УДК: 004.056

АКТИВНЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ SYN-flood АТАК

© 2012 **Д. А. Корнев**[1], **В. Н. Лопин**[2], **В. Г. Лузгин**[3]

[1]*аспирант каф. программного обеспечения и администрирования информационных систем* (ПОиАИС),
e-mail: northcomm@mail.ru

[2]*докт. техн. наук, профессор каф.* ПОиАИС,
e-mail: kzis3@yandex.ru

[3]*ст. преподаватель каф.* ПОиАИС,
e-mail: a543b210@mail.ru

Курский государственный университет

Рассмотрена технология активного метода обнаружения распределённых атак типа SYN flood на серверы информационных систем на ранних стадиях. Посредством подобных технологий возможно обеспечить эффективное раннее обнаружение, а также значительно влиять на интенсивность атаки и соответствующее значение риска, разворачивая масштабируемую, независимую и нетребовательную к ресурсам систему обнаружения на стороне защищаемого сервера.

Ключевые слова: активные методы обнаружения, обнаружение на ранних этапах, анализ полуоткрытых соединений.

Введение

Атаки типа SYN-flood являются наиболее распространенным типом распределенных атак отказа в обслуживании (Distributed Denial of Service attack, DDoS-атак). Наиболее желательным является *раннее* выявление подобных атак, однако традиционные (пассивные) методы обнаружения зачастую неточны именно на ранних стадиях из-за их зависимости от пассивного перехвата атакующих сигнатур. Метод, предлагаемый в данной статье, перехватывает атакующие сигнатуры, используя *схему активного зондирования*, что обеспечивает эффективность раннего обнаружения.

Вопросу защиты от распределённых атак отказа в обслуживании уделяется много внимания в последние несколько лет. Как было указано выше, традиционные пассивные методы обнаружения неточны на ранних стадиях и эффективны лишь на поздних этапах, когда присутствие атакующих сигнатур очевидно. Эффективные методы для обнаружения таких атак на ранней стадии до недавнего времени отсутствовали.

Одной из схем активного зондирования является *метод* DARB (the **D**el**A**y **P**ro**B**ing method), выполняющий анализ полуоткрытых соединений. Эксперименты показывают, что метод исследования задержек способен распознать те полуоткрытые соединения, которые вызваны SYN flood-атакой, и те, которые вызваны другими причинами, точно и на ранней стадии [1]. Метод получает задержки маршрутизаторов (узлов), посылая пакеты, содержащие специально установленные значения TTL (Time To Live, время жизни пакета) в заголовке IP. Результаты зондирования используются затем для надёжного (и с минимальными накладными расходами) обнаружения SYN-флуда. Такой подход более независим, нежели другие методы, требующие содействия сетевых устройств.

Чтобы обнаруживать SYN-flood-атаку на её ранней стадии, в данном подходе на

рассмотрение вносятся следующие основные идеи:

- подход базируется на том факте, что *нормальные полуоткрытые* соединения, обрабатываемые внутри сервера, существуют как результат перегрузки сетевого трафика, в то время как полуоткрытые соединения, вызванные SYN-флудом, исходят исключительно от злоумышленника;
- метод обнаружения является *активным* и *независимым*. Подход использует более активный механизм для поиска признаков SYN-флуда вместо пассивного перехвата атакующих сигнатур. Такой активный подход требует своего развёртывания лишь на стороне защищаемого сервера и не зависит от взаимодействия с другими сетевыми устройствами;
- используется надёжный и не требующий больших накладных расходов механизм для исследования (зондирования) задержек между сервером и клиентом. Рассматриваемый алгоритм DARB менее затратен и более надёжен, чем прямые методы исследования, такие как ping.

