Уфа: УГАТУ, 2005 Вестник УГАТУ Т. 6, № 1 (12). С. 116–120

УДК 681.324.067

В. И. ВАСИЛЬЕВ, А. Ф. ХАФИЗОВ

НЕЙРОСЕТЕВЫЕ СИСТЕМЫ ОБНАРУЖЕНИЯ АТАК НА WWW-СЕРВЕР

Излагается подход к применению нейросетевых технологий для создания систем обнаружения атак прикладного уровня модели OSI. С целью оценки эффективности приводится математический расчет вероятностей ошибок системы для различных методов обнаружения атак. Предлагается комбинированный метод обнаружения атак, основанный на объединении метода поиска сигнатур и метода обнаружения аномалий. Исследуются особенности реализации данного метода с помощью нейронных сетей. *Информационная безопасность; системы обнаружения атак; нейронные сети*

ВВЕДЕНИЕ

Эффективность функционирования современных информационных систем в значительной мере связана с проблемой обеспечения безопасности обрабатываемой в них информации. Вместе с тем анализ современных систем защиты информации показывает, что их возможности не позволяют достичь достаточного уровня информационной безопасности. Злоумышленник может получить несанкционированный доступ в результате проведения атаки, которая заключается в отправке в систему таких входных данных, обработка которых приведет к нештатным ситуациям: выполнению недопустимых команд, удалению либо модификации критичной информации [1].

Для защиты от атак применяются специализированные программно-аппаратные средства, функциями которых являются обнаружение и блокирование попыток несанкционированного доступа — системы обнаружения атак (СОА) и системы межсетевого экранирования (СМЭ) соответственно. Межсетевые экраны имеют достаточно простой механизм функционирования, поэтому их создание не представляет большой трудности, хотя эффективность их функционирования далека от желаемой. Алгоритм функционирования систем обнаружения атак, в отличие от межсетевых экранов, сложен (так как современные информационные системы имеют сложный алгоритм функционирования) и характеризуется наличием ряда проблем, не решенных по настоящее время. Данный факт приводит к тому, что разработчики систем обнаружения атак вынуждены упрощать алгоритмы анализа информации, следствием чего является снижение числа обнаруживаемых атак.

Одним из перспективных направлений по совершенствованию систем обнаружения атак является применение нейросетевого математического аппарата, позволяющего проводить анализ информации на качественно новом уровне, а следовательно, значительно повысить эффективность защиты информации.

1. ПОСТАНОВКА ЗАДАЧИ ОБНАРУЖЕНИЯ АТАКИ

Доказательство преимущества нейронных сетей для создания системы обнаружения атак можно представить следующим образом. Под *задачей обнаружения атаки* будем понимать задачу поиска из всего потока входных значений той группы данных, которые не относятся к «безопасным» данным, а следовательно, соответствуют атаке.

Теоретически, любые поступающие в информационную систему данные можно представить в виде некоторой функции $y = f(x)$ — аппроксиматора входных значений. Чем точнее данная функция описывает множество «безопасных» данных, тем качественнее система обнаружения атак будет обнаруживать атаки.

В классической теории аппроксимации, составляющей математическую основу задачи обнаружения атак, аппроксимируемая функция представляется в следующем виде [2]:

$$f(x) = \sum_i \alpha_i \psi_i(x). \qquad (1)$$

Набор базисных функций $\psi_i(x)$, как правило, выбирается априори, исходя из конкретного метода аппроксимации и некоторых свойств данных функций, доказываемых и используемых в процессе аппроксимации.

При использовании нейронных сетей базовое выражение для аппроксимируемой функции выглядит несколько иначе, например, для трехслойной нейронной сети с последовательными связями

$$f(x) = \psi\left(\sum_i \alpha_i \psi\left(\sum_j \ldots\left(\sum_k \alpha_{kj} x_k\right)\right)\right), \quad (2)$$

где ψ — активационная функция нейрона.

