полиморфные технологии на службе спаммеров

крис касперски ака мыщъх, ака nezumi, ака souriz, ака elraton, ака толстый хомяк, no-email

ожесточенная борьба со спамерами не приводит к их вымиранию, напротив — побуждает разрабатывать новые виды оружия. все что нас не убивает — делает сильнее! извечная проблема меча и щита, превратившая каменные топоры в атомные бомбы. если в остальных статьях мыщъх говорил, как фильтры борются со спаммерами, то сейчас развернет свой хвост на 180 градусов, показав откуда от торчит.

введение

Как ни печально об этом говорить, но рынок "веерных" рекламных рассылок уже сложился, одиночки уши в туман, а на арену вышли крупные игроки, борющиеся не только с фильтрами, но и со своими коллегами по "цеху". Дилетантам здесь не место и в конкуренткой борьбе побеждает либо сильнейший (в смысле ширины каналов), либо умнейший (причем использовать _свои_ мозги может только законченный идиот). Спаммеры активно пользуются вирусными наработками и тщательно изучают все образцы оружие, предназначенного для борьбы с ними. Но, как учит военная мудрость, "не рой яму другому, чтобы он не использовал ее как окоп".

Прежде, чем приступать к рассылке, опытный спаммер обязательно установит у себя последние версии всех фильтрующих систем и будет "рихтовать" письмо до тех пор, пока оно не обретет достойный вид, ничем не выделяющийся среди общего потока корреспонденции. Затем начнет поиск подходящего proxy или релеея, пригодного для массовой рассылки, но отсутствующего в DRBL-базах (банках данных, хранящие сведения об серверах и узлах, хотя бы однажды замеченных в спаммерской активности или допускающим массовую рассылку без авторизации).

Естественно, чтобы рассылка не была накрыта баллистической ракетой через несколько минут после ее начала, необходимо предпринять ряд дополнительных шагов, например, постоянно менять дислокацию, используя распределенную сеть дронов (обыкновенных пользовательских компьютеров, подключенных к Интернет и предварительно зараженных червем, установившим back-door). Тогда DRBL-базы окажутся бессильны. Ведь артиллерийским огнем весь Инетернет не накроешь, а пользователи, занесенные в "черные списки" еще и в суд подать могут — с чего это вдруг их лишили электронной почты?!

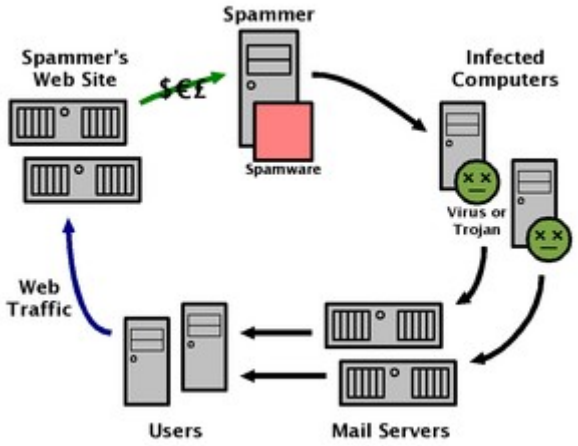

Рисунок 1 использование компьютеров-дронов для обхола DRBL-баз

Впрочем, DRBL-базы выявляют и отсеивают всего 20% - 30% спама. Намного более серьезную угрозу представляют сигнатурные фильтры. Даже если рассылка велась с тысячи разных IP-адресов, но рассылалось одно и тоже письмо, любой нормальный почтовый сервер классифицирует его как спам и адресат не получит рекламы. Какая жалость! (для спамера и той сволочи, что заказала рассылку).

Следовательно, рассылаемые письма должны отличаться друг от друга, если не по смыслу, то хотя бы по форме. А это не так-то просто сделать! Содержимое письма уникально и может меняться произвольным образом. Ну какому рекламодателю понравится, если женские прокладки с крылышками будут заменены чугунными трубами с левой резьбой?! А телефоны и контактные адреса?! Для сигнатурного поиска — это самое то!

