УДК 519.682.1+681.3

Н. В. Медведев, Г. А. Гришин

МОДЕЛИРОВАНИЕ СЦЕНАРИЕВ СЕТЕВЫХ АТАК НА ОСНОВЕ СТОХАСТИЧЕСКИХ КОНТЕКСТНО-НЕЗАВИСИМЫХ ГРАММАТИК

Рассмотрены формальные подходы к описанию сетевых атак на основе формальных грамматик. Разрабатываемые системы защиты информации следует проверять экспериментально, для этого предложена формальная модель сетевых атак.

Важность проблемы информационной безопасности компьютерных сетей существенно возрастает с увеличением числа и сложности компьютерных сетей и масштабов внедрения перспективных информационных технологий [1]. Сейчас злоумышленники обладают целым арсеналом изощренных атак на распределенные информационные системы, стремясь получить доступ к конфиденциальной информации, нарушив ее целостность [2].

В настоящее время усилия в области информационной безопасности направлены, главным образом, на разработку механизмов защиты. Значительно меньше внимания уделяется исследованию теоретических вопросов реализации атак, в том числе их представлению и моделированию [3–5].

Практически не существует приемлемых программных средств моделирования атак. Поэтому многие исследователи используют упрощенные или частные модели атак, составляющие лишь малую часть возможных типов атак.

Практическая потребность в средствах моделирования атак обусловливается, по крайней мере, двумя причинами.

Во-первых, разрабатываемые системы защиты информации следует проверять экспериментально, а для этого нужно иметь среду моделирования атак.

Во-вторых, для практического внедрения систем защиты информации (в частности, систем обнаружения атак) требуется первоначальное наполнение и последующее наращивание их базы знаний. Для этого можно использовать среду моделирования атак, которая служит средством генерации прецедентов атак (обучающих данных).

В настоящей работе дано представление о предлагаемых авторами формальных моделях атак, предназначенных для реализации в разрабатываемом имитаторе атак. Во второй части статьи проведен анализ атак на компьютерные сети. Третья часть посвящена заданию концептуальной модели атак на компьютерные сети, четвертая часть является центральным и посвящен собственно формальным моделям атак, в пятой части изложены направления будущих исследований.

Анализ атак на компьютерные сети. Под атакой на компьютерные сети принято понимать несанкционированное информационное воздействие на объекты компьютерной сети, осуществляемое как по каналам связи, так и при непосредственном контакте с поражаемым объектом сети.

В настоящее время в работах по защите информации [6–9] предложено большое количество таксономий атак на компьютерные сети: списки терминов атак, списки категорий атак, категории результатов атак, эмпирические списки атак, матрицы уязвимости, таксономии, базирующиеся на действиях атакующей программы, таксономии дефектов и вариантов уязвимости защиты, таксономии атак, основанные на их сигнатурах, таксономии вторжений, базирующиеся на процессах (таксономии инцидентов). В частности, таксономия списков категорий атак, введенная Ранумом [8], основана на использовании восьми классов сетевых атак: 1) “социальная инженерия” — введение в заблуждение жертвы; 2) “заимствование прав” — захват прав доступа авторизованных пользователей; 3) “использование” — использование “дыр” в программном обеспечении или операционных системах; 4) “транзитивное доверие” — использование доверия между хостами (хост–хост) и сетями (сеть–сеть); 5) “атаки, управляемые данными”, — “троянцы”, лазейки, вирусы; 6) “инфраструктура” — использование ошибок и особенностей сетевых протоколов или инфраструктуры сети; 7) “отказ в обслуживании” — препятствование использованию системы; 8) “волшебство” — новые атаки, которые никто еще не фиксировал.

В соответствии с одной из наиболее распространенных классификаций (рис. 1) атаки могут структурироваться по семи основным признакам [10]: характеру воздействия (пассивное, активное); цели воздействия (нарушение конфиденциальности, целостности, доступности информации); условию начала осуществления воздействия (по запросу от атакуемого объекта, по наступлению ожидаемого события, безусловная атака); наличию обратной связи с атакуемым объектом (с обратной связью и без нее); расположению субъекта атаки относительно атакуемого объекта (внутрисегментное, межсегментное); уровню эталонной модели ISO/OSI, на котором осуществляется воздействие

