Динамический метод фильтрации интернет сайтов с агрессивным содержанием.

К.В. Моисеев.

Защита пользователей Интернета от агрессивного контента (порнография, пропаганда и распространение наркотиков, экстремизм) является чрезвычайно важной задачей. Важность этой задачи обусловлена несколькими факторами. Во-первых, это необходимость защитить детей от просмотра сайтов, содержание которых может нанести им вред. Во-вторых, это аспекты, связанные с информационной безопасностью. Хорошо известно, что Интернет ресурсы с порнографическим содержанием, ресурсы, позволяющие скачивать нелицензионные копии программного обеспечения или игры, зачастую, содержат немалое количество компьютерных вирусов. Таким образом, посещение даже одной единственной Интернет страницы с соответствующим содержанием, может привести к заражению всей локальной компьютерной сети, и как следствие, к потере или утечке критически важной (или конфиденциальной) информации и потребует существенных усилий для восстановления нормальной работоспособности находящихся в сети компьютеров. В-третьих, это желание ограничить нерациональное использование Интернета. Ведь не секрет, что, находясь на рабочем месте, сотрудники тратят заметное количество времени на просмотр не имеющих никакого отношения к работе Интернет ресурсов. Примерами таких ресурсов могут служить: развлекательные сайты, сайты, посвященные продаже автомобилей, сайты для поиска работы и, конечно же, социальные сети.

Проблемой блокировки Интернет сайтов с агрессивным содержанием занимается целый ряд российских и зарубежных компаний. Существует целый ряд обзоров, например [1], которые исследуют вопросы, связанные с блокировкой ресурсов Интернета как в России, так и во всем мире.

Основным технологическим направлением, которое используют практически все компании для ограничения доступа пользователей Интернет к сайтам с агрессивной тематикой, является URL фильтрация - так называемые «черные» и «белые» списки доменных адресов сайтов (URL). Эти списки создаются компаниями в процессе мониторинга Интернет ресурсов, а также мониторинга поведения пользователей в Интернете. Например, если в черный список включить доменное имя www.xxx.ru, то доступ к соответствующему сайту будет заблокирован. «Белый» список доменных имен сайтов, обычно, используется для предоставления доступа лишь к фиксированному

набору Интернет сайтов, которые включены в этот список, доступ ко всем остальным ресурсам Интернета запрещен.

Здесь следует сказать, что блокирование доступа к Интернет сайтам только по спискам имеет свои серьезные недостатки. Во-первых, изменения в Интернете происходят слишком быстро, и обновление списков не успевает за этими изменениями. Во-вторых, существуют стандартные процедуры, которые позволяют опытному пользователю такие списки обходить, например, используя IP адреса вместо доменного имени, или используя публично доступные сервисы обеспечения анонимности в Интернете. И, наконец, в-третьих, существует проблема предоставления или ограничения доступа к сайтам, часть содержимого которых не является агрессивным контентом, а другая часть содержит нежелательную информацию. При использовании алгоритмов, основанных только на «белых» и «черных» списках встает дилемма – блокировать ли такой сайт целиком или же игнорировать то, что часть информации на этом сайте относится к нежелательной категории, и разрешать доступ к этому сайту. Один из самых показательных примеров, иллюстрирующих эту ситуацию – работа с сайтами поисковых систем. При обычном поисковом запросе результаты поиска не будут содержать никакого нежелательного контента. В то же время, при попытке получить доступ, например, к сайтам для взрослых, а именно, при отправке соответствующего поискового запроса даже среди результатов поиска может оказаться достаточное количество нежелательной информации. Очевидно, что ограничивать по спискам доступ к сайтам поисковых машин в большинстве случаев не представляется возможным.

Технология фильтрации Интернет ресурсов на основе блокировки списков доменных имен, помимо ограничения доступа к сайтам с агрессивным содержанием, достаточно часто используется для Интернет цензуры. В настоящий момент более 25 стран практикуют Интернет-цензуру. Жители этих стран постоянно сталкиваются с проблемами доступа к разнообразным Интернет ресурсам, которые, в основном, посвящены вопросам защиты прав человека и освещению деятельности оппозиционных политических движений. Правозащитные организации во всем мире пытаются бороться различными способами с такого рода цензурой. Один из вариантов – чисто технологический – предложили специалисты Гражданской лаборатории университета Торонто (Канада) (www.citizenlab.org). Специалистами этой лаборатории был подготовлен «Справочник по обходу Интернет цензуры для всех» [2]. Все методы, предлагаемые в этом справочнике, прекрасно подходят не только для борьбы с Интернет цензурой, но и системами фильтрации агрессивного Интернет контента, особенно с теми из них, которые построены на идее URL-фильтрации.

