

УДК 681.3.06

А.С.Кадыров (асп., каф. ПМ), В.Е.Клавдиев, к.т.н., доц.

СИСТЕМА ИНТЕРАКТИВНОЙ ВЕБ-СТАТИСТИКИ

Разработка системы интерактивной веб-статистики обусловлена отсутствием на рынке систем анализа веб-трафика качественных и недорогих пакетов, позволяющих получать важные для маркетингового анализа результаты (такие, например, как отслеживания пользовательских потоков (маршрутов) через веб-сайт, расшифровка запросов поисковых систем, в случае прихода пользователей с них, определение посещений сайта роботами и другие). Кроме того, подобные системы всегда ориентированы на анализ веб-трафика как такового, тогда как в разработанной системе упор делается на конкретных посетителей (отслеживаются страны, города, конкретные IP-адреса, что зачастую бывает важно при определении и мониторинге целевой аудитории).

Целью данной работы являлась реализация такого программного продукта.

Опишем возможности системы, многие из которых не встречались автору ранее в различных программных продуктах для сбора и анализа веб-статистики.

В системе предусмотрена возможность он-лайн просмотра адресов текущих пользователей и страниц, на которых они находятся в настоящий момент.

Возможность просмотра статистики в двух режимах - графическом и потоковом.

Под потоком в системе понимается пользовательская сессия (последовательность перемещений по сайту, начиная от точки входа и заканчивая точкой выхода) посещения сайта; момент ухода пользователя определяется заданным тайм-аутом.

Графический режим включает в себя следующие отчеты:

1. Отчет по пользовательским доменам и IP-адресам,
2. Отчет по конкретному пользователю с возможностью поиска пользователя по IP-адресу,
3. Отчет по ссылкам (Referer report) с разбором запросов поисковых систем,
4. Отчет по пользовательским агентам, включая распознавания приходов спайдеров поисковых систем и редакторов с каталогов Yahoo, DMOZ и других,
5. Отчет по хитам на локальные страницы сайта.

Ко всем отчетам применяется система графических фильтров, позволяющая определять промежуток времени, для которого будут отображаться результаты.

Потоковый режим включает в себя следующие отчеты:

1. Отображение пользовательских потоков, отсортированное по времени посещений,
2. Отображение пользовательских потоков, отсортированное по пользователям,
3. Возможность выбора пользователя и поиска пользователя по IP-адресу и показа его потоков,
4. Возможность выбора локальной страницы сайта и показ потоков сквозь нее.

Ко всем потоковым отчетам также применяется система фильтров, позволяющая определять промежуток времени, для которого будут отображаться результаты.

Кроме того, в системе имеются:

- подсистема оповещения о приходе пользователя, работа которой заключается в уведомлении о приходе конкретного пользователя (Авторизованная персона имеет собственную базу данных IP-адресов, для обновления которой реализован веб-интерфейс).

- E-mail автоответчик. Посылая письмо на адрес автоответчика, с IP (или его частью) в поле письма Subject, через некоторое время авторизованная персона получает ответ с отчетом о посещениях пользователей с подобными адресами.

Система реализована в виде набора CGI-скриптов и CGI-интерфейсов к БД MySQL.

Язык реализации Perl 5. Запись данных в БД реализуется путем вставки в каждую страницу сайта SSI-включения для исполнения программ.

Выводы. Предложенная система интерактивной веб-статистики является инструментом, ориентированным на предоставление пользователю аналитических онлайн и офф-лайн отчетов о веб-трафике, направленных прежде всего на анализ трафика конкретных посетителей, и решение нетривиальных задач, таких как, например, анализ запросов поисковых систем и отслеживание пользовательских потоков.

На основе анализа этих данных можно делать выводы о самых разных вещах, начиная от выявления ошибочных ссылок и заканчивая оценкой предпринимаемых усилий по развитию и продвижению узла.

Система написана на языке Perl, использует базу данных MySQL и механизм включений на стороне сервера (SSI), что делает ее независимой от типа веб-сервера и легко переносимой как на UNIX-подобные платформы, так и на Windows.