аппаратные антивирусы – серые кардиналы магистральных каналов

крис касперски, ака мыщъх, a.k.a. nezumi, no-email

активность червей, едва не заваливших Интернет, за последние несколько лет резко сошла на нет, вольному хакерству пришел конец — провайдеры развернули распределенную систему обнаружения вторжения с датчиками, сенсорами и адаптивными сигнатурными базами. теперь все больше и больше атак захлебываются не успев даже начаться. условия игры изменились, да и не игра это уже а битва за выживание! что ж! мы принимаем вызов!

введение

Рядовые пользовали ни разу не озабочены своей безопасностью, они не скачивают заплаток, не обновляют антивирусов, а на все предупреждения защиты не задумываясь отвечают "yes", становясь рассадниками заразы, генерирующей огромное количество паразитного трафика, изрядно напрягающего крупных провайдеров, вынужденных расширять каналы, пропускные способности которых тут же съедаются спаммерами и бонтетами.

Базовые программное обеспечение катастрофически ненадежно и дыры обнаруживаются буквально каждый день, причем каждая вторая-третья дыра критическая. Антивирусы, построенные по древним технологиям, восходящим к почившему на лозе некогда знаменитому AIDSTEST'у, совершенно неэффективны против троянских коней и прочей малвари. Эвристические анализаторы обнаруживают лишь "пионерские" поделки, написанные на Дельфинах и запротекченные UPX'ом, а для всего остального сначала необходимо получить образец малвари, выделить сигнатуру, занести ее в базу, выложить на сервер и дождаться пока пользователи не удосужатся обновится. А с какой им радости обновляться?! Они компьютер совсем не для того покупали, чтобы над ним корячиться!

Примем как факт: ни домашние, ни корпоративные пользователи не обновляются и обновляться не будут. Никто не хочет вкладывать деньги в системы защиты и держать целый штат специалистов по безопасности. Почему же тогда глобальные эпидемии больше не возникают?! Что мешает червям свободно распространяться от машины к машине?!

Оказывается, что помимо KAV'a, DrWeb'a, NOD'a (техника обхода которых постоянно обсуждается на хакерских форумах) существуют еще и аппаратные антивирусы, встроенные в железки от CISCO, DLINK'a и прочих производителей. Особенности реализации сами по себе не делают аппаратный антивирус крутым и могущественным. В них нет ничего загадочного, таинственного или сверхестественного — эвристика, сигнатурный поиск и прочие классические техники каменного века. Почему же тогда аппаратные антивирусы работают, а программные тормозят со страшной силой (попробовали бы они так тормозить на магистральных каналах!), но ничего не ловят?!

А все потому, что сигнатуры аппаратных антивирусов описывают не конкретный экземпляр малвари, а _сценарий_ атаки, эксплуатирующей ту или иную уязвимость, причем сигнатуры берутся у независимых поставщиков (крупнейшим из которых является Endeavor Security Inc), а не собираются каждым производителем индивидуально. Распределенные сенсорные сети (они же "гриды" от английского "sensor grind") детектят всякую аномальную активность, фиксируя хакерские атаки, эксплуатирующие еще неизвестные дыры, которые тут же описываются разработчиками антивируса на языке регулярных выражений и отправляются в базу, на что требуется время, причем, довольно значительное, поскольку, дизассемблирование малвари/шеллкода приходится выполнять вручную, а специалистов по реверсингу традиционного не хватает.

Почему же тогда при всех своих недостатках аппаратные антивирусы оказались настолько эффективны, что прищемили всех червей, хакеров и установили в Интернете круглосуточный комендантский час?! Причина в том, что большинство атакующих даже не подозревают о существовании "серых кардиналов" и потому никак от них не защищаются, да и как защищаться, если принципы работы аппаратных антивирусов неизвестны, а сам антивирус доступен лишь сотрудникам крупных IT-компаний, среди которых хакеры встречаются намного реже, чем крокодилы в Сахаре, да и тем, что встречаются никто не даст вскрывать коробку ценой в несколько тысяч долларов.

