МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Учреждение образования «БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ

ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет информационных технологий

Кафедра информационных систем и технологий

Специальность Информационные системы и технологии

**ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №14 НА ТЕМУ:**

**СОГЛАСОВАНИЕ КРИПТОГРАФИЧЕСКИХ КЛЮЧЕЙ НА ОСНОВЕ ТЕХНОЛОГИЙ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ**

Выполнил студент 3 курса 4 группы

Раченок Илья Александрович

2023 г.

**Цель:** изучение основ построения и функционирования искусственных нейронных сетей (ИНС), а также использования ИНС в криптографии; приобретение практических навыков программной реализации алгоритма согласования ключевой информации на основе технологии ИНС.

**Задачи:**

1. Закрепить теоретические знания по основам построения и функционирования ИНС.

2. Усвоить особенности построения, основные алгоритмы взаимного обучения двух связанных нейронных сетей на основе модели ТРМ.

3. Разработать приложение для реализации модели ТРМ с целью согласования двумя сторонами совместного тайного ключа.

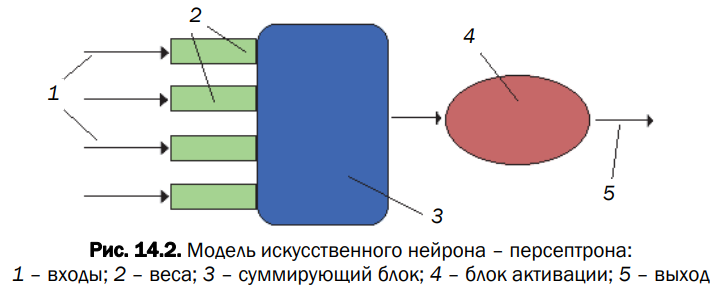
4. Познакомиться с методиками оценки криптостойкости алгоритма на основе ТРМ.

5. Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

**Теоретические сведения**

Был спроектирован искусственный нейрон, называемый также персептроном или перцептроном (англ. *perceptron*), симулирующий работу биологического нейрона (рис. 14.2). Впервые модель искусственного нейрона была представлена У. МакКаллоком и У. Питтсом в 1943 г. в исторической работе.

Эта модель содержит главные элементы биологического нейрона. Вот их эквиваленты: входные значения – дендриты, весовые коэффициенты – синапсы, суммирующий блок – ядро, функция активации – основание аксона, а выходное значение – это аксон.



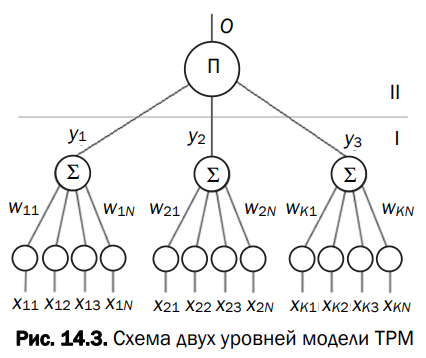
Представленная выше модель нейронной сети в принципе бесполезна, так как она не способна решать поставленные задачи до тех пор, пока не будет соответствующим способом «натренирована». Этот процесс называется обучением и основан на соответствующем подборе коэффициентов вектора весов в контексте решаемой задачи.

Взаимное обучение двух ИНС касается простейшей однонаправленной модели, в которой участвуют два персептрона. Роли таких нейронных сетей заранее не определены, каждая может выполнять функции как учителя, так и ученика (в зависимости от этапа обучения). Это значит, что сети учатся друг у друга, используя для этого полученные результаты, и, стремясь к «общей цели», находят общие элементы в результате своих вычислений. Эта модель обучения, благодаря своим свойствам, может быть использована в криптографии, а именно в определении и согласовании двумя сторонами (двумя ИНС) криптографического ключа (аналогично известному нам протоколу Диффи – Хеллмана).

Именно общее значение векторов весов обоих персептронов, отличающееся только знаком, может быть использовано в качестве секретного ключа. Протокол этого типа выполняет ту же роль, что и хорошо известный протокол обмена ключами Диффи – Хеллмана. Однако следует обратить внимание на два обстоятельства. Первый из них касается безопасности всей системы. Поскольку если у третьей стороны, наблюдающей за процессом обмена, появится возможность получить внутреннее состояние вектора весов одного из персептронов, участвующих в процессе обмена ключами, то она сможет синхронизироваться с ним. Вторая проблема касается особенностей практического применения всей системы. Обмен информацией в компьютерных системах базируется на пространстве битов, или на дискретных величинах. Поэтому обоснованным является вопрос: будет ли правильным (или сходящимся к общему вектору) процесс обмена ключами для векторов весов с дискретными (битовыми) величинами?

Задача решается с помощью многослойной ИНС. Это сеть, называемая машиной четности (англ. *parity machine*) или, более точно, древовидной машиной четности (англ. *Tree Parity Machine*, ТРМ). Нейроны представляют собой персептроны с дискретными векторами весов.

