УДК 681.326

М. Н. Дмитриев, О. Д. Соколова

ПРИМЕНЕНИЕ ГИПЕРСЕТЕЙ ДЛЯ МОДЕЛИРОВАНИЯ АТАК НА РАСПРЕДЕЛЕННЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ

В работе рассматриваются методы обнаружения типовых атак на распределенные вычислительные системы. Приводится обзор механизмов реализации основных видов удаленных атак как на протоколы TCP/IP, так и на инфраструктуру сети. В качестве математической модели распространения атаки «ложный агент» использован объект «гиперсеть».

Введение

Большинство сетевых атак на распределенные вычислительные системы (РВС) используют уязвимости, прямо связанные с протоколами или с их реализациями. Существует множество вариантов атак, однако подавляющее их большинство принято делить на четыре известных типа [1]:

1) анализ сетевого трафика;
2) подмена доверенного объекта РВС;
3) ложный объект РВС, в том числе:
   a) внедрение в РВС ложного объекта путем навязывания ложного маршрута;
   b) внедрение в РВС ложного объекта используя недостатки алгоритмов удаленного поиска;
4) отказ в обслуживании.

Все эти типы атак многократно описаны в соответствующей литературе [1,7,8,10–12]. В данной работе авторы решили уделить особое внимание атаке типа «ложный объект», так как для ее описания можно использовать модели, основанные на графах и гиперсетях. Современные глобальные сети представляют собой совокупность сегментов, связанных между собой через сетевые узлы. При этом под маршрутом понимается последовательность узлов сети, по которой данные передаются от источника к приемнику, а под маршрутизацией — выбор маршрута. Все маршрутизаторы (роутеры) имеют специальную таблицу, называемую таблицей маршрутизации, в которой для каждого адресата указывается оптимальная последовательность прохождения узлов. Отметим, что такие таблицы существуют не только у маршрутизаторов, но и у любых хостов в глобальной сети. Для обеспечения эффективной маршрутизации в РВС применяются специальные управляющие протоколы, позволяющие роутерам обмениваться информацией друг с другом: RIP (Routing Internet Protocol), OSPF (Open Shortest Path First); уведомлять хосты о новом маршруте: ICMP (Internet Control Message Protocol); удаленно управлять маршрутизаторами: SNMP (Simple Network Management Protocol). Все эти протоколы позволяют удаленно изменять маршрутизацию в сети Интернет, т. е. являются протоколами управления сетью.

ISSN 1818-7897. Вестник НГУ. Серия: Математика, механика, информатика. 2008. Т. 8, вып. 1. С. 26–37

Использование ложного объекта для организации удаленной атаки на РВС

Очевидно, что процесс маршрутизации в глобальных сетях играет важнейшую роль и, как следствие, может подвергаться атаке. Основная цель атаки, связанной с навязыванием ложного маршрута — изменить исходную маршрутизацию на объекте РВС так, чтобы новый маршрут проходил через ложный объект, которым является хост атакующего. При передаче потоков в РВС могут возникать проблемы идентификации сетевых управляющих устройств (например, маршрутизаторов) при взаимодействии их с объектами системы. Если операционная система не решает подобных проблем, то РВС может подвергнуться типовой удаленной атаке, связанной с изменением маршрутизации и внедрением в систему ложного объекта. Внедрить такой объект можно и в том случае, если инфраструктура предусматривает использование алгоритмов удаленного поиска.

Реализация типовой угрозы «внедрение в РВС ложного объекта путем навязывания ложного маршрута» состоит в несанкционированном использовании протоколов управления сетью для изменения исходных таблиц маршрутизации. При этом атакующий посылает по сети специальные служебные сообщения, определенные данными протоколами, от имени сетевых управляющих устройств. В результате успешного изменения маршрута атакующий получает полный контроль над потоком информации, которой обмениваются два объекта РВС. Атака при этом переходит во вторую стадию ее реализации, связанную с приемом, анализом и передачей сообщений, получаемых от дезинформированных объектов РВС. Второй способ реализации этой же угрозы — «внедрение в РВС ложного объекта путем использования недостатков алгоритмов удаленного поиска». Объекты РВС обычно имеют не всю необходимую для адресации сообщений информацию, под которой понимаются аппаратные (адрес сетевого адаптера) и логические адреса (например, IP-адрес) объектов РВС. Для получения подобной информации в РВС используются различные алгоритмы удаленного поиска, заключающиеся в передаче по сети специального вида поисковых запросов и в ожидании ответов на них. Полученных таким образом сведений об искомом объекте запросившему их субъекту РВС достаточно для последующей адресации к нему. Примером сообщений, на которых базируются алгоритмы удаленного поиска, могут служить SAP-запрос в ОС Novell NetWare, а также ARP- и DNS-запросы в Интернет.