Аппаратные антивирусы окутаны плотным мраком секретности. И даже доступ к сигнатурным базам лицензируется на весьма жестким условиях. Подписка о неразглашении, работа только с юридическими лицами — вот и все хакерство! Информацию приходится собирать буквально по крупицам. Мыщъх, имеющий доступ к коробке, и сотрудничающий с лидерами отрасли, очень хорошо знает все зубчатые шестеренки и рычаги управления, приводящие в движение грандиозный механизм быстрого реагирования, стоящий на страже Интернета. Давно хотел написать на эту тему статью, но… все упирается в эти пресловутые подписки о неразглашении, которые приходится обходить весьма хитрым путем, апеллируя к отрытым проектам, которых намного больше одного. Ситуация осложняя тем, что служебное положение с некоторых пор обязывает мыщъх'а предоставлять своим боссом все материалы, прямо или косвенно связанные с вирусами, на предмет проверки — а не сболтнул ли пещерный грызун чего лишнего и только после всех правок, согласований и жестокой цензуры, сопровождаемой непрерывной нервотрепкой, изрядно покоцанный материал пускается в печать. Пикантность ситуации заключается в том, что боссы говорят на английском, а русского никак не понимают и потому приходится делать пословный перевод всех русских статей, что ужасно напрягает, высаживая на измену, снижающую мотивацию публикаций до абсолютного нуля по Фаренгейту. Ну у них на западе Фаренгейт. Ничего, со временем и не к таким штукам привыкаешь.

Удачи, конечно, тоже бывают. Эта статья изначально представляла собой внутрифирменный отчет, написанный по заказу определенной фирмы, аппаратный антивирус которой мыщъх обстоятельно анализировал, оценивая его достоинства и соображая чтобы такого хорошего свиснуть у конкурентов, чтобы покрыть недостатки. Урезанную версию отчета предполагалась опубликовать на сайте фирмы, чему мыщъх был несказанно рад, но в последний момент фирма пошла в отказ ("да нас за такое засудят, да это же против антивирусной этики"), но согласилась на публикацию в "Хакере", поскольку русский сегмент рынка ею совсем не окучен (зачем медведям антивирусы?!), так что мыщъх получил зеленый свет и вот…

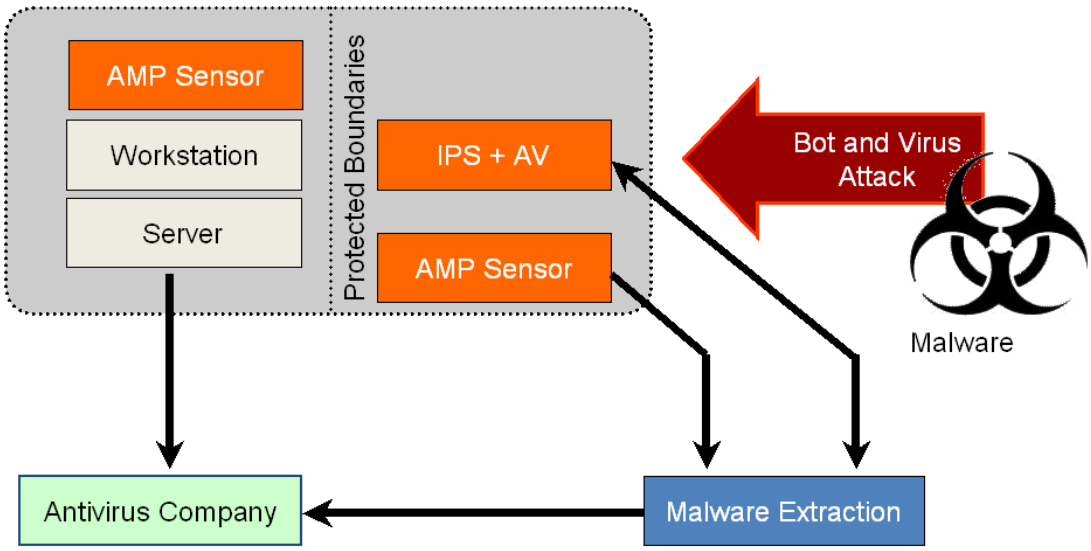

Рисунок 1 блок-схема распределенной сети раннего предупреждения и предотвращения атак

внутри коробки

Аппаратный антивирус представляет собой гибрид системы обнаружения вторжений с пакетным сканером, работающим на определенном сетевом уровне и опирающимся на более или менее развитый сигнатурный "движок". Простейшие антивирусы, встраиваемые в дешевое оборудование, работают либо на Ethernet, либо на IP уровне.