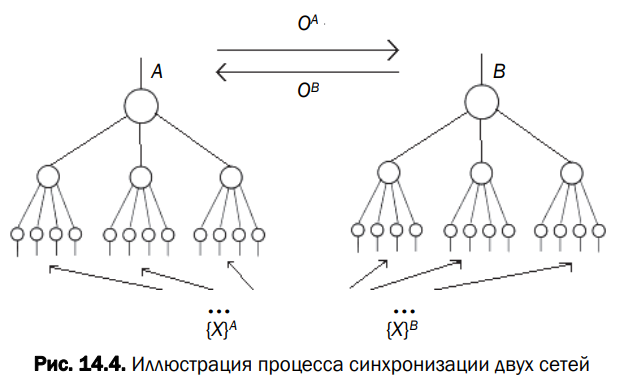
Архитектуру сети составляют K классических персептронов, принадлежащих внутреннему уровню; внешний уровень (обозначен римской цифрой II) объединяет выходы персептронов внутреннего уровня (I) (рис. 14.3, здесь K = 3).



Каждый из персептронов имеет N входов. Следовательно, вектор входных значений всей сети равен KN. В этой архитектуре приняты следующие обозначения: xkj – j-тый вход k-того персептрона, yk – выход k-того персептрона, где k = 1, 2, ..., K и j = 1, 2, ..., N. Для упрощения принято, что величины входного вектора бинарны: xkj = {–1, +1}. Каждый из персептронов имеет веса, которые получают j-тую величину входа в k-том персептроне. Следовательно, величины векторов весов обозначаются как wkj. Выход всей архитектуры О – это произведение всех выходных величин внутренних персептронов. Для простоты в дальнейшем будем говорить о рассматриваемой архитектуре как о TPM-архитектуре с тремя внутренними элементами. На величины вектора весов налагаются следующие ограничения: wkj = {–L, –L + 1, ..., 0, ... L – 1, L}.

Представленная выше архитектура, как и все ее параметры, идентичны для обеих сетей, участвующих в процессе синхронизации. Все структурные элементы (как и в других криптографических системах) известны, однако это не представляет угрозы для всей системы. Секретные элементы, на которых основывается конструкция системы, – это начальные состояния векторов весов. Естественно, сети, участвующие в процессе синхронизации, также взаимно «не знают» векторов весов.

Общая схема процесса синхронизации двух сетей в графическом виде представлена на рис. 14.4. Кратко проанализируем особенности процесса синхронизации сетей на основе архитектуры ТРМ.



Две ИНС, обученные на основании вышеуказанной схемы, впоследствии достигают состояния синхронизации. Это означает, что их векторы весов имеют идентичные величины (wA = wB).

Число шагов, за которое векторы весов достигают одинаковых значений, зависит от начальных состояний векторов весов, входных значений и от таких параметров системы, как K и N. Вся процедура взаимного обучения никак не влияет на два первых параметра, так как выбор начальных состояний векторов весов и значение входных величин генерируются случайным способом. Однако параметры K и N пропорционально влияют на увеличение времени синхронизации.

Самый простой способ атаки на синхронизирующиеся сети А и В (со стороны С) основывается на том, что оппонент может знать архитектуру обеих сетей, а также иметь всю информацию, которой обмениваются между собой сети.

**Практическое задание**

1. Разработать приложение, реализующее модель ТСР, – эмулятор процесса синхронизации весовых коэффициентов двух ИНС.

Для начала установим входные значения в соответствии с вариантом:

|  |  |  |  |
| --- | --- | --- | --- |
| Алгоритм обучения | K | N | ±L |
| Случайного блуждания | 4 | 8 | 4 |

На основе разработанного приложения можно осуществить синхронизацию двух TPM-машин с заданными параметрами, используя правило обучения случайного блуждания. В процессе каждой синхронизации приложение подсчитывает количество шагов, необходимых для завершения процедуры, и вычисляет среднее значение шагов на данном этапе. Затем на основе полученных данных строится график, иллюстрирующий зависимость среднего числа шагов, требуемых для полной синхронизации обеих машин. В сумме приложение производит 70 таких синхронизаций.

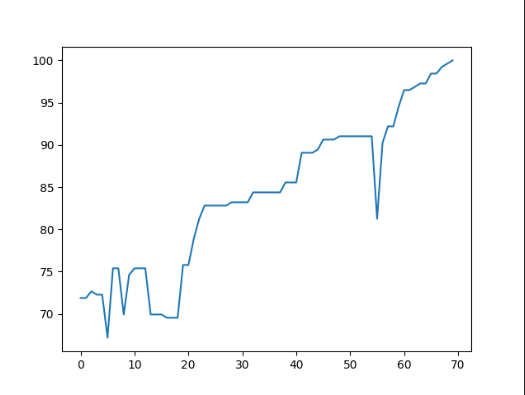


Рисунок 14.5 – График зависимости среднего числа шагов для полной синхронизации от числа синхронизаций

