УДК 004.891.3

На правах рукописи

ЛИ АЛЕКСАНДР АЛЕКСАНДРОВИЧ

Метод обнаружения вторжений в информационную систему на основе нейронных сетей

05.13.01 – Системный анализ, управления и обработка информации

Автореферат

диссертации на соискание ученой степени
кандидата технических наук

Республика Казахстан
Алматы, 2010

Работа выполнена в Казахском национальном техническом университете имени К. И. Сатпаева

Научный руководитель:	кандидат технических наук Ермаков А. С.
Официальные оппоненты:	доктор технических наук Сыздыков Д. Ж.
	кандидат технических наук Буранбаева А. И.
Ведущая организация:	Институт математики Министерства образования и науки Республики Казахстан

Защита состоится 28 октября 2010 г. в 14-30 ч. на заседании объединенного диссертационного совета ОД 14.13.03 при Казахском национальном техническом университете имени К.И. Сатпаева по адресу: Республика Казахстан, 050013, г. Алматы, ул. Сатпаева, 22, нефтяной корпус, конференц-зал.

С диссертацией можно ознакомиться в библиотеке Казахского Национального Университета имени К. И. Сатпаева.

Автореферат разослан «_____»_____2010 г.

Ученый секретарь
диссертационного совета,
доктор технических наук  Б.Х. Айтчанов

ВВЕДЕНИЕ

Актуальность темы исследования. Эффективность функционирования современных информационных систем в значительной мере связана с проблемой защиты обрабатываемой в них информации. Сложность процессов обработки информации, сопутствующая современным системам, приводит к появлению большого числа ошибок в программном коде информационной системы. Данные ошибки способствуют появлению уязвимостей в программном коде, а, следовательно, разнообразных атак, то есть способов получения несанкционированного доступа к ресурсам информационной системы.

Анализ существующих систем защиты информации показывает, что их возможности не позволяют обеспечить безопасность информационной системы на достаточном уровне. Причиной этого является то, что процесс создания систем обнаружения атак сопряжен с рядом нерешенных научно – технических задач. Существующие системы обнаружения атак используют простейшие алгоритмы обработки поступающей информации, что не позволяет обнаружить значительное количество атак на информационные системы.

Цель и задачи исследования. Целью диссертационной работы является разработка теоретических основ и практических подходов к созданию нейросетевой системы обнаружения атак на информационную систему, а также создание исследовательского прототипа нейросетевой системы обнаружения атак.

В процессе работы над диссертацией решались следующие задачи:

1. Провести исследования и разработать прототип нейросетевой системы обнаружения атаки на ИС.
2. Разработать метод выявления нарушений безопасности.
3. Разработать рекомендации по применению разработанного метода в системе обнаружения вторжений.
4. Создание программного модуля, разработанного на полученных рекомендациях с целью апробации результатов.
5. Проведения анализа полученных результатов.

В работе использовались методы системного анализа и информатики, теории принятия решений, теории распознавания образов, теории нейронных сетей и нечеткой логики, а также теории математической статистики. Широко использовалось моделирование на персональных ЭВМ.

Теоретическую и методологическую основу исследования составляют труды отечественных и зарубежных ученых, связанных с проблемами защиты и безопасности информации, а также материалы различных научных конференций, отчетов, авторские публикации.

Научная новизна диссертационной работы заключается в следующем:

1. Предложен и обоснован метод обнаружения атак, основанный на комбинированном применении методов поиска сигнатуры атаки и обнаружения аномалий в работе пользователя, позволяющий существенно улучшить характеристики обнаружения атак.

2. Предложен новый поход к обработке поступающих в информационную систему данных, заключающийся в представлении их в виде сигнатур, что позволяет использовать преимущества аппарата нейронных сетей для решения задачи распознавания атак.

3. На основании проведенных исследований показана возможность применения нейронной сети для обработки входных данных с целью достижения высокой эффективности обнаружения атак.

4. Предложены рекомендации применения разработанного метода в системах обнаружения вторжений.

Основные положения, выносимые на защиту:

1. Подход к построению системы обнаружения атак, основанный на комплексировании существующих методов обнаружения атак.

2. Модель безопасной работы пользователя информационной системы.

3. Подход к обработке информации, основанный на применении математического аппарата нейронных сетей к решению задач распознавания сигнатур атаки и безопасной работы пользователя.

4. Результаты исследований по формированию репрезентативного обучающего множества и поиску оптимальных параметров обучения нейронной сети.

5. Структура и средства программной реализации исследовательского прототипа нейросетевой системы обнаружения атак.

Теоретическая и практическая значимость диссертационного исследования:

1. Предложенный метод обнаружения атак может быть использован для построения систем обнаружения атак, превосходящим по своим характеристикам большинство коммерческих систем.

2. Предложенный подход к формированию обучающей выборки нейронной сети позволяет достичь высоких характеристик системы обнаружения атак за счет обеспечения требования репрезентативности выборки.

3. Выбранный подход, основанный на применении комбинированного метода обнаружения атак, позволяет упростить реализацию системы обнаружения атак.

