УДК 004.094

КОМБИНИРОВАНИЕ МЕХАНИЗМОВ ЗАЩИТЫ ОТ СКАНИРОВАНИЯ В КОМПЬЮТЕРНЫХ СЕТЯХ

А. А. Чечулин,
младший научный сотрудник
И. В. Котенко,
доктор техн. наук, профессор
Санкт-Петербургский институт информатики и автоматизации РАН

Предлагается подход к комбинированию различных механизмов обнаружения сетевого сканирования в компьютерных сетях, который позволяет существенно повысить эффективность обнаружения за счет уменьшения количества ложных срабатываний и повышения точности обнаружения сканирования. Дается представление об отдельных частных методиках обнаружения сканирования. Рассматриваются основные принципы их комбинирования, а также дополнительные архитектурные улучшения общей модели обнаружения сканирования. Предлагается подход к автоматической настройке параметров используемых механизмов на основе статистических данных об анализируемом трафике.

Ключевые слова *— защита информации, компьютерные сети, сетевое сканирование, обнаружение сетевых атак, комбинирование механизмов защиты.*

Введение

Актуальной задачей защиты информации в компьютерных сетях является исследование перспективных механизмов обнаружения и предотвращения сканирования сети. Сканирование сетей может использоваться как злоумышленниками для сбора информации об атакуемой компьютерной сети, так и сетевыми червями в процессе их распространения.

Представляется, что при большом количестве узлов в контролируемой компьютерной сети и многообразии работающих в ней приложений перспективным подходом к обнаружению и защите от сетевых атак является согласованное использование комплекса различных механизмов защиты [1, 2]. Это объясняется, в первую очередь, тем, что отдельные методы защиты ориентированы главным образом на определенный тип трафика и виды атак, а комбинированные механизмы позволяют объединить преимущества отдельных методов и нивелировать их недостатки.

В статье предлагается и исследуется подход к комбинированию различных механизмов обнаружения сетевого сканирования в компьютерных сетях, который позволяет значительно повысить эффективность обнаружения. Хотя в существующих работах (например, в [3–9]) задача разработки отдельных частных механизмов обнаружения сканирования в достаточной степени исследована, проблема повышения эффективности этих механизмов и возможности их автоматической настройки в соответствии с текущей сетевой обстановкой (используемыми в сети приложениями, параметрами трафика и т. п.), в том числе за счет комбинирования, остается не решенной.

Решение поставленной проблемы в статье представлено на основе последовательного выполнения следующих задач: определения нескольких эффективных механизмов защиты; выделения наиболее важных параметров трафика и классификации трафика по ним; определения показателей эффективности обнаружения; разработки общего подхода к комбинированию механизмов защиты, в том числе разработки и обучения алгоритма подбора оптимальных параметров для механизмов на каждом из выделенных классов трафика и разработки и обучения алгоритма комбинирования механизмов защиты.

Механизмы обнаружения сканирования

За основу предлагаемого комбинированного подхода к обнаружению сетевого сканирования в компьютерных сетях было принято использование нескольких семейств механизмов, базирующихся

на следующих методиках: методике «дросселирования/регулирования вирусов» (*Virus Throttling*) и ее модификации (VT-S и VT-C) [4, 5]; методиках, основанных на анализе неудачных соединений (*Failed Connection* — FC) [6]; методиках, использующих метод «порогового случайного прохождения» (*Threshold Random Walk* — TRW) [7, 8]; методиках ограничения интенсивности соединений на основе кредитов доверия (*Credit Based Rate Limiting* — CB) [9]. Представим основные элементы реализации нескольких из указанных механизмов.

Методика «дросселирования/регулирования вирусов» для реализации на коммутаторах (VT-S) основывается на следующей модели обработки сетевого трафика. Для каждого узла существует список длиной N для отметки значений хэш-функции на последние адреса, к которым приходили запросы на соединения. Хэш-функция принимает значения от 0 до $N - 1$, при получении значения k на k-е место в списке записывается 1. Запрос на соединение идентифицируется по пакету TCP-SYN. Задается минимальный интервал между поступлениями запросов на соединение к новым адресам (rate threshold). Если запрос поступил спустя время, меньшее, чем заданный порог, адрес получателя хэшируется и происходит запись в таблицу хэшей для этого отправителя. В противном случае список отметок хэшей очищается и добавляется только информация по текущему запросу. Если таблица хэшей полна, то все запросы от отправителя считаются вредоносными на время T. Используется протокол TCP, анализируются пакеты TCP SYN. В пакетах контролируются поля «source IP» и «destination IP».

