Динамические системы

Использование методов хаотической динамики для обнаружения атак на ресурсы распределенных информационных систем

Н. А. Магницкий

Аннотация. В работе предложен оригинальный метод анализа, классификации и распознавания информационных атак в компьютерных сетях с использованием аппроксимаций траекторий атак нелинейными хаотическими системами обыкновенных дифференциальных уравнений. Предлагаемый подход позволяет свести большеразмерную задачу классификации и распознавания векторных временных рядов атак в исходном пространстве параметров распределенной информационной системы к маломерной задаче классификации и распознавания наборов параметров (коэффициентов) систем ОДУ.
Ключевые слова: *компьютерные атаки, нелинейные хаотические системы, распознавание образов.*

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. 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.

Состояние РИС в момент времени t характеризуется вектором наблюдаемых параметров $x(t) = (x_1(t), x_2(t), \dots, x_m(t))^T$. В работах [3, 4] каждая конкретная атака определенного класса рассматривалась как траектория в m-мерном пространстве параметров: $L_j = (x(t), x(t+\tau), \dots, x(t+n\tau))$, где τ — период замеров значений параметров. Множество траекторий всех атак данного класса обозначалось через $L = (L_1, L_2, \dots, L_l)$. Пусть L^* – наблюдаемая в процессе функционирования РИС траектория. Тогда если $L^* \in L$, то на РИС осуществлена атака. Таким образом, для обнаружения атаки необходимо построить множество L и определить принадлежность наблюдаемой траектории L^* к этому множеству.

Поставленную задачу обнаружения атак переформулируем следующим образом: в пространстве параметров R^m задана обучающая выборка J временных векторных рядов $x^j = (x_1^j, x_2^j, \dots, x_n^j)$, $x_k^j = x^j(t_k)$, $k = 1, \dots, n$, $j = 1, \dots, J$ такая, что для любого x^j известно, что $x^j \in L$ или $x^j \notin L$. Пусть данному классу атак L принадлежит l временных векторных рядов обучающей выборки. Требуется построить алгоритм, решающий задачу принадлежности (или не принадлежности) наблюдаемого векторного временного ряда x^* классу атак L.

Таким образом, задача сводится к анализу, классификации и распознаванию временных векторных рядов. В течение длительного времени к анализу временных рядов подходили с позиций математической статистики. Использовался соответствующий

математический аппарат, включающий понятия последовательностей случайных величин, случайных процессов, статистических моделей, стохастических дифференциальных уравнений. При этом сначала методами регрессионного анализа решалась задача идентификации, т. е. делалась попытка ответить на вопрос, каковы параметры системы, породившей данный временной ряд. Считалось, что эти параметры могут помочь идентифицировать (распознать) систему, отличить ее от других. При этом, как совершенно справедливо отмечено в [5], оставалось совершенно непонятным, имеют ли полученные регрессионные уравнения какое-либо отношение к действительным уравнениям динамики системы или нет.

С начала 80-х годов после публикации знаменитой работы Такенса [6] (см. также [7]) широкое распространение получили алгоритмы анализа временных рядов методами нелинейной динамики. Идея Такенса состоит в представлении скалярного временного ряда простой хаотической динамической системой, однако неизвестно, какую координатную ось или скалярную функцию координат многомерной хаотической системы следует выбрать в качестве «наблюдаемой» для сопоставления с исходным временным рядом. Неизвестен также способ выбора размерности пространства динамической системы. В настоящей работе предложен свободный от этих недостатков метод аппроксимации векторного временного ряда решением специальной нелинейной хаотической системы обыкновенных дифференциальных уравнений (ОДУ). При этом векторные временные ряды, соответствующие одному классу атак на ресурсы РИС, аппроксимируются хаотическими системами ОДУ с близкими наборами своих параметров (коэффициентов), что дает возможность свести большеразмерную задачу классификации и распознавания векторных временных рядов в исходном пространстве параметров РИС и временных отсчетов к задаче классификации и распознавания наборов параметров (коэффициентов) систем ОДУ в маломерном пространстве этих параметров.