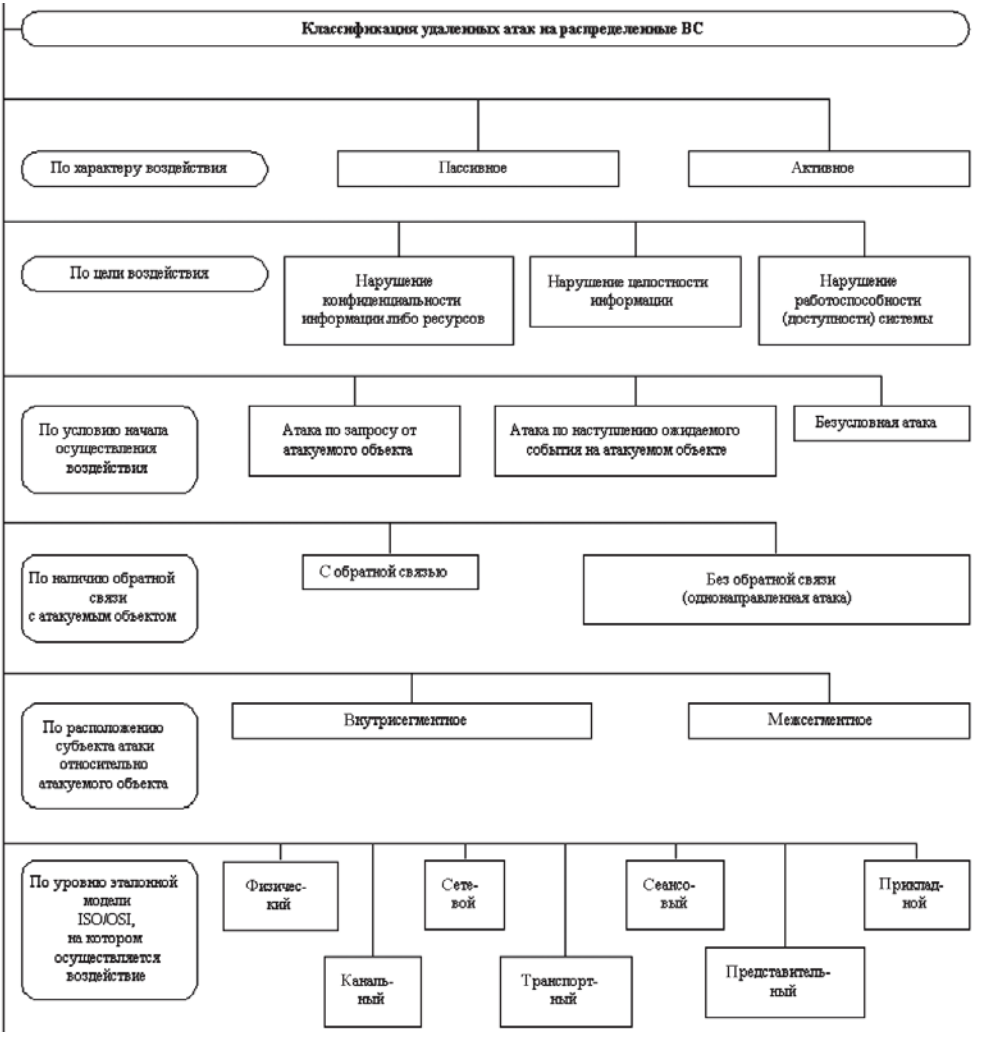

Рис. 1. Классификация атак

(физический, канальный, сетевой, транспортный, сеансовый, представительный, прикладной); объекту, на который направлено воздействие (на сетевые сервисы, на инфраструктуру сети).

К типовым сценариям реализации удаленных атак можно отнести: 1) анализ сетевого трафика, 2) сканирование сети, 3) подмену доверенного объекта сети и передачу по каналам связи сообщений от его имени с присвоением его прав доступа, 4) внедрение ложного объекта в сеть, 5) отказ в обслуживании, 6) неавторизованный доступ с удаленного хоста посредством подбора пароля, 7) неавторизованное повышение привилегий доступа, 8) удаленный запуск приложений.

Концептуальная модель атак на компьютерные сети. Анализ удаленных атак показывает, что каждая атака предварительно планируется на макроуровне в виде частично упорядоченного множества

шагов, которые вместе составляют сценарий атаки. Каждый шаг направлен на достижение частной цели (например, анализ атакуемой компьютерной сети, преодоление системы аутентификации, повышение прав, получение доступа к информации, выполнение операции с объектом доступа, "заметание следов"). Эти шаги могут быть реализованы в различном, хотя и не в произвольном, порядке, могут повторяться и выполняться с различных удаленных компьютеров. Каждый шаг сценария атаки реализуется путем последовательных простых командами и операций микроуровня.