В данном выражении функция $f(x)$ представлена в нейросетевом логическом базисе с множеством коэффициентов α_i, α_{kj}, которые являются настраиваемыми в процессе поиска наилучшей аппроксимации.

Как видно из (2), применение нейронных сетей позволит не только реализовать возможности классических систем, но и улучшить точность аппроксимации, а следовательно, в конечном итоге улучшить характеристики обнаружения атак.

С точки зрения математической статистики, задачу обнаружения атак можно сформулировать как задачу различения гипотез: соответствуют ли текущие входные данные атаке или нет [3]. Все множество входных данных M можно разбить на два непересекающихся подмножества A и B, соответствующих множеству данных атаки и множеству безопасных данных соответственно. Существуют две вероятности ошибок обнаружения атак:

- ошибка первого рода («пропуск цели») P_I — вероятность того, что запрос, являющийся атакой, будет отнесен к множеству B;
- ошибка второго рода («ложное срабатывание») P_{II} — вероятность того, что безопасный запрос будет отнесен к множеству атак A.

Чем меньше вероятности ошибок P_I и P_{II}, тем эффективнее система обнаружения атак способна обнаружить попытки несанкционированного доступа в информационную систему. Чтобы упростить построение системы обнаружения атак, введем величину *риска* от принятия решения W — средневзвешенную вероятность ошибок первого и второго рода:

$$W = \delta_1 P_I + \delta_2 P_{II}, \quad (3)$$

где δ_1 — вес ошибки первого рода; δ_2 — вес ошибки второго рода, $(\delta_1 + \delta_2 = 1)$.

Значения весов δ_1 и δ_2 устанавливаются разработчиком системы исходя из конкретной реализации с учетом требуемого соотношения важности каждой из указанных ошибок.

Таким образом, задача построения нейросетевой системы обнаружения атак может быть сформулирована следующим образом: необходимо выбрать такие параметры аппроксимации в формуле (2), чтобы получить минимальное значение риска (3).

2. МЕТОДЫ ОБНАРУЖЕНИЯ АТАК

Классически считается, что существуют два принципиально противоположных подхода к обнаружению атак [4]: 1) «разрешено все, кроме того, что запрещено», и 2) «запрещено все, кроме того, что разрешено». Каждый из этих методов имеет свои особенности, выраженные в разном соотношении вероятностей ошибок 1-го и 2-го рода.

Первый из них, носящий название *метода поиска сигнатур атаки*, считает атакой любые данные, подходящие под описание (сигнатуру) известных атак. Если же текущие данные не соответствуют любой из известных сигнатур атак, то считается, что они безопасны.

Во втором методе — *методе обнаружения аномалий* — предварительно строится так называемая модель безопасных действий пользователя, представляющая собой все множество входных данных («шаблонов безопасных данных»), которые считаются безопасными для защищаемой системы. Присутствие во входном потоке данных, не соответствующих модели безопасных действий, считается аномалией, т. е. атакой.

Для метода обнаружения атак, основанного на поиске сигнатур атаки, указанные выше вероятности P_I и P_{II} можно записать в следующем виде:

$$P_I = \frac{\psi(A) - \psi\left(\bigcup_{x=0}^{n} (A \cap A_{\text{упр}.x})\right)}{\psi(A)}, \quad (4)$$

$$P_{II} = \frac{\psi\left(\bigcup_{x=0}^{n} (A_{\text{упр}.x} \cap B)\right)}{\psi(B)}, \quad (5)$$

где A — множество входных данных, которые характеризуют атаку; B — множество безопасных данных; $A_{\text{упр}.x}$ — множество сигнатур атак, которые будут использованы при создании системы обнаружения атак; $\psi(Z)$ — функция, определяющая число элементов конечного множества Z.

Увеличение вероятностей ошибок первого и второго рода здесь происходит за счет вынужденного упрощения описания множества атак A, что приводит к ложному описанию элементов множества B множеством $A_{\text{упр}.x}$, а также к неполному описанию множества A множеством $A_{\text{упр}.x}$.

