

Глава 2

СЕРВЕРЫ СТАТИСТИКИ И АВТОНОМНЫЕ ПРОГРАММЫ- АНАЛИЗАТОРЫ

Все программные средства для анализа посещаемости можно разделить на две категории — счетчики посещаемости и программы-анализаторы журнальных файлов сайта, на профессиональном сленге — анализаторы логов. Весьма условно названные программные средства можно сравнить с почтовыми ящиками на сайтах, предоставляющих услуги электронной почты (счетчики), и почтовым клиентом на собственном компьютере (анализатор логов). У каждого из методов анализа есть свои достоинства и недостатки. Но, прежде чем говорить о них, нужно сказать несколько слов о том, какие основные параметры посещаемости они оценивают, какие данные используют для анализа и откуда их получают.

ЖУРНАЛЬНЫЕ ФАЙЛЫ — ВСЕ ПОД КОНТРОЛЕМ

Как известно, в основе функционирования Интернета (WWW, “Всемирной паутины”) лежит технология “клиент-сервер”. Для того чтобы пользователь мог посмотреть веб-страницу, на его компьютере должна быть установлена программа-клиент (браузер), на компьютере, где размещен сайт, соответственно, программа-сервер. Просмотр страницы возможен только при слаженной работе клиента и сервера. Для этого они должны обмениваться информацией. Клиент должен сообщить серверу, что именно интересует пользователя. Сервер должен передать файлы выбранной пользователем¹ страницы. Для этого он должен как минимум знать интернет-адрес компьютера пользователя (его часто называют *IP-адрес*, потому что передача информации осуществляется в соответствии с протоколом IP). Естественно, он должен знать и адрес страницы, запрошенной пользователем (URL — универсальный указатель информационного ресурса; на сленге — *урл*). Различие в наименовани-

¹ Поскольку все посетители вашего сайта— это одновременно и пользователи компьютеров, далее эти два термина будут использоваться как синонимы.

ях интернет-адресов вызвано тем, что в общем случае компьютер пользователя, в отличие от сайта, не предоставляет свои ресурсы для других пользователей). Как только эта информация поступает на сервер, она немедленно фиксируется в *журнальном файле (лог-файле)*. По IP-адресу посетителя зачастую можно определить, в какой стране или даже в каком городе расположен его компьютер. Так у владельца сайта появляется важная информация об одном из посетителей.

Помимо IP-адреса пользователя и URL заинтересовавшей его страницы браузеры часто передают другую любопытную для владельца сайта информацию. Например, это может быть URL страницы, которую пользователь просматривал *перед* тем, как перейти на ваш сайт, тип браузера, разрешение экрана монитора пользователя т.д.

Если на странице имеются фотография, рисунок, текст, содержащийся в отдельной области страницы, называемой *фреймом*, то каждый такой элемент страницы запрашивается браузером посетителя отдельно. Соответственно, в журнальном файле делается отдельная запись.

Некоторые веб-страницы напоминают “сборную солянку”. Текст может находиться на одном сервере, изображение на другом, музыкальный клип на третьем. Соответственно, в лог-файлах каждого из серверов появится запись о том, что тот или иной элемент был востребован. Так что каждый шаг пользователя в Сети контролируется тысячами невидимых компьютерных глаз. Но для того, чтобы извлечь пользу из миллиардов битов информации, содержащейся в журнальных файлах Интернета, ее нужно предварительно обработать. Именно этим и занимается автономная программа-анализатор — извлекает полезную информацию из журнальных файлов сайта.

ВНЕШНИЕ СЧЕТЧИКИ (СЕРВЕРЫ СТАТИСТИКИ)

На страницах многих сайтов можно увидеть небольшие логотипы, иногда с цифрами (рис. 2.1). Это и есть зримое отображение *внешних* (иногда их называют *сторонними*) счетчиков посещений. Механизмы их работы различны, но в общем это выглядит примерно так.



Рис. 2.1. Логотипы внешних счетчиков

Когда сайт получает от какого-либо пользователя запрос на просмотр страницы, на которой установлен счетчик, он обращается к специализированному серверу (его называют *сервер статистики*, иногда — *статистический сервер*, логотип этого сервера и отображается на страницах сайта). Сервер, получив запрос, выполняет особую *серверную* программу-сценарий (на сленге *программа-*

сценарий — это *скрипт*). В результате в базе данных сервера статистики сохраняется информация о посещении пользователем страницы и динамически формируется изображение (логотип сервера статистики). HTML-код изображения (иногда с числами, соответствующими количеству просмотров страницы) отправляется на компьютер пользователя, сделавшего запрос. Таким образом, “картинка” счетчика, которую пользователь видит на вашей странице, на самом деле формируется не вашим сайтом, а сервером статистики.