Предлагаемая модификация методики Virus Throttling на основе метода CUSUM (VT-C) [3] заключается в следующем. Для каждого узла существует счетчик C. $CMax$ — это порог количества запросов к новым адресам подряд. Счетчик увеличивается, если прошло менее $RMin$ секунд с момента прихода последнего запроса с незарегистрированным адресом получателя. Запрос на соединение идентифицируется по пакету TCP-SYN. При превышении счетчиком C заданного порога запросы на соединение от данного узла считаются вредоносными. Так же как и для методики Virus Throttling, используется протокол TCP и анализируются пакеты TCP SYN. В пакетах контролируются поля «source IP» и «destination IP».

В *методиках* FC, основанных на анализе неудачных соединений, используется и хранится следующая информация: адрес узла, инициировавшего соединение, частота ошибок, время соединения и счетчик ошибочных соединений. При превышении порога частоты добавления новых записей в хэш-таблицу создается запись для индивидуального узла. В дальнейшем при увеличении значения счетчика ошибочных соединений на фиксированную величину (например, 100) для этой записи обновляется значение частоты ошибок f (failure rate). При превышении величины failure rate заданного порога запросы на соединение от данного узла считаются вредоносными. При реализации данных методик используется протокол TCP, анализируются пакеты TCP RESET. В пакетах контролируются поля «source IP» и флаг TCP RST.

Методика TRW базируется на следующей модели вычислений. Для анализа узла, демонстрирующего большую сетевую активность, на предмет возможности проведения сканирования используется метод последовательного тестирования гипотез (*Sequential Hypothesis Testing*).

Пусть H_1 — гипотеза, что узел r демонстрирует повышенную сетевую активность (проводит сканирование), H_0 — гипотеза, что узел не демонстрирует повышенной сетевой активности (не проводит сканирование), а Y_i — переменная, характеризующая результат попытки соединения с i-м узлом, к которому направлен запрос соединения. Эта переменная может иметь следующие значения: 0, если соединение установлено; 1, если соединение не установлено.

Предполагается, что условия наступления гипотез H_i — случайные величины $Y_i \,|\, h_i = 1, 2, \ldots$ — независимы и распределены равномерно. В этом случае можно выразить распределение бернуллиевой случайной величины Y_i так:

$$P_r = [Y_i = 0 \,|\, H_0] = \theta_0,\ P_r = [Y_i = 1 \,|\, H_0] = 1 - \theta_0,$$
$$P_r = [Y_i = 0 \,|\, H_1] = \theta_1,\ P_r = [Y_i = 1 \,|\, H_1] = 1 - \theta_1.$$

Предположение, что попытка установления соединения от неинфицированного узла имеет большую вероятность успешного исхода, чем от инфицированного, подразумевает выполнение условия $\theta_0 > \theta_1$.

Учитывая эти две гипотезы, при принятии решения о вредоносности/безопасности узла возможны следующие четыре варианта:

1) «обнаружение сканирования» — выбирается гипотеза H_1 при реальном сканировании с узла;

2) «пропуск атаки (false negative)» — выбирается гипотеза H_0 при реальном сканировании с узла;

3) «отсутствие сканирования» — выбирается гипотеза H_0 при реальном отсутствии сканирования с узла;

4) «ложное срабатывание (false positive)» — выбирается гипотеза H_1 при реальном отсутствии сканирования с узла.

Введем обозначения P_D — вероятность обнаружения сканирования с узла и P_F — вероятность

ложного срабатывания с целью определить требования к выполнению методики. Для максимальной эффективности выполнения методики предполагается, что верны следующие условия: $P_D \geq \beta$, $P_F \leq \alpha$, $\alpha = 0{,}01$, $\beta = 0{,}99$.

