ИССЛЕДОВАНИЕ И МОДЕЛИРОВАНИЕ НЕЙРОСЕТЕВОГО МЕТОДА ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ СЕТЕВЫХ АТАК

Адиль Тимофеев, Александр Браницкий

***Аннотация**: Рассматриваются возможности и перспективы применения нейросетевых технологий для распознавания сетевых атак. Значительное внимание уделяется методологии имитационного моделирования нейронных сетей для обнаружения и классификации сетевых атак и экспериментальным результатам.*

***Keywords**: нейронные сети, методы обнаружения и классификации, сетевые атаки.*

***ACM Classification Keywords:** E.4 CODING AND INFORMATION THEORY*

Введение

Одним из наиболее эффективных средств массового распараллеливания и ускорения процессов обработки и передачи потоков данных в задачах обнаружения закономерностей, распознавания образов и классификации данных являются искусственные нейронные сети (НС). Естественным прототипом искусственных НС является биологический мозг и центральная нервная система человека и животных. Возможности искусственных и биологических НС могут значительно расшириться при коллективном (мультиагентном) решении сложных интеллектуальных задач (data mining, knowledge discovery и т.п.).

Высокая сложность и размерность многих задач обнаружения закономерностей, распознавания образов и классификации данных, а также часто возникающая необходимость их решения в реальном времени требуют массового параллелизма и самоорганизации распределённых вычислений на базе НС. С этой точки зрения особый интерес представляют самоорганизующиеся гомогенные и гетерогенные полиномиальные нейронные сети (ПНС) и их разновидности [1–9].

Основные идеи, математические модели, методы обучения и принципы самоорганизации гомогенных и гетерогенных НС были описаны и развиты в [2–9]. К гомогенным НС прежде всего относятся трёхслойные перцептроны и НС Хопфилда [1]. Особенности ПНС заключаются в следующем: архитектура ПНС гетерогенна и многослойна; наличие слоя полиномиальных нейронных элементов (П-нейронов); возможность и целесообразность самоорганизации архитектуры ПНС различных типов; детерминированные и вероятностные методы обучения и самоорганизации гетерогенных ПНС; принципы минимальной сложности и высокой экстраполяции гетерогенных ПНС; алгебраическое требование диофантовости (целочисленности синаптических весов) гетерогенных ПНС [2–9].

Архитектура НС представляет собой иерархическую последовательность нескольких однородных слоёв (непересекающихся подмножеств) параллельно работающих нейроэлементов (НЭ) различных типов. В различных слоях НС могут использоваться разные НЭ, но каждый слой (подмножество НЭ) является однородным (гомогенным). При этом обработка информации в каждом слое НЭ осуществляется параллельно.

Каналы связи между предыдущим и последующим слоями гетерогенной НС являются однонаправленными и имеют регулируемые веса (синаптические параметры). Эти веса каналов связи настраиваются в процессе обучения и самоорганизации архитектуры НС по имеющимся экспериментальным данным или прецедентам.

Традиционно гомогенные или гетерогенные НС используются для автономного принятия решений в задачах обнаружения закономерностей, распознавания образов, диагностики состояний, классификации данных и т.п. По существу эти НС являются обучаемыми интеллектуальными агентами, которые настраиваются на индивидуальное (одно-агентное) решение конкретных задач по обучающим базам данных (ОБД).

В то же время существует большой класс интеллектуальных задач, требующий не только индивидуальных (одно-агентных), но и коллективных (мульти-агентных) решений. Классическим примером этого могут служить особенно сложные и ответственные задачи медицинской диагностики, когда врачи вынуждены прибегать к помощи своих коллег для совместной постановки окончательного диагноза. При этом формируется “консилиум” , т.е. профессиональная группа врачей, интегрирующая знания и опыт входящих в неё членов для коллективного принятия наиболее правильных и сбалансированных диагностических решений.

Другим примером сложных задач, требующих коллективных решений, являются глобальные задачи, допускающие естественную (например, иерархическую или мультифрактальную) декомпозицию на множество локальных задач. В этом случае решение сложной (глобальной) задачи может быть распределено между интеллектуальными НС-агентами, специализирующимися на решении N частных (локальных) задач. Параллельная работа N таких НС-агентов может значительно ускорить обработку информации и повысить надежность решения общей (глобальной) задачи.

