Модель с замерзанием: подавление интернет-эпидемий

Новости.

Вирусы последнее время совсем совесть потеряли. Эпидемия за эпидемией. Главная причина проистекает из "свободной" идеологии передачи данных. И поправить не слишком просто, косметическая правка приведет ко вполне косметическим результатам.

Относительно качества ПО, производимого монстрами рынка есть хорошая иллюстрация на эмигрантском форуме privet.com. Да и вообще, для относительно сложных программ есть принцип Д.Кнута: доказательство правильности программы - структура более сложная, поэтому более подверженная ошибкам. ("Beware of bugs in the above code; I have only proved it correct, not tried it"). Так что с этой стороны надежды на защиту нет никакой.

И следует смириться с тем, что черви неистребимы, как грипп. Хуже того - после появления в сети достаточно распространенной денежной системы они станут ещё более неистребимы и опасны, чем сейчас спам. Но об этом ниже - в статье Виктора С. Грищенко.

Долго искал место на сайте, куда поместить следующую новость. Но вот она - 30-го августа 2003 года с 12:00 до 9:00 31-го августа 2003 под Подольском состоится мега акция Moscow HomeNET Point. Сайт тусовки - тут.

Небольшие обновлениям разделов.

- Выложен третий параграф третьей главы второй части **книги о домашних сетях**.
- По разделу **оборудование** - осеннее снижение цен на линейку модемных пулов USR TotalControl MP/16.
- Небольшие дополнения в коллекции . Высылка фотографий на nag@nag.ru активно приветствуется.

Модель с замерзанием: подавление интернет-эпидемий.

Abstract:

Последние годы ознаменовались расцветом интернет-червей [4] [3]. После того, как было теоретически доказано, что поражение миллиона компьютеров за 15 минут вполне возможно, возникли и некоторые (теоретические, опять же) наработки по подавлению таких сверхбыстрых эпидемий на уровне сетевой инфраструктуры [5]. В частности, данная статья написана в ответ на идею M.Williamson [1] об ограничении скорости соединения с новыми IP-адресами.

В статье вскрываются некоторые недостатки подхода Вильямсона, сильно ограничивающие его область применения. Также предлагается более пластичный подход (модель с замерзанием), сильно ограничивающий возможности червя (и не только червя) по сканированию Сети.

Модель замерзания подразумевает подавление источников, имеющих большое количество "неотвеченных" IP-потоков. Сугубый ad-hoc к технологии IP, конечно.

Об ограничении сканирования.

Главная идея [1] - это замедление сканирования Интернет через ограничение количества новых IP-адресов, к которым может обратиться владелец данного IP-адреса за определенный промежуток времени. Ориентировочно этот порог предполагается в районе 1-2Hz. В случае повсеместного внедрения ограничение сделает сверхбыстрые эпидемии (10-15 минут на тотальное заражение) технически (почти) невозможными[1].

Помимо прочего в этой работе утверждается, что в случае применения на достаточно высоком уровне ограничение Вильямсона может серьезно повредить базовым интернет-сервисам, а именно DNS.

Чтобы рассуждать практически, автор взял 24-часовой лог трафика DNS-сервера средней руки (не корневого, нет). При рассмотрении абсолютного числа различных IP-адресов, *к которым обращался* сервер, а именно 1834 штуки, ограничение в 1-2Hz кажется применимым (см. Рис. 1):

