Web Analytics: Enhancing Customer Relationship Management

Nabil Alghalith
Truman State University

The Web is an enormous source of information. However, due to the disparate authorship of web pages, this information is buried in its nebulous and chaotic structure. Due to the pervasiveness of Web access, increasing numbers of users are relying on Web search engines for finding information. Users are interested in identifying how pieces of information are related as they are presented on the Web. Because of its interactiveness, simplicity, and unobtrusiveness, using a web site as a data collection tool is now a commonplace. The results of the web analytics - the rules that will uncover fraudulent transactions or show which customers are likely to buy what products at the same time, or who is about to switch to a competitor - would ideally be integrated into a dynamic website. This would provide an automated, end-to-end, targeted marketing tool geared towards customer relationship management. This paper explores the following areas: Web analytics, Dimensions of web analytics, Web content analytics, Web usage analytics, Web structure analytics, Techniques of web analytics, Benefits of web analytics, Web Analytics Business Applications, Conclusions.

INTRODUCTION

The advent of e-commerce revolutionized every industry. Every aspect of commerce, from sales pitch to final delivery, could be automated and made available around the clock, all over the world. Business-to-Business solutions carried this one step further, allowing vertical partnerships and co-branding. Businesses found a new incentive to bring their data into the digital age. Dynamic content allowed the first truly personalized, interactive websites to come into being, all through the magic of e-commerce. The sources of data were of the old realm of business point-of-sale terminals, inventory databases, transaction records. Attempts to understand the data, first with statistical tools, later with OLAP systems, met with limited success – until the introduction of Web mining and Web analytics. As defined by Delen, Sharda, and Turban (2015), Web analytics is "the application of business analytics activities to Web-based processes, including e-commerce". Using machine-learning algorithms, and artificial neural networks, Web analytics software finds useful hidden information and patterns in the data and uses them to form new rules and predict the future behavior of customers, turning that mountain of data "Big Data" into valuable knowledge and untapped business opportunities and to enhance customer experience management (CEM) and customer relationship management (CRM). Customer relationship management system is an information system that provide an integrated approach to all aspects of interaction a company has with its customers (Brown, 2012)

What is Web Analytics?

Web analytics is information technology tool that collects, stores, analyzes, and graphically presents data collected from websites (Laudon, 2014).

Web analytics is one of the hottest topics in information technology. It automatically and thoroughly explores "Big Data" very large datasets, consequently uncovering otherwise concealed relationships among data (Moxon, 1996). This technology has been successfully applied in all fields including science, health, marketing and finance (Solheim, 1996) to aid in managerial planning and decision making and help organization gain a competitive advantage. In addition, Web analytics techniques are being applied to find and organize information from the web that is useful for data visualization and dash boarding. Web analytics itself is in the second generation of artificial intelligence; the core concept of web mining is focused on machine learning (Chapnick, 1996). This machine learning ability allows modification of search criteria automatically before the next execution, once particular patterns or trends have been discovered in the data searched (Yevich, 1997). It is important to understand that web mining is a discovery-oriented data analysis technology and not a single product or a system. It is a highly focused data transformation framework. This transformation process uses a series of analytical techniques, such as clustering, association, pattern recognition and classification (Etzioni, 1996). These techniques are taken from the fields of mathematics, cybernetics, pattern recognition and genetics, and can be used independently or cooperatively. The function is to extract high quality information to identify facts and draw conclusions based on relationships or patterns among the data. Most importantly, Web analytics can ask a processing engine to show answers to questions we do not know how to ask. For example, business customers' data are kept in different databases; thus, they are isolated from each other. Web analytics technology can search all the different databases together, and provide a better customer view so that the business can concentrate more on potentially good customers. The rationale is that when asking for specific relationships, more important relationships might be missed. Asking to find relationships that we do not know exist will yield data that are more meaningful or business knowledge. The combination of these two areas, data mining and the Web, is known as Web analytics.

Customer Relationship Management

Business and information technology that focus on collecting, processing and coordinating information about business interactions with customers including business transactions, marketing, sales, and service.

A company's Web site use can provide a wealth of marketing information. Web page visits generate three types of data:

- Data that is automatically generated and stored in server access logs (Sharda, 2015)
- User profiles that are collected by the web site profiling technology.
- Metadata about usage activities.