Специальные агенты-координаторы могут принимать коллективные (мультиагентные) решения на основе локальных (одно-агентных) решений остальных НС-агентов с помощью мажоритарных принципов или процедур голосования (например, по “большинству голосов”) [7–9]. При этом все локальные решения принимаются параллельно, что ускоряет принятие глобального (коллективного) решения в N раз.

В ряде случаев глобальная самоорганизация НС-агентов обеспечивается иерархической, фрактальной или мультифрактальной декомпозицией общей задачи на N подзадач. При этом степень внешнего (глобального) параллелизма в мульти-агентной нейросетевой системе определяется параметром N, характеризующем одновременную работу N локальных НС-агентов. Предлагаемые иерархические гетерогенные архитектуры и быстрые алгоритмы обучения ПНС разных типов обеспечивают высокий параллелизм и самоорганизацию нейровычислений в процессе решения интеллектуальных задач. Они успешно применялись для решения ряда прикладных задач (распознавание деталей на конвейере, классификация дорожных ситуаций, диагностика и оценка эффективности лечения артритов, векторная диагностика и расшифровка гастритов, прогнозирование исхода черепно-мозговых травм и т.д.) и нейросетевого представления генетического кода [1–9].

Аккумулируемые в гомогенных и гетерогенных НС с самоорганизующейся архитектурой “нейрообразы” и решающие (классифицирующие и идентифицирующие) правила обеспечивают массовый параллелизм, хорошую экстраполяцию и высокое быстродействие при принятии оптимальных или субоптимальных решений. Коллективное (мультиагентное) использование гетерогенных ПНС в качестве нейросетевых

агентов позволяет дополнительно распараллелить и распределить между локальными НС-агентами процессы решения сложных (глобальных) задач распознавания образов, анализа изображений и сцен, расширенной (векторной) диагностики состояний и адаптивной маршрутизации и классификации информационных потоков.

1. Проблемы и методы защиты информации

Бурное развитие компьютерных сетей и информационных технологий порождает множество проблем, связанных с безопасностью информационных ресурсов. В связи с несовершенством существующих методов защиты компьютерных систем от сетевых атак разработка новых методов защиты информации, позволяющих повысить уровень защищенности компьютерных систем от несанкционированного воздействия, является актуальной и востребованной [10].

Существует три основных подхода, используемых при обнаружении и классификации сетевых атак [10–13]:

- статистический анализ;
- экспертные системы;
- нейронные сети.

Кроме того, развиваются подходы, основанные на и генетических алгоритмах и иммуноклеточных методах.

Статистический анализ находит применение, как правило, при обнаружении аномального поведения. Отклонение от среднего значения (т. е. дисперсия) профиля нормального поведения дает сигнал администратору о том, что зафиксирована атака. Средние частоты и величины переменных вычисляются для каждого типа нормального поведения (например, количество входов в систему , количество отказов в доступе, время суток и т. д.). О возможных атаках сообщается, когда наблюдаемые значения выпадают из нормального диапазона, т. е. превышают заданный порог [10–13].

Экспертная система — это система, которая в контексте обнаружения атак принимает решение о принадлежности того или иного события к классу атак на основании имеющихся правил. Эти правила основаны на опыте специалистов и хранятся в специальном хранилище. В большинстве случаев правила экспертной системы опираются на так называемые сигнатуры, которые и ищутся в контролируемом пространстве [10].

Во введении к данной работе представлен краткий обзор основных типов НС и их особенностей при решении проблем интеллектуального характера. К этим проблемам относятся и задачи обнаружения и классификации сетевых атак. Значительный интерес представляют описываемые ниже исследования и средства их реализации и методом анализа сетевого трафика на наличие аномальных соединений с целью обнаружения и классификации атак на базе трёхслойных НС, обучаемых с помощью алгоритма обратного распространения ошибки.

2. Сетевые атаки и их классификация.

