в форме структурных автоматов на основе имеющихся частично определенных моделей подсистем и компонент исходной системы; разработка методов выбора мест получения контрольной и диагностической информации (используемой при построении множества базовых точек интерполяции), учитывающих прикладное предназначение рассматриваемой системы; разработка классификации методов интерполяции по эффективности использования как средства доопределения математических автоматных моделей систем в зависимости от реализуемых системой алгоритмов и так далее.

Список литературы

[1] *Твердохлебов В. А.* Геометрические образы законов функционирования автоматов. — Саратов: Научная книга, 2008.

[2] *Твердохлебов В. А., Епифанов А. С.* Представление автоматных отображений геометрическими структурами. — Саратов: Наука, 2013.

[3] *Половко А. М., Бутусов П. Н.* Интерполяция. Методы и компьютерные технологии их реализации. — СПб.: БХВ-Петербург, 2004.

[4] *Резчиков А. Ф., Твердохлебов В. А., Иващенко В. А.* Человек. Машина. Среда. — Саратов: Наука, 2013.

[5] *Резчиков А. Ф., Твердохлебов В. А., Иващенко В. А.* Критические ситуации в человеко-машинных системах. — Саратов: Наука, 2015.

[6] *Epifanov A. S.* Regularization and recognition methods of automata models of technical systems // Contemporary Engineering Sciences. — 2015. — Vol. 8, no. 20. — P. 919–926.

[7] *Epifanov A. S.* Methods of interpolation of automatons models of systems // Applied Mathematical Sciences. — 2014. — Vol. 8, no. 81. — P. 4025–4030.

[8] *Епифанов А. С.* Разработка методов распознавания и доопределения автоматных моделей ПЛИС с использованием аппарата геометрических образов автоматов // Компьютерные науки и информационные технологии: матер. Междунар. науч. конф. — Саратов: Наука, 2016. — С. 157–161.

[9] *Резчиков А. Ф., Твердохлебов В. А.* Техническое диагностирование мехатронных систем // Мехатроника, автоматизация, управление. — 2003. —№ 2. — С. 2–6.

Исследование прогнозных моделей динамической системы на примере прогноза инцидентов информационной безопасности

Ермакова А. Ю.[1], Лось А. Б.[2]

[1]*a.alla1105@yandex.ru*, [2]*alexloss2011@mail.ru*

[1]*МТУ(МИРЭА), Москва, Россия*, [2]*МИЭМ НИУ ВШЭ, Москва, Россия*

В статье продолжены исследования методов прогнозирования изменения состояний динамической системы, первоначально заданных в виде табличной функции (узловые точки). Основной метод построения прогнозной модели — поиск непрерывной «аппроксимирующей» функции, наиболее близко отстоящей от заданных узловых точек. В работе рассматриваемый метод применяется для прогнозирования динамики появления инцидентов, приводящих к нарушению информационной безопасности. Полученные результаты могут применяться для оценки рисков нарушения информационной безопасности и вероятного ущерба, которые, в свою очередь, служат для обоснования уровня защищенности информационной системы.

Ключевые слова: прогнозная модель, динамическая система, аппроксимирующая функция, инциденты информационной безопасности

Введение

В данной работе продолжены, начатые ранее в работах [1–5], исследования методов прогнозирования динамики изменения состояний динамической системы, первоначально заданных в виде табличной функции. Основным рассматриваемым методом построения прогнозной модели является поиск непрерывной «аппроксимирующей» функции, наиболее близко отстоящей от заданных табличных состояний системы (узловых точек). В указанных работах прогнозные модели строились для исследования поведения ряда технических и экономических показателей таких, как рост производительности средств вычислительной техники, котировки курсов мировых валют и ряда других показателей. В настоящей статье указанный выше метод применяется к задачам защиты информации, в частности, для исследования динамики появления различных инцидентов, приводящих к нарушению информационной безопасности системы. Прогноз динамики появления инцидентов может быть далее использован при оценке рисков нарушения информационной безопасности (ИБ) конкретной информационной системы (ИС) и, в конечном итоге, для оценки уровня ее защищенности, в частности, для оценки времени ее безопасной эксплуатации [6].

1. Построение прогнозной модели динамической системы

Метод построения прогнозной модели состоит в следующем. Пусть на отрезке $[a, b]$, задана одномерная сетка

$$\{x_i / x_i = x_{i-1} + h_i,\ h_i > 0,\ i = 1, \ldots, n;\quad x_0 = a,\ x_n = b\},$$

в узлах x_i которой заданы значения $y_i = f(x_i)$, $i = 0, 1, \ldots, n$ — соответствующие значения функции $f(x)$. Будем далее рассматривать данную таблицу как множество состояний некоторой динамической системы. При этом величины x_i будут означать моменты времени при наблюдении состояния рассматриваемой системы, а величины y_i — сами эти состояния.

