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

Выполнила:

Студентка 3 курса 1 группы ФИТ

Шимчёнок Елизавета Константиновна

**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**

Решением проблемы распространения ключа в симметричных системах шифрования занимались такие ученые, как А. Климов, А. Митягин, А. Шамир, Б.Р. Приянка, А. Мартин. Ими было предложено использование криптографического алгоритма, основанного на синхронизации нейронных сетей.

Основные принципы построения и основы моделирования ИНС:

Человеческий мозг состоит. Главным образом, из большого числа соединенных друг с другом элементарных нервных клеток – нейронов.

Основная способность нейронов – возможность управления и выработки импульсов.

Нейрон:

– дендрит (тело клетки, имеющее древовидную структуру);

– аксон (длинный отросток нейрона (длинное волокно), который руководит импульсами от тела клетки к другим нейронам).

Синопсисы – биохимические стыки, соединяющие аксон с дендритами других нейронов.

У. МакКаллок и У.Питтсом представили модель искусственного нейрона, называемый также персептроном или перцептроном, симулирующий работу биологического нейрона.

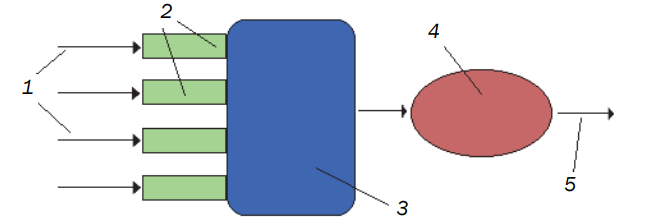


Рисунок 1 – Модель искусственного нейрона – персептрона: 1 – входы; 2 – веса; 3 – суммирующий блок; 4 – блок активации; 5 – выход

Главные элементы искусственного нейрона, являющиеся эквивалентами биологического нейрона:

Входные значения – дендриты;

Весовые коэффициенты – синапсы;

Суммирующий блок – ядро;

Функция активации – основание аксона;

А выходное значение – аксон.

Значение выхода нейрона = скалярному произведению входных значений на векторы весовых коэффициентов (нейрон с линейной функцией активации). Действие такой сети основывается на вычислении выходов каждого из нейронов на основе общего выходного вектора *X*.

**Архитектура TPM**

В основе нейросетевой криптографии лежат TPM.

Древовидная машина чётности (Tree Parity Machine, TPM) – многоуровневая нейронная сеть прямого распространения, состоящая из:

– K\*N входных нейронов,

– K скрытых нейронов,

– одного выходного нейрона.

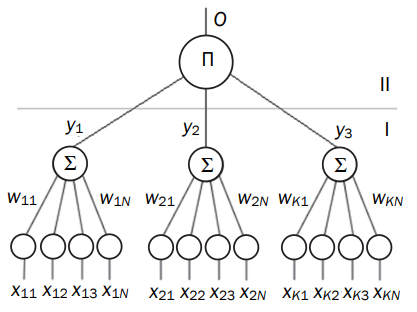


Рисунок 2 – Схема двух уровней модели TPM

Входные значения являются случ. и двоичными значениями ∈{–1, +1}.

Веса – дискретными числами между -L и +L, где L – заданный параметр.

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

Как и в других НС, в TPM взвешенная сумма используется для определения выходных значений, а состояние скрытого нейрона определяется его суммарным входом.

Просто так такая модель бесполезна, так как не способна решать задачи.

Процесс обучения нейронной сети – обучение – основан на соответствующем подборе коэффициентов вектора весов в контексте решаемой задачи.

Основной концепцией синхронизации TPM является взаимное обучение. Процесс синхронизации начинается со случайно заданных входных векторов и случайно выбранных весовых коэффициентов. Обновление весов происходит только в том случае, если выходные значения двух TPM равны.

Для достижения равенства весовнеобходимо выполнить следующие действия:

0. У двух абонентов есть многослойная сеть TPM.

1. Инициализация нейронной сети. В начале процесса синхронизации значения весовых коэффициентов задаются случайным образом.