При наступлении анализируемого события вычисляется следующее вероятностное отношение:

$$\Lambda(\mathbf{Y}) \equiv \frac{\Pr[\mathbf{Y} \mid H_1]}{\Pr[\mathbf{Y} \mid H_0]} = \prod_{i=1}^{n} \frac{\Pr[Y_i \mid H_1]}{\Pr[Y_i \mid H_0]},$$

где $\mathbf{Y}$ — вектор наблюдаемых событий; $P_r[Y_i \mid H_i]$ — условная вероятность наступления события, при котором последовательность $\mathbf{Y}$ полностью соответствует гипотезе H_i.

Затем выполняются сравнения вероятностного отношения Λ с верхним (η_1) и нижним (η_0) пороговыми значениями и по результатам сравнения принимается решение о наличии/отсутствии сканирования с узла: $\Lambda(\mathbf{Y}) \leq \eta_0 \Rightarrow H_0$, $\Lambda(\mathbf{Y}) \geq \eta_1 \Rightarrow H_1$. В случае если $\eta_0 < \Lambda(\mathbf{Y}) < \eta_1$, продолжается ожидание дополнительных событий для более точной идентификации. Пороговые значения выбираются следующим образом:

$$\eta_1 = \frac{\beta}{\alpha} \text{ и } \eta_0 = \frac{1-\beta}{1-\alpha}.$$

Выбор параметров сетевого трафика

Для классификации трафика в данной статье выделено около 30 параметров. В качестве основных были выбраны следующие: интенсивность соединений; процент запросов на установление соединения (TCP SYN пакеты) от общего количества пакетов; продолжительность соединений; интервалы между приходами пакетов в соединении; среднее количество пакетов на один хост-источник; отношение количества запросов на установление соединений (TCP SYN) с количеством подтверждений на установление соединений (TCP SYN-ACK) и др.

Задача обучения, заключающаяся в поиске оптимальных конфигураций используемых отдельных механизмов защиты на основе параметров трафика, объективно осложнена высокой корреляцией этих параметров в реальных сетях.

Выбранные параметры трафика (средняя частота приходящих пакетов, процент TCP SYN запросов и др.) представляют собой некоторые статистические данные, собранные за определенный период времени.

Существует несколько подходов к вычислению таких параметров. Наиболее простой подход — это хранение в памяти всех пакетов за анализируемый период времени и вычисление статистических параметров по этой истории. Основным недостатком его является большая ресурсоемкость, особенно для высокоскоростных сетей.

Другой вариант — хранение только самих значений статистических параметров. Например, значение средней частоты трафика за период можно вычислить, имея количество пришедших пакетов с начала периода. Этот путь значительно менее ресурсоемкий, однако по истечении разумно небольшого периода статистические значения приходится обнулять, так как средние значения показателя (например, за неделю) уже использовать бессмысленно. После обнуления параметров и до их последующего расчета использование методов комбинирования невозможно.

Для преодоления этой ситуации предлагается считать статистику параметров трафика сразу в нескольких временных «окнах», образуемых со сдвигом по времени. При вычислениях используются параметры трафика самого «старого» окна. Количество окон может задаваться произвольно при настройке системы.

Показатели эффективности механизмов обнаружения

Показатели эффективности механизмов обнаружения хостов со сканирующим трафиком основаны на бинарной классификации хостов вида «содержит/не содержит» вредоносный трафик. Данные показатели основываются на следующей матрице классификации:

	Реально содержит	Реально не содержит
Распознан как вредоносный	*a* (true positive)	*b* (false positive)
Распознан как не вредоносный	*c* (false negative)	*d* (true negative)

Здесь *a* — количество хостов, трафик которых содержит сканирующий трафик и правильно распознанных системой; *b* — количество хостов, трафик которых не содержит сканирующий трафик, но распознанных системой как вредоносные; *c* — количество хостов, трафик которых содержит сканирующий трафик, но распознанных системой как безопасные; *d* — количество прочих хостов.

В проводимых авторами исследованиях по анализу механизмов обнаружения сканирования, наряду с показателями ошибок первого и второго рода — степенью ложных срабатываний (false positive) и степенью пропусков атак (false negative), используются также следующие интегрированные показатели, вычисляемые на основе показателей ошибок первого и второго рода: полнота, точность, аккуратность, *F*-мера и ошибка.