Потоковый анализ TCP-пакетов — это уже совсем другой ценовой класс, поскольку, парсинг TCP-пакетов весьма ресурсоемкое дело, пожирающее оперативную память со сверхсветовой скоростью, особенно если атакующий умышленно посылает IP-пакеты в обратном порядке, то есть пакет, находящийся в конце TCP-сегмента, идет первым, и, чтобы применить сигнатуру, антивирус должен созвать полную ассамблею, то есть собрать весь TCP-

сегмент, откладывая пакеты в память, которая, между прочим, не резиновая, а злоумышленник (вот гад какой!) шлет пакеты с предельно низкой скоростью, такой, чтобы его только не отрубило по тайм-ауту. Да и не только он один. Лишь в исключительных случаях пакеты следуют в том порядке в котором они отправлялись. На Интернет-перекрестах они многократно перемешиваются с другими, переупорядочиваются, кое-кто теряется по дороге… А куда антивирусу деваться?! Приходится складировать пакеты в память и ждать прихода всего сегмента целиком или же расширять базу сигнатур, доводя ее до состоянии при котором атака однозначно идентифицируется по любому фрагменту TCP-пакета, что опять-таки требует памяти — сигнатуры нужно где-то хранить.

Антивирусы первых поколений использовали фиксированные последовательности байт, иногда "привязанные" к определенной точке — смещению от начала пакета или другой структуры. Затем появились подстановочные символы "*" и "?" (известные еще со времен MS-DOS), а за ними пришли и регулярные выражения типа REGEX/PRCE (впрочем, продвинутые антивирусы поддерживают сразу оба стандарта). Разбор регулярных выражений требует значительных вычислительных мощностей, к тому же, регулярные выражения в общем случае невозможно откомпилировать (во всяком случае эффективно). Хуже того, они обладают существенными ограничениями, не позволяющими в частности, распознавать полиморфный код, для которого в обычных программных антивирусах пишутся специальные модули, использующие самые невероятные алгоритмы — от простого подсчета энтропии до натягивая ветвлений на графы с переименованием регистров и ячеек памяти в псевдопеременные.

Крутые полиморфные вирусы распознаются с большим трудом и огромным количеством ложных срабатываний (и это с учетом специально заточенных под них модулей детекции!). Регулярные выражения здесь вообще отдыхают. Все что могут разработчики — это создать сигнатурную базу, перечисляющую _все_ возможные варианты следования байт в мутированном вирусе. Несколько тысяч регулярных выражений на _один_ полиморфный вирус — явление вполне нормальное, хотя очень хреново работающее. Вручную набить (и отладить!) столько регулярных выражений — нереально, а потому процесс их создания полностью автоматизирован. Отсюда и качество детеции (вернее, его отсутствие). В среднем таким путем распознается от 75% до 95% штаммов, когда KAV и Dr.Web ловят до 99,6%. Уровень в 98% — для них уже катастрофа и явный лаг детектора, который устраняется как только поднимается крик "почему ваш антивирус ни хвоста не ловит?!".