Апробация результатов исследования. Практическая ценность и новизна работы, в виде метода обнаружения атак и разработки системы обнаружения вторжений, подтверждаются актами внедрения от ТОО «ТекомСервис» и ТОО «VEGA Partners».

Публикации. Результаты работы опубликованы в 4 печатных трудах и 3 трудах конференций.

Структура диссертационной работы. Диссертация состоит из введения, трех разделов, заключения, списка использованных источников, приложения.

ОСНОВНАЯ ЧАСТЬ

Рассматривается структура современных систем обнаружения вторжений (СОВ). Характеризуются основные направления распознавания нарушений безопасности защищаемых систем в современных СОВ. Выполнен анализ используемых методов и моделей структуры СОВ в соответствии с выделенными основными группами. Приведены основные недостатки существующих СОВ и обоснованы направления их совершенствования.

Системы обнаружения вторжения (СОВ) – это системы, собирающие информацию из различных точек защищаемой компьютерной системы (вычислительной сети) и анализирующие эту информацию для выявления как попыток нарушения, так и реальных нарушений защиты (вторжений).

Недостатки существующих систем обнаружения

Недостатки современных систем обнаружения можно разделить на две группы – недостатки, связанные со структурой СОВ, и недостатки, относящиеся к реализованным методам обнаружения (см. таблицу 1).

Таблица 1. Недостатки современных СОВ

Недостаток	Описание
Отсутствие общей методологии построения	Частично это можно объяснить недостаточностью общих соглашений в терминологии, так как СОВ – это достаточно новое направление, основанное Андерсоном в 1980 г.
Эффективность	Часто методы системы пытаются обнаружить любую понятную атаку, что приводит к ряду неудовлетворительных последствий.
Переносимость	До сих пор большинство СОВ создается для использования на конкретном оборудовании, и достаточно трудно использовать их в другой системе, где требуется реализовать похожую политику безопасности.
Возможности обновления	Очень сложно обновить существующие системы новыми технологиями обнаружения.
Производительность	Трудно оценить производительности СОВ в реальных условиях.

Методы обнаружений вторжений

В настоящее время для построения систем обнаружения атак используются, в основном, два метода:

– метод обнаружения атак, основанный на поиске сигнатуры атаки;

– метод обнаружения атак, основанный на поиске аномалий.

Под сигнатурой атаки понимается некоторый шаблон, который однозначно соответствует данной атаке. Шаблон атаки может быть представлен в

различных формах, например, в виде символьной строки, regexp–шаблона, граничной частоты повторения определенного события и т.д.

Аномалией считается любое отклонение от нормальной (допустимой) модели действий пользователя.

Существующим метода обнаружения вторжений присущи следующие недостатки:

- недопустимо высокий уровень ложных срабатываний и пропусков атак;
- слабые возможности по обнаружению новых атак;
- большинство вторжений невозможно определить на начальных этапах;
- трудно, иногда невозможно, определить атакующего, цели атаки;
- отсутствие оценок точности и адекватности результатов работы;
- невозможно определять «старые» атаки, использующие новые стратегии;
- сложность обнаружения вторжений в реальном времени с требуемой полнотой в высокоскоростных сетях;
- слабые возможности по автоматическому обнаружению сложных координированных атак;
- значительная перегрузка систем, в которых функционируют СОВ, при работе в реальном времени;

На основе изложенного можно сделать вывод о том, что в практической деятельности накоплен значительный опыт решения проблем обнаружения вторжений. Применяемые СОВ в значительной степени основаны на эмпирических схемах процесса обнаружения вторжений, дальнейшее совершенствование СОВ связано с конкретизацией методов синтеза и анализа сложных систем, теории распознавания образов в применении к СОВ.

Комбинированный метод

Как было показано ранее, способом, гарантирующим обнаружение атаки в 100% случаев, является метод поиска аномалий. Поскольку естественным желанием разработчика является создание высокоэффективной системы обнаружения атак, то напрашивается вывод о том, что все современные системы обнаружения атак должны быть основаны на методе поиска аномалий в работе пользователя. Однако не все так просто. Практически, каждое внедрение даже одного и того же программного комплекса уникально. Это означает, что в каждом конкретном случае безопасная модель действий пользователя также уникальна.

Согласно законам рынка, стоимость создания программного продукта (в нашем случае, это система обнаружения атак), должна распределяться на число внедрений разработанного продукта. То есть, чем больше число внедрений продукта в информационные системы заказчика, тем дешевле она обойдется каждому из заказчиков. Если же разработчик создает уникальную систему об-

наружения атак, которая полностью будет соответствовать модели действий конкретного пользователя, но она будет неприменима в остальных случаях, и, следовательно, число ее внедрений будет единичным. Таким образом, стоимость ее внедрения будет значительно выше стоимости внедрения универсальной системы.

Согласно общей концепции внедрения систем обеспечения безопасности, стоимость средств защиты какой–либо информации не должна превышать стоимости ущерба, который может возникнуть в результате ее несанкционированного использования. Таким образом, получается, что разработка уникальной системы обнаружения атак может быть экономически оправданной только в исключительных случаях, требующих гарантированного обеспечения безопасности информации, имеющей высокую стоимость.