Следуя такому представлению атак, можно предложить двухуровневую концептуальную модель атак (рис. 2).

На первом (макро) уровне задается общий сценарий атаки. Даже единичный прецедент такого сценария позволяет эксперту идентифицировать намерения атакующего, особенности выполнения атаки, ее варианты и переменные параметры, ведущие к той же самой цели. Каждый сценарий описывается множеством допустимых последовательностей шагов, определяющих класс атак на макроуровне.

Второй (микро) уровень определяет более детальную спецификацию атаки. Каждый шаг сценария (макроуровня) на микроуровне состоит из последовательных событий. Этими событиями являются конкретные команды операционной системы, вызываемые стандартные приложения и программы атакующего (эксплоиты) с конкретными параметрами вызова.

Формальные модели атак на компьютерные сети. Описанные ранее уровни концептуальной модели атак на компьютерные сети можно задать формально.

На макроуровне каждая последовательность может рассматриваться как "слово", принадлежащее формальному языку, специфицируемому посредством некоторой формальной грамматики [11–14].

Цепочка принадлежит языку сценария атак, порождаемому грамматикой, только в том случае, если существует ее вывод из цепи этой грамматики. Процесс построения такого вывода (а, следовательно, и определения принадлежности цепочки языку) называется *разбором*.

С практической точки зрения наибольший интерес представляет разбор по контекстно-свободным (КС) грамматикам. Их порождающей мощности достаточно для описания большей части синтаксической структуры сетевых атак для различных подклассов КС-грамматик имеются хорошо разработанные способы решения задачи разбора.