Web usage mining, or Web analytics, uses various technologies to dig through these large amounts of data to return useful information for evaluating and designing marketing strategies. Analyzing the data, called clickstream analysis, helps business owners better understand user behavior. Web analytics and clickstream analysis provide insight for improving business processes, increasing customer value, and improving Web sites themselves.

Web analytics is qualified as either off-site or on-site:

> **Off-site Web analytics:** analyzes a company's Internet presence outside of its own Web site.
> **On-site Web analytics**: considers visitor activity on a company's own Web site.

There are two technical ways to collect data for on-site Web analytics. The first is server log file analysis; the second is page tagging. Server log file analysis is the more traditional method, which records visitors' file requests. Page tagging uses embedded JavaScript to report page use or mouse clicks to a Web analytics third party. Both methods, when aggregated with other data sources such as e-mail or

social media-originating data, produce useful reports revealing ways to increase Web site traffic and improve performance.

In applying Web analytics to business objectives, four main categories of metrics are used:

- Web site usability
- Traffic sources
- Visitor profiles
- Conversion statistics (TWG, 2013)

Web Site Usability

Evaluates items such as page views, time on sight, and click paths to determine how user-friendly or user-relevant a web site is.

Traffic Source

Metrics identify traffic origination points, such as referral Web sites or even offline advertising campaigns.

Visitor Profiles

Data from visitor profiles can provide information such as geographical origination of traffic, the time of day users most frequently visit, or what keywords are used in reaching the sight.

Conversion Statistics

Measure which visitors are new, returning, or abandoning the site, as well as which are actually completing sales. Web analytic programs can arrange relevant data into a convenient dashboard for regular metric monitoring.

Dimensions of Web Analytics

There are three primary dimensions of Web analytics:

Web content analytics
Web usage analytics
Web structure analytics

Web Content Analytics

Web content analytics is the process of information or resource discovery from huge number of sources across the Web. There are two approaches in Web content analytics: the agent-based and database approaches. The agent-based approach involves artificial intelligence systems that can act autonomously or semi-autonomously on behalf of a particular user, to discover and organize Web-based information. Some intelligent Web agents can utilize a user profile to search for relevant information, then organize and interpret the discovered information. Some use various information retrieval techniques and the characteristics of open hypertext documents to organize and filter retrieved information.

The database approach focuses on integrating and organizing the heterogeneous and semi-structured data on the Web into more structured and high-level collections of resources. These organized resources can then be accessed and analyzed. These metadata, or generalizations, are then organized into structured collections (e.g., relational or object-oriented databases) and can be analyzed (Bamshad, 1997).

Web Usage Analytics

The other dimension of Web analytics is Web usage analytics. This is the process of discovering user access patterns (or user habits), as data are automatically collected in daily access logs. Referrer logs, which collect information about referring pages for each reference and user registration, have recently been included. Web usage analytics is crucial in establishing user profiles for a better structured website.

As the manner in which the Web is used continues to expand, there is a continual need to figure out new kinds of knowledge about user behavior.

Web Structure Analytics

Web structure analytics is the process of finding and developing useful information from the links found on web sites or web documents. It is used by search engines to identify web sites popularity and portals ranks and authority.

Techniques of Web Analytics

- The common techniques for Web analytics are:
- Clustering/classification
- Association rules
- Path analysis, and
- Sequential patterns

Clustering/Classification

Clustering, or classification, is a means to developing profiles of items with similar characteristics. This ability enhances the discovery of relationships that are otherwise not obvious. For example, classification of Web access logs allows a business to discover the average age of customers who order a certain product. This information can be valuable when developing advertising strategies.

Association Rules

Association rules are the ones that govern data warehouses “Big Data” of transactions where each transaction consists of a set of items. This technique is used to predict the correlation of items "where the presence of one pattern of items in a transaction implies (with a certain degree of confidence) the presence of other items. For example, association rules can provide a prediction of the percentage of clients assessing a particular site who will place online orders for a certain product.

Path Analysis

Path analysis is a technique that involves the generation of some form of graph that represents relation[s] defined on Web pages. This can be the physical layout of a web site in which the web pages are nodes and the hypertext links between these pages are directed edges. Most graphs are involved in determining frequent traversal patterns or large reference sequences from physical layout, such as the most frequently visited paths in a website. Another example is what paths users travel before they go to a particular URL.