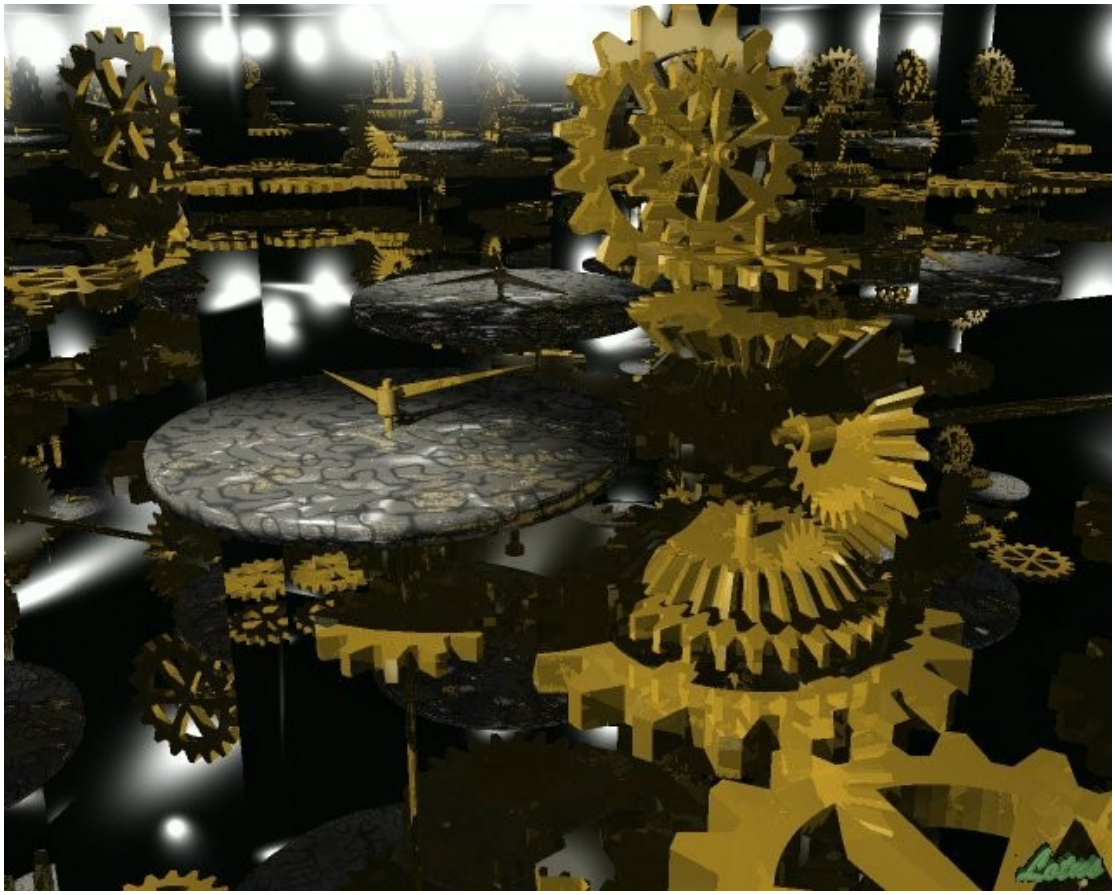

Рисунок 2 шестеренки, приводящие антивирус в движение

Ограничения регулярных выражений приходится компенсировать дополнительными средствами. В частностями — пороговыми датчиками (threshold sensor/detector). Что это значит? Допустим, мы имеем сигнатуру, описывающую последовательность NOP'ов, за которой идет JMP ESP (классический сценарий передачи управления на shell-код при стековом переполнении). Может ли такая последовательность встретиться в "честном" потоке данных? А почему бы и нет?! NOP'ы вообще очень распространенное явление, а JMP ESP представляет собой двухбайтовую команду и потому вероятность ложных позитивных срабатываний весьма велика и чтобы Интернет не погрузился в пучину репрессий, у антивируса имеется определенный порог ниже которого атака не фиксируется. И хотя грамотно написанному shell-коду для захвата управления достаточно послать всего один пакет (в идеале), аппаратные антивирусы целенаправленные атаками не интересуются и просыпаются лишь когда в сети появляется червь или злобный хакер, забрасывающий shell-код на все узлы без разбора.

Выходит, что аппаратные антивирусы годятся лишь для предотвращения глобальных эпидемий?! Не совсем. Ряд атак однозначно описывается языком регулярных выражений. В частности, если мы имеем ошибку переполнения в графической библиотеке IE, неправильно обрабатывающего теги gif-файлов, на прикладом уровне атака однозначно идентифицируется парсингом gif-заголовков. Но до прикладного уровня еще дотянуться надо! Хорошо, если gif лежит на WEB-сервере "как он есть", а если его послали мылом в одной из многочисленных кодировок, которую только поддерживают почтовые клиенты, да еще в упакованном виде. Ни один антивирус, работающий на сетевом уровне, ее не распарсит, правда, почтовые антивирусы справляются с такой ситуацией без труда. Некоторые производители в борьбе за рейтинги пытаются парсить прикладные протоколы с сетевого уровня, формально поддерживая сигнатуры, описывающий заданный тип атак, однако, их очень легко обломать, если, конечно, знать об их существовании, в идеале имея доступ к базе сигнатур,

