УДК 004.056:5

Обнаружение компьютерных атак на основе функционального подхода

А. С. Шабуров, А. А. Миронова
Пермский национальный исследовательский политехнический университет
Россия, 614990, Пермь, Комсомольский пр., 29
shans@at.pstu.ru; 89128876457

Приводится краткий обзор существующих методов противодействия и обнаружения компьютерных атак на информационные системы. Предполагается, что в большинстве систем защиты решение задачи сводится не к обнаружению атак, а лишь к обнаружению их последствий, а адаптивность к неизвестным атакам в большинстве существующих систем обнаружения атак, отсутствует. Предлагается математическая модель обнаружения компьютерных атак на основе функционального подхода. Использование функционального подхода предполагает построение полного множества безопасных состояний информационной системы и обнаружение на этой основе признаков компьютерных атак.

Ключевые слова: *компьютерная атака; защита информации; нейронные сети; метод обнаружения аномалий; анализ сигнатур; теория распознавания образов; функциональный подход.*

Введение

В последнее время значительно обострилась проблема компьютерных атак на информационные системы. За последний год около 67 % компаний по всему миру столкнулись с подобными ситуациями. При этом более трети (37 %) не обладают достаточными возможностями и ресурсами для борьбы с такими угрозами, а злоумышленники стали более изощренными, чем когда-либо и находятся в постоянном поиске уязвимостей во всей технологической цепочке, включая людей и процессы [1].

По данным индекса критичности утечек данных BLI (Breach Level Index), основной целью киберпреступников при осуществлении атак в 2014 г. стали персональные данные. На долю подобных атак пришлось 54 % всех инцидентов, что больше, чем в любой другой категории, в том числе больше числа инцидентов с кражей финансовых данных. Кроме того, на долю утечек, преследовавших цель хищения, пришлось около трети наиболее значимых взломов, которые были классифицированы в рамках индекса BLI как катастрофические [2].

Большая часть кибератак совершается в отношении финансового сектора и госструктур, отмечают в Group-IB. На 4 % российских банков за год были совершены успешные хакерские атаки. Жертвами стали сайты Центрального банка России, ВТБ24, Альфа-Банка, Бинбанка. В отношении госсектора компьютерные атаки чаще всего совершаются в целях промышленного шпионажа, а самые масштабные за последние годы были связаны с конкретными событиями, такими как Олимпийские игры в Сочи, референдум в Крыму, ситуацией в Украине.

Анализ публикаций подтверждает значительную актуальность тематики в области разработки методов противодействия компьютерным атакам на информационные и телекоммуникационные системы, а также необходимость поиска наиболее эффективных способов защиты информации, основанных на оптимальных и постоянно совершенствующихся алгоритмах противодействия угрозам безопасности [3].

Анализ и проблема существующих методов

В целом, большинство применяемых на сегодняшний день методов противодействия компьютерным атакам можно разделить на два класса: методы обнаружения существующих компьютерных атак и методы прогнозирования возникновения потенциальных компьютерных атак [4].

В свою очередь, методы, направленные на обнаружение существующих компьютерных атак, также подразделяются на методы обнаружения аномалий (или аномальных отклонений) и методы анализа сигнатур (методы обнаружения злоупотреблений).

Методы обнаружения аномалий основываются на выявлении отклонений от нормального поведения системы и позволяют выявить неизвестные ранее компьютерные атаки. К данной группе методов относятся:

1. *Статистический анализ.* В течение некоторого заданного промежутка времени для рассматриваемой информационной системы (ИС) формируется набор статистических характеристик, описывающих нормальное поведение данной системы. В случае если поведение системы отклоняется от имеющихся характеристик, ее поведение считаются аномальным и рассматривается как атака.

2. *Кластерный анализ.* Методы данной группы основываются на разбиении множества наблюдаемых векторов - свойств системы на кластеры. А затем среди полученных кластеров выбирают те, которые описывают нормальное поведение исследуемой системы.

3. *Нейронные сети.* В течение некоторого периода времени происходит обучение нейронной сети, когда поведение ИС считается нормальным. После процесса обучения происходит запуск нейронной сети в режим е распознавания. Наличие атаки определяется отклонением в распознавании нормального поведения во входном потоке.

4. *Иммунные сети.* Аналогично с нейронной сетью иммунную сеть можно использовать для распознавания образов. В процессе применения данного метода формируются антитела, которые сопоставляют свойства атак с характеристиками, заложенными в них, и распознают данное событие как атакующее воздействие.

5. *Экспертные системы.* В экспертных системах информация о нормальном поведении хранится в виде правил, а наблюдаемое поведение представляется в виде фактов. На основании этих фактов и правил принимается решение о соответствии наблюдаемого поведения "нормальному" либо о наличии аномалии.