Удалённой сетевой атакой будем называть информационное разрушающее воздействие на распределённую компьютерную сеть, осуществляемое программно по доступным каналам связи. Конкретные разновидности сетевых атак представлены в базе данных (БД) KDD-99 [11]. В качестве обучающего множества выступает база KDD-99 [4]. Эта БД содержит около 5000000 записей о соединениях. Каждая запись представляет собой образ сетевого соединения, включает 41 параметр сетевого трафика и промаркирована как „атака" или „не атака". В базе представлены 22 типа атаки. При этом атаки делятся на 4 основные категории: DoS, U2R, R2L и Probe [10–12].

DoS атаки — это сетевые атаки, направленные на возникновение ситуации, когда на атакуемой системе происходит отказ в обслуживании. Данные атаки характеризуются генерацией большого объема трафика, что приводит к перегрузке и блокированию сервера. Выделяют шесть DoS атак: back, land, neptune, pod, smurf, teardrop.

U2R атаки предполагают получение зарегистрированным пользователем привилегий локального суперпользователя (сетевого администратора). Выделяют четыре типа U2R атак: buffer_overflow, loadmodule, perl, rootkit.

R2L атаки характеризуются получением доступа незарегистрированного пользователя к компьютеру со стороны удаленного компьютера. Выделяют восемь типов R2L атак: ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster.

Probe атаки заключаются в сканировании сетевых портов с целью получения конфиденциальной информации. Выделяют четыре типа Probe атак: ipsweep, nmap, portsweep, satan. Согласно источнику [11] для обнаружения и классификации 9 из 22 типов атак достаточно 29 параметров, характеризующих сетевые соединения. Список этих параметров приведен в табл. 1.

Параметр	Описание параметра
1. duration	Продолжительность соединения
3. service	Служба
4. flag	Флаг терминального состояния IP-соединения
5. src_byte	Количество байт, переданных от источника к приемнику
6. dst_byte	Количество байт, переданных от приемника к источнику
7. land	Равенство порта отправителя порту получателя
8. wrong fragment	Количество отброшенных пакетов
9. urgent	Число пакетов с флагом URG
10. hot	Количество hot-индикаторов
11.count	Количество соединений между удаленным хостом и локальным хостом
12. srv_count	Количество соединений к локальной службе
13. serror rate	Процентное число соединений с ошибкой типа syn для данного

	хоста-источника
14. srv serror rate	Процентное число соединений с ошибкой типа SYN для данной службы источника
15. rerror rate	Процентное число соединений с ошибкой типа REJ для данного хоста-источника
16. srv rerror rate	Процентное число соединений с ошибкой типа REJ для данной службы источника
17. same_srv_rate	Процентное число соединений к службе
18. diff_srv_rate	Процентное число соединений к различным службам
19. srv_diff_host_rate	Процентное число соединений к различным хостам
20. dst_host_count	Количество соединений к локальному хосту, установленных удаленной стороной
21. dst_host_srv_count	Количество соединений к локальному хосту, установленных удаленной стороной и использующих одну и ту же службу
22. dst_host_same_srv_rate	Процентное число соединений к локальному хосту, установленных удаленной стороной и использующих одну и ту же службу
23. dst_host_diff_srv_rate	Процентное число соединений к локальному хосту, установленных удаленной стороной и использующих различные службы
24. dst_host_same_src_port_rate	Процентное число соединений к данному хосту при текущем номере порта источника
25. dst host srv diff host rate	Процентное число соединений к службе разных хостов
26. dst_host_serror_rate	Процентное число соединений с ошибкой типа syn для данного хоста-приемника
27. dst_host_srv_serror_rate	Процентное число соединений с ошибкой типа SYN для данной службы приемника
28. dst_host_rerror_rate	Процентное число соединений с ошибкой типа REJ для данного хоста-приемника
29. dst_host_srv_rerror_rate	Процентное число соединений с ошибкой типа REJ для данной службы приемника

Таблица 1. Параметры сетевых соединений.

3. Исследование нейросетевой технологии обнаружения и классификации сетевых атак

Обычно для обучения и тестирования НС имеющиеся экспериментальные данные разбиваются на обучающую БД и контрольную БД. В проведённых исследованиях в качестве обучающей БД с параметрами сетевых соединений, представляющая собой 10% subset KDD CUP. 99. Она представляет собой обычный текстовый файл, разбитый на несколько файлов, каждый из которых содержит набор

параметров определённого типа атаки или нормального соединения. Полученные файлы преобразуются в файлы, содержащие числовые эквиваленты, соответствующие определённому параметру соединения.