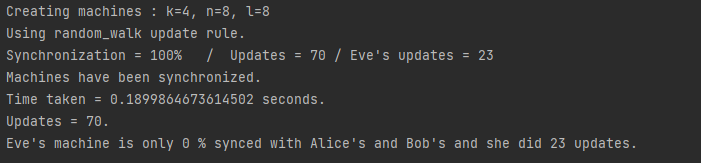
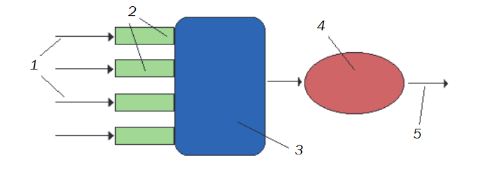


Рисунок 14.6 – Вывод в консоли

**Вывод:** в ходе лабораторной работы мы смогли синхронизировать две машины с значением 70 шагов, при этом наблюдающая сторона – третья машина, не синхронизировалась с другими. На сто процентную синхронизацию машина затратила 0.19 секунды.

**Контрольные вопросы**

**1. Изобразить схематично структуру персептрона и пояснить аналогии между его компонентами и частями биологического нейрона.**



1 – Входы перцептрона (дендриты);

2 – Веса (синапсы);

3 – Суммирующий блок (ядро);

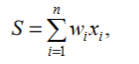
4 – Блок активации (основание аксона);

5 – Выход (аксон).

**2. Охарактеризовать (и показать на примерах) области использования ИНС.**

ИНС активно используются при анализе и обработке больших массивов информации (задачи классификации и прогнозирования) и при обучении искусственного интеллекта. Некоторые методы на основе ИНС могут быть использованы в криптографии для согласования ключевой информации.

**3. Как в простейшем виде записывается формальное представление персептрона?**



Где S – выходной сигнал;

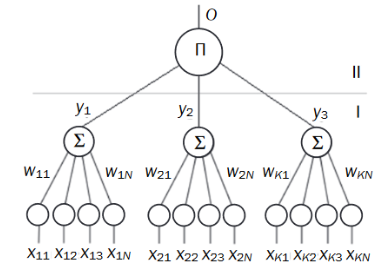
wi – вес входного сигнала;

xi – значение входного сигнала.

**4. Охарактеризовать (и показать на примерах) области использования ИНС в криптографии.**

Некоторые методы на основе ИНС могут быть использованы в криптографии для согласования ключевой информации. Примером может выступать метод на основе синхронизации TPM-машин.

**5. Дать пояснение к структуре и функционалу информационной системы на основе ИНС, предназначенной для согласования ключевой информации.**



На входные нейроны подаются сигналы, которые суммируются скрытыми нейронами. Сигналы с выходов скрытых нейронов суммируются выходным нейроном.

**6. Дать характеристику известным алгоритмам обучения ИНС.**

Все алгоритмы обучения ИНС можно разделить на методы обучения с учителем и без учителя. Обучение без учителя предусматривает наличие только входных данных (такие методы могут решать задачи кластеризации или масштабирования входных данных).

**7. Какие алгоритмы используются для обучения ИНС, предназначенных для согласования ключевой информации между двумя сторонами?**

**Правило обучения Хебба** – правило для синхронизации TPM-машин, которое может быть сформулировано следующим образом: если нейрон получает входной сигнал другого нейрона и выходные значения обоих нейронов имеют одинаковый знак, веса входных значений должны быть увеличены.

**Правило Анти-Хебба** – если нейрон получает входной сигнал другого нейрона и выходные значения обоих нейронов имеют одинаковый знак, веса входных значений должны быть уменьшены.

**Обучение на основе случайного блуждания** – строится на запоминании и забывании значений из определённого диапазона с корректировками на каждом шаге в зависимости от степени синхронизации.

**8. Могут ли легитимно участвовать в процессе синхронизации более трех сетей? Мотивируйте ответ.**

Да. Если эти сети выступают в качестве посредников передачи информации, можно создать цепочку из TPM-машин, которые на каждом шаге будут синхронизироваться между собой. Однако подобный способ является крайне ресурсоёмким по вычислительной мощности.

**9. Дать характеристику криптостойкости системы на основе двух взаимодействующих ИНС.**

Так как при синхронизации используются случайные вектора входных значений, системы обладает высокой криптостойкостью, однако она должна быть достаточно быстрой, чтобы при потенциальная атакующая нейронная сеть не успела синхронизироваться с одной из машин. При равных вычислительных мощностях атакующая нейросеть никогда не будет успевать синхронизироватся, так как ей, как правило, не известны сгенерированные случайные значения входных параметров.

**10. Какие виды атак на нейрокриптографические системы вам известны? В чем заключается их сущность?**

**Брутфорс** – атакующий должен проверить все возможные варианты весов.

**Обучение собственной TPM** – атакующий имеет такую же TPM, которые есть в атакуемой системе, и пытается синхронизировать её с системными машинами. Для подобных атак нужны большие вычислительные мощности, так как атакующая TPM обучается медленнее, чем синхронизируются системные машины.