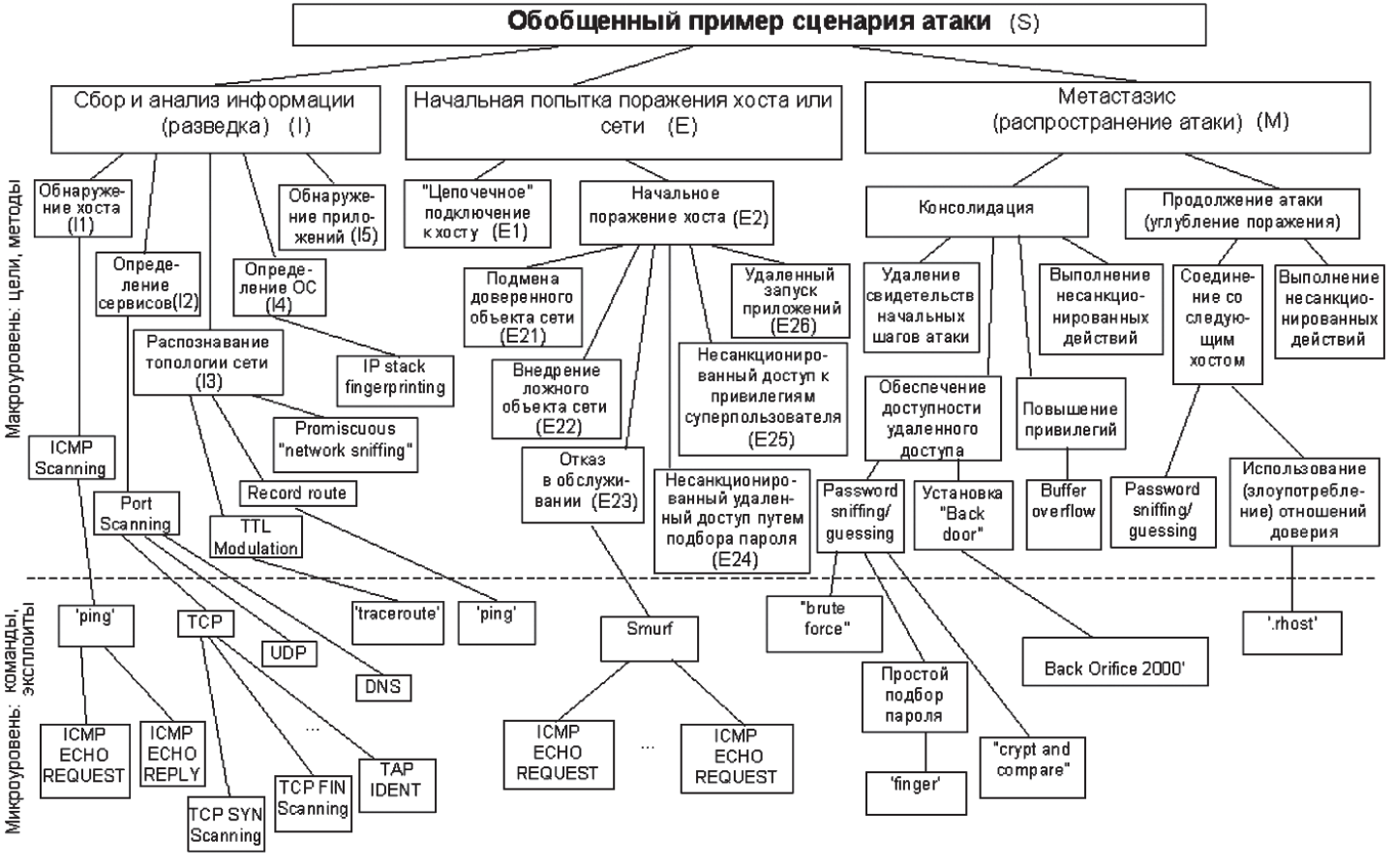

Рис. 2. Дерево разбора сетевой атаки на двух уровнях представления

Для КС-грамматик можно ввести удобное графическое представление вывода, называемое деревом вывода, причем для всех эквивалентных выводов деревья вывода совпадают.

Дерево называется деревом вывода (или деревом разбора) (рис. 3) в КС-грамматике $G_A = \langle V_N, V_T, S, P\rangle$, если выполнены следующие условия.

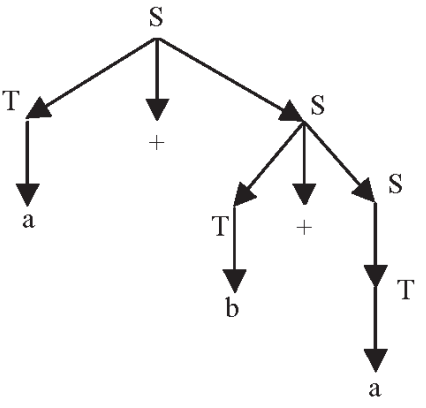

Рис. 3. Дерево вывода для цепочки $a + b + a$ в грамматике G_A

1. Каждая вершина дерева помечена символом из множества $(V_N \cup V_T \cup \varepsilon)$, при этом корень дерева помечен символом S; листья — символами из $(V_T \cup \varepsilon)$.

2. Если вершина дерева отмечена символом $A \in V_N$, а ее непосредственные потомки — символами $a_1, a_2, \ldots, a_n$, где каждое $a_i \in (V_T \cup V_N)$, то $A \to a_1 a_2 \ldots a_n$ — правило вывода в этой грамматике.

3. Если вершина дерева отмечена символом $A \in V_N$, а ее единственный непосредственный потомок помечен символом ε, то $A \to \varepsilon$ — правило вывода в этой грамматике.

Дерево вывода (рис. 3) можно строить нисходящим либо восходящим способами.

При нисходящем разборе дерево вывода формируется от корня к листьям; на каждом шаге до вершины, отмеченной нетерминальным символом, пытаются найти такое правило вывода, чтобы имеющиеся в нем терминальные символы “проектировались” на символы исходной цепочки.

Метод восходящего разбора заключается в том, что исходную цепочку пытаются “свернуть” к начальному символу S; на каждом шаге ищут подцепочку, которая совпадает с правой частью какого-либо правила вывода; если такую подцепочку находят, то ее заменяют нетерминалом из левой части этого правила.

Если грамматика однозначная, то при любом способе построения будет получено одно и то же дерево разбора.

Дерево разбора, описывающее обобщенный сценарий атаки посредством стохастической грамматики, имеет следующий вид (рис. 2):

$$G_A = \langle V_N, V_T, S, P\rangle,$$

где V_N — множество терминальных символов, которые обозначают шаги атаки нижнего уровня, $V_N = \{I, E, M, I, \ldots, E2, E21, E22, E23, E24, E25, E26, \ldots\}$; V_T — множество нетерминальных символов, которые ставятся в соответствие верхним и промежуточным уровням

представления шагов сценария атаки; S — начальный символ сценария атаки; P — множество правил вывода, описывающих операции детализации сценария атаки посредством замены символов. Каждая замена осуществляется с заданной вероятностью:

$$\alpha_i \xrightarrow{P_{ij}} \beta_{ij}, \quad i = 1, \ldots, n_i, \quad i = 1, \ldots, k.$$

Пример цепочки вывода:

$S \xrightarrow{P_{11}} I$, $S \xrightarrow{P_{12}} IE$, $S \xrightarrow{P_{13}} IEM$, …, $E2 \xrightarrow{P_{21}} E21$, $E2 \xrightarrow{P_{22}} E22$, $E2 \xrightarrow{P_{23}} E23$, $E2 \xrightarrow{P_{24}} E24$, $E2 \xrightarrow{P_{25}} E25$, $E2 \xrightarrow{P_{26}} E26$.