При использовании механизмов удаленного поиска реализация рассматриваемой типовой угрозы состоит в перехвате поискового запроса и передаче в ответ на него ложного сообщения, в котором указываются данные, использование которых приведет к адресации на атакующий ложный объект. Таким образом, в дальнейшем весь поток информации между субъектом и объектом взаимодействия будет проходить через ложный объект РВС.

Получив контроль над проходящим информационным потоком между объектами, ложный объект РВС может применять различные методы воздействия на перехваченную информацию. К их числу, например, относятся: селекция потока информации и сохранение ее на ложном объекте РВС; модификация информации.

Методы выявления атак

Мониторинг информационной безопасности РВС выполняет сбор, анализ данных с целью выявления в системе событий, характеризующих атаки или этапы их подготовки [2]. Кроме того, системой вырабатываются ответные действия на выявленные атаки. Ответная реакция на атаку может быть разной в зависимости от исполняемого приложения, так как конечной целью является исполнение системой ее функций, а не противодействие атакам.

Процесс принятия решения о том, какие состояния считать атакой, предполагает определение множества признаков контролируемых объектов, измерение значений этих признаков и установление их пороговых величин, построение решающих правил распознавания классов состояний объектов по векторам признаков, представляющих эти состояния. Методы решения задач распознавания в литературе принято делить [2] на **лингвистические** (синтаксические, структурные) и **геометрические**. Такие же подходы можно использовать и для распознавания атак на РВС.

Лингвистические методы используют в качестве признаков некоторые заранее определенные непроизводные (исходные) элементы. К таким, например, относятся поля пакетов, атрибуты базы данных применительно к задачам активного аудита. Лингвистические методы применяются, например, при сигнатурном анализе. Сигнатурный метод анализа основан на том, что большинство атак на систему известны и развиваются по схожим сценариям. В данном подходе сигнатуры вторжений определяют характерные особенности, условия, устройства и взаимосвязь событий, которые ведут к попыткам или собственно к вторжению. Простейшим методом реализации сигнатурного анализа является поддержание системой безопасности базы данных сигнатур вторжений. Последовательность действий, выполняемая пользователем или программой во время выполнения, сравнивается с известными сигнатурами. Признаком попытки нарушения безопасности может служить частичное соответствие последовательности событий сигнатуре. Типичными представителями, реализующими данную идею, являются антивирусные сканнеры (работают с базой данных сигнатур вирусов) и системы обнаружения сетевых атак (работают с базой данных сигнатур удаленных атак).

При использовании геометрических методов состояния распознаваемых объектов представляются точками в многомерном пространстве признаков, число измерений которого равно числу признаков. Среди данных методов наиболее распространены те, которые базируются на статистических моделях. Такие модели определяют статистические показатели, характеризующие параметры штатного поведения системы. Если с течением времени наблюдается определенное отклонение данных параметров от заданных значений, то фиксируется факт обнаружения атаки, который также именуется вторжением. Как правило, в качестве таких параметров могут выступать: уровень загрузки процессора; нагрузка на каналы связи; штатное время работы пользователей системы; количество обращений к сетевым ресурсам и другие.

Приведем некоторые известные модели, которые могут применяться при статистическом анализе безопасности поведения программ и в системах обнаружения [3–5]:

1. **Операционная модель** основывается на том, что каждое новое наблюдение пе-

ременной должно укладываться в некоторых границах. Если этого не происходит, то констатируем, что имеем дело с отклонением. Допустимые границы определяются на основании анализа предыдущих значений переменной. Данная модель может использоваться, если некоторое значение метрики можно аргументированно связать с попыткой вторжения (например, количество попыток ввода пароля более 10).

2. **Модель среднего значения и среднеквадратичного отклонения** для каждого статистического параметра на основе математического ожидания и дисперсии определяет доверительный интервал, в пределах которого должен находиться данный параметр. Если текущее значение параметра выходит за эти границы, то фиксируется осуществление атаки. Например, если для каждого пользователя сети определен доверительный интервал для времени его работы в системе, то факт регистрации пользователя вне этого интервала может рассматриваться как попытка получения несанкционированного доступа к ресурсам системы.