6. *Поведенческая биометрия.* В основе данных методов лежит гипотеза о различии "почерка" работы с интерфейсами ввода/вывода для различных пользователей. На базе построенного профиля нормального поведения для данного пользователя обнаруживаются отклонения от этого профиля, вызванные попытками других лиц работать с клавиатурой или другими физическими устройствами ввода.

Методы анализа сигнатур используются для распознавания известных компьютерных атак. Основой данных методов является сравнение поведения информационной системы с описанием известной атаки. Если оно совпадает, то поведение объекта считается атакой. К данной категории методов относятся:

1. *Анализ состояний систем.* Данный метод предполагает описание процесса функционирования исследуемой системы как ориентированный граф, вершинами которого являются состояния системы, а ребрами – переходы между ними. Некоторые из путей в рассматриваемом графе помечаются как недопустимые. В этом случае конечное состояние каждого из таких путей представляет собой потенциальную угрозу, а обнаружение подобного рода недопустимых путей означает успешное обнаружение атаки.

2. *Графы сценариев атак.* Для построения графа атак необходимо формализовать понятие атаки, разработать формальный язык представления атак и ИС в целом. После этого необходимо построить и проанализировать граф атак и при наличии последовательности наблюдаемых в системе действий сигнализировать об атаке.

3. *Нейронные сети.* Нейронные сети могут быть использованы для обнаружения атак в ИС на основе метода анализа сигнатур. При этом сначала происходит обучение данной нейронной сети на примерах существующих атак на защищаемую систему, а затем осуществляется сравнение и выявление принадлежности наблюдаемого поведения к одному из классов изученных атак.

4. *Иммунные сети.* Анализ сигнатур может быть рассмотрен как один из возможных способов использования иммунных систем, как и в случае с нейронными сетями.

5. *Методы, основанные на спецификациях.* Для функционирования данного метода необходимо сформировать множество всех возможных атакующих воздействий в виде спецификаций атак. Совпадение текущего события со спецификацией расценивается как атака.

6. *Сигнатурные методы.* Для данного метода необходимо формирование некоторого алфавита для описания наблюдаемых событий системы и построение множества правил – сигнатур с использованием сформированного алфавита. Совпадения характеристик события ИС с одной из сигнатур свидетельствует о наличии атаки.

Наибольший интерес для исследования представляют собой методы обнаружения аномалий, основанные на выявлении отклонений от нормального поведения системы и позволяющие выявить неизвестные ранее компьютерные атаки.

Как правило, распознавание компьютерных атак в динамике функционирования информационной системы может быть представлено на основе системного анализа пространства параметров процессов в системе по установленным правилам и выявление тех параметров, которые характеризуют действие атаки. В свою очередь, системное описание способов и средств защиты информации, направленных на противодействие компьютерным атакам, рационально осуществить на основе теории распознавания образов [5], в соответствии с которой объекты компьютерных атак могут быть интерпретированы распознаваемыми образами пространства их признаков.

В соответствии с данной теорией, компьютерная атака является образом необходимо распознаваемым в ходе процесса сбора, хранения, обработки и передачи информации в информационной системе. При попытках нарушителя воздействовать на нее с целью вывода из строя или снижения эффективности применения.

Словарь признаков компьютерных атак может содержать количественные и качественные признаки, которые декомпозируются на детерминированные признаки атак, распознаваемые сигнатурными методами обнаружения атак, вероятностные признаки атак, распознаваемые методами анализа аномальных отклонений в ИС, логические признаки атак, распознаваемые методами функционального анализа. При этом исследование методов обнаружения компьютерных атак на информационные системы показывает, что зачастую решение задачи сводится не к обнаружению атак, а лишь к обнаружению их последствий. Таким образом, адаптивность к неизвестным атакам в большинстве существующих систем обнаружения атак в целом, отсутствует [6].

Теоретико-множественная модель

Решение проблемы обнаружения компьютерной атаки предполагается на основе функционального подхода [6],

Функциональное представление информационной системы предполагает ее рассмотрение с точки зрения выполнения элементарных функций, представляющих собой алгоритмы преобразования агрегированного пространства состояний самой системы.

Сложность решения подобной задачи может быть обусловлена многогранностью и многофункциональностью описываемой информационной системы, для чего требуется детальный анализ алгоритмов ее безопасного функционирования через перечисление множества всех допустимых состояний.

Процесс агрегирования, являющийся ключевым понятием функционального подхода, заключается в построении агрегированного пространства состояний информационной системы – *Re*, которое отличается от настоящего рядом упрощений (укрупнений), но при определенных допущениях может рассматриваться как реальное.