1. Аппроксимация векторного временного ряда решением нелинейной системы ОДУ

Пусть исходный векторный временной ряд $x(t)$ задан своими значениями x_k в точках t_k, $k=1,\dots,n$. Фазовое пространство R^m предлагается ввести из переменных — координат вектора временного ряда. При этом численные значения $x_{ik} = x_i(t_k)$, $i = 1,\dots,m$ задают в фазовом пространстве R^m некоторую сеточную траекторию. Эту траекторию будем аппроксимировать в фазовом пространстве R^m решением $x(t,\mu) = (x_1(t,\mu),\dots,x_m(t,\mu))^{\mathrm{T}}$ системы нелинейных хаотических ОДУ

$$\dot{x} = F(x,\mu),\ x \in R^m,\ \mu \in R^p, \tag{1}$$

на промежутке $t \in [0, t_n]$. Заметим, что в отличие от теории Такенса, в этом случае не возникают проблемы, связанные как с выбором наблюдаемой, так и с выбором пространства вложения и его размерности. Единственным условием является условие $m > 2$, что обеспечивает существование хаотической динамики в системе (1). Заметим, что в случае выбора правой части системы (1) с зависящими от времени периодическими коэффициентами метод будет работать даже при $m = 2$ (см. [8]).

Вектор параметров системы $\mu \in R^p$ и вид аппроксимирующей правой части $F(x,\mu)$ в уравнении (1) можно определить следующим образом. В качестве критерия близости сеточной функции $x(t_k)$ и решения $x(t,\mu)$ уравнения (1) используем расстояние между этими функциями в сеточной норме. Тогда для определения вектора параметров $\mu \in R^p$ воспользуемся необходимым условием экстремума функционала

$$\Phi(x,\mu) = \sum_{i=1}^{m}\sum_{k=1}^{n}\bigl(x_{ik} - x_i(t_k,\mu)\bigr)^2,$$

которое равносильно системе из p уравнений

$$\sum_{i=1}^{m}\sum_{k=1}^{n}\bigl(x_{ik} - x_i(t_k,\mu)\bigr)\cdot\frac{\partial x_i(t_k,\mu)}{\partial \mu_j} = 0,\quad j = 1,\dots,p, \tag{2}$$

нелинейных относительно неизвестного вектора параметров μ. Входящие в последнее уравнение функции

$$u_{ij} = \frac{\partial x_i(t_k,\mu)}{\partial \mu_j}$$

являются решением матричного линейного неоднородного дифференциального уравнения

$$\dot{U} = PU + Q,$$

с начальным условием $U(0) = O_{m\times p}$, где $O_{m\times p}$ — нулевая матрица, а матрицы $P_{m\times m}$ и $Q_{m\times p}$ — производные от функции $F(x,\mu)$ в (1) соответственно по переменным x и μ. Алгоритмы решения подобных уравнений разработаны в [9, 10].

Векторная функция $F(x,\mu) = (f_1,\ f_2,\ \dots,\ f_m)^{\mathrm{T}}$ в правой части системы (1) может быть выбрана, например, в виде полиномов по переменным x_i

$$f_i = a_i + \sum_{j=1}^{m} a_{ij}x_j + \sum_{j=1}^{m}\sum_{k=1}^{m} a_{ijk}\,x_jx_k,\quad i = 1,\dots,m.$$

Тогда координатами вектора параметров μ являются коэффициенты a_i, a_{ij}, a_{ijk}. Возможны и другие классы функций для правой части аппроксимирующего уравнения (1).

В частном случае, когда решение $x(t, \mu)$ системы (1) может быть представлено в виде полинома по переменным x_i, линейного по параметрам системы, задача нахождения этих параметров сводится к решению системы из p линейных алгебраических уравнений

$$A\mu = b, \tag{3}$$

где элементы

$$a^*_{ij} = \sum_{k=1}^{n} \frac{\partial x_i(t_k, \mu)}{\partial \mu_i} \cdot \frac{\partial x_i(t_k, \mu)}{\partial \mu_j}$$

матрицы A и вектор b с координатами

$$b_j = \sum_{i=1}^{m} \sum_{k=1}^{n} x_{ik} \cdot \frac{\partial x_i(t_k, \mu)}{\partial \mu_j}$$

не зависят от параметров и выражаются лишь через переменные x_i, $i = 1, \ldots, m$.