Для того чтобы сервер статистики мог получить подробные данные о посетителе вашего сайта, вместе с HTML-кодом изображения счетчика на его компьютер загружается *клиентская* программа-сценарий. Она собирает необходимые для работы счетчика сведения и передает их на сервер статистики. Эта программа не способна ни считывать файлы с жесткого диска компьютера пользователя, ни записывать их на диск, опасаться ее не следует.

Но для того, чтобы отследить повторные посещения этим же пользователем этой же страницы через несколько дней, программы-сценария недостаточно. Поэтому многие счетчики статистики записывают на жесткий диск компьютера пользователя так называемые *файлы cookie*. В буквальном переводе *cookie* — это “печенье”. Объем такого файла не может превышать 4 Кбайт, зачастую это — всего лишь несколько строк кода. Почему именно так назвали файлы, нам не известно. В литературе вместо непонятного *cookie* иногда используют термин *ключик*, на сленге такие файлы называют просто *куки*.

Файлы *cookie* могут быть временными и постоянными. Первые действуют только до тех пор, пока работает браузер. Вторые хранятся на компьютере пользователя заранее заданное время. При повторном обращении пользователя к странице, оставившей на его компьютере файл *cookie*, счетчик может определить, что этот посетитель уже был на сайте, и включить данный факт в соответствующий отчет. Файлы *cookie* используются также для запоминания данных, используемых при аутентификации пользователя, благодаря им ему не приходится вводить логин и пароль при каждом обращении, допустим, к почтовому веб-серверу.

Полученная с помощью вышеперечисленных средств информация о посетителе заносится в базу данных сервера статистики. При обращении к ней владельца файла необходимые данные извлекаются, обрабатываются и предоставляются в виде отчетов. По желанию владельца эти отчеты могут быть доступны всем либо только тем, кто знает соответствующие логин и пароль.

ПРОБЛЕМА ВЫБОРА

Исходные данные, получаемые автономными анализаторами логов и внешними счетчиками посетителей, собираются различными способами. Поэтому возможности этих программ существенно отличаются. Что лучше? Рассмотрим основные преимущества и недостатки каждого из методов анализа посещаемости сайтов.

АНАЛИЗАТОРЫ ЛОГОВ: ПРЕИМУЩЕСТВА И НЕДОСТАТКИ

Считается, что самую полную статистику дают анализаторы логов. Сервер, на котором расположен ваш сайт, фиксирует все обращения к нему.

Автономная программа для сбора статистики может стоить несколько сотен долл., хотя встречаются и бесплатные варианты. Некоторые предоставляются в аренду за относительно небольшую плату. Программа может быть установлена на сервере, где размещен сайт. Настроить ее сложнее, чем установить счетчики; могут понадобиться консультации специалиста. Второй вариант — установить программу на своем компьютере и регулярно копировать на него журнальные файлы. Срок хранения журнальных файлов в таком случае определяет владелец сайта, а не хозяин сервера статистики. Достоверность результатов достаточно высокая, поскольку данные не теряются из-за перегрузки сервера статистики. Страницы вашего сайта на мониторах посетителей отображаются быстрее, поскольку браузеру не приходится загружать логотипы счетчиков с серверов статистики.

К недостаткам можно отнести то, что автономные программы не позволяют сравнивать посещаемость с сайтами-конкурентами — у них нет для этого данных. Кроме того, с помощью анализатора логов сложнее отслеживать географию посетителей — для этого нужна дополнительная информация по привязке IP-адресов к географическому расположению, т.е. постоянно пополняемая база данных.

Но анализаторы логов не используют файлы cookie. Это означает, что они не способны идентифицировать посетителей, заходящих на сайт из корпоративных сетей, из-за прокси-серверов и в некоторых других случаях. И уж тем более они не могут распознавать посетителей, повторно заходящих на сайт, что крайне важно при проведении рекламных компаний.