Пусть также для аппроксимации табличных данных выбран некоторый класс функций $F(x, c_0, c_1, \ldots, c_m)$, $m < n$, где $c_0, c_1, \ldots, c_m$ — коэффициенты, выбор значений которых позволяет определить конкретную функцию из выбранного класса. Требуется найти значения коэффициентов $c_0, c_1, \ldots, c_m$, для которых выполнено условие:

$$\Phi(c_0, c_1, \ldots, c_m) = \min \sum_{i=0}^{n} (y_i - F(x_i, c_0, c_1, \ldots, c_m))^2 . \tag{1}$$

Выбранные в соответствии с критерием (1) значения коэффициентов, позволяют определить среди множества функций конкретную функцию, наиболее согласованную с табличными (экспериментальными) данными или, иначе говоря, обеспечивающую наилучшее среднеквадратическое приближение.

Функция $F(x, c_0, c_1, \ldots, c_m)$ называется моделью, а искомые коэффициенты $c_0, c_1, \ldots, c_m$ — параметрами модели. В дальнейшем ограничимся рассмотрением случая, когда модель линейно зависит от параметров и ее можно представить в виде:

$$F(x, c_0, c_1, \ldots, c_m) = c_0\varphi_0(x) + c_1\varphi_1(x) + \cdots + c_m\varphi_m(x) . \tag{2}$$

Здесь $\{\varphi_i(x)\}_{i=1}^{n}$ — множество, так называемых, базисных функций.

Базисные функции могут быть как линейными, так и нелинейными функциями переменной x. Независимо от этого модель (2) остается линейной, поскольку она линейно зависит от модельных параметров $c_0, c_1, \ldots, c_m$.

В качестве базисных функций могут быть выбраны, например, степенные функции:

$$\varphi_0(x) = 1; \quad \varphi_1(x) = x; \quad \varphi_2(x) = x^2; \quad \ldots; \quad \varphi_m(x) = x^m .$$

Тогда модель будет представлять собой полином степени m:

$$F(x, c_0, c_1, \ldots, c_m) = c_0 + c_1x + c_2x^2 + \cdots + c_mx^m. \tag{3}$$

Очевидно, что в качестве базисных функций могут быть использованы и другие функции, необходимо лишь, чтобы они были линейно независимыми.

Таким образом, для линейной модели (2) требуется найти значения параметров $c_0, c_1, \ldots, c_m$, обеспечивающих выполнение условия (1).

С математической точки зрения поставленная задача является задачей поиска минимума функции нескольких переменных, который можно искать, исходя из необходимых условий экстремума для функций нескольких переменных.

С целью построения «аппроксимирующей» функции $F(x, c_0, c_1, \ldots, c_m)$ разработано специализированное программное обеспечение, подробное описание которого приведено в работе [4].

2. Построение прогнозной модели динамики появления инцидентов, приводящих к нарушению информационной безопасности системы

Как было отмечено выше, с точки зрения решения задач защиты информации, весьма важным направлением исследований является изучение динамики появления различных инцидентов, приводящих к нарушению ИБ, в том числе, построение ее прогнозных моделей. На основе прогнозных данных могут быть получены оценки рисков и величина ущерба в результате нарушения ИБ, а также получены оценки уровня защищенности рассматриваемой ИС.

Для построения требуемой модели воспользуемся данными о появлении инцидентов, размещенными на сайте Лаборатории Касперского [7].

В табл. 1 представлена статистика веб-угроз за февраль–март 2018 года, а на рис. 1 — динамика абсолютного числа зафиксированных веб-угроз за этот период.

В эксперименте по прогнозированию веб-угроз для построения прогнозной функции $F(x) = F_1(x)$ использовались данные с 13.02.2018 по 12.03.2018, с целью удобства представления исходные данные поделены на 100, за нулевое значение по оси Ox принято 01.02.2018. В данном эксперименте «аппроксимирующая» функция $F_1(x)$ имеет вид:

$$F_1(x) = 46007.4546 + \frac{3460901.5541}{x^2} + 1336.59334\,x - 5589.6969 \sin\left[\frac{x}{2}\right] . \tag{4}$$

График функции $y = F_1(x)$ представлен ниже на рис. 2.

