Магницкий Н. А.

Институт системного анализа РАН, Москва; e-mail: nmag@isa.ru

ИСПОЛЬЗОВАНИЕ БИНАРНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ОБНАРУЖЕНИЯ АТАК НА РЕСУРСЫ РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ*

Определен спектр информационных угроз и предложен оригинальный нейросетевой метод обнаружения и анализа информационных атак в распределенных информационных системах.

1. Введение

Распределенная информационная система (далее РИС) представляет собой совокупность взаимосвязанных программных и аппаратных ресурсов, которые необходимо защищать от злоумышленных действий (атак) нарушителей. Обнаружение атак — процесс выявления таких злоумышленных действий, нацеленных на РИС.

Будем предполагать, что каждый ресурс РИС характеризуется состоянием, которое известно в любой момент времени. Для каждого состояния известен способ оценки его влияния на информационную безопасность РИС. Будем также предполагать, что для каждого ресурса смена его состояния наблюдаема и что наблюдатель имеет средства сбора информации о состояниях всех ресурсов РИС.

Состояние ресурса считается безопасным в текущий момент времени, если все использующие его объекты имеют права доступа к данному ресурсу и характеристика загруженности ресурса меньше порога загруженности. Состояние ресурса считается опасным в текущий момент времени, если хотя бы один из использующих его объектов не имеет права доступа к данному ресурсу или характеристика загруженности ресурса равна емкости ресурса. Информационно безопасное состояние — это состояние РИС, при котором все ресурсы РИС находятся в безопасном состоянии. Атака на РИС — это последовательность действий, производимых нарушителем

* Работа поддержана программой фундаментальных научных исследований ОИТВС РАН, проект № 2.4.

и ведущих к нарушению информационной безопасности системы. Цель атаки — перевод системы в опасное состояние, при котором для одного или нескольких ресурсов нарушена конфиденциальность, целостность или доступность. Нарушение конфиденциальности — ситуация, когда информацией обладает посторонний субъект. Нарушение целостности — ситуация, когда произошло умышленное искажение (модификация или удаление) данных, хранящихся в РИС или передаваемых из одной РИС в другую. Нарушение доступности — ситуация, когда доступ к данному ресурсу для легальных пользователей блокирован. Блокирование может быть постоянным, либо вызывать задержку, достаточно долгую для того, чтобы ресурс стал бесполезным.

Атаки, направленные на различные объекты РИС и различающиеся по своей реализации, можно разделить на пять классов [1, 2]: атаки на сетевые ресурсы — связанные группы хостов, коммутационных элементов и каналов передачи данных; атаки на файловые ресурсы — данные, представленные в форме файлов; атаки на программные ресурсы — программы, находящиеся в стадии исполнения; атаки на ресурсы баз данных — точки доступа, дисковое пространство, файловые и системные утилиты, утилиты администратора; атаки на вычислительные ресурсы — процессоры, память.

2. Математическая постановка задачи обнаружения атак

Состояние РИС в момент времени t определим вектором наблюдаемых параметров $x(t) = \big(x_1(t), x_2(t), \ldots, x_n(t)\big)$. Будем рассматривать конкретную атаку определенного класса как траекторию в n-мерном пространстве параметров:

$$L_i = \big(x(t), x(t+\tau), \ldots, x(t+k\tau)\big),$$

где τ — период замеров значений параметров. Обозначим через

$$L = (L_1, L_2, \ldots, L_m)$$

множество траекторий всех атак данного класса. Пусть L_i — наблюдаемая в процессе функционирования РИС траектория, тогда если $L_i \in L$, то на РИС осуществлена атака. Таким образом, для обнаружения атак необходимо построить множество L и определить принадлежность наблюдаемой траектории L_i к этому множеству.

Пусть T — конечная длительность атаки заданного класса и τ — период времени замера наблюдаемых параметров, т. е. на траектории проведения атаки задано $k = T/\tau$ состояний РИС. Для каждого замера параметров в пространстве параметров существует набор несвязных областей таких, что попадание замера в одну из этих областей означает принадлежность траектории множеству L на данном замере (т. е. возможность

принадлежности наблюдаемой траектории к одной из возможных траекторий атак при условии, что все предшествующие замеры принадлежали множеству L). Полагаем, что если в ходе k замеров состояний РИС траектория на каждом замере попадала в одну из таких областей, то она принадлежит множеству L — множеству траекторий соответствующего класса атак.