СЧЕТЧИКИ: ПРЕИМУЩЕСТВА И НЕДОСТАТКИ

Многие счетчики (серверы статистики) предоставляют свои услуги бесплатно. Иногда помимо бесплатного счетчика предлагается платный с расширенными возможностями. Процедура установки счетчика несложна — на каждой странице сайта нужно разместить фрагмент кода; сделать это может рядовой пользователь. Результаты обработки предоставляются быстро и в достаточно наглядной форме (например, в виде графиков). Счетчики позволяют сравнивать посещаемость вашего сайта с посещаемостью конкурента (если он тоже пользуется услугами такого же счетчика). Как правило, счетчики предоставляют подробные данные относительно географии посетителей (т.е. об их распределении по странам и регионам). За счет использования файлов cookie счетчики позволяют достаточно точно идентифицировать пользователя и определить, «старый» он или впервые посетил сайт. Могут также определить язык, на который настроен браузер пользователя, часовой пояс, параметры экрана. Профессионалы ценят счетчики за то, что они позволяют определить, разрешено ли на компьютере пользователя выполнение программ Java, JavaScript, применяется ли прокси-сервер и т.д.

Основные недостатки счетчиков обусловлены тем, что информация хранится и обрабатывается не на вашем сайте, а на сервере статистики. Как правило, исходные данные хранятся не очень долго, максимум один год. Достоверность данных невелика из-за того, что они довольно часто теряются. Происходит это в основном по двум причинам: из-за перегрузки сервера статистики и потому, что пользователь покидает страницу или сайт еще до того, как счетчик успеет зафиксировать его присутствие.

Для продвижения сайта бывает важно отслеживать посещаемость его роботами поисковых систем. Счетчики этого либо не делают вообще, либо делают плохо.

Наличие счетчика замедляет доступ посетителей к сайту. Чем больше на странице счетчиков, тем медленнее она открывается в браузере пользователя.

Ни один счетчик не позволяет регистрировать загрузку с сайта мультимедийных файлов, флеш-объектов, изображений и т.п. Специалистам известны и другие недостатки счетчиков.

ДВА В ОДНОМ

Таким образом, счетчики и автономные программы дополняют друг друга. Поэтому некоторые владельцы сайтов используют одновременно и счетчик, и анализатор логов.

Такое решение вряд ли можно назвать совершенным. Ведь на анализ результатов придется тратить значительно больше времени. А если еще и сопоставлять их, да выявлять причины несовпадения, времени понадобится еще больше. Поэтому недавно на рынке появились программы, сочетающие возможности счетчика посещений и анализатора логов. То есть, на сайте устанавливаются одновременно и собственный (*внутренний*) счетчик, и программа-анализатор. При этом полученные с их помощью данные обрабатываются одновременно и выдаются в виде единого отчета (точнее, многих отчетов; их число может достигать до 400). Стоимость подобной программы российского производства может составлять 25 тыс. руб., последующие версии — бесплатно. Но такие программы лишь приступили к завоеванию рынка.

Один из важных моментов при использовании такой программы — степень достоверности данных для потенциального рекламодателя. Он прекрасно знает, что внешние счетчики, как правило, занижают число посетителей. Но знает он и другое: счетчик внутренний может быть “подкручен” нечестным владельцем сайта. Каким образом будет решена эта проблема, какое место под солнцем завоюют комбинированные анализаторы-счетчики, покажет ближайшее будущее. Далее мы в книге подробно опишем один из наиболее популярных бесплатных счетчиков, а также один из бесплатных анализаторов логов. Мы полагаем, что, прежде чем приобретать платную комбинированную программу, способную выдать до 400 отчетов, владельцу сайта целесообразно потренироваться на программах бесплатных, выдающих 30–40 отчетов. И, скорее всего, одна из них его вполне устроит.

ПОПУЛЯРНЫЕ ВНЕШНИЕ СЧЕТЧИКИ (КРАТКИЙ ОБЗОР)

RAMBLER'S TOP 100

Первым сервером статистики Рунета принято считать Rambler's TOP 100 (<http://top100.rambler.ru/top100>), появившийся еще в 1997 году. Он и до сих пор весьма популярен, но уже не как внешний счетчик, а как рейтинг, который применяют в основном для продвижения сайтов. Именно как рейтинг он известен многим пользователям Интернета. Его используют для поиска наиболее популярных сайтов на ту или иную тему. Попасть на первую страницу Rambler's TOP 100 мечтают многие “раскрутки” сайтов, хотя это — скорее вопрос престижа.

У огромной популярности Rambler's TOP 100 есть и обратная сторона. Этот счетчик ежедневно загружается более 100 миллионов раз. Интенсивность загрузки неравномерна; в часы пик серверы Rambler's TOP 100 могут не успевать обрабатывать все запросы на загрузку страниц Рунета, из-за чего данные статистики будут занижены. Но самый главный недостаток Rambler's TOP 100 как средства анализа посещаемости сайтов — малое количество отчетов. С помощью этого счетчика вы не узнаете ни географию пользователей, ни точки входа-выхода на ваш сайт, ни многое другое.

