

К вопросу об учёте виртуальных читателей и посетителей сайтов

Карауш А. С., Майнагашев А. А.

Муниципальная информационная библиотечная система, Томск, Россия

Любая библиотека с каждым годом всё больше сталкивается с вопросами подсчёта посетителей полнотекстовых электронных ресурсов. Причём, вне зависимости от регистрации пользователей библиотека несёт материальные затраты на подготовку и предоставление этих данных. Однако в отчётах, которые ежегодно предоставляются в вышестоящие организации, такие данные либо отсутствуют, либо не принимаются во внимание, поскольку имеют незначительную ценность для анализа. В публикации определяются вопросы, и делается попытка оценки виртуального пользователя для возможности в дальнейшем «приравнивания» его к «реальному» посетителю библиотеки.

Отчеты в библиотеках с каждым годом вызывают чувство недоделанности и недосказанности за всё увеличивающийся для библиотеки объём работ – виртуальное обслуживание пользователей. Многие библиотеки предоставляют на своих локальных (доступных для определенных пользователей) и открытых сайтах многочисленные электронные ресурсы. И то, что современный человек уже не ходит в библиотеку так, как это было раньше, ни для кого не секрет. При этом количество посещений сайтов библиотек растёт, но в отчётах эти цифры имеют разное значение, поскольку посещение сайта в каждой библиотеке трактуется и оценивается по своим законам. При попытке провести анализ данных посещений сайта всегда будет стоять задача «привязать» количество виртуальных посещений и «реальных» читателей.

При каких же условиях, и какие значения должны иметь анализируемые цифры, чтобы можно было «приравнять» виртуальных посетителей к «реальным» читателям. Основными составляющими могут быть разного рода показатели.

1. Использование только «типовых библиотечных» ресурсов, т.е. не чтение любых страниц сайта пользователем, а процесс чтения только тех страниц, где расположены ресурсы, ради которых пользователи приходили бы в библиотеку, но появилась новая возможность всё это читать с сайта.
2. Продолжительность сеанса чтения полнотекстовых ресурсов.
3. Возможность исключения из отчётов посещения поисковых роботов и сканеров.
4. Возможность и требование регистрации пользователя на сайте.
5. Стабильность показателей из года в год для количества посещений и виртуальных пользователей.
6. Возможность подсчёта статистических сведений о пользователях (возраст, пол, категория, место работы, место учёбы и пр.).
7. Согласованность получаемых данных со стандартами библиотечной статистики.

Для реализации модели, в которой имеется возможность отдельно подсчитывать посещения пользователей для разных категорий ресурсов, в МИБС города Томска было решено использовать несколько вариантов предоставления полнотекстовых ресурсов:

1. Файлы полнотекстовых ресурсов дублируются для сайтов разного уровня доступа. Таким образом, на сайте www.library.tomsk.ru имеются ссылки на файлы полнотекстовых ресурсов, которые должны учитываться отдельно от переходов пользователей на полнотекстовый ресурс при поиске через электронный каталог. Это позволило отделить запросы поисковых роботов от запросов виртуальных читателей и при анализе log-файла электронного каталога <http://www.library.tomsk.ru:8000> учитывать только те запросы, которые были из результатов поиска в электронном каталоге.
2. Возможности современного программного обеспечения позволяют анализировать время посещений конкретного пользователя электронных ресурсов сайта.

3. Для анализа используются только полнотекстовые ресурсы, представленные в формате PDF.
4. В отделе обслуживания подготовлена тетрадь учета виртуальной точки обслуживания для соблюдения правил библиотечной статистики.

Для подсчёта количества пользователей и посещений сегодня можно использовать разное программное обеспечение (ПО). В МИБС используется ПО AWStats и Web Log Explorer. Каждая из программ имеет свои достоинства и недостатки.

ПО **Web Log Explorer** – развитое средство для анализа файлов отчётов (log-файлов) известных серверов и построения на их основе сравнительных таблиц.

Достоинства:

- быстрая настройка;
- графический интерфейс, который упрощает работу;
- идентификация большинства «роботов» для корректировки данных посещений;
- гибкая в настройке система фильтрации.

Недостатки:

- нет возможности для определения и анализа продолжительности времени сессии пользователя;
- для каждого фильтра может быть получен только один отчет.

ПО **AWStats** – средство анализа log-файлов, позволяющее получить развернутый отчет в формате web-страницы множества данных о посетителях сайта и его отдельных страниц.

Достоинства:

- возможность подсчёта и анализа времени сессии пользователя;
- все результаты генерируются в виде web-страниц и связаны гиперссылками;
- AWStats может быть встроена в web-сайт.

Недостатки:

- первичная сложность настройки параметров;
- отсутствие графического интерфейса.

Используя вышеперечисленные ПО можно получать сведения и значения показателей, на основе которых создать модель виртуального пользователя максимально приближенного к «реальному» читателю библиотеки. В результате анализа в Муниципальной информационной библиотечной системе города Томска были получены параметры, требования и предложения для конвертации показателей библиотечной статистики.

Параметры, которые предлагаются для обсуждения при «приравнении» виртуального пользователя к «реальному» читателю:

1. Ресурс был получен пользователем путём перехода по ссылке на файл электронного документа из результатов поиска этого пользователя в электронном каталоге и только так.
2. Один и тот же IP-адрес, используемый в течение отчетного периода (года), удовлетворяющий п.1, считается одним пользователем.
3. Запросы в течение отчетного периода (года), сделанные согласно п.1, считаются книговыдачей и приравниваются (добавляются) к общей книговыдаче библиотеки.
4. Данные за отчетные периоды вносятся в тетрадь учета вновь созданной точки обслуживания (Web-сервер) для полноценной работы со статистикой.

Вопросы, остающиеся нерешёнными:

1. Определение сведений о виртуальном пользователе (пол, категория, возраст и пр.) путём внесения этих сведений самостоятельно пользователем.
2. Получение реальных данных из «мусора», который пользователи оставляют о себе в регистрационных данных сервера.
3. Невозможность предоставления таким пользователям лицензионного контента с учетом особенностей гл.4 Гражданского кодекса РФ.

Выводы

Обе рассматриваемые системы анализа log-файлов вполне пригодны для определения виртуальных пользователей. Система AWStats имеет несколько большую функциональность, но Web Log Explorer более качественно фильтрует запросы поисковых роботов, что является важным критерием для определения реальных адресов пользователей.

Результаты анализа можно использовать для «приравнивания» виртуальных пользователей к реальным читателям библиотек с учетом соблюдения описанных выше условий.