Активный метод обнаружения DARB

Активный подход к обнаружению начинается с определения того, является ли полуоткрытое соединение результатом SYN-флуда. *Нормальные полуоткрытые* соединения обычно вызваны простой перегрузкой сети (или ошибками, в результате которых теряется высланная клиентом комбинация SYN+ACK), в то время как полуоткрытые соединения, вызванные SYN-флудом, к перегруженности трафика не имеют никакого отношения. Если задержка между сервером и клиентом гораздо больше, чем нормальная задержка, вероятной причиной является перегруженный маршрутизатор, однако, если признаков перегрузки нет, полуоткрытое соединение рассматривается как *ненормальное* и может быть результатом SYN-flood-атаки.

Задержка между сервером и клиентом оценивается с помощью исследующего (зондирующего) метода DARB. Устанавливая различное время жизни (TTL) в IP-заголовке пакета, можно заставить пакет исчезать («умирать») на разных маршрутизаторах. Информация о «смерти» пакетов отправляется соответствующим маршрутизатором и обеспечивает возможность оценки задержки по пути пакета. Эти значения позволяют затем оценить вероятность того, является ли данное полуоткрытое соединение результатом SYN-flood-атаки или нет.

Полуоткрытое соединение – это такое состояние соединения, при котором оно является установленным на одном конце, в то время как другой конец либо недоступен, либо обладает некорректной информацией для соединения. Такое состояние обычно бывает вызвано незавершённостью процедуры «трёхкратного рукопожатия», обеспечиваемой протоколом TCP. Сервер в течение некоторого периода времени обрабатывает полуоткрытое соединение, пытаясь довести процесс установки соединения до конца, посылая клиенту SYN+ACK пакеты. В случае когда полуоткрытые соединения являются результатом перегруженности сети или ошибки, они могут классифицироваться как *нормальные полуоткрытые соединения*.

Однако полуоткрытые соединения могут наблюдаться и на атакуемом сервере, подверженном SYN-флуду – злоумышленник посылает жертве SYN-пакеты с подставными IP-адресами; сервер принимает эти SYN-запросы и переходит в состояние ожидания ответного ACK-пакета, посылая SYN+ACK-пакет по месту назначения, соответствующему IP источника. Серверу, однако, сложно отличить такой тип полуоткрытого соединения, вызваного SYN-флудом, от нормального полуоткрытого соединения. В результате атакуемый сервер будет ожидать истечения некоторого срока, и пытаться несколько раз заново посылать источнику SYN+ACK-пакеты. Такой тип полуоткрытых соединений следует классифицировать как *ненормальные*.

Центральной проблемой, таким образом, становится отличить нормальные полуоткрытые соединения от ненормальных. Базовым различием является то, что нормальные полуоткрытые соединения возникают из-за перегрузок сети, в то время как ненормальные к перегрузкам не имеют отношения, то есть задержка между узлами, вероятнее всего, такая же, как и при нормальном состоянии сети. Если полуоткрытое соединение вызвано перегрузкой, маршрут между сервером и клиентом имеет признаки перегруженности, например увеличенную задержку пакетов, повышенный уровень потерь и высокий объём очереди на перегруженном узле.

В рассматриваемом методе задержка на маршруте между сервером и клиентом исследуется подобно тому, как это делается с помощью утилиты traceroute. Слишком большая задержка рассматривается как признак перегруженности, и при таких обстоятельствах полуоткрытое соединение считается нормальным. В остальных случаях мы имеем дело с ненормальным полуоткрытым соединением.