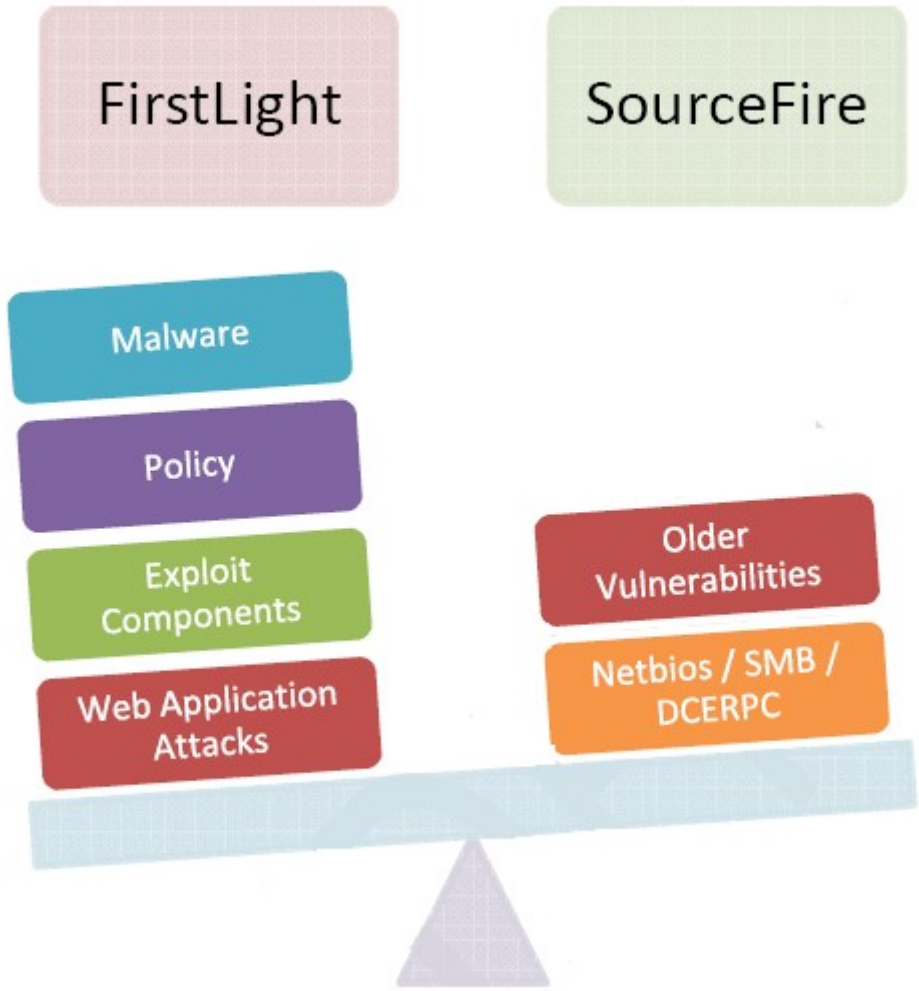

Рисунок 3 два крупнейших независимых поставщика сигнатур — FirstLight и SourceFire

>>> врезка стандарты регулярных выражений

REGEX в широком смысле этого слова означает регулярные выражения (Regular Expressions) вообще, но в определенном контексте ассоциируется с библиотекой "REGEX", написанной Генри Спенсером (Henry Spencer), синтаксис которой перекочевал во многие скриптовые языки (Perl, Tcl и т.д.), подробнее о которых можно прочитать на Вике: http://en.wikipedia.org/wiki/Regular_expression;

Двигаясь ходом эволюционного развития, регулярные выражения образовали библиотеку **PRCE**, что расшифровывается как ***P**erl **C**ompatible **R**egular **E**xpressions* — Perl-совместимые регулярные выражения, постепенно ставшие стандартом де-факто, практический пример применения которого приведен ниже: /\w+?\s\w+?\(([\w\s=]+,*|[\w\s=]+|(?R))*\);/

Естественно, помимо REGEX/PRCE существуют и другие библиотеки. Ряд антивирусов использует свои собственные ни с чем не совместимые "стандарты", и потому поставщикам сигнатур приходится не по детски извращаться, чтобы удовлетворить изыски разработчиков защитных механизмов.

Рисунок 4 интенсивный трафик магистральных каналов

как это ломают

Рассмотрим основные, можно даже сказать фундаментальные, механизмы обхода аппаратных антивирусов, широко обсуждаемые в закрытых кругах и уже вырвавшиеся на свободу в виде весьма агрессивных вторжений в чужие системы. Держать информацию под колпаком больше не имеет смысла. Если кто от этого и выиграет — так только вандалы, нападающие на ничего не ведающих пользователей. Пользователи должны быть предупреждены! Короче…

Покажем как обхитрить threshold sensor на примере "отравления" DNS сервера поддельными пакетами. Microsoft выпускает уже четвертую по счету заплатку, затрудняющую атаку, но отнюдь не делающую ее невозможной. До недавнего времени номер порта-отправителя UDP-пакета с DNS-запросом и 16-битный номер последовательности Transaction ID (TXID) были легко предсказуемыми и хакеры без труда генерировали подложные DNS-ответы, воспринимаемые системой как правильные, в результате чего жертву удавалось заманить на совершенно посторонний узел, которому она сообщала конфиденциальные данные (от номера кредитки, до пароля на почтовый ящик).