Таблица 1. Веб-угрозы за период с 13.02.2018 по 12.03.2018

Тип веб-угрозы	% от общего числа угроз
Trojan.Script.Generic	58.8 %
Trojan.Script.Agent.gen	20.6 %
Trojan.JS.Miner.m	8.0 %
Trojan-Clicker.HTML.Iframe.dg	2.4 %
Trojan.Win32.Miner.ays	1.5 %
Trojan.Win64.Shelma.a	0.9 %
Trojan.JS.Agent.eak	0.9 %
Trojan.JS.Miner.o	0.4 %
Trojan-Dropper.VBS.Agent.bp	0.4 %
Packed.Multi.MultiPacked.gen	0.3 %

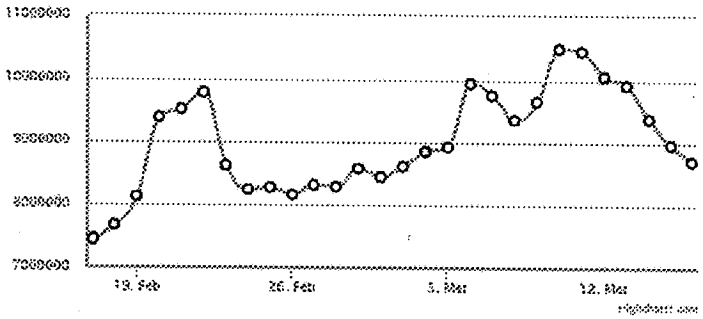

Рис. 1. График динамики абсолютного числа зафиксированных Лабораторией Касперского веб-угроз за период с 13.02.2018 г. по 12.03.2018 г. График построен для всего мира

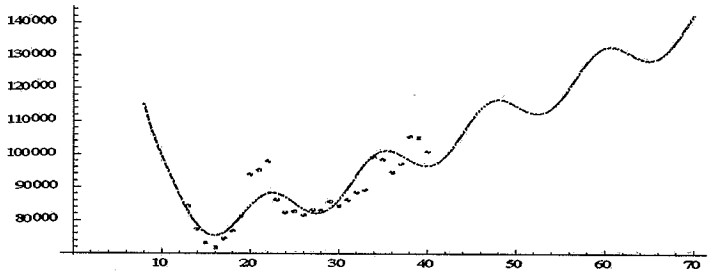

Рис. 2. График функции $F_1(x)$

Таблица 2. Прогноз веб-угроз на период с 13.03.2018 по 11.04.2018

Период	Прогнозное значение	Период	Прогнозное значение	Период	Прогнозное значение
13.03.2018	9729460	23.03.2018	11349700	02.04.2018	13290300
14.03.2018	9942970	24.03.2018	14252800	03.04.2018	13203500
15.03.2018	10271600	25.03.2018	11260500	04.04.2018	13061563
16.03.2018	10665540	26.03.2018	11402500	05.04.2018	12931200
17.03.2018	11058645	27.03.2018	11675600	06.04.2018	12876500
18.03.2018	11385640	28.03.2018	12044656	07.04.2018	12942800
19.03.2018	11597300	29.03.2018	12450900	08.04.2018	13146100
20.03.2018	11673800	30.03.2018	12826800	09.04.2018	13468700
21.03.2018	11624700	31.03.2018	13112100	10.04.2018	13863800
22.03.2018	11496100	01.04.2018	13268700	11.04.2018	14566900

В табл. 2 приведены результаты эксперимента по построению прогнозных значений веб-угроз на период с 13.03.2018 по 11.04.2018.

В следующем эксперименте прогнозная модель строилась для изучения динамики появления уязвимостей ИС. В табл. 3 приведена статистика уязвимостей, размещенная на сайте Лаборатории Касперского, а на рис. 3 представлена динамика абсолютного числа зафиксированных уязвимостей.

В эксперименте для построения прогнозной функции $F'(x) = F'_2(x)$ использовались данные за период с 13.02.2018 г. по 12.03.2018 г., за нулевое значение по оси Ox принята дата 01.02.2018 г.

В данном эксперименте аппроксимирующая функция $F_2(x)$ имела вид:

$$F'_2(x) = 2463152.8794 - \frac{6.9581 \times 10^7}{x^2} + 19435.0639\,x - 801271.8964\,Log[x] + 20916.829\sin[x]\,.$$

График функции $y = F_2(x)$ представлен ниже на рис. 4.