В данном случае разработка функционального представления информационной системы осуществляется на основе анализа пространства параметров процессов в системе по установленным правилам и выявлении тех параметров, которые характеризуют действие атаки.

Элементарной функцией $f_i \in F$ будем называть математическое описание соответствующего ей элементарного действия или композицию элементарных действий минимальной длины, в виде алгоритма преобразования, определенного на всем пространстве агрегированных состояний информационной системы.

Описание алгоритма соответствующего преобразования требует построения области определения и области значений для каждой из элементарных функций. При этом областью определения элементарной функции системы будем называть полное подмножест-

во ее состояний, каждое из которых для данного преобразования имеет образ с ним не совпадающий.

Областью значений элементарной функции будем называть полное подмножество состояний информационной системы, каждое из которых для данного преобразования есть прообраз. При этом если:

$$S_f^{/} \subset S \tag{1}$$

– область определения функции *f* в *Re*, а

$$S_f^{//} \subset S \tag{2}$$

– область значений функции *f* в *Re,* то преобразование в *Re* согласно *f* опишется отображением:

$$G_f : S_f^{/} \to S_f^{//}. \tag{3}$$

В то же время предполагается достоверным, что в процессе конкретного элементарного действия изменению подвергаются лишь часть компонентов пространства, то есть во внимание принимаются только те компоненты агрегированного пространства состояний, которые имеют смысл для данного преобразования. Таким образом, с целью упрощения формализации задачи исследования информационной системы для каждой элементарной функции f_i строится абстрактное пространство состояний *Im* мерности $\dot{M} = \left\{\dot{X}_{\dot{k}}; \dot{k} = \overline{1, \left|\dot{M}\right|}\right\}$, с компонентами $\dot{X}_{j1}, j = \overline{1, \dot{M}_{i1}}$ в котором она полностью определена.

Следовательно, функция f_i в общем случае не всюду определена на декартовом произведении

$$\dot{S} = x_1 \times x_2 \times \ldots \times x_{\dot{M}_{i1}}, \tag{4}$$

или всюду определена на множестве состояний:

$$\dot{S}_f^{/} \subset \dot{S}, \tag{5}$$

где $S_f^{/}$ – есть область определения функции *f.*

$$\dot{S}_f^{//} \subset \dot{S} \text{ -} \tag{6}$$

область значения функции *f* в абстрактном пространстве *Im.* При этом отображением

$$G : \rho_i^M \leftrightarrow \dot{M}_i, \; s_n^{/} \in S \tag{7}$$

определяется состояние информационной системы на каждом шаге функционирования.

Значение компонентов абстрактного пространства в общем случае принимает значения

$$\dot{X}_j = x_{j1} | x_{j2} | \ldots | x_{jk} | \ldots | x_{jk_j^{\max}}, \tag{8}$$

при этом $x_j = \left\{x_{jk}; k = \overline{1, K_j^{\max}}\right\}$.

Подпространства области определения и области значения элементарной функции принимают смысл отображений:

$$G_i^{/} : \rho_{i1}^M \leftrightarrow \dot{M}_i, \tag{9}$$

$$G_i^{//} : \rho_{i2}^M \leftrightarrow \dot{M}_i, \tag{10}$$

$$G_i : \rho_i^M = \rho_{i1}^M \cup \rho_{i2}^M \leftrightarrow \dot{M}_i = M_{i1}^{/} \cup M_{i2}^{/}. \tag{11}$$

На рис. 1 представлена теоретико-множественная модель, описывающая процесс преобразования состояния в функциональной системе.

Процесс построения абстрактного пространства для элементарного преобразования заключается в последовательности этапов:

1. Переходе из *Re* в *Im* для выполнения элементарной функции *f*:

$$S_{\rho_f^M}^{/} \to S_{\rho_f^{\dot{M}}}^{/}, \tag{12}$$

который описывается отображением:

$$G_f^{/} : S_f^{/} \to \dot{S}_f^{/}. \tag{13}$$

2. Преобразовании в *Im* согласно *f* как отображения:

$$\dot{G}_f : \dot{S}_f^{/} \to \dot{S}_f^{//}. \tag{14}$$

3. Переходе из *Im* в *Re* после выполнения *f*:

$$S_{\rho_f^{\dot{M}}}^{//} \to S_{\rho_f^M}^{//}, \tag{15}$$

который описывается отображением:

$$G_f^{//} : \dot{S}_f^{//} \to S_f^{//}. \tag{16}$$

Критерий уровня наблюдаемости системы *N* определяет уровень абстракции анализируемых событий в защищаемой системе и определяет границы применимости метода для обнаружения атак в сетях. Традиционно, в подобного класса системах, рассматриваются следующие уровни наблюдаемости:

