

ЯРОСЛАВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
им. П. Г. ДЕМИДОВА

на правах рукописи

Лукиянов Александр Владимирович

**Исследование нейроподобных сетей, работающих со
средним значением стохастического потока**

специальность 05.13.17 — теоретические основы информатики

Диссертация на соискание ученой степени кандидата физико-математических наук

научный руководитель:

доктор физико-математических наук,
профессор Тимофеев Е. А.

научный консультант:

кандидат физико-математических наук,
профессор Соколов В. А.

Ярославль

2000

Содержание

1. Введение	3
1.1. Формальная модель нейрона	4
1.2. Некоторые подходы к аппаратной реализации искусственных нейронных сетей	5
1.3. Импульсное кодирование информации в биологических нейронных сетях	7
1.4. Клеточные нейронные сети	8
1.5. Нейроны с альтернативными синапсами	8
1.6. Дискретное преобразование Фурье	9
1.7. Обзор диссертации	10
2. Потокное представление информации	21
2.1. Общие определения	21
2.2. Представление значений из $[0; 1]$	22
2.3. Представление значений из $[-1; 1]$	25
2.4. Представление комплексных значений	27
3. Поточный нейрон	30
3.1. Основные элементы нейрона	30
3.2. Описание работы нейрона	32
3.3. Вычисление средних значений	33
3.4. Обоснование перехода к линейной модели	35
3.5. Полносвязная сеть и ее обучение	37
3.6. Результаты эксперимента	39
4. Поточный нейрон с альтернативными синапсами	44
4.1. Основные элементы нейрона	44
4.2. Описание работы нейрона	46
4.3. Значения на выходе нейрона	47
4.4. Ассоциативная память на сети Хопфилда	49
4.5. Обучение по методу Хебба	50
4.6. Оптимизационное обучение	51
4.7. Результаты эксперимента	52

5. Потокое устройство, выполняющее дискретное преобразование Фурье	57
5.1. Значения и их представление	58
5.2. Описание схемы ДПФ	59
5.3. Обоснование	60
5.4. Сходимость к среднему значению	61
5.5. Результаты моделирования	62
5.6. Сравнение с обычной реализацией	63
5.7. Выводы	63
6. Заключение	66
Список рисунков	67
Список таблиц	68
Литература	69

1. Введение

Искусственные нейронные сети в последние десятилетия применяются для решения большого класса задач, для которых неизвестны эффективные алгоритмы, или требуется быстрая параллельная обработка данных. В этот класс входят задачи обработки изображений [4, 30], задачи распознавания оптических образов [44, 63], звуковых сигналов [57], организации ассоциативной памяти [9, 10, 48], предсказания показателей биржевых рынков [5], синтеза речи [60] и многие другие.

В основу искусственных нейронных сетей (ИНС) положены следующие черты биологических нейронных сетей, позволяющие им хорошо справляться со сложными задачами с неизвестными принципами решения: имеется простой обрабатывающий элемент — нейрон; очень большое число нейронов участвует в обработке информации; один нейрон связан с большим числом других нейронов (глобальные связи); веса связей между нейронами могут изменяться; информация обрабатывается параллельно.

Сложность нейронной сети определяется количеством нейронов, количеством связей между ними и сложностью отдельного нейрона. В диссертации разработаны новые модели нейронов, которые являются более простыми для аппаратной реализации по сравнению с другими моделями. Связи между разработанными нейронами состоят всего из двух физических линий, по которым передается два бита за единицу времени. Это достигается за счет использования кодирования информации в виде среднего значения стохастической последовательности.

При разработке искусственной нейронной сети, как правило, строится формальная модель нейрона, которая изучается математическими методами, и для которой разрабатывается алгоритм обучения. На основе формальной модели может быть создана аппаратная реализация ИНС, которая обладает свойствами изученной формальной модели и обучается теми же методами, что и формальная модель. Первая формальная модель нейрона была предложена У. Мак-Каллоком и В. Питтсом [52]. Другие формальные модели нейронов и нейронных сетей предлагались Ф. Розенблаттом [59] (перцептрон), Дж. Хопфилдом [48] и другими.

Поскольку искусственные нейросети разрабатывались на основе принципов работы биологических нейронных сетей, они унаследовали их некоторые свойства: нечеткую логику, выраженную в виде алгебраических уравнений, возможность обучения, параллельность выполнения операций. При обучении сеть адаптируется для решения

конкретной задачи, при этом выделяются неявно выраженные закономерности [3, 39].

Обучение является существенным элементом в разработке нейронной сети. Выбор метода обучения может сильно влиять на эффективность работы нейросети. В диссертации предложены два метода обучения разработанных нейронов: модифицированный метод Хебба и метод оптимизации приближенной функции ошибки. Проведенные эксперименты на имитационной модели позволили сравнить эффективность этих методов обучения.

Аппаратная реализация ИНС обладает свойством массового параллелизма, что позволяет ей обрабатывать данные существенно быстрее обычного компьютера [40, 49]. Кроме того, в некоторых специализированных управляющих устройствах может быть выгодно использовать простую аппаратную нейросеть вместо достаточно сложного универсального компьютера. Это делает актуальной разработку аппаратных нейросетей с небольшими аппаратными затратами.

Одним из методов, позволяющих уменьшить аппаратные затраты на реализацию нейронных сетей, является импульсное кодирование информации, которое обсуждается в разделе 1.2. Импульсное кодирование присутствует также и в биологических нейронных сетях, этот вопрос рассматривается в разделе 1.3. Исследование искусственных нейронных сетей, основанных на импульсном кодировании, представлено работами А. Ф. Мюррея, М. Томлинсона, Дж. Томберга, Ю. А. Маматова, Г. П. Штерна, А. К. Карлина, А. Н. Малкова и Е. А. Тимофеева. В диссертации предложена модель усовершенствованного и упрощенного нейрона, работающего с потоками импульсов, количество генераторов случайных чисел в котором не зависит от количества синапсов.

Опишем более подробно некоторые аспекты нейронных сетей, необходимые в дальнейшем: формальную модель нейрона, подходы к аппаратной реализации ИНС, импульсное кодирование информации в биологических нейросетях, клеточные нейронные сети, формальную модель нейрона с альтернативными синапсами, а также дискретное преобразование Фурье.

1.1. Формальная модель нейрона

Наиболее распространенная формальная модель нейрона, впервые предложенная У. Мак-Каллоком и В. Питтсом [52], имеет N входов и один выход. На входы поступают действительные значения x_i , которые умножаются на весовые коэффициенты

w_i и суммируются, эту взвешенную сумму называют состоянием нейрона:

$$s(\mathbf{x}, \mathbf{w}) = w_0 + \sum_{i=1}^N x_i w_i, \quad (1)$$

где w_0 — некоторое пороговое значение. На выход нейрона поступает состояние нейрона, преобразованное нелинейной функцией активации:

$$y = f(s(\mathbf{x}, \mathbf{w})). \quad (2)$$

В качестве функции активации, как правило, выбирается сигмоидальная, ступенчатая или кусочно-линейная функция, ограничивающая значение на выходе нейрона.

Разработанная в диссертации модель потокового нейрона при определенных условиях можно считать приближенно эквивалентной модели нейрона Мак-Каллока–Питтса.

1.2. Некоторые подходы к аппаратной реализации искусственных нейронных сетей

Аппаратная реализация, в отличие от моделирования нейронной сети на обычном компьютере, полностью использует заложенный в нейронных сетях параллелизм. Поэтому аппаратная реализация нейронной сети может работать существенно быстрее, чем ее модель на обычном последовательном компьютере. Имеется несколько подходов к аппаратной реализации ИНС, различающихся способом кодирования информации. Далее приводится краткий обзор различных реализаций нейросетей, использующих аналоговое, цифровое и импульсное кодирование.

Аналоговая реализация нейрона выглядит привлекательно, так как операции над величинами тока или напряжения реализуются достаточно просто, но аналоговые схемы имеют ряд недостатков. Как отмечает М. Томлинсон в статье [62], аналоговые схемы неустойчивы к помехам, к изменению температуры. Вариации при изготовлении аналоговых элементов делают трудным создание идентичных устройств на их основе. Существуют трудности при соединении нескольких аналоговых устройств. Нет надежной и простой долговременной аналоговой памяти. Тем не менее, аналоговые реализации искусственных нейронных сетей широко известны [41, 48].

Цифровая реализация нейрона свободна от указанных недостатков, но аппаратные затраты в ряде случаев оказываются существенными. Модель нейрона (1) удобна для моделирования нейронных сетей на обычном компьютере, но сложна для реализации в виде специализированного устройства. В полностью параллельной цифровой

аппаратной реализации данного нейрона все арифметические операции выполняются на отдельных устройствах умножения и сложения, что становится существенным препятствием для увеличения количества синапсов одного нейрона и количества нейронов в сети. Передача многобитовых значений также увеличивает количество связей между нейронами.

Модель нейрона с более простой цифровой аппаратной реализацией была предложена в статье У. Мак-Каллока и В. Питтса [52]. В нейроне Мак-Каллока–Питтса используются бинарные значения на входах и выходе нейрона, а также ступенчатая функция активации. Бинарность входных значений устраняла необходимость в сложных схемах умножения, ступенчатая функция активации также проста в реализации. Но сохранялась необходимость суммировать N чисел, полученных от умножения весов на бинарные входы. Кроме того, бинарность входов несколько сужала область применения таких нейронов.

Импульсное кодирование информации, предложенное в работах [55, 56], сочетает простоту аппаратной реализации и отсутствие многих недостатков аналогового кодирования. Приведем краткий обзор нейросетей, использующих импульсное кодирование.

А. Мюррей, А. Гамильтон и др. рассматривали аналоговое представление импульсов и выделяли четыре способа кодирования информации в виде потока аналоговых импульсов: модуляция амплитуды импульсов, модуляция длительности импульсов, модуляция частоты импульсов и модуляция фазового сдвига между двумя потоками. В статьях [45, 54–56] рассматривались аппаратные схемы для сложения и умножения таких аналоговых потоков импульсов, а также реализация аналогового нейрона, основанного на таком представлении информации.

М. Томлинсон и Д. Уокер разработали цифровое представление потока импульсов в виде последовательности из 0 и 1, а также нейрон, основанный на этом представлении [62]. В отличие от аналогового представления, в цифровом потоке импульсов имеется синхронизация, и импульсы всегда имеют одинаковую амплитуду. Синхронизация, в частности, существенно упрощает операцию умножения между двумя потоками по сравнению с аналоговыми схемами, оно реализуется одним логическим элементом «И» (для независимых потоков).

Дж. Томберг и К. Каски также рассматривали аппаратную реализацию нейрона, использующего модуляцию частоты импульсов. В статье [61] рассматривается аналоговая и цифровая реализация таких нейронов. Цифровые потоки, использованные

в этой статье, состоят из значений -1 и 1 , что позволяет представить в виде потока значения из отрезка $[-1; 1]$.

В статьях [27–29] Маматов Ю. А., Булычев С. Ф., Карлин А. К. и др. предложили модель цифрового вероятностного нейрона, работающего со значениями, представленными в виде стохастических потоков. Данный нейрон принимает на входах потоки, среднее значение которых лежит на отрезке $[-1; 1]$. Значения весовых коэффициентов хранятся в регистрах и преобразуются в поток с помощью нескольких независимых генераторов случайных чисел $0,1$. Приближенное суммирование производится на двух однобитовых шинах. Состояние нейрона накапливается в суммирующем регистре и его среднее значение зависит от средних значений входных значений и весовых коэффициентов нелинейно. Общее количество генераторов случайных чисел имеет порядок $O(N)$.

В статье [7] была рассмотрена модификация данного нейрона, в которой количество генераторов случайных чисел $0,1$ сокращено до $O(\log N)$, и среднее значение состояния нейрона линейно зависит от средних значений входных значений и весовых коэффициентов.

В модели нейрона, предложенной в диссертации, была устранена необходимость преобразования весовых коэффициентов в поток импульсов, что позволило сократить количество генераторов случайных чисел. В предложенной модели оно не зависит от количества входов нейрона.

Таким образом, в диссертации предложена модель усовершенствованного и упрощенного нейрона, работающего с потоками импульсов, количество генераторов случайных чисел в котором не зависит от количества синапсов.

1.3. Импульсное кодирование информации в биологических нейронных сетях

Исследованиями в нейробиологии было установлено, что нервные импульсы, передаваемые между нейронами, приблизительно одинаковы по длительности и амплитуде [1, 33], при передаче последовательности импульсов по длинным нервным волокнам интервалы между импульсами выравниваются, и в результате остается только информация о средней плотности импульсов [35].

Эти биологические данные показывают, что информация в биологических нейронных сетях может передаваться в форме плотности приблизительно одинаковых

нервных импульсов.

1.4. Клеточные нейронные сети

Потоковые нейроны решают проблему сложности каждого нейрона в отдельности, а также уменьшают количество электрических соединений между парами нейронов за счет уменьшения количества битов, передаваемых одновременно. Однако сложность нейронных сетей заключается также и в количестве связей между нейронами. Так, например, в полносвязных сетях Хопфилда [48] из N нейронов имеется $C_N^2 = N(N - 1)$ связей между нейронами.

Одним из способов решения проблемы большого количества связей являются клеточные нейронные сети (КНС) с локальными связями между нейронами. КНС обладают многими свойствами полносвязных сетей и в то же время намного проще для аппаратной реализации. Прототипом КНС являются клеточные автоматы, впервые введенные в работах Джона фон Неймана. В классической работе [41] была предложена аналоговая клеточная нейронная сеть, представлявшая собой двумерный массив однотипных элементов, соединенных с ближайшими соседями, и работавшая в непрерывном времени. В работе [42] доказывается, что КНС обладают сложным поведением и эквивалентны машине Тьюринга. В ряде работ рассматриваются КНС, работающие в дискретном времени и реализующие ассоциативную память [37,38,51,53]. КНС, реализованные на потоковых нейронах, занимают промежуточное положение между дискретными сетями и сетями с непрерывным временем.

1.5. Нейроны с альтернативными синапсами

При сохранении локальности связей КНС можно улучшить способности сети за счет использования более сложных моделей нейронов. Одной из таких моделей является нейрон с альтернативными синапсами, формальная модель которого была предложена в статье А. А. Короткина и А. В. Панкратова [8]. Особенностью нейрона с альтернативными синапсами (А-нейрона) является то, что входные значения суммируются с разными весовыми коэффициентами в зависимости от их знаков. Состояние А-нейрона определяется в соответствии со следующей формулой:

$$s(\mathbf{x}, \mathbf{w}^+, \mathbf{w}^-) = \sum_{\substack{i \\ x_i > 0}} x_i w_i^+ + \sum_{\substack{i \\ x_i < 0}} x_i w_i^-. \quad (3)$$

В той же статье [8] доказана теорема о том, что А-нейрон с N синапсами способен классифицировать $2N + 1$ случайных образов с вероятностью, стремящейся к 1 при $N \rightarrow \infty$.

Данная формальная модель А-нейрона была предложена на основе открытия пар согласованно функционирующих нейронов головного мозга [43, 50, 64], важную роль этих нейронных структур отмечает Ю. Д. Кропотов [12–15].

В диссертации разработана и изучена модель нейрона с альтернативными синапсами, работающего со средними значениями стохастических потоков. Количество генераторов случайных битов в этой модели не зависит от количества синапсов. При определенных условиях разработанную модель А-нейрона можно считать приближенно эквивалентной модели А-нейрона (3).

1.6. Дискретное преобразование Фурье

Представление информации в виде среднего значения потока может быть также применено для аппаратной реализации дискретного преобразования Фурье с низкими аппаратными затратами.

Дискретное преобразование Фурье (далее ДПФ) широко применяется в цифровой технике для спектрального представления информации. Прямое ДПФ выполняется над последовательностью дискретных величин конечной длительности. В результате получается также конечной длительности последовательность дискретных величин, дающих частотно-спектральное представление указанной входной последовательности. При обратном ДПФ (ОДПФ) по второй последовательности находится первая из них. Каждая из этих последовательностей содержит одинаковое количество N дискретных величин.

ДПФ применяется для ряда задач обработки данных, куда входят обработка цифровых сигналов и изображений [31], адаптивное предсказание речи [32], техническое зрение [16], цифровая голография, сейсморазведка и многие другие. Несколько сотен ссылок на работы, посвященные ДПФ, приведены в книге [6].

Обозначим исходный набор величин x_n , а результат дискретного преобразования Фурье — y_k ($n, k \in \{0, \dots, N - 1\}$). Тогда ДПФ и ОДПФ могут быть записаны следующим образом:

$$y_k = \sum_{n=0}^{N-1} x_n e^{-i \frac{2\pi}{N} nk},$$

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} y_k e^{i \frac{2\pi}{N} nk}.$$

Аппаратная реализация ДПФ рассматривалась многими авторами. Предлагались как цифровые, так и аналоговые устройства (например, акустооптические и оптоэлектронные), выполняющие ДПФ. В полностью параллельной цифровой реализации, использующей граф соединений быстрого ДПФ, имеется $N \log_2 N$ комплексных умножений и столько же комплексных сложений [31].

В диссертации предлагается схемотехническая модель устройства, выполняющего дискретное преобразование Фурье и использующего для представления чисел стохастические потоки. При аппаратной реализации этого устройства количество элементов имеет порядок $N \log_2 N$, причем используются только $2N - 3$ сумматора разрядности $\log_2 N$, $N \log_2 N$ логических элементов «И» и $\log_2 N$ генераторов случайных чисел 0, 1.

1.7. Обзор диссертации

В главе 2 приводится обзор способов представления информации в виде среднего значения цифровых стохастических последовательностей (потоков). Рассматриваются некоторые операции с потоками и два метода оценки среднего значения потока.

В разделе 2.1 приводятся основные определения стохастических потоков и их свойств.

Определение. Стохастическим потоком называется случайный процесс с дискретным временем и конечным множеством возможных значений.

Определение. Будем говорить, что поток (x_1, \dots, x_t, \dots) имеет среднее значение x , или число x представлено в виде потока (x_1, \dots, x_t, \dots) , если следующий предел сходится по вероятности и равен x :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n x_t = x.$$

Определение. Два потока (x_1, \dots, x_t, \dots) и (y_1, \dots, y_t, \dots) со средними значениями x и y будем называть независимыми, если следующий предел сходится по вероятности к xy :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n x_t y_t = xy.$$

Данное определение отличается от стандартного определения независимых случайных процессов, но в данной работе рассматривается только независимость в смысле умножения соответствующих элементов потоков.

В разделе **2.2** рассматривается представление в виде потока действительных значений из отрезка $[0; 1]$, методы оценки среднего значения потока, основные операции с потоками.

В разделе **2.3** рассматриваются два способа представления в виде потока действительных значений из отрезка $[-1; 1]$ и операции с такими потоками.

В разделе **2.4** предлагается представление комплексных значений в виде потока и операции с комплексными потоками.

В главе **3** предлагается модель цифрового нейрона, работающего со средними значениями потоков. Схема этого цифрового нейрона состоит из небольшого количества простых логических элементов, нескольких сумматоров и нескольких генераторов случайных чисел. Это обеспечивает низкую стоимость аппаратной реализации и возможность наращивания количества синапсов.

В разработанной модели нейрона информация представляется в виде среднего значения стохастических последовательностей (потоков). Следуя [28, 62], такие нейроны будем называть потоковыми.

Потоковые нейроны разрабатываются с целью уменьшения аппаратных затрат и стоимости аппаратной реализации нейронных сетей. Их применение позволяет создавать сети больших размеров и увеличивать число синапсов отдельных нейронов. Положительным свойством потоковых нейронов также является относительная устойчивость к ошибкам передачи данных.

В статье [28] рассматривается схемотехническая модель потокового нейрона, в котором число генераторов случайных битов пропорционально числу синапсов нейрона. В работе [7] число генераторов случайных битов было сокращено до $\log_2 N$, где N — количество синапсов.

Изложенная в настоящей работе модель потокового нейрона имеет число генераторов случайных битов, не зависящее от количества синапсов.

В разделе **3.1** перечисляются основные элементы потокового нейрона.

Рассматриваемый нейрон имеет N входов, на которые поступают последовательности $(x_{i1}, \dots, x_{it}, \dots)$, где $x_{it} \in \{-1, 0, 1\}$, $(N + 1)$ синаптических коэффициентов w_0, \dots, w_N и регистр z , в котором накапливается некоторое суммарное значение. На

выход нейрона подается последовательность y_t , формируемая в соответствии со знаком значения регистра z и коэффициентом разреженности b .