В отличие от traceroute, DARB выбирает отдельные узлы по пути пакета для исследования задержек, а не все узлы на каждом транзитном участке («прыжке», *англ.* hop). DARB отслеживает сетевой маршрут к месту назначения, посылая пакеты со специальным полем TTL в заголовке IP и затем фиксируя время «смерти» пакетов. Значение поля TTL, ограничивающее время жизни пакета, передаваемого по сети (Интернету), уменьшается на каждом пересылающем устройстве (маршрутизаторе, узле). Если значение TTL достигает нуля до того, как пакет попадет на хост-получатель, узел (маршрутизатор) отбрасывает данный пакет и посылает ICMP-сообщение (Internet Control Message Protocol) TTL exceeded in transit error хосту-источнику, информируя его об истечении времени жизни пакета. Если пакет сформирован надлежащим образом, хост назначения возвращает завершающее сообщение хосту-источнику, если пакет достигает места назначения. Моменты отсылки пакетов и получения ответных пакетов ICMP фиксируются для вычисления задержки между хостом-источником и каждым из узлов. Указанный нами метод DARB похож на процедуру traceroute, которая действует, посылая пакеты с постепенно увеличиваемым значением TTL. Она посылает хосту назначения случайные пакеты со значением TTL, равным 1, и затем монотонно увеличивает это значение после каждого ответного пакета. Предлагаемый метод DARB устанавливает значение TTL согласно алгоритму, представленному на листинге. Схема выбора представляет собой алгоритм двоичного поиска. Поскольку данная схема не гарантирует успешного исследования всех узлов на маршруте, она стремится исследовать как можно более дальние узлы [1, 2]. Значение far_hop определяет наибольшее значение TTL, и пакет с этим значением TTL будет передан на самый дальний узел при зондировании. Значение near_hop определяет ближайший узел. Процесс зондирования может начаться со шлюза, работающего непосредственно на защищаемом сервере, так как последний может быть в курсе состояния перегрузки в рамках своей автономной системы. Переменная current_hop содержит текущий «прыжок», который совершит пакет. Значение TTL в самом начале устанавливается достаточно большим, чтобы гарантировать возможность доставки пакета хосту назначения, которое равно значению, используемого в операции ping. Если было возвращено сообщение ICMP echo reply, это означает, что пунктом назначения является активный («живой») хост. При SYN-flood-атаках злоумышленник, чтобы гарантировать эффективность атаки, использует недостижимые IP-адреса. Когда активный хост получает неопознанный SYN+ACK-пакет от сервера, он отбрасывает его или посылает серверу RST-пакет. Таким образом, активный пункт назначения можно рассматривать как безопасного легитимного пользователя. Если, с другой стороны, прямого ответа от хоста назначения нет, задержки будут исследоваться на узлах вдоль маршрута. Зондирующий пакет со значением current_hop в качестве TTL сначала

отыскивает успешный отклик от коррелятивных узлов. При получении такого отклика он исследует дальнейшие узлы. Однако, если получить ответ не удаётся, то есть время ожидания запроса истекает, метод пытается исследовать ближайшие узлы. Процесс исследования продолжается, пока есть возможность исследовать доступные узлы.

Важным допущением в данном методе является то, что на ранних этапах поблизости от защищаемого сервера нет серьёзных сетевых перегрузок. Как показывают исследования, большинство DDoS атак длятся от 3 до 20 минут. Даже если присутствует вспышка флуд-трафика в начале атаки, большинство атакуемых серверов всё же могут продержаться в течение нескольких минут, ибо серьёзная сетевая нагрузка возникает постепенно, а не внезапно.

```
Probe(Input: A Half-Open Connection; Output: Delay Value)
    Сформировать исследующий пакет P
    Set P.TTL = 128 // Такое же значение, как и в операции «ping»
    Send(P)
    if Получено сообщение ICMP echo reply Me then
        // Исследуемый узел является активным хостом и может
    // рассматриваться как легитимный пользователь
        return 0
    end if
    // Если нельзя получить прямой результат пинга, выполняем
  // инициализацию процесса зондирования
    Set far_hop = 32
    Set near_hop = Число «прыжков» до шлюза жертвы
    Set current_hop = (far_hop + near_hop) / 2
    Set delay = ∞
    while far_hop > near_hop do
        Set P.TTL= current_hop
        Send(P)
        if Получено сообщение ICMP TTL exceeded Mt then
            Set delay = Время отклика Mt
            Set near_hop = current_hop + 1
            Set current_hop = (far_hop + near_hop) / 2
        else
            Set far_hop = current_hop – 1
            Set current_hop = (far_hop + near_hop) / 2
        end if
    end while
    if delay = ∞ then
        return –1 // Неудача при исследовании каждого узла на пути
    else
        return delay
    end if
```

Схема выборочного зондирования DARB

Метод DARB может получить результаты зондирования за несколько секунд,

что обеспечивает достаточно времени, чтобы дать сигнал тревоги и принять соответствующие меры до того, как DDoS-атака примет серьёзный оборот.

