Chapter VIII

A Review of Methodologies for Analyzing Websites

Danielle Booth
Pennsylvania State University, USA

Bernard J. Jansen
Pennsylvania State University, USA

ABSTRACT

This chapter is an overview of the process of Web analytics for Websites. It outlines how visitor information such as number of visitors and visit duration can be collected using log files and page tagging. This information is then combined to create meaningful key performance indicators that are tailored not only to the business goals of the company running the Website but also to the goals and content of the Website. Finally, this chapter presents several analytic tools and explains how to choose the right tool for the needs of the Website. The ultimate goal of this chapter is to provide methods for increasing revenue and customer satisfaction through careful analysis of visitor interaction with a Website.

INTRODUCTION

Web analytics is the measure of visitor behavior on a Website. However, what kind of information is available from Website visitors, and what can be learned from studying such information? By collecting various Web analytics metrics, such as number of visits, visitors, and visit duration, one can develop key performance indicators (KPIs) — a versatile analytic model that measures several metrics against each other to define visitor trends. KPIs use these dynamic numbers to get an in-depth picture of visitor behavior on a site. This information allows businesses to align their Websites’ goals with their business goals for the purpose of identifying areas of improvement, promoting popular parts of the site, testing new site functionality, and ultimately increasing revenue. This chapter covers the most common metrics, different methods for gathering metrics, how to utilize key performance indicators, best
key practices, and choosing the right Web analytics tool.

The first section addresses metrics, information that can be collected from visitors on a Website. It covers types of metrics based on what kind of data is collected as well as specific metrics and how they can be utilized. The following section discusses the two main methods for gathering visitor information -- log files and page tagging. For each method, this section covers the advantages and disadvantages, types of supported information, and examples for data format. Following this is a section on how to choose the key performance indicators (KPIs). This includes outlining several business strategies for integrating Web analytics with the rest of an organization as well as identifying the type of Website and listing several specific KPIs for each site type. The following section provides the overall process and advice for Web analytics integration, and the final section deals with what to look for when choosing analytics tools as well as a comparison of several specific tools. Finally, the conclusion discusses the future of Web analytics.

**METRICS**

In order to understand the benefits of Website analysis, one must first understand metrics – the different kinds of available user information. Although the metrics may seem basic, once collected, they can be used to analyze Web traffic and improve a Website to better meet its overall goals. According to Panalysis (http://www.panalysis.com/), an Australian Web analytics company, these metrics generally fall into one of four categories: site usage, referrers (or how visitors arrived at your site), site content analysis, and quality assurance. Table 1 shows examples of types of metrics that might be found in these categories.

Although the type and overall number of metrics varies with different analytics vendors, there is still a common set of basic metrics common to most. Table 2 outlines eight widespread types of information that measure who is visiting a Website and what they do during their visits, relating each of these metrics to specific categories.

Each metric is discussed below.

**Visitor Type**

Since analyzing Website traffic first became popular in the 1990s with the Website counter, the measure of Website traffic has been one of the most closely watched metrics. This metric, however, has evolved from merely counting the number of hits a page receives into counting the number of individuals who visit the Website.

There are two types of visitors: those who have been to the site before, and those who have not. This difference is defined in terms of repeat and new visitors. In order to track visitors in such a way, a system must be able to determine individual users who access a Website; each individual visitor is called a unique visitor. Ideally, a unique visitor is just one visitor, but this is not always the case.

<table>
<thead>
<tr>
<th>Site Usage</th>
<th>Referrers</th>
<th>Site Content Analysis</th>
<th>Quality Assurance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers of visitors and sessions</td>
<td>Which websites are sending visitors to your site</td>
<td>Top entry pages</td>
<td>Broken pages or server errors</td>
</tr>
<tr>
<td>How many people repeatedly visit the site</td>
<td>The search terms people used to find your site</td>
<td>Most popular pages</td>
<td>Visitor response to errors</td>
</tr>
<tr>
<td>Geographic information</td>
<td>How many people place bookmarks to the site</td>
<td>Top pages for single page view sessions</td>
<td></td>
</tr>
<tr>
<td>Search Engine Activity</td>
<td></td>
<td>Top exit pages</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Top paths through the site</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effectiveness of key content</td>
<td></td>
</tr>
</tbody>
</table>
Related Content

Building Elastic Java Application Services Seamlessly in the Cloud: A Middleware Framework
www.irma-international.org/chapter/building-elastic-java-application-services-seamlessly-in-the-cloud/140825/

Novel Approaches for Integrating MART1 Clustering Based Pre-Fetching Technique with Web Caching
www.irma-international.org/article/novel-approaches-for-integrating-mart1-clustering-based-pre-fetching-technique-with-web-caching/89327/

A Layered Communication Architecture for Power Routing in the Smart Grid
www.irma-international.org/article/a-layered-communication-architecture-for-power-routing-in-the-smart-grid/113321/

Design of a Least Cost (LC) Vertical Search Engine Based on Domain Specific Hidden Web Crawler
Sudhakar Ranjan and Komal Kumar Bhatia (2018). The Dark Web: Breakthroughs in Research and Practice (pp. 319-333).
www.irma-international.org/chapter/design-of-a-least-cost-lc-vertical-search-engine-based-on-domain-specific-hidden-web-crawler/185880/

Deployment and Adoption Strategy of Cloud Computing for Blended Learning in Higher Education Institutions in Sub-Saharan Africa
www.irma-international.org/chapter/deployment-and-adoption-strategy-of-cloud-computing-for-blended-learning-in-higher-education-institutions-in-sub-saharan-africa/140847/