Модель применима для измерения счетчиков событий, временных интервалов и используемых ресурсов. Преимуществом модели по сравнению с операционной является независимость оценки аномальности поведения от априорных знаний.

3. **Многовариационная модель** аналогична модели среднего значения и среднеквадратичного отклонения, но учитывает корреляцию между двумя или большим количеством метрик (количество операций ввода-вывода, количество выполненных процедур входа в систему и время сессии).

4. **Модель Марковского процесса** применима только к счетчикам событий, рассматривает каждый тип событий как переменную состояния и использует матрицу переходов для характеристики частот переходов между состояниями. Наблюдение является аномальным, если вероятность перехода, определенная предыдущим состоянием и матрицей перехода очень мала. Модель применима в том случае, если рассматривается множество команд, последовательность которых важна.

5. **Модель временных серий** использует временные периоды вместе со счетчиками событий и измерениями ресурса. Учитываются как значения наблюдений, так и временные интервалы между ними. Новое наблюдение является аномальным, если вероятность его появления с учетом времени низка. Преимуществом данной модели является учет временного сдвига между событиями.

Моделирование атаки «внедрение ложного агента»

Для более подробного изучения атаки «внедрение ложного агента» рассмотрим сети передачи данных. Для передачи данных от источника к приемнику используется таблица маршрутизации, в которой для каждого адресата указывается оптимальная последовательность узлов. Опишем с помощью модели гиперсети атаку «внедрение в сеть ложного агента путем навязывания ложного маршрута».

Сеть передачи данных представляет собой иерархическую структуру. Так как каждый уровень при этом можно моделировать графом, то для отображения взаимодействий различных уровней удобно использовать объект гиперсеть, давно применяемый

для моделирования различных сетей. Определение абстрактной гиперсети приведено в [6].

Сеть передачи данных состоит из объектов (узлов) и каналов связи между ними. В рассматриваемом случае, а именно — для моделирования информационных потоков, в иерархической структуре сети достаточно выделить два уровня: первичную и вторичную сеть.

Первичная сеть предназначена для образования каналов и трактов между узлами связи, используемых вторичными сетями. Первичная сеть в решающей мере определяет ряд важнейших качественных характеристик, такие как надежность, живучесть, пропускную способность и другие показатели.

Вторичная сеть непосредственно обеспечивает передачу данных или обмен заданными видами сообщений. Каждая вторичная сеть использует каналы первичной сети для передачи потоков между абонентами — вершинами вторичной сети. Потоки данных при этом проходят по ребрам вторичной сети с учетом их вложения в ветви первичной.

Описание модели работы РВС до и после внедрения агента

С учетом описанных выше подходов к построению модели, задаются два графа: первичная сеть $PS = (X, V)$ и вторичная сеть $WS = (Y, R)$, где

X — множество вершин первичной сети, $|X| = n$;

V — множество ветвей первичной сети;

$Y \subseteq X$ — множество вершин вторичной сети;

R — множество ребер вторичной сети (информационные потоки).

По каждому ребру $(i, j) \in R$ передаются информационные пакеты с интенсивностью λ_{ij} (считаем, что потоки имеют Пуассоновское распределение).

В первичной сети для каждой ветви $(i, j) \in V$ задано время c_{ij} доставки единицы информации от вершины i к вершине j, где $i, j = \overline{1, n}$.

В ходе работы сети в каждой вершине $x_i \in X$ собираются статистические показатели, характеризующие параметры штатного поведения подконтрольной системы. В данном случае таким показателем является суммарный поток информации, входящий в каждую вершину.

Опишем далее предлагаемую авторами схему работы РВС при нормальном режиме ее функционирования и при обнаружении вторжения.

1) Для каждой вершины вторичной сети требуется составить таблицу маршрутизации, т. е. оптимальным образом распределить потоки по каналам первичной сети. При нормальном функционировании обмен информацией осуществляется по таблице маршрутизации. Система мониторинга при этом контролирует входящие потоки в каждой вершине.

2) Если в ходе мониторинга состояния сети обнаружено, что суммарный поток в вершину x_i выходит за границы, определенные при нормальном функционировании сети, то эту вершину удаляем (запрещаем потокам передавать данные через эту вершину).