Синаптические коэффициенты w_i кодируются w' битами и битом знака, регистр z кодируется $(z' + w' + 2)$ битами и представляется в форме дополнения до 1. Коэффициент b кодируется b' битами.

В состав нейрона всего входит:

- 1) $3Nw' + 2N + 2 + b'$ битовых логических элементов «И»;
- 2) две шины квазисуммирования разрядности w' ;
- 3) четыре сумматора разрядностей $w' + 1$, $w' + 2$, $w' + 3$ и $z' + w' + 2$;
- 4) b' генераторов случайных битов.

На рис. 3 (стр. 31) изображена схема описываемого нейрона. Блоки $\mathbf{W}_1 - \mathbf{W}_4$ содержат регистры с соответствующими синаптическими коэффициентами и устройства выбора. Блок \mathbf{Z} содержит регистры z и w_0 , а также ряд сумматоров. Блок \mathbf{B} содержит регистр b и осуществляет вероятностную фильтрацию.

В разделе **3.2** описывается работа нейрона на одном шаге.

На шаге t очередные значения x_{it} из входных последовательностей умножаются на соответствующие весовые коэффициенты w_i , и модули результатов поступают на одну из двух шин в зависимости от знака. На шинах происходит побитовое логическое «ИЛИ», и результаты накапливаются в суммирующем регистре z .

Обозначим поступающие на положительную шину значения $w_i^+(t)$, а поступающие на отрицательную шину — $w_i^-(t)$:

$$w_i^+(t) = \begin{cases} |w_i|, & \text{если } w_i x_{it} > 0; \\ 0, & \text{если } w_i x_{it} \leq 0, \end{cases}$$

$$w_i^-(t) = \begin{cases} |w_i|, & \text{если } w_i x_{it} < 0; \\ 0, & \text{если } w_i x_{it} \geq 0. \end{cases}$$

На шаге t значение регистра z изменяется следующим образом:

$$z_{t+1} = z_t - z_t 2^{-z'} + \Delta z_t,$$

$$\Delta z_t = \bigvee_{i=1}^N w_i^+(t) - \bigvee_{i=1}^N w_i^-(t) + w_0.$$

Символом \bigvee обозначена операция побитового логического «ИЛИ».

На каждом шаге на выход нейрона подается с вероятностью b значение 0 и с вероятностью $1 - b$ знак регистра z .

В разделе **3.3** вычисляется математическое ожидание и дисперсия величины z_t при $t \rightarrow \infty$, при условии независимости величин x_{it} в совокупности для всех i, t . Вычисляется также математическое ожидание z_t для одного частного случая, необходимого для обучения.

В разделе **3.4** получены неравенства для математического ожидания величины $\sqrt{\prod_{i=1}^N w_i^+(t)}$ при условии независимости величин x_{it} . Показано, что функция состояния нейрона приближается к линейной от значений весов и средних значений входных последовательностей при $b \rightarrow 1$.

В разделе **3.5** формулируется задача обучения полносвязной нейронной сети, и предлагаются методы обучения: метод Хебба, модифицированный метод Хебба и метод оптимизации приближенной функции ошибки. Каждый нейрон обучается независимо, номер обучаемого нейрона обозначен j .

Приведем здесь модифицированный метод Хебба и метод оптимизации. Обучение производится на множество изображений $\{\mathbf{X}^1, \dots, \mathbf{X}^P\}$, $\mathbf{X}^m = (X_1^m, \dots, X_N^m)$, где $X_i^m \in \{-1, 1\}$.

Модифицированный метод Хебба. Предлагается следующий вариант правила Хебба. Для нейрона j (предполагая $0/0 = 1/2$):

$$\begin{aligned} A_{ij}^+ &= \frac{1}{4} \sum_{m=1}^P (1 + X_i^m)(1 + X_j^m), \\ A_{ij}^- &= \frac{1}{4} \sum_{m=1}^P (1 - X_i^m)(1 + X_j^m), \\ B_{ij}^+ &= \frac{1}{4} \sum_{m=1}^P (1 + X_i^m)(1 - X_j^m), \\ B_{ij}^- &= \frac{1}{4} \sum_{m=1}^P (1 - X_i^m)(1 - X_j^m), \\ w_{ij} &= \left(2^{w'} - 1\right) \left(\frac{A_{ij}^+}{A_{ij}^+ + B_{ij}^+} - \frac{A_{ij}^-}{A_{ij}^- + B_{ij}^-} \right), \\ i &\in \{1, \dots, N\}, \end{aligned}$$

где A_{ij}^+ , A_{ij}^- , B_{ij}^+ , B_{ij}^- — количество элементов в подмножествах множества образов.

Оптимизационное обучение. Оптимизационная задача заключается в нахождении коэффициентов w_{1j}, \dots, w_{Nj} , при которых функция E принимает наименьшее значение:

$$E(w_{1j}, \dots, w_{Nj}) = \sum_{m=1}^P (S(\bar{Z}(\mathbf{X}^m, \mathbf{w})) - X_j^m)^2,$$

$$\bar{Z}(\mathbf{x}, \mathbf{w}) = \frac{1 - b^N}{N} \sum_{i=1}^N w_i x_i.$$

$$S(z) = \frac{2}{1 + e^{-z}} - 1.$$

где j — номер обучаемого нейрона. Параметры w_{0j} и w_{jj} принимаются равными нулю. Здесь величины w_{ij} являются вещественными. После оптимизации они нормируются и округляются до ближайшего целого.

В разделе **3.6** описывается эксперимент на имитационной модели полносвязной сети размера 8×8 из потоковых нейронов, метод статистической обработки результатов, и приводятся средние статистические значения для всех образов и всех уровней помех от 1 до 10 при разных способах обучения. Произведено сравнение методов обучения между собой и с традиционным методом Хебба, показавшее преимущество метода оптимизации над модифицированным методом Хебба и преимущество модифицированного метода Хебба над традиционным методом Хебба.

В главе 4 рассматривается модель цифрового нейрона с альтернативными синапсами, работающего со средними значениями потоков.

Потоковый нейрон с альтернативными синапсами обрабатывает положительные и отрицательные значения, поступающие на его синапсы, независимо, с различными в общем случае весовыми коэффициентами. На выход данного нейрона поступает две последовательности y_t^+ и y_t^- , для положительных и отрицательных значений. Формальная модель нейрона с альтернативными синапсами (или А-нейрона) была впервые предложена в статье [8]. В той же статье показано, что А-нейроны в среднем обладают лучшими классифицирующими способностями по сравнению с обычными нейронами. Формальная модель А-нейрона была разработана на основе биологических данных о существовании пар согласованно функционирующих нейронов [43, 50, 64], важную роль этих нейронных структур отмечает Ю. Д. Кропотов в статьях [12–15].

Схема потокового А-нейрона, как и схема потокового нейрона, описанного в третьей главе, имеет количество генераторов случайных битов, не зависящее от количества синапсов. Количество связей между нейронами не возрастает, так как в схеме потокового нейрона для кодирования значений $-1, 0, 1$ требуется два бита, столько же, сколько передается между потоковыми А-нейронами.

В разделе **4.1** перечисляются основные элементы потокового нейрона с альтернативными синапсами.

Рассматриваемый нейрон имеет $2N$ входов, на которые поступают последовательности $(x_{i1}^+, \dots, x_{it}^+, \dots)$ и $(x_{i1}^-, \dots, x_{it}^-, \dots)$, где $x_{it}^+, x_{it}^- \in \{0, 1\}$, $i \in \{1 \dots N\}$, t — дискретное время. Выполняется условие $x_{it}^+ x_{it}^- = 0$. Имеется $2N$ синаптических коэффициентов $w_1^+, w_1^-, \dots, w_N^+, w_N^-$, два пороговых коэффициента w_0^+, w_0^- и регистр z , в котором накапливается некоторое суммарное значение. На выход нейрона подаются две последовательности $(y_1^+, \dots, y_t^+, \dots)$ и $(y_1^-, \dots, y_t^-, \dots)$, где $y_t^+, y_t^- \in \{0, 1\}$, формируемые в соответствии со знаком значения регистра z и коэффициентом разреженности b .

Синаптические и пороговые коэффициенты w_i^+ и w_i^- кодируются w' битами и битом знака, регистр z кодируется $(z' + w' + 2)$ битами и представляется в форме дополнения до 1. Коэффициент b кодируется b' битами.

В состав нейрона всего входит:

- 1) $(2N + 2)$ регистров разрядности $w' + 1$, один регистр разрядности $z' + w' + 2$, один регистр разрядности b' ;
- 2) $4Nw' + 2N + 2 + b'$ битовых логических элементов «И»;
- 3) две шины квазисуммирования разрядности w' ;
- 4) четыре сумматора разрядностей $w' + 1$, $w' + 2$, $w' + 3$ и $z' + w' + 2$;
- 5) b' генераторов случайных битов.

На рис. 7 (стр. 45) изображена схема описываемого нейрона. Блоки $\mathbf{W}_1 - \mathbf{W}_N$ содержат регистры с синаптическими коэффициентами $(w_1^+, w_1^-, \dots, w_N^+, w_N^-)$ и устройства выбора. Блок \mathbf{Z} содержит регистры z , w_0^+ и w_0^- , а также ряд сумматоров. Блок \mathbf{B} содержит регистр b и осуществляет вероятностную фильтрацию.

В разделе 4.2 описывается работа нейрона с альтернативными синапсами на одном шаге.

Введем необходимые обозначения. Определим следующие функции знака:

$$\begin{aligned} \operatorname{sgn}^+(z) &= \begin{cases} 0, & \text{если } z < 0; \\ 1, & \text{если } z \geq 0, \end{cases} \\ \operatorname{sgn}^-(z) &= \begin{cases} 1, & \text{если } z < 0; \\ 0, & \text{если } z \geq 0, \end{cases} \\ \operatorname{sgn}(z) &= \operatorname{sgn}^+(z) - \operatorname{sgn}^-(z). \end{aligned} \tag{4}$$

Будем использовать операцию $[x]$ как округление к ближайшему целому.

Опишем работу нейрона в момент времени t .

На входы нейрона поступают значения x_{it}^+ и x_{it}^- , $i \in \{1, \dots, N\}$.

В зависимости от входных значений x_{it}^+ , x_{it}^- и знаков синаптических коэффициентов w_i^+ и w_i^- , на одну из шин квазисуммирования подается значение $|w_i^+|$ или $|w_i^-|$. Обозначим поступающие из блока \mathbf{W}_i в момент времени t на положительную шину значения $w_i^+(t)$, а на отрицательную шину — $w_i^-(t)$:

$$\begin{aligned} w_i^+(t) &= x_{it}^+ |w_i^+| \operatorname{sgn}^+(w_i^+) + x_{it}^- |w_i^-| \operatorname{sgn}^+(w_i^-), \\ w_i^-(t) &= x_{it}^+ |w_i^+| \operatorname{sgn}^-(w_i^+) + x_{it}^- |w_i^-| \operatorname{sgn}^-(w_i^-), \\ k &\in \{1, \dots, w'\}. \end{aligned}$$

Следует заметить, что в вышеприведенных выражениях сумма может быть заменена на побитовое «ИЛИ», так как хотя бы одно из двух слагаемых равно нулю, а умножение — на операцию «И», так как $x_{it}^\pm \in \{0, 1\}$, $\operatorname{sgn}^\pm(w_i^\pm) \in \{0, 1\}$

На шинах квазисуммирования происходит побитовое логическое «ИЛИ», и на них образуются значения $w^+(t)$ и $w^-(t)$:

$$\begin{aligned} w^+(t) &= \bigvee_{i=1}^N w_i^+(t), \\ w^-(t) &= \bigvee_{i=1}^N w_i^-(t). \end{aligned}$$

Значения $w^+(t)$ и $w^-(t)$ поступают на вход блока \mathbf{Z} . На шаге t значение регистра z изменяется следующим образом:

$$\begin{aligned} z_{t+1} &= z_t - \left[z_t 2^{-z_t} \right] + \Delta z_t, \\ \Delta z_t &= w^+(t) - w^-(t) + w_0^+ \operatorname{sgn}^+(z_t) - w_0^- \operatorname{sgn}^-(z_t). \end{aligned} \quad (5)$$

На выход блока \mathbf{Z} подаются значения $\operatorname{sgn}^+(z_t)$ и $\operatorname{sgn}^-(z_t)$, определяемые в соответствии с (4). Эти значения поступают в блок \mathbf{B} .

Будем называть $S_t = \operatorname{sgn}(z_t)$ состоянием нейрона в момент времени t .

На выход блока \mathbf{B} с вероятностью b подаются значения $0, 0$, и с вероятностью $(1 - b)$ значения $\operatorname{sgn}^+(z_t)$ и $\operatorname{sgn}^-(z_t)$:

$$\begin{aligned} \mathbf{P} \{ y_t^+ = 0 \wedge y_t^- = 0 \} &= b, \\ \mathbf{P} \{ y_t^+ = \operatorname{sgn}^+(z_t) \wedge y_t^- = \operatorname{sgn}^-(z_t) \} &= 1 - b, \end{aligned}$$

При выполнении условия $w_i^- = -w_i^+$ данный нейрон эквивалентен нейрону, описанному во второй главе.

В разделе 4.3 вычисляется математическое ожидание величины $w^+(t) - w^-(t)$, при условии независимости величин x_{it}^\pm в совокупности для всех i . Вычисляется также математическое ожидание этой величины для одного частного случая, необходимого для обучения.

В разделе 4.4 формулируется задача обучения полносвязной нейронной сети, построенной на нейронах с альтернативными синапсами. Описывается двухшаговый метод обучения, на первом шаге вычисляются коэффициенты w_1^\pm, \dots, w_N^\pm , на втором шаге вычисляются w_0^+ и w_0^- . Предлагается правило вычисления коэффициентов w_0^\pm для полносвязной сети, при котором все эталонные изображения гарантированно являются устойчивыми. Каждый нейрон обучается независимо, номер обучаемого нейрона обозначен j . Обучение производится на множество изображений $\{\mathbf{X}^1, \dots, \mathbf{X}^P\}$, $\mathbf{X}^m = (X_1^m, \dots, X_N^m)$, где $X_i^m \in \{-1, 1\}$.

В разделе 4.5 формулируются правила обучения нейрона с альтернативными синапсами по методу Хебба и по модифицированному методу Хебба, а также аналогичные методы для обычных нейронов.

Приведем здесь методы обучения для А-нейронов.

Обучение с помощью стандартного метода Хебба, с использованием альтернативных синапсов (описано в статье [8]):

$$w_{ij}^+ = \left[\frac{2^{w'} - 1}{P} \sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^+(X_i^m) \right],$$

$$w_{ij}^- = \left[\frac{2^{w'} - 1}{P} \sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^-(X_i^m) \right].$$

Обучение с помощью модифицированного метода Хебба, с использованием альтернативных синапсов (здесь $0/0=0$):

$$w_{ij}^+ = \left[\frac{\sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^+(X_i^m)}{(2^{w'-1}) \sum_{m=1}^P \operatorname{sgn}^+(X_i^m)} \right],$$

$$w_{ij}^- = \left[\frac{\sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^-(X_i^m)}{(2^{w'-1}) \sum_{m=1}^P \operatorname{sgn}^-(X_i^m)} \right].$$

В разделе 4.6 формулируются правило обучения А-нейрона методом оптимизации приближенной функции ошибки, и аналогичное правило для обычных нейронов.

Приведем здесь метод обучения для А-нейронов.

Для обучения А-нейрона будем оптимизировать функцию E_a :

$$\begin{aligned}
 E_a(\boldsymbol{\omega}^+, \boldsymbol{\omega}^-) &= \sum_{m=1}^P \left(s(f_a(\mathbf{X}^{m(+)}, \mathbf{X}^{m(-)}, \boldsymbol{\omega}^+, \boldsymbol{\omega}^-)) - X_j^m \right)^2, \\
 f_a(\boldsymbol{\chi}^+, \boldsymbol{\chi}^-, \boldsymbol{\omega}^+, \boldsymbol{\omega}^-) &= \sum_{i=1}^N \chi_i^+ \omega_i^+ + \sum_{i=0}^N \chi_i^- \omega_i^-, \\
 \mathbf{X}^{m(+)} &= (X_1^m \operatorname{sgn}^+(X_1^m), \dots, X_N^m \operatorname{sgn}^+(X_N^m)), \\
 \mathbf{X}^{m(-)} &= (X_1^m \operatorname{sgn}^-(X_1^m), \dots, X_N^m \operatorname{sgn}^-(X_N^m)), \\
 s(x) &= \frac{1 - e^{-x}}{1 + e^{-x}}.
 \end{aligned}$$

После нахождения локального минимума $\boldsymbol{\omega}^{+*}$, $\boldsymbol{\omega}^{-*}$ весовые коэффициенты вычисляются следующим образом:

$$\begin{aligned}
 w_i^+ &= \left[\frac{(2^{w'} - 1) \omega_i^{+*}}{\max_k |\omega_k^{+*}|} \right], \\
 w_i^- &= \left[\frac{(2^{w'} - 1) \omega_i^{-*}}{\max_k |\omega_k^{-*}|} \right].
 \end{aligned}$$

В разделе 4.7 описывается эксперимент на имитационной модели полносвязной сети размера 8×8 из потоковых нейронов с альтернативными синапсами, метод статистической обработки результатов, приводятся средние статистические значения для всех образов и всех уровней помех от 1 до 10 при разных способах обучения. Произведено сравнение методов обучения между собой, показавшее преимущество метода оптимизации над модифицированным методом Хебба и преимущество модифицированного метода Хебба над традиционным методом Хебба. Сравнение результатов моделирования сети из потоковых А-нейронов и из обычных потоковых нейронов показало, что эффективность потоковых А-нейронов зависит от способа обучения, в случае стандартного метода Хебба эффективность даже ухудшается. В случае модифицированного метода Хебба и при оптимизационном обучении потоковые А-нейроны показывают лучшие результаты по сравнению с обычными потоковыми нейронами.

В главе 5 рассматривается модель устройства, выполняющего дискретное преобразование Фурье и использующего потоковое представление информации.

Схема этого цифрового устройства состоит из небольшого количества простых логических элементов, сумматоров и нескольких генераторов случайных чисел. Это обеспечивает низкую стоимость аппаратной реализации.

Предлагается представление комплексных чисел в виде стохастических потоков, дается описание схемотехнической модели устройства ДПФ, приводится обоснование, результаты моделирования и сравнение с традиционной цифровой схемой.

В разделе **5.1** рассматривается представление комплексных чисел в виде стохастических потоков и кодирование таких потоков.

В разделе **5.2** описывается потоковое устройство дискретного преобразования Фурье и оценивается количество элементов аппаратной реализации.

Данное устройство, имеет $N = 2^n$ входов, на которые поступают последовательности, состоящие из значений $\exp \frac{i2\pi k}{N}$, закодированных значениями $k \in \{0, \dots, N - 1\}$. Последовательности генерируются случайным образом так, чтобы среднее значение было равно значению функции в соответствующей точке. Имеется также N выходов, на которые подаются аналогичные последовательности, среднее значение которых равно дискретному спектру Фурье в соответствующей точке. Входные значения обозначены $X_{i,t}$, выходные значения — $Y_{k,t}$. Их соответствующие коды обозначены $x_{i,t}$ и $y_{k,t}$.

На рис. 8 (стр. 59) буквой G обозначен равномерный генератор последовательности случайных чисел g_t из множества $H = \{0, \dots, N - 1\}$.

Устройство выбора С принимает на входе N кодов $x_{i,t}$, номер $g_t \in H$ и выдает $x_{g_t,t}$.

Устройство умножения М принимает на входе код $x_{g_t,t}$ и номер $g_t \in H$ и выдает N кодов $y_{k,t}$:

$$y_{k,t} = (x_{g_t,t} + g_t k) \bmod N \quad (6)$$

$$k \in H.$$

Устройство умножения содержит N сложений и умножения на константы $k \in H$. Умножения на константы можно заменить сложениями, причем количество сложений равно $N - 1$.

Один такт работы схемы, обозначенный t , можно описать следующим образом:

- 1) получается число g_t от генератора G, которое подается на устройства С и М;
- 2) производится выбор кода $x_{g_t,t}$ из входных последовательностей;

- 3) код $x_{g_t,t}$ передается на умножитель M , где параллельно получают коды $y_{k,t}$ в соответствии с (6);
- 4) на входы подаются следующие коды значений входных последовательностей, и процесс повторяется.