Альтернативой URL фильтрации является динамическая фильтрация контента, когда содержимое, запрашиваемого Интернет сайта, анализируется в момент обращения к этому ресурсу. Под содержимым Интернет сайта понимается, вообще говоря, доменное имя, различные метаданные, текст, изображения и т.д. Загрузка страниц сайта в браузер блокируется, если содержимое определяется как нежелательное. Одним из самых важных элементов динамической контентной фильтрации являются алгоритмы, анализирующие текстовую составляющую. Задача этих алгоритмов – с максимально возможной скоростью и точностью принимать решение о том, соответствует ли текст анализируемой страницы определенной тематике. Очевидно, что в случае соответствия текста страницы одной из нежелательных тематик, доступ пользователей к этой странице блокируется.

Во всем мире достаточно активно проводятся исследования, направленные на создание систем динамической контентной фильтрации. Однако, как правило, весь накопленный мировым сообществом опыт достаточно плохо применим к русскоязычной части Интернета. Тем не менее, для того, чтобы получить представление о состоянии и развитии данного направления за рубежом имеет смысл ознакомиться с результатами проекта POESIA Project [3].

С точки зрения способа установки и использования системы контентной фильтрации также существует несколько возможных вариантов. Самый простой – установка системы контентной фильтрации в качестве локального приложения на каждый компьютер. Второй вариант – использование выделенного фильтрующего сервера внутри организации. Ну и, наконец, третий вариант – использование внешнего фильтрующего сервера, который установлен либо у провайдера услуг Интернет, либо у сервис-провайдера, который предоставляет услугу фильтрации Интернет контента.

Первый вариант наиболее часто используется для реализации систем родительского контроля. Фильтр в качестве локального приложения является сравнительно недорогим решением в случае использования на одном или нескольких компьютерах. Но в тоже время это решение наиболее уязвимо и наименее удобно. Действительно, в случаях, когда у пользователя есть администраторские права на доступ к компьютеру или физический доступ к системному блоку компьютера, обойти или сломать систему контентной фильтрации не является сложной задачей. Этот вариант является не очень удобным по сравнению с прочими потому, что процедура администрирования и управления настройками такого решения является достаточно трудоемкой.

Второй вариант наиболее удобен для крупных организаций. Использование выделенного фильтрующего сервера характеризуется наибольшими возможностями по настройке и администрированию системы фильтрации. Администрирование системы

достаточно удобно, но требует высокой квалификации системного администратора по настройке сервера фильтрации и Интернет-трафика внутри организации. Из всех представленных вариантов это самое дорогое решение. Оно подходит для организаций с большим количеством Интернет-трафика, предъявляющих повышенные требования к сохранению конфиденциальности в процессе посещения Интернет ресурсов.

Внешний фильтрующий сервер является некоторым компромиссом между первым и вторым вариантами. При сравнительно невысокой стоимости он обеспечивает достаточно простой и удобный процесс администрирования системы и высокую надежность работы системы фильтрации. Во многих случаях этот вариант предполагает использование системы фильтрации Интернет контента без каких бы то ни было настроек на персональных компьютерах пользователей. Например, в случаях установки фильтрующего сервера у Интернет провайдера, все необходимые настройки по перенаправлению Интернет трафика через систему фильтрации контента выполняются на стороне Интернет провайдера.

Компания «Технологии управляемого хаоса» предлагает программный продукт CCT Internet Content Filter, который комбинирует внутри себя оба типа фильтрации контента – как динамическую фильтрацию, которая основана на уникальном эвристическом алгоритме анализа тематики текстовой информации Интернет страниц, так и фильтрацию на основе управления белыми и черными списками доменных имен (URL фильтрацию). Созданный в компании эвристический алгоритм анализа тематики текста основан на алгоритме поиска похожей информации, который реализован в рамках метода корреляционной индексации. Основы этого метода изложены в [4].

Необходимость использования обоих типов фильтрации обусловлена тем, что каким бы интеллектуальным не был эвристический алгоритм анализа тематики страницы, как и любой эвристический алгоритм, он будет определять тематику текста с определенной вероятностью. Существует два типа ошибок при работе эвристического алгоритма: ошибочный пропуск Интернет страниц с агрессивным содержанием и блокировка Интернет страниц, ошибочно отнесенных к одной из запрещенных тематик. Обычно суммарная ошибка алгоритмов динамической фильтрации колеблется в пределах 5% – 10%, что приводит к неправильным пропускам или блокировкам каждой 20-ой (10-ой) Интернет страницы (при указанных значениях суммарной ошибки алгоритмов, когда вероятность пропуска одной страницы еще достаточно велика, так называемый «серфинг» по сайтам с нежелательным содержанием уже невозможен). Следует обратить внимание, что процедура перенастройки алгоритма динамической фильтрации для исправления ошибок является достаточно продолжительной и выполняемой всегда силами

разработчика программного обеспечения. Таким образом, именно использование технологии «черных» и «белых» списков сайтов в CCT Internet Content Filter дает возможность оперативного исправления ошибок при работе системы фильтрации содержимого Интернет страниц.

