Вестник ЦЭМИ РАН. 2018

ISSN 1111-1111

URL - http://cemi.jes.su

Выпуск 2 Том . 2018

Проблемы защиты от спама: таргетированный спам

Ляпичева Н. Г.
Центральный экономико-математический институт РАН
Российская Федерация, Москва, Нахимовский проспект, 47

Аннотация

Целью описанной в статье работы является решение проблемы защиты от таргетированного спама, использующего механизмы контекстной рекламы в массовых рассылках. Анализ спам-трафика проводился на как на основе протоколирования заголовков входящей почты, проводившегося антивирусным/антиспамовым шлюзом , так и экспертной оценки собственного полученного почтового трафика. В результате предложена схема использования почтовых клиентов, позволяющая минимизировать количество получаемого спама. Схема может быть рекомендована для пользователей стандартных почтовых клиентов в частично обновленной аппаратно-программной среде

Ключевые слова: почтовый сервис, защита от спама, информационная безопасность, таргетированная реклама

Дата публикации: 13.12.2018

Ссылка для цитирования:

Ляпичева Н. Г. Проблемы защиты от спама: таргетированный спам // Вестник ЦЭМИ РАН. 2018. Выпуск 2 [Электронный ресурс]. Доступ для зарегистрированных пользователей. URL: http://cemi.jes.su/s111111110000062-1-1 (дата обращения: 17.12.2018). DOI: 10.18254/S0000062-1-1

1 ВВЕДЕНИЕ

Контекстная реклама, одним из вариантов которой является персональная (таргетированная) реклама – представляет собой часть информационного пространства, с которой в настоящее время сталкивается практически каждый пользователь интернета при обращении к наиболее значимым поисковым системам[1] [2] [3]. Конструкторам сайтов активно предлагаются программные продукты для сбора данных из поисковых запросов, конечной целью сбора является предоставление контекстной рекламы отправителям. Это направление развития общедоступного информационного пространства, в свою очередь, породило явление, которое можно назвать: «персонализация спама» или «таргетированный (целевой) спам». Внешне оно выражается в нарастании объемов спама, нефильтруемого централизованными средствами очистки почтового трафика.

2 ОСНОВНАЯ ЧАСТЬ

В течение довольно длительного периода в ЦЭМИ РАН обеспечивался достаточно высокий уровень очистки почтового трафика от спама >= 90% [4] [5] [6], за счет использования многоуровневой схемы антиспамовой обработки. Функционирование централизованных средств очистки почтового трафика (антиспамового/антивирусного почтового шлюза) отслеживается как в оперативном режиме, так и в виде ежемесячных отчетов о состоянии почтового трафика.

3 В ежемесячных отчетах зафиксировано падение производительности антиспамового шлюза до ~70%, т.е. % принятых писем повышается с 10-15% до 20-30% (см. Рис. 1, Процент писем, принятых антиспамовым шлюзом за период 2013-2018 г).

