IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Volume-5, Issue-10 E-ISSN: 2347-2693

A Study on Raise of Web Analytics and its Benefits

U. Padma Jyothi[1*], Sridevi Bonthu[2], B. V. Prasanthi[3]

[1*]Department of Computer Science and Engineering, Vishnu Institute of Technology, Bhimavaram, India
[2] Department of Computer Science and Engineering, Vishnu Institute of Technology, Bhimavaram, India
[3]Department of Computer Science and Engineering, Vishnu Institute of Technology, Bhimavaram, India

[*]*Corresponding Author: padmajyothi64@gmail.com, Tel.: +91-9550993996*

Available online at: www.ijcseonline.org

Received: 17/Sep/2017, Revised: 30/Sep/2017, Accepted: 13/Oct/2017, Published: 30/Oct/2017

***Abstract*—** Internet is expanding day by day in terms of its users and websites. Web Analytics is the process of measuring statistics of the website and analyzing the behaviour of traffic. 39 percent of the companies present now do not use any web analytics. This paper studies evolution of web analytics, the strategic methodologies that allows us to assess online activities, process, and the tools used. This paper also discusses the benefits obtained to small and large scale businesses if they incorporate web analytics to their websites.

***Keywords*—** Web analytics, Java script, e-commerce, web usage, log file, page tagging.

I. INTRODUCTION

Analytics is a surrounding and multidimensional field that uses mathematics, statistics, predictive modelling and various machine-learning techniques to find significant patterns and knowledge in the available data. Because of powerful computers with lots and lots of storage which run good algorithms, recording data has become an easier task. By applying data analytics, we can find out lot things and produce answers to questions which one never thought of asking. That is the power of data analytics. Every application we use is a web application now-a-days. Every time someone clicks, views, downloads, submits, visits, shares, or otherwise engages with some kind of your digital attendance, they add to analytics. Concerning to web, there are six high level areas in analytics. They are web analytics, App Analytics, Social Media Analytics, Advertising Analytics, Audience Analytics and Big Data Analytics. The paper studies the web analytics, which is becoming more useful now.

Now-a-days, all the data related to any kind of business, education [11], health care, government data etc are all accessed with the help of web applications and serviced from any objet storage clouds [9]. In this digital era, the usage of website is increasing drastically, a website not only provides cost effective and timely platform for communication with stakeholders, but also helps to shape and present its image on the internet [12]. Different methods are used by the designers of the Website, to make a website more pleasing, usable, effective,robust, comprehensive, and competitively better.

The usage of the website can be defined as a "quality attribute that describes how easy it is for a user to navigate through the website" [10]. Apart from traditional methods of website evaluation like expert based testing, user based testing. web analyatics came to existence which performs analytics on the