В разделе **5.3** приводится доказательство следующей теоремы.

Теорема. Пусть последовательность g_t и последовательности $X_{n,t}$ независимы для всех $n \in H$. Тогда последовательности $Y_{k,t}$, полученные в данной схеме, будут иметь средние значения, равные $1/N$ спектра Фурье.

В разделе **5.4** рассматривается скорость сходимости среднего значения выходных последовательностей.

В разделе **5.5** приводится описание эксперимента на имитационной модели устройства ДПФ. Моделирование подтвердило работоспособность данного устройства.

В разделе **5.6** производится сравнение потокового устройства ДПФ с цифровым полностью параллельным устройством, использующим граф соединений быстрого ДПФ.

2. Потокое представление информации

Потоком будем называть последовательность случайных величин с конечным совокупным множеством возможных значений. В простейшем случае эти случайные величины независимы и одинаково распределены. Поток передает информацию в виде своего среднего значения. Бинарные цифровые потоки (состоящие из 0 и 1) были впервые применены для построения потокового нейрона в статье М. Томлинсона [62]. Свойства бинарных потоков, операции с ними и схемотехнические модели реализации были подробно рассмотрены в статьях [27–29]. Представление отрезка $[-1; 1]$ в виде потока было рассмотрено в статьях [28, 61]. Представление комплексных чисел рассматривалось в статьях [19, 20].

Далее приводится обзор, в котором рассматривается представление чисел в виде потоков, операции с потоками и преобразование потока в число. Представление действительных чисел в виде потока приводится для полноты обзора, представление комплексных чисел является новым. Оно было разработано для устройства дискретного преобразования Фурье, описанного в главе 5.

2.1. Общие определения

Термин «поток» является буквальной переводом английского «stream». Впервые английский термин был употреблен в статье А. Ф. Мюррея [55] в контексте «pulse stream». В статье Дж. Томберга [61] использовался термин «stochastic bit stream». В статье М. Томлинсона [62] использовался немного другой термин «stochastic pulse train». Русский термин впервые встречается в статье [27].

Определение. Стохастическим потоком (x_1, \dots, x_t, \dots) будем называть последовательность случайных величин x_t , с конечным совокупным множеством возможных значений H :

$$\forall t \quad x_t \in H.$$

Определение. Будем говорить, что поток $\mathbf{X} = (x_1, \dots, x_t, \dots)$ имеет среднее значение x , или число x представлено в виде потока \mathbf{X} , если следующий предел сходится по вероятности и равен x :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n x_t = x. \quad (7)$$

Над потоками можно производить некоторые действия. При умножении соответствующих элементов двух потоков получается третий поток, при определенных условиях его среднее значение равно произведению средних значений двух исходных потоков. В следующем определении предлагается называть такие потоки независимыми.

Определение. Два потока (x_1, \dots, x_t, \dots) и (y_1, \dots, y_t, \dots) со средними значениями x и y будем называть независимыми, если следующий предел сходится по вероятности к xy :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n x_t y_t = xy. \quad (8)$$

В простейшем случае величины x_t взаимно независимы и одинаково распределены. Например, такой поток получается в результате серии испытаний по схеме Бернулли. У таких потоков всегда существует среднее значение, равное математическому ожиданию одного элемента потока.

2.2. Представление значений из $[0; 1]$

Вначале рассмотрим бинарные потоки как самый простой случай потоков. Для представления числа x из отрезка $[0; 1]$ в виде бинарного потока (x_1, \dots, x_t, \dots) на каждом шаге проводятся независимые испытания с вероятностью успеха x :

$$\begin{aligned} P\{x_t = 0\} &= 1 - x, \\ P\{x_t = 1\} &= x, \\ Mx_t &= x, \\ Dx_t &= x(1 - x). \end{aligned} \quad (9)$$

В потоке, полученном таким способом, величины x_t взаимно независимы и $Mx_t = x$ для всех t .

Рассмотрим свойства бинарных потоков и некоторые операции с ними более подробно. Некоторые очевидные свойства, вытекающие из независимости испытаний при генерации потока:

- 1) поток является стационарным, то есть количество единиц не зависит от времени начала измерения, а только от его длительности;
- 2) отсутствует последствие, то есть количество единиц, попадающих в непрерывающиеся отрезки времени взаимно независимо;

3) число единиц на отрезке времени длины n подчиняется биномиальному распределению:

$$\begin{aligned} P\left\{\sum_{t=1}^n x_t = m\right\} &= C_n^m x^m (1-x)^{n-m}, \\ M\left(\sum_{t=1}^n x_t\right) &= nx, \\ D\left(\sum_{t=1}^n x_t\right) &= nx(1-x). \end{aligned}$$

При достаточно большом отрезке времени количество единиц на этом отрезке можно приближенно считать нормально распределенным с теми же математическим ожиданием и дисперсией.

Для оценки неизвестной плотности потока используется традиционный способ оценки вероятности по частоте или метод экспоненциально-взвешенного среднего:

$$p^* = \frac{1}{n} \sum_{t=1}^n x_t, \quad (10)$$

$$p_\mu^* = \mu \sum_{t=1}^n x_t (1-\mu)^{n-t}, \quad (11)$$

где μ — небольшое положительное число. Эти величины также являются случайными со следующими параметрами (для потока, полученного серией независимых одинаковых испытаний):

$$\begin{aligned} Mp^* &= x, \\ Mp_\mu^* &= x(1 - (1-\mu)^n), \\ Dp^* &= Dx_t/n = x(1-x)/n, \\ Dp_\mu^* &= Dx_t(1 - (1-\mu)^{2n}) \frac{\mu}{2-\mu} = \\ &= x(1-x)(1 - (1-\mu)^{2n}) \frac{\mu}{2-\mu}. \end{aligned}$$

Для аппаратной реализации удобно использовать значения $n = 2^k$ и $\mu = 2^{-k}$.

Рассмотрим операции с бинарными потоками.

Для умножения двух величин x и y , представленных в виде независимых бинарных потоков (x_1, \dots, x_t, \dots) и (y_1, \dots, y_t, \dots) , умножаются соответствующие элементы потоков, что эквивалентно операции логическое «И»:

$$z_t = x_t y_t = x_t \wedge y_t.$$

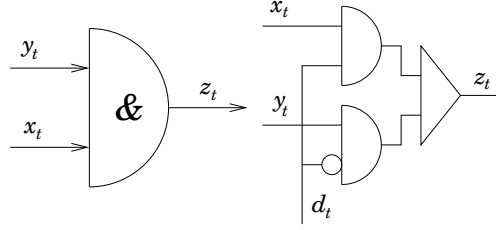


Рис. 1. Реализация умножения и полусуммы бинарных потоков

Среднее значение потока (z_1, \dots, z_t, \dots) будет равно произведению средних значений потоков (x_1, \dots, x_t, \dots) и (y_1, \dots, y_t, \dots) , по определению независимых потоков. Если величины x_t и y_t независимы, и $Mx_t = x$, $My_t = y$, то:

$$P\{z_t = 1\} = P\{x_t = 1\}P\{y_t = 1\} = xy,$$

$$P\{z_t = 0\} = 1 - P\{x_t = 1\}P\{y_t = 1\} = 1 - xy,$$

$$Mz_t = Mx_t My_t = xy,$$

$$Dz_t = M(x_t y_t - xy)^2 = xy(1 - xy).$$

Для получения полусуммы двух значений x и y , представленных в виде бинарных потоков $X = (x_1, \dots, x_t, \dots)$ и $Y = (y_1, \dots, y_t, \dots)$, применяется случайный выбор элементов с использованием третьего бинарного потока (d_1, \dots, d_t, \dots) , независимого от потоков X и Y :

$$z_t = x_t d_t + y_t (1 - d_t) = (x_t \wedge d_t) \vee (y_t \wedge \bar{d}_t),$$

$$P\{d_t = 1\} = P\{d_t = 0\} = 1/2.$$

Если величины $\{x_t, d_t\}$ независимы, и $\{y_t, d_t\}$ независимы, и выполняется $Mx_t = x$, $My_t = y$, то:

$$\begin{aligned} P\{z_t = 1\} &= P\{x_t = 1\}P\{d_t = 1\} + P\{y_t = 1\}P\{d_t = 0\} = \\ &= \frac{x + y}{2}, \end{aligned}$$

$$P\{z_t = 0\} = 1 - \frac{x + y}{2},$$

$$Mz_t = \frac{x + y}{2}.$$

На рис. 1 изображены схемы из логических элементов, реализующие умножение и полусумму бинарных потоков.

Если бинарные потоки достаточно разрежены ($P\{x_t = 1\} \ll 1$), то для приближенного суммирования нескольких потоков можно применить квазисуммирование

на общей шине. Пусть даны m независимых бинарных потоков $(x_1^{(i)}, \dots, x_t^{(i)}, \dots)$ ($i \in \{1 \dots m\}$), со средними значениями $x^{(1)} \dots x^{(m)}$ соответственно. Тогда поток (z_1, \dots, z_t, \dots) , образованный на общей шине, будет получен следующим образом:

$$z_t = \bigvee_{i=1}^m x_t^{(i)}.$$

Если величины $\{x_t^{(1)}, \dots, x_t^{(m)}\}$ независимы, то:

$$\begin{aligned} \mathbb{P}\{z_t = 0\} &= \prod_{i=1}^m \mathbb{P}\{x_t^{(i)} = 0\}, \\ \mathbb{M}z_t = \mathbb{P}\{z_t = 1\} &= 1 - \prod_{i=1}^m (1 - \mathbb{P}\{x_t^{(i)} = 1\}). \end{aligned}$$

Для несовместных бинарных потоков ($\forall i \neq j \quad x_t^{(i)} x_t^{(j)} = 0$) данный метод дает поток, среднее значение которого в точности равно сумме средних значений исходных потоков.

2.3. Представление значений из $[-1; 1]$

Рассмотрим представление в виде потока числа из интервала $[-1; 1]$. Для этого необходимо ввести функцию кодирования K . Будем говорить, что число A имеет код a , если $a = K(A)$. Функция K должна иметь обратную функцию, чтобы по коду можно было восстановить исходное значение.

Первый вариант. Для представления действительного числа x из отрезка $[-1; 1]$ в виде потока x_t достаточно двух базовых значений -1 и 1 , из которых формируется поток. Для формирования потока используется генератор случайных чисел, на каждом шаге проводится независимое испытание и выбирается значение -1 или 1 с определенной вероятностью:

$$\begin{aligned} \mathbb{P}\{x_t = -1\} &= \frac{1-x}{2}, \\ \mathbb{P}\{x_t = 1\} &= \frac{1+x}{2}, \\ \mathbb{M}x_t &= \frac{x+1}{2} - \frac{1-x}{2} = x, \\ \mathbb{D}x_t &= 1 - x^2 \leq 1. \end{aligned}$$

При аппаратной реализации передаются и обрабатываются коды значений -1 и 1 :

$$K(-1) = 0,$$

$K(x_t)$	$K(y_t)$	$K(z_t)$
0	0	1
0	1	0
1	0	0
1	1	1

Таблица 1. Операция сравнения кодов потоков

$$\begin{aligned}
K(1) &= 1, \\
K(x) &= \frac{1+x}{2}, \\
K^{-1}(y) &= 2y - 1.
\end{aligned}$$

Для умножения двух величин x и y , представленных в виде независимых потоков (x_1, \dots, x_t, \dots) и (y_1, \dots, y_t, \dots) (где $x_t, y_t \in \{-1, 1\}$), умножаются соответствующие элементы потоков, что эквивалентно операции сравнения кодов (таблица 1):

$$\begin{aligned}
z_t &= x_t y_t, \\
K(z_t) &= (K(x_t) \wedge K(y_t)) \vee (\overline{K(x_t)} \wedge \overline{K(y_t)}).
\end{aligned}$$

Если величины x_t и y_t независимы, и выполняется $Mx_t = x$, $My_t = y$, то:

$$\begin{aligned}
P\{z_t = 1\} &= P\{x_t = -1\}P\{y_t = -1\} + P\{x_t = 1\}P\{y_t = 1\} = \\
&= \frac{1+xy}{2}, \\
P\{z_t = -1\} &= \frac{1-xy}{2}, \\
Mz_t &= xy, \\
Dz_t &= 1 - (xy)^2.
\end{aligned}$$

Операция полусуммы строится аналогично полусумме бинарных потоков.

Второй вариант. Для представления числа $x \in [-1; 1]$ можно также формировать поток из значений $-1, 0, 1$. Данный метод обладает бóльшими аппаратными затратами, но в оптимальном случае дисперсия оценки среднего p^* меньше, чем в предыдущем варианте.

Для представления числа $x \in [-1; 1]$ в виде потока (x_1, \dots, x_t, \dots) ($x_t \in \{-1, 0, 1\}$) на каждом шаге будем проводить независимые испытания, получая очередное значение потока:

$$P\{x_t = 0\} = 1 - |x|,$$

$$\begin{aligned} \mathbf{P}\{x_t = 1\} &= \begin{cases} |x|, & \text{если } x > 0; \\ 0, & \text{если } x \leq 0, \end{cases} \\ \mathbf{P}\{x_t = -1\} &= \begin{cases} |x|, & \text{если } x < 0; \\ 0, & \text{если } x \geq 0. \end{cases} \end{aligned}$$

Поток, построенный таким методом, обладает следующими параметрами:

$$\begin{aligned} \mathbf{M}x_t &= x, \\ \mathbf{D}x_t &= |x|(1 - |x|) \leq 0.25. \end{aligned}$$

Введем следующие функции кодирования:

$$\begin{aligned} K(-1) &= (K_1(-1), K_2(-1)) = (1, 0), \\ K(0) &= (K_1(0), K_2(0)) = (0, 0), \\ K(1) &= (K_1(1), K_2(1)) = (0, 1), \\ K^{-1}(k_1, k_2) &= k_2 - k_1. \end{aligned}$$

Тогда операция умножения двух независимых потоков (x_1, \dots, x_t, \dots) и (y_1, \dots, y_t, \dots) , дающая в результате поток (z_1, \dots, z_t, \dots) , будет выражена следующими формулами:

$$\begin{aligned} z_t &= x_t y_t, \\ K_1(z_t) &= (K_1(x_t) \wedge K_2(y_t)) \vee (K_2(x_t) \wedge K_1(y_t)), \\ K_2(z_t) &= (K_1(x_t) \wedge K_1(y_t)) \vee (K_2(x_t) \wedge K_2(y_t)). \end{aligned}$$

Если величины x_t и y_t независимы, и выполняется $\mathbf{M}x_t = x$, $\mathbf{M}y_t = y$, то:

$$\begin{aligned} \mathbf{P}\{z_t = 1\} &= \mathbf{P}\{x_t = -1\}\mathbf{P}\{y_t = -1\} + \mathbf{P}\{x_t = 1\}\mathbf{P}\{y_t = 1\}, \\ \mathbf{P}\{z_t = -1\} &= \mathbf{P}\{x_t = -1\}\mathbf{P}\{y_t = 1\} + \mathbf{P}\{x_t = 1\}\mathbf{P}\{y_t = -1\}, \\ \mathbf{P}\{z_t = 0\} &= \mathbf{P}\{x_t = 0\} + \mathbf{P}\{y_t = 0\} - \mathbf{P}\{x_t = 0\}\mathbf{P}\{y_t = 0\}, \\ \mathbf{M}z_t &= xy, \\ \mathbf{D}z_t &= 1 - (xy)^2 - \mathbf{P}\{z_t = 0\}. \end{aligned}$$

Операция полусуммы строится аналогично полусумме бинарных потоков.

2.4. Представление комплексных значений

Рассмотрим представление в виде потока комплексного числа из правильного N -угольника, вписанного в единичную окружность на комплексной плоскости, у кото-

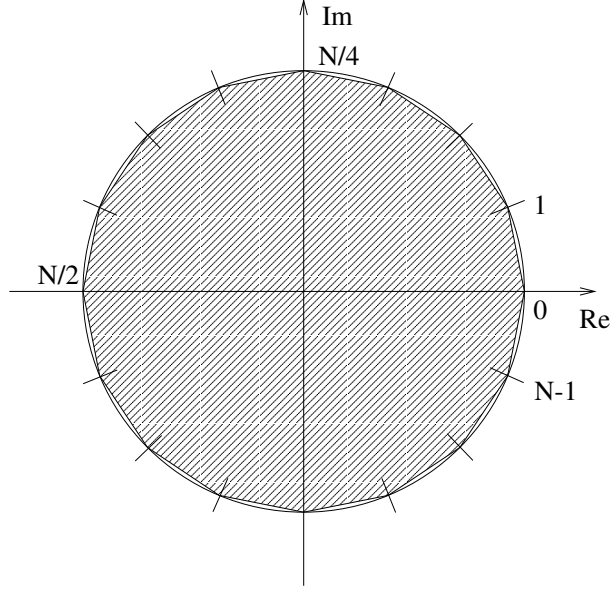


Рис. 2. Представимые комплексные значения и коды ($N = 16$)

рого действительная единица является вершиной. Для простоты будем считать, что $N = 2^n$. Пример такой области показан на рис. 2.

Значения, из которых формируется поток, принадлежат множеству \hat{C} :

$$\begin{aligned}\hat{C} &= \{W_k \mid k \in \{0, \dots, N-1\}\}, \\ W_k &= \exp \frac{i 2\pi k}{N}.\end{aligned}$$

Поток (x_1, \dots, x_t, \dots) , сформированный из таких значений с заданными вероятностями $\mathbf{P}\{x_t = W_k\} = p_k$ с помощью генератора случайных чисел, обладает следующими параметрами:

$$\begin{aligned}\mathbf{M}x_t &= \sum_{k=0}^{N-1} p_k W_k, \\ |\mathbf{D}x_t| &\leq 1.\end{aligned}$$

Для любого заданного комплексного числа x из N -угольника необходимо не более трех ненулевых p_k , чтобы выполнялось равенство $\mathbf{M}x_t = x$. К сожалению, в общем случае значения p_k вычисляются нетривиально.

В случае, когда $x = W_k(2p - 1)$ ($0 \leq p \leq 1$), возможно легко получить комплексный поток с $\mathbf{M}x_t = x$, положив $p_k = p$ и $p_l = 1 - p$ (где $l = (k + N/2) \bmod N$), и проводя серию независимых испытаний:

$$\mathbf{P}\{x_t = W_k\} = p,$$

$$\begin{aligned}
\mathbb{P}\{x_t = W_l\} &= 1 - p, \\
\mathbb{P}\{x_t \notin \{W_k, W_l\}\} &= 0, \\
Mx_t &= pW_k + (1 - p)W_l = x, \\
Dx_t &= 4W_{2k}p(1 - p) = 1 - x^2.
\end{aligned}$$

В частности, таким способом можно представить действительные числа из $[-1; 1]$ в виде комплексного потока.

Введем функцию кодирования K :

$$\begin{aligned}
K(W_k) &= k, \\
K^{-1}(k) &= W_k.
\end{aligned}$$

Для представления и передачи кода k требуется n битов, так как $N = 2^n$ и $0 \leq k \leq N - 1$.

Рассмотрим операцию умножения комплексных потоков. Для умножения независимых комплексных потоков (x_1, \dots, x_t, \dots) и (y_1, \dots, y_t, \dots) со средними значениями x и y соответственно будем умножать соответствующие элементы потоков:

$$\begin{aligned}
z_t &= x_t y_t, \\
K(z_t) &= (K(x_t) + K(y_t)) \bmod N.
\end{aligned}$$

Построенный таким способом комплексный поток (z_1, \dots, z_t, \dots) будет иметь среднее значение $z = xy$. Таким образом, операция умножения двух комплексных потоков реализуется одним сумматором разрядности n .

Операция полусуммы строится аналогично полусумме бинарных потоков.

3. Поточковый нейрон

Предлагаемая здесь схема цифрового нейрона состоит только из небольшого количества простых логических элементов, нескольких сумматоров и нескольких генераторов случайных чисел. Это обеспечивает низкую стоимость аппаратной реализации и возможность наращивания количества синапсов.

В разработанной модели нейрона информация представляется в виде среднего значения стохастических последовательностей (потоков). Следуя [28, 29, 34, 62], такие нейроны будем называть потоковыми.