Исходя из выражений (4) и (5), видно, что для данного метода вероятность ошибки первого рода P_I значительно больше вероятности ошибки второго рода P_{II}, что может быть использовано для достижения минимального значения риска W в конкретных прикладных задачах.

Для уменьшения вероятностей ошибок принятия решения предлагается комбинированный метод, заключающийся в дополнительном описании части элементов множества B. Несмотря на то, что при введении дополнительного множества повышаются накладные расходы на обработку и хранение этого множества, данный метод позволяет снизить вероятность ошибки второго рода (рис. 1), поскольку здесь:

$$P_I = \frac{\psi\left(A - A \cap \left(A_{\text{упр}} \cup A_{\text{точн}}\right)\right)}{\psi(A)} = \frac{\psi\left(\left(A - A \cap A_{\text{упр}}\right) - A \cap A_{\text{точн}}\right)}{\psi(A)}, \quad (6)$$

$$P_{II} = \frac{\psi(B \cap A_{\text{упр}}) - \psi(B_{\text{точн}} \cap A_{\text{упр}})}{\psi(B)}, \quad (7)$$

где $A_{\text{точн}}$ — множество сигнатур атак, описывающее только элементы множества A; $B_{\text{точн}}$ — множество безопасных данных, описывающее элементы множества B.

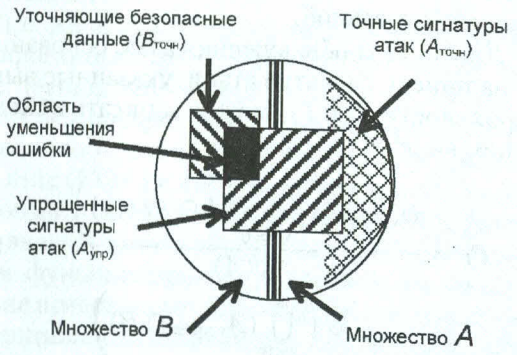

Рис. 1. Комбинирование метода поиска сигнатур атаки с уточняющими безопасными данными

Из рис. 1 и выражений (6) и (7) видно, что чем больше уточняющие безопасные данные $B_{\text{точн}}$ перекрывают множество упрощенных сигнатур атаки, тем меньше будет величина ошибки второго рода и, следовательно, улучшится качество обнаружения атак.

Для метода обнаружения атак, основанного на поиске аномалий, выражения для вероятностей ошибок 1-го и 2-го рода можно записать следующим образом:

$$P_I = \frac{\psi\left(\bigcup_{x=0}^{n}\left(B_{\text{упр}.x} \cap A\right)\right)}{\psi(A)}, \quad (8)$$

$$P_{II} = \frac{\psi(B) - \psi\left(\bigcup_{x=0}^{n}\left(B_{\text{упр}.x} \cap B\right)\right)}{\psi(B)}, \quad (9)$$

где $B_{\text{упр.х}}$ — множество всех шаблонов безопасных действий, которые будут использованы при создании системы обнаружения атак.

Сложность алгоритмов обработки информации приводит к тому, что множество шаблонов безопасных действий $B_{\text{упр.х}}$, которые может обработать система обнаружения атак, будет значительно различаться от того набора данных B, которые будут являться безопасными в системе. Следствием этого является значительно большая величина вероятности ложного обнаружения атаки P_{II} по сравнению с величиной P_I.

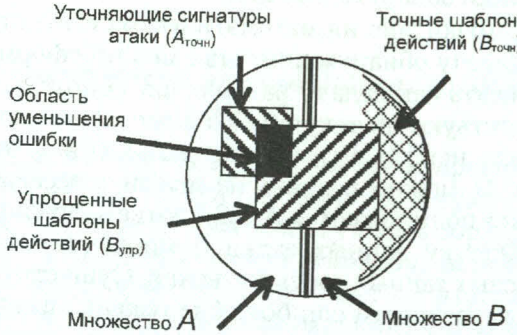

Рис. 2. Комбинирование метода обнаружения аномалий с уточняющими сигнатурами атак

Снижение вероятности ошибки может быть достигнуто путем комбинирования методов обнаружения атак, как это было рассмотрено выше, но в данном случае метод обнаружения сигнатур атак будет использоваться как уточняющий (рис. 2).