Кроме того, данный счетчик может использоваться лишь для основного сервера, но не для его зеркал. В общем и целом же получается, что хромает как количество, так и качество полученных данных, а говорить об их дальнейшем серьезном использовании и вовсе не стоит.

РЕЙТИНГ@MAIL.RU

Этот рейтинг сменил несколько названий. В конце прошлого века он назывался TopList, потом поменял фамилию на TopMail, сейчас на главной странице (<http://top.mail.ru>) значится Рейтинг@mail.ru.

Счетчик этого рейтинга дает несколько больше статистических данных о посещаемости сайта, чем Rambler's TOP 100. По состоянию на начало 2007 года количество отчетов составляло 11. В частности, описываемый счетчик позволял узнать, из каких стран чаще “приходят” посетители, какие поисковые машины и каталоги ими использовались для поиска нужных страниц на вашем сайте, как изменялось количество посетителей в течение месяца, какие страницы сайта сколько раз посещались в течение того или иного дня. При этом можно полученные данные отфильтровывать: показывать сведения *только* по определенным страницам или *все, за исключением* определенных страниц.

Но для полноценного анализа посещаемости этих данных недостаточно.

СЧЕТЧИК SPYLOG

Счетчик SpyLOG (www.spylog.ru) стал первым на просторах Рунета, который начал позиционировать себя не как рейтинг, заодно дающий немножко статистических данных. Его главной целью стала именно статистика. По состоянию на начало

2007 года счетчик обслуживал свыше полумиллиона сайтов. Он предлагает как платные услуги для бизнеса, так и бесплатные для некоммерческих организаций.

Среди платных сервисов, помимо счетчика с расширенными возможностями, предлагаются мониторинг потребительской активности интересующей вас отрасли, семинары для сотрудников компаний, продвигающих бизнес через Интернет, и заказные исследования различных рынков.

В начале 2007 года бесплатный пакет для получения статистических данных назывался SpyLOG Tracker. У него были две разновидности — “Кроха” для всех сайтов с минимальным набором услуг (всего 14 отчетов) и “Солидарность”, предоставляющий для некоммерческих сайтов почти такие же возможности, как и соответствующий платный пакет (он назывался SpyLOG Tracker Лайт и предоставлял до 50 отчетов). Решение о причислении сайта к коммерческим или некоммерческим принимал модератор системы SpyLOG.

Что касается самого дешевого пакета SpyLOG Tracker Лайт, то ежемесячная плата за его услуги составляла 450 руб., плюс от 15 до 0,3 руб. за каждую тысячу загрузок в сутки. Чем больше загрузок, тем меньшая оплата за каждую тысячу (за 500 и менее загрузок в сутки — 12 руб., за 100 000 и менее — 3 руб.). Помимо пакета “Лайт” предлагались более дорогие пакеты “Бизнес” и “Премиум” (абонентская плата за последний 3000 руб./мес., но 20 000 загрузок за день входили в абонентскую плату).

Теперь, когда вы примерно представляете стоимость платных пакетов SpyLOG, поговорим немного о бесплатных.

Пакет “Кроха” предоставляет почти такой же набор отчетов, как и Рейтинг@mail.ru. К полезным дополнениям можно отнести то, что этот счетчик определяет не только типы браузеров пользователей, но и типы операционных систем, параметры дисплеев их компьютеров и другие данные, которые могут заинтересовать веб-мастера сайта или его владельца.

Пакет “Солидарность” — это уже вполне серьезное решение, способное удовлетворить многих владельцев сайтов. Например, с его помощью можно определять распределение аудитории сайта не только по странам, но и по регионам России и Украины. Самые важные из регистрируемых счетчиком данных — точки входа на сайт и точки выхода, пути посетителя по сайту, ссылки на вашем сайте, которые на самом деле никуда не ведут, и многое другое. Мы не будем далее на этом останавливаться, поскольку спешим описать следующий популярный счетчик — HotLog.

СЧЕТЧИК HOTLOG

Счетчик SpyLOG, описанный в предыдущем разделе, первым начал предоставлять услуги статистики за деньги. Его конкурент, счетчик HotLog (<http://hotlog.ru>), появившийся несколько позже, поначалу предоставлял практически такие же статистические данные бесплатно. И быстро вышел на уровень примерно 0,5 млн. зарегистрированных сайтов. Однако время шло, и нравы создателей HotLog менялись вместе с ним. Сейчас HotLog, помимо бесплатного пакета услуг, предоставляет два платных. Сервис “Что? Где? Когда?” позволяет получить “портрет” каждого посети-

теля сайта, т.е. узнать, с какой страницы он перешел на сайт, из какого города и даже из какой организации, параметры браузера и операционной системы его компьютера и т.п. Стоила эта услуга в начале 2007 года 150 руб./мес.