4

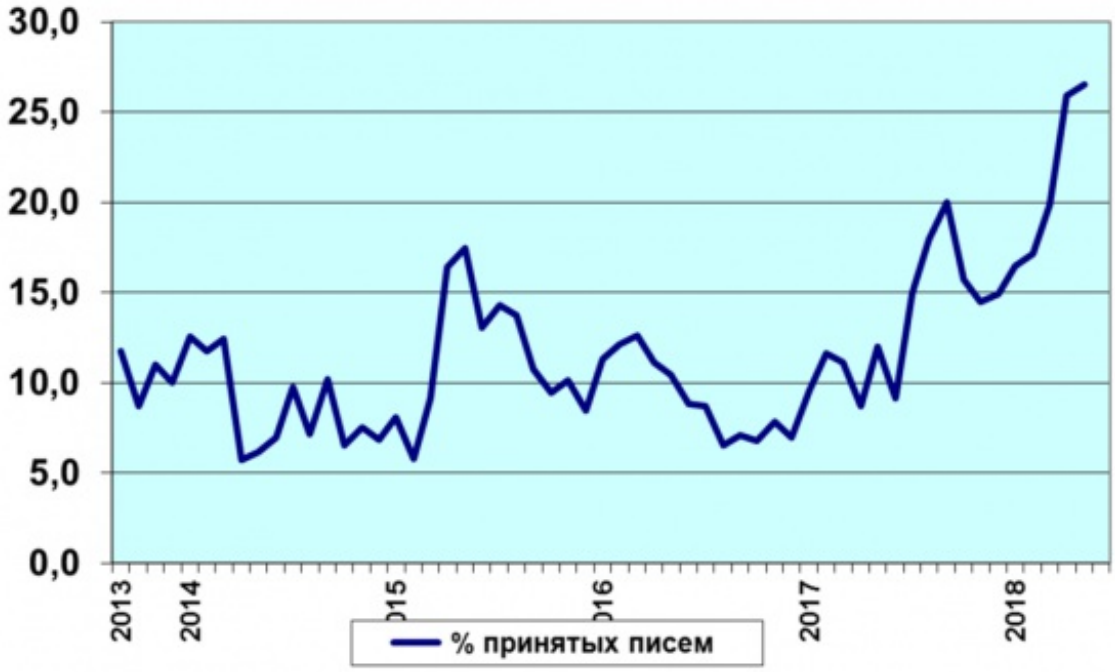

Рис. 1 Почтовый трафик антиспамового шлюза за период 2013-2018 г.

5 Одновременно с этим зафиксировано увеличение количества спама, приходящего в почтовые ящики и изменение его состава, что привело к изучению проблемы, возникшей в работе антиспамового шлюза.

6 При изучении заголовков и содержания приходящего спама использовался собственный (автора) почтовый трафик, ежедневно приходящий на ряд служебных почтовых ящиков и личный ящик на сайте mail.ru: -Персональный (личный) почтовый ящик расположен на сервере inbox.ru. В нем количество постороннего спама минимально, почти весь спам представляет собой рассылки, запрошенные при регистрации на различных сайтах или не запрошенные, но «по умолчанию» включенные администраторами сайтов. -Два служебных почтовых ящика расположены на сервере в домене cemi.rssi.ru, используются давно и существуют во многих спам-базах, по которым проводятся массовые рассылки. В результате в эти ящики постоянно поступает некоторое количество спама, отражающее общее состояние функционирования почтовой системы института.

7 Пришедшие в почтовые ящики письма рассматривались и классифицировались в период 29 мая – 18 июня 2018 г. При этом выделялись реальные деловые письма, служебные письма (информационные сообщения от сервисов узла ЦЭМИ РАН) и спам.

8 В спаме по полю «Тема», а иногда и по содержанию определялось, случайная это реклама или же таргетированная, созданная в результате сбора информации о поисковых запросах автора и предназначенная только для интересующихся данной тематикой. Основным критерием, по которому спам может быть отнесен к целевому, является личное мнение спамополучателя при на рассмотрении поля «Тема:»(«Subject:»). Можно рассматривать и содержание полученных спам-писем, но это небезопасно из-за еще не распознанных почтовых вирусов и возможного фишинга – ссылки на вредоносный сайт.

9 Из выборки за упомянутые даты можно выделить по критерию «Тема:» отдельные группы спам-писем.

10 **Нецелевой спам:** темы этой группы (и предмет рекламы) случайны, никогда не упоминались автором при поиске или в другом контексте. Пример нецелевой тематики:

11 Таблица 1 Темы нецелевого спама

Восстановитель для авто Bright New
Все говорят ака а как поднять бабла
Компактная рыболовная сеть
Превосходное звучание от Harman Kardon!

Развивающая и увлекательная игрушка. Кукольный эко-домик

Справка о валютных операциях

Фирменные японские воблеры с системой дальнего заброса.

Щенячий Патруль. Конструктор для детей от 1 года до 5 лет

12 (Можно отметить, что наиболее одиозные темы практически не проскакивают через антиспамовый шлюз, например, неприкрытые предложения секс-услуг).

13 **Таргетированный (целевой) спам** в рассмотренный период был представлен несколькими группами тем:

14 Темы **первой группы**, сформулированные приблизительно в таких терминах, одноразово или крайне редко использовались автором в поисковых системах (обычно – на Яндексе).

15 Таблица 2 Темы первой группы

Магнитная щетка для окон

200 руководителей нефтегаза Дальнего Востока и Восточной Сибири

Шикарный газон на вашем участке на всё лето всего за 1 неделю на любой почве

Уникальный инструмент для сада и огорода

16 Темы **второй группы** приблизительно относятся к рабочей тематике автора, но не являются деловыми письмами от легальных отправителей.

17 Таблица 3 Темы второй группы

Как усилить Wi-Fi сигнал

Science conference, Bulgaria

Ляпичева Нина Григорьевна Ищете научный журнал РИНЦ? Это наш!Срок приема статей продлен

Семинар "Управление электроприводом с помощью MATLAB, Simulink" (27 июня, Москва) и сертификация по MATLAB/Simulink (21-22 июня, Москва)

18 Темы **третьей группы** привязаны к самым свежим поисковым запросам, которые в период проведения исследования были использованы при ведении производственной практики у студентов ГАУГН (например, было сделано несколько запросов по Трудовому кодексу относительно кадрового учета и начисления зарплаты).