3) В случае, если мы удалили какую-то вершину, требуется перераспределить потоки, которые использовали данную вершину в качестве промежуточного маршрута, т. е. составить новую таблицу маршрутизации. Конечно, может случиться так, что после действий, определенных в 2), какие-то вершины в сети перестанут быть связными. Будем считать, что структура сети избыточна и ее ключевые элементы существуют в нескольких экземплярах, т. е. при отказе одного из них функционирование сети обеспечивают другие. Таким образом, в подконтрольной сети имеются альтернативные пути следования потоков, и неразрешимость задачи передачи потока от одной вершины к другой может возникнуть только в случае, если удалена одна из концевых вершин потока (отправитель или получатель).

Рассмотрим действия агента, который хочет получить контроль над всеми информационными потоками, проходящими между объектами сети передачи данных. Их суть состоит в следующем.

Агент выбирает вершину вторичной сети $x_A \in Y$, через которую проходит наибольшее количество потоков, и затем модифицирует таблицы маршрутизации начальных вершин потоков, которые не используют в качестве промежуточного маршрута вершину x_A.

Цель агента — изменить исходную маршрутизацию на объектах сети таким образом, чтобы все потоки проходили через вершину x_A.

Если поток при нормальном функционировании сети проходил через маршрут $i \to x_1 \to \cdots, \to x_k \to j$, где $i, x_1, \cdots, x_k, j \in Y$, то после действий агента данный маршрут будет иметь вид: $i \to y_1 \to \cdots, \to x_A \to \cdots \to y_p \to j$, где $i, y_1, \cdots, x_A, \cdots, y_p, j \in Y$. При этом минимизируется время доставки от вершины i до вершины x_A и от вершины x_A до вершины j (для поиска маршрута с минимальным временем используется алгоритм Дийкстры).

Метод обнаружения вторжения в сеть агента

Для мониторинга моделируемых потоков применяем метод, основанный на статистическом анализе данных. Он базируется на том факте, что мы знаем о предыдущих наблюдениях величины $F_i = (f_1^i, \cdots, f_n^i)$ в промежутке времени $[t_k, t_{k+1}]$, $\bigcup_{k=0}^{p} = [0, T]$. Здесь f_l^i — величина входящего в вершину i информационного потока, передаваемого из вершины l. Интервал $[0, T]$ — достаточно большой промежуток времени мониторинга сети, а p — количество разбиений этого интервала.

Для параметра «суммарный входящий поток» строится доверительный интервал (с учетом интенсивностей λ_{ij}) в каждом промежутке времени $[t_k, t_{k+1}], k = \overline{0, p}$. Новое наблюдение является аномальным, если оно не укладывается в границах доверительного интервала.

Рассмотрим на примере моделирование атаки «внедрение ложного агента» и обнаружение этой атаки. Пусть информационная сеть моделируется гиперсетью S с графом первичной сети PS (рис. 1) и графом вторичной сети WS (рис. 2).