Пакет услуг HotLog GOLD, помимо сервиса “Кто? Где? Когда?”, позволяет получать подробные отчеты, в том числе электронной почтой, осуществлять резервное копирование данных и т.д., стоит все это 999 руб./мес.

Создатели счетчика HotLog утверждают, что он — “самый объемный в Интернете бесплатный сервис интернет-статистики”. Действительно, в начале 2007 года этот счетчик бесплатно предоставлял 21 отчет, в том числе такие важные, как ссылающиеся страницы, пути по сайту и глубина его просмотра, точки входа и выхода. Однако, судя по отзывам пользователей, размещенным в Сети, этот счетчик довольно часто не справляется с потоком запросов. Кроме того, нарекания вызывали рекламные компании, проводившиеся создателями счетчика HotLog для его популяризации. Поэтому в последующих главах мы будем описывать не HotLog, а ставший за последние месяцы более популярным счетчик LiveInternet. Но вначале упомянем украинские и зарубежные счетчики.

УКРАИНСКИЕ СЧЕТЧИКИ

Все рассмотренные счетчики не используют систему опознавания “свой-чужой”. Они предоставляют статистику как для российских сайтов, так и для украинских, белорусских и прочих, главное — чтобы они были русскоязычными.

Украинские счетчики проявляют национальную сознательность. Они озабочены в основном украинским сектором Интернета (по аналогии с Рунетом — Уанетом). К наиболее популярным на момент написания книги относились два — TopPing (<http://mytop-in.net>) и счетчик с потенциально неудобнопроизносимым названием bigmir)net (<http://top.bigmir.net>). Впрочем, все его называют просто “Бигмир”. По количеству зарегистрированных страниц они на один-полтора порядка уступают лидерам Рунета и позиционируют себя, в общем-то, как рейтинги, т.е. никаких гарантий относительно полноты статистических данных они не дают. Если на одной и той же странице установлены, допустим, счетчик bigmir)net и SpyLog, последний может показать количество посетителей на 25% большее, чем bigmir)net. Администраторы последнего оправдывали это так: “мы не можем гарантировать стабильную работу каналов в Штатах (серверы, обчитывающие статистику Бигмира, расположены за океаном), и... мы не гарантировали стабильную работу бесплатного сервиса (счетчика), поэтому если вы считаете невозможным пользоваться услугами сервиса, то снимите счетчик с сайта...” А время от времени в “Новостях” счетчика появляются торжествующие объявления “На данный момент функционирование системы статистики bigmir)net полностью восстановлено!”.

ЗАРУБЕЖНЫЕ СЧЕТЧИКИ

Внешние счетчики на платной основе предоставляют компании WebSideStory, SPSS, CoreMetrics и др. Несомненным лидером здесь является счетчик Google Analytics.

Большинству зарубежных счетчиков присущи следующие серьезные недостатки:

- они плохо (или вообще не) поддерживают кодировки кириллицы;
- они недостаточно осведомлены о российских поисковых системах, каталогах и рейтингах;
- они не поддерживают базы данных, позволяющие определить географическую расположенность пользователей из стран СНГ.

Сервис статистики Google Analytics был создан на основе разработок компании Urchin Web Analytics. После приобретения последней гигантом Google стоимость использования ее программы снизилась с \$400 до \$199 в месяц. В число клиентов сервиса входят десятки компаний из числа крупнейших (участниц списка Fortune 500).

Ситуация на зарубежном рынке счетчиков посещений изменилась после того, как Google, продолжая свои традиции, сделал счетчик Google Analytics совершенно бесплатным. Поскольку произошло данное событие (по отношению к моменту написания книги) недавно, пока есть лишь предположения относительно того, к чему это может привести. Очевидно, другим игрокам на рынке придется или последовать примеру Google, или существенно расширить возможности своих программ.

Вообще-то, за 2005-2006 годы Google выпустил множество программ и продуктов для пользователей, большинство из них бесплатно. Почти все они используются Google для сбора данных о поведении пользователей на тех или иных сайтах. На основе этой информации создаются алгоритмы учета популярности сайтов для определения их рейтинга в результатах поиска.