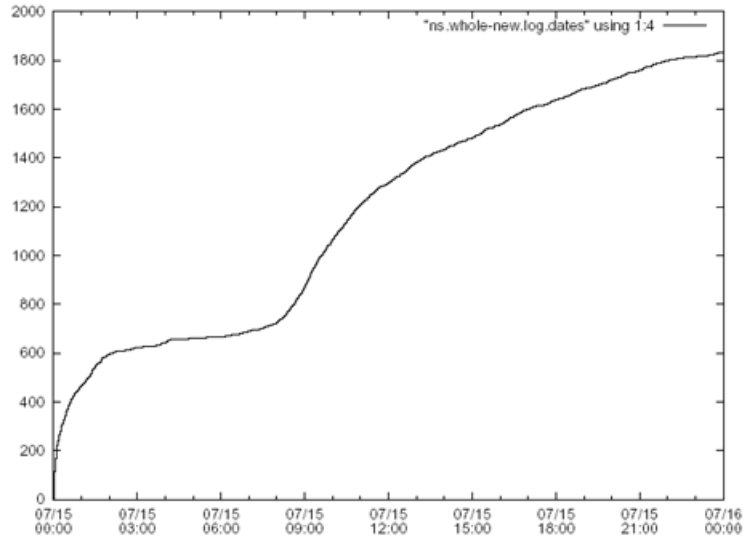


Figure 1: Кумулятивное число новых dstip, DNS сервер, 24-х часовой период

Рис. 1: Кумулятивное число новых dstip, DNS сервер, 24-х часовой период

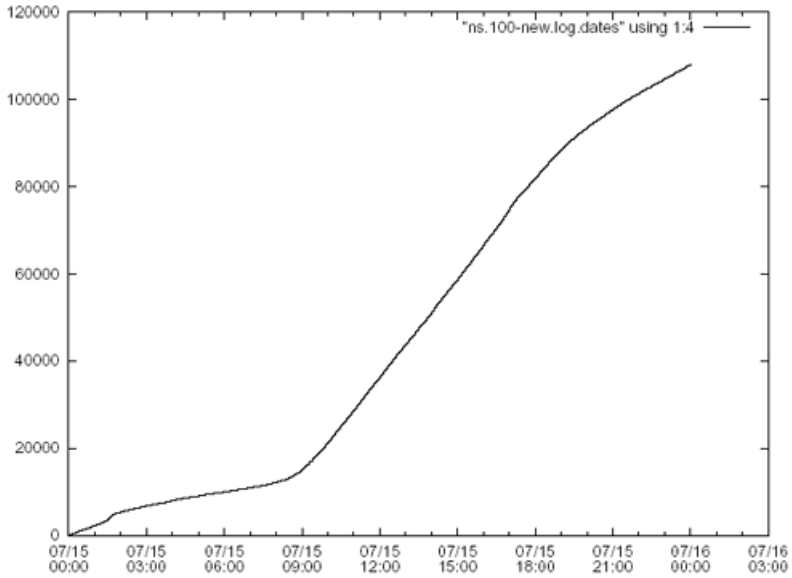


Рис. 2: Кумулятивное число новых dstip, история ограничена ста последними адресами, тот же лог.

Но настоящая проблема даже не в этом, а в том, что 1.25Hz - это средняя температура по больнице, поскольку DNS-трафик идет всплесками. Например, для разрешения www.hpl.hp.com[2] DNS сервер должен сделать три последовательных запроса к трем другим серверам[2]. На загруженном сервере под 1-герцовым ограничением Вильямсона это может привести к 3-х секундной задержке. Очевидным образом это может привести к ещё большей задержке при нескольких одновременных запросах, более длинных запросах, и прочая, и прочая.

Таким образом, DNS-сервер не имеющий совершенно никаких причин, чтобы придерживаться даже 1,25Hz в среднем в сутки, в малом масштабе времени может продемонстрировать куда большие частоты запросов новым адресатам. Уж не говоря про то, что все эти показатели зависят от таких переменных, как количество пользователей, время дня и день недели, размещенное у клиентов ПО и, опять же, прочая, прочая...

Таким образом, ограничение Вильямсона реально применимо где-то на уровне корпоративных фэйрволов, не выше, поскольку конструкция ограничиваемой сети должна быть примерно известна (как минимум, знать IP-адреса DNS-серверов).

О фильтровании содержимого.

В случае, если автор червя располагает данными о доселе неизвестной уязвимости и умеет программировать (изучил опыт предыдущих червей), на настоящий момент ничто не мешает ему

отформатировать пол-интернета за 10-15 минут (в крайнем случае - пол-дня). Д.Мур и другие [5] приходят к заключению, что наиболее перспективным путем борьбы с червями является content filtering, размещенная на транзитной инфраструктуре система удаления пакетов, выпущенных червем.

