УДК 004.93.1

ПРИМЕНЕНИЕ МЕТОДА ИНДУКТИВНОГО ПРОГНОЗИРОВАНИЯ СОСТОЯНИЙ ДЛЯ ОБНАРУЖЕНИЯ КОМПЬЮТЕРНЫХ АТАК В ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

Лаптев Владимир Николаевич
к.т.н., доцент
Кубанский государственный аграрный университет, Краснодар, Россия

Сидельников Олег Васильевич
Филиал Военной академии связи, Краснодар, Россия

Шарай Вячеслав Александрович
Институт информационных технологий и безопасности Кубанского государственного технологического университета, Краснодар, Россия

Рассматривается задача обнаружения компьютерных атак на ИТКС. Показано, что она сводятся к решению комбинаторных задач, и их решение связано с ветвлением решающих процессов, с перебором вариантов, число которых быстро растет при усложнении системы закономерностей. Такой перебор неизбежен, но его можно сокращать до приемлемой величины, позволяющей решать задачи обнаружения компьютерных атак на ИТКС. Применение индуктивного прогнозирования состояний позволит модифицировать базу знаний СОА ИТКС в автоматическом режиме, формировать новые правила и удалять старые

Ключевые слова: ИНДУКТИВНОЕ ПРОГНОЗИРОВАНИЕ СОСТОЯНИЙ, ОБНАРУЖЕНИЕ КОМПЬЮТЕРНЫХ АТАК, ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННАЯ СИСТЕМА, БАЗА ЗНАНИЙ

UDC 004.93.1

APPLICATION OF THE METHOD OF THE INDUCTIVE FORECASTING OF STATES FOR DETECTION OF COMPUTER ATTACKS IN INFORMATION-TELECOMMUNICATION SYSTEMS

Laptev Vladimir Nikolayevich
Cand.Tech.Sci., associate professor
Kuban State Agrarian University, Krasnodar, Russia

Sidelnikov Oleg Vasilevich
Filial of the Military Academy communication, Krasnodar, Russia

Sharaj Vyacheslav Aleksandrovich
Institute of information technology and safety of the Kuban state technological university, Krasnodar, Russia

The problem of acquisition of computer attacks is considered. It is demonstrated that it is a solution of combinatorial problems, and their solution is bound to fork of solving processes, with search of alternatives that grows fast at thickening of system of regularities. Such search is foregone, but it can be reduced to the reasonable magnitude, allowing solving problems of acquisition of computer attacks. Application of inductive forecasting of statuses will allow to inoculate knowledge base in an automatic mode, to shape the new rules and to delete the old ones

Keywords: INDUCTIVE PREDICTION OF STATES, DETECTION OF COMPUTER ATTACKS, INFORMATION AND TELECOMMUNICATION SYSTEMS, KNOWLEDGE BASE

Национальная безопасность Российской Федерации существенным образом зависит от обеспечения информационной безопасности. При этом, актуальность создания надежных и производительных систем обнаружения компьютерных атак на информационно-телекоммуникационные системы (ИТКС) и противодействия компьютерному нападению вытекает из Стратегии национальной безопасности, Доктрины информационной

безопасности РФ, Приоритетных проблем научных исследований в области обеспечения информационной безопасности РФ, требований руководящих документов [1-4].

Обнаружение компьютерных атак на ИТКС является одной из задач, решение которой позволяет повысить защищенность ИТКС в процессе их функционирования и развития [5].

Решению общих проблем обнаружения компьютерных атак посвящены работы В.И.Городецкого, И.В.Котенко, А.В.Лукацкого, С.М.Климова, С.С.Корт и др., а также ряда зарубежных авторов, в том числе Д.Деннига, Д.Андерсона, С.Кумара, С.Норткатт и др.

Используемые в настоящее время методики, методы и алгоритмы обнаружения компьютерных атак на ИТКС не в полной мере разрешают противоречие между увеличением времени на обнаружение компьютерных атак существующими методами за счет увеличения времени анализа признаков атак с одной стороны, и увеличением времени, отводимое на анализ признаков атак, которое приводит к возрастанию вероятности пропуска атаки с другой стороны [6].

Цель статьи – обоснование метода индуктивного прогнозирования состояний для обнаружения компьютерных атак на ИТКС.

1. Анализ известных методов обнаружения компьютерных атак

Компьютерной атакой на ИТКС считаются действия, направленные на реализацию угроз несанкционированного доступа к информации, воздействия на нее или на ресурсы ИТКС с применением программных и (или) технических средств [7].