Полнота r вычисляется как отношение правильно распознанных вредоносных хостов к обще-

му количеству вредоносных хостов. Этот параметр характеризует способность системы распознавать вредоносные хосты, но не учитывает количество неправильно распознанных безопасных хостов.

Точность p определяется как отношение правильно распознанных вредоносных хостов к общему количеству хостов, распознанных как вредоносные. Точность характеризует способность системы распознавать только реально вредоносные хосты.

Аккуратность вычисляется как отношение правильно принятых системой решений к общему числу решений.

F-мера используется как единая метрика, объединяющая метрики полноты и точности; вычисляется по формуле $F = 2pr/(p + r)$.

Ошибка определяется как отношение неправильно принятых системой решений к общему числу решений.

В статье для определения эффективности механизмов обнаружения используются показатели false positive, false negative, *F*-меры и аккуратности.

Сущность подхода к комбинированию механизмов обнаружения

Комбинирование механизмов обнаружения и реагирования основывается на использовании в процессе работы тех механизмов, которые проявили себя наилучшим образом при обучении на трафике, максимально близком по всем параметрам к тому, на котором происходит реальное обнаружение сканирования.

Для решения задачи комбинирования применялись следующие методы интеграции конечных результатов отдельных механизмов:

1) выбор наилучшего (оптимального) механизма и отключение остальных;

2) простого большинства голосов;

3) взвешенного большинства голосов;

4) комбинирование с помощью методов Data Mining.

Все предложенные методы предполагают первоначальное тестирование каждого метода защиты и нахождение лучшего метода (с заданными параметрами) при определенных характеристиках трафика. Предполагается, что данные методы позволяют использовать наиболее сильные стороны каждого механизма для каждого класса трафика.

Сложностью реализации *первого метода* является то, что не всегда можно осуществить полное упорядочение механизмов защиты по их степени применимости, в первую очередь, в силу разнообразия возможного трафика и его характеристик.

Второй метод использует данные от всех механизмов защиты, но при этом предполагает, что механизмы работают с одинаковой степенью точности и, как и первый метод, не учитывает степень применимости механизмов защиты для различных ситуаций.

Представим более детально предлагаемый вариант реализации *третьего метода*, основанного на вычислении весовых коэффициентов для каждого механизма.

Весовые коэффициенты определяются на основе параметров трафика, которые должны быть выбраны заранее. Функция $W(d, x_1, ..., x_n)$ вычисления весовых коэффициентов принимает в качестве аргументов тип (класс) d механизма защиты и точку декартова n-мерного гиперпространства, заданного параметрами $(x_1, ..., x_n)$. Размерность гиперпространства определяется количеством выбранных параметров трафика.

По результатам предварительного тестирования (обучения) для каждого параметра строится зависимость значения *F*-меры при использовании различных механизмов защиты от значения определенного параметра.

Определим функцию $F(x)$ зависимости значения *F*-меры от параметра трафика x следующим образом: на основе данных, полученных на этапе обучения, строится интерполяционный полином Лагранжа, который позволяет получить предположительные значения *F*-меры для любого значения параметра x.

Обозначим через $F_i(d, x)$ функцию зависимости значения *F*-меры от i-го параметра для механизма защиты d. Тогда функцию вычисления весовых коэффициентов определим следующим образом:

$$W(d, x_1, \ldots, x_n) = \frac{\sum_{i=1}^{n} F_i(d, x)}{n}.$$

Значение функции W тем меньше, чем меньше значение *F*-меры для данного механизма при данных характеристиках трафика.

При вычислении данной функции предполагается, что все параметры трафика являются равнозначными. Можно также снабдить параметры трафика весами согласно их важности для принятия решения. Тогда слагаемые суммы в числителе следует умножить на соответствующий весовой коэффициент.

Пусть функция $HF(f, d)$ принятия решения о вредоносности хоста, которая получает значение после получения пакета f механизмом защиты d, равна –1, если хост признан вредоносным, и +1, если хост признан безопасным.

Тогда функция голосования для m механизмов защиты при принятии решения о вредоносности хоста после получения пакета f при характеристиках трафика $x_1, ..., x_n$ выглядит следующим образом:

$$V(f, x_1, \ldots, x_n) = \sum_{j=1}^{m} (W(d_j, x_1, \ldots, x_n) HF(f, d_j)).$$

Если значение $V(f, x_1, ..., x_n)$ отрицательно, то хост признается вредоносным, в противном случае — безопасным.

Четвертый метод комбинирования использует данные от всех механизмов защиты, и эти данные вместе с основными параметрами трафика передаются в обученный заранее классификатор (например, Naïve Bayes [10]). В результате на основе входных данных и предварительного обучения классификатор принимает решение о вредоносности анализируемого трафика.

Эти же методы комбинирования применимы и к автоматическому выбору оптимальных параметров механизмов обнаружения сканирования. Основная идея данного подхода заключается в том, что целесообразно также распространить комбинирование на разные конфигурации отдельных механизмов, меняя параметры настройки механизма в зависимости от текущей ситуации в сети.

Основные требования к разрабатываемым комбинированным механизмам защиты в целом соответствуют функциональным требованиям разработанного проактивного подхода к обнаружению сканирования в целом [2], а именно адекватность и оперативность обнаружения, эффективность использования системных ресурсов, автоматическое выполнение, возможность обнаружения различных видов сетевых червей.

Среда для проведения экспериментов

Для моделирования и оценки методов комбинирования механизмов защиты и отдельных механизмов защиты, а также выбора для них оптимальных параметров разработана автоматизированная методика и программные средства исследования (моделирования и анализа) механизмов защиты.

Суть практической оценки методов комбинирования и механизмов защиты сводится к проведению комплекса экспериментов на основе использования программной среды для различных значений входных параметров и измерению показателей эффективности исследуемых механизмов защиты.

При реализации программных средств моделирования и анализа отдельных механизмов, методов комбинирования механизмов и методики выбора оптимальных параметров для механизмов использовалась *архитектура*, включающая следующие компоненты:

- *модели источников трафика,* предназначенные для предоставления сетевого трафика исследуемым механизмам защиты, включая как модели трафика сканирования, так и модели обычного сетевого трафика;
- *модели предобработки и синхронизации источников трафика,* служащие для приведения трафика из формата источников в формат, удобный для анализа исследуемыми механизмами защиты. Этот модуль также предназначен для синхронизации трафика при одновременном использовании нескольких источников трафика;
- *модели механизмов обнаружения сканирования*. Входными параметрами каждого механизма обнаружения являются те поля сетевого пакета, полученного от источника трафика, которые им обрабатываются. В качестве управляющих параметров вводятся различные внутренние параметры каждого механизма, которые влияют на его эффективность.

Реализованные программные средства моделирования и анализа механизмов защиты позволяют оценивать следующие основные показатели эффективности методов комбинирования и выбора оптимальных параметров:

- долю блокированного и (или) задержанного легитимного трафика (степень ложных срабатываний, false positives);
- долю пропущенного злонамеренного трафика (степень пропусков атак, false negatives);
- интегрированные показатели (полноту, точность, аккуратность, *F*-меру и ошибку);
- время реакции на атаку.

При проведении экспериментов использовался комбинированный способ моделирования трафика, заключающийся в применении в качестве входных данных различных записей реального трафика с дополнением их необходимым для исследования трафиком. В данном случае необходимый для исследования трафик — это трафик сканирования и трафик «быстрых» приложений, таких как P2P или NetBIOS/NS. Исследования проводились как на записях реального трафика, так и на сгенерированном трафике с подключением различных модулей генерации моделей трафика.

Для сравнения механизмов использовались трафики двух типов:

- трафик сети уровня предприятия с большим набором различных приложений (в том числе сканеров уязвимостей) без преобладания трафика какого-либо из них;
- трафик локальной сети с преобладанием приложений P2P.

Выбор таких типов трафика обоснован тем, что механизмы обнаружения и реагирования против сетевого сканирования должны их точно обнаруживать как в трафике легитимных приложений, так и в трафике приложений, создающих большое количество соединений с различными узлами. Последнее особенно важно, так как трафик приложений, создающих большое количе-

ство соединений с различными узлами, похож на трафик, появляющийся при сканировании.

Результаты экспериментов

Для проведения экспериментов указанные виды трафика были смешаны с трафиком сканирования. Использовались следующие параметры сканирования: скорость отправки запросов на соединение — 50 запросов/с, вероятность успешного соединения — 30%, вероятность получения TCP-RST пакета — 30%, выбор IP-адресов для сканирования — случайный.

Приведем сначала результаты тестирования отдельных механизмов обнаружения (VT-S, VT-C, FC, TRW и CB) для различных видов трафика.

Средние значения процента ложных срабатываний (FP) и процента пропуска атак (FN) для отдельных механизмов на обычном и P2P-трафике без трафика сканирования представлены в табл. 1.

Средние значения процента ложных срабатываний (FP), процента пропуска атак (FN), объема памяти (V) и аккуратности (A) для отдельных механизмов на обычном и P2P-трафике, смешанном с трафиком сканирования, показаны в табл. 2.

Отметим, что во многих случаях для трафика P2P предложенный механизм Virus throttling на основе метода CUSUM (VT-C) превзошел используемый в настоящее время на сетевых коммутаторах механизм VT-S. Этот факт свидетельствует о целесообразности его реализации на коммутаторах и использовании его вместо VT-S или совместно с VT-S в сетях с трафиком P2P.

В настоящее время авторы продолжают проводить эксперименты для различных реализаций методов комбинирования механизмов защиты.

Проведенные эксперименты показали, что использование методов комбинирования, основанных на выборе наилучшего (оптимального) механизма, применении взвешенного большинства голосов и методов Data Mining для различных типов трафика приводит в большинстве случаев к улучшению показателей эффективности.

Эти эксперименты были проведены на смеси обычного трафика, трафика, содержащего соединения P2P, и трафика сетевого червя таким образом, чтобы основные параметры хостов заметно менялись. Это осуществлялось, например, путем добавления в трафик хостов с некоторого момента времени трафика P2P или, наоборот, удаления трафика P2P с хоста, на котором до этого времени этот трафик был.

При использовании метода комбинирования, основанного на выборе оптимального механизма, для обычного трафика был автоматически выбран метод VT-C, а для трафика, содержащего P2P-трафик — метод CB. К сожалению, реализованный метод комбинирования не учитывал изменений в параметрах трафика, и при появлении в обычном трафике трафика P2P и, наоборот, в случае удаления трафика P2P комбинированный механизм начинал показывать результаты, заметно худшие даже в сравнении с некоторыми отдельными методами обнаружения сканирования.

Метод комбинирования на базе простого большинства голосов лучше реагировал на изменение параметров трафика, чем отдельные методы обнаружения, но результаты в каждом конкретном случае оказались хуже работы отдельных методов, оптимальных для данных параметров трафика.

■ ***Таблица 1.*** *Результаты работы механизмов защиты на трафиках без сканирования*

Механизм защиты	Обычный трафик		P2P-трафик	
	FP	FN	FP	FN
VT-S	0,022600	0	0,089913	0
VT-C	0,023400	0	0,061528	0
FC	0,005950	0	0,002946	0
TRW	0,004300	0	0,002311	0
CB	0,021800	0	0,080051	0

■ ***Таблица 2.*** *Результаты работы механизмов защиты на трафиках, смешанных с трафиком сканирования*

Вид трафика	Механизм защиты	FP	FN	V	A
Обычный трафик	VT-S	0,023300	0,000663	23496	0,998929
	VT-C	0,024100	0,000530	36224	0,998983
	FC	0,009680	0,054291	48572	0,972348
	TRW	0,004980	0,002416	28912	0,998216
	CB	0,025600	0,000344	9608	0,987114
P2P-трафик	VT-S	0,301541	0,0178923	34820	0,96833
	VT-C	0,101948	0,0317885	105272	0,964804
	FC	0,103761	0,013333	74108	0,98227
	TRW	0,0599373	0,0972275	127656	0,905997
	CB	0,200021	0,001258	22632	0,993393

При использовании метода комбинирования, основанного на взвешенном большинстве голосов, в качестве отслеживаемых параметров были выбраны интенсивность запросов на установление соединений, процент запросов на установление соединения (пакеты TCP SYN) от общего количества пакетов, продолжительность соединений, среднее количество пакетов на один хост-источник. В результате проведения экспериментов были получены следующие средние значения: процента ложных срабатываний FP = 0,009421 и процента пропуска атак FN = 0,008596.

Наилучшие результаты были показаны при применении для комбинирования методов Data Mining (в экспериментах использовался наивный байесовский подход). В результате проведенных экспериментов получены средние значения процента ложных срабатываний FP = 0,003529 и процента пропуска атак FN = 0,001286.

На процент ложных срабатываний влияет то, что после обнаружения сканирования с хоста хост блокируется, при этом блокируется и его нормальный (не вредоносный) трафик.

Заключение

В данной статье предложен подход к комбинированию механизмов обнаружения сканирования и выбору оптимальных параметров для этих механизмов. Представлено разработанное программное средство для проведения экспериментов и определения эффективности средств защиты на записях реального трафика.

По результатам экспериментов проведена оценка отдельных и комбинированных механизмов обнаружения сканирования. По полученным данным можно судить о том, что использование предложенных методов комбинирования, в частности метода комбинирования, основанного на взвешенном большинстве голосов, и методов Data Mining, а также настройка параметров для отдельных механизмов защиты в зависимости от статистических показателей трафика позволяет существенно улучшить эффективность работы этих механизмов.

В дальнейшей работе планируется проведение большого количества экспериментов для анализа эффективности методов комбинирования для различных смесей исследуемого сетевого трафика, разработка новых и совершенствование существующих отдельных механизмов обнаружения, исследование других схем кооперации отдельных механизмов, развитие разработанного программного средства в целях создания среды моделирования для механизмов защиты от сканирования.

Работа выполняется при финансовой поддержке РФФИ (проект № 10-01-00826-а), программы фундаментальных исследований ОНИТ РАН (проект № 3.2) и при частичной финансовой поддержке, осуществляемой в рамках проектов Евросоюза SecFutur и MASSIF.

Литература

1. **Чечулин А. А.** Обнаружение и противодействие сетевым атакам на основе комбинированных механизмов анализа трафика // Информационная безопасность регионов России (ИБРР-2009): Материалы VI Санкт-Петербургской межрегион. конф., 28–30 октября 2009 г. / СПОИСУ. СПб., 2009. С. 143–144.
2. **Котенко И. В., Воронцов В. В., Чечулин А. А., Уланов А. В.** Проактивные механизмы защиты от сетевых червей: подход, реализация и результаты экспериментов // Информационные технологии. 2009. № 1. С. 37–42.
3. **Чечулин А. А., Котенко И. В.** Исследование механизмов защиты от сетевых червей на основе методик Virus Throttling // Защита информации. Инсайд. 2008. № 3. С. 68–73.
4. **Williamson M.** Throttling Viruses: Restricting propagation to defeat malicious mobile code // Proc. of ACSAC Security Conf., Las Vegas, Nevada. 2002. P. 61–68.
5. **Twycross J., Williamson M.** Implementing and testing a virus throttle // Proc. 12th USENIX Security Symp., 2003. http://www.hpl.hp.com/techreports/2003/HPL-2003-103.html (дата обращения 09.10.2010).
6. **Chen S., Tang Y.** Slowing Down Internet Worms // 24th Intern. Conf. on Distributed Computing Systems (ICDCS'04), Tokyo, Japan, Mar. 2004. P. 312–319.
7. **Jung J., Paxson V., Berger A. W., Balakrishnan H.** Fast portscan detection using sequential hypothesis testing // Proc. of the 2004 IEEE Symp. on Security and Privacy, Oakland, California, USA, May 9–12, 2004 / IEEE Computer Society, 2004. P. 211–225.
8. **Weaver N., Staniford S., Paxson V.** Very fast containment of scanning worms // Proc. of the 13th USENIX Security Symp., Aug. 9–13, 2004. http://www.icsi.berkeley.edu/~nweaver/containment/containment.pdf (дата обращения 09.10.2010).
9. **Schechter S., Jung J., Berger A. W.** Fast Detection of Scanning Worm Infections // Proc. of the Seventh Intern. Symp. on Recent Advances in Intrusion Detection, French Riviera, France, Sept. 2004. P. 59–81.
10. **Zuev D., Moore A. W.** Traffic Classification using a Statistical Approach: Proc. of the 2005 ACM SIGMETRICS Intern. Conf. on Measurement and Modeling of Computer Systems, Banff, Alberta, Canada, June 06–10, 2005//Lecture Notes in Computer Science. Springer, 2005. Vol. 3431. P. 321–324.