В целях ускорения обучения, тестирования и дальнейшего запуска применялся алгоритм, реализующий метод главных компонент. Вычисление вариационной матрицы главных компонент сводится к вычислению собственных векторов и собственных значений ковариационной матрицы исходных данных. Найденные собственные векторы представляют собой коэффициенты, задающие линейное преобразование для исходных векторов, а соответствующие им собственные числа выступают в роли критерия (меры) информативности новой системы. Из набора сжатых векторов с помощью генератора случайных чисел формируется обучающая БД, мощность которой составляет 10% от полной KDD CUP. 99.

Элементы этой БД поочередно подаются на вход НС для настройки весовых (синаптических) параметров. Весовые коэффициенты каждой НС сохраняются в соответствующих файлах для возможной их дальнейшей загрузки при очередном запуске системы обнаружения вторжений.

По завершении этих этапов НС полностью готовы к этапу запуска, который заключается в обнаружении и классификации активных аномальных и нормальных IP-соединений.

При эмулировании сетевых соединений данные считываются из полной базы KDD Cup 99, которая рассматривалась как контрольная БД. Записи из полной контрольной БД подавались на обученные НС. На осно- вании результатов этого этапа определяется статистика и даётся анализ эффективности НС по критериям качества распознавания типов аномальных соединений и наличию ложных срабатываний (когда нормальное соединение принимается за атаку).

Процесс обучения и тестирования НС показан на рис. 1, а процесс запуска НС представлен на рис. 2.