3. РАЗРАБОТКА ПРОТОТИПА НЕЙРОСЕТЕВОЙ СИСТЕМЫ ОБНАРУЖЕНИЯ АТАК

Предложенный комбинированный метод обнаружения атак был использован при создании исследовательского прототипа нейросетевой системы обнаружения атак на WWW-сервер.

Таблица

Сравнительная таблица результатов сканирования СОА

Тестируемая СОА	Общее число атак	Число правильно распознанных атак	Вероятность ошибки первого рода	Общее число безопасных запросов	Число ложных срабатываний	Вероятность ошибки второго рода	Риск
Сканирование Nessus							
Нейросетевая СОА	1247	1149	0,08	132	7	0,05	0,21
RealSecure Network Sensor	1247	732	0,41	132	3	0,02	0,85
Snort	1247	821	0,34	132	24	0,18	0,87
Сканирование Nikto							
Нейросетевая СОА	3125	2986	0,04	83	8	0,10	0,19
RealSecure Network Sensor	3125	1029	0,67	83	2	0,02	1,37
Snort	3125	936	0,70	83	13	0,16	1,56
Сканирование XSpider							
Нейросетевая СОА	732	707	0,03	62	12	0,19	0,26
RealSecure Network Sensor	732	632	0,14	62	3	0,05	0,32
Snort	732	598	0,18	62	5	0,08	0,45
Сканирование Retina							
Нейросетевая СОА	648	617	0,05	17	0	0,00	0,10
RealSecure Network Sensor	648	586	0,10	17	0	0,00	0,19
Snort	648	572	0,12	17	3	0,18	0,41
Сканирование ISS Internet Scanner							
Нейросетевая СОА	421	403	0,04	28	1	0,04	0,12
RealSecure Network Sensor	421	410	0,03	28	3	0,11	0,16
Snort	421	401	0,05	28	2	0,07	0,17
Сканирование N-Stealth HTTP							
Нейросетевая СОА	27244	26579	0,02	892	40	0,04	0,09
RealSecure Network Sensor	27244	3875	0,86	892	5	0,01	1,72
Snort	27244	3027	0,89	892	12	0,01	1,79

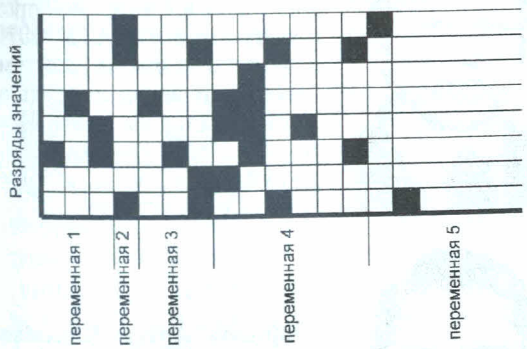

Рис. 3. Представление входных данных в виде пифограммы

Обмен между пользователем и WWW-сервером осуществлялся посредством сложного протокола прикладного уровня HTTP1.1 (HyperText Transfer Protocol) [5]. В целях упрощения задачи построения прототипа системы обнаружения атак было принято, что для обеспечения безопасности информационной системы необходимо контролировать допустимость значений переменных CGI и Cookie, т. е. переменных, обработка которых может привести к нештатной ситуации в системе (т. е. «взлому»).

Нейронные сети по своей специфике могут обрабатывать входные данные ограниченного диапазона. Поэтому при построении исследовательского прототипа значения Cookie и CGI переменных были представлены в виде графического образа – пифограммы (рис. 3).

Специфика комбинированного метода обнаружения атак выражается в том, что задачей нейронной сети является распознавание входных данных и отнесение их к одному из двух классов – классу атаки либо классу безопасных данных. Произведенный сравнительный анализ характеристик нейронных сетей различных топологий показал, что наилучшие результаты обработки данных могут быть получены при использовании гибридной нейронной сети структуры Counterpropagation, представляющей собой последовательно соединенные сети структур Кохонена и многослойного персептрона [6] (рис. 4). В качестве базового пакета для построения

и обучения нейросетевой системы обнаружения атак использовался нейросимулятор SNNS (Stuttgart Neural Network Simulator) 4.2.