Программный продукт может быть использован как для установки на выделенный фильтрующий сервер внутри организации, так и при использовании внешнего фильтрующего сервера. Программный продукт реализован как надстройка над прокси сервером. Такая архитектура делает систему фильтрации легко администрируемой и минимально зависящей от сетевой инфраструктуры конкретной организации. Схема использования CCT Internet Content Filter представлена на Рис. 1

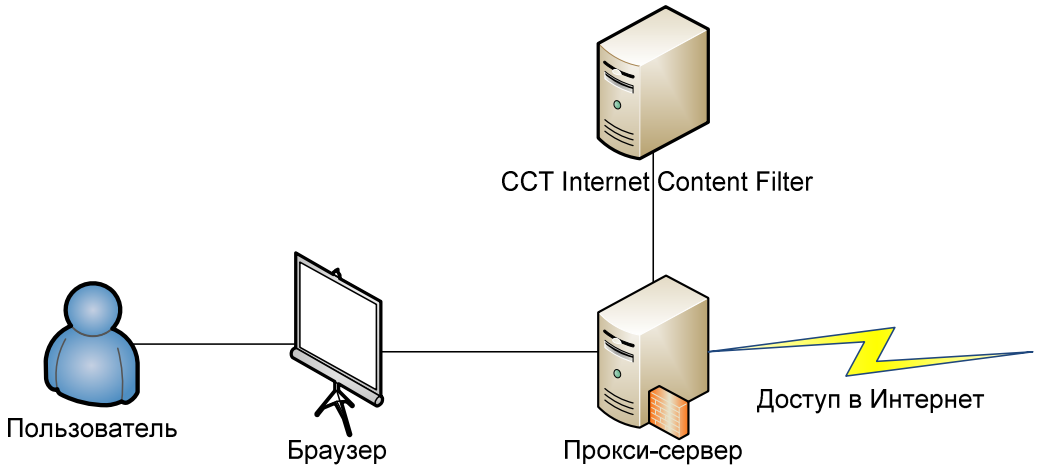

Рис. 1. Схема использования CCT Internet Content Filter.

Компания «Технологии управляемого хаоса» готова предоставить для всех читателей журнала тестовый доступ к фильтрующему серверу нашей компании. Для этого, всего лишь, необходимо в настройках веб браузера указать настройки для доступа к Интернету с использованием прокси-сервера:

- ✓ Адрес прокси-сервера: www.controlchaostech.com
- ✓ Порт: 8080
- ✓ Имя пользователя: modern_lib
- ✓ Пароль: htsyprsd

Эти имя пользователя и пароль будут действительны до 31 октября 2009 года.

На сегодняшний день программный продукт CCT Internet Content Filter для анализа содержимого Интернет сайтов использует лишь текстовое содержимое страницы. Дальнейшее развитие программного продукта предполагает интеграцию в систему контентной фильтрации решения, которое способно анализировать не только текстовую,

но и графическую информацию. Последнее важно для фильтрации Интернет ресурсов, содержащих агрессивный контент. Ведь не секрет, что в большом числе случаев такие сайты не содержат или содержат минимальное количество текстовой информации, а вся информация (в том числе текстовая) представлена изображениями. Для фильтрации графической информации будут использованы разрабатываемые компанией технологии по распознаванию образов на изображениях.

Список литературы:

1. Торчинский Ф.И., Регулирование Интернет и фильтрация «языка вражды», V Всероссийская объединенная конференция "Технологии информационного общества - Интернет и современное общество" (IST/IMS-2002)
2. Справочник по обходу Интернет цензуры для всех. Проект Civisec, The Citizen Lab, The University of Toronto. Сентябрь, 2007. http://www.civisec.org/sites/all/themes/civisec/guides/everyone's-guide-russian.pdf.
3. Website of POESIA Project. http://www.poesia-filter.org/.
4. Y.D.Kalafati, K.V.Moiseyev , S.O.Starkov, S.A.Shushkova, Technology of Storage and Processing of Electronic Documents with Intellectual Search Properties. International Journal Information Theories and Applications (IJ ITA) V.15, 184, 2008