Для того, чтобы эта система приобрела терапевтическую, а не патологоанатомическую ценность, предполагается создание системы автоматического обнаружения в сети аномалий и выявления сигнатур червя[3] - ведь через 10 минут лечить будет поздно.

Исключительно чтобы обозначить сложности данного подхода заметим:

- Эксплойты, действующие по принципу переполнения буфера, как правило, имеют в себе тексты вроде "AAAAA...AAAA" - чтобы попасть полезным текстом в необходимую точку памяти, ненужную память сначала заполняют чем попало. Ничего не стоит вместо классического A использовать случайные символы - тогда задача машинного распознавания эксплойта (выявления сигнатуры) по сложности поднимается до уровня создания искусственного интеллекта. Как видим, простейшая хитрость усложнила проблему на несколько порядков.
- Вообще-то не обязательно весь эксплойт едет в одном пакете - можно специально разложить код в несколько пакетов, случайными порциями. Жертва сама их соберет необходимым образом. А вот задача сборки пакетов на транзитной инфраструктуре в общем случае невыполнима (а в частных случаях требует огромных затрат).

Итак, поразмыслив пол-часа, мы создали глобальному перлюстратору проблем в лучшем случае на много миллиардов долларов.

Модель охлаждения.

Правила

- каждый IP адрес (скажем, ip_a) имеет две истории - для исходящих и входящих потоков.
- получая поток с нового IP адреса, ip_a получет +1 градус тепла
- отправляя поток на новый адрес, ip_a теряет 1 градус тепла
- потоки с IP-адресов холоднее -F градусов "замораживаются"' (задерживаются или удаляются)
- "холодные" IP со временем теплеют, например +1 градус с частотой Вильямсона (допустим, ν_w = 1 Hz)
- слишком горячие IP могут как-нибудь остывать.

Предположения Предполагается, что IP-адреса отправителя не могут фальсифицироваться случайным образом (IP spoofing), что, на данный момент, в большинстве случаев неправда.

Так же предполагается, что маршруты пакетов более-менее симметричны, то есть если пакет от ip_a к ip_b проходит через наше счетное устройство C то и ответный пакет $ip_b \to ip_a$ через него пройдет.

DNS в среде с нагреванием/охлаждением

На первый взгляд кажется, что DNS-сервер должен иметь примерно нулевую температуру: он получает запрос клиента (+1), запрашивает корневой DNS-сервер (-1), получает ответ (+1), запрашивает и получает ответ ещё от n DNS-серверов (-1+1 $\times$ n) и отвечает клиенту (-1).

Чтобы проверить эту гипотезу я набросал программу, которая по логам сетевой активности вычисляет температуру IP адреса (заметим: ни замерзание-остановка, ни потеря излишнего тепла не были реализованы за невозможностью и ненадобностью соответственно). Уже упоминавшийся DNS-сервер накопил +3148 градусов тепла за все тот же 24-х часовой период.

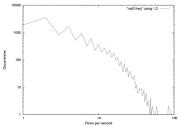
Причина этого явления кроется, видимо, в деталях протокола, который я поленился досконально изучить.

NAT в среде H/O

В нашей среде NAT - это очень интенсивный клиент, делающий запросы множества пользователей от своего "имени" (IP-адреса). Вышеупомянутой программе был дан 24-х часовой лог NAT'а, всего 74373 записей. Весьма интересно, что NAT продемонстрировал почти идеальный баланс: +2 градуса. (на Рис. 3) можно посмотреть количества потоков в секунду - что характерно, чаще они четные).

Упомянем случай бота поисковой машины (Google, например). Шарящий по сети бот может рассматриваться, как клиент очень большой интенсивности. От интенсивности зависит краткосрочная амплитуда, в то время как среднесрочный баланс примерно равен нулю. При условии, что частота попадания на "мертвые" сервера не превосходит частоты Вильямсона, бот замерзнуть не должен.