В начале июля 2008 года, Microsoft выпустила MS08-037 патч, радикально меняющий стратегию назначения локальных портов. Если раньше номер порта каждого отправляемого пакета тупо увеличивался на единицу, то теперь используется… нет, даже не rand(), а довольно серьезная криптографическая функция, генерирующая ну очень случайные 16 бит, настолько случайные и непредсказуемые насколько это вообще возможно, хотя вообще-то на трезвую голову хватило бы и rand(), но Microsoft не ищет простых путей. Предыдущий патч (MS08-020) исправлял вполне предсказуемый TXID, основанный на простой временной функции, которая, однако, была предсказуема только в лабораторных условиях, и до промышленных хакерских стандартов явно не дотягивала, но специалисты по безопасности написали кучу умняковых статей с серьезными математическими выкладками (надо же как-то отрабатывать гранты, если кроме теоретических знаний у мальчиков из колледжа ни хвоста нет). Обиженная до глубины хвоста Microsoft разозлилась настолько, что всобачила криптостойкую функцию CryptGenRandom() в dnsapi.dll, живописно описав все ее преимущества на своем же блоге о blogs.technet.com/swi/archive/2008/04/09/ms08-020-how-predictable-is-the-dns-transaction-id.aspx,

дипломатично "забыв" упомянуть, что если вызов CryptGenRandom() провалится, то вызывается обычный rand().

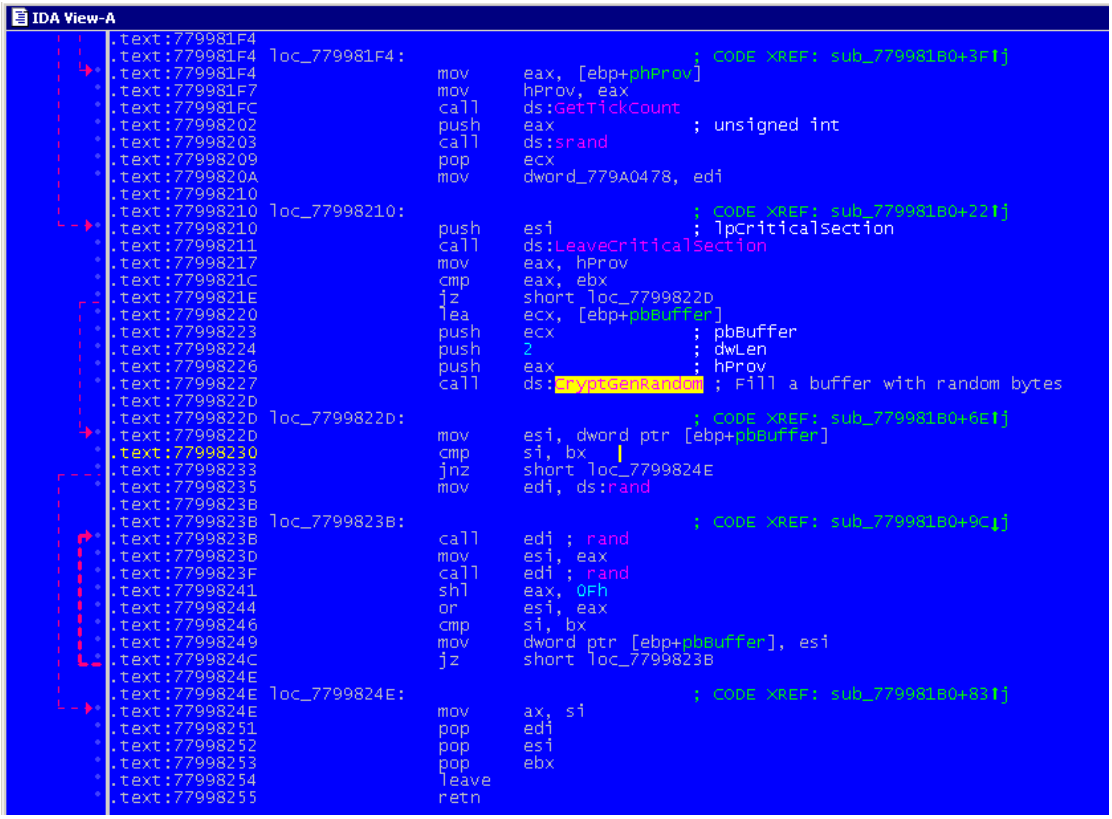

Рисунок 5 криптографически стойкий TXID

Короче, теперь сплошной…. (покусано цензурой). Оба поля совершенно случайны и абсолютно непредсказуемы. Во всяком случае на первый взгляд. А если копнуть вглубь? Причем даже не дизассемблером и не отладчиком, а головой? По умолчанию пропатченный TCPIP.SYS драйвер использует ограниченный набор портов (49152-65535), что в пересчеты на хакерскую валюту дает нам 20 бит (точнее, даже 19 с хвостиком, но хвостик мы округляем в большую строну), плюс 16 битное поле TXID. Итого мы имеем 20 бит. Следовательно, для успешной атаки, хакеру необходимо в среднем послать 2^20/2 = 524.288 подложных пакетов. Да это же настоящий шторм, который распознает любая IDS! У нее все датчики зашкалят!