2. Генерация случайного входного вектора, который подается на вход НС.

3. Вычисление значения скрытых и выходного нейронов.

4. Проверка равенства выходных значений обоих НС.

5. В случае несовпадения выходных значений необходимо вернуться ко второму шагу. Если выходные значения равны, тогда для каждой НС используем одно из трех правил обучений, представленных ниже.

6. После достижения полной синхронизации НС на основе полученных весовых коэффициентов генерируется ключ шифрования.

Для достижения полной синхронизации древовидных машин четности можно использовать три правила обучения:

Правило обучения Хебба – если сигнал перцептрона неверен и равен нулю, то необходимо увеличить веса тех входов, на которые была подана единица:



где wi,j – весовые значения скрытого слоя; xi,j – входные значения; τA – выходное значение TPMА; τB – выходное значение TPMВ; τ – общий выход.

Правило анти-Хеба – если сигнал перцептрона неверен и равен единице, то необходимо уменьшить веса тех входов, на которые была подана единица.



Случайное блуждание – это математическая модель, описывающая случайное движение частицы или объекта в пространстве. В случайном блуждании, на каждом шаге, объект перемещается в случайном направлении или случайно изменяет свою позицию в зависимости от вероятностей.



Представленные правила обучения должны гарантировать, нахождение весовых коэффициентов в допустимом диапазоне между –L и +L.

С учетом вышесказанного можно сделать вывод о том, что конечным результатом процесса синхронизации является получение одинаковых значений весовых коэффициентов двух НС.

Синхронизация двух TPM-машин может использоваться для обмена ключевой информацией, аналогично обмену ключами в алгоритме Диффи-Хеллмана.

Алгоритм синхронизации двух TPM-машин:

1.Сгенерировать для каждой машины вектор случайных весовых коэффициентов;

2. Сгенерировать случайных вектор из K\*N входных значений;

3. Вычислить значения на выходах скрытых нейронов;

4. Вычислить значения на выходах выходных нейронов;

5. Применить подходящее правило обучения;

6. Сравнить полученные выходные значения обоих машин: если значения равны, значит синхронизация достигнута; если значения отличаются, перейти к шагу.

**Ход работы**

Входе лабораторной работы было разработано приложение, реализующее 500 синхронизаций двух TPM-машин с заданными параметрами и применением правила обучения Хебба.

При выполнении скрипта main.py будет проведено *iterations* итераций эксперимента, и в конце будет выведена средняя статистика, такая как среднее количество обновлений и среднее время синхронизации. Также будет отображен график зависимости среднего количества обновлений от номера итерации.

Скрипт tmp.py содержит определение класса TPM. В конструкторе класса инициализируются параметры *k*, *n* и *l*, а также создается матрица весов *W* размерности [*k*, *n*], заполненная случайными целыми числами от *-l* до *l*. Метод get\_output принимает входные данные *X*, изменяет их форму и вычисляет выход *tau* по формулам, используя матрицу весов *W*. Метод \_\_call\_\_ является оберткой для метода get\_output и позволяет вызывать объект класса как функцию, передавая ему входные данные. Метод update обновляет матрицу весов *W* в зависимости от значения *tau1* и *tau2* с помощью одного из трех правил обновления: hebbian, anti\_hebbian или random\_walk.

Скрипт main.py содержит код для выполнения главной программы. Он импортирует класс TPM из скрипта tmp.py, определяет несколько вспомогательных функций, включая функцию random\_vector, которая генерирует случайный вектор входных данных, и функцию sync\_score, которая вычисляет показатель синхронизации между двумя экземплярами класса TPM. Затем определены параметры, такие как *k*, *n*, *l*, update\_rule, steps, times, updates, updates\_sum, updates\_avg\_vector и iterations. В цикле от 1 до *iterations* создаются два экземпляра класса TPM, инициализируются переменные, и происходит итеративное обновление и синхронизация двух экземпляров до достижения полной синхронизации. Замеряется время выполнения и количество обновлений для каждой итерации. Результаты сохраняются в соответствующих списковых переменных. В конце выводятся среднее количество обновлений и среднее время синхронизации для всех итераций, а также строится график зависимости среднего количества обновлений от номера итерации.