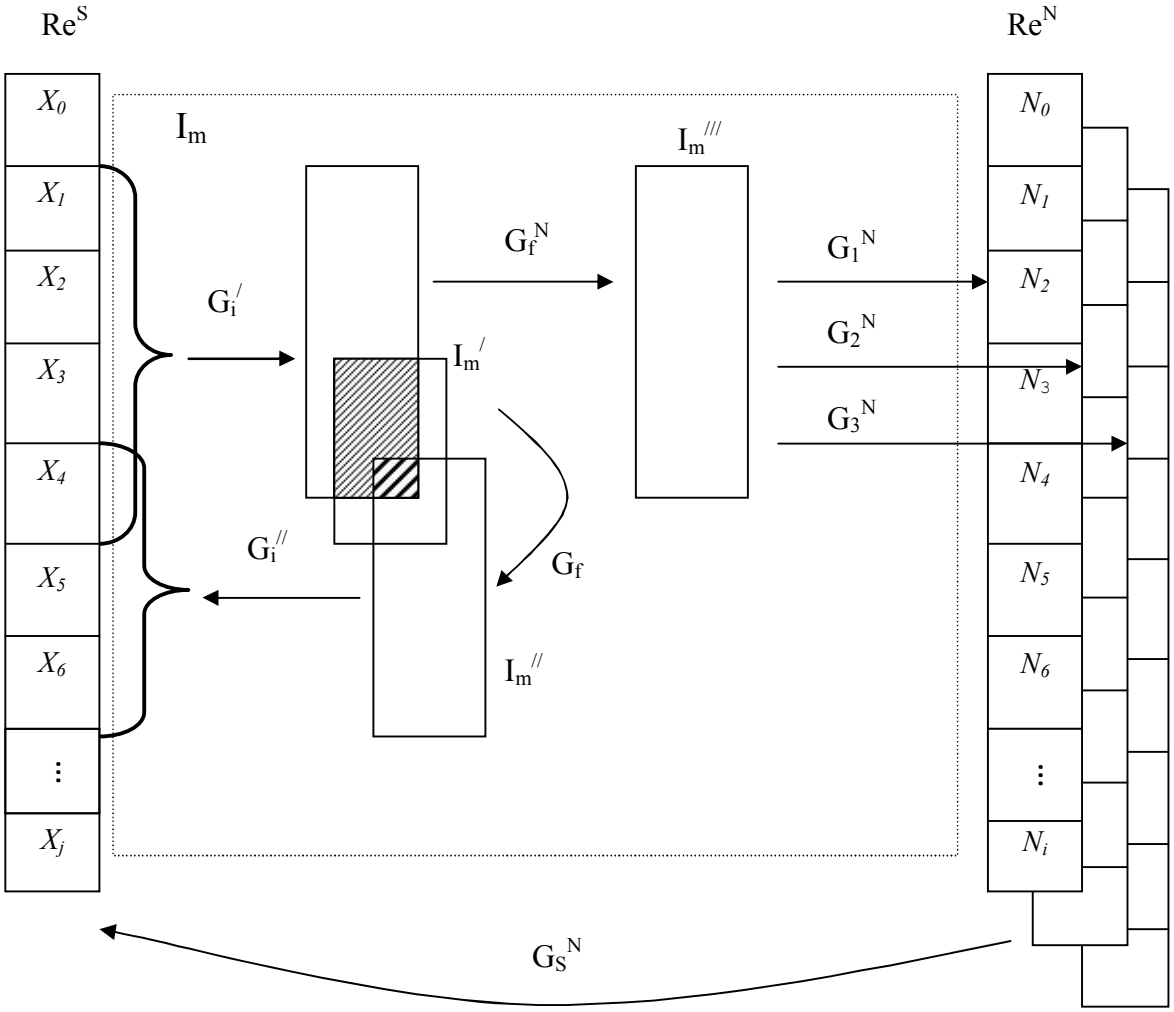

Рис. 1. *Теоретико-множественная модель, иллюстрирующая процессы преобразования в функциональной системе*

– наблюдение на уровне операционной системы отдельного узла сети;
– наблюдение на уровне сетевого взаимодействия объектов на узлах сети;
– наблюдение на уровне отдельных приложений узла сети;
– комбинация наблюдателей разных уровней.