Потоковые нейроны разрабатываются с целью уменьшения аппаратных затрат и стоимости аппаратной реализации нейронных сетей. Это позволяет создавать сети больших размеров и увеличивать число синапсов. Положительным свойством потоковых нейронов также является относительная устойчивость к ошибкам передачи данных.

Потоковые нейроны были разработаны на основе биологических данных о возможности передачи информации между нейронами в виде серий импульсов определенной плотности [1, 33, 35].

В статье [28] рассматривается реализация потокового нейрона, в котором число генераторов случайных битов пропорционально числу синапсов нейрона. В работе [7] число генераторов случайных битов было сокращено до $\log_2 N$, где N — количество синапсов.

Изложенная в настоящей работе схема имеет количество генераторов случайных битов, не зависящее от количества синапсов.

Далее приводится описание схемы, математическая модель, результаты экспериментов на программной модели. Для обучения используется обычный метод Хебба, модифицированный метод Хебба и метод оптимизации приближенной функции ошибки. Для оптимизации используется алгоритм Rprop [58]. Результаты обучения проверяются с помощью точного вычисления функции нейрона, а также на программной модели.

3.1. Основные элементы нейрона

Рассматриваемый нейрон имеет N входов, на которые поступают последовательности $(x_{i1}, \dots, x_{it}, \dots)$, где $x_{it} \in \{-1, 0, 1\}$, $(N + 1)$ синаптических коэффициентов w_0, \dots, w_N и регистр z , в котором накапливается некоторое суммарное значение. На

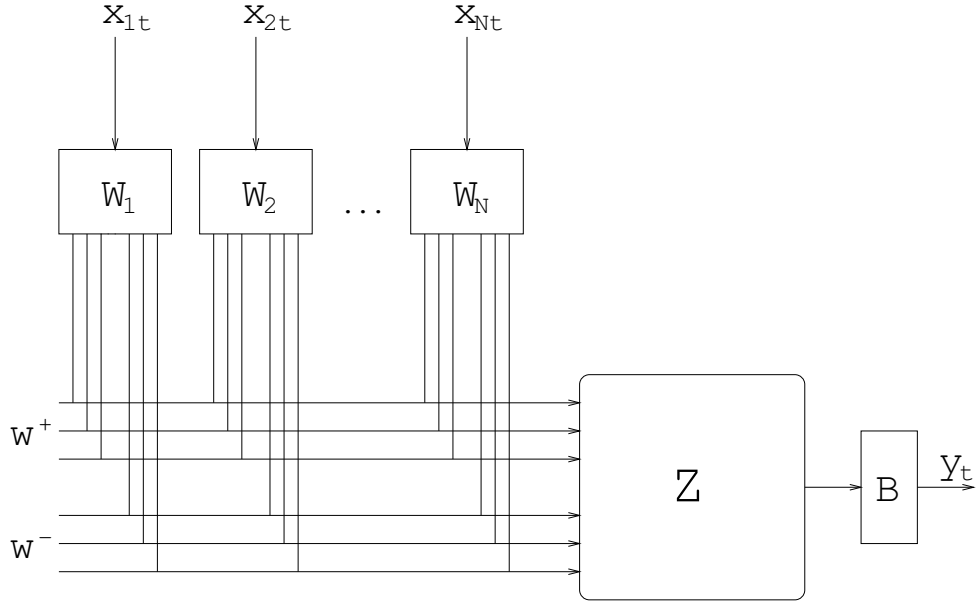


Рис. 3. Схема потокового нейрона

выход нейрона подается последовательность y_t , формируемая в соответствии со знаком значения регистра z и коэффициентом разреженности b .

Синаптические коэффициенты w_i кодируются w' битами и битом знака, регистр z кодируется $(z' + w' + 2)$ битами и представляется в форме дополнения до 1. Коэффициент b кодируется b' битами. Будем считать, что величины принимают следующие значения:

$$\begin{aligned}
 w_i &\in \{-2^{w'} + 1, \dots, 2^{w'} - 1\}, \\
 z &\in \{-2^{z'+w'+1}, \dots, 2^{z'+w'+1} - 1\}, \\
 b &\in [0; 1).
 \end{aligned}$$

В состав нейрона всего входит:

- 1) $3Nw' + 2N + 2 + b'$ битовых логических элементов «И»;
- 2) две шины квазисуммирования разрядности w' ;
- 3) четыре сумматора разрядностей $w' + 1$, $w' + 2$, $w' + 3$ и $z' + w' + 2$;
- 4) b' генераторов случайных битов.

На рис. 3 изображена схема описываемого нейрона. Блоки $\mathbf{W}_1 - \mathbf{W}_4$ содержат регистры с соответствующими синаптическими коэффициентами и устройства выбора. Блок \mathbf{Z} содержит регистры z и w_0 , а также ряд сумматоров. Блок \mathbf{B} содержит регистр b и осуществляет вероятностную фильтрацию.

3.2. Описание работы нейрона

Введем необходимые обозначения. Определим функцию $\text{bit}_k w$, которая принимает значение k -го бита неотрицательного числа w . Пусть

$$w = \sum_i 2^{i-1} b_i, \quad b_i \in \{0, 1\}.$$

Тогда положим

$$\text{bit}_k w = b_k.$$

Определим функцию знака $\text{sgn}(z)$:

$$\text{sgn}(z) = \begin{cases} -1, & \text{если } z < 0; \\ 1, & \text{если } z \geq 0. \end{cases} \quad (12)$$

Опишем работу нейрона в момент времени t .

На входы нейрона поступают значения x_{it} , $i \in \{1, \dots, N\}$.

В зависимости от знака входного значения x_{it} и знака синаптического коэффициента w_i , на одну из шин квазисуммирования подается значение $|w_i|$. Обозначим поступающие на положительную шину значения $w_i^+(t)$, а поступающие на отрицательную шину — $w_i^-(t)$:

$$w_i^+(t) = \begin{cases} |w_i|, & \text{если } w_i x_{it} > 0; \\ 0, & \text{если } w_i x_{it} \leq 0, \end{cases}$$

$$w_i^-(t) = \begin{cases} |w_i|, & \text{если } w_i x_{it} < 0; \\ 0, & \text{если } w_i x_{it} \geq 0. \end{cases}$$

$$\text{bit}_k w_i^+(t) = \begin{cases} 1, & \text{если } w_i x_{it} \text{ bit}_k |w_i| > 0; \\ 0, & \text{если } w_i x_{it} \text{ bit}_k |w_i| \leq 0, \end{cases} \quad (13)$$

$$\text{bit}_k w_i^-(t) = \begin{cases} 1, & \text{если } w_i x_{it} \text{ bit}_k |w_i| < 0; \\ 0, & \text{если } w_i x_{it} \text{ bit}_k |w_i| \geq 0. \end{cases} \quad (14)$$

$$k \in \{1, \dots, w'\}.$$

Биты значений $w^+(t)$ и $w^-(t)$, получаемые на шинах в момент времени t , будут равны:

$$\text{bit}_k w^+(t) = \bigvee_{i=1}^N \text{bit}_k w_i^+(t),$$

$$\text{bit}_k w^-(t) = \bigvee_{i=1}^N \text{bit}_k w_i^-(t),$$

$$k \in \{1, \dots, w'\}.$$

Значения $w^+(t)$ и $w^-(t)$ поступают на вход блока \mathbf{Z} . На шаге t значение регистра z изменяется следующим образом:

$$\begin{aligned} z_{t+1} &= z_t - z_t 2^{-z'} + \Delta z_t, \\ \Delta z_t &= w^+(t) - w^-(t) + w_0. \end{aligned} \quad (15)$$

Переполнение регистра z невозможно, так как:

$$\begin{aligned} \Delta z &\leq 2^{w'+1} - 2, \\ \max(z - z 2^{-z'} + \Delta z) &= \max(z - z 2^{-z'}) + \max(\Delta z) = \\ &= 2^{z'+w'+1} - 1 - 2^{w'+1} + 2^{-z'} + 2^{w'+1} - 2 = \\ &= 2^{z'+w'+1} + 2^{-z'} - 3 \leq 2^{z'+w'+1} - 1. \end{aligned}$$

Аналогично для нижнего предела.

На выход блока \mathbf{Z} подается знак значения регистра z , определяемый в соответствии с (12).

На выход блока \mathbf{B} с вероятностью b подается значение 0 и с вероятностью $(1 - b)$ значение $\text{sgn}(z_t)$:

$$\begin{aligned} \mathbf{P}\{y_t = 0\} &= b, \\ \mathbf{P}\{y_t = \text{sgn}(z_t)\} &= 1 - b. \end{aligned}$$

3.3. Вычисление средних значений

Вычислим среднее значение счетчика z при заданном распределении случайных величин x_{it} и заданных значениях w_i .

Пусть величины x_{it} независимы для всех i и t . Тогда, учитывая (13–14), получаем:

$$\begin{aligned} P_k^+ &= \mathbf{P}\{\text{bit}_k w^+(t) = 1\} = 1 - \prod_{i=1}^N \mathbf{P}\{x_{it} w_i \text{bit}_k |w_i| \leq 0\}, \\ P_k^- &= \mathbf{P}\{\text{bit}_k w^-(t) = 1\} = 1 - \prod_{i=1}^N \mathbf{P}\{x_{it} w_i \text{bit}_k |w_i| \geq 0\}, \\ &k \in \{1, \dots, w'\}. \end{aligned}$$

Соответственно, математическое ожидание w^+ , w^- :

$$\begin{aligned} \mathbf{M}w^+ &= \sum_{k=1}^{w'} 2^{k-1} P_k^+, \\ \mathbf{M}w^- &= \sum_{k=1}^{w'} 2^{k-1} P_k^-. \end{aligned}$$

Итерационное выражение (15) можно преобразовать в сумму последовательности:

$$z_{t+1} = \sum_{\tau=0}^{t-1} (1 - 2^{-z'})^\tau \Delta z_{t-\tau}$$

При вычислении $\mathbb{M}z_t$ и дисперсии $\mathbb{D}z_t$ при $t \rightarrow \infty$ получаем сумму бесконечной геометрической прогрессии с коэффициентами $(1 - 2^{-z'})$ и $(1 - 2^{-z'})^2$ соответственно.

Среднее значение счетчика z_t при $t \rightarrow \infty$:

$$\begin{aligned} \mathbb{M}z_t &= 2^{z'} \mathbb{M}\Delta z + O((1 - 2^{-z'})^t) = \\ &= 2^{z'} (w_0 + \mathbb{M}w^+ - \mathbb{M}w^-) + O((1 - 2^{-z'})^t) = \\ &= 2^{z'} \left(w_0 + \sum_{k=1}^{w'} 2^{k-1} (P_k^+ - P_k^-) \right) + O((1 - 2^{-z'})^t). \end{aligned} \quad (16)$$

Дисперсия z_t при $t \rightarrow \infty$, при условии независимости Δz_t :

$$\begin{aligned} \mathbb{D}z_t &= \frac{2^{z'}}{2 - 2^{-z'}} \mathbb{D}\Delta z + O((1 - 2^{-z'})^{2t}) \leq \\ &\leq \frac{2^{z'} (2^{w'} - 1)^2}{2 - 2^{-z'}} + O((1 - 2^{-z'})^{2t}). \end{aligned}$$

Вычислим среднее значение z_t при $t \rightarrow \infty$ в одном частном случае. Пусть

$$\begin{aligned} X^+ &= \{ i \mid \mathbb{P}\{x_{it} = 0\} = b \wedge \mathbb{P}\{x_{it}w_i > 0\} = 1 - b \}, \\ X^- &= \{ i \mid \mathbb{P}\{x_{it} = 0\} = b \wedge \mathbb{P}\{x_{it}w_i < 0\} = 1 - b \}, \\ X^0 &= \{ i \mid w_i = 0 \}, \\ X^+ \cup X^- \cup X^0 &= \{1, \dots, N\}. \end{aligned}$$

Тогда значения P_k^+, P_k^- можно получить следующим образом (предполагая $0^0 = 1$):

$$\begin{aligned} N_k^+ &= \sum_{i \in X^+} \text{bit}_k |w_i|, \\ N_k^- &= \sum_{i \in X^-} \text{bit}_k |w_i|, \\ P_k^+ &= 1 - b^{N_k^+}, \\ P_k^- &= 1 - b^{N_k^-}. \end{aligned}$$

Соответственно,

$$\mathbb{M}z_t = 2^{z'} \left(w_0 + \sum_{k=1}^{w'} 2^{k-1} (b^{N_k^-} - b^{N_k^+}) \right). \quad (17)$$

3.4. Обоснование перехода к линейной модели

Для оптимизационного обучения нейрона необходимо перейти от дискретной функции, зависящей от отдельных битов значений весовых коэффициентов, к какой-либо дифференцируемой функции, зависящей от значений весовых коэффициентов.

Будем считать, что на вход нейрона подаются последовательности, сформированные другими нейронами с таким же параметром b , то есть

$$P\{x_{it} = 0\} = b.$$

Далее будет показано, что при увеличении параметра b отклонение функции нейрона (16) от некоторой линейной функции уменьшается.

Для простоты будем считать, что:

$$\forall i, t \quad w_i x_{it} \geq 0.$$

При этом будем рассматривать только Mw^+ , так как $Mw^- = 0$. Рассмотрим связь значения квазисуммы Mw^+ и суммы S :

$$S = \sum_{i=1}^N |w_i|.$$

В соответствии с (17) имеем:

$$Mw^+ = \sum_{k=1}^{w'} 2^{k-1} (1 - b^{N_k^+}),$$

$$S = \sum_{k=1}^{w'} 2^{k-1} N_k^+.$$

Можно вывести следующие ограничения на Mw^+ :

$$\left. \begin{aligned} 1 - b^{N_k^+} &\geq \frac{1 - b^N}{N} N_k^+, \\ Mw^+ &\geq \frac{1 - b^N}{N} S, \\ Mw^+ &\leq W(1 - b^{S/W}), \\ Mw^+ &\leq (1 - b)S, \end{aligned} \right\} \quad (18)$$

где $W = 2^{w'} - 1$,

$$N_k^+ \in \{0, \dots, N\},$$

$$S \in \{0, \dots, NW\}.$$

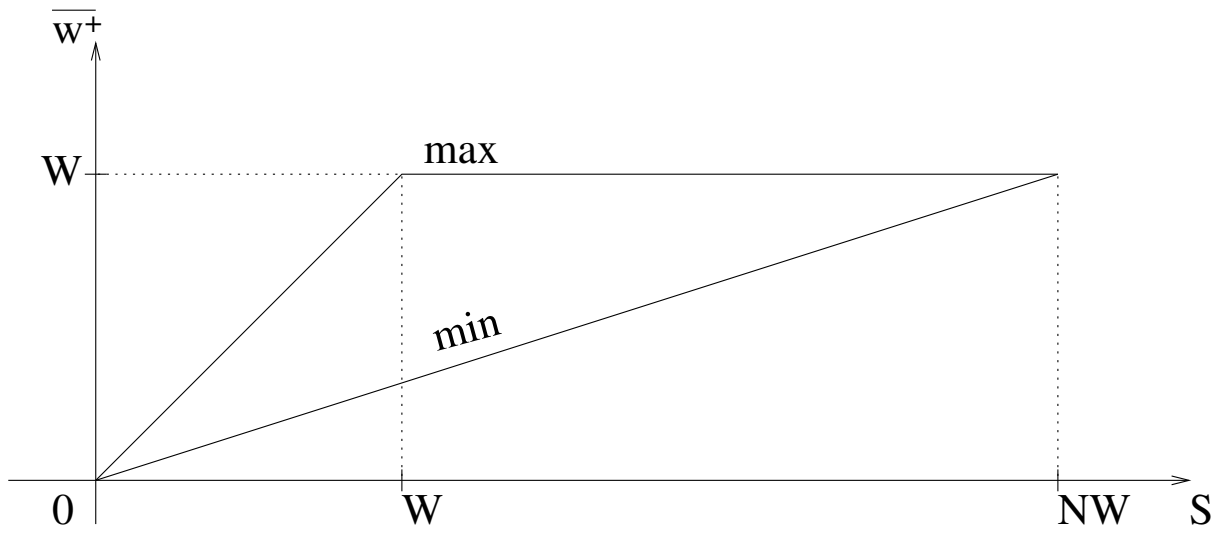


Рис. 4. Зависимость Mw^+ от S при $b = 0$.

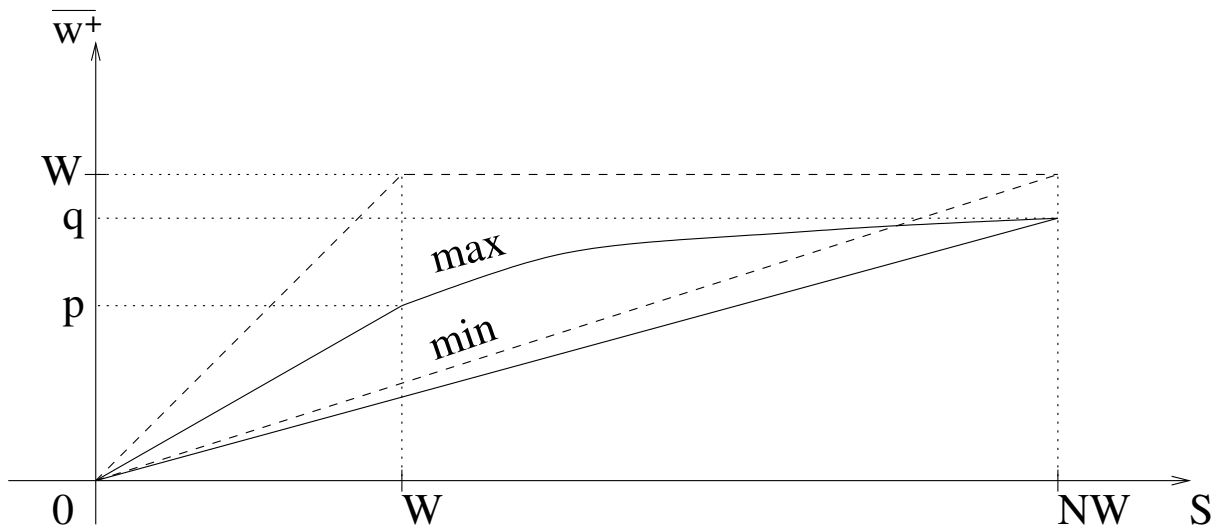


Рис. 5. Зависимость Mw^+ от S при $0 < b < 1$.

Таким образом, можно обеспечить достаточное приближение квазисуммы к линейной функции за счет выбора параметра b близким к 1.

Области, образованные системой ограничений (18), показаны на рис. 4–5. Верхняя линия на этих рисунках соответствует верхнему пределу Mw^+ при данном S , а нижняя линия — нижнему. Обозначены следующие величины:

$$\begin{aligned} W &= 2^{w'} - 1, \\ p &= W(1 - b), \\ q &= W(1 - b^N). \end{aligned}$$

3.5. Полносвязная сеть и ее обучение

Будем рассматривать множество из N нейронов, в котором выход каждого нейрона поступает на входы всех остальных нейронов, связь на себя отсутствует. Будем обучать данную сеть для распознавания изображений из заданного множества $\{\mathbf{X}^1, \dots, \mathbf{X}^P\}$, $\mathbf{X}^m = (X_1^m, \dots, X_N^m)$, где $X_i^m \in \{-1, 1\}$.

При обучении полносвязной нейронной сети будем обучать каждый нейрон независимо. Обозначим обучаемый нейрон индексом j .

На первом шаге обучения вычисляются весовые коэффициенты w_{ij} по методу Хебба, по модифицированному методу Хебба или методом оптимизации. Эти методы рассмотрены ниже.

На втором шаге обучения вычисляется w_{0j} . Для каждого изображения вычисляется $z = Z(\mathbf{X}^m)$ в соответствии с (17), и берутся предельные значения (предполагая $\min \emptyset = 2^{w'}$ и $\max \emptyset = -2^{w'}$):

$$\begin{aligned} M_j^+ &= \min_{\substack{m \in \{1, \dots, P\} \\ X_j^m > 0}} 2^{-z'-1} Z(\mathbf{X}^m), \\ M_j^- &= \max_{\substack{m \in \{1, \dots, P\} \\ X_j^m < 0}} 2^{-z'-1} Z(\mathbf{X}^m), \\ w_{0j} &= -\frac{1}{2}(M_j^+ + M_j^-). \end{aligned}$$

Критерием правильного обучения является условие $M_j^+ > M_j^-$.