Скрипт update\_rules.py содержит реализацию трех различных правил обновления матрицы весов *W* в соответствии с TPM: hebbian, anti\_hebbian и random\_walk. Каждое из этих правил обновляет веса в зависимости от значений *tau1*, *tau2*, *sigma* и *X*, используя функцию *theta* для определения коэффициента обновления. Значения весов также ограничиваются в пределах от *-l* до *l*.

Таким образом, данный код реализует модель TPM и проводит эксперименты по синхронизации и обновлению матрицы весов двух экземпляров класса TPM. График зависимости среднего количества обновлений от номера итерации позволяет оценить эффективность синхронизации между двумя TPM.

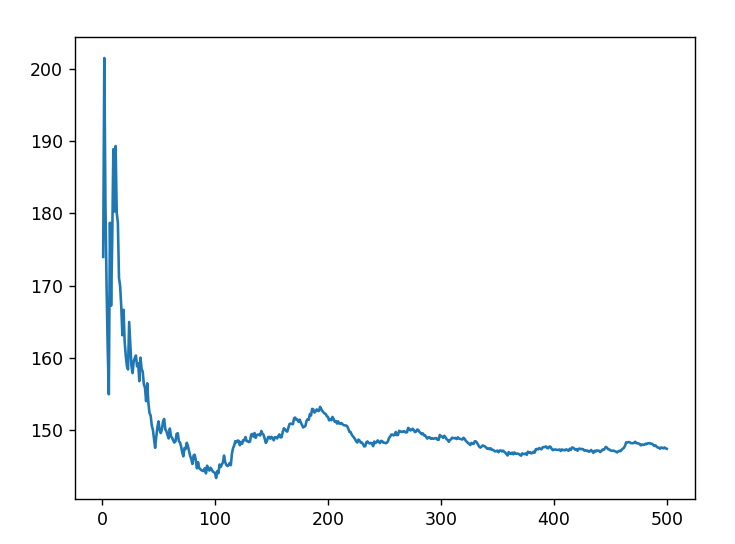


Рисунок 3 – График зависимости среднего числа шагов для достижения синхронизации от числа синхронизаций

В контексте данного кода, синхронизация означает достижение одинаковых состояний у двух экземпляров класса TPM. Когда два экземпляра TPM становятся синхронизированными, это означает, что их матрицы весов *W* становятся одинаковыми.

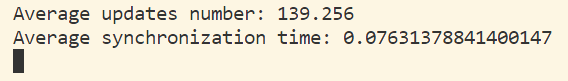


Рисунок 4 – Среднее число обновлений и среднее время синхронизации

В криптографии синхронизация нейронных сетей может использоваться для обеспечения безопасной коммуникации и шифрования данных. Например, можно использовать синхронизацию нейронных сетей для создания общего секретного ключа между двумя сторонами, который затем может быть использован для шифрования и дешифрования сообщений. При этом секретный ключ является результатом согласованности весов нейронных сетей и неизвестен злоумышленникам. Таким образом, синхронизация нейронных сетей может обеспечить надежность и безопасность в области криптографии.

**Ответы на вопросы**

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

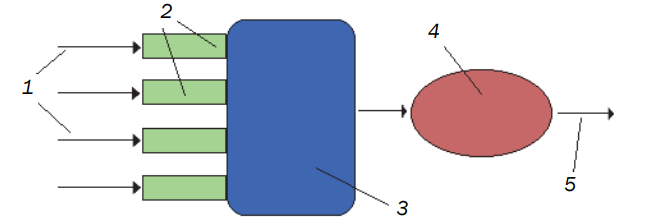


Рисунок 5 – Модель искусственного нейрона – персептрона: 1 – входы; 2 – веса; 3 – суммирующий блок; 4 – блок активации; 5 – выход

Главные элементы искусственного нейрона, являющиеся эквивалентами биологического нейрона:

Входные значения – дендриты;

Весовые коэффициенты – синапсы;

Суммирующий блок – ядро;

Функция активации – основание аксона;

А выходное значение – аксон.