В разработанной модели отображение G_S^N есть семантика наблюдаемости состояния, возникающая как образ состояния информационной системы посредством средств отображения информации и представляющий собой семантический код результата наблюдения. При этом каждый из элементов вектора наблюдаемости N несет информацию, в зависимости от того, насколько наблюдаемость обеспечивает оценку состояния всей системы. Под наблюдаемостью информационной системы понимается степень соответствия семантического кода образа состояния информационной системы ее истинному состоянию.

Отображение G_f^N отдельных элементарных функций в абстрактном пространстве формирует конкретные образы отдельных компонентов как значения их наблюдаемости, причем каждая из них имеет свое проявление $G_{1,2,3}^N$ для различных уровней наблюдаемости разными средствами отображения.

Множества элементарных функций, объединенные в правильные композиции, представляют собой цепочки, отражающие поведение информационной системы во введенном пространстве состояний. При этом поведение $l \in F^{\div}$ длины $|l|$ представляет собой кортеж:

$$l = f_{l_1} \circ f_{l_2} \circ \ldots \circ f_{l_{|l|}} \in F^*, \quad (17)$$

а $F_{mp} = \left\{ f^n ; n = \overline{1, |l|} \right\}$ – множество элементарных функций, соответствующих требуемому поведению информационной системы на каждом шаге n. При этом l можно обозначить отображением

$$G_l = F_{mp} \to F, \qquad (18)$$

$$(\forall f^n)(\exists! f_i) P(G_l(f^n) = f_i) \qquad (19)$$

Последний предикат утверждает о существовании единственно верного элементарного действия f_i на каждом шаге функционирования информационной системы. Отклонения от набора элементарных действий, в свою очередь, свидетельствует о наличии признака компьютерной атаки.

Заключение

Таким образом, анализ проблемы увеличения компьютерных атак на информационные системы различного назначения требует поиска наиболее эффективных способов их обнаружения и применения имеющегося арсенала средств защиты информации. Разработанная модель позволяет представить информационную систему на основе функционального подхода, предполагающего решение задачи нахождения полного множества безопасных состояний информационной системы.

Нахождение полного множества подобных состояний, в свою очередь, позволит опреде-лить признаки компьютерной атаки. Последующее распознавание характера атаки может осуществляться на основе системного анализа пространства параметров процессов в системе по установленным правилам и выявление тех параметров, которые характеризуют действие подобной компьютерной атаки.

Список литературы

1. *Абашев А.Н, Пазухин В.А, Слышкин А.С.* На шаг впереди киберпреступников // Журнал "Information Security / Информационная безопасность" № 1, 2015. С. 8–12.
2. *Gemalto* Releases Findings of 2014 Breach Level Index // URL: http://www.gemalto.com/press/Pages/Gemalto-Releases-Findings-of-2014-Breach-Level-Index.aspx (дата обращения: 25.09.2015).
3. *Мазин А.В., Клочко О.С.* Анализ методов противодействия угрозам и атакам на вычислительные системы. Наукоемкие технологии в приборо- и машиностроении и развитие инновационной деятельности в вузе // Матер. Всеросс. науч.-технич. конф. Т. 3. 2014. С. 71–75.
4. *Климов С.М., Сычёв М.П., Астрахов А.В.* Противодействие компьютерным атакам. Методические основы // Электронное учебное издание. М.: МГТУ им. Н.Э. Баумана, 2013. 108 с.
5. *Фор А.* Восприятие и распознавание образов / пер. с фр. А.В. Серединского; под ред. Г.П. Катыса. М.: Машиностроение, 1989. 272 с.
6. *Харитонов В.А.* Основы теории живучести функционально избыточных систем. С.-Пб.: Ин-т информатики и автоматизации Российской академии наук, 1993. 60 с.
7. *Гамаюнов Д.Ю.* Обнаружение компьютерных атак на основе анализа поведения сетевых объектов: дис… канд. физ.-мат. наук. М: МГУ им. Ломоносова, 2007.

The detection of computer attacks based on the functional approach

A. S. Shaburov, A. A. Mironova

Perm National Research Polytechnic University, Russia, 614990, Perm, Komsomolsky pr., 29

shans@at.pstu.ru, 89128876457

A brief summary of existing countermeasure methods and detection of computer attacks on information systems are stated. It is suggested that in the majority of defense systems the solution of the problem comes not to detection of computer attacks but simply to detection of its consequences while adaptability to unknown attacks in the majority of existing detection systems is missing. The mathematical model of computer attack detection based on the functional approach is proposed. The use of the functional approach presupposes the full set construction of secure conditions of an informational system and its further detection of computer attacks signs.

Key words: *computer attack; information defense; neural network; anomaly detection method; signature analyses; theory of image recognition; functional approach.*