Метод Хебба. Весовые коэффициенты вычисляются по методу Хебба [47] следующим способом:

$$w_{ij} = \frac{2^{w'} - 1}{P} \sum_{m=1}^P X_j^m X_i^m,$$

Модифицированный метод Хебба. Предлагается следующий вариант правила Хебба. Для нейрона j (предполагая $0/0 = 1/2$):

$$\begin{aligned}
 A_{ij}^+ &= \frac{1}{4} \sum_{m=1}^P (1 + X_i^m)(1 + X_j^m), \\
 A_{ij}^- &= \frac{1}{4} \sum_{m=1}^P (1 - X_i^m)(1 + X_j^m), \\
 B_{ij}^+ &= \frac{1}{4} \sum_{m=1}^P (1 + X_i^m)(1 - X_j^m), \\
 B_{ij}^- &= \frac{1}{4} \sum_{m=1}^P (1 - X_i^m)(1 - X_j^m), \\
 w_{ij} &= \left(2^{w'} - 1\right) \left(\frac{A_{ij}^+}{A_{ij}^+ + B_{ij}^+} - \frac{A_{ij}^-}{A_{ij}^- + B_{ij}^-} \right), \\
 i &\in \{1, \dots, N\},
 \end{aligned}$$

где A_{ij}^+ , A_{ij}^- , B_{ij}^+ , B_{ij}^- — количество элементов в подмножествах множества образов.

Оптимизационное обучение. При оптимизации будем считать величины w_{ij} вещественными. После оптимизации они нормируются и округляются до ближайшего целого. Будем также считать, что на вход нейрона поступают вещественные величины x_i .

Заменяем функцию активации $\text{sgn}(z)$ на сигмоид $S(z)$:

$$S(z) = \frac{2}{1 + e^{-z}} - 1.$$

Заменяем также сложную функцию квазисуммирования (17), реализованную данным нейроном, более простой линейной функцией \bar{Z} :

$$\bar{Z}(\mathbf{x}, \mathbf{w}) = \frac{1 - b^N}{N} \sum_{i=1}^N w_i x_i.$$

При достаточно большом значении параметра b квазисумма оказывается достаточно близкой к линейной функции. Это позволяет успешно использовать стандартные оптимизационные методы.

Оптимизационная задача заключается в нахождении коэффициентов w_{1j}, \dots, w_{Nj} , при которых функция E принимает наименьшее значение:

$$E(w_{1j}, \dots, w_{Nj}) = \sum_{m=1}^P (S(\bar{Z}(\mathbf{X}^m, \mathbf{w})) - X_j^m)^2,$$

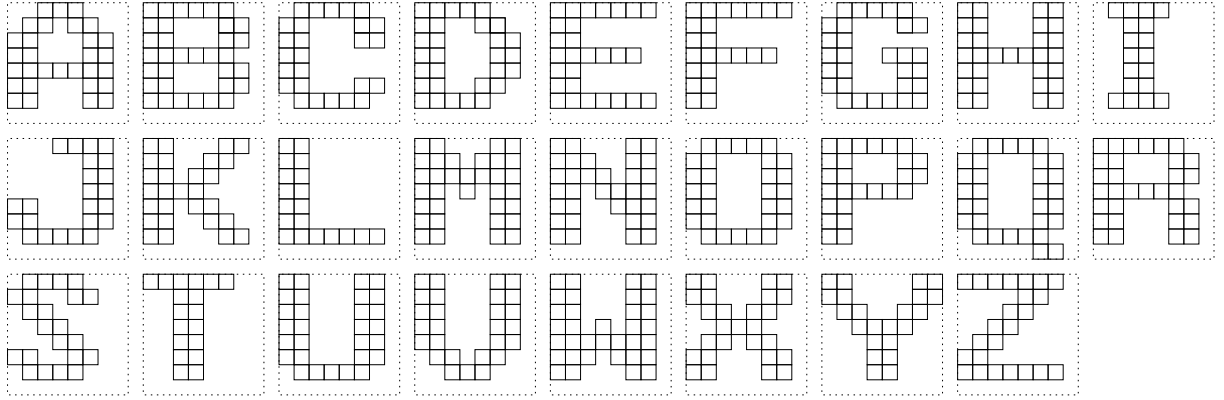


Рис. 6. Эталонные изображения.

где $S(x)$ — функция активации, j — номер обучаемого нейрона. Параметры w_{0j} и w_{jj} принимаются равными нулю.

При моделировании используется оптимизационный алгоритм Rprop [58].

3.6. Результаты эксперимента

Моделировалась полносвязная сеть 8×8 , решавшая задачу восстановления изображения по образцу с помехами.

Сеть обучалась вышеописанными методами на заглавные буквы латинского алфавита A–Z, заданные матрицами 8×8 из стандартного шрифта. Матрицы изображений показаны на рис. 6.

К каждому эталонному изображению добавлялось от 1 до 10 помех (изменений состояния ячейки на противоположное). Положение первой помехи определялось по номеру эксперимента, остальные выбирались случайно. Сеть работала максимум 10000 шагов, при совпадении состояния сети с оригинальным изображением моделирование прекращалось, так как каждое эталонное изображение является устойчивым состоянием сети. Устойчивость проверялась заранее с помощью функции (17). Для каждой буквы и каждого количества помех эта процедура повторялась сериями по 64 испытания. Моделирование прекращалось между сериями по условию $\zeta^* \leq 1\%$, где ζ^* — измеренное среднеквадратичное отклонение процента правильных восстановлений изображения. Таким образом, по правилу 3ζ погрешность измеренного процента восстановлений не превышает 3% с достаточной степенью достоверности.

Эксперимент проводился с различными значениями параметров b , w' и z' . Достаточно хорошие результаты дают значения $w' = 8$, $z' = 7$, $b > 0.5$.

Обучение с помощью стандартного метода Хебба позволяет запомнить только 5

букв А–Е. Таблицы 2 и 3 позволяют сравнить стандартный метод Хебба и модифицированный метод Хебба. В этих таблицах приведены проценты восстановления исходных изображений по отношению к общему числу экспериментов. Из полученных данных видно, что модифицированный метод Хебба превосходит обычный метод Хебба по стабильности эталонных состояний.

Модифицированный метод Хебба позволяет запомнить только 7 букв А–Г. В таблицах 4 и 5 показаны результаты моделирования сети, обученной на 7 букв с помощью модифицированного метода Хебба и метода оптимизации соответственно. Можно заметить, что метод оптимизации дает более сбалансированное обучение.

С помощью метода оптимизации удастся запомнить все буквы А–Z. В таблице 6 показаны результаты моделирования сети, обученной на 26 букв методом оптимизации. По вертикали таблицы указаны буквы, по горизонтали — количество помех от 1 до 10.

Кол-во помех	A	B	C	D	E	Среднее
1 (1.6%)	99.2	100	100	100	100	99.84
2 (3.1%)	98.8	100	100	100	100	99.77
3 (4.7%)	96.0	100	98.4	99.5	100	98.79
4 (6.3%)	93.1	99.2	97.8	96.6	98.4	97.02
5 (7.8%)	90.0	100	95.6	100	98.2	96.77
6 (9.4%)	85.9	100	94.0	96.5	96.6	94.59
7 (10.9%)	83.7	95.0	92.1	93.8	95.8	92.09
8 (12.5%)	79.2	93.8	88.5	92.1	95.2	90.73
9 (14.0%)	75.2	89.3	83.9	90.0	92.5	86.17
10 (15.7%)	71.8	85.2	78.5	86.8	87.4	81.95

Таблица 2. Результат эксперимента при обучении по стандартному методу Хебба, 5 изображений.

Кол-во помех	A	B	C	D	E	Среднее
1 (1.6%)	100	100	100	100	100	100.0
2 (3.1%)	100	100	100	100	100	100.0
3 (4.7%)	100	100	100	100	100	100.0
4 (6.3%)	100	100	100	98.4	100	99.69
5 (7.8%)	98.8	99.5	97.8	100	100	99.22
6 (9.4%)	98.8	98.4	98.8	95.9	99.2	98.23
7 (10.9%)	100	96.4	98.2	93.8	98.8	97.46
8 (12.5%)	98.8	94.1	100	93.0	98.0	96.78
9 (14.0%)	98.0	90.9	95.1	89.8	95.7	93.88
10 (15.7%)	97.0	89.3	93.1	89.4	94.7	92.69

Таблица 3. Результат эксперимента при обучении по модифицированному методу Хебба, 5 изображений.

Кол-во помех	A	B	C	D	E	F	G	Среднее
1 (1.6%)	100.0	93.9	96.9	100.0	100.0	99.5	96.9	98.16
2 (3.1%)	100.0	88.9	93.3	100.0	100.0	100.0	93.3	96.49
3 (4.7%)	100.0	82.8	89.7	100.0	100.0	100.0	89.7	94.60
4 (6.3%)	100.0	77.5	85.7	95.7	98.4	100.0	84.3	91.66
5 (7.8%)	100.0	73.4	81.9	92.9	98.4	98.4	80.6	89.37
6 (9.4%)	100.0	68.9	77.6	89.7	96.4	96.0	77.0	86.50
7 (10.9%)	100.0	64.6	74.1	85.8	96.6	95.7	71.5	84.03
8 (12.5%)	100.0	58.4	68.5	81.7	95.0	94.7	67.9	80.88
9 (14.0%)	100.0	54.8	63.9	77.4	91.3	92.9	61.2	77.37
10 (15.7%)	100.0	50.1	60.4	71.4	91.4	90.1	55.7	74.17

Таблица 4. Результат эксперимента при обучении по модифицированному методу Хебба, 7 изображений.

Кол-во помех	A	B	C	D	E	F	G	Среднее
1 (1.6%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.00
2 (3.1%)	100.0	99.2	100.0	100.0	99.5	100.0	100.0	99.81
3 (4.7%)	97.9	99.5	98.8	97.8	100.0	98.2	97.7	98.54
4 (6.3%)	97.2	97.7	98.8	98.4	96.4	96.2	98.4	97.60
5 (7.8%)	96.7	94.4	97.0	96.0	94.7	94.7	95.8	95.63
6 (9.4%)	94.1	91.8	95.1	94.3	93.3	91.9	93.2	93.38
7 (10.9%)	90.7	89.0	91.4	90.0	89.8	88.2	91.6	90.10
8 (12.5%)	88.1	85.2	89.8	87.5	85.5	85.6	88.6	87.19
9 (14.0%)	85.2	80.0	87.2	84.0	81.1	80.0	86.1	83.37
10 (15.7%)	79.9	77.1	83.6	79.4	76.5	77.4	83.6	79.65

Таблица 5. Результат эксперимента при обучении методом оптимизации, 7 изображений.

	1	2	3	4	5	6	7	8	9	10
A	98.4	94.2	90.3	81.4	75.2	67.8	60.5	53.4	46.4	41.1
B	98.4	96.6	93.1	88.2	83.6	76.4	70.7	65.6	52.3	45.6
C	98.2	97.0	92.1	87.7	79.8	74.0	67.7	59.3	41.0	35.3
D	97.0	93.3	88.2	82.6	75.3	67.5	60.5	53.1	44.7	35.3
E	96.6	92.8	86.7	81.4	74.9	70.1	62.1	56.1	48.6	41.6
F	97.6	94.9	90.5	85.0	78.0	70.3	62.4	53.0	41.7	35.1
G	96.7	90.9	86.4	79.9	73.5	65.6	59.1	52.1	39.6	34.5
H	99.5	98.0	95.3	92.3	88.2	84.3	79.5	76.2	67.5	60.3
I	98.2	94.7	88.7	82.8	75.9	66.4	58.9	49.7	41.0	32.8
J	98.4	96.5	94.3	88.9	86.1	81.1	76.5	69.8	56.2	47.8
K	98.8	97.6	93.5	87.8	81.4	74.9	66.6	58.4	55.4	48.1
L	94.3	88.0	81.9	73.0	66.2	59.2	51.3	43.9	25.6	21.3
M	98.2	95.4	91.5	86.0	80.8	72.5	66.1	58.9	44.7	37.6
N	99.4	98.8	95.3	91.3	86.9	82.9	76.1	70.4	61.0	53.4
O	97.8	97.5	93.8	90.5	85.8	80.4	75.5	71.3	61.0	56.6
P	98.8	98.4	93.5	88.9	84.0	78.1	70.9	63.8	47.6	39.2
Q	98.4	97.8	93.1	90.7	86.8	81.2	76.0	69.7	60.0	55.2
R	99.2	98.0	96.0	93.5	90.9	86.3	82.8	78.2	68.4	62.4
S	98.8	95.8	91.3	86.1	80.7	74.7	67.3	60.3	50.8	43.1
T	98.8	95.0	89.0	82.4	74.5	66.2	55.7	49.1	36.3	28.2
U	98.2	96.7	92.8	89.0	86.1	79.2	74.3	68.5	52.7	44.9
V	98.2	96.0	91.4	86.2	80.6	73.3	67.6	58.8	53.6	46.2
W	95.7	90.0	83.9	77.6	69.4	59.1	50.8	42.2	35.4	27.1
X	99.5	98.2	95.3	90.9	86.2	79.5	74.0	67.4	40.4	32.9
Y	98.4	95.7	92.1	85.5	77.0	69.5	61.5	53.7	32.2	25.3
Z	97.0	95.0	89.3	85.2	80.4	74.7	68.2	61.2	54.1	48.8
Среднее	98.05	95.49	91.13	85.96	80.32	73.66	67.03	60.16	48.39	41.53

Таблица 6. Результат эксперимента при обучении методом оптимизации, 26 изображений.

4. Поточковый нейрон с альтернативными синапсами

В предыдущей главе 3 предлагалась схема потокового нейрона, в которой имеется не зависящее от количества синапсов число генераторов случайных битов. На выход этого нейрона поступает последовательность значений из множества $\{-1, 0, 1\}$.

В этой главе рассматривается модификация этого потокового нейрона, в которой на выход поступает две последовательности y_t^+ и y_t^- , для положительных и отрицательных значений. Каждый синапс имеет два весовых коэффициента, соответствующих положительным и отрицательным значениям, поступающим на вход нейрона. Следуя [8], такие нейроны будем называть А-нейронами или нейронами с альтернативными синапсами. В той же статье показано, что А-нейроны в среднем обладают лучшими классифицирующими способностями по сравнению с обычными нейронами. Формальная модель А-нейрона была разработана на основе биологических данных о существовании пар согласованно функционирующих нейронов [43, 50, 64], важную роль этих нейронных структур отмечает Ю. Д. Кропотов [12–15].

Далее исследуется математическая модель потокового А-нейрона. Для обучения используется метод Хебба и оптимизация на приближенной линейной модели. Результаты обучения на приближенной модели проверяются с помощью точного вычисления функции нейрона, а также на программной модели. Для тестирования нейрона на программной модели используется задача ассоциативной памяти на полносвязной сети.

4.1. Основные элементы нейрона

Рассматриваемый нейрон имеет $2N$ входов, на которые поступают последовательности $(x_{i1}^+, \dots, x_{it}^+, \dots)$ и $(x_{i1}^-, \dots, x_{it}^-, \dots)$, где $x_{it}^+, x_{it}^- \in \{0, 1\}$, $i \in \{1 \dots N\}$, t — дискретное время. Выполняется условие $x_{it}^+ x_{it}^- = 0$. Имеется $2N$ синаптических коэффициентов $w_1^+, w_1^-, \dots, w_N^+, w_N^-$, два пороговых коэффициента w_0^+, w_0^- и регистр z , в котором накапливается некоторое суммарное значение. На выход нейрона подаются две последовательности $(y_1^+, \dots, y_t^+, \dots)$ и $(y_1^-, \dots, y_t^-, \dots)$, где $y_t^+, y_t^- \in \{0, 1\}$, формируемые в соответствии со знаком значения регистра z и коэффициентом разреженности b .

Синаптические и пороговые коэффициенты w_i^+ и w_i^- кодируются w' битами и битом знака, регистр z кодируется $(z' + w' + 2)$ битами и представляется в форме дополнения до 1. Коэффициент b кодируется b' битами. Будем считать, что величины

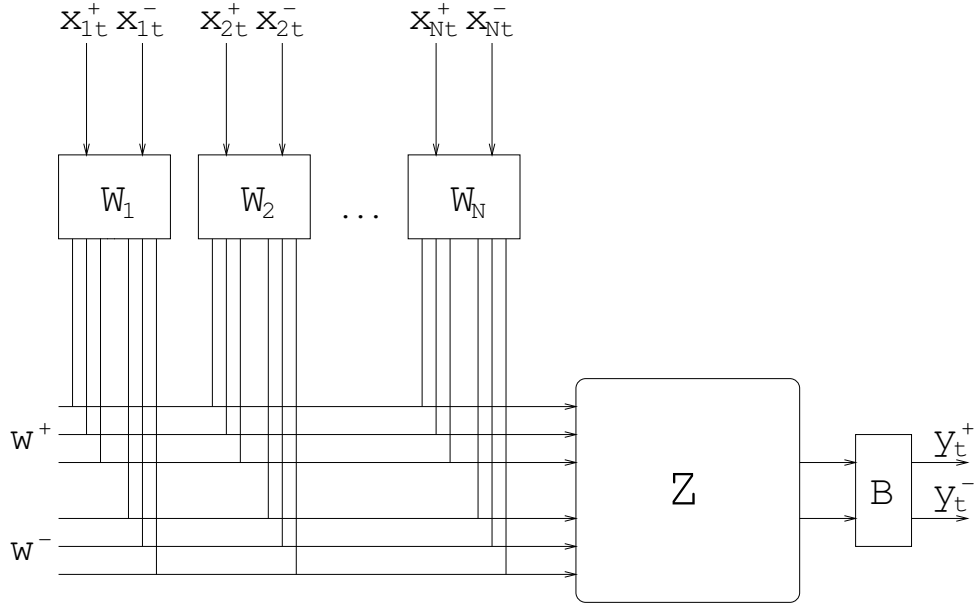


Рис. 7. Схема потокового А-нейрона

принимают следующие значения:

$$w_i^+, w_i^- \in \{-2^{w'} + 1, \dots, 2^{w'} - 1\},$$

$$z \in \{-2^{z'+w'+1}, \dots, 2^{z'+w'+1} - 1\},$$

$$b \in [0; 1).$$

В состав нейрона всего входит:

- 1) $(2N + 2)$ регистров разрядности $w' + 1$, один регистр разрядности $z' + w' + 2$, один регистр разрядности b' .
- 2) $4Nw' + 2N + 2 + b'$ битовых логических элементов «И»;
- 3) две шины квазисуммирования разрядности w' ;
- 4) четыре сумматора разрядностей $w' + 1$, $w' + 2$, $w' + 3$ и $z' + w' + 2$;
- 5) b' генераторов случайных битов.

На рис. 7 изображена схема описываемого нейрона. Блоки $\mathbf{W}_1 - \mathbf{W}_N$ содержат регистры с синаптическими коэффициентами $(w_1^+, w_1^-, \dots, w_N^+, w_N^-)$ и устройства выбора. Блок \mathbf{Z} содержит регистры z , w_0^+ и w_0^- , а также ряд сумматоров. Блок \mathbf{B} содержит регистр b и осуществляет вероятностную фильтрацию.

4.2. Описание работы нейрона

Введем необходимые обозначения. Определим функцию $\text{bit}_k w$, которая принимает значение k -того бита неотрицательного числа w . Пусть

$$w = \sum_i 2^{i-1} b_i, \quad b_i \in \{0, 1\}.$$

Тогда положим

$$\text{bit}_k w = b_k.$$

Определим следующие функции знака:

$$\begin{aligned} \text{sgn}^+(z) &= \begin{cases} 0, & \text{если } z < 0; \\ 1, & \text{если } z \geq 0, \end{cases} \\ \text{sgn}^-(z) &= \begin{cases} 1, & \text{если } z < 0; \\ 0, & \text{если } z \geq 0, \end{cases} \\ \text{sgn}(z) &= \text{sgn}^+(z) - \text{sgn}^-(z). \end{aligned} \tag{19}$$

Будем использовать операцию $[x]$ как округление к ближайшему целому.

Опишем работу нейрона в момент времени t .

На входы нейрона поступают значения x_{it}^+ и x_{it}^- , $i \in \{1, \dots, N\}$.