Некоторые специалисты лучшей из *новых* программ гиганта считают именно Google Analytics. С одной стороны, этот счетчик предоставляет в распоряжение веб-мастеров мощный инструмент анализа посещаемости сайтов с широкими возможностями. С другой, предоставляет Google множество данных о вашем сайте. После установки на страницы сайта счетчиков Google может отслеживать, как пользователи находят и каким образом используют ваш сайт. Более того, Google способен отслеживать так называемую *конверсию*, т.е. определять, какая доля посетителей сайта превращается в постоянных посетителей (например, покупателей).

Мы так подробно говорим об этой программе потому, что в самом конце 2006 года появились сообщения о том, что Google решил выйти на российский рынок Интернет-рекламы. А огромная популярность на Западе счетчика Google Analytics вызвана в том числе и тем, что он тесно интегрирован со службой рекламы AdWords. Надо полагать, что в скором времени этот пока еще малопопулярный в России счетчик преодолет все присущие зарубежным счетчикам недостатки и выйдет на лидирующие позиции.

КАКОЙ СЧЕТЧИК ПОСЕЩЕНИЙ ВЫБРАТЬ

Выбрать счетчик посещений, в общем-то, не трудно. В первую очередь потому, что их можно установить несколько. Счетчики не мешают работе друг друга, но следует помнить о следующем. Экспериментально установлено, что каждый следую-

щий счетчик (код которого находится в коде веб-страницы ниже, чем код предыдущего счетчика) учитывает примерно на 5% меньше посещений, чем предыдущий. Вызвано это тем, что пользователи, не дождавшись полной загрузки страницы, переходят к просмотру следующей. Поэтому не следует устанавливать слишком много счетчиков. Они замедляют загрузку страниц и потребляют дополнительный трафик. Руководствоваться следует принципом разумной достаточности. Например, для сбора достаточно подробных статистических данных можно воспользоваться бесплатным счетчиком LiveInternet, для продвижения сайта и оценки его популярности — рейтингом Rambler's TOP 100. Если вы владелец крупной компании, SEO-специалист может потребовать установки еще и платного счетчика. Какого именно — ему виднее.

ПОПУЛЯРНЫЕ АНАЛИЗАТОРЫ ЛОГОВ

Подобно тому как помимо распространенных внешних счетчиков посещений существуют счетчики внутренние, анализаторы логов также можно разделить на две категории. Многие из них предназначены для установки на том же сервере, на котором расположен ваш сайт. Но существуют и такие, которые могут быть установлены на вашем собственном компьютере, иногда их называют *локальные* анализаторы логов.

Анализаторы логов, размещаемые на сервере, могут быть очень эффективными средствами анализа посещаемости для сайтов, которые размещаются на выделенном сервере. Такие серверы обычно создают крупные и средние компании. Журнальные файлы (логи) могут храниться на таком сервере длительное время, период хранения определяет владелец сервера, являющийся одновременно и владельцем сайта.

Но гораздо чаще сайты размещаются не на выделенном сервере, а на таком, который предоставляет платный или бесплатный хостинг (размещение) десяткам, сотням и тысячам владельцев. Владельцы такого сервера (на сленге — *хостеры*) хранят статистику, собираемую сервером, недолго, обычно от 1 недели до 1 месяца. Для владельцев сайтов, пользующихся хостингом, более приемлемым может оказаться локальный анализатор логов, размещаемый на собственном компьютере. На него можно периодически копировать данные из лог-файлов и хранить их на собственном компьютере столько, сколько понадобится. Но синхронизация новых и уже скопированных ранее данных — задача не простая.

Предложений на рынке программ-анализаторов логов очень много. Некоторые из них бесплатные; однако мощные профессиональные программы могут стоить очень дорого. В рамках одного из исследований (http://blog.eugeniavlasova.com/archives/publikatsii/tochnost_vezhlivost_statistiki_chast_ii.html) был проведен опрос на тему “Какими лог-анализаторами Вы предпочитаете пользоваться?”. Пришло двадцать четыре ответа, причем ни один из них не повторился. Автор данного исследования выделил из общего ряда три программы, представленные в нижеследующем перечне первыми.

THE WEBALIZER

Эта программа бесплатна. Получить ее можно по адресу <http://www.mrunix.net/webalizer>. Она не предъявляет высоких требований к быстродействию компьютера. Результаты анализа представляются в виде удобных таблиц и простых графиков. Принцип “ничего лишнего” выдержан разработчиком этой программы и в дизайне, и в предоставляемой информации. Простота и быстродействие программы обусловлены в первую очередь тем, что она не комбинирует полученные из логов данные, не сопоставляет одни данные с другими. Детальный анализ посещаемости возложен на владельца сайта. Тем не менее с помощью Webalizer можно получить данные о посещаемости сайта, узнать, какие его страницы являются популярными, какие из них дают наибольший трафик, с каких страниц приходят на сайт посетители, по каким поисковым фразам, какими браузерами пользуются и где живут.