Упрощение систем обнаружения атак, а, следовательно, снижение стоимости их реализации, как было показано выше, может быть достигнуто за счет упрощения формализации сигнатур атак (при применении метода поиска сигнатур) либо за счет упрощения формализации безопасной модели действий пользователя (при применении метода обнаружения аномалий). Но следствием данного упрощения, неизменно является ухудшение характеристик обнаружения атак.

Можно сделать вывод о том, что комбинированное использование различных методов обнаружения атак, как в случае добавления уточняющих сигнатур к методу поиска аномалий, так и при добавлении уточняющих шаблонов «безопасной» модели пользователя к сигнатурному методу, обеспечивает улучшение характеристик систем обнаружения атак.

Поэтому для разработки метода необходимо использовать именно комбинированный метод обнаружения атак, состоящий из сигнатурного метода с добавлением уточняющих шаблонов «безопасной» модели поведения пользователя**.**

Преимущества нейронных сетей

Любая нейронная сеть используется в качестве самостоятельной системы представления знаний, которая в практических приложения выступает, как правило, в качестве одного из компонентов системы управления либо модуля принятия решений, передающих результирующий сигнал на другие элементы, не связанные непосредственно с искусственной нейронной сетью (см. таблицу 2).

Таблица 2. Преимущества нейронных сетей

Преимущество	**Описание**
Возможность обучения	Натренированная на ограниченном множестве обучающих выборок, она обобщает накопленную информацию и вырабатывает ожидаемую реакцию применительно к данным, не обрабатывавшимся в процессе обучения.
Параллельная обра-	Важнейшее свойство нейронных сетей, свидетельст-

ботка данных	вующее об их огромном потенциале и широких прикладных возможностях, состоит в параллельной обработке информации одновременно всеми нейронами. Благодаря этой способности при большом количестве межнейронных связей достигается значительное ускорение процесса обработки информации. Во многих ситуациях становится возможной обработка сигналов в реальном масштабе времени.
Классификация образов	В задачах классификации образов сеть обучается таким важнейшим их признакам. В процессе обучения выделяются признаки, отличающие образы друг от друга, которые и составляют базу для принятия решений об отнесении образов к соответствующим классам.
Прогнозирование	При решении задач прогнозирования роль нейронной сети состоит в предсказании будущей реакции системы по ее предшествующему поведению.

Подводя итог, можно утверждать, что нейросетевая система в состоянии не только выполнять набор функций, которые реализованы в классических системах обнаружения атак, но и выполнять их со значительно лучшими характеристиками, что позволит устранить ряд проблем, присущих существующим классическим системам обнаружения атак. Следовательно, проведение исследований в области анализа возможностей применения нейросетей для обнаружения атак является актуальным для современной науки и техники.

Описание метода

Разрабатываемый метод основывается на выполнении следующих подзадач:

1. Преобразование входящего запроса в сигнатуру;
2. Поиск сигнатуры в базе данных;
3. Обнаружения аномалий;
4. Противодействие обнаруженной атаке;
5. Расширение базы данных сигнатур.

На рисунке 1 представлен общий алгоритм разрабатываемого метода.