В зависимости от входных значений x_{it}^+ , x_{it}^- и знаков синаптических коэффициентов w_i^+ и w_i^- , на одну из шин квазисуммирования подается значение $|w_i^+|$ или $|w_i^-|$. Обозначим поступающие из блока \mathbf{W}_i в момент времени t на положительную шину значения $w_i^+(t)$, а на отрицательную шину — $w_i^-(t)$:

$$w_i^+(t) = x_{it}^+ |w_i^+| \text{sgn}^+(w_i^+) + x_{it}^- |w_i^-| \text{sgn}^+(w_i^-),$$

$$w_i^-(t) = x_{it}^+ |w_i^+| \text{sgn}^-(w_i^+) + x_{it}^- |w_i^-| \text{sgn}^-(w_i^-),$$

$$\text{bit}_k w_i^+(t) = x_{it}^+ \text{bit}_k |w_i^+| \text{sgn}^+(w_i^+) + x_{it}^- \text{bit}_k |w_i^-| \text{sgn}^+(w_i^-), \tag{20}$$

$$\text{bit}_k w_i^-(t) = x_{it}^+ \text{bit}_k |w_i^+| \text{sgn}^-(w_i^+) + x_{it}^- \text{bit}_k |w_i^-| \text{sgn}^-(w_i^-), \tag{21}$$

$$k \in \{1, \dots, w'\}.$$

Следует заметить, что в вышеприведенных выражениях сумма может быть заменена на побитовое «ИЛИ», так как хотя бы одно из двух слагаемых равно нулю, а умножение — на операцию «И», так как $x_{it}^\pm \in \{0, 1\}$, $\text{sgn}^\pm(w_i^\pm) \in \{0, 1\}$

Биты значений $w^+(t)$ и $w^-(t)$, получаемые на шинах в момент времени t , будут равны:

$$\begin{aligned}\text{bit}_k w^+(t) &= \bigvee_{i=1}^N \text{bit}_k w_i^+(t), \\ \text{bit}_k w^-(t) &= \bigvee_{i=1}^N \text{bit}_k w_i^-(t), \\ k &\in \{1, \dots, w'\}.\end{aligned}$$

Значения $w^+(t)$ и $w^-(t)$ поступают на вход блока **Z**. На шаге t значение регистра z изменяется следующим образом:

$$\begin{aligned}z_{t+1} &= z_t - \left[z_t 2^{-z'} \right] + \Delta z_t, \\ \Delta z_t &= w^+(t) - w^-(t) + w_0^+ \text{sgn}^+(z_t) - w_0^- \text{sgn}^-(z_t).\end{aligned}\tag{22}$$

Переполнение регистра z невозможно, так как:

$$\begin{aligned}\Delta z &\leq 2^{w'+1} - 2, \\ \max(z - \left[z 2^{-z'} \right] + \Delta z) &= \max(z - \left[z 2^{-z'} \right]) + \max(\Delta z) = \\ &= 2^{z'+w'+1} - 1 - 2^{w'+1} + 2^{w'+1} - 2 = \\ &= 2^{z'+w'+1} - 3 \leq 2^{z'+w'+1} - 1.\end{aligned}$$

Аналогично для нижнего предела.

На выход блока **Z** подаются значения $\text{sgn}^+(z_t)$ и $\text{sgn}^-(z_t)$, определяемые в соответствии с (19). Эти значения поступают в блок **B**.

Будем называть $S_t = \text{sgn}(z_t)$ состоянием нейрона в момент времени t .

На выход блока **B** с вероятностью b подаются значения $0, 0$, и с вероятностью $(1 - b)$ значения $\text{sgn}^+(z_t)$ и $\text{sgn}^-(z_t)$:

$$\begin{aligned}\mathbb{P}\{y_t^+ = 0 \wedge y_t^- = 0\} &= b, \\ \mathbb{P}\{y_t^+ = \text{sgn}^+(z_t) \wedge y_t^- = \text{sgn}^-(z_t)\} &= 1 - b,\end{aligned}$$

При выполнении условия $w_i^- = -w_i^+$ данный нейрон эквивалентен нейрону, описанному в статье [21].

4.3. Значения на выходе нейрона

Оценим вероятности $\mathbb{P}\{z_t \geq 0\}$ и $\mathbb{P}\{z_t < 0\}$ при заданном распределении случайных величин x_{it}^+ , x_{it}^- и заданных значениях w_i^+ , w_i^- .

Пусть величины x_{it}^+, x_{it}^- независимы для всех i и t . Тогда, учитывая (20), (21), получаем:

$$P_k^+ = \mathbb{P} \{ \text{bit}_k w^+(t) = 1 \} = \quad (23)$$

$$= 1 - \prod_{i=1}^N \mathbb{P} \{ x_{it}^+ w_i^+ \text{bit}_k |w_i^+| \leq 0 \} \mathbb{P} \{ x_{it}^- w_i^- \text{bit}_k |w_i^-| \leq 0 \},$$

$$P_k^- = \mathbb{P} \{ \text{bit}_k w^-(t) = 1 \} = \quad (24)$$

$$= 1 - \prod_{i=1}^N \mathbb{P} \{ x_{it}^+ w_i^+ \text{bit}_k |w_i^+| \geq 0 \} \mathbb{P} \{ x_{it}^- w_i^- \text{bit}_k |w_i^-| \geq 0 \},$$

$$k \in \{1, \dots, w'\}.$$

Соответственно, математическое ожидание w^+ , w^- :

$$Mw^+ = \sum_{k=1}^{w'} 2^{k-1} P_k^+,$$

$$Mw^- = \sum_{k=1}^{w'} 2^{k-1} P_k^-,$$

$$Mw = Mw^+ - Mw^- = \sum_{k=1}^{w'} 2^{k-1} (P_k^+ - P_k^-).$$

Имеется несколько случаев, зависящих от значений w_0^+ и w_0^- :

- 1) при $Mw + w_0^+ \geq 0$ и $Mw - w_0^- \geq 0$, $\mathbb{P} \{ z_t \geq 0 \} \geq 1/2$, $Mz_t \geq 0$ при достаточно большом t ;
- 2) при $Mw + w_0^+ \leq 0$ и $Mw - w_0^- \leq 0$, $\mathbb{P} \{ z_t \geq 0 \} \leq 1/2$, $Mz_t \leq 0$ при достаточно большом t ;
- 3) при $Mw + w_0^+ < 0$ и $Mw - w_0^- > 0$ значение z_t колеблется около 0, при этом

$$\mathbb{P} \{ z \geq 0 \} \approx \frac{w_0^- - Mw}{w_0^- + w_0^+};$$

- 4) при $Mw + w_0^+ > 0$ и $Mw - w_0^- < 0$ регистр z имеет два стационарных состояния, больше и меньше нуля, при небольших t значение z_t будет сильно зависеть от начального состояния регистра z .

При выполнении условия $w_0^+ = -w_0^-$ можно легко вычислить Mz_t при $t \rightarrow \infty$:

$$\begin{aligned} Mz_t &= 2^{z'} M\Delta z + O((1 - 2^{-z'})^t) = \\ &= 2^{z'} (w_0^+ + Mw) + O((1 - 2^{-z'})^t). \end{aligned}$$

Вычислим Mw в одном частном случае, необходимом для обучения. Пусть

$$\begin{aligned}\pi_i^+ &= \mathbf{P}\{x_{it}^+ = 0\}/b, \\ \pi_i^- &= \mathbf{P}\{x_{it}^- = 0\}/b,\end{aligned}$$

и выполняется условие $(\pi_i^+ = 1 \wedge \pi_i^- = 0) \vee (\pi_i^+ = 0 \wedge \pi_i^- = 1)$. Тогда значение Mw можно получить следующим образом (предполагая $0^0 = 1$):

$$\begin{aligned}N_k^+ &= \sum_{i=1}^N \pi_i^+ \text{bit}_k |w_i^+| \text{sgn}^+(w_i^+) + \sum_{i=1}^N \pi_i^- \text{bit}_k |w_i^-| \text{sgn}^+(w_i^-), \\ N_k^- &= \sum_{i=1}^N \pi_i^+ \text{bit}_k |w_i^+| \text{sgn}^-(w_i^+) + \sum_{i=1}^N \pi_i^- \text{bit}_k |w_i^-| \text{sgn}^-(w_i^-), \\ Mw &= \sum_{k=1}^{w'} 2^{k-1} (b^{N_k^-} - b^{N_k^+}).\end{aligned}\tag{25}$$

4.4. Ассоциативная память на сети Хопфилда

Для демонстрации возможностей потокового А-нейрона рассмотрим задачу ассоциативной памяти на сети Хопфилда.

Сеть состоит из N нейронов. Выход каждого нейрона поступает на входы всех остальных нейронов, при этом связь на себя отсутствует. Требуется обучить данную сеть на заданное множество изображений $\{\mathbf{X}^1, \dots, \mathbf{X}^P\}$, $\mathbf{X}^m = (X_1^m, \dots, X_N^m)$, $X_i^m \in \{-1, 1\}$.

Для успешного обучения необходимо выполнение условия устойчивости, то есть обученная сеть, начальное состояние которой было установлено в соответствии с любым эталонным изображением, остается в этом состоянии в течение достаточно длительного времени. Желательно, чтобы при наличии помех на изображении сеть возвращалась бы в состояние, соответствующее исходному изображению без помех с достаточно большой вероятностью.

Обозначим синаптические коэффициенты j -того нейрона w_{ij}^\pm , входные последовательности — x_{it}^\pm , выходные — y_{it}^\pm . На синапс с коэффициентом w_{ij}^\pm поступает выходная последовательность i -того нейрона y_{it}^\pm . Выполняется условие $w_{jj}^\pm = 0$.

Будем обучать каждый нейрон независимо. Обозначим обучаемый нейрон индексом j . Требуется найти коэффициенты w_{ij}^\pm ($0 \leq i \leq N$), такие что

$$\begin{aligned}\text{sgn}(W(\mathbf{X}^m) + w_0^+ \text{sgn}^+(X_j^m) - w_0^- \text{sgn}^-(X_j^m)) &= \text{sgn}(X_j^m) \\ (\forall m \in \{1, \dots, P\}),\end{aligned}\tag{26}$$

где $W(\mathbf{X}^m)$ вычисляется в соответствии с (25), полагая

$$\begin{aligned}\pi_i^+ &= \operatorname{sgn}^+(X_i^m), \\ \pi_i^- &= \operatorname{sgn}^-(X_i^m).\end{aligned}$$

На первом шаге обучения вычисляются величины w_i^\pm , ($1 \leq i \leq N$) методом Хебба или методом оптимизации. На втором шаге вычисляются величины w_0^\pm следующим способом:

$$\begin{aligned}w_0^+ &= \left[2^{w'} \xi - \min_{X_j^m > 0} W(\mathbf{X}^m) \right], \\ w_0^- &= \left[\max_{X_j^m < 0} W(\mathbf{X}^m) - 2^{w'} \xi \right],\end{aligned}$$

где ξ — небольшое положительное число. От значения ξ зависит, насколько легко нейрон меняет состояние.

4.5. Обучение по методу Хебба

Обучение по стандартному методу Хебба без использования альтернативных синапсов:

$$\begin{aligned}w_{ij}^+ &= \left[\frac{2^{w'} - 1}{P} \sum_{m=1}^P X_j^m X_i^m \right], \\ w_{ij}^- &= -w_{ij}^+.\end{aligned}$$

Обучение с помощью модифицированного метода Хебба, описанного в статье [21], без альтернативных синапсов (здесь $0/0=0$):

$$\begin{aligned}a_{ij}^+ &= \frac{\sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^+(X_i^m)}{\sum_{m=1}^P \operatorname{sgn}^+(X_i^m)}, \\ a_{ij}^- &= \frac{\sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^-(X_i^m)}{\sum_{m=1}^P \operatorname{sgn}^-(X_i^m)}, \\ w_{ij}^+ &= \left[(2^{w'} - 1)(a_{ij}^+ - a_{ij}^-)/2 \right], \\ w_{ij}^- &= -w_{ij}^+.\end{aligned}$$

Обучение с помощью стандартного метода Хебба, с использованием альтернативных синапсов (описано в статье [8]):

$$w_{ij}^+ = \left[\frac{2^{w'} - 1}{P} \sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^+(X_i^m) \right],$$

$$w_{ij}^- = \left[\frac{2^{w'} - 1}{P} \sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^-(X_i^m) \right].$$

Обучение с помощью модифицированного метода Хебба, с использованием альтернативных синапсов (здесь $0/0=0$):

$$w_{ij}^+ = \left[\frac{\sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^+(X_i^m)}{(2^{w'-1}) \sum_{m=1}^P \operatorname{sgn}^+(X_i^m)} \right],$$

$$w_{ij}^- = \left[\frac{\sum_{m=1}^P X_j^m |X_i^m| \operatorname{sgn}^-(X_i^m)}{(2^{w'-1}) \sum_{m=1}^P \operatorname{sgn}^-(X_i^m)} \right].$$

4.6. Оптимизационное обучение

Для обучения нейрона без А-синапсов будем оптимизировать функцию E :

$$E(\boldsymbol{\omega}) = \sum_{m=1}^P (s(f(\mathbf{X}^m, \boldsymbol{\omega})) - X_j^m)^2,$$

$$f(\boldsymbol{\chi}, \boldsymbol{\omega}) = \sum_{i=1}^N \chi_i \omega_i,$$

$$s(x) = \frac{1 - e^{-x}}{1 + e^{-x}}.$$

После нахождения локального минимума $\boldsymbol{\omega}^*$ весовые коэффициенты вычисляются следующим образом:

$$w_i^+ = \left[\frac{(2^{w'} - 1) \omega_i^*}{\max_k |\omega_k^*|} \right],$$

$$w_i^- = -x_i^+.$$

Для обучения А-нейрона будем оптимизировать функцию E_a :

$$E_a(\boldsymbol{\omega}^+, \boldsymbol{\omega}^-) = \sum_{m=1}^P \left(s(f_a(\mathbf{X}^{m(+)}, \mathbf{X}^{m(-)}, \boldsymbol{\omega}^+, \boldsymbol{\omega}^-)) - X_j^m \right)^2,$$

$$\begin{aligned}
f_a(\boldsymbol{\chi}^+, \boldsymbol{\chi}^-, \boldsymbol{\omega}^+, \boldsymbol{\omega}^-) &= \sum_{i=1}^N \chi_i^+ \omega_i^+ + \sum_{i=0}^N \chi_i^- \omega_i^-, \\
\mathbf{X}^{m(+)} &= (X_1^m \operatorname{sgn}^+(X_1^m), \dots, X_N^m \operatorname{sgn}^+(X_N^m)), \\
\mathbf{X}^{m(-)} &= (X_1^m \operatorname{sgn}^-(X_1^m), \dots, X_N^m \operatorname{sgn}^-(X_N^m)), \\
s(x) &= \frac{1 - e^{-x}}{1 + e^{-x}}.
\end{aligned}$$

После нахождения локального минимума $\boldsymbol{\omega}^{+*}$, $\boldsymbol{\omega}^{-*}$ весовые коэффициенты вычисляются следующим образом:

$$\begin{aligned}
w_i^+ &= \left[\frac{(2^{w'} - 1) \omega_i^{+*}}{\max_k |\omega_k^{+*}|} \right], \\
w_i^- &= \left[\frac{(2^{w'} - 1) \omega_i^{-*}}{\max_k |\omega_k^{-*}|} \right].
\end{aligned}$$

При оптимизации используется алгоритм Rprop [58].

4.7. Результаты эксперимента

Моделировалась полносвязная сеть 8×8 , решавшая задачу восстановления изображения по образцу с помехами.

Сеть обучалась вышеописанными методами на 10 заглавных букв латинского алфавита А–J, заданных матрицами 8×8 из стандартного шрифта. Изображения букв приведены на рис. 6.

На основе предыдущих экспериментов были выбраны следующие параметры: $w' = 8$, $z' = 7$, $b = 0.8$.

К каждому эталонному изображению случайным образом добавлялись помехи, состояние нескольких клеток изображения менялось на противоположное. Первая помеха выбиралась по номеру эксперимента, остальные — случайным образом. Сеть работала максимум 10000 шагов, при совпадении состояния сети с оригинальным изображением моделирование прекращалось, так как каждое эталонное изображение является устойчивым состоянием сети. Для каждой буквы эта процедура повторялась сериями по 64 раза, после каждой серии оценивалось среднеквадратичное отклонение, моделирование прекращалось по условию $\varsigma < 0.01$. Таким образом, среднеквадратичное отклонение вероятности восстановления исходного изображения составляет 1%, для среднего по всем 10 буквам — около 0.3%. По правилу 3 ς можно считать, что средний результат по 10 буквам имеет погрешность не более 1%.

В таблицах 7–12 приведены результаты экспериментов с разными методами обучения обычных нейронов и А-нейронов. Указано количество восстановлений исходного изображения в процентах по отношению к общему количеству экспериментов.

Результаты экспериментов показывают, что преимущество потоковых А-нейронов зависит от способа обучения, в случае стандартного метода Хебба эффективность даже ухудшается. В случае модифицированного метода Хебба и при оптимизационном обучении потоковые А-нейроны показывают лучшие результаты по сравнению с обычными потоковыми нейронами.

Кол-во помех	A	B	C	D	E	F	G	H	I	J	Среднее
1 (1.6%)	88.4	95.5	93.0	88.9	96.7	96.7	92.1	94.8	89.0	93.7	92.87
2 (3.1%)	75.5	91.7	85.0	78.3	92.9	91.9	83.4	89.0	79.6	85.2	85.24
3 (4.7%)	60.7	86.0	77.2	68.8	89.0	88.6	75.2	83.2	70.6	73.8	77.32
4 (6.3%)	48.4	81.1	69.1	56.6	85.0	85.0	66.2	74.4	61.0	61.3	68.81
5 (7.8%)	35.9	73.7	60.0	46.3	79.8	80.7	55.8	67.6	52.9	48.2	60.11
6 (9.4%)	25.8	68.0	50.2	36.4	74.0	77.3	47.4	59.0	45.2	36.3	51.94
7 (10.9%)	18.0	59.8	40.9	27.6	65.3	71.2	38.9	51.0	39.5	27.2	43.92
8 (12.5%)	14.1	53.5	30.9	20.1	58.0	66.3	29.9	41.9	32.7	19.0	36.64
9 (14.0%)	9.0	43.4	22.0	13.0	50.1	60.3	21.8	34.9	28.7	11.5	29.46
10 (15.7%)	5.7	36.9	15.7	8.7	40.8	53.6	14.0	26.7	23.9	6.3	23.25

Таблица 7. Результат эксперимента при обучении по методу Хебба без А-синапсов, $\xi = 0.005$.

Кол-во помех	A	B	C	D	E	F	G	H	I	J	Среднее
1 (1.6%)	95.8	85.9	86.1	84.3	96.7	91.7	83.8	92.3	90.5	87.8	89.48
2 (3.1%)	92.1	73.1	72.9	70.1	92.7	84.7	67.3	82.6	81.0	77.4	79.40
3 (4.7%)	87.6	60.0	60.0	55.3	91.2	77.5	54.0	75.6	71.2	65.6	69.80
4 (6.3%)	83.0	48.6	48.7	44.4	87.1	69.5	41.5	65.9	64.8	56.4	60.97
5 (7.8%)	78.7	38.8	37.5	33.1	84.2	63.8	31.7	58.1	55.6	45.6	52.72
6 (9.4%)	73.9	28.6	28.4	25.3	79.8	56.8	23.5	49.8	49.4	36.2	45.17
7 (10.9%)	65.8	21.9	21.1	15.6	75.5	51.2	16.8	43.1	43.2	28.9	38.32
8 (12.5%)	60.8	15.4	14.1	10.9	70.7	45.3	11.5	36.6	37.5	22.9	32.58
9 (14.0%)	54.4	10.6	8.8	7.5	64.3	39.6	7.2	29.1	32.6	15.8	27.00
10 (15.7%)	48.3	7.7	5.2	4.6	59.9	34.3	5.1	23.5	28.6	11.1	22.83

Таблица 8. Результат эксперимента при обучении по методу Хебба для А-нейрона, $\xi = 0.005$.