Осуществление компьютерной атаки происходит при наличии точек несанкционированного доступа к информационным ресурсам и

коммуникационному оборудованию ИТКС или при наличии потенциального внутреннего нарушителя с полномочиями штатного оператора в территориально-распределенной вычислительной сети.

Последствиями воздействий компьютерных атак могут стать блокирование управляющей информации и внедрение ложной информации, нарушение установленных регламентов сбора, обработки и передачи информации в комплексах средств автоматизации, отказы и сбои в работе ИТКС, а также компрометация получаемой потребителями информации [5, 8].

Процесс обнаружения компьютерных атак начинается со сбора данных, необходимых для определения факта атаки на ИТКС [5, 8-11]. В частности, можно анализировать сведения о пакетах данных, поступающих в ИТКС, производительность программно-аппаратных средств (вычислительная нагрузка на хосты, загруженность оперативной памяти, скорость работы прикладного ПО), сведения о доступе к определенным файлам системы и т. д.

Для сбора исходной информации традиционно используют специализированные датчики, размещаемые на разных элементах ИТКС. Существуют два типа таких датчиков – сетевые и хостовые. Анализ данных, собранных сетевыми и хостовыми датчиками, проводится в ИТКС с использованием специальных методов обнаружения атак.

Эффективность обнаружения компьютерных атак во многом зависит от применяемых методов полученной информации. В первых системах обнаружения компьютерных атак, разработанных в 80-х годах, использовались статистические методы обнаружения атак. В настоящее время к статистическому анализу добавился ряд новых методик применения интеллектуальных систем обнаружения атак, в качестве

интеллектуального инструмента в которых, используются нейронные сети, системы нечеткой логики и экспертные системы [5, 8-11].

Процесс обучения с применением нейросетевых технологий начинается с предъявления системе набора обучающих примеров, состоящих из входных и выходных сигналов. Затем нейронная сеть автоматически подстраивает свои синоптические веса таким образом, что при последующем предъявлении входных сигналов на выходе получаются требуемые сигналы. Недостатками данного подхода являются: сложность построения; трудность подобрать обучающую выборку, адекватно, описывающую предметную область; длительный период обучения; непонятность (непрозрачность) результатов; нехватка адекватного обучающего материала [12].

Указанные недостатки отсутствуют в системах на основе баз знаний, использующих для обучения логический вывод (ЛВ). При этом под способностью к обучению понимается возможность создания базы знаний, а также пополнение и модификация правил в базе знаний под влиянием вновь полученной информации [13].

Большинство современных интеллектуальных систем, использующих ЛВ, позволяет модифицировать базу знаний только в ручном режиме. Известные методы формирования знаний (или методы машинного обучения), позволяющие автоматически изменять базу знаний, основаны на применении индуктивного ЛВ [13]. Индукция подразумевает наличие достаточно представительной выборки обучающих примеров, которая обобщается посредством сгенерированных правил.

Перспективным методом обнаружения компьютерных атак на ИТКС является технология обнаружения компьютерных атак на основе метода индуктивного прогнозирования состояний [14].

2. Обоснование метода индуктивного прогнозирования состояний для обнаружения компьютерных атак

В [13] предлагается процесс распознавания признаков объекта разделить на два этапа: обучение и собственно распознавание. Первый этап – индуктивный, второй – дедуктивный. На первом из них, обрабатываются данные многочисленных наблюдений над исследуемым классом объектов, и на основе полученных результатов строится некоторое решающее правило. На втором этапе описанное правило применяется для распознавания интересующих нас, но непосредственно не измеряемых свойств других объектов этого же класса.

Рассмотрим множество объектов и обозначим его через U, полагая, что оно состоит из отдельных элементов, обозначаемых через $u_2: U=\{u_1,u_2,\ldots,u_m\}$. Множество всех признаков, используемых при описании этих объектов, обозначим через $S=\{s_1,s_2,\ldots,s_m\}$. Множество всех объектов, обладающих некоторым конкретным признаком s_j, обозначим через U_j, называя его классом с признаком s_j, а его дополнение, т. е. множество всех объектов, не обладающих признаком s_j, – через $\overline{U}_j$.