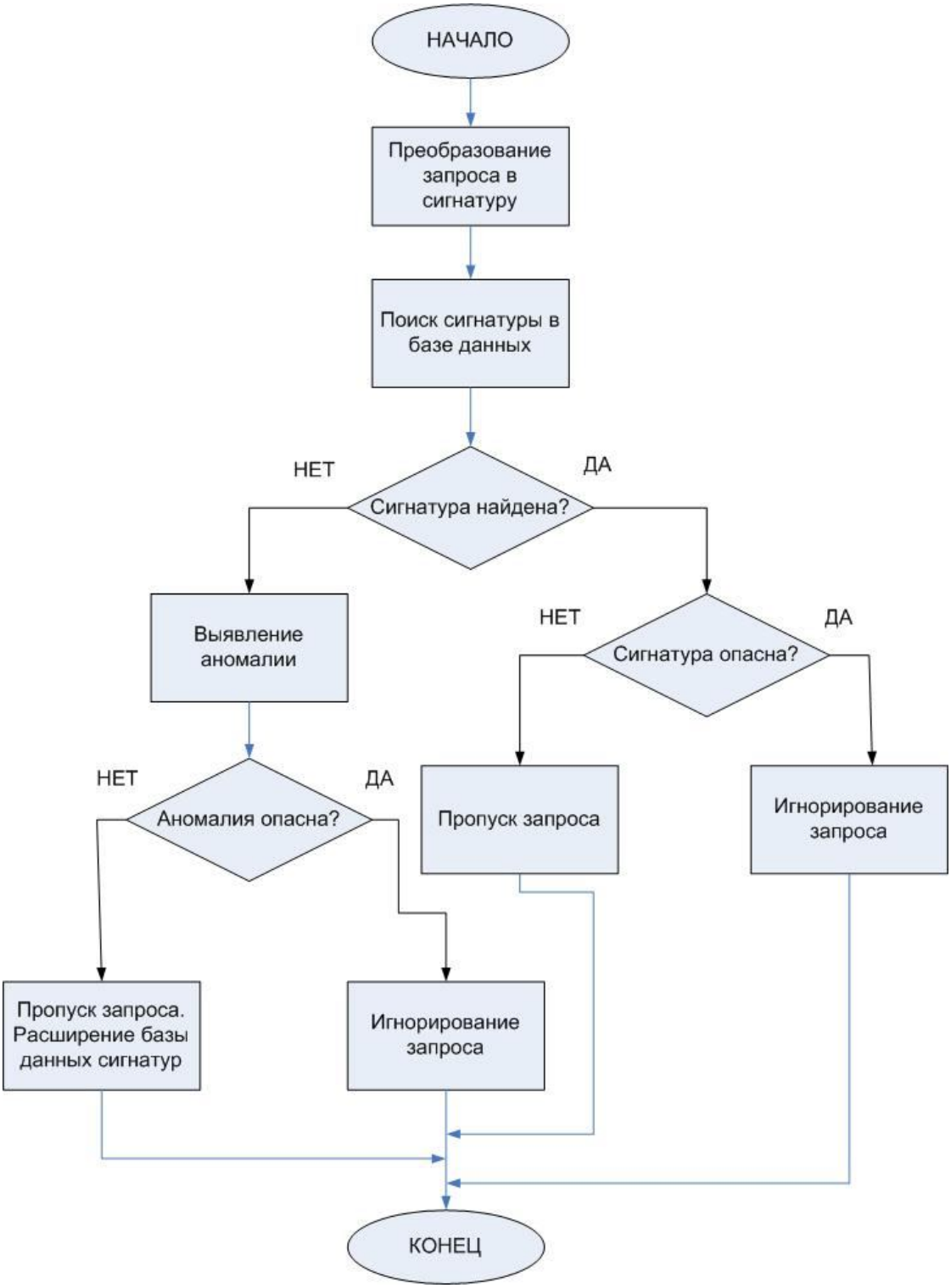

Рисунок 1 – Блок-схема алгоритма

Рассмотрим подробнее каждую подзадачу.

Преобразование входных данных в сигнатуру

Сигнатура атаки – характерные признаки атаки или вируса, используемые для их обнаружения. Большинство современных антивирусов, сканеров уязвимостей и систем обнаружения вторжений используют синтаксические сигнатуры, взятые непосредственно из тела атаки (файла вируса или сетевого пакета). Также существуют сигнатуры, основанные на поведении или аномали-

ях – например, слишком агрессивное обращение к какому-либо сетевому порту на компьютере.

Поиск сигнатуры в базе данных

На сегодняшний день существуют несколько наиболее продуктивных алгоритмов поиска сигнатур атак (см. таблицу 3). Это следующие алгоритмы:

Таблица 3. Виды поисков сигнатур

Виды поисков	**Описание**
Существование	распознавание факта является основанием для регистрирования попытки атаки.
Последовательность	распознавание строго определенной последовательности событий достаточно для обнаружения атаки.
Частичный порядок	поводом для сигнализации атаки служит распознавание сигнатуры, состоящей из частично упорядоченных событий.
Интервал времени	учитываются временные соотношения между событиями.
Период	учитываются непосредственно моменты времени, в которые происходят события.

В настоящее время для реализации сигнатурного метода используют в основном первые два способа.

Под **поиском** механизм нейросетей подразумевает результат обработки входящих данных и сравнение его с эталонным результатом, при полном совпадении результатов, сигнатура в базе данных существует, в противном случае возникает ситуация «неизвестности» сигнатуры.

Обнаружение аномалий

В случае необнаружения сигнатуры в базе данных, осуществляется проверка входящей сигнатуры на аномалию. Для этого необходимо найти «опасную» сигнатуру наиболее близкую к оцениваемой, при помощи меры близости. Чем меньше мера близости, тем опаснее аномалия. В случае если оценка равна нулю, можно говорить о полном совпадении и соответственно о наибольшей опасности. На рисунке 6 представлена классификация выходных данных разрабатываемого прототипа.