Кол-во помех	A	B	C	D	E	F	G	H	I	J	Среднее
1 (1.6%)	100	91.9	93.4	97.2	95.9	98.4	90.7	100	100	98.2	96.58
2 (3.1%)	99.5	83.5	87.6	93.4	91.9	96.2	79.6	100	100	95.0	92.67
3 (4.7%)	100	75.7	78.1	89.5	86.1	96.2	70.2	100	100	93.0	88.88
4 (6.3%)	100	68.7	70.7	85.7	80.6	93.1	62.2	100	100	89.8	85.07
5 (7.8%)	100	59.3	64.6	80.3	76.1	90.8	51.8	100	100	85.8	80.86
6 (9.4%)	99.5	51.3	57.5	75.1	70.4	88.8	43.6	100	100	81.6	76.78
7 (10.9%)	96.1	45.2	49.8	68.7	66.5	86.2	34.6	100	100	78.0	72.50
8 (12.5%)	95.5	37.5	42.4	63.5	59.7	84.1	28.5	100	100	73.8	68.51
9 (14.0%)	91.2	32.0	34.6	55.4	56.8	82.7	22.3	98.2	100	69.4	64.27
10 (15.7%)	88.7	26.3	31.0	49.4	50.1	80.0	17.8	100	100	64.3	60.77

Таблица 9. Результат эксперимента при обучении по модифицированному методу Хебба без А-синапсов, $\xi = 0.05$.

Кол-во помех	A	B	C	D	E	F	G	H	I	J	Среднее
1 (1.6%)	100	93.2	96.9	100	97.0	98.4	95.2	100	100	98.4	97.92
2 (3.1%)	100	86.1	93.6	100	94.7	96.6	89.3	100	100	97.2	95.75
3 (4.7%)	100	80.7	87.4	100	92.1	94.7	83.6	100	100	95.2	93.37
4 (6.3%)	100	72.7	83.9	98.8	88.4	93.4	77.9	100	100	92.2	90.73
5 (7.8%)	100	66.2	78.8	96.6	85.3	89.2	72.9	100	100	90.1	87.91
6 (9.4%)	100	58.9	73.5	95.5	83.9	86.0	67.4	100	100	86.4	85.17
7 (10.9%)	100	52.5	68.2	92.2	78.9	83.1	59.8	100	98.8	81.3	81.49
8 (12.5%)	98.0	47.8	61.8	88.9	76.4	80.0	55.3	100	96.9	78.2	78.33
9 (14.0%)	96.6	40.9	58.1	82.1	73.4	74.2	50.5	98.4	92.9	72.6	73.97
10 (15.7%)	95.7	35.8	52.9	77.5	72.2	70.4	44.1	98.2	85.9	66.2	69.89

Таблица 10. Результат эксперимента при обучении по модифицированному методу Хебба для А-нейрона, $\xi = 0.05$.

Кол-во помех	A	B	C	D	E	F	G	H	I	J	Среднее
1 (1.6%)	98.8	99.5	100	100	98.2	100	98.8	100	98.8	100	99.41
2 (3.1%)	96.1	100	100	100	96.2	100	100	100	100	100	99.23
3 (4.7%)	94.2	96.5	97.0	98.8	94.9	100	95.1	98.4	95.5	100	97.05
4 (6.3%)	92.1	95.0	100	98.2	90.9	100	95.1	97.4	91.8	100	96.04
5 (7.8%)	88.8	93.9	94.6	95.3	88.4	100	92.3	93.9	88.9	96.7	93.29
6 (9.4%)	85.5	90.4	91.0	93.5	86.0	97.5	90.0	92.1	84.9	95.4	90.63
7 (10.9%)	82.3	87.6	87.9	93.6	80.1	95.8	88.4	89.3	80.2	92.3	87.76
8 (12.5%)	78.7	84.0	84.7	89.9	78.8	94.6	82.9	84.5	73.8	91.2	84.30
9 (14.0%)	73.8	78.7	80.9	84.5	73.4	92.8	81.9	80.3	69.0	87.6	80.31
10 (15.7%)	68.9	75.0	75.3	80.2	69.3	89.4	79.1	76.8	61.2	83.5	75.89

Таблица 11. Результат эксперимента при оптимизационном обучении без А-синапсов, $\xi = 0.05$.

Кол-во помех	A	B	C	D	E	F	G	H	I	J	Среднее
1 (1.6%)	100	100	100	100	98.4	100	100	100	100	100	99.84
2 (3.1%)	100	100	100	100	96.4	100	99.5	100	100	100	99.59
3 (4.7%)	100	99.5	100	100	95.6	100	98.8	100	98.0	100	99.19
4 (6.3%)	100	100	100	100	93.5	100	100	100	93.7	100	98.72
5 (7.8%)	100	100	99.5	100	91.5	100	97.8	99.2	90.0	100	97.80
6 (9.4%)	100	100	100	99.5	89.3	100	97.0	99.2	83.2	100	96.82
7 (10.9%)	100	100	100	100	87.5	98.8	96.4	100	76.8	100	95.95
8 (12.5%)	98.8	97.8	100	98.2	85.4	100	95.4	95.5	69.1	100	94.03
9 (14.0%)	99.2	96.1	97.8	97.8	82.9	100	94.4	94.7	60.0	99.5	92.23
10 (15.7%)	97.8	94.4	95.3	94.9	80.3	98.4	92.2	93.6	50.5	100	89.74

Таблица 12. Результат эксперимента при оптимизационном обучении для А-нейрона, $\xi = 0.05$.

5. Потокое устройство, выполняющее дискретное преобразование Фурье

Дискретное преобразование Фурье (ДПФ) широко применяется в цифровой технике для спектрального представления информации. ДПФ используется для ряда задач обработки данных, куда входят обработка цифровых сигналов и изображений [31], адаптивное предсказание речи [32], техническое зрение [16], цифровая голография, сейсморазведка и многие другие.

Предлагается схемотехническая модель устройства, выполняющего дискретное преобразование Фурье и использующего для представления чисел стохастические потоки.

Предлагаемая модель устройства преобразует поступающие на ее N ($N = 2^n$) входов последовательности значений X_g

$$X_g = (X_{g,1}, \dots, X_{g,t}, \dots), \\ g \in H = \{0 \dots N - 1\}$$

в N новых последовательностей Y_k

$$Y_k = (Y_{k,1}, \dots, Y_{k,t}, \dots)$$

так, что если средние значения входных последовательностей равны \bar{X}_g , то средние значения выходных последовательностей будут иметь вид

$$\bar{Y}_k = \frac{1}{N} \sum_{g \in H} \bar{X}_g \exp \frac{i 2\pi g k}{N} \quad (27) \\ k \in H.$$

Среднее значение последовательности $A = (A_1, \dots, A_t, \dots)$ определяется как предел

$$\bar{A} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T A_t.$$

С использованием такого представления информации в виде последовательностей была построена достаточно простая схема, которая дает статистическое приближенное значение спектра Фурье. При аппаратной реализации количество элементов имеет порядок $N \log_2 N$, причем используются только $2N - 3$ сумматора разрядности

$\log_2 N$, $N \log_2 N$ логических элементов “И” и $\log_2 N$ генераторов случайных чисел 0, 1.

Впервые подобное представление информации в виде потоков импульсов было использовано для построения моделей нейронов с целью уменьшения аппаратных затрат. Это также соответствует работе биологических нейронов [33]. В работе [62] рассматривается цифровая схема стохастической нейронной сети, в работах [45, 54] используется цифро-аналоговый подход. Здесь используется чисто цифровое представление значений, что исключает накопление погрешностей.

Далее описывается представление комплексных чисел в виде стохастических потоков, описание схемотехнической модели устройства ДПФ, приводится обоснование, результаты моделирования и сравнение с традиционной цифровой схемой.

5.1. Значения и их представление

В данной схеме дискретного преобразования Фурье (ДПФ) значения входных и выходных последовательностей принадлежат конечному множеству комплексных чисел \hat{C} специального вида

$$X_{g,t}, Y_{k,t} \in \hat{C} = \left\{ \exp \frac{i 2\pi k}{N} \mid k \in H \right\}.$$

В аппаратной схеме используются коды чисел. Число $x \in \hat{C}$ кодируется бинарным представлением числа $k \in H$, функция кодирования $K : \hat{C} \rightarrow H$ и обратная к ней определяются следующим образом:

$$\begin{aligned} K(c) &= \frac{N}{i 2\pi} \ln c, \\ K^{-1}(k) &= \exp \frac{i 2\pi k}{N}. \end{aligned}$$

Множество \hat{C} с операцией умножения образует группу, изоморфная операция на множестве кодов H будет сложением по модулю N . То есть, для любых значений $a, b \in \hat{C}$ верно равенство:

$$K(ab) = (K(a) + K(b)) \bmod N.$$

Данный факт облегчает аппаратную реализацию умножения двух значений из множества \hat{C} , позволяя заменить умножение сложением кодов.

Будем обозначать коды входных последовательностей $x_{g,t}$, а выходных — $y_{k,t}$:

$$\begin{aligned} x_{g,t} &= K(X_{g,t}) \\ y_{k,t} &= K(Y_{k,t}). \end{aligned}$$

5.2. Описание схемы ДПФ

При описании схемы будем использовать коды значений для оценки количества элементов схемы. Множество кодов имеет мощность N , двоичное представление числа из множества кодов занимает $\log_2 N$ битов. Сумматор разрядности $\log_2 N$ имеет порядка $\log_2 N$ элементов.

Как уже было сказано, имеется N входов и N выходов. На входы поступают коды значений входных последовательностей, производится их преобразование, и на выход поступают коды значений выходных последовательностей.

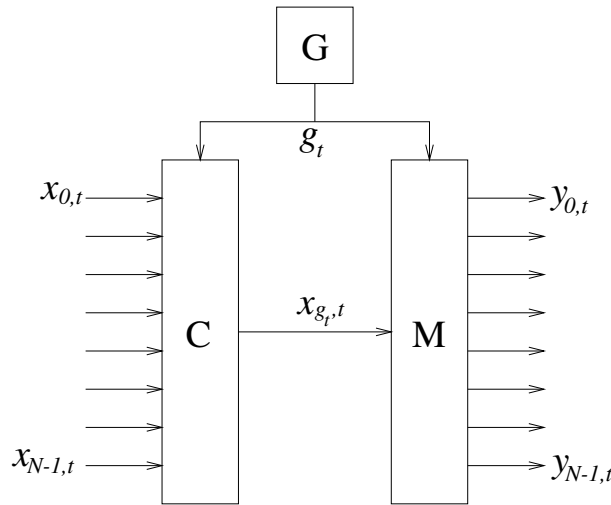


Рис. 8. Структура схемы ДПФ

На рис. 8 буквой G обозначен равномерный генератор последовательности случайных чисел g_t из множества H . Выполняется условие равенства частот элементов g_t :

$$\forall m \in H \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \delta_m(g_t) = \frac{1}{N},$$

где

$$\delta_y(x) = \begin{cases} 0 & \text{если } x \neq y \\ 1 & \text{если } x = y. \end{cases}$$

Этот генератор может быть реализован как $\log_2 N$ независимых генераторов случайных чисел $0, 1$ с вероятностями $1/2$. Таким образом, порядок количества элементов генератора составляет $\log_2 N$. Значения, выдаваемые генератором, поступают на устройство выбора C и на устройство умножения M .

Устройство выбора C принимает на входе N кодов $x_{i,t}$, номер $g_t \in H$ и выдает $x_{g_t,t}$. Порядок количества элементов составляет $N \log_2 N$, так как каждый из N кодов

на входе кодируется с помощью $\log_2 N$ битов и для выбора одного бита необходимо $N - 1$ логических элементов “И”.

Устройство умножения M принимает на входе код $x_{g_t,t}$ и номер $g_t \in H$ и выдает N кодов $y_{k,t}$:

$$y_{k,t} = (x_{g_t,t} + g_t k) \bmod N \quad (28)$$

$$k \in H.$$

Устройство умножения содержит N сложений и умножения на константы $k \in H$. Умножения на константы можно заменить сложениями, причем количество сложений будет иметь порядок N . Все сложения имеют разрядность $\log_2 N$, поэтому порядок количества элементов в множителе будет $N \log_2 N$.

Получили, что общее количество элементов в схеме ДПФ будет иметь порядок $N \log_2 N$.

Один такт работы схемы, обозначенный t , можно описать следующим образом:

- 1) получается число g_t от генератора G , которое подается на устройства C и M ;
- 2) производится выбор кода $x_{g_t,t}$ из входных последовательностей;
- 3) код $x_{g_t,t}$ передается на умножитель M , где параллельно получают коды $y_{k,t}$ в соответствии с (28);
- 4) на входы подаются следующие коды значений входных последовательностей, и процесс повторяется.

Коды $y_{k,t}$, получаемые на выходе схемы, соответствуют значениям $Y_{k,t}$:

$$Y_{k,t} = K^{-1}(y_{k,t}) = X_{g_t,t} \exp \frac{i 2\pi g_t k}{N}. \quad (29)$$

5.3. Обоснование

Аппаратная реализация работает с кодами, но поскольку существует взаимно однозначное соответствие между кодами и значениями, то можно теоретически рассматривать сами значения. В обосновании будем использовать значения из множества \hat{C} .

Определение. Будем называть две последовательности A_t и B_t независимыми, если

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T A_t B_t = \left(\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T A_t \right) \left(\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T B_t \right).$$

Данное условие естественным образом выполняется, если последовательности сформированы независимыми генераторами случайных чисел.

Теорема. Пусть последовательность g_t и последовательности $X_{n,t}$ независимы для всех $n \in H$. Тогда последовательности $Y_{k,t}$, полученные по (29) будут иметь средние значения, определенные в (27).

Доказательство. Средние значения последовательностей $Y_{k,t}$ будут равны

$$\begin{aligned}\bar{Y}_k &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T Y_{k,t} = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T X_{g_t,t} \exp \frac{i 2\pi g_t k}{N}.\end{aligned}\quad (30)$$

Величины g_t могут принимать только значения $0 \dots N - 1$. Разделим сумму на N частей для каждого возможного значения g_t . Тогда

$$\begin{aligned}\bar{Y}_k &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{n \in H} \delta_n(g_t) X_{n,t} \exp \frac{i 2\pi n k}{N} = \\ &= \sum_{n \in H} \exp \frac{i 2\pi n k}{N} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \delta_n(g_t) X_{n,t}.\end{aligned}$$

Используя условие независимости последовательностей $X_{n,t}$ и $\delta_n(g_t)$, и учитывая равенство частот различных элементов последовательности g_t , получаем:

$$\begin{aligned}\bar{Y}_k &= \sum_{n \in H} \left(\bar{X}_n \exp \frac{i 2\pi n k}{N} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \delta_n(g_t) \right) = \\ &= \frac{1}{N} \sum_{n \in H} \left(\bar{X}_n \exp \frac{i 2\pi n k}{N} \right).\end{aligned}$$

Это соответствует (27), что и требовалось доказать.

5.4. Сходимость к среднему значению

Оценим величины отклонения на шаге T :

$$\Delta_{T,k} = \left| \frac{1}{T} \sum_{t=1}^T Y_{k,t} - \bar{Y}_k \right|.$$

Пусть для некоторого значения T

$$\exists \xi \quad \forall n \in H : \left| \frac{1}{T} \sum_{t=1}^T \delta_n(g_t) X_{n,t} - \frac{1}{N} \bar{X}_n \right| < \xi.$$

Тогда

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T Y_{k,t} - \bar{Y}_k &= \frac{1}{T} \sum_{t=1}^T X_{g_t,t} \exp \frac{i 2\pi g_t k}{N} - \frac{1}{N} \sum_{n \in H} \bar{X}_n \exp \frac{i 2\pi n k}{N} = \\ &= \sum_{n \in H} \exp \frac{i 2\pi n k}{N} \left(\frac{1}{T} \sum_{t=1}^T \delta_n(g_t) X_{n,t} - \frac{1}{N} \bar{X}_g \right), \\ \left| \frac{1}{T} \sum_{t=1}^T Y_{k,t} - \bar{Y}_k \right| &< N\xi. \end{aligned}$$

Таким образом, среднее значение выходных последовательностей сходится в худшем случае в N раз медленнее, чем среднее значение последовательности $\delta_t(g_t) X_{g_t,t}$.

Чтобы оценить «среднюю» скорость сходимости, предположим, что каждая последовательность $Y_{k,t}$ состоит из одинаково распределенных независимых случайных величин, принимающих значения из \hat{C} . Тогда дисперсия одной такой случайной величины не будет превышать 1. На шаге T

$$\left| \mathbb{D} \left(\frac{1}{T} \sum_{t=1}^T Y_{k,t} \right) \right| \leq \frac{1}{T^2} \sum_{t=1}^T |\mathbb{D}(Y_{k,t})| \leq \frac{1}{T}.$$

В соответствии с правилом 3σ , можно считать, что погрешность не превышает $3/\sqrt{T}$. Соответственно, для абсолютной точности ξ нужно затратить не более $9\xi^{-2}$ шагов.

5.5. Результаты моделирования

Проводилось моделирование на имитационной модели для исходных дискретных функций, определенных на множестве $\{0, \dots, 127\}$:

$$\begin{aligned} f_1(x) &= (\sin(2\pi k/5) + \cos(2\pi k/11))/2, \\ f_2(x) &= \begin{cases} \sin(2\pi k/5), & \text{при } x < 64; \\ \sin(2\pi k/11), & \text{при } x \geq 64. \end{cases} \end{aligned}$$

Моделирование показало, что данный метод пригоден для вычисления спектра с абсолютной точностью порядка 0.001 за приемлемое время.

На рис. 9–10 показана исходная функция (вверху), модуль спектра, вычисленного обычным методом, и модуль спектра, вычисленного предложенным здесь методом. Два спектра наложены для сравнения.

Показаны некоторые статистические показатели.

- st — количество шагов с начала моделирования.
- $s1$ — максимальное по всем выходам среднеквадратичное отклонение одного значения выходных последовательностей от значения спектра (по модулю).
- $s = s1/\sqrt{st}$
- d — максимальное отклонение среднего значения выходных последовательностей на шаге st от значений спектра (по модулю).

Проценты указаны по отношению к максимальному значению модуля спектра.

5.6. Сравнение с обычной реализацией

Число элементов обычной цифровой реализации быстрого преобразования Фурье (БПФ) имеет порядок $P^2 N \log_2 N$, где P — число битов, необходимых для кодирования комплексного значения функции с требуемой точностью. Коэффициент P^2 возникает из реализации умножения на коэффициенты.

Приведенная схема также имеет порядка $N \log_2 N$ элементов, однако реализация оказывается намного более простой за счет замены умножения сложением.

Точность результата в обычном случае зависит от числа битов P , причем погрешность устранить в принципе невозможно, кроме некоторых частных случаев.

Точность приведенной схемы зависит от времени сбора статистики, причем неустранимая погрешность отсутствует. Для получения абсолютной точности ξ необходимо затратить не более $9\xi^{-2}$ шагов.

В числе достоинств данной схемы — относительно небольшое число элементов, отсутствие неустранимой погрешности, относительная устойчивость к ошибкам.

5.7. Выводы

Разработан метод представления комплексных чисел в виде последовательностей и схема дискретного преобразования Фурье на их основе.

Рассмотрена структура схемы ДПФ, приведена оценка количества элементов схемы и обоснована точность получаемых в пределе результатов.