Например, U может представлять адреса источника IP–датаграммы, собираемые сетевыми датчиками для аудита, $s_1,s_2,\ldots,s_m$ – такие признаки, как октеты диапазона адресов: 192.168.1.16. U_3 – множество всех разрешенных адресов источника IP–датаграммы, содержащиеся в третьем октете, $\overline{U}_3$ – множество не разрешенных адресов источника IP–датаграммы, содержащиеся в третьем октете.

При исследовании реальных объектов мощность множества U, т. е. число элементов в нем, оказывается обычно очень большой, и, как правило, известна лишь относительно малая его часть. Предположим, что

нам доступна вся информация об этом множестве и удалось описать каждый его элемент, перечислив признаки, которыми последний обладает, например в виде $u_1 - (s_1, s_2, s_4)$. Это означает, что объект представляет собой комбинацию признаков s_1, s_2 и s_4, т. е. обладает этими признаками и никакими другими. Доступную информацию можно представить иначе — строкой из нулей и единиц – булевым вектором. Символы строки соответствуют признакам s_1, s_2 ... и следует, что если объект обладает признаками: «1» – обладает, «0» – не обладает.

Например, описание $u_1 - (s_1, s_2, s_4)$ можно заменить на вектор: $[1101]$ (в данном случае число признаков ограничено четырьмя; при большем их числе вектор дополняется справа нулями).

Пользуясь такими средствами, можно представить множество U в целом. Для этого следует расположить друг под другом булевы векторы; представляющие последовательно объекты u_1, u_2 и т. д., получить булеву (из нулей и единиц) матрицу. Обозначим ее через R. Матрица содержит в удобной для обозрения форме информацию об отношении принадлежности признаков объектам: если объект U_j обладает признаком s_j, то на пересечении i-й строки и j-го столбца ставится «1», в противном случае – «0».

При больших ограничениях на используемые измерительные средства индивидуальность объектов может быть потеряна. Тогда отдельные строки матрицы R будут служить уже не моделями каких-то единичных объектов, а представлять их целыми группами, говоря о том, что в множестве U существуют объекты с заданными комбинациями признаков:

$$R = \begin{array}{c} \begin{matrix} a & b & c & d \end{matrix} \\ \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix}$$

Строки матрицы являются R – группы объектов, неразличимых в данной системе признаков. Столбцы задают классы.

Множество всех таких комбинаций булевых векторов образуют так называемое булево пространство M. Множество U допустимых комбинаций является подмножеством из $U \subseteq M$. Назовем это подмножество область существования объектов исследуемого класса, а его дополнение $\overline{U} = M \setminus U -$ область запрета, поскольку данное множество образуется запрещенными признаками. Всего их будет 12 ($2^4 = 16, 16 - 4 = 12$).

Таким образом, знание даже одного элемента множества запрета $\overline{U}$, т.е. информация о том, что некоторый объект не существует, позволяет иногда решить задачу распознавания, в то время, как аналогичные сведения о существовании некоторого объекта оказываются недостаточными для этого. Поэтому формулировку закономерностей, позволяющих решить задачи распознавания признаков компьютерных атак, будем связывать с запретами, т.е. запрет на некоторые комбинации признаков, что позволит осуществлять не весь перебор возможных признаков атак в базе данных, а ограничиться сокращенным перебором.

При использовании метода индуктивного прогнозирования состояний компьютерная атака рассматривается как последовательность действий, приводящих систему из начального состояния в скомпрометированное (конечное) состояние. Таким образом, атака моделируется как множество состояний и переходов между ними. Состояние системы рассматривается как набор переменных, описывающих объекты, представленных в сигнатуре атаки. Начальное состояние ассоциируется с состоянием до выполнения атаки, а

скомпрометированное состояние соответствует состоянию по окончании атаки. Переходы из состояния в состояние ассоциируются с новыми событиями в системе, влияющими на исполнение сценария атаки (например, запуск приложения, открытие TCP–соединения и т.д.). Типы допустимых событий (состояний системы) хранятся в базе данных. Между начальным и скомпрометированным состояниями существует множество переходов.

Такой способ позволяет:

описать атаку более абстрактно, чем на уровне системных вызовов, и более точно, чем с использованием неформального текстового описания;

выделить основные события в ходе выполнения атаки.