Как же все-таки спамерам удалось перехитрить систему фильтров?! А в том, что это им удалось, никаких сомнений не остается, даже сам Касперский в этом честно признается: http://www.kaspersky.ru/faq?chapter=168403418&qid=169455884 (см. раздел "Kaspersky Anti-Spam 2.0 пропускает очень много СПАМа").

каменный век — первые эксперименты

Давным-давно, когда Интернет был медленным, а письма рассылались преимущественно в "голом" текстовом формате, "химичить" с форматом особо не получалось. Какое там творчество! Какой там полет хакерской мысли! Ладно, берем номер телефона и думаем как бы его видоизменить так, чтобы и клиент смог дозвониться, и в то же время фильтр не съел. Меняем ноль на букву "О", единицу — на "I", тройку — на "З". Так же можно добавлять пробелы, скобки и тире в разных местах. Тоже самое можно сделать и с текстом письма, тут даже появляется больше свободы, поскольку помимо замены сходных по начертанию букв, можно заменять слова их синонимами, менять блоки текста местами и т. д.

Еще один хитрый прием — не указывать кодировку письма, а предоставить получателю или его почтовому клиенту определить это автоматически, правда, для автоматического определения кодировки требуется достаточно длинное письмо, а если оно будет коротким, справиться с этой задачей сможет только человек. Как следствие, вместо одной сигнатуры фильтр получает целую кучу. Вероятность ложных срабатываний — увеличивается, а качество распознавания спама — ухудшается. Кстати говоря, немногословные рекламные рассылки в стиле "нары новые, само вывозом, звонить шесть-шесть-девять с кодом чукотки" практически не распознаются никаким фильтрами, поскольку объем значимой информации в них минимален (да и та может быть видоизменена на любой манер) а сверху и снизу легко наклеить заголовки с приветствиями/поздравлениями.

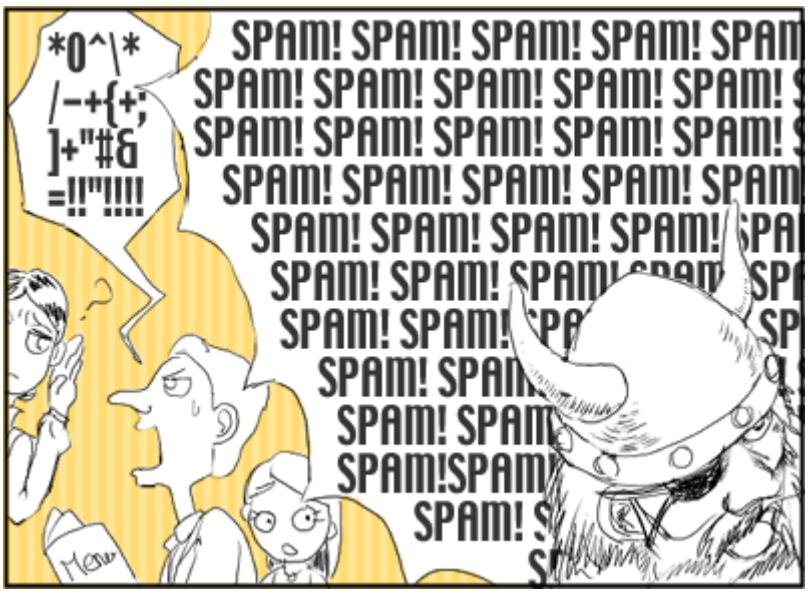

Рисунок 2 технологии полиморфного спама каменного века

Используя "движки", выдернутые из программ, имитирующих некоторое подобие диалога с человеком, американские спамеры сумели создать генераторы, передающие одно и тоже сообщение бесконечным множеством вариантов. Грубо говоря как бы "компилятор" текста. Обратный же процесс, выделения ключевой "мысли" уже требует применения искусственного интеллекта или сложного лингвистического анализа, который в полной мере не реализован и до сих пор. Кое-что имеется у Касперского, но… могучий русский язык снимает проблему "компиляторов текстов" сам собой. Уж очень сложно написать программу, транслирующую исходное сообщение более чем десяток вариантов. В английском с этим