На рис. 1 показана сеть каналов, по которым будут реализованы ребра вторичной

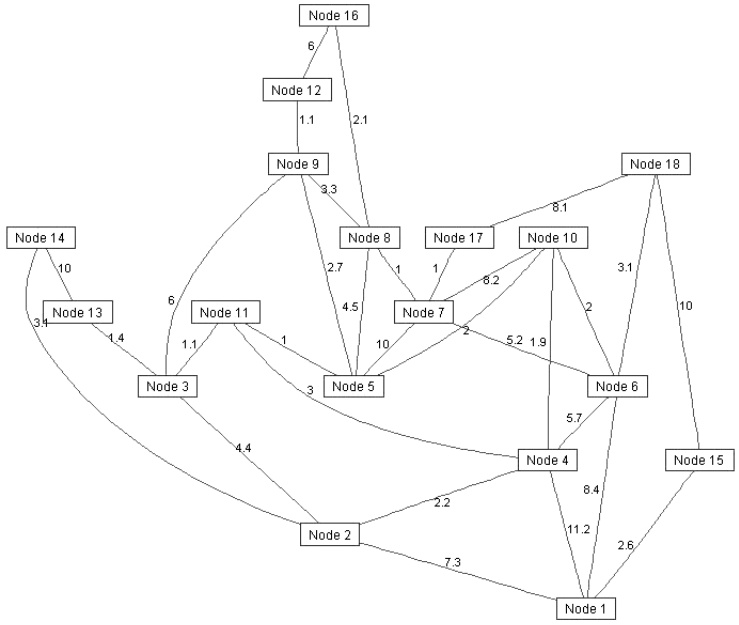

Рис. 1. Граф первичной сети

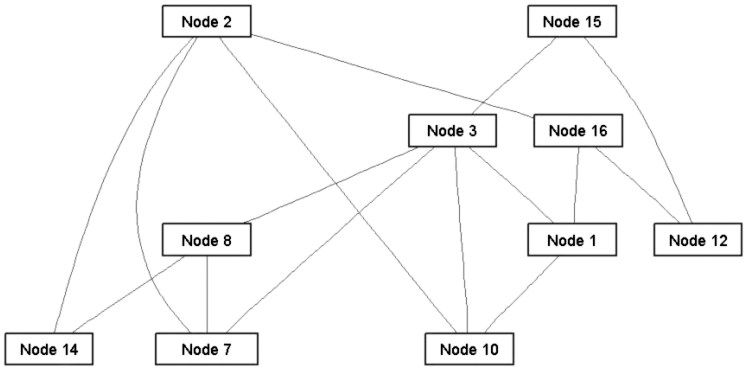

Рис. 2. Граф вторичной сети

сети. Указана скорость прохождения потоков информации по каждой ветви графа PS. Несмотря на то, что в реальных сетях кроме этого существуют ограничения на емкость каналов, в данной задаче емкость не учитывается. Причина в том, что такой учет приводит лишь к дополнительным ограничениям при поиске возможного маршрута реализации, однако не меняет существенных свойств моделируемой атаки на PBC.

Для каждого узла вторичной сети формируется маршрутная таблица, в которой записываются кратчайшие пути доставки сообщений от этого узла ко всем другим, с которыми эта вершина связана в графе вторичной сети. Маршруты проходят по существующим ребрам графа первичной сети (т. е. решается задача оптимального вложения вторичной сети в первичную).

Рассмотрим эту задачу на примере формирования маршрутной матрицы в узле 2. Для каждого ребра (2–7, 2–10, 2–14, 2–16) ищем кратчайший путь по ветвям первичной

сети (с учетом времени прохождения сообщений по каждому каналу первичной сети):

2 — 7: time_of_delivery = 11,3, path = (2, 4, 10, 6, 7)
2 — 10: time_of_delivery = 4,1, path = (2, 4, 10)
2 — 14: time_of_delivery = 3,1, path = (2, 14)
2 — 16: time_of_delivery = 12,7, path = (2, 4, 10, 5, 8, 16).

На рис. 3 показано вложение ребра 2–16 в первичную сеть по оптимальному маршруту, минимизирующему время доставки.

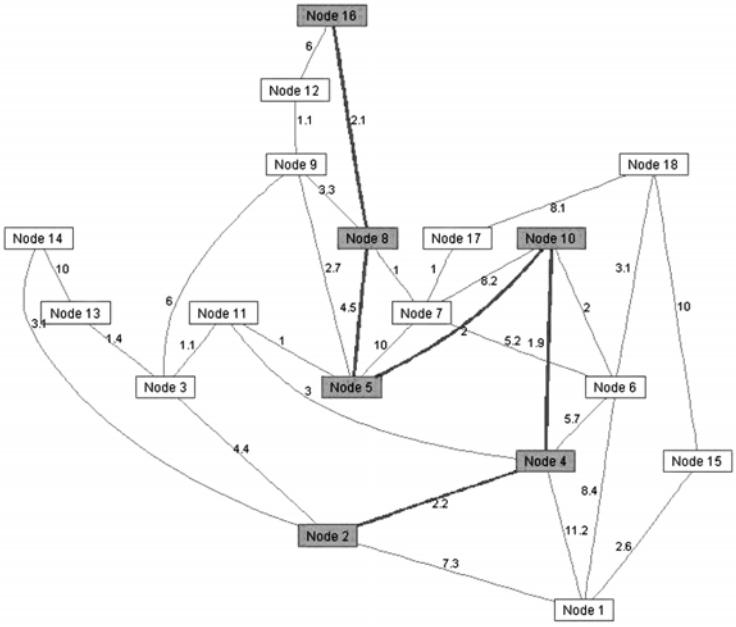

Рис. 3. Маршрут потока 2–16

Применяя подобный алгоритм действий для всех вершин вторичной сети, сформируем таблицу маршрутизации 1 для нормального режима функционирования сети (работа сети до атаки):