2. Решение задачи классификации и распознавания атак

Выше задача обнаружения атаки данного класса на РИС сведена к задаче определения принадлежности некоторого набора параметров (коэффициентов) системы ОДУ, аппроксимирующей наблюдаемый векторный временной ряд параметров атаки, к множеству G в R^p, к которому принадлежат параметры (коэффициенты) систем ОДУ для векторов атак данного класса из обучающей выборки. Для решения последней задачи в пространстве R^p параметров (коэффициентов) систем ОДУ достаточно малой размерности может быть успешно использована предложенная автором в [3] бинарная нейронная сеть, применимая к решению задач, множества входной информации которых имеют сложную многосвязную и даже фрактальную структуру. В [3] дано подробное описание бинарной нейронной сети, метода ее обучения на обучающей выборке и алгоритм распознавания принадлежности некоторого вектора параметров (коэффициентов) системы ОДУ, аппроксимирующей новый наблюдаемый векторный временной ряд, множеству G атак данного класса L.

3. Заключение

В работе предложен метод анализа, классификации и обнаружения информационных атак на ресурсы РИС с использованием аппроксимаций векторов параметров атак решениями нелинейных хаотических систем ОДУ. При этом векторные временные ряды, соответствующие одному классу атак на ресурсы РИС, аппроксимируются хаотическими системами ОДУ с близкими наборами своих параметров (коэффициентов), что дает возможность свести большеразмерную задачу классификации и распознавания векторных временных рядов атак в исходном пространстве параметров РИС и временных отсчетов к задаче классификации и распознавания наборов параметров (коэффициентов) систем ОДУ в маломерном пространстве этих параметров (коэффициентов).

Работа поддержана программой ОНИТ РАН 4 (проект 2.5).

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.* Использование бинарной нейронной сети для обнаружения атак на ресурсы распределенных информационных систем // Труды ИСА РАН, 2008,33 с. 200–205.
4. *Магницкий Н. А.* Использование иммунной сети для обнаружения атак на ресурсы распределенных информационных систем // Информационные технологии и вычислительные системы, 2009,3 с. 22–26.
5. *Малинецкий Г. Г.*, *Потапов А. Б.* Современные проблемы нелинейной динамики. М.: URSS, 2004. 320 с.
6. *Takens F.* Detecting strange attractors in turbulence Commun.– Lect. Notes in Math. Berlin: Springer. 1981, 898, p. 336–381.
7. *Takens F.* Estimation of dimension and order of time series. Nonlinear dynamical systems and chaos, 19, 1996.
8. *Магницкий Н. А.* Теория динамического хаоса. М.: Ленанд/URSS, 2011, 320с.
9. *Магницкий Н. А.*, *Сидоров С. В.* Управление хаосом в нелинейных динамических системах. Дифференциальные уравнения, 1998, т. 34, № 11. с. 1501–1509.
10. *Магницкий Н. А.*, *Сидоров С. В.* Локализация и стабилизация неустойчивых решений хаотических динамических систем. В сб. Нелинейная динамика и управление. Под ред. С. В. Емельянова и С. К. Коровина, Вып. 1, М.: ФИЗМАТЛИТ, 2001 с. 217–246.

Магницкий Николай Александрович. Зав. лабораторией ИСА РАН. Д. ф.-м. н., профессор. Окончил в 1974 г. МГУ. Количество печатных работ: более 200 (в т. ч. 6 монографий). Область научных интересов: интегральные и дифференциальные уравнения, нелинейные хаотические динамические системы, нейронные сети, математическое моделирование. E-mail: nikmagn@gmail.com