проще. Жестко заданный порядок слов в предложении, простейшие лексические правила, скромный лингвистический набор (в следствии которого каждое слово имеет множество синонимом), легкая стыковка слов друг с другом, позволяющая (с некоторыми ухищрениями) обходится без предлогов…

Русский язык обладает развитой системой сложных правил с кучей исключений. Одна мышь, две (три, четыре) мыш**и**, пять мыш**ей**! Вот и попробуй все это заложить в программу!!! Тем не менее, работа над созданием "компиляторов" русского текста ведется и весьма активно. Взять хотя бы разработчиков игр. Чтобы персонажи не выкрикивали одни и те же навязшие в зубах фразы, необходимо научить машину генерировать произвольные фразы на основе заданной мысли. А в игровой индустрии замешаны совсем не малые деньги и есть все основания предполагать, что такие генераторы когда-нибудь да появиться. Тогда… ни синтаксический, ни лексический анализ ни за что не сможет отличить спам от простого письма.

>>> врезка оружие возмездия

Сейчас в научных институтах всего мира идет интенсивная работа по созданию "смысловых" анализаторов для русского и английского языков (русский мыщъх поставил первым, поскольку он для него родной, хоть с технической точки зрения и находится в хвосте прогресса). Разбивать предложение на части речи (во всяком случае для английского языка) научились уже давно, затем объяснили машине (пускай, не без набора шаблонов) как эти части связаны друг с другом, благодаря чему приблизительный смысл удается восстановить даже если значительная часть слов отсутствует в машинном словаре.

Коммерческие компании тоже не отстают (например, у всем известной ABBY над этим целый этаж работает, причем не первый год). Главная цель — машинный перевод (ну нельзя качество переводить текст, зная только значения отдельных слов, но не их роль в предложении), создание реферативных систем ("заглатывающим" простыню текста и в двух словах объясняющих о чем он), ну и поиск стоит не на последнем месте. Это сегодня мы вынуждены мучительно перебирать все комбинации ключевых фраз, а через несколько [десятков] лет будет достаточно сказать: "проблемы тушканчиков средней полосы" и дальше машина уже сама.

Побочным эффектом создания таких анализаторов станет окончательная победа над спамом, поскольку независимо от формы рекламного сообщения, его суть остается прежней — рекламной. Конечно, полностью смап не исчезнет, просто притихнет на некоторое время, а потом разгорится новый виток борьбы: составить письмо, богатое идиоматикой, чтобы человек его понял, а машина нет!

Рисунок 3 уже столько раз борцы со спамом провозглашали полную и окончательную победу, а борьба тем временем все накаляется и накаляется…

HTML – начало конца

Массовое внедрение поддержки формата HTML в почтовые клиенты необычной расширило границы спамерской активности и серьезно напрягло фильтры, поскольку теперь

прежде чем начинать какой бы то ни было анализ, необходимо "распарсить" HTML, выделив из него текст, по обыкновению тесно перемешанный с тегами, словно это фаршированный голубец. А парсинг требует времени и процессорных ресурсов, а вместе с ними еще и знания психофизических моделей и особенностей зрительной системы человека. Иначе можно очень просто разместить между символами сообщения "невидимый" текст: мелкий шрифт или шрифт по цвету полностью или практически полностью совпадающий с фоном. Это все просто и понятно. А вот то, что ярко-желтый плохо различим на фоне ярко-зеленого знает уже не каждый (фильтр).