Sequential Patterns

This technique is applied to Web access server transaction logs. The purpose is to discover sequential patterns that indicate user visit patterns over a certain period. For example, 25% of clients who visited a site had performed a search within the past week on a specific keyword or a product.

Benefits of Web Analytics

Web Analytics has numerous benefits including the following:

- Matching available resources to visitor interests
- Increasing the value of each visitor
- Improving the visitor's experience at the website
- Performing targeted resource management
- Collecting information in new ways
- Testing the relevance of content and web site architecture
- Web site optimization

Matching Available Resources to Visitor Interests

Available resources can be products a business sells, information fragments distributed online, banner ads from client advertisers, e-mail fragments from a mailing list, or anything else distributed online. Metadata of these resources are then stored in a database. Web analytics help learn visitor interests by collecting and analyzing information generated by interactions with a website, such as click stream data, search requests, intelligent agents and cookies. Web analytics tools can use the gleaned knowledge to rank resources by their relevance to users' interests. Servicing a user request for information, with the best matching resources, results in a higher visitor-to-customer conversion rate for e-commerce.

Increasing the Value of Each Visitor

Upon carrying out collaborative filtering, a Web analytics tool can predict what information a visitor might be interested in, and what products she might consider purchasing. These predictions are used to present the visitor with related products and resources. This knowledge significantly increases the value of a customer for an e-business when used in individualized cross-selling and up-selling promotions, and thus increases revenue.

Improving the Visitor's Experience at the Website

A sound combination of data and text analytics techniques can help determine user interests early in the process of the visitor's interaction with a website. This allows the web site to act interactively and proactively to deliver the most relevant customized resources to the visitor. In the Internet world, where the cost of switching to a different vendor is virtually zero, easy access to relevant information might become the difference between a profitable customer and a lost opportunity. By increasing the customer's satisfaction, business owners reduce attrition and build brand loyalty.

Performing Targeted Resource Management

Not all visitors are equal. Some are the best potential customers, ready to click and buy. Others are prospecting for information, simultaneously familiarizing themselves with the brand. They might become very important and profitable customers in the future. However, there is a final group of visitors who enjoy only free rides. These folks will use promotional resources that a business offers to the fullest extent, but will never purchase anything. All of these visitors come through a single pipeline to a web site and are in a common queue for web site resources. Can a business owner tell them apart? Web site performance is limited and requests coming from the best prospects should be prioritized. If a business is distributing promotional resources of high value, it should spend its promotional budget wisely by offering and delivering promotional materials only to the best prospects - not to every Web surfer that visits the site. A Web analytics tool can work with load-balancing products to provide the best quality of service to the best customers.

Collecting Information in New Ways

While for the majority of e-vendors the task of collecting data is just an intermediate step necessary for improved target marketing, for others this task is the main motivation for creating a web site itself. Traditional data collection methods - promotions, surveys, focus groups, etc., have many well-known problems, including low response rates, poor accuracy, and high costs. Imagine the ability to offer promotional items online through a content-rich website, where visitors can find useful information in addition to submitting their contact information and requesting the promotion. Web analytic tools can learn a visitor's preferences (at virtually no cost) based on the content that the user was browsing. Of course, Web analytics is designed to work hand-in-hand with a privacy management system, allowing a business to collect valuable data while respecting the privacy of its web site visitors.

Testing the Relevance of Content and Website Architecture

Business owners sometimes need to take a close look at the characteristics of their websites content and architecture. Perhaps they would like to increase usability, or optimize the web site for the eyes of the

best prospects. Log analyzers can help visualize the most navigated paths through a website, averaged for all visitors. But is that really what an owner is looking for? When optimizing web site structure, the main concern should be to improve the experience for the most promising prospects, not for everybody. Small minority of web site visitors comprise really valuable prospects. The remaining majority hold little value other than sustaining the brand recognition traffic. Thus, business owners have to segregate the least important prospects and subtract their contribution from the overall picture of the site navigation. What is left represents the real quality of a website. This is the picture that can help improve the bottom line.

Web Site Optimization

The process of collecting and analyzing quantitative and qualitative information about web site including activities including web site traffic, activities, visits, and crowds. The idea of web site optimization is to automatically improve the site performance and quality.