Пусть $G(t) = G_1(t), G_2(t), \ldots, G_l(t)$ множество областей таких, что попадание замера $x(t)$ в одну из этих областей означает принадлежность траектории к множеству L на данном замере. Таким образом, задача обнаружения атак сводится к задачам:

1) построения несвязного множества областей G для каждого интервала k-го замера;
2) определения принадлежности замеренного состояния РИС к одной из этих областей $x \in G$.

Следовательно, задача обнаружения атак нейросетями для одного замера параметров может быть сформулирована следующим образом: задана обучающая выборка x^j, $j = 1, \ldots, J$, такая что для любого x^j известно, что $x^j \in G$ или $x^j \notin G$. Требуется построить нейросеть, решающую задачу принадлежности наблюдаемого состояния РИС к одной из областей множества G.

3. Решение задачи методом бинарной нейронной сети

Выше задача обнаружения атак сведена к задаче определения принадлежности некоторого заданного вектора $x \in R^n$ некоторому множеству $G \subset R^n$, причем основная трудность в определении множества G по векторам обучающей выборки состоит в том, что оно может быть многосвязным, невыпуклым и даже иметь фрактальную структуру. Поэтому традиционные подходы к решению аналогичных задач, использующие классические искусственные нейронные сети и опирающиеся на алгоритмы аппроксимации множества G выпуклыми многогранниками, в нашем случае неприменимы. Ниже для решения поставленной задачи предлагается использовать бинарную нейронную сеть [3].

Предположим, что множество G, которому принадлежат векторы атак, является либо объединением конечного числа связных n-мерных не обязательно выпуклых подмножеств n-мерного пространства параметров или, как минимум, измеримым множеством. В этом случае для любого $\varepsilon > 0$ существует покрытие множества G n-мерными параллелепипедами V_m, такое что

$$G \subset \cup V_m = V, \quad V_m \cap V_j = \varnothing$$

при $m \neq j$, и евклидов n-мерный объем всех параллелепипедов, покрывающих множество G, отличается от объема (меры) множества G на величину, не превышающую ε.

Пусть число таких параллелепипедов равно M. Запишем координаты m-го параллелепипеда V_m в виде

$$b_{mi} \leqslant x_i < c_{mi}, \quad i = 1, \ldots, n; \quad m = 1, \ldots, M. \tag{1}$$

Рассмотрим нелинейную бинарную функцию Хевисайда: $\sigma(s) = 1$ при $s \geqslant 0$ и $\sigma(s) = 0$ при $s < 0$. С ее помощью построим бинарную нейронную сеть, позволяющую с любой степенью точности отделить точки множества G от остальных точек пространства R^n. Пусть вектор $x = (x_1, \ldots, x_n)^{\mathrm{T}}$ является входом сети. Тогда первый скрытый слой будет содержать nM узлов y, имеющих следующие выходы

$$\begin{gathered} y_{mi} = y_{mi}(x_i) = \sigma(x_i - b_{mi}) - \sigma(x_i - c_{mi}), \\ i = 1, \ldots, n; \quad m = 1, \ldots, M. \end{gathered} \tag{2}$$

Ясно, что $y_{mi} = 1$ тогда и только тогда, когда имеет место (1). В противном случае $y_{mi} = 0$. Второй скрытый слой содержит M узлов z, имеющих следующие выходы

$$z_m = z_m(x) = \sigma\left(\sum_{i=1}^{n} y_{mi} - n + \frac{1}{2}\right), \quad m = 1, \ldots, M. \tag{3}$$

Из выражений (2) и (3) видно, что $z_m = 1$ тогда и только тогда, когда вектор $x = (x_1, \ldots, x_n)^{\mathrm{T}} \in V_m$. В противном случае $z_m = 0$. Другими словами, второй скрытый слой сети отделяет параллелепипед V_m от других точек пространства R^n. Выходом сети является в данном случае единственный выход u, такой что

$$u = u(x) = \sigma\left(\sum_{m=1}^{M} z_m - \frac{1}{2}\right). \tag{4}$$

Из выражений (2)–(4) видно, что $u = 1$ тогда и только тогда, когда вектор

$$x = (x_1, \ldots, x_n)^{\mathrm{T}} \in V.$$

В противном случае $u = 0$. Другими словами, бинарная сеть (2)–(4) обладает способностью отделять множество V, являющееся объединением параллелепипедов V_m и аппроксимирующее с любой точностью исходное множество G, от других точек пространства R^n. Таким образом, нами доказана теорема существования бинарной сети, обладающей требуемыми свойствами.