Рис. 3:



NAT, число потоков в секунду. Потоки-в-секунду - по горизонтали, число - по вертикали

Сканирующий червь в H/O среде

Достаточно очевидным образом сканирующий хост очень быстро зарабатывает рекордно низкую температуру. Был взят лог из 2028180 записей (24-х часовой период) трафика с IP-адреса, осуществляющего сканирование на порту netbios-ns 137. Температура в результате достигла - 1,628,474 градусов (см. Рис. 4) [5].

"Ядерная зима" и её возможность.

Вопрос: можно ли использовать замерзание для косвенной DoS атаки? Например, некоторым образом вывести из строя часть DNS-серверов. Произойдет цепная реакция: другие сервера, обращаясь к "мертвым", будут в свою очередь охлаждаться и замерзать, пока не замерзнет решительно все.

Стоит отметить, что все протоколы, действующие по принципу "запрос клиента - ответ сервера", а это FTP, HTTP, POP и другие, неуязвимы для замерзания. В дальнейшем будет рассматриваться DNS, как протокол с достаточно нетривиальным поведением.

Теоретически, достаточно знать об одном неотвечающем DNS-сервере - в том случае, если атакуемый DNS-сервер достаточно загружен (что-что, а загрузку серверу обеспечить просто). "Достаточно" сервер будет загружен в том случае, если наше обращение вымывается из истории соединений достаточно быстро. Т.е. зная длину истории (l_h) на считающем устройстве (C), можно после запроса о "мертвом" домене соответствующее количество раз $(l_h - 1)$ запрашивать информацию о других доменах. Если такой цикл будет происходить с частотой, большей ν_w, то DNS-сервер начнет замерзать. При длине истории 100 и ν_w=1Hz сценарий вполне реальный. Другой вопрос, что как только "обстрел" прекратится, сервер возобновит работу (оттает).

Вне зависимости от H/O модели мы видим, что злоумышленник может заставить DNS-сервер демонстрировать довольно странное поведение. Чтобы ограничить такие возможности, следует во-первых разделить клиентов DNS-сервера на своих и остальных (свои - те что находятся в "своих" сетях, для них данный сервер - сервер DNS по умолчанию). Возможности "остальных" клиентов следует ограничить запросами по зонам, поддерживаемым данным сервером (запрос-ответ). Таким образом, возможности манипулирования DNS-сервером существенно сократятся.

Безопасность инфраструктуры.

Потеря информации

Общая проблема всех методов подавления червей на транзитной инфраструктуре - то, что IP - это пакетный протокол общения точка-точка[6]. Поэтому, вполне естественным образом, чем дальше от точек происхождения и назначения, тем меньше полезной информации мы можем извлечь из проходящего трафика. Уже в машине происхождения кусок данных разбивается на пакеты, которые

затем добираются до точки назначения кто как сможет. Т.е. если нитка протянута между двумя гвоздиками, то наиболее предсказуемо её географическое положение - вблизи гвоздиков, наименее предсказуемо - в середине.

Отсюда вывод: наиболее выгодное место для борьбы с червем - сами машины-жертвы (тут в большинстве случаев достаточно своевременно пропатчить ПО), затем фэйрволы (тут хватает простеньких правил пускать/не пускать), затем edge routers (они гарантированно имеют представление о трафике и к клиенту и от клиента, т.е. условие симметрии ещё соблюдается, также в большинстве случаев есть гарантия, что данный путь к клиенту - единственный, обхода нет).

Наименее выгодно с данной точки зрения бороться с червем на транзитной инфраструктуре, поскольку с огромной скоростью мимо пролетают обрывки каких-то потоков - анализ затруднен; потоки в противоположном направлении летят где-то в другом месте - нет симметрии; данный путь практически гарантированно не является единственным путем, ведущим к любому из клиентов - т.е. чтобы обезопасить клиента, нужно перекрыть все пути, частичные меры не дают значимого эффекта [5].

С другой стороны, представление об общей обстановке в сети (распознавание эпидемии) легче всего получить, наблюдая за транзитной инфраструктурой и сложнее всего - в конечной точке (туда придёт один поток, и его вам хватит).

