ПОДХОД К ОБЕСПЕЧЕНИЮ БЕЗОПАСНОСТИ КОМПЬЮТЕРНЫХ СЕТЕЙ С ПРИМЕНЕНИЕМ ЭЛЕМЕНТОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Войцехович Л.Ю.

Брестский государственный технический университет, г. Брест

Оперативный обмен информацией становится неотъемлемым атрибутом успешной деятельности в любой сфере. В последнее время прорыв в этой области обеспечили компьютерные технологии: компьютерные сети, электронная коммерция, корпоративные web-сайты и др. Однако наряду с необходимостью повышения надежности и скорости коммуникации, остро встал вопрос обеспечения защиты информационных ресурсов [1].

Для защиты компьютерных систем применяются различные подходы. Все подходы можно разбить на две основные категории: организационные и технические. В свою очередь технические подходы подразделяются на сетевые и хостовые. Далее в статье речь пойдет о сетевых средствах обеспечения безопасности, а именно о системах обнаружения вторжений.

Задачей Систем Обнаружения Вторжений (Intrusion Detection Systems - IDS) является защита компьютерных сетей.

Наряду с правильной политикой безопасности, архитектурой межсетевых фильтров, антивирусным программным обеспечением и другим средствам IDS часто отводится роль основного элемента защиты. IDS используются в качестве средства раннего оповещения о сетевых проблемах. Это обусловлено размещением IDS в общей схеме обороны на сетевом уровне, на котором подозрительные действия могут быть обнаружены раньше, чем на более высоких уровнях. Кроме того, IDS способна предоставлять необходимые доказательства злоумышленных действий, а также выявлять скрытые тенденции, что становится возможным при анализе большого количества данных, обрабатываемых IDS.

К недостаткам существующих моделей IDS в первую очередь можно отнести уязвимость к новым атакам, низкую точность и скорость работы. Современные системы обнаружения вторжений плохо приспособлены к работе в реальном режиме времени, в то время как возможность обрабатывать большой объем данных в реальном времени – это определяющий фактор практического использования систем IDS. Указанные недостатки трудно устранить, используя только классические методы в области компьютерной безопасности. Поэтому в последнее время системы IDS активно изучаются.

В данной статье предлагается подход к обнаружению атак на компьютерные сети, основанный на нейронных сетях.

Процесс обработки информации в IDS приведен на рис. 1. Он включает три этапа.

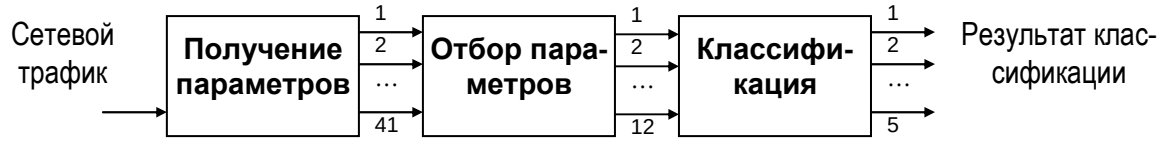

Рис. 1. Процесс обнаружения

На первом этапе осуществляется захват трафика сети (feature selection). Сбор необходимых данных выполняет специальное программное средство (sniffer). В этой работе мы использовали базу данных KDD-99 [2]. Эта база содержит около 5 000 000 записей. Каждая запись представляет собой образ сетевого соединения, который включает 41 параметр трафика и промаркирован как “атака” или “не атака”. Отдельная запись состоит из около 100 байт данных. Например, первый параметр определяет длительность соединения, второй – указывает используемый протокол, третий – целевую службу и т.д.

Второй этап связан с уменьшением размерности входного вектора данных и получением главных компонент (feature extraction). Между используемыми параметрами суще-

ствуют сложные взаимосвязи, которые достаточно тяжело проследить. Некоторые данные являются избыточными. Большое количество параметров может значительно увеличить время вычислений, поэтому этап получения главных компонент является важным этапом в процессе функционирования предлагаемых IDS.

Третий этап состоит в обнаружении и распознавании атак (classification). Атаки в базе KDD-99 делятся на четыре основные класса: DoS, U2R, R2L и Probe.

Рассмотрим различные архитектурные решения для построения систем обнаружения атак. Они основаны на применении модулярных нейронных сетей. В качестве входных данных используется 41-размерный вектор, который характеризует параметры соединения сети. Задачей IDS является обнаружение и распознавание атак. Поэтому в качестве выходных данных используется 5-мерный вектор, где 5 - это количество классов атак плюс нормальное состояние.