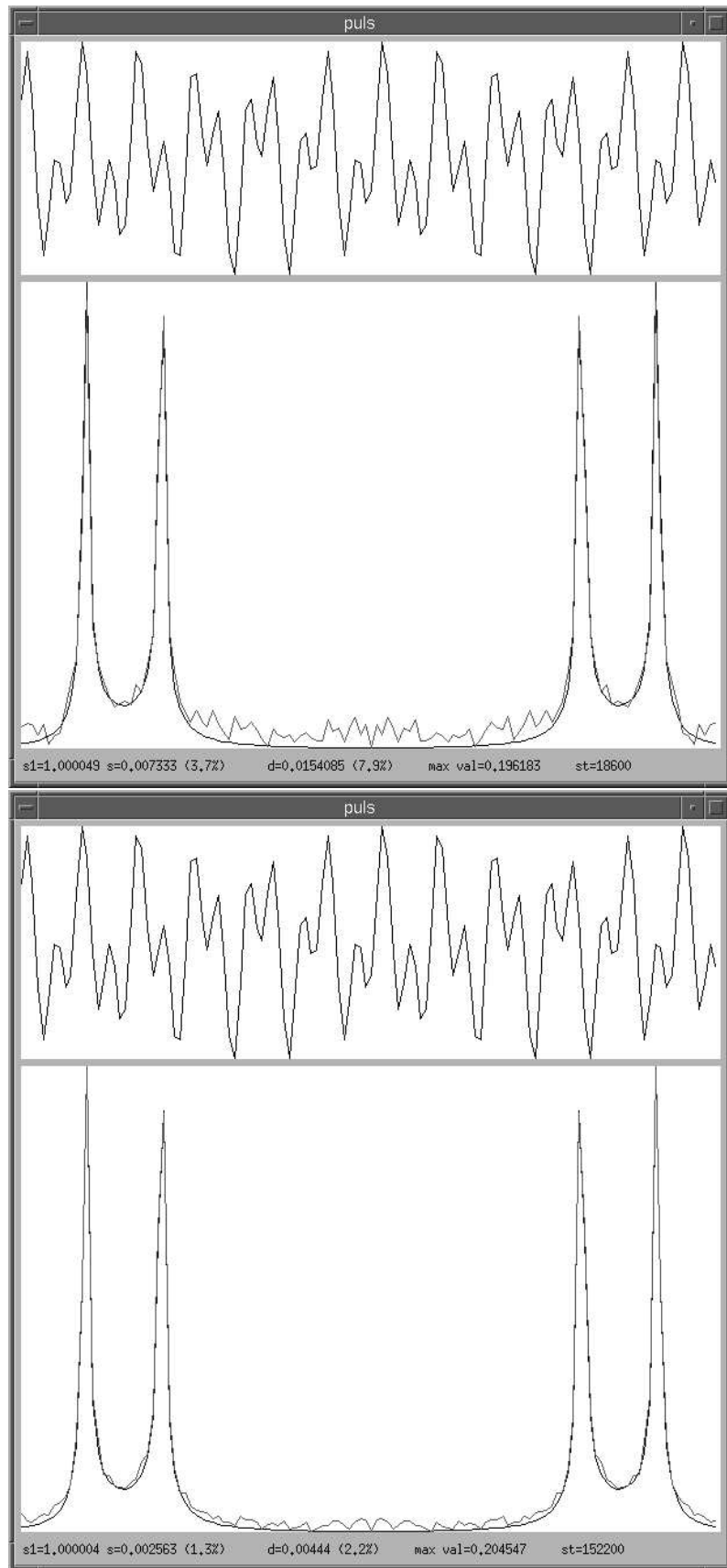


Рис. 9. Моделирование ДПФ для функции f_1

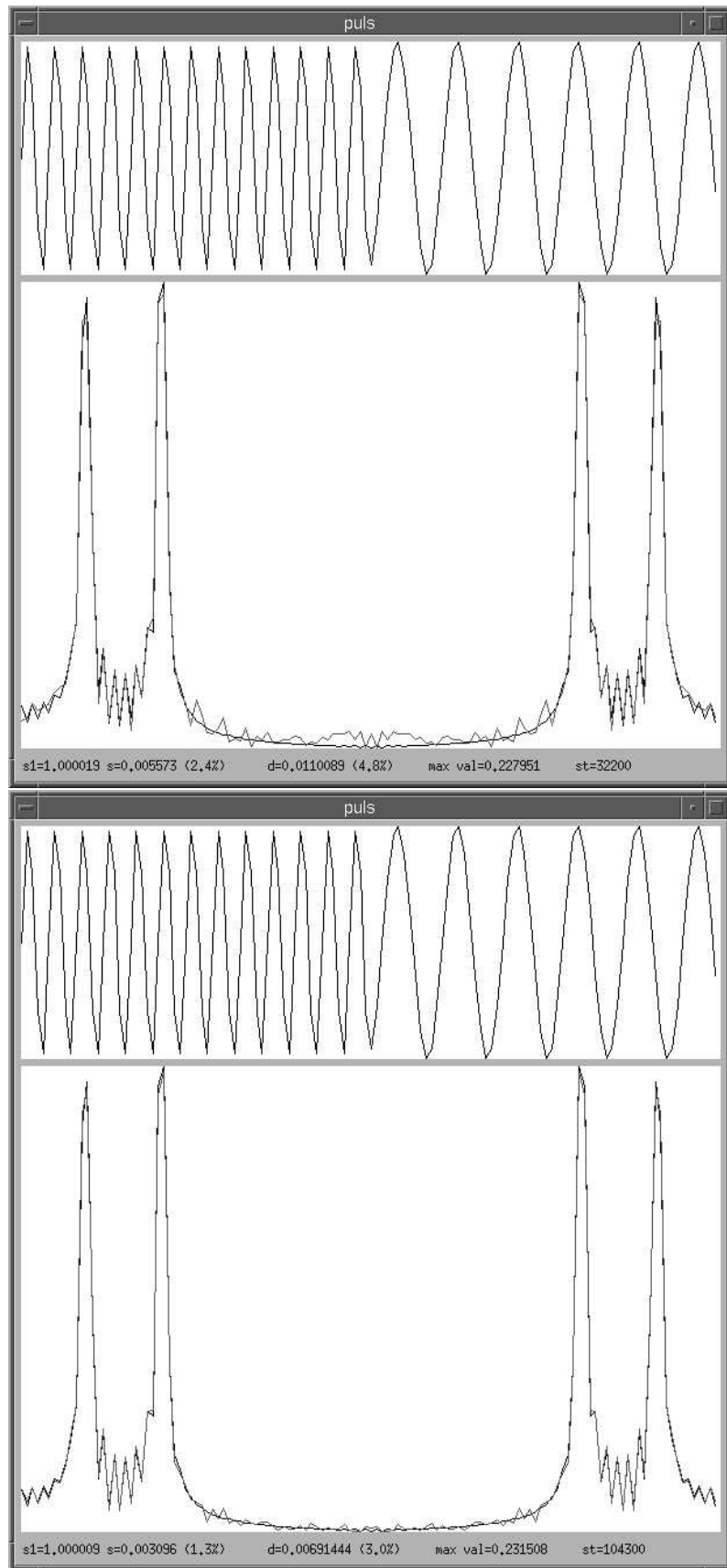


Рис. 10. Моделирование ДПФ для функции f_2

6. Заключение

В настоящей работе предложены новые модели устройств, работающих со средним значением стохастического потока: нейрона; нейрона с альтернативными синапсами; устройства, выполняющего дискретное преобразование Фурье. Данные устройства состоят из небольшого количества простых логических элементов, ячеек памяти и содержат небольшое количество генераторов случайных чисел. Это делает их привлекательными для аппаратной реализации.

Построены математические модели данных устройств и получены основные характеристики распределения состояния нейрона и выходов устройства, выполняющего ДПФ. Доказана работоспособность разработанных устройств.

Для потоковых нейронов предложены два метода обучения: модифицированный метод Хебба и метод оптимизации приближенной функции ошибки.

Для исследования данных устройств и методов обучения была создана программная имитационная модель, позволяющая проводить эксперименты с потоковыми схемами. На основе этой имитационной модели было проведено сравнение методов обучения нейронов, результаты экспериментов подтвердили работоспособность всех трех устройств.

Список рисунков

1.	Реализация умножения и полусуммы бинарных потоков	24
2.	Представимые комплексные значения и коды ($N = 16$)	28
3.	Схема потокового нейрона	31
4.	Зависимость Mw^+ от S при $b = 0$	36
5.	Зависимость Mw^+ от S при $0 < b < 1$	36
6.	Эталонные изображения.	39
7.	Схема потокового А-нейрона	45
8.	Структура схемы ДПФ	59
9.	Моделирование ДПФ для функции f_1	64
10.	Моделирование ДПФ для функции f_2	65

Список таблиц

1.	Операция сравнения кодов потоков	26
2.	Результат эксперимента при обучении по стандартному методу Хебба, 5 изображений.	41
3.	Результат эксперимента при обучении по модифицированному методу Хебба, 5 изображений.	41
4.	Результат эксперимента при обучении по модифицированному методу Хебба, 7 изображений.	42
5.	Результат эксперимента при обучении методом оптимизации, 7 изображений.	42
6.	Результат эксперимента при обучении методом оптимизации, 26 изображений.	43
7.	Результат эксперимента при обучении по методу Хебба без А-синапсов, $\xi = 0.005$	54
8.	Результат эксперимента при обучении по методу Хебба для А-нейрона, $\xi = 0.005$	54
9.	Результат эксперимента при обучении по модифицированному методу Хебба без А-синапсов, $\xi = 0.05$	55
10.	Результат эксперимента при обучении по модифицированному методу Хебба для А-нейрона, $\xi = 0.05$	55
11.	Результат эксперимента при оптимизационном обучении без А-синапсов, $\xi = 0.05$	56
12.	Результат эксперимента при оптимизационном обучении для А-нейрона, $\xi = 0.05$	56

Литература

- [1] Бернс Б. *Неопределенность в нервной системе.* // М.: Мир, 1969. 252 с.
- [2] Бехтерева Н.П. *Здоровый и больной мозг человека* // М.:Наука, 1980, 208 с.
- [3] Горбань А. Н. *Обучение нейронных сетей* // М.: СП "ParaGraph". 1990. 160 с.
- [4] Горбань А. Н. *Проекционные сетчатки для обработки бинарных изображений* // Математическое обеспечение и архитектура ЭВМ: Материалы науч.-техн. конф. «Проблемы техники и технологий XXI века», 22–25 марта 1994 г. — Красноярск: изд. КГТУ, 1995. С. 6–9.
- [5] Горбань А. Н., Россиев Д. А. *Нейронные сети на персональном компьютере* // Новосибирск: Наука. Сибирская издательская фирма РАН. 1996. 276 с.
- [6] Залманзон Л. А. *Преобразование Фурье, Уолша, Хаара и их применение в управлении, связи и других областях* // М.: Наука. 1989. 493 с.
- [7] Карлин А. К., Малков А. Н., Маматов Ю. А., Тимофеев Е. А., Штерн Г. П. *Вероятностный нейрон, работающий с плотностью потока бинарных импульсов* // Микроэлектроника, т. 27, № 3, 1998, с. 170-175.
- [8] Короткин А. А., Панкратов А. В. *Классифицирующие свойства нейронов с альтернативными синапсами* // Моделирование и анализ информационных систем. Ярославль. 1997. Вып. 4. С. 118–123.
- [9] Кохонен Т. *Ассоциативная память* // М.:Мир, 1980.
- [10] Кохонен Т. *Ассоциативные запоминающие устройства* // М.:Мир, 1982.
- [11] Красненко Н. П., Федоров В. А. *Применение временных и корреляционных (спектральных) окон для оценивания параметров спектральной плотности стационарного случайного процесса* // Изв. вузов. Радиоэлектроника. 1985. № 7. С. 79-82.
- [12] Кропотов Ю. Д. *Мозговая организация восприятия и памяти: гипотеза о программировании действий* // Физиология человека. 1989. Т. 15. № 3. С. 19–27.
- [13] Кропотов Ю. Д. *Нейроинформатика: Основы, современное состояние и перспективы* // Физиология человека, 1989. Т. 15. С. 130–149.

- [14] Кропотов Ю. Д., Пономарев В. А. *Нейрофизиология целенаправленной деятельности* // СПб.: Наука. Санкт-Петербург изд. фирма, 1993. 171 с.
- [15] Кропотов Ю. Д., Пахомов С. В. *Математическое моделирование механизмов обработки сигналов нейронными популяциями в головном мозге: Сообщ. II. Влияние синаптической пластичности на свойства нейронной сети с локальными связями в стационарном режиме* // Физиология человека, 1994. Т. 10. С. 405–410.
- [16] Кузьмин С. А. *Методы определения ориентации объектов в системах технического зрения* // Измерения, контроль, автоматизация. 1986. Вып. 2. С. 36-45.
- [17] Лебедев А. Н. *Память человека, ее механизмы и границы* // Исследование памяти. М.:Наука, 1990. С. 104–118.
- [18] Лебедев А. Н. *О физиологических основах восприятия и памяти* // Психол. журн. 1992. № 2. С. 30–41.
- [19] Лукьянов А. В. *Представление комплексных чисел в потоковой форме* // Сборник «Моделирование и анализ информационных систем». Ярославль, 1996. № 3. С. 57-61.
- [20] Лукьянов А. В. *Схемотехническая модель преобразования Фурье, работающая со средним значением стохастического потока* // Сборник «Моделирование и анализ информационных систем». Ярославль, 1998. № 4. С. 123-133.
- [21] Лукьянов А. В. *Схемотехническая модель цифрового нейрона, работающая со средним значением стохастического потока* // Моделирование и анализ информационных систем. 1999. Т. 6, № 1. С. 29-35.
- [22] Лукьянов А. В. *Оптимизационное обучение цифрового нейрона, работающего со средним значением стохастического потока* // Моделирование и анализ информационных систем. 1999. Т. 6, № 2. С. 39-42.
- [23] Лукьянов А. В. *Потоковый нейрон с альтернативными синапсами* // Моделирование и анализ информационных систем. 2000. Т. 7, № 1. С. 6-15.
- [24] Лукьянов А. В. *Потоковый нейрон с альтернативными синапсами.* // VIII Всероссийский семинар «Нейроинформатика и ее приложения», материалы семинара. Красноярск. 2000.

- [25] Лукьянов А. В. *Схемотехническая модель цифрового нейрона, работающая со средним значением стохастического потока* // Микроэлектроника. 2001. № 1. (в печати)
- [26] МакКаллок У. С., Питтс В. *Логическое исчисление идей, относящихся к нервной активности* // Нейрокомпьютер. 1992. № 3, 4. С. 40–53.
- [27] Маматов Ю. А., Булычев С. Ф., Карлин А. К., Малков А. Н. *Потоковый нейрон на цифровых элементах* // Нейрокомпьютер. 1993. № 3,4. С. 23-31.
- [28] Маматов Ю. А., Булычев С. Ф., Карлин А. К. и др. *Цифровая реализация потокового нейрона* // Радиотехника и электроника. 1995. № 11. С. 1652-1660.
- [29] Маматов Ю. А., Булычев С. Ф., Карлин А. К. и др. *Схемотехнические модели построения вероятностных нейронов на базе цифровой техники* // Микроэлектроника. 1996. Т. 25, № 1. С. 3-8.
- [30] Престон К., Дафф Дж. Б., Левьяльди С. и др. *Основы клеточной логики с приложениями к обработке изображений в медицине* // Тр. Ин-та инж. по электротехн. и электронике. М., 1979. Т. 67. № 5. С. 149–185.
- [31] Рабинер Л., Гоулд Б. *Теория и применение цифровой обработки сигналов* // М.: Мир. 1978. 848 с.
- [32] Рабинер Л. Р., Шафер Р. В. *Цифровая обработка речевых сигналов* // М.: Радио и связь. 1981. 496 с.
- [33] Соколов Е. Н., Вайткявичус Г. Г. *Нейроинтеллект. От нейрона к нейрокомпьютеру.* // М.: Наука. 1989. 240 с.
- [34] Тимофеев Е. А. *Моделирование нейрона, передающего информацию плотностью потока импульсов* // Автоматика и телемеханика. № 3. 1997. С. 190-199.
- [35] Экклс Дж. *Физиология синапсов.* // М.: Мир. 1966. 396 с.
- [36] Amit D.J. *Modelling Brain Functions. The World of Attractor Neural Networks* // Cambridge University Press. 1989. PP. 504.
- [37] Bandman O.L. *Cellular-neural computations: formal model* // Bulletin of the Novosibirsk Computing Center, Computer Science, issue 3, 1995, p. 1-17.

- [38] Bandman O.L. *Stability of stored patterns in cellular-neural associative memory* // Bulletin of the Novosibirsk Computing Center, Computer Science, issue 4, 1996, p. 1-16.
- [39] Barlow H.B. *Unsupervised learning* // Neural Computation, 1989, No. 1, pp. 295-311.
- [40] Carpenter G.A., Grossberg S. *A massively parallel architecture for a self-organizing neural pattern recognition machine* // Computer Vision, Graphics, and Image Processing, 1987. Vol. 37, pp. 54-115.
- [41] Chua L.O., Yang L. *Cellular Neural Networks: Theory and Applications* // IEEE Transactions on Circuits and Systems. 1988. V. 35. No 10. P. 1257–1290.
- [42] Chua L.O., Roska T., Venetianer P.L. *The CNN is universal as the Turing machine* // IEEE Transactions on Circuits and Systems. 1993. V. 40. No 4. P. 289–291.
- [43] Friesen W.O. *Antifacilitation and facilitation in the cardiac ganglion of the spiny lobster *Punulirus interuptus** // J. Comp. Physiol. 1975. Vol. 101. P. 207-224.
- [44] Fukushima K. *A Neural Network for Visual Pattern Recognition* // Computer, Vol. 21, No. 3, March 1988, pp. 65-75.
- [45] Hamilton A., Murray A.F., Baxter D.J., Churcher S., Reekie H.M., Tarassenko L. *Integrated Pulse Stream Neural Networks: Results, Issues, and Pointers* // IEEE Transactions on Neural Networks. 1992. Vol. 3, No. 3. P. 385.
- [46] Hands M.A., Pfeiffer W., Thienpont H., Kirk A., Hall T.J., Pignon D., and Parmiter P. *Proposal for stochastic bit stream processing using optoelectronic smart pixels: A neural network architectural case study* // Journal of Parallel and Distributed Computing, 1997, Vol. 41, No. 1, pp. 92-108.
- [47] Hebb D.O. *The organization of Behavior* // New York: Wiley, 1949.
- [48] Hopfield J.J., Tank D.W. *Computing with Neural Circuits: a Model* // Science, Vol. 233, 1986, p. 625.
- [49] Jordan M.I. *Attractor dynamics and parallelism in a connectionist sequential machine* // In Proceedings of the Eighth Annual conference of the Cognitive Science Society, 1986, pp. 531-546.

- [50] Kidd M. *Electron microscopy of the inner plexiform layer of the retina in the cat and pigeon* // J. Anat. (London). 96. P. 179–187 (1962).
- [51] Liu D., Michel A. *Sparsely interconnected artificial neuron networks for associative memories* // Lecture notes in Comp. Sci., Vol. 606, p. 155.
- [52] McCulloch W.S., Pitts W. *A logical calculus of ideas immanent in nervous activity* // Bull. Math. Biophys., 1943. No. 5. PP. 115–133.
- [53] Michel A.M., Farrell J.A., Sun H. *Analysis and synthesis techniques for Hopfield type synchronous discrete time neural networks with application to associative memory* // IEEE Transactions, 1990, Vol. 37, No. 11, p. 1356.
- [54] Murray A.F., Corso D.D., Tarassenko L. *Pulse-Stream VLSI Neural Networks Mixing Analog and Digital Techniques* // IEEE Transactions on Neural Networks. 1991. Vol. 2. No. 2. P. 193.
- [55] Murray A.F., Smith A.V.W. *Asynchronous arithmetic for VLSI neural system* // Electron. Lett. Vol. 23. No. 12. PP. 642-643. June 1987.
- [56] Murray A.F., Smith A.V.W. *A novel computational and signaling method for VLSI neural networks* // Proc. European Solid State Circuits Conf. 1987. PP. 19-22.
- [57] Pratt L.Y., Mostow J., Kamm C.A. *Direct transfer of learned information among neural networks.* // Proceedings of the 9th National Conference on Artificial Intelligence (AAAI-91), 1991, pp. 584-580, Anaheim, California.
- [58] Riedmiller M., Braun H. *A direct adaptive method for faster backpropagation learning: The RPROP algorithm* // Proceedings of the IEEE International Conference on Neural Networks (ICNN). P. 586-591. San Francisco, 1993.
- [59] Rosenblatt F. *The Perceptron: A probabilistic model for information storage and organization in the brain* // Psychological review. 1958. No. 65. PP. 386–408.
- [60] Sejnowski T.J., Rosenberg C.R. *Parallel networks that learn to pronounce English text* // Complex Systems, 1987. No. 1, pp. 145-68.
- [61] Tomberg J.E., Kaski K.K.K. *Pulse-Density Modulation Technique in VLSI Implementations of Neural Network Algorithms* // IEEE Journal of solid-state circuits. Vol. 25, No. 5, October 1990.

- [62] Tomlinson M.S., Walker D.J. *DNNA: A digital neural networks architecture* // Proc. Int. Neural Networks Conf. (INNC-90). Vol. 2. 1990. P. 589-592.
- [63] Wang S.S., Lin W.G. *A New Self-Organizing Neural Model for Invariant Pattern-Recognition* // Pattern Recognition, Vol. 29, No. 4, April 1996, pp. 677-687.
- [64] Wässle H., Boycott B.B., Illing R.B. *Morphology and mosaic of on- and off-beta cells in the cat retina and some functional considerations* // Proc. Roy. Soc. London. B. 1981. Vol. 212. PP. 177–195.