Стоп! А куда нам спешить?! Сядем, покурим, а пока курим, будем посылать пакеты. По одному в минуту. Ведь как устроена IDS? Она садится на канал, ловя все пролетающие пакеты, и, если с одной стороны появляется большое количество DNS ответов, которые не запрашивались и которые имеют совершенно левый номер порта с не менее левым TXID, выставляется флаг атаки. Естественно, поскольку левые пакеты сыплются и без всякой хакерской помощи (у какого провайдера маршрутизатор идеально настроен?), то IDS для предотвращения ложных позитивных срабатываний реагирует не на количество левых пакетов вообще, а именно на их интенсивность. Один пакет в минуту это не шторм. Это вполне нормальное явление, даже в мелкой подсети. Конечно, такими темпами атака будет длиться в среднем 364 дня, что вполне сопоставимо с вечностью, однако, если атакуется не какая-то конкретная машина, ситуация резко меняется. Допустим, у хакера имеется 100 потенциальных жертв, которым рассылаются пакеты. Тогда, как нетрудно рассчитать, среднее время атаки сокращается до 3,6 дня. Причем, атакующий может свободно менять свой собственный IP, ведь UDP работает без установки соединения, а ловить ответ хакеру не нужно. В пересчете же на каждый используемый IP, интенсивность посылки пакетов находится ниже порога чувствительности сенсоров, а потому IDS вместе с аппаратными антивирусами сидят тихо и не возникают.

Хорошо, с сенсорами мы разобрались. Займется теперь сигнатурами. Их тоже несложно одолеть. Даже без привлечения полиморфизма. Отбросим бесполезную мелочь и сосредоточимся на антивирусах, установленных на магистральных каналах, обладающих достаточной мощью для сбора всего TCP-пакета и несущие на своем борту обширные базы

оперативно обновляемых сигнатур, бьющие массированную атаку буквально через считанные часы (а то и минуты) после ее начала. Существует ли _универсальный_ способ обхода _заранее_ _неизвестного_ антивируса? Оказывается, существует. Причем такой, которому никакой _магистральный_ антивирус _принципиально_ не может противостоять. Прежде, чем придумывать убийственный контраргумент вернемся к истокам и вспомним с чего все начиналось. Интернет — сеть, продолжающая функционировать даже после начала атомной войны (тогда она казалась неизбежной), когда большинство узлов разрушено. Сеть, в которой пакеты самостоятельно (ну не совсем самостоятельно, конечно), прокладывают себе маршрут. Сеть, в которой два фрагмента одного TCP-пакета из пункта A в пункт B могут идти разными путями… Гм, а ведь на счет путей это идея! IP-пакеты, пущенные разными маршрутами, окончательно собираются в TCP только на целевом узле, и никакой отдельно взятый магистральный антивирус не в состоянии собрать полный TCP-пакет, поскольку через него физически "прокачивается" только небольшая его часть!

Рисунок 6 хакер-невидимка

Как реализовать такую систему на практике? Имея домашний компьютер с несколькими сетевыми интерфейсами (ADSL-модемом и сотовым телефоном с GPRS) нетрудно написать утилиту, разбивающие исходное послание на IP-пакеты, пускаемые через разные интерфейсы. Но это неинтересно. Смысла нет. С целенаправленными атаками магистральные антивирусы не борются, а если у жертвы (или у ее провайдера) установлен хотя бы простенький

программный антивирус или брандмауэр, то подобное дробление ничем не поможет атакующему, поскольку TCP-пакет будет собран на узле, где установлен программный антивирус/брандмауэр или во всяком случае _все_ IP-пакеты пройдут через него и он сможет собрать полный TCP.

А вот для червей это очень даже хорошая стратегия. Допустим червь уже заразил два узла, находящиеся в различных подсетях и теперь хочет кинуть свою тушку на третий. Посылая пакеты с двух узлов одновременно (не забывая о том, что при этом придется реализовать определенный протокол синхронизации, т.к. TCP работает с установкой соединения, маркируя пакеты номерами последовательности) червь пройдет сквозь аппаратный антивирус без всяких преград, даже не заметив, что тут кто-то был!

Главное — добиться того, чтобы червя нельзя было отождествить по одному отдельно взятому пакету. А как это можно сделать? Да очень просто! Достаточно написать тривиальный криптор, "размазанный" по всем пакетам и шифрующих их содержимое произвольным ключом. Очевидно, что не зная ключа (который можно получить только собрав все пакеты воедино), антивирус не сможет расшифровать вирусное тело, а, значит, не сможет и отождествить его по сигнатурам. Естественно, XOR с константой палится еще на излете (если ключ представляет из себя 1 байт, то 256 сигнатур детектят червя по любому произвольно взятому IP-пакету), но вот уже RC4 (реализуемый ничуть не сложнее) таким способом уже не словить и победа остается за червями!