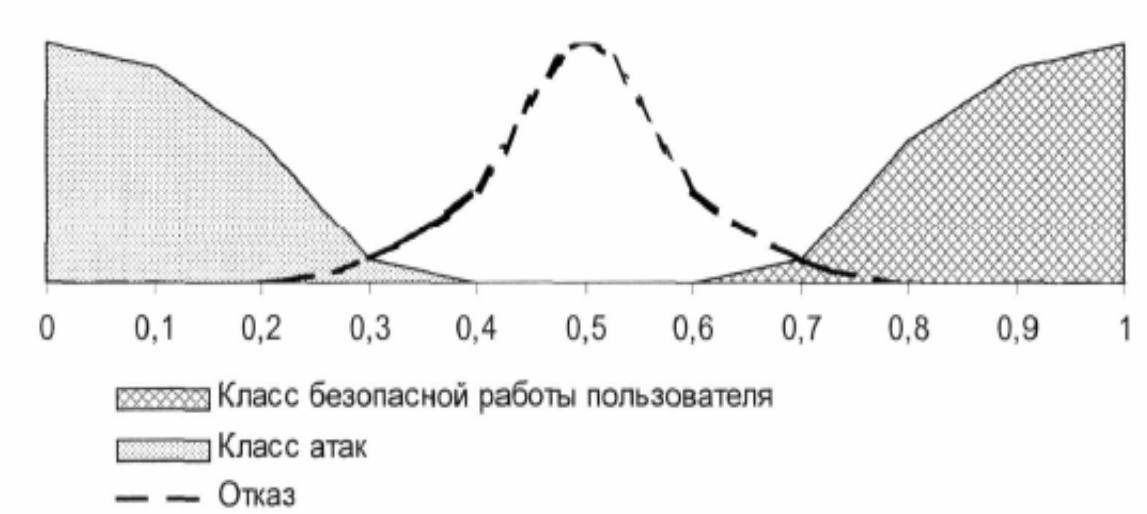

Рисунок 2 – Классификация значений на выходе нейросети

Так как выходное значение нейронной сети после обработки входных данных представляет собой непрерывную величину в диапазоне от 0 до 1, необходимо предусмотреть механизм, который сопоставлял данное значение одному из 3 классов значений - классу атаки, классу безопасного действия пользователя и классу отказа нейронной сети. Под отказом нейронной сети будем считать ситуацию, когда нейронная сеть выдает ложный результат, т.е. может иметь место, как ошибки первого, так и второго родов.

Выступая в качестве эксперта, предположим, что нейронная сеть относит сигнатуру к классу атаки, если выходное значение находится в пределах от 0 до 0.3 включительно; к образу безопасного действия пользователя, если оно в пределах от 0.7 до 1 включительно; и к классу отказа нейронной сети при остальных значениях.

С учетом постановки задачи исследования, нам необходимо получить наименьшее значение ошибки первого рода. Поэтому, значение сигнала на выходе из множества «отказа» нейронной сети мы будем также относить к классу атаки. Таким образом, в случае, если на выходе нейронной сети присутствует значение от 0.7 до 1 включительно, то считается, что в информационную систему поступают безопасные данные, а при всех остальных значениях — на информационную систему осуществляется атака.

Исходя из особенности формирования структуры нейронной сети, поиск значений параметров обучения, а также оптимальной структуры нейронной сети, которая сможет классифицировать сигнатуру как атака либо безопасные действия пользователя, будем осуществлять экспериментальным путем.

Противодействия обнаруженным атакам

Основной целью создания системы защиты информации является обеспечение информационной безопасности информационной системы. Оно не может состоять только из пассивной стороны, то есть обнаружения попыток несанкционированного доступа. Важна именно активная сторона защиты – вы-

полнение комплекса мер, направленных на снижение последствий от действия атаки.

Под противодействием будем понимать реакцию системы защиты информации, которая может включать в себя следующие действия:

– сигнализацию о несанкционированном доступе;

– блокировку (отключение терминала, группы терминалов, элементов вычислительной сети и т.д.);

– отказ в запросе.

Подведя итоги вышесказанному, можно сделать вывод о том, что автоматизация процесса выбора необходимых действий в случае атаки должна быть основана на формализации связей «атака – противодействие ей». В процессе выбора должно учитываться, что ущерб от противодействия атаке должен быть не более того ущерба, который будет получен от взлома системы в результате отсутствия противодействия атаке.

Разрабатываемая система будет использовать пассивный метод противодействия атаки, то есть игнорирование выявленного «опасного» пакета и занесения его сигнатуры в базу данных.

Расширение базы данных сигнатур

Для эффективного функционирования СОВ необходимо постоянно пополнять базу данных сигнатур новыми данными, иначе через определенное время система столкнется с неизвестностью новых атак. Все вышесказанное более наглядно представлено на рисунке 4.

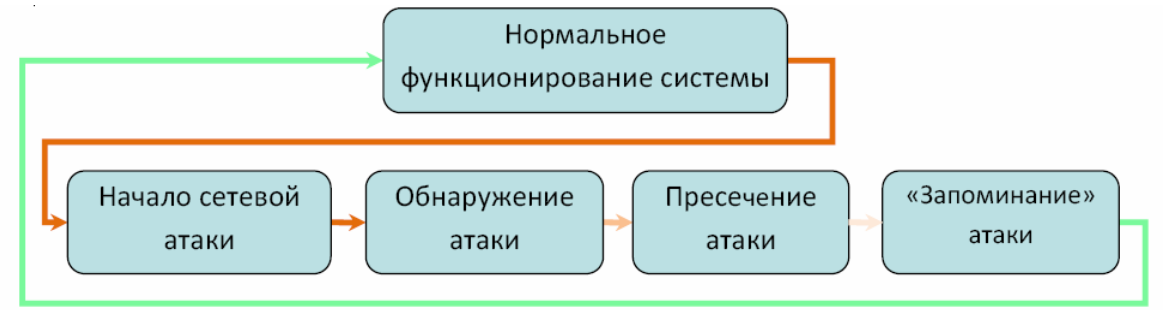

Рисунок 4 – Алгоритм расширения базы данных сигнатур

Реализация данной функции не представляется сложной, так как одной из основной функций математического аппарата нейросети является, именно, обучение. При появлении неизвестной сигнатуры достаточно переключить нейросеть в режим обучения и вновь подать на вход обнаруженную неизвестную сигнатуру. Необходимо ввести критерий оценки «опасности», иначе со временем база данных сигнатур будет содержать в себе всевозможные сигнатуры, что приведет к ее расширению и, как следствие, к резкому снижению производительности.