В табл. 4 приведены результаты эксперимента по построению прогнозных значений уязвимостей на период с 13.03.2018 по 11.04.2018

3. Оценка времени безопасной эксплуатации информационной системы

Применим полученные выше результаты для оценки уровня защищенности информационной системы. С этой целью воспользуемся, введенным ранее в работе [6], понятием времени безопасной эксплуатации ИС.

Пусть

- $y_1, y_2, \ldots, y_m$ — множество угроз ИС;
- $p_1, p_2, \ldots, p_m$ — вероятности успешной реализации угроз ИС;
- $z_1, z_2, \ldots, z_m$ — ущерб ИС от реализации данных угроз.

Таблица 3. Уязвимости за период с 13.02.2018 по 12.03.2018

Тип уязвимости	
Exploit.Win32.ShadowBrokers.a	10,0 %
Exploit.Win32.ShadowBrokers.z	2,6 %
Exploit.Win64.ShadowBrokers.c	6,1 %
Exploit.Win32.ShadowBrokers.ab	6,0 %
Exploit.Win32.ShadowBrokers.aa	5,9 %
Exploit.Win64.ShadowBrokers.d	5,9 %
Exploit.Win32.ShadowBrokers.ad	5,9 %
Exploit.Java.Generic	5,8 %
Exploit.Script.Generic	4,5 %
Exploit.Script.Blocker	3,5 %

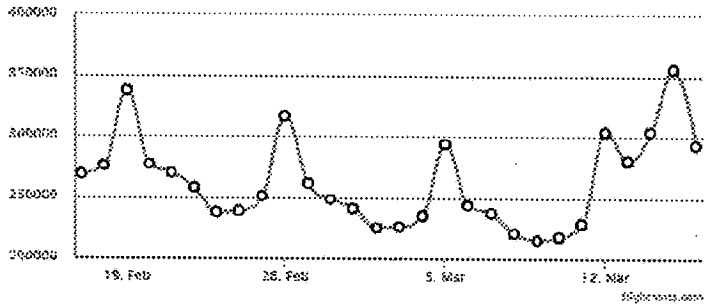

Рис. 3. График динамики абсолютного числа зафиксированных Лабораторией Касперского уязвимостей за период с 13.02.2018 г. по 12.03.2018 г. График построен для всего мира

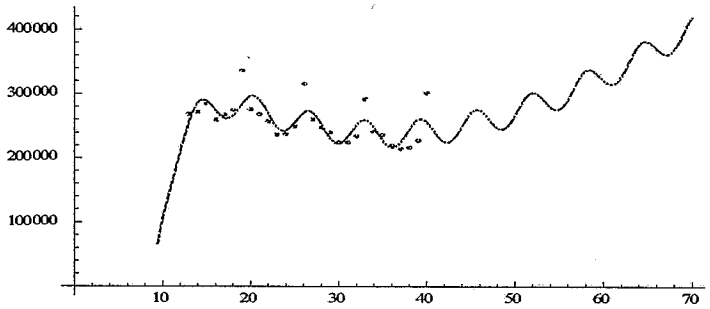

Рис. 4. График функции $y = F_2(x)$

Таблица 4. Прогноз уязвимостей на период с 13.03.2018 по 11.04.2018

Период	Прогнозное значение	Период	Прогнозное значение	Период	Прогнозное значение
13.03.2018	239699	23.03.2018	291147	02.04.2018	315856
14.03.2018	225920	24.03.2018	302660	03.04.2018	327607
15.03.2018	230086	25.03.2018	295438	04.04.2018	353754
16.03.2018	250560	26.03.2018	280835	05.04.2018	376857
17.03.2018	270996	27.03.2018	277204	06.04.2018	382438
18.03.2018	275362	28.03.2018	293018	07.04.2018	372285
19.03.2018	262671	29.03.2018	319075	08.04.2018	362805
20.03.2018	247881	30.03.2018	336951	09.04.2018	369934
21.03.2018	248135	31.03.2018	335935	10.04.2018	394486
22.03.2018	266991	01.04.2018	322870	11.04.2018	421394

В этом случае величина риска $R_{риск}$ при нарушении информационной безопасности (ИБ) определяется соотношением:

$$R_{риск} = \sum_{i=1}^{m} p_i z_i \,. \tag{5}$$

ИС считается защищенной, если выполнено условие:

$$R_{риск} \leqslant R_0 \,, \tag{6}$$

где R_0 — величина максимально допустимых потерь при нарушении ИБ.

