам.

The
**MATH
WORKS**
Inc.

MathSoft

$$\Sigma + \sqrt{\quad} = f + \delta$$



StatSoft

WOLFRAM
RESEARCH

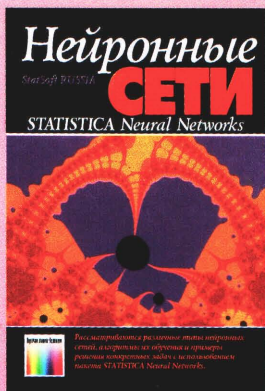
Waterloo Maple



Electronics

ORIGINATOR OF THE ELECTRONICS WORKBENCH

CambridgeSoft
Internet Chemistry Software Leader



ISBN 5-93517-015-9

Нейронные сети. STATISTICA Neural Networks: Пер. с англ.

В книге популярно изложены методы анализа данных, основанные на построении нейросетевых моделей. Рассматриваются различные типы нейронных сетей, алгоритмы их обучения и примеры решения конкретных задач с использованием пакета. *STATISTICA Neural Networks*

Заключительная глава содержит краткое руководство для пользователей. *STATISTICA Neural Networks*

Книга представляет интерес в первую очередь для тех, кто начинает изучать нейросетевые методы и хочет научиться грамотно решать практические задачи, используя современные подходы к обработке данных.



ISBN 5-93517-031-0

Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. 2-е издание.

Книга посвящена одному из современных направлений в области информатики и вычислительной техники — нейрокompьютерным технологиям. Достоинством книги является то, что в ней рассмотрены не только вопросы теории искусственных нейронных сетей, но и большое внимание уделено современным программным оболочкам-имитаторам нейронных сетей, а также решению с их помощью практических задач распознавания образов, кластеризации, прогнозирования, оптимизации, построения и использования нейросетевых экспертных систем. Книга содержит обширный справочный материал.

Для научных и инженерно-технических работников в области информатики и вычислительной техники, занимающихся созданием и использованием интеллектуальных систем, а также аспирантов и студентов разных специальностей в области компьютерных технологий.

Books[®]

Харьков, ул. Сумская, 51 т. 7-140-470
www.books.ua

НДС 20.00 % 1762-23325



9 785935 170943
Нейронные сети и их
прим. в сис. упр. и св.

Цена 21.50 грн.

издательства
«ОМ» на сервере:

k.ru

ISBN 5-93517-094-9