Расчеты

Рассчитаем производительность разрабатываемой нейросети при различных параметрах, зависящих от:

– от размера обучающей выборки;

– от размера входящих данных;

– от количества слоев.

Рассмотрим размер обучающей выборки.

В данном эксперименте количество входящих нейронов было равно 51, изменялся размер обучающей выборки. Результаты сведены в таблицу 4.

Таблица 4 – Время, потраченное на обучения

Размер выборки	Шаг	Средняя оценка	Время, с
100	0.004924354	1.570827E-15	2.53
500	0.0005900296	1.763657E-15	27.10
1000	0.00243	1.100726E-11	70.9

При увеличении выборки в 2 раза, время обучения вырастает примерно в 4 раза, это происходит из-за алгоритма обучения, который учитывает результаты предыдущих эпох обучения.

Рассмотрим размер входящих данных.

В данном эксперименте изменялось количество входящих нейронов и соответственно количество нейронов скрытого слоя, размер обучающей выборки остался на уровне 500. Результаты сведены в таблицу 5.

Таблица 5 – Время, потраченное на обучение

Количество нейронов (количество нейронов скрытого слоя)	Шаг	Средняя оценка	Время, с
32 (65)	0.00195070	5.579407E-17	11.75
51 (103)	0.00059002	1.763657E-15	27.10
83 (167)	0.00045698	2.785903E-15	83.60

При увеличении количества нейронов время обучения резко увеличивается, это происходит из-за увеличения межнейронных связей.

Рассмотрим количество внутренних слоев.

В данном эксперименте количество входящих нейронов было равно 51, количество обучающей выборки 500, изменялось количество внутренних слоев. Результаты сведены в таблицу 6.

Таблица 6 – Время, потраченное на обучение

Количество внутренних слоев	Шаг	Средняя оценка	Время, с
1	0.0005900296	1.763657E-15	27.10
2	0.0003965783	8.131954E-15	25.50
3	0.0007476374	1.753376E-14	22.93

Можно сделать вывод, что при увеличении скрытых слоев и распределении нейронов по добавленным слоям количество связей между нейронами уменьшается, поэтому время обучения меньше, но из-за уменьшения межнейронных связей точность обучения также уменьшается.

Сравнение нейросетевой и последовательной реализаций решения систем линейных уравнений

В данном параграфе рассмотрены вопросы сопоставления и сравнения нейросетевой и последовательной реализаций решения систем линейных алгебраических уравнений и даны сравнительные характеристики обоих вариантов реализации.

Метод Гаусса – широко известный прямой алгоритм решения систем линейных уравнений, для которых матрицы коэффициентов являются плотными.

Несложно показать, что время выполнения алгоритма составляет:

$$T_{общ} = (2 * N^3 / 3 + N^2) * t_n, \quad (3.5)$$

где t_n – время выполнения одной условной операции последовательного алгоритма, N – размерность матрицы.

Для нашей системы в качестве данных для обучения служат матрицы коэффициентов СЛАУ и матрицы корней уравнений. А наиболее удобным вариантом топологии сети являются самоорганизующиеся карты Кохонена.

Число M в блок–схеме – число слоев в получившейся сети. Используя для моделирования программное обеспечение Matlab® Version 6.1.0.450 Release 12.1 с пакетом расширения Neural Networks Toolbox и работая с топологией «карты Кохонена» количество слоев для матрицы любой размерности равно 2.

Также найдем количество нейронов в полученной сети. Количество нейронов во всех слоях, кроме последнего, равно количеству элементов входной матрицы, т.е. $N * (N + 1)$, количество нейронов в последнем слое равно размерности выходного вектора N. Получим, что

$$K = N * (N + 1) + N = N^2 + 2N \ , \quad (3.6)$$

где K – количество нейронов

Вычислим время выполнения нейросетевого варианта. Общее время работы нейросети состоит из времени обучения и времени решения задачи, причем время обучения равно времени решения, умноженного на количество учебных выборок. Для достижения должного уровня обучения, количество учебных выборок рекомендуется брать на **порядок** больше количества нейронов.

$$T_{общ} = T_{обуч} + T_{решения} \quad (3.7)$$

$$T_{обуч} = T_{решения} * 10 * K \quad (3.8)$$

Текущее состояние нейрона определяется, как взвешенная сумма его входов:

$$s = \sum_{i=1}^{n} x_i w_i \,, \tag{3.9}$$

где x – входной сигнал, w – вес синапса, n – количество входов.