19 Таблица 4 Темы тетьей группы

В состав АВАНСА должны включаться: ПРЕМИИ, Надбавки и Компенсации

Почему в 2018г «СГОРАЮТ» НЕИСПОЛЬЗОВАННЫЕ ОТПУСКА и денежная компенсация после увольнения: судебная практика-2018г

ОТПУСК вовремя НЕ ОТГУЛЯННЫЙ через 21 месяц АННУЛИРУЕТСЯ без компенсационной выплаты

Отказать в ПРИЕМЕ НА РАБОТУ лицам СТАРШЕ 60 ЛЕТ будет СЛОЖНО с Октября 2018г (новый пакет законов в рамках пенсионной реформы)

20 В рассмотренной выборке почтового трафика наиболее обильно

представлены письма третьей группы - по несколько в один день и в течение ряда дней.

21 Так как антиспамовому шлюзу затруднительно определить, является некое письмо целевым спамом или же оно действительно было адресовано конкретному получателю, то письмо считается легитимным (в худшем случае - подозрительным на спам, с соответствующей отметкой) и принимается на доставку.

22 Результат исследования отображен на графике «Процент целевого спама за период 29 мая – 18 июня 2018 г.» (см. Рис.2), - общее количество пришедшего спама и в нем процент целевого спама.

23

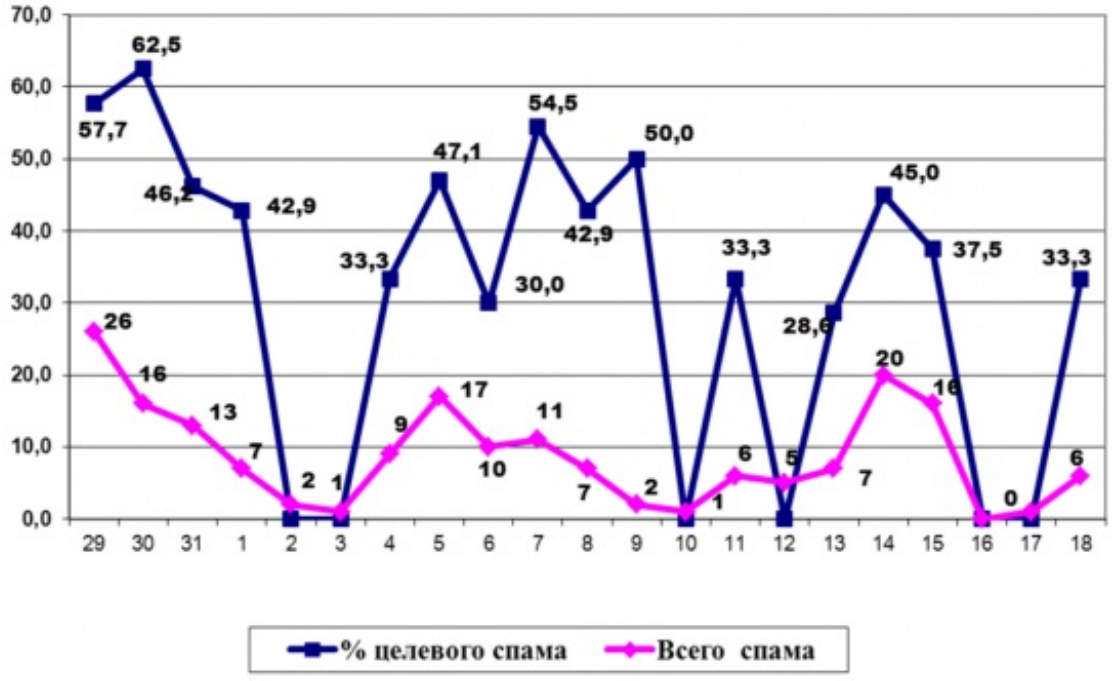

Рис.2. Процент целевого спама за период 29 мая – 18 июня 2018 г.

24 Процент целевого спама, как можно видеть из графика, колеблется между 30% и 60%, причем чем количественно больше спама, тем выше и процент целевого спама. Цикличность графика связана с рабочими днями и выходными, в которые общий объем спама падает почти до нуля – до почти полной очистки трафика антиспамовым шлюзом. В эти же дни соответственно падает и процент целевого спама.

25 РЕЗУЛЬТАТЫ.