В случае, если удается построить модель риска, зависящую от времени

$$R_{риск} = R(t) \,,$$

то соотношение (6) позволяет вычислить величину T_0 — время безопасной эксплуатации ИС из уравнения:

$$R(t) = \sum_{i=1}^{m} p_i(t) z_i(t) = R_0 \,. \tag{7}$$

В настоящее время проблема применения указанного подхода заключается, в основном, в отсутствии данных по инцидентам для конкретной ИС. Практика показывает, что многие руководители организаций не дают разрешения на публикацию указанных данных, считая, что это может повредить имиджу организации.

Применим имеющиеся данные Лаборатории Касперского для иллюстрации предлагаемого подхода к оценке уровня защищенности ИС. В соответствии с рассматриваемым выше методом построения прогнозной модели появления инцидентов, функция их числа имеет вид:

$$f(t) = \alpha_1 + \frac{\alpha_2}{t^2} + \alpha_3 t + \alpha_4 \sin\frac{t}{2}, \tag{8}$$

где α_1, α_2, α_3, α_4 — соответствующие константы (с учетом деления исходных данных на 100).

Обозначим через N — общее число информационных систем, для которых проводилось наблюдение. Тогда среднее число инцидентов, приходящееся на одну ИС в момент времени t можно оценить величиной:

$$\frac{1}{N} f(t) = \frac{1}{N}\left(\alpha_1 + \frac{\alpha_2}{t^2} + \alpha_3 t + \alpha_4 \sin\frac{t}{2}\right). \tag{9}$$

Для простоты вычислений будем считать, что величина ущерба при возникновении инцидентов одинаковая и равна z.

Тогда, уравнение (7) принимает вид:

$$\frac{1}{N}\left(\alpha_1 + \frac{\alpha_2}{t^2} + \alpha_3 t + \alpha_4 \sin\frac{t}{2}\right) z = R_0. \tag{10}$$

Заметим далее, что для функции $f(t)$, при $t \geqslant 1$, имеет место неравенство:

$$f(t) \leqslant \alpha_5 + \alpha_3 t,$$

где $\alpha_5 = \alpha_1 + \alpha_2 + \alpha_4$.

Тогда, очевидно, оценка для величины T_0 — времени безопасной эксплуатации ИС может быть найдена из уравнения:

$$\frac{1}{N}(\alpha_5 + \alpha_3 t) z = R_0,$$

откуда получаем выражение для оценки величины T_0:

$$T_0 = \frac{NR_0 - z\alpha_5}{z\alpha_3}.$$

При наличии данных по инцидентам для конкретной ИС указанный подход может быть применен для оценки уровня ее защищенности и оценки времени ее безопасной эксплуатации.

Заключение

В настоящей статье исследуется метод построения прогнозной модели состояний динамической системы на примере прогнозирования интенсивности инцидентов, возникающих в информационной системе и приводящих к нарушению информационной безопасности. По данным Лаборатории Касперского, построена аппроксимирующая функция, позволяющая прогнозировать среднее число инцидентов, приводящих к нарушению безопасности информационной системы. На основании полученных данных построена оценка для времени безопасной эксплуатации информационной системы.

Список литературы

[1] *Ермакова А. Ю.* Оценка качества прогнозирования динамики изменения валютных курсов на основе построения аппроксимирующих функций // Качество. Инновации. Образование. — 2013. — № 2(93). — С. 71–79.

[2] *Ермакова А. Ю.* Исследование качества прогнозирования биржевых курсов драгоценных металлов // Качество. Инновации. Образование. — 2014. — № 1(104). — С. 49–56.

[3] *Ермакова А. Ю.* Построение прогнозной модели динамики изменения цен на древесину // Вестник Московского государственного университета леса — Лесной Вестник. — 2016. — Т. 20, № 6. — С. 88–96.

[4] *Ермакова А. Ю.* Разработка методов прогнозирования на примере анализа средств вычислительной техники // Промышленные АСУ и контроллеры. — 2017. — № 1. — С. 28–34.

[5] *Ермакова А. Ю.* Построение прогнозной модели изменения показателей на примере прогнозирования мировых цен на древесину // Традиционная и инновационная наука: история, современное состояние, перспективы: сб. ст. Междунар. науч.-практ. конф. В 5 ч. — Ч. 3. — Уфа: Аэтерна, 2018. — С. 37–42.

[6] *Кабанов А. С., Лось А. Б., Трунцев В. И.* Временная модель оценки риска нарушения информационной безопасности // Доклады ТУСУР. — 2012. — № 1, Ч. 2. — С. 87–91.

[7] Сайт Лаборатории Касперского. — URL: `https://securelist.ru/statistics/`.