Из (3.5) следует, что время выполнения работы каждого нейрона равно:

$$T_{нейрона} = N * t_н \tag{3.10}$$

$$T_{решения} = K * T_{нейрона} = (N^2 + 2N) * N * t_н = (N^3 + 2N^2) * t_н \tag{3.11}$$

где $t_н$ – время выполнения одной условной операции нейросетевого алгоритма.

Полученная нейросеть, в отличие от последовательного алгоритма, хорошо распараллеливается, т.к. формирование выходных сигналов каждого нейрона происходит независимо от других нейронов текущего слоя. В данном случае сложность параллельного алгоритма будет иметь порядок N^3 / p, но в отличие от последовательного алгоритма здесь $p \le (N^2 + N)$

Последовательный алгоритм был реализован на встроенном языке программирования в системе Matlab® 6.12.

Результаты вычислительных экспериментов для последовательного метода реализации сведены в таблицу 7.

Таблица 7. Сравнение нейросетевой и последовательной реализаций

Размер матрицы	*T* выполнения последовательного алгоритма, сек	*T* выполнения нейросетевого алгоритма, сек	Ускорение
10	0.0160	0.0150	1.0666
100	1.8750	1.6215	1.1563
200	13.2810	11.1184	1.1945
300	44.4070	35.9545	1.2350
400	103.4070	86.6594	1.1932

Из сравнения видно, что нейросетевой вариант реализации решения СЛАУ методом Гаусса дает 1,2–кратный выигрыш по времени.

В диссертационной работе поставлена и решена задача создания теоретических и практических основ построения систем обнаружения атак на сервер на основе применения математического аппарата нейронных сетей.

ЗАКЛЮЧЕНИЕ

К основным выводам и результатам исследований можно отнести следующее:

1. Подавляющее большинство современных коммерческих систем обнаружения атак на сервер не способны обеспечить их эффективную защиту, что требует применения принципиально новых подходов к обработке информации.

2. С применением системного подхода проведена формализация задачи обнаружения атак, произведена классификация атак на информационные сис-

темы, анализ возможных уязвимостей информационных систем, а также возможных мер противодействия атакам, позволившие выявить наиболее эффективные подходы к решению поставленной задачи. Предложен и обоснован метод обнаружения атак, основанный на комбинированном применении методов поиска сигнатуры атаки и обнаружения аномалий в работе пользователя, позволяющий существенно улучшить характеристики обнаружения атак.

3. На основе модели безопасной работы пользователя в информационной системе и предложенного подхода к упрощению задачи обработки информации, синтезирована структура нейросетевой системы обнаружения атак.

4. Предложен подход к решению задачи классификации образов, заключающийся в представлении входных данных в виде сигнатур и отнесения их с использованием нейронной сети к классам атаки либо безопасным действиям пользователя.

5. Проведены исследования по определению оптимальных параметров алгоритмов обучения нейронной сети, включающие в себя:

– выбор методов формирования репрезентативных множеств, используемых в процессе обучения нейронной сети, а также методов оценки качества ее функционирования;

– поиск оптимальных значений параметров обучения нейронной сети.

Показано, что оптимизация параметров обучения сети снижает величину ошибки обнаружения атак.

6. Результаты сравнительной оценки характеристик разработанного исследовательского прототипа нейросетевой системы обнаружения атак свидетельствуют о том, что:

– ошибки обнаружения атак, соответствующие разработанному прототипу, в 4–5 раз меньше аналогичных значений для современных систем обнаружения атак;

– внедрение созданного прототипа позволит снизить финансовый ущерб от возможных атак.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Ермаков А. С., Ли А. А. Вопросы нейросетевой реализации решения систем линейных алгебраических уравнений // Тезисы докладов международной научной конференции «Ломоносов 2008»: –Астана: Казахстанский филиал МГУ им. М.В. Ломоносова; 2008, I часть, С. 108–110;

2. Ермаков А. С., Ли А. А., Алимсеитова Ж. К. Сигнатурный метод обнаружения вторжений на основе метода Гаусса // Материалы IV Международной научно–методической конференции «Математическое моделирование и информационные технологии в образовании и науке» (ММ ИТОН): посвященной 80–летию КазНПУ имени Абая – Алматы, КазНПУ им. Абая, 2008, Т1, С. 51–54;

3. Ли А. А. Анализ последовательной и нейросетевой реализаций решения систем линейных уравнений // Вестник КазНТУ имени К.И. Сатпаева. – 2008. –№5 (68) 2008. С. 156–169;

4. Ли А. А. Использование нейросетевых технологий в проблеме создания систем обнаружения вторжений // Вестник КазНТУ имени К.И. Сатпаева. –2008. –№6 (69) 2008. С. 154–157;

5. Ли А. А. Объединение нейронных сетей и генетических алгоритмов // Вестник КазНТУ имени К.И. Сатпаева. –2008. –№6/1 (70) 2008. С. 102–104. ISSN 1680–9211;

6. Ли А. А., Склонина А. С. Возможность применения нейросетей в системах обнаружения вторжений // Труды III Международной научно–практической конференции молодых ученых «ЖАС ГАЛЫМ – 2009». В 15 томах. Том 11. – Тараз, Таразский государственный педагогический институт, 16–18 апреля 2009 г., С. 153–156.