Таблица маршрутизации 1

Поток	Маршрут	Время доставки (мс)
2—16	2—4—10—5—8—16	12,7
2—10	2—4—10	4,1
2—7	2—4—10—6—7	11,3
2—14	2—14	3,1
3—8	3—11—5—8	6,7
3—7	3—11—5—8—7	7,7
3—10	3—11—5—10	4,2
3—15	3—2—1—15	14,3
1—16	1—6—7—8—16	16,7

1—10	1—6—10	10,4
12—15	12—9—5—10—6—1—15	18,8
12—16	12—16	6,0
8—7	8—7	1,0
8—14	8—5—10—4—2—14	13,7

Агент, имеющий доступ к таблицам маршрутизации всех вершин просматриваемой сети, отыскивает наиболее удобную вершину для внедрения. В данном случае вершина, через которую проходит наибольшее количество потоков — это вершина 10. Целью работы агента является перераспределение всех потоков, которые не проходят через эту вершину, так, чтобы их маршрут следовал через выбранную вершину 10.

Результатом этого становится следующая таблица маршрутизации 2 (т. е. работа сети после внедрения агента):

Таблица маршрутизации 2

Поток	Маршрут	Время доставки (мс)
2—16	2—4—10—5	12,7
2—10	2—4—10	4,1
2—7	2—4—10—6—7	11,3
2—14	2—4—10—4—2—14	11,3
3—8	3—11—5—10—5—8	10,7
3—7	3—11—5—10—6—7	11,4
3—10	3—11—5—10	4,2
3—15	3—11—5—10—6—1—15	17,5
1—16	1—6—10—5—8—16	19,0
1—10	1—6—10	10,4
12—15	12—9—5—10—6—1—15	18,8
12—16	12—9—5—10—5—8—16	14,4
8—7	8—5—10—6—7	13,7
8—14	8—5—10—4—2—14	13,7

Теперь все потоки проходят через вершину, в которой агент считывает информацию. Мониторинг сети, осуществляемый на каждом отрезке времени, позволяет заметить отклонение от нормального функционирования системы. Обнаружив агента (методом, описанным выше), удаляем вершину, в которой он находится, и перераспределяем все информационные потоки вторичной сети, так как они все проходили через вершину с агентом. Таким образом, в новом графе (без вершины с агентом) строится новая таблица маршрутизации. Естественно, что время доставки пакетов внутри сети меняется — на некоторых маршрутах оно увеличилось.

Результатом для описанного выше примера будет следующая таблица маршрутизации 3:

Таблица маршрутизации 3

Поток	Маршрут	Время доставки (мс)
2—16	2—4—11—5—8—16	12,9
2—10	————	
2—7	2—4—11—5—8—7	11,8
2—14	2—14	3,1
3—8	3—11—5—8	6,7
3—7	3—11—5—8—7	7,7
3—10	—————	
3—15	3—2—1—15	14,3
1—16	1—6—7—8—16	16,7
1—10	—————	
12—15	12—9—5—11—4—2—1—15	20,0
12—16	12—16	6,0
8—7	8—7	1,0
8—14	8—5—11—4—2—14	13,9

Видно, что для тех ребер вторичной сети, где вершина 10 была концевой (ребра 2–10, 3–10, 1–10), реализации путей теперь нет, так как нет вершины-получателя. В реальной ситуации такая проблема решается либо заменой удаленной вершины на другую (меняются графы первичной и вторичной сетей), либо восстановлением через некоторое время удаленной вершины, так как после обнаружения агента его действия уже не могут представлять опасности.