Поскольку программа бесплатна, многие владельцы серверов, предоставляющие хостинг, устанавливают ее у себя и предлагают владельцам сайтов пользоваться ее услугами.

WEBLOG EXPERT

Эта программа (<http://www.weblogexpert.com>) доступна в трех вариантах.

- WebLog Expert Professional стоимостью \$125.
- WebLog Expert Standard стоимостью \$75.
- WebLog Expert Lite — бесплатно.

Очевидно, возможности этих версий несколько различаются. Сравнительные параметры приведены на странице <http://weblogexpert.com/lite.htm>. В частности, бесплатная версия позволяет получить 16 таблиц и 16 графиков, платные версии — по 37 таблиц, 48 графиков и 20 сравнительных графиков.

Для установки WebLog Expert не нужен сервер, как этого часто требуют программы-анализаторы логов. Все, что требуется для работы, — это получить с сервера, на котором установлен сайт, архив журнальных записей (логов) и подождать некоторое время, пока программа создаст отчет в виде HTML- или PDF-файла (последний вариант выбирается, когда отчет нужно напечатать).

Перед тем как программа начнет обрабатывать данные, запускается мастер настроек. С помощью фильтров можно получить подробные статистические данные по интересующей категории посетителей или событий — выделить пользователей, зашедших из определенных поисковых машин, стран, посмотреть отдельно поведение роботов различных поисковых серверов и т.д. Напомним: внешние счетчики могут выявлять приходы роботов поисковых машин или плохо, или после определенных усилий.

Благодаря своей гибкости, дружественному интерфейсу и наглядности отчетов WebLog Expert стал весьма популярен у владельцев сайтов и маркетологов. Однако для владельцев крупных сайтов нужен инструмент посерьезнее.

SPYLOG FLEXOLYZER

Программный пакет SpyLOG Flexolyzer (<http://www.flexolyzer.ru>), как заявляют производители, ориентирован на решение сложных задач с любой степенью детализации. Его идеологическое отличие от всех прочих инструментов анализа статистики состоит в том, что он позволяет не просто увидеть, кто и зачем приходит на сайт, но и точно узнать, что интересно целевой аудитории.

Достигается это за счет того, что SpyLOG Flexolyzer позволяет конструировать собственные отчеты, произвольно оперируя данными из разных отчетов. Чтобы пользователи смогли разобраться в системах настроек, на сайте даются примеры анализа эффективности рекламной кампании, оптимизации и т.д.

Кроме того, программа интегрируется с системой статистики SpyLOG, и те, кто подписаны на один из пакетов SpyLOG Tracker, получают доступ к информации о поведении своей целевой аудитории *вне* сайта. SpyLOG распознает пользователей с помощью файлов cookie. Благодаря им он может накапливать информацию относительно того, какие сайты (разумеется, оснащенные счетчиком SpyLOG) и как часто посещают пользователи Рунета.

Стоимость базовой версии SpyLOG FlexoLyzer составляет 20 000 рублей, включая годовую поддержку продукта, что подразумевает обновления продукта в течение одного года, обновление баз данных по географии и баз данных по поисковым машинам, каталогам и рейтингам. Для его эффективной работы применительно к сайту со средней посещаемостью (свыше 10 000 обращений в сутки) требуется довольно мощный компьютер. Самым оптимальным решением, по мнению создателей SpyLOG, будет выделение для анализа логов отдельного компьютера.

Как уже говорилось в разделе “Два в одном”, интеграция в одной программе двух методов статистики может оказаться очень эффективным решением. Однако, прежде чем приобретать программу SpyLOG Flexolyzer, неплохо бы посчитать, удастся ли за счет продвижения сайта и повышения прибыли компенсировать немалые средства, необходимые для приобретения и внедрения этой программы (вместе со счетчиком SpyLOG Tracker).

AWSTATS

Эта программа была разработана коллективом программистов из разных стран мира и изначально предназначалась для бесплатной установки.

Программу можно загрузить с официального сайта группы ее разработчиков, его адрес <http://awstats.sourceforge.net>.



Сайт AWStats official web site рассчитан на применение браузера Firefox. Это вполне логично: названный браузер также разрабатывался сообществом программистов и предназначен для бесплатного использования. Считается, что у него меньше “дыр” в системе защиты, чем у Internet Explorer 6, и что возможности его шире. Так что, прежде чем установить программу AWStats, придется установить браузер Firefox.