На основании предыдущих результатов экспериментов [3] мы отобрали три наиболее удачные модели систем обнаружения атак.

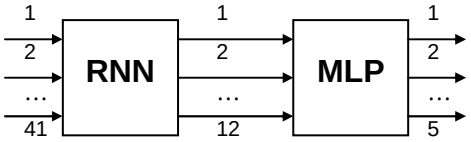

Рис. 2. Первый вариант IDS

На рис. 2 приведена простейшая система обнаружения атак, которая состоит из рециркуляционной нейронной сети (RNN) и многослойного персептрона (MLP) [4], которые соединены последовательно. Задачей RNN является сжатие входного 41-размерного вектора в 12-размерный выходной вектор. Многослойный персептрон осуществляет обработку сжатого пространства входных образов (главных компонент) с целью распознавания класса атаки.

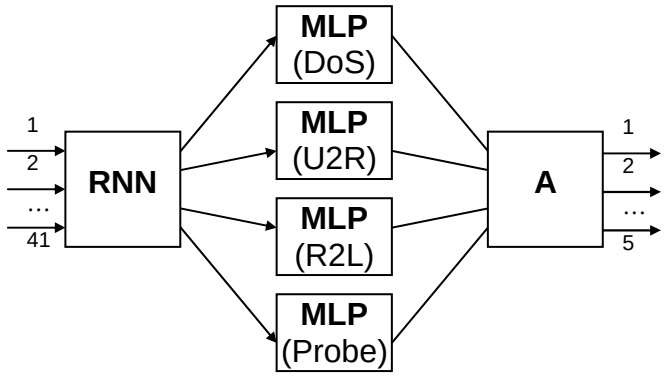

Рис. 4. Третий вариант IDS

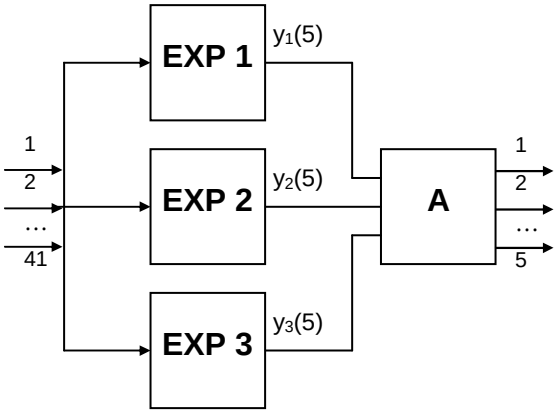

Рис. 3. Второй вариант IDS

На рис. 3 приведена вторая схема системы обнаружения атак. Она характеризуется тем, что главные компоненты с выходов RNN одновременно поступают на 4 отдельных многослойных персептрона, каждый из которых соответствует определенному классу атаки: DoS, U2R, R2L и Probe. С выходов MLP данные поступают на арбитр, который и принимает окончательное решение о состоянии системы. В качестве арбитра может использоваться линейный или многослойный персептрон.

Рассмотрим Ансамблевую нейронную сеть (рис. 4). Каждый эксперт представлен отдельной системой классификации (в нашем случае в качестве эксперта применена модель 1). Арбитр (многослойный персептрон) осуществляет процедуру голосования для формирования совместного решения всех трех экспертов. Такая нейронная сеть обучается по алгоритму усиления за счет фильтрации [5], который предполагает обучение каждого последующего эксперта на множестве данных, формируемых на основании результатов обучения предыдущих экспертов.

После обучения нейронных сетей они способны обнаруживать враждебную активность в сети.

Эксперименты проводились в соответствии с данными, приведенными в таблице 1.

Таблица 1. Структура обучающей выборки и тестовых данных

	DoS	U2R	R2L	Probe	Normal	Всего
обучающая выборка	3571	37	278	800	1500	6186
тестовая выборка	391458	52	1126	4107	97277	494020

Обучающая выборка использовалась для настройки нейронных сетей. После этапа обучения на каждую из предложенных моделей подавались образы из тестовой выборки и рассчитывались показатели эффективности (доля обнаруженных, доля распознанных атак и число ложных срабатываний системы).

Сводные данные по каждому из вариантов построения системы обнаружения атак приведены в таблице 2:

Таблица 2. Сводные данные по результатам тестирования

Модель	Обнаруженные атаки	Распознанные атаки	Ложные срабатывания	Общая доля распознанных %
мод. 1	396696 (99.98%)	375522 (94,65%)	46446 (47.75%)	86.30%
мод. 2	395949 (99.80%)	375391 (94.61%)	13398 (13.77%)	92.97%
мод. 3	396549 (99.95%)	375730 (94.70%)	12549 (12.90%)	93.21%

Таким образом, модель 3 характеризуется высокой точностью (93,21%) и наименьшим числом ложных срабатываний. При использовании модели 1 были распознаны 86,3% входных образов, а модели 2 – 92,97%. Модели 2 и 3 могут успешно применяться для работы с большими наборами сложных по структуре данных.