Ключевыми факторами применимости метода индуктивного прогнозирования состояний являются следующие:

в связи с увеличением количества параметров (признаков) атаки, учитываемых в модели анализа состояний используется метод индуктивного вывода, основанный на распознавании признаков атак, связанных с запретами, т.е. запрет на некоторые комбинации признаков, что позволит осуществлять не весь перебор возможных признаков атак в базе данных, а ограничиться сокращенным перебором;

индукция подразумевает наличие достаточно представительной выборки обучающих примеров, которая обобщается посредством сгенерированных правил, позволяющая модифицировать базу знаний системы обнаружения атак (СОА) ИТКС в автоматическом режиме и сформировывать новые правила и удалять старые;

скомпрометированное состояние должно быть распознано без использования внешних знаний о намерениях нарушителя (трудно

обнаружить атаку маскарада с использованием учетной записи легитимного пользователя и корректного пароля).

Таким образом, задача обнаружения компьютерных атак на ИТКС сводятся к решению комбинаторных задач, и их решение связано с ветвлением решающих процессов, с перебором вариантов, число которых быстро растет при усложнении системы закономерностей. Такой перебор неизбежен, но его можно сокращать до приемлемой величины, позволяющей решать задачи обнаружения компьютерных атак на ИТКС. Применение индуктивного прогнозирования состояний позволит модифицировать базу знаний СОА ИТКС в автоматическом режиме и сформировывать новые правила и удалять старые.

Литература

1. Указ Президента Российской Федерации от 12 мая 2009 г. № 537 «О Стратегии национальной безопасности Российской Федерации до 2020 года» // Российская газета. — 19 мая 2009 года. - Режим доступа: http://www.rg.ru/2009/05/19/strategia-dok.html

2. Доктрина информационной безопасности Российской Федерации (от 09.09.2000) // Российская газета. - Режим доступа: http:// www.gov.ru.

3. Приоритетные проблемы научных исследований в области обеспечения информационной безопасности Российской Федерации [Текст]: [Утверждены Исполняющим обязанности Секретаря Совета Безопасности Российской Федерации, председателя научного совета при Совете Безопасности Российской Федерации 7 марта 2008 г] – М. : 2007. – Режим доступа: http://www.scrf.gov.ru/documents/93.html.

4. Руководящий документ Государственной технической комиссии при Президенте Российской Федерации. Концепция защиты средств вычислительной техники и автоматизированных систем от несанкционированного доступа к информации. // Сборник Руководящих документов по защите информации от несанкционированного доступа [Текст]: Государственная техническая комиссия при Президенте Российской Федерации. – М.: СИП РИА, 1998. – 120с.

5. Климов СМ., Методы и модели противодействия компьютерным атакам.- М: Люберцы.: КАТАЛИТ, 2008. – 316 с.

6. Сидельников О.В., Коробка А.А. Анализ существующих методов обнаружения удаленных сетевых атак. // Перспективы развития средств и комплексов связи. Подготовка специалистов связи: Материалы межвузовской научной конференции. В 2 ч. Ч. 2 / Новочеркасское высшее военное командное училище связи. – Новочеркасск, 2009, С. 56-61.

7. Рекомендации по стандартизации Р 50.1.053-2005 Информационные технологии. Основные термины и определения в области технической защиты информации. – М.: Стандартинформ, 2005.

8. Сердюк В.А. Новое в защите от взлома корпоративных систем. – М.: Техносфера, 2007. – 306 с.

9. Корт С.С. Теоретические основы защиты информации: Учебное пособие. - М.: Гелиос АРВ, 2004. – 240 с.

10. Шангин В.Ф. Защита компьютерной информации. Эффективные методы и средства / Шангин В.Ф. - М.: ДМК Пресс, 2010. – 544 е.: ил.

11 Лукацкий А. В. Обнаружение атак. – 2-е изд. – СПб.: БХВ-Петербург, 2003. – 608 с.

12. Котельников, Е. В Абдуктивный метод модификации посылок в исчислении высказываний / Е.В. Котельников // Вестник Вятского научного центра Верхне-Волжского отделения Академии технологических наук Российской Федерации. Серия: Проблемы обработки информации. Вып. 1(6)/2006. - Киров, 2006 – С. 18-28.

13. Закревский А.Д. Логика распознавания. Изд. 2-е, доп. – М.: Едиториал УРСС, 2003. – 200. – 144 с.

14. Дорохов И.Н. Формализация выбора и принятия решений в изобретающей эксперт-ной системе // Тр. XX-ой Международной научно-технической конференции «Интел-лектуальные САПР» (CAD – 2007), 3-10 сентября 2007, Дивноморское, в 3- х томах. Т.1. – М.: Физматлит, 2007. С. 289-296.