Таким образом, если даже Д.Муру сотоварищи улыбнется фортуна в части создания искусственного интеллекта, наиболее выгодное место для размещения фильтрующих устройств - периферия сети, но никак не ядро. Так и масштабируется лучше. С другой стороны, как координировать и защищать все это распределенное хозяйство, не очень-то ясно. В сложном протоколе много проколов, а протокол обещает быть сложным, особенно учитывая отсутствие человеческого вмешательства.

Пропускная способность

Модели с ограничением количества потоков от одного клиента решают одну проблему, растущую год от года - сети все быстрей, а средний эксплойт как не превышал нескольких килобайт последние 15 лет, так и до сих пор не превышает. Соответственно, возможности червей все выше и выше. И если червь Морриса полз от соседа к соседу часами и днями, то гипотетический червь Уорхола [7], бомбя правого и виноватого, справляется за 15 минут с количеством машин, в 100 раз большим. Sapphire так и за 10 минут управился без особых хитростей.

Упрощенные варианты модели H/O.

Легко видеть, что реализация модели H/O *verbatim*, т.е. согласно правил отнюдь не обязательна. Хотя алгоритм прост и легко реализуется в некоем счетном устройстве, работающем параллельно с настоящим обработчиком пакетов, тем не менее для дешевых устройств можно пойти на упрощения. В случае ATM, наблюдаемой автором, можно, например, отказаться от истории и считать входящие/исходящие потоки на каждый ATM интерфейс.

Выводы.

Такие маленькие опции к сетевому оборудованию, как упрощения модели нагревания-охлаждения, в случае локального внедрения "спасут сеть" от работающего в ней сканирующего червя, в случае же глобального распространения сделают сверхбыстрые эпидемии маловероятными - люди будут иметь шанс вмешательства.

Городить же искусственный интеллект постольку, поскольку пользователи редко пользуются естественным - занятие сомнительное, к тому же с отрицательной обратной связью. Первый же червь, перехитривший глобальную систему фильтрования, может получить море беспечных жертв, не заштопавших дыры годичной давности.

(c) Виктор С. Грищенко

Bibliography

- 1Matthew M. Williamson. Throttling Viruses: Restricting propagation to defeat malicious mobile code.
- 2see www.dnsstuff.com
- 3Stuart Staniford, Vern Paxson, Nicholas Weaver. How to 0wn the Internet in Your Spare Time
- 4 David Moore and others. The Spread of the Sapphire/Slammer Worm
- 5David Moore, Colleen Shannon, Geoffrey Voelker and Stefan Savage. internet Quarantine: Requirements for Containing Self-Propagating Code
- 6Stuart Staniford, Gary Grim, Roelof Jonkman. Flash Worms: Thirty Seconds to Infect the Internet, at

http://www.silicondefense.com/flash/

- 7Nicholas C Weaver. Warhol Worms: The Potential for Very Fast Internet Plagues, http://www.cs.berkeley.edu/~nweaver/warhol.html
- 8Киножурнал "Ералаш"

Footnotes

... невозможными[1]
: При ряде условий, конечно - например, невозможности IP address spoofing - подделки адреса отправителя

... www.hpl.hp.com[2]
: там работает Вильямсон

... червя[3]
: Не говорите об этом людям из Касперского, а то вместо паспортизации интернета они начнут продвигать глобальную перлюстрацию! Денег-то сколько...

... антивирусы![4]
: Это как мальчик с двумя абонементами - если потеряет оба, то покажет проездной. [8]

...worm) [5]
: "Когда я познакомился с этим человеком, он уже достиг дна. С тех пор он принялся копать."

... точка-точка[6]
: забудем про броадкаст

Анонс

- Тезисы о широкополосном подключении конечных пользователей - пришлось их перенести в следующий выпуск;
- Ссылки на интересные материалы в Сети. Обязательно пишите;
- Фотогалерея кроссов;
- Еще одно испытание HomePNA 2.0 (и его пришлось перенести);
- Публикация схем и примеров грозозащит задержалась, но это совсем не значит что присылать свои варианты уже поздно.