Предлагаемый подход к построению сетевой системы обнаружения вторжений, основанный на взаимодействии рециркуляционной нейронной сети и многослойного персептрона, позволяет достичь высоких показателей. Нейронная сеть способна учесть сложные нелинейные зависимости, существующие между рассчитываемыми параметрами трафика, и распознать действия злоумышленника в общем потоке данных, поступающих из сети. Кроме того, обученная нейронная сеть функционирует достаточно быстро. Дальнейшие исследования связаны с повышение эффективности нейросетевых моделей и проведением детальных испытаний в условиях реальной сети.

Литература

1. Web Application Security Consortium. Классификация угроз [Электрон. ресурс]. - Режим доступа: www.webappsec.org.
2. 1999 KDD Cup Competition [Electronic resource]. - Mode of access: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

3. Golovko V., Vaitsekhovich L. Neural Network Techniques for Intrusion Detection // Proceedings of International Conference on Neural Networks and Artificial Intelligence (ICNNAI-2006). - 2006. - P. 65-69.
4. Головко В.А. Нейронные сети: обучение, организация и применение. Кн. 4: Учеб. пособие для вузов / Общая ред. А.И. Галушкина. – М.: ИПРЖР, 2001. – 256 с.
5. Drucker H., Schapire R. and Simard P. Improving performance in neural networks using a boosting algorithm // In S.J.Hanson, J.D.Cowan and C.L.Giles eds., Advanced in Neural Information Processing Systems 5, Denver, CO, Morgan Kaufmann, San Mateo, CA. - 1993. - P. 42-49.

НЕЙРОСЕТЕВАЯ АППРОКСИМАЦИИ ПРИ МОДЕЛИРОВАНИИ И АНАЛИЗЕ РЕЗУЛЬТАТОВ ФЛУОРЕСЦЕНТНЫХ ЭКСПЕРИМЕНТОВ

Горошко В.В.
Белорусский государственный университет, г. Минск

Введение

Фотофизические процессы играют важную роль в природных и искусственных преобразователях энергии (хлоропласты растений, солнечные батареи). Часто фотофизические процессы в реальных системах весьма сложны и не могут быть адекватно описаны аналитически. В этом случае используют имитационное моделирование, позволяющее моделировать и анализировать поведение сколь угодно сложной системы при использовании соответствующих вычислительных ресурсов. Однако при применении на практике имитационного моделирования для анализа данных, исследователь сталкивается с рядом проблем, среди которых наиболее значимыми являются: (i) значительные временные затраты, так как имитационное моделирование, как правило, чрезвычайно громоздко в вычислительном плане; (ii) наличие не одного, а множества локальных минимумов ошибки (вследствие стохастичности имитационной модели), что делает необходимым многократный анализ данных с различными начальными приближениями неизвестных параметров.

Для решения этих проблем было предложено аппроксимировать имитационную модель некоторой гладкой зависимостью, которая бы отражала поведение, как экспериментальной системы, так и ее имитационной модели в зависимости от входных переменных и скрытых (искомых) параметров (Nazarov, 2004). В качестве аппроксиматора, способного заменить имитационную модель, было решено использовать искусственные нейронные сети (ИНС). Являясь универсальным аппроксиматором (Cybenko, 1989; Hornik, 1989), ИНС позволяют решать многие трудно формализуемые задачи.

Нами предложен метод замены имитационной модели системы искусственной нейронной сетью (ИНС) для ускорения алгоритмов оптимизации. Разработанный метод был применён для задачи моделирования флуоресценции молекулярной системы с двумя возбуждаемыми уровнями.

Теория

Основной проблемой анализа данных с использованием имитационной модели (simulation-based fitting) являются высокие временные затраты при подгонке параметров. Эта проблема может быть решена путем замены имитационной модели моделью "черного ящика" – нейронной сетью (Nazarov, 2004).

Большинство зависимостей в химии и физике могут быть представлены гладкими кривыми (если конкретная экспериментальная реализация такой зависимости будет носить стохастический характер – будем рассматривать средние значения). Это дает возможность аппроксимировать такие зависимости, а значит и имитационную модель, многослойным персептроном.