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

Искусственные нейронные сети (ИНС) широко применяются для анализа и обработки больших объемов информации, включая задачи классификации и прогнозирования, а также для обучения и развития искусственного интеллекта. В криптографии некоторые методы, основанные на использовании ИНС, могут быть применены для согласования ключевой информации, обеспечивая безопасную передачу данных.

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

,

где S – выходной сигнал, – вес входного сигнала, – значение входного сигнала.

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

В криптографии можно применять определенные методы, основанные на использовании искусственных нейронных сетей (ИНС), с целью согласования ключевой информации. Один из примеров таких методов может быть связан с использованием 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.

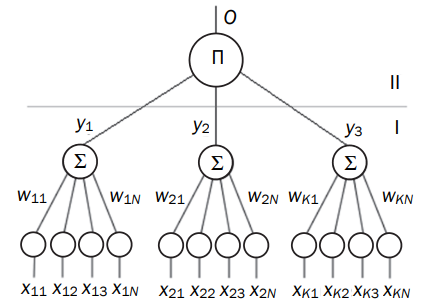


Рисунок 6 – Схема двух уровней модели TPM

Синхронизацию нейронных сетей (далее НС) можно представить следующим образом: первоначально на вход НС поступают общие входные данные, являющиеся случайным набором. Спустя короткое время достигается полная синхронизация, в результате которой весовые коэффициенты нейронных сетей оказываются равными. Полученные значения весовых коэффициентов в дальнейшем используются в качестве ключа шифрования.

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

Алгоритмы обучения искусственных нейронных сетей можно классифицировать на методы обучения с учителем и без учителя. В случае обучения без учителя используются только входные данные, и такие методы могут быть применены, например, для задач кластеризации или масштабирования входных данных.

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

1. Правило обучения Хебба: Этот алгоритм основан на идее, что связи между нейронами усиливаются, если активность этих нейронов совпадает во времени. В контексте согласования ключевой информации, нейроны, отвечающие за обработку и передачу информации, могут обучаться усиливать свои связи, если активируются одновременно.

2. Правило Анти-Хебба: в отличие от правила Хебба, это правило основано на уменьшении связей между нейронами, когда они активируются одновременно. Правило Анти-Хебба может быть полезным для согласования информации, позволяя снизить связь между нейронами, которые могут быть несовместимыми или конфликтными.

3. Обучение на основе случайного блуждания: Этот алгоритм основан на идее случайного блуждания, где информация передается между сторонами на основе случайно выбранных путей. В контексте согласования ключевой информации, ИНС может обучаться на основе случайных взаимодействий и выбирать наиболее эффективные пути передачи информации.

Эти алгоритмы могут быть комбинированы или адаптированы в зависимости от конкретных требований и задач согласования ключевой информации между двумя сторонами.

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

Могут. 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.**

Системы, использующие случайные векторы входных значений для синхронизации, обладают высокой степенью криптостойкости. Они должны быть достаточно быстрыми, чтобы предотвратить возможность атакующей нейронной сети успеть синхронизироваться с одним из узлов. При равных вычислительных мощностях атакующая нейронная сеть не сможет успеть синхронизироваться, так как ей, как правило, неизвестны сгенерированные случайные значения входных параметров.

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

Атаки перебора (Brute Force Attacks): это тип атаки, при котором злоумышленник пытается перебрать все возможные ключи или параметры модели, чтобы обнаружить или подобрать правильные значения, несмотря на использование нейронных сетей и криптографических методов.

Атаки посредника (Man-in-the-Middle Attacks): в таких атаках злоумышленник позиционируется между отправителем и получателем информации, перехватывает и изменяет данные, передаваемые между ними. Это позволяет злоумышленнику обойти криптографическую защиту и получить доступ к конфиденциальной информации.

**Вывод:** в ходе лабораторной работы были изучены основы построения и функционирования искусственных нейронных сетей (ИНС), а также использования ИНС в криптографии. Данный алгоритм может быть взломан при помощи генетической атаки.