Теорема. *Любое измеримое множество $G \in R^n$ может быть отделено от своего дополнения в R^n с помощью бинарной нейронной сети* (2)–(4) *с любой степенью точности, т. е. для любого $\varepsilon > 0$ существуют целое M и вещественные b_{mi} и c_{mi}, $i = 1, \ldots, n$; $m = 1, \ldots, M$, такие что мера множества ошибок первого и второго рода меньше ε, т. е.*

$$\operatorname{mes}\,(x \in G : u = 0 \cup x \notin G : u = 1) < \varepsilon. \tag{5}$$

Кроме того, формулы (2)–(4) дают алгоритм построения такой сети. Заметим, что веса переходов от входа к первому скрытому слою, от первого скрытого слоя ко второму скрытому слою и от второго скрытого слоя к выходу заданы. Неизвестными являются только сдвиги b_{mi} и c_{mi} между входом и первым скрытым слоем, а также число узлов M во втором скрытом слое.

4. Алгоритм обучения бинарной нейронной сети

Разобьем всю обучающую выборку из N векторов на два множества векторов: множество P, состоящее из M векторов $x^m = (x_1^m, \ldots, x_n^m)^{\mathrm{T}}$, $m = 1, \ldots, M$, принадлежащих изучаемому классу атак; и множество Q, состоящее из K векторов $v^k = (v_1^k, \ldots, v_n^k)^{\mathrm{T}}$, $k = 1, \ldots, K$, принадлежащих либо другим классам атак, либо вообще не являющихся атаками. Количество узлов во втором скрытом слое бинарной нейронной сети (2)–(4) примем равным M. Для каждого вектора x^m, $m = 1, \ldots, M$ из множества P, характеризующего изучаемый класс атак, определим его расстояние до второго множества Q:

$$r_m = \operatorname{dist}(x^m, Q) = \min \rho(x^m, v^k), \quad k = 1, \ldots, K. \tag{6}$$

В соответствии с формулой (6) определим следующим образом сдвиги b_{mi} и c_{mi} между входом и первым скрытым слоем бинарной нейронной сети (2)–(4):

$$b_{mi} = \frac{x_i^m - \alpha r_m}{\sqrt{n}}, \quad c_{mi} = \frac{x_i^m + \alpha r_m}{\sqrt{n}}, \qquad 0 < \alpha \leqslant 1, \quad i = 1, \ldots, n; \quad m = 1, \ldots, M. \tag{7}$$

Значение параметра α в формулах (7) может быть выбрано достаточно произвольно. Оно должно определяться особенностями решаемой задачи и конкретными значениями векторов обучающей выборки. Можно положить также $\alpha = 1$.

Решение задачи о принадлежности некоторого вектора $x \in R^n$ множеству G атак заданного класса с использованием обученной бинарной нейронной сети (2)–(4) происходит следующим образом. Координаты вектора x подаются на входы сети, и по формулам (2)–(4) вычисляется выход сети $u(x)$. Если $u(x) = 1$, то считаем, что вектор x принадлежит заданному классу атак. В противном случае, когда $u(x) = 0$, считаем, что вектор x не принадлежит заданному классу атак, т. е. в этом случае он характеризует либо атаку другого класса, либо отсутствие атаки вообще.

5. Заключение

Построенная бинарная нейронная сеть (2)–(4) обладает по крайней мере двумя важными преимуществами по сравнению с классическими нейронными сетями (например, с многослойными персептронами):

1) она применима к решению задач, множества входной информации которых имеют сложную многосвязную и даже фрактальную структуру;
2) метод ее обучения является прямой вычислительной процедурой и не сводится к поиску глобального экстремума какой-либо сложной нелинейной функции, что не накладывает никаких принципиальных ограничений на размерность задачи.

Литература

1. *Зима В.*, *Молдовян А.*, *Молдовян Н.* Безопасность глобальных сетевых технологий. СПб.: БХВ-Петербург, 2001.
2. *Лукацкий А.* Обнаружение атак. СПБ.: БХВ-Петербург, 2000.
3. *Магницкий Н. А.* Некоторые новые подходы к построению и обучению искусственных нейронных сетей // Нелинейная динамика и управление. Вып. 1 / Под ред. Емельянова С. В., Коровина С. К. М.: ФИЗМАТЛИТ, 2001. С. 87–98.