В результате изучения почтового трафика были проведены меры по настройке фильтров почтового клиента Thunderbird, который давно используется для работы с электронной почтой. Определенную трудность вызывает то, что каждый новый вид проскочившего спама следует помечать соответствующим значком, после чего данный вид спама переносится в специальный ящик и в дальнейшем заносится прямо туда. Эти меры позволяют

довести очистку входящей почты до приемлемой степени. Самым легким способом, который хорошо совмещается с проверкой входящей почты, является установка отметок «спам» на тех письмах, которые не распознаются стандартными анализаторами спама – т.е. на письмах «целевого спама».

26 ВЫВОДЫ

Не углубляясь в вопрос, как именно спамеры получают данные о предпочтениях пользователей, приходится принять как данность то, что разбираться в легитимности писем, принятых антиспамовым шлюзом на доставку, приходится самому получателю. Глобальное решение проблемы ограничения доступа пользователя к информации в Интернете лежит вне темы защиты от спама и носит не только технологический характер. **Решение проблемы находится в руках пользователя** – настройка почтового клиента на удаление паразитного почтового трафика., что так же не отменяет необходимости в остальных достижениях информационной безопасности [7] [8] [9] [10]. В процессе чтения почты реализуется окончательная очистка почтового ящика, пользователя что обеспечивается настройкой спам – фильтров почтового клиента[11].: Каждый разработчик современных операционных систем и офисных приложений предусматривает средства анализа и удаления нежелательных сообщений. Собственно, именно так и поступают все популярные бесплатные почтовые сервисы (mail.ru, yandex.ru, и др.) – в них сам пользователь доводит спам-фильтры персональных почтовых ящиков до необходимой степени очистки.

Примечания:

1. Сбор данных о каждом пользователе, умное сегментирование и таргетинг — комбо маркетолога на 2016-й. Сервис по автоматизации интернет-маркетинга Carrot Quest [Электронный ресурс]. 2015. URL:https://habr.com/company/carrotquest/blog/297326/ (Дата обращения 16 декабря 2015)

2 . Контекстная реклама. Википедия [Электронный ресурс].URL: https://ru.wikipedia.org/wiki/ (Дата обращения 16 марта 2018)

3 . Софт и сервисы для профессионального поиска. HRазведка:блог [Электронный ресурс]. URL:http://hrazvedka.ru/poisk_soft/soft-i-servisy-dlya-professionalnogo-poiska.html

4. Ляпичева Н.Г. Анализ вирусной активности в почтовом трафике на узле ЦЭМИ РАН../Труды Третьей Всероссийской научно-практической конференции «Научное, экспертно-аналитическое и информационное обеспечение стратегического управления, разработки и реализации приоритетных национальных проектов и программ», 31 мая - 1 июня 2007 года, ИНИОН РАН, М:2007, сс. 538-542

5 . Ляпичева Н.Г.,Никонова О.М. Математико-статистический анализ объёмов спама на узле ЦЭМИ РАН. /Обозрение прикладной и промышленной математики, т.15, Выпуск 4. ТВП, 2008. сс.670-671

6 . Ляпичева Н.Г., Никонова О.М., Современные проблемы почтового сервиса. /В сб. Обозрение прикладной и промышленной математики. Десятый Всероссийский симпозиум по прикладной и промышленной математике, Санкт-Петербург, 19-24 мая 2009 г. Тезисы докладов. М.:ОПиПМ, 2009. т.16, вып.2. сс.362-363

7. Терентьев А.М. Выбор адекватных средств информационной защиты персонального компьютера в России / Журнал «Национальные интересы. Приоритеты и безопасность» М., ООО «Издательский дом Финансы и кредит», N33(174)-2012, с. 37-42

8. Терентьев А.М. Корпоративный вариант реализации антивирусных пакетов Doctor Web в научных учреждениях: предпосылки. / Журнал «Национальные интересы. Приоритеты и безопасность» М., «Издательский дом Финансы и

кредит», N17(206), 2013, с.41-48.