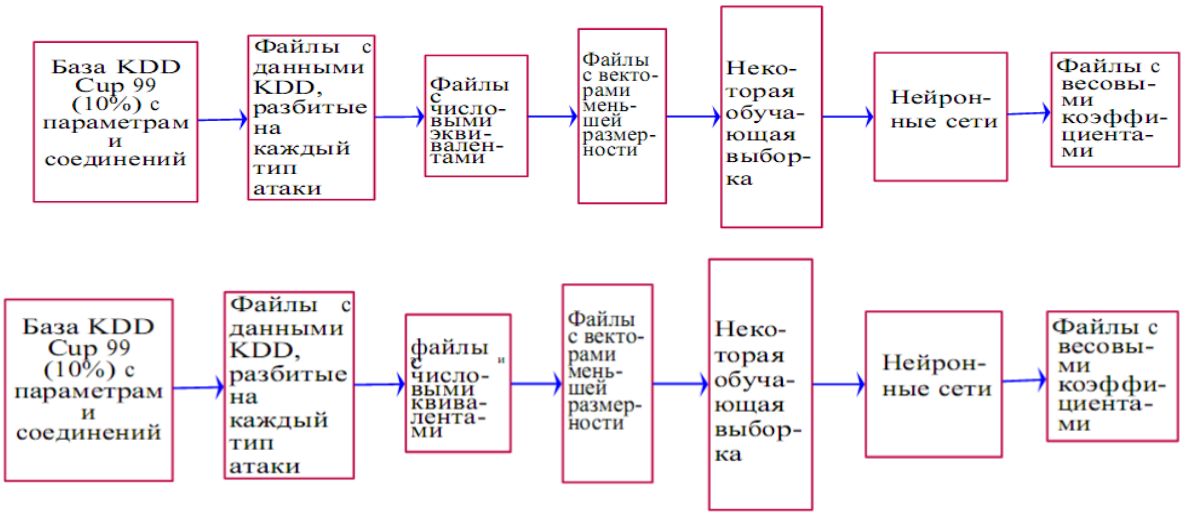

Рис. 1: Процесс обучения и тестирования НС

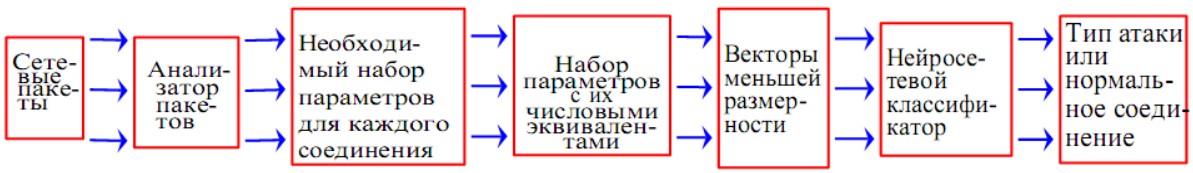

Рис. 2: Процесс запуска НС

Пакеты, поступающие на сетевую карту, перехватываются и обрабатываются анализатором пакетов (сниффером). Из поля данных и заголовка каждого пакета выделяются и вычисляются необходимые для классификации соединений параметры. Собираются некоторые статистические сведения об активных в данный момент соединениях. После сбора всех необходимых показаний о каждом запущенном соединении набор вычисленных параметров “подвергается” текстовой обработке: каждая отдельная составляющая набора параметров отображается в соответствующий ей числовой эквивалент согласно таблице, заданной в программе. Опционально (т. е. необязательно) полученный набор параметров сжимается по методу главных компонент.

Преобразованные векторы подаются на вход обученных НС, которые в свою очередь формируют линейную комбинацию каждого вектора со своими весовыми параметрами, выступающей в роли аргумента функции активации НС. На основании полученного значения принимается решение о принадлежности конкретного соединения тому или иному классу атак.

Следует отметить, что существенное значение в успешности обнаружения и определения зло-употреблений является правильно заданная планка, выступающая в роли “водораздела" между типичными атаками и нормальными соединениями. Слишком малое значение этого барьера означает возможность пропуска атак и принятия их за нормальные соединения, а слишком большое значение увеличивает число ложных срабатываний. Тем самым большое количество типично нормальных соединений будет приниматься за аномальные.

Каждая НС в отдельно созданном для нее потоке обучается параллельно с остальными для распознавания нормального соединения и одного типа атаки. Для их обучения используются равные по размеру выборки положительного и отрицательного трафиков. По завершении стадии обучения происходит барьерная синхронизация всех потоков (основной поток дожидается завершения выполнения созданных потоков).

Программная реализация НС имеет вид:

```
struct neural_network {
struct fann *ann;
enum type_attack attack;
pthread_t thread;
int index;
};
enum type_attack {back, neptune, pod, smurf, teardrop, ipsweep, nmap, portsweep, satan, normal};
```

Она состоит из следующих полей:

• указателя ann на структуру struct fann, которая определена в библиотеке FANN (Fast Artificial Neural Network) [13];

• переменной attack типа перечисления enum type_attack, задающей тип атаки, которую обучена распознавать данная нейросеть;

• переменной thread типа pthread_t, которая служит в роли идентификатора потока, в котором исполняется данная нейронная сеть;

• переменной index типа int, которая задает положение (индекс) данной переменной типа

struct neural_network в статическом массиве.

Для обучения НС применяется алгоритм обратного распространения ошибки (итеративный градиентный алгоритм, который используется с целью минимизации ошибки работы многослойного перцептрона и получения желаемого выхода) с симметричной сигмоидальной функцией активации.

4. Результаты вычислительных экспериментов

НС, которые обнаружили подозрительную сетевую активность, сигнализируют об этом и записывают параметры аномального соединения в журнал регистрации.

При проведении этапа эмулирования сетевых атак были получены следующие результаты, представленные в табл. 2. Здесь левый столбец – тип эмулируемого соединения, верхняя строка – тип нейронной сети, их пересечение – процентное количество правильно распознанных атак соответствующего типа соответствующими НС.

	back	neptune	pod	smurf	teardrop	ipsweep	nmap	portsweep	satan
back	99.7%	0%	0%	0%	0%	0%	0%	0%	0.2%
neptune	7.9%	100%	0%	0%	0%	0%	80.8%	100%	19.1%
pod	0.4%	0.4%	99.6%	0%	32.1%	1.5%	1.9%	0.4%	1.5%
smurf	0%	0%	0%	99.8%	0%	0%	99.9%	0%	100%
teardrop	0%	0%	77.3%	0.1%	99.9%	0%	0.1%	0%	0.1%
ipsweep	0%	0%	0%	0%	0%	99.8%	92.2%	1.5%	1.7%
nmap	0%	44.6%	0%	0%	0%	44.1%	99%	44.6%	3%
portsweep	0.4%	9.9%	0%	0%	0%	0%	0.1%	99.6%	90.2%
satan	0%	88.6%	0%	0%	0.1%	0.1%	3.6%	88.7%	99.5%
normal	0.4%	0%	0%	0%	0%	0.5%	1.1%	0.1%	1.3%