Рисунок 7 Интернет после очередной эпидемии червей

заключение

Интернет не умрет никогда. Глобальные эпидемии в исторической перспективе — явление вполне закономерное, можно даже сказать неизбежные. Всех нас будет колбасить и плющить, так что не стоит сопротивляться, а лучше расслабится и спокойно наблюдать за гонкой вооружений двух противоборствующих сторон — хакеров и создателей защитных механизмов, дополняющих друг друга как инь и янь, как свет и тьма, как день и ночь, как добро и зло…

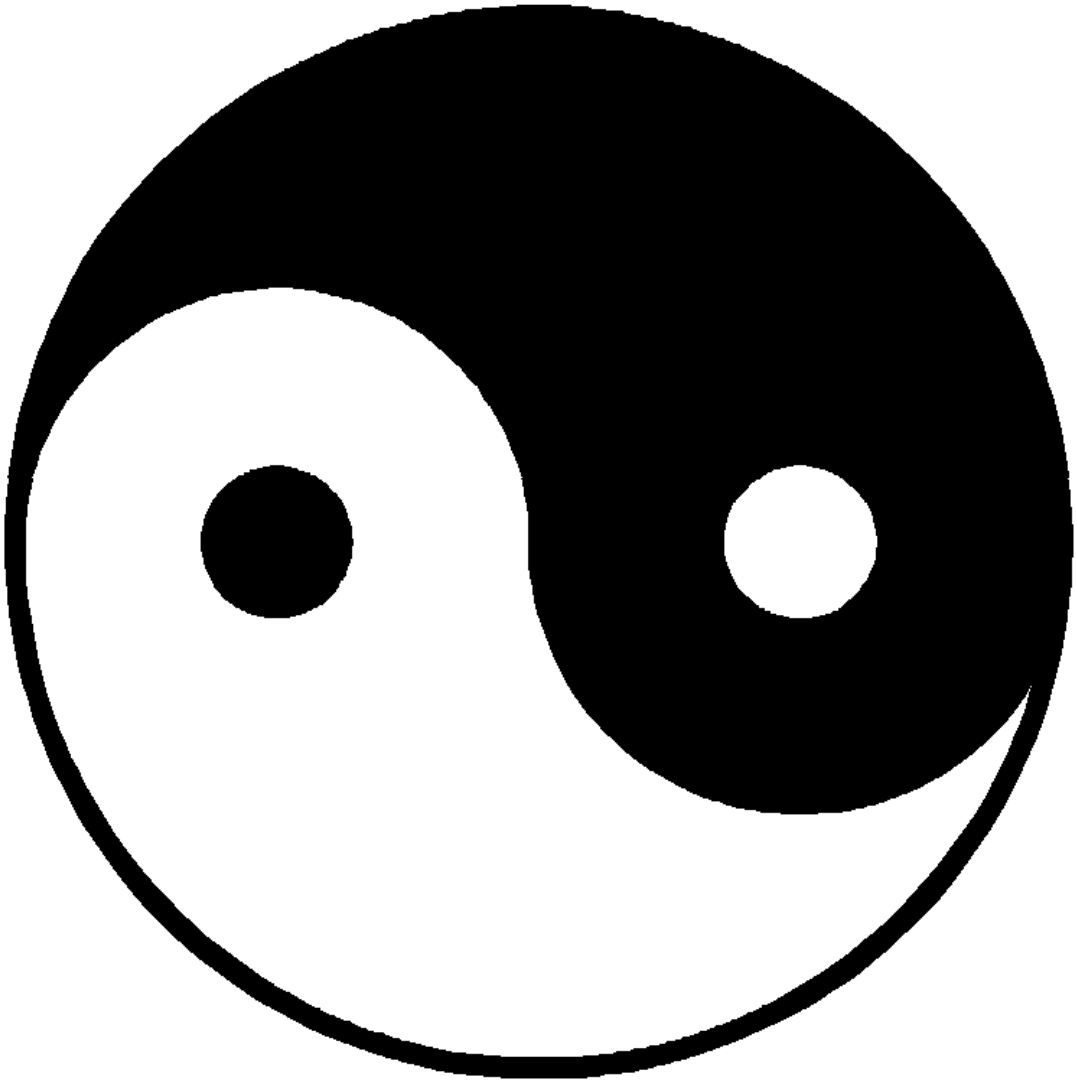

Рисунок 8 гармония противоборствующих сторон