9. Терентьев А.М. Корпоративный вариант реализации антивирусных пакетов Doctor Web в научных учреждениях: реализация. / Журнал «Национальные интересы. Приоритеты и безопасность» М., «Издательский дом Финансы и кредит», N19(208), 2013, с.40-45

10. Терентьев А.М. Корпоративный вариант реализации антивирусных пакетов Doctor Web в научных учреждениях: результаты. / Журнал «Национальные интересы. Приоритеты и безопасность» М., «Издательский дом Финансы и кредит», N20(209), 2013, с.41-46 mail service, anti-spam protection, information security, targeted advertising

11. Ляпичева Н.Г., Никонова О.М.,Левенко Е.С. Избавление от спама: «последняя миля» почтового сервиса. / Сб.. «Обозрение прикладной и промышленной математики». М., ОПиПМ, 2012. т.19, вып.4. с. 581-582

Библиография:

1. Контекстная реклама [Электронный ресурс] 2015 https://ru.wikipedia.org/wi

2. Сбор данных о каждом пользователе, умное сегментирование и таргетинг — комбо маркетолога на 2016-й. Сервис по автоматизации интернет-маркетинга Carrot Quest [Электронный ресурс] 2015 https://habr.com/company/carrotquest/blog/297326/

3. Софт и сервисы для профессионального поиска [Электронный ресурс] 2018 http://hrazvedka.ru/poisk_soft/soft-i-servisy-dlya-professionalnogo-poiska.html

4. Ляпичева Н.Г.Анализ вирусной активности в почтовом трафике на узле ЦЭМИ РАН.. // Труды Третьей Всероссийской научно-практической конференции «Научное, экспертно-аналитическое и информационное обеспечение стратегического управления, разработки и реализации приоритетных национальных проектов и программ», 31 мая - 1 июня 2007 года, ИНИОН РАН, М:2007, сс. 538-542 (рус.), http://cemi.socionet.ru/files/inion2007_ngl.doc.

5. Ляпичева Н.Г.,Никонова О.М. Математико-статистический анализ объёмов спама на узле ЦЭМИ РАН. // Обозрение прикладной и промышленной математики, т.15, Выпуск 4. ТВП, 2008. сс.670-671

6. Ляпичева Н.Г., Никонова О.М., Современные проблемы почтового сервиса. // В сб. Обозрение прикладной и промышленной математики. Десятый Всероссийский симпозиум по прикладной и промышленной математике, Санкт-Петербург, 19-24 мая 2009 г. Тезисы докладов. М.:ОПиПМ, 2009. т.16, вып.2. сс.362-363 http://cemi.socionet.ru/files/doklad2009_ngl-nik.pdf

7. Ляпичева Н.Г., Никонова О.М.,Левенко Е.С. Избавление от спама: «последняя миля» почтового сервиса. // Сб. «Обозрение прикладной и промышленной математики». М., ОПиПМ, 2012. т.19, вып.4. с. 581-582

8. Терентьев А.М. Выбор адекватных средств информационной защиты персонального компьютера в России // Журнал «Национальные интересы.

Приоритеты и безопасность» М., ООО «Издательский дом Финансы и кредит», N33(174)-2012, с. 37-42

9. Терентьев А.М. Корпоративный вариант реализации антивирусных пакетов Doctor Web в научных учреждениях: предпосылки. // Журнал «Национальные интересы. Приоритеты и безопасность» М., «Издательский дом Финансы и кредит», N17(206), 2013, с.41-48.

10. Терентьев А.М. Корпоративный вариант реализации антивирусных пакетов Doctor Web в научных учреждениях: реализация. // Журнал «Национальные интересы. Приоритеты и безопасность» М., «Издательский дом Финансы и кредит», N19(208), 2013, с.40-45.

11. Терентьев А.М. Корпоративный вариант реализации антивирусных пакетов Doctor Web в научных учреждениях: результаты. // Журнал «Национальные интересы. Приоритеты и безопасность» М., «Издательский дом Финансы и кредит», N20(209), 2013, с.41-46.

Problems of the Anti-Spam Protection: Targeted Spam

N. Lyapicheva
CEMI RAS
Russian Federation, Moscow, Nakhimovsky prospect, 47

Abstract

The purpose of the work described in the article is to solve the problem of protection against targeted spam, which uses the mechanisms of contextual advertising in mass mailings. Analysis of spam traffic was carried out on the basis of logging incoming mail headers, conducted by anti-virus/anti-spam gateway , and expert evaluation of own received mail traffic. As a result, a scheme of using e-mail clients is proposed, which allows to minimize the amount of spam received. The scheme can be recommended for users of standard mail clients in a partially updated hardware and software environment.

Keywords: mail service, anti-spam protection, information security, targeted advertising

Date of publication: 13.12.2018

Citation link:

Lyapicheva N. Problems of the Anti-Spam Protection: Targeted Spam // Vestnik CEMI RAS. 2018. Issue 2 [Electronic resource]. Access for registered users. URL: http://cemi.jes.su/s111111110000062-1-1 (circulation date: 17.12.2018). DOI: 10.18254/S0000062-1-1