Современные фильтры, конечно, HTML знают как свой собственный хвост, а сам факт наличия "невидимого" текста трактуют как спам даже не прибегая к сигнатурному поиску! К тому же, "продвинутые" почтовые клиенты типа The Bat! имеют режим "упрощенного HTML", игнорирующий цвета, шрифты и прочую дребедень подобного типа. Очень удобно для чтения писем от респондентов изображающих из себя гениев дизайна на уровне третьего класса. Естественно, в упрощенном режиме отображения, весь невидимый текст вылезает на поверхность, делая сообщение совершенно нечитаемым. Тоже самое относится и к обычным почтовым клиентам. Пускай, спамер перемешал номер контактного телефона невидимыми символами. Заинтересованный клиент копирует его в буфер обмена (ну не в ручную же его переписывать!) и… к своему удивлению вместо телефона видит какую-то невменяемую хрень.

Короче, от всех этих фокусов с HTML'ом спамеры постепенно отказались, поскольку они себя не оправдали ни с какой стороны. Фильтры подтянули качество распознания HTML-спама до прежней отметки (и даже перешагнули ее, с учетом нетипичных для "честных" писем "извращений"), а пользователи, даже те, что заинтересовались рекламой, не всегда могли ей воспользоваться. Плюс ко всему рассылка HTML'а длиться дольше и обходится гораздо дороже (в плане трафика). А скорость рассылки определяет все! Как только образцы непрошеной корреспонденции попадают в DRBL-базы, то даже при условии 100% полиморфизма (совершенно недостижимого в HTML'е) IP-адреса начинают давить один за другим и даже очень крупная армия дронов гибнем за считанные десятки минут, ну от силы — часы.

король палитра первый

Эпидемии графического спама то вспыхивают, то затухают. Вначале это были просто картинки, вставленные в "честный" HTML-текст с номерами контактных телефонов и прочей уникальной информацией, однозначно идентифицирующей спам. Фильтры первых поколений игнорировали картинки, но были быстро доработаны и полноводный поток спама, захлестнувший Интернет, сразу иссяк. Тогда спамеры применили готовые генераторы изображений, предотвращающие автоматическую регистрацию на почтовых серверах, и вносящие в начертания символов некоторые искажения. Казалось бы, фильтры, не рыпаясь, должны были дружным строем идти сдаваться на мясокомбинат, однако, все вышло совсем не так…

Незначительные искажения (или невысокую степень зашумленности изображения) фильтры распознают чисто статистическим методом по кривой Гаусса (кто изучал метрологию — знает что это за штука). Да! Фильтр не в состоянии OCR'ить изображение, но это ему и не нужно! Имея в своем распоряжении большое количество случайным образом искаженных изображений, он просто выделяет свойственные им "родственные черты" и палит их на месте!

Значительные искажения уже не распознаются фильтрами, но чтобы их разобрать, получателю приходится совершать значительные насилия над собой, натягивая глаза на жопу. Это же насколько его должна заинтересовать реклама, чтобы он так извращался?! Так что, независимо от количества успешно доставленных писем, эффективность такого спама близка к нулю, следовательно, и популярность то же.

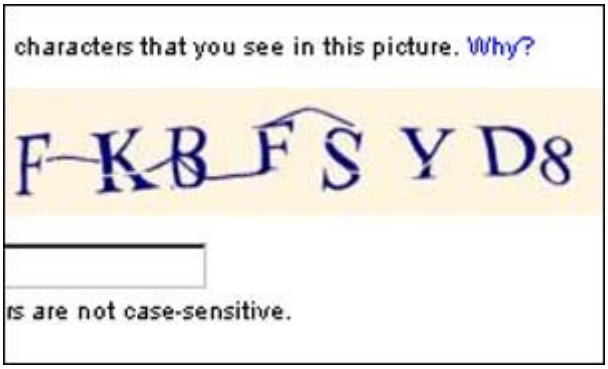

Рисунок 4 такие искажения легко распознаются человеком, но обманывают простейшие сигнатурные фильтры