Эта грамматика может быть восстановлена индуктивно по множеству слов-прецедентов формальными методами. Данная грамматика может играть двоякую роль, а именно: использоваться и как модель генерации прецедентов, и как модель распознавания атак на базе синтаксического анализа.

С математической точки зрения формальная модель атаки на микроуровне может определяется также в терминах формальных грамматик или в терминах марковских цепей.

Выводы. В настоящей работе проведен анализ атак на компьютерные сети, задана концептуальная модель атак на компьютерные сети, предложены формальные модели атак. Формальные модели атак приведены с использованием аппарата стохастических контекстно-свободных грамматик.

Направления будущих исследований.

1. Разработка представительного множества удаленных атак нескольких классов и их описание в терминах сценариев.

2. Разработка алгоритма восстановления формальной грамматики, задающей модель атаки данного класса, на основе множества прецедентов (примеров).

3. Разработка моделей сценариев атак в терминах формальных грамматик.

4. Разработка стохастических моделей фрагментов атак на микроуровне.

5. Разработка объектно-ориентированного проекта программного прототипа системы моделирования атак и его программная реализация.

6. Экспериментальное исследование разработанной системы моделирования атак и определение ее возможностей при проектировании и эксплуатации систем защиты компьютерных сетей.

СПИСОК ЛИТЕРАТУРЫ

1. З а к о н о д а т е л ь н о - правовое и организационно-техническое обеспечение информационной безопасности автоматизированных систем и информационно-вычислительных сетей / Под ред. И.В.Котенко. – СПб.: ВУС, 2000.

2. K a r s a e v O. Framework for Ontology-based Representation of Distributed Knowledge in Multiagent Network Security System // Proceedings of the 4th World Multi-conference on Systems, Cybernetics and Informatics (SCI-2000), Vol. III: “Virtual Engineering and Emergent Computing”. Orlando, USA, 2000.
3. M o u n j i A. Languages and Tools for Rule-Based Distributed Intrusion Detection. PhD thesis, Faculties Universitaires Notre-Dame de la Paix Namur (Belgium), September 1997.
4. D o y l e J. Some representational limitations of the Common Intrusion Specification Language. Laboratory for Computer Science, Massachusetts Institute of Technology, October 26, 1999.
5. E c k m a n n S. T., V i g n a G., K e m m e r e r R. A. STATL: An Attack Language for Statebased Intrusion Detection. Dept. of Computer Science, University of California, Santa Barbara. 2000.
6. S. K u m a r. Classification and Detection of Computer Intrusions. PhD thesis, Purdue University, West Lafayette, Indiana, USA, Aug. 1995.
7. H o w a r d J. D. An Analysis of Security Incidents on the Internet, 1989–1995, Ph.D. Dissertation, Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, April, 1997.
8. R a n u m M. A Taxonomy of Internet Attacks. Web Security Sourcebook. John Wiley & Sons. 1997.
9. K o r b a J. Windows NT Attacks for the Evaluation of Intrusion Detection Systems, M.Eng. Thesis, MIT Department of Electrical Engineering and Computer Science, May, 2000.
10. М е д в е д о в с к и й И. Д., С е м ь я н о в П. В., Л е о н о в Д. Г. Атака на Internet. – М.: ДМК, 1999.
11. Д ж. Ф о с т е р. Автоматический синтаксический анализ. – М.: Мир, 1975.
12. Ф у К. Структурные методы в распознавании образов. – М.: Мир, 1977.
13. А х о А., У л ь м а н Д ж. Теория синтаксического анализа, перевода и компиляции. Т. 1, 2. – М.: Мир, 1978.
14. L a m m e l R., V e r h o e f C. Semiautomatic Grammar Recovery. Centrum voor Wiskunde en Informatica, Amsterdam, The Netherlands. 2000.

Статья поступила в редакцию 8.06.2005