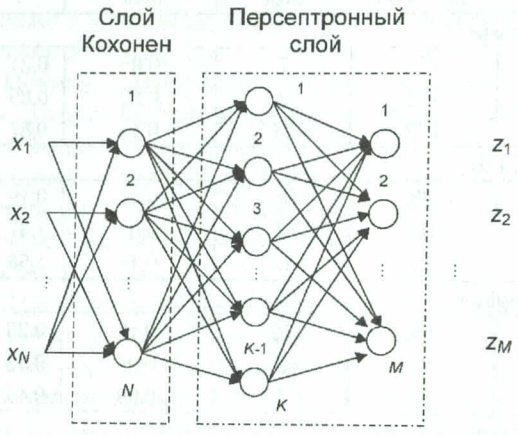

Рис. 4. Структура сети Counterpropagation

Обучение заключалось в последовательном поиске группы оптимальных значений каждой из группы обучаемых параметров нейронной сети. После обучения прототип нейросетевой системы обнаружения атак подвергся тестированию. Для этого WWW-сервер–«жертва» был атакован с помощью сканеров уязвимостей различных типов (Nessus, Nikto, XSpider, Retina, ISS IS, N-Stealth HTTP). Характеристики обнаружения атак представлены в таблице. Для сравнения, помимо исследовательского прототипа, были протестированы методом сканирования современные системы обнаружения атак Snort и RealSecure Network Sensor. Как видно из таблицы, разработанная система обнаружения атак обеспечивает значительное снижение величины риска.

ЗАКЛЮЧЕНИЕ

Таким образом, в качестве результатов проведенных исследований можно отметить следующее:

- предложенный метод обнаружения атак может быть использован для построения систем обнаружения атак, превосходящих по своим характеристикам классические системы;
- предлагаемый подход к представлению входных данных в виде графического образа позволил эффективно использовать возможности нейронной сети;
- используемая топология нейронной сети позволила достичь показатели обнаружения атак, превышающие характеристики современных систем обнаружения атак.

СПИСОК ЛИТЕРАТУРЫ

1. **Лукацкий А. В.** Обнаружение атак. СПб.: БХВ-Петербург, 2001. 624 с.
2. **Нейроматематика:** Нейро-компьютеры и их применение. Кн. 6. Учеб. пособие для вузов / А. Д. Агеев, А. Н. Балухто, А. В. Бычков и др.; Общ. ред. А. И. Галушкина М.: ИПРЖР, 2002. 448 с.
3. **Васильев В. И.** Распознающие системы. К.: Наукова Думка, 1969. 291 с.
4. **Graham R.** FAQ: Network Intrusion Detection Systems. http://www.robertgraham.com/pubs/network-intrusion-detection.html
5. **Fielding R., Gettys J.** Request for Comments 2616 — HyperText Transfer Protocol — HTTP/1.1. Network Working Group, 1999.
6. **Hecht-Nielsen R.** Counterpropagation networks // Proc. of the IEEE First Int. Conf. of Neural Networks. IEEE Press, 1987. P. 19–32.

ОБ АВТОРАХ

Васильев Владимир Иванович, проф., зав. каф. выч. техники и защиты информации. Дипл. инж. по промэлектронике (УГАТУ, 1970). Д-р техн. наук по системн. анализу и автоматич. управлению (ЦИАМ, 1990). Иссл. в обл. многосвязных, многофункциональных и интеллектуальных систем.

Хафизов Артем Фоэлевич, вед. специалист банка «УралСиб». Дипл. инж. (УГАТУ, 2000). Канд. техн. наук по сист. анализу, управлению и обр. информации (УГАТУ, 2004). Иссл. в обл. безопасности инф. систем и нейронных технологий.