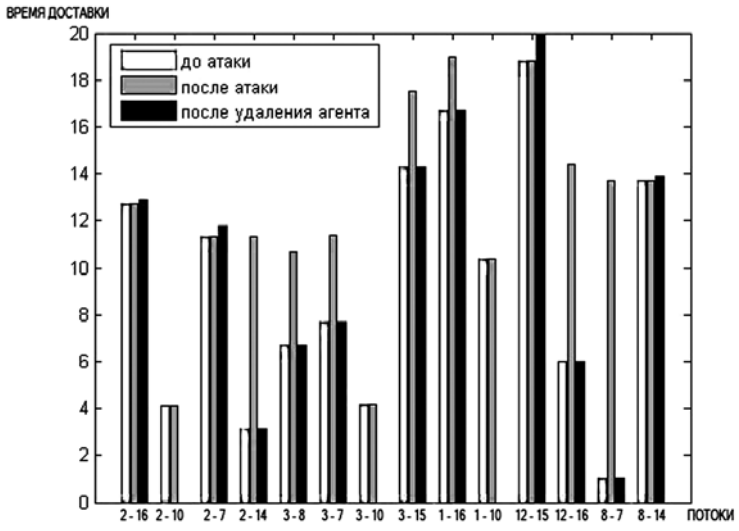

Рис. 4. Изменение времени доставки

На рис. 4 показано изменение времени доставки пакетов после внедрения агента и после его удаления. Видно, что после атаки «внедрение агента» на некоторых потоках увеличилось время доставки сообщений (маршрут увеличился — теперь сообщения до-

ставляются не по кратчайшему пути, а с заходом в вершину с ложным агентом). После удаления вершины с агентом некоторые потоки вернулись на прежние маршруты, и время доставки для них стало тем же, что и до атаки. На некоторых же маршрутах (например, 2–7 и 12–15) время доставки увеличилось, так как кратчайший путь между этими вершинами раньше проходил через вершину с агентом, а ее удаление привело, соответственно, к увеличению длины маршрута.

Заключение

Процесс обнаружения информационных атак на ресурсы РВС является весьма сложным технологическим процессом, который связан со сбором необходимых данных о процессе функционирования РВС, их анализом и выявлением факта атаки.

Применяя модель гиперсети в качестве модели сети передачи данных, с помощью мониторинга различных характеристик (например, передаваемых потоков информации) можно отслеживать несанкционированный доступ в сеть.

Для эффективного обнаружения атаки на всех стадиях ее жизненного цикла необходимо комбинированное использование как лингвистических, так и геометрических методов. Реализация такого комплексного подхода к проблеме выявления атак позволит значительно снизить риск успешного вторжения в РВС.

Список литературы

1. *Медведовский И.Д., Семьянов П.В., Леонов Д.Г.* Атака на Internet. М.: Изд-во ДМК, 2000.

2. *Васенин В. А., Галатенко А. В., Корнеев В. В. и др.* Математическое и программное обеспечение активного аудита больших распределенных систем // Математика и безопасность информационных технологий. М.: МЦНМО, 2005.

3. *Сердюк В.* Вы атакованы — защищайтесь! // Безопасность. 2003. № 9, http://www.bytemag.ru

4. *Лукацкий А.* Обнаружение атак. СПб.: БХВ-Петербург, 2001.

5. *Корт С. С.* Методы выявления нарушений безопасности: http://www.kiev-security.org.ua/box/12/113.shtml

6. *Попков В. К.* Математические модели связности. Новосибирск: Изд-во ИВМиМГ СО РАН, 2006.

7. *Классификация* сетевых атак: http://webdocs.ru/content-364.html.

8. *Ту Дж., Гонсалес Р.* Принципы распознавания образов. М.: Мир, 1978.

9. *Котенко И. В.* Многоагентные модели противоборства злоумышленников и системы защиты в сети Интернет // Математика и безопасность информационных технологий. М.: МЦНМО, 2005.

10. *Detoisien E.* External Attacks // LinuxFocus. 2003. March, No. 282: http://linuxfocus.org.

11. *Hansman S, Hunt R.* A Taxonomy of Network and Computer Attacks // Computers

and Security 2005. Vol. 24. No. 1. P. 31–43.

12. *Mircovic J., Martin J., Reiher P.* A Taxonomy of DDos Attacks and DDos Defense Mechanisms // Computer Communication Review. 2004. Vol. 34. Is. 2. P. 39–53.

Материал поступил в редколлегию 05.09.2007

Адреса авторов

СОКОЛОВА Ольга Дмитриевна
РОССИЯ, 630090, г. Новосибирск, 90
Пр. Акад. Лаврентьева, 6
ИВМ и МГ СО РАН
e-mail: olga@rav.sscc.ru

ДМИТРИЕВ Максим Николаевич
РОССИЯ, 630090, г. Новосибирск, 90
Ул. Пирогова, 2
Новосибирский государственный университет
e-mail: mnd@ngs.ru