царствие жабье

Подлинный полиморфизм стал достижим только с появлением в HTML скриптовых языков (в частности, Java Script), проникших даже в популярные почтовые клиенты. Казалось бы, зачем электронному письму тащить на своем борту какой-то там язык?! Это же ведь не сайт в конце концов! Тем не менее, применение ему все-таки нашлось. Например, сотрудники компании получают по мылу некоторую форму, а Java тут же проверяет корректность заполнения полей, исключая наиболее глупые ошибки. Боле разумного применения скриптам придумать, пожалуй, невозможно. Тем не менее они есть и мало того, что служат неиссякаемым источником ошибок, приводящих к возможности захвата управления компьютером или утечке конфиденциальных данных, так они еще и спамерам помогают!

Полиморфный Java/VBasic-спам делится на две категории. Первая (самая многочисленная и самая простая в реализации) основана на функции (функциях), расшифровывающих зашифрованный текст послания и выводящий его в окно почтового клиента "на лету". Поскольку, ключ шифрования может меняться с каждым письмом, то сигнатурному фильтру необходимо иметь на своем борту полноценный виртуальный интерпретатор, "перерабатывающий" скрипты и анализирующий выдаваемое ими содержимое. Это же какие аппаратные мощности иметь надо, чтобы выполнять такой анализ в реальном времени?! Поставишь такой фильтр и бумажная почта будет ходить быстрее электронной! Исходный текст подобных скиптов здесь не приводится, чтобы не облегчать спамерам жизнь. К тому же есть одна очень веская зацепка. Шифруя содержимое письма, сам код шифратора остается неизменным и может быть использован в качестве сигнатуры.

Полиморфики второй категории не только генерируют случайный ключ, но и произвольным образом модифицируют сам расшифровщик, препятствуя выделению устойчивой сигнатуры. Для этого спамеру даже не требуется рвать себе задницу, поскольку появилось множество готовых Java-обсускаторов, запутывающих исходный код скрипта до такой степени, что в ней не остается ни одной устойчивой сигнатуры и все фильтры отдыхают. Правда, сам факт наличия запутанного Java-кода указывает на явную ненормальность письма, выдавая его спамерскую принадлежность, поскольку, у него _совсем_ другое частотное соотношение java-команд (ах, для этого фильтр должен еще и жабу знать, но… что поделаешь! жизнь еще и не на то вынуждает!)

Вот и приходится хитрить, создавая готовые генераторы функций-шифраторов/дешифраторов, ни статистически, ни "лингвистически" неотличимых от прочих java-функций, которые все чаще и чаще встречаются в обычных письмах и здесь фильтры уже вынуждены проявлять осторожность.

Естественно, создание генератора подобного рода является серьезнейшей инженерной задачи и требует от программиста высокой квалификации, но… если постараться и затараниться пивом с питтцей и хот-догами, за несколько дней можно и вручную создать пару тысяч совершенно различных пар шифровщиков/дешифровщиков, с легкостью проходящих сквозь системы фильтров, поскольку процент "совпадающих" писем в рассылке окажется ниже того порога, начиная с которого фильтрация становится эффективной.

заключение

Посмотрим с оптимизмом в завтрашний день!!! Хм… и эту ночь они называют днем?! Ну и нравы у людей пошли! С приходом в Сеть коммерции ее накрыл рекламный мрак и от спама нам уже никуда не уйти. Единственный позитив, который нельзя не отметить, спам становится все более качественным и контекстно-чувствительным. Над созданием сообщений работают если не дизайнеры, то все-таки пыхтят так, чтобы нужная информация сразу же бросалась в глаза и оседала в мозгу даже после того, как человек рефлекторно нажмет <DEL>. К тому же, зная IP-адрес получателя (а в большинстве случаев его можно установить тем или иным путем), спампер определяет его географическую принадлежность и шлет рекламу соответствующую ареалу обитания "жертвы" (туркам шерстяные одеяла не предлагать!).

Рисунок 5 рассчитывать на победу над спамом — бесполезно, но если с ним не бороться, в нем можно утонуть