7. А. Ермаков, А. Ли Выявление сетевых атак на основе нейросетевого анализа//Ізденіс - Поиск. –2010. –№2 2010. С. 229–232.

Александр Александрович Ли

Нейрондық желілер негізінде ақпараттық желіге қол сұғуларды табу әдісі

ТҮЙІН

05.13.01 – Ақпаратты жүйелі талдау, басқару және өңдеу

Зерттеу саласы – қол сұғуларды табу жүйесі, нейронды желілер.

Диссертациялық жұмыстың мақсаты ақпараттық жүйеге жасалатын шабуылдарды нейрожелілік табу жүйесін жасаудың теориялық негіздері мен іс жүзіндегі қолдану жолдарын әзірлеу, сондай-ақ шабуылдарды нейрожелілік табу жүйесінің зерттеу прототипін жасау болып табылады.

Зерттеу әдістері. Осы еңбекте жүйелі талдау және информатика әдістері, шешімдерді қабылдау теориясы, бейнелерді тану теориясы, нейронды желілер мен дәл емес логика теориясы, сондай-ақ математикалық статистика теориясы қолданылды. Дербес ЭЕМ-де моделдеу әдісі кеңінен қолданылды.

Зерттеудің теориялық және іс жүзіндегі маңыздылығын отандық және шетелдік ғалымдардың ақпаратты қорғау және қауіпсіздігінің проблесмасына байланысты еңбектері, сондай-ақ әртүрлі ғылыми конференциялардың, есептердің авторлық жарияланымдардың материалдары құрайды.

Еңбектің ғылыми жаңашылдығы:

1. Шабуылдардың сигнатурасын іздеу және пайдаланушының жұмысында шабуылдарды табу сипаттарын едәуір жақсартуға мүмкіндік беретін ауытқуларды табу әдістерін құрамдастырып қолдануға негізделген табу әдістері ұсынылды және негізделіп түсіндірілді.

2. Ақпараттық жүйеге келіп түсетін сигнатуралар түрінде ұсынылатын деректерді өңдеуге жаңа амал ұсынылды, бұл нейронды желілердің аппаратының артықшылықтарын шабуылдарды тану мәселелерін шешу үшін пайдалануға мүмкіндік береді.

3. Жүргізілген зерттеулердің негізінде нейронды желіні шабуылдарды табудың жоғарғы тиімділігіне қол жеткізу мақсатында кіретін деректерді өңдеу үшін қолдану мүмкіндігі көрсетілген.

4. Шабуылдарды табу жүйелерінде әзірленген әдістерді қолдану ұсыныстары берілген.

Диссертациялық зерттеу жұмысының тиеориялық және пратикалық маңыздылығы:

1. Ұсынылған шабуылдарды табу әдісі өзінің сипаттамалары бойынша көптеген коммерциялық жүйелерден асып түсетін шабуылдарды табу жүйелерін қолдану үшін пайдаланылуы мүмкін.

2. Нейронды желіні іріктеуге үйретуды қалыптастыруға қолданылатын ұсынылған тәсіл арқылы іріктеудің қайта көрсетушілік қасиетінің талаптарын

қамтамасыз ету есебінен шабуылдарды табу жүйесінің жоғарғы сипаттарына қол жеткізуге болады.

3. Шабуылдарды табудың құрамдасқан әдісін қолдануға негізделіп, таңдалған тәсілі шабуылдарды табу жүйесін оңай іске асыруға мүмкіндік береді.

Alexandr Alexandrovich Li

Method of information system intrusion detection based on neural networks

SUMMARY

05.13.01 – Systems analysis, information management and handling

Research area – intrusion detection systems, neural networks

The purpose of the dissertation work is to develop theoretical bases and practical approaches to creating a neural network system for detecting information system attacks and to create a research prototype of a neural network-based system for attack detection.

Research methods. Methods of systems analysis and information science, decision-making theories, pattern recognition theories, neural network and fuzzy logic theories and theories of mathematical statistics were used in the work. Modeling on a PC was extensively used.

The theoretical and practical importance of the research is the work of national and foreign scientists relating to problems of information protection and security and materials of various scientific conferences, reports and authors' publications.

Scientific novelty of the work:

1. An attack detection method based on the combined use of the methods of attack signature search and detection of anomalies in a user's work, which allows for substantial improvement in attack detection characteristics, is proposed and substantiated.

2. A new approach to handling data entering an information system is proposed, which consists in representing the data in the form of signatures, allowing one to use the advantages of neural networks to solve attack recognition problems.

3. Based on the research, the potential is shown of using a neural network to process incoming data in order to achieve highly effective attack detection.

4. Recommendations for applying the developed method in intrusion detection systems are proposed.

Theoretical and practical value of the dissertation research:

1. The proposed attack detection method may be used for building attack detection systems with characteristics superior to most commercial systems.

2. The proposed approach to generating a neural network training set will make it possible to achieve high attack detection system performance by ensuring that the set is representative.

3. The chosen approach based on the use of a combined attack detection method will help simplify implementation of an attack detection system.