Таблица 2: Показатели эффективности обнаружения атак НС

Заключение

Исследование и имитационное моделирование гомогенных НС на экспериментальных данных KDD Cup 99 в задачах обнаружения и классификации сетевых атак свидетельствует об эффективности нейросетевых технологий. Естественно ожидать, что при использовании гетерогенных НС полиномиального и диофантового типа, а также их коллективов результаты распознавания сетевых атак могут быть улучшены. Перспективным представляется также частично исследованный авторами метод обнаружения и классификации сетевых атак, основанный на иммуно-клеточном подходе.

Работа выполнена при частичной поддержке гранта по Программе № 14 Президиума РАН (GRID) и издательского гранта РФФИ № 12-08-07022-д.

Литература

[1] С.Хайкин. Нейронные сети. Полный курс.– М.: Издательский дом “Вильямс”, 2006, 1104 с.

[2] Каляев А. В.. Тимофеев А. В. Методы обучения и минимизации сложности когнитивных нейро-модулей нейрокомпьютера с программируемой архитектурой. - Доклады АН, 1994, т. 237, с. 180-183.

[3] Тимофеев А.В. Методы синтеза диофантовых нейросетей минимальной сложности. - Доклады АН , 1995, т.301, № 3, с.1 106-1109.

[4] Тимофеев А.В., Шибзухов З.М. Методы синтеза и минимизации сложности диофантовых нейронных сетей над конечным полем. Автоматика и телемеханика, 1997, № 4, с. 204-212.

[5] Тимофеев А. В. Оптимальный синтез и минимизация сложности генно-нейронных сетей по енетическим базам данных. Нейрокомпьютеры: разработка и применение, № 5-6, 2002, с. 34-39.

[6] Тимофеев А. В., Шибзухов З. М., Шеожев А. М. Синтез нейросетевых архитектур по многозначному дереву решений. – Нейрокомпьютеры:разработка и применение, № 5-6, 2002, с. 44-49.

[7] Timofeev A. V. Polynomial Neural Network with Self-Organizing Architecture. - International Journal on Optical Memory and Neural Networks, 2004, N 2.

[8] Timofeev A. V. Parallelism and Self-Organization in Polynomial Neural Networks for Image Recognition. - Pattern Recognition and Inage Analysis, 2005, vol. 15, No.1, pp. 97 - 100.

[9] Тимофеев А.В. Иерархические самоорганизующиеся нейронные сети и алгоритмы мульти-агентного принятия решений – Материалы 4-й Всероссийской мульти-конференции по проблемам управления МКПУ-2011 (Россия, Дивноморское, 28 сентября – 7 октября 2011 г.), т.1, с. 43–45.

[10] А.В. Лукацкий. Обнаружение атак. 2-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2003.

[11] KDD Cup 1999 Data http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[12] Технологии обнаружения сетевых атак. http://www.bstu.by/~opo/ru/uni/bstu/science/ids/.

[13] Fast Artificial Neural Network.http://leenissen.dk/fann/up

Сведения об авторах

Тимофеев Адиль Васильевич – заведующий лабораторией информационных технологий в управлении и робототехнике Санкт-Петербургского института информатики и автоматизации Российской академии наук, Профессор кафедры информатики математико-механического факультета Санкт-Петербургского государственного университета, доктор технических наук, профессор, Заслуженный деятель науки РФ, 199178, Россия, Санкт-Петербург, 14-я линия, д. 39, СПИИРАН, tav@iias.spb.su

Браницкий Александр Александрович *– студент кафедры информатики Математико-механического факультета Санкт-Петербургского государственного университета, дипломник, alexander.branitskiy@gmail.com*