Построенная таблица значений задержек T_{delay} сохраняет значения задержек сети в соответствии с автономными системами Интернета (Autonomous Systems, ASes, AC). Через каждый промежуток времени T, в таблицу T_{delay} вносятся значения задержек сети при последовательных «трёхкратных рукопожатиях» протокола TCP. Можно также активно зондировать задержки сети в то время, когда находятся в режиме ожидания и отсутствуют последовательные «рукопожатия» для некоторых AC. Эти задержки собираются в T_{delay} и обновляются через каждый период времени T. Сбор значений задержек и оценка перегруженности выполняются в соответствии с автономными системами, поскольку состояние перегруженности не изменяется внутри одной AC. Используя T_{delay}, DARB сначала отыскивает в таблице определённые IP-адреса, перед тем как начнёт зондировать сеть. Если найдутся соответствующие значения задержек в рамках одной AC, процесс зондирования может быть пропущен ради экономии времени.

В качестве T_{delay} используется таблица с фиксированным размером. Таблица с динамически изменяемым размером может быть более гибкой, но при этом также может стать объектом DDoS-атаки. Каждая из 10 000 записей в таблице занимает 32 бита, 24 из которых используются в первых трёх октетах для IP-адреса, который содержит достаточно информации для идентификации AC. Остальные восемь бит записи содержат значение задержки при зондировании. Общий размер таблицы будет составлять около 40 КБайт, что приводит к невысокой стоимости хранения такого количества дополнительной информации на современных компьютерах.

Преимуществом использования активного метода типа DARB является то, что он не испытывает негативного влияния файерволов или шлюзовых фильтров. Какой-либо простой метод должен посылать сообщение ICMP echo request месту назначения как, например, это делает ping. Однако файеволы и шлюзы, установленные на некоторых конечных хостах, могут отфильтровывать такие сообщения, что делает обычный механизм пинга ненадёжным. Вместо него следует использовать механизм DARB.

Значения задержек, полученные с помощью метода DARB, используются для того, чтобы классифицировать полуоткрытые соединения либо как нормальные, либо как ненормальные. Если возвращаемый методом результат имеет большое значение задержки, полуоткрытое соединение с высокой долей вероятности является нормальным. В то же время маленькая задержка указывает, что трафик сети находится в нормальном состоянии и не вызывает полуоткрытых соединений.

Вероятностная оценка

Получим среднюю задержку сетевого трафика, когда сеть работает без сбоев. На промежутке времени t возьмём выборку S некоторых полуоткрытых соединений на сервере. Пусть x_i будет задержкой, возвращаемой методом DARB для i-го полуоткрытого соединения из S. Среднее значение тогда рассчитывается следующим образом:

$$\bar{x} = \frac{\sum_{i \in S} x_i}{|S|}.$$

Рассмотрим функцию $f(x)$ для оценки задержек при зондировании. Пусть $\beta = \bar{x}$ – случайная величина, выражающая значение задержки полуоткрытого соединения, исследуемого с помощью DARB. Определим функцию $f(x)$, выражающую вероятность того, что полуоткрытое соединение принадлежит к числу ненормальных, следующим образом:

.

Данная функция моделирует *экспоненциальное распределение*, поскольку именно оно отражает соотношение между задержками сетевого трафика и шансом существования ненормальных полуоткрытых соединений. Когда исследуемая методом DARB задержка мала для данного полуоткрытого соединения, то есть близка к 0, такое соединение имеет высокую вероятность быть ненормальным.