Web Analytics Business Applications

Web analytics has numerous business applications including the following (Nasraoui, 2006):

- Evaluate promotional campaigns.
- Provide users dynamic information based on their interests.
- Target electronic ads based on user patterns.
- Design more effective marketing strategies.
- Predict user behavior.
- Provide better customer care and customer intimacy.
- Maintain and manage customer base.

CONCLUSION

Web analytics can be a promising tool to enhance customer relationship management and to address ineffective search engines that produce incomplete indexing, retrieval of irrelevant information. It is essential to have a system that helps the user find relevant and reliable information easily and quickly on the Web. Web analytics discovers information from mounds of data on the Web, but it also monitors and predicts user visit habits. This gives designers more reliable information in structuring and designing a website. Web analytics technology can help technicians design websites with paths that can be navigated easily by end users, saving time and effort. The world of Web analytics is simultaneously a minefield and a gold mine. By saving data associated with visitors, content, and interactions, a business can at least ensure that it will be available for use later. This could lead to web site optimization. Despite the difficulties, business owners might consider evaluating and incorporating Web analytics applications now. The sooner one starts learning from the data, the sooner one can gain a competitive advantage.

REFERENCES

Bamshad, M., Cooley, R., & Jaideep, S. (1997). *Web mining: Information and pattern discovery* on the World Wide Web at http://www- users.cs.umn.edu/~mobasher/webminer/survey/survey.html.

Brin, S., and L. Page. (2012). "*Reprint of the Anatomy of a Large-Scale Hypertextual Web Search Engine.*" Computer Networks, Vol. 56, No. 18, pp. 3825-3833.

Brogan, C., and Bastone, J. (2011). "*Acting on Customer Intelligence from Social Media: The New Edge for Building Customer Loyalty and Your Brand*." SAS white paper. sas.com/resources/whitepaper/wp_21122.pdf (accessed March 2013).

Brown et al. (2012). *Managing Information Technology*, 7th edition, Prentice Hal.

Chapnick, P. (1996, September). Data mining at redux. Database Programming & Design, 9(9), S5.

Davydov, M. M. (1997, March). *Exploiting data mining at the application level*. Wall Street & Technology 15(3), 59.

Delen, D., Sharda, R., & Turban, E. (2015) *Business Intelligence and Analytics: Systems for Decision Support*. Pearson Education.

Etzioni, O. (1996, November). *The world wide web: Quagmire or gold mine*? Communications of the ACM, 36, 65-68.

Johnston, S. J. (1996, September). *How to get a better return on data*. Information Week, (596), 82-84.

Kaplan, A. M., and M. Haenlein. (2010). "*Users of the World, Unite! The Challenges and Opportunities of Social Media*." *Business Horizons*, Vol. 53, No. 1, pp. 59-68.

Kleinberg, J. (1999). "*Authoritative Sources in a Hyperlinked Environment*." Journal of the ACM, Vol. 46, No. 5, pp. 604-632.

Laudon & Traver. (2014). *E-commerce: Business, Technology, Society*. Prentice Hall.

Moxon, B. (1996, August). *Defining data mining*. DBMS, 9(9), S-10.

Nastraoui, O., M. Spiliopoulou, J. Srivastava, B. Mobasher, and B. Masand. (2006). "*WebKDD 2006: Web Mining and Web Usage Analysis Post-Workshop Report*." ACM SIKDD Explorations Newsletter, Vol. 8, No. 2, pp. 84-89.

Paine, K. D., and M. Chaves. (2012). "*Social Media Metrics*." SAS white paper. sas.com/resources/whitepaper/wp_19861.pdf (accessed February 2013).

Peterson, E. T. (2008). *The Voice of Consumer: Qualitative Data as a Critical Input to Web Site Optimization*. Forseeresults.com/Form_peterson_webanalytics.html

Solheim, H. G. (1996). *Specific data mining applications*. at http://www.pvv.unit.no/ ~hgs/project/ report/node80.html

The Westover Group. (2013). "*20 Key Web Analytics Metrics and How to Use Them*." www.thewestovergroup.com (accessed February 2013).

Turban et al. (2011). *Decision Support and Business Intelligence Systems*. Prentice Hall.

Yevich, R. (1997). *Data mining, Data Warehouse: Practical Advice from the Experts*. Prentice Hall.