На момент написания книги самой свежей версией программы была 6.6. Загрузить ее можно было в разных вариантах:

- `awstats-6.6.zip` — в виде файла, архивированного программой Winzip. После сохранения файла на жестком диске его нужно разархивировать, а потом уже установить программу на сервер;
- `awstats-6.6.exe` — в виде выполняемого файла.

Предлагаются и другие варианты дистрибутивов, рассчитанных под операционные системы, отличные от Windows.

Программа позволяет получать такие группы отчетов о посещаемости, как:

- **When** (Когда) — по месяцам, по дням, по неделям и по часам;
- **Who** (Кто) — страны, хосты, последние посещения, нераспознанные IP-адреса, аутентифицированные пользователи, посещения роботов поисковых машин;
- **Navigation** (Навигация) — распределение посещений, точки входа, точки выхода, типы файлов, операционные системы, браузеры;
- **Referers** (Рефереры) — поисковые машины, ссылающиеся сайты, поисковые слова, поисковые фразы.

Эта программа достаточно популярна как за рубежом, так и в странах СНГ (видимо, в силу полной и бесповоротной бесплатности). Но устанавливать ее следует на сервере, и процедура установки довольно сложна. Поэтому в части III книги мы будем описывать не эту программу, а все же Webalizer. Хотя ее возможности меньше, она установлена на многих веб-серверах и также бесплатна.

ANALOG

Создатели этого бесплатного анализатора логов рекламируют его как “The most popular logfile analyser in the world” — наиболее популярный анализатор лог-файлов в мире. Ознакомиться с его возможностями можно на странице <http://www.analog.cx>.

Считается, что это — один из самых быстрых анализаторов лог-файлов веб-сервера. Успешно справляется с журнальными записями большого объема (свыше 1 Гбайт), экономно расходуя при этом память. Отчеты не содержат ничего лишнего, только необходимый минимум.

Отчеты предоставляются на 32-х языках. Совместим с любой операционной системой. Настораживает, однако, то, что последнее обновление на сайте появилось довольно давно — в середине 2005 года.

ALTERWIND LOG ANALYZER

Это — детище российских разработчиков (www.alterwind.ru). Программа предоставляется в трех вариантах — Lite (бесплатно), Standart (\$69) и Professional (\$99).

AlterWind Log Analyzer Lite — это бесплатный анализатор лог-файлов веб-сайта. С его помощью вы сможете определить все основные характеристики посещаемости сайта, узнать, с каких поисковых систем и по каким фразам приходят к вам посетители, узнать, какие ссылающиеся сайты приносят вам наибольший трафик, обнаружить и устранить ошибки в работе сайта. Всего генерируется 23 отчета. Вариант Standart предоставляет 33 отчета и 20 фильтров, вариант Professional — 52 отчета, 20 фильтров и множество дополнительных услуг — например, отчеты, предназначенные для оптимизации и продвижения сайтов.

КАКОЙ АНАЛИЗАТОР ЛОГОВ ВЫБРАТЬ

Ответ на этот вопрос мало зависит от *владельца* сайта.

Почти все владельцы серверов, предоставляющих хостинг, ныне устанавливают какой-то анализатор логов и предоставляют клиентам собранные им данные. Делается это строго по желанию, поскольку база данных с журнальными файлами имеет свойство быстро расти, а свою базу данных владелец сервера (хостер) сохраняет хорошо если месяц, а то и неделю. Клиенту нужно или копировать базу данных с лог-файлами на свой компьютер, уничтожив на сервере, или платить за хостинг больше. Синхронизация (в данном случае — “бесшовная” стыковка) баз данных на сервере и на своем компьютере — задача не из тривиальных. Ее можно поручить веб-мастеру, если у вас имеется таковой, но самостоятельно браться за это дело неспециалисту не стоит.

Установка на сервере приглянувшейся программы-анализатора логов также требует квалификации. Это — работа для специалиста. К тому же нужно получить согласие владельца сервера, предоставившего вам хостинг.

Самый простой вариант — это когда у вас свой, выделенный сервер, на котором размещен(ы) ваш(и) сайт(ы). Но в этом случае в компании (на фирме) наверняка есть сисадмин, веб-мастер, компьютерный гуру, который может провести все работы по установке на *ваш* сервер анализатора логов. Все, что вам остается, — это научиться работать с уже *выбранной* и *установленной* программой. Точнее — научиться анализировать полученные с ее помощью результаты. Это, в общем-то, не сложно. Конкретный пример рассматривается в части III нашей книги.