Подать сигнал тревоги при DDoS-атаке можно на довольно ранней стадии, выполняя проверку полуоткрытых соединений, хранимых сервером. При этом не каждое полуоткрытое соединение, обрабатываемое защищаемым сервером, необходимо исследовать с помощью метода DARB. Лишь подозрительные соединения, находящиеся в обработке дольше предопределённого времени t, которое можно регулировать, должны подвергаться исследованию. Нужно производить периодические «снимки» состояния полуоткрытых соединений через каждый определённый период времени. Обозначим через S_T снимок множества полуоткрытых соединений в момент T, а через S_{T-t} – соответственно снимок момент $T–t$ и сравним их. Если некоторые из соединений присутствуют в обоих снимках, следует выполнить исследование этих соединений, то есть выполнить зондирование элементов пересечения $S_T \cap S_{T-t}$. Если пересечение содержит слишком много полуоткрытых соединений, из него нужно взять на исследование некоторую выборку.

Должен быть также определён некоторый порог T, используемый при сравнении с вероятностью, вычисляемой посредством функции $f(x)$. Если $f(x)>T$, соединение считается подозрительным и классифицируется как ненормальное. В остальных случаях соединение является легитимным и рассматривается как нормальное.

Заключение

Представленная активная схема раннего обнаружения является масштабируемой и надёжной в том смысле, что она получает задержки, используя технику зондирования DARB. Схема раннего обнаружения может быть масштабирована для применения на сетях большого размера. В целом, технику DARB не придётся применять в большинстве случаев, поскольку задержки пакетов на путях следования можно будет получать из таблицы, хранящей историю этих значений. Это становится особенно приемлемым в тех случаях, когда сервер собирает информацию о трафике периодически (ради поддержки управления состоянием перегрузок сетевого трафика) и когда предполагается, что сетевой трафик не изменяется резко в течение короткого периода. В конечном счёте лишь небольшая доля хостов назначения должны будут подвергнуться проверке с помощью DARB.

В данной статье был введён на рассмотрение оригинальный механизм ответных мер против SYN-flood-атак. Подход действует посредством классификации полуоткрытых соединений как нормальных либо ненормальных. Нормальные соединения, те, что вызваны перегрузкой сети, демонстрируют черты, не присущие соединениям ненормальным. Перегрузки сети идентифицируются посредством использования зондирующего метода DARB, который собирает данные о задержках между сервером и клиентами. Если задержка значительно превышает среднее значение, это считается признаком перегруженности, и соответствующее полуоткрытое соединение относится к числу нормальных. В остальных случаях соединение считается ненормальным.

Указанные выше пороговые параметры *t*, *T*, *N* становятся важными для оценки эффективности рассматриваемого метода. При проведении экспериментов данные параметры выбираются на основании опытных данных и могут быть не оптимальными. Оптимизация этих параметров может стать отдельной задачей в будущей работе.

Библиографический список

1. *Bin Xiao*, *Wei Chen*, *Yanxiang He*, *Edwin H.-M. Sha*. An Active Detecting Method Against SYN Flooding Attack. Department of Computing The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.
2. *Bin Xiao*, *Wei Chen*, *Yanxiang He*. An autonomous defense against SYN flooding attacks: Detect and throttle attacks at the victim side independently. Department of Computing, Hong Kong Polytechnic University, Hong Kong.
3. *Changhua Sun*, *Chengchen Hu*, *Yachao Zhou*, *Xin Xiao and Bin Liu*. A More Accurate Scheme to Detect SYN Flood Attacks. Department of Computer Science and Technology, Tsinghua University, Beijing, China
4. *Haining Wang*, *Danlu Zhang*, *Kang G. Shin*. Detecting SYN Flooding Attacks. EECS Department, The University of Michigan.
5. *Sun Qibo, Wang Shangguang, Yan Danfeng,Yang Fangchun*. An Early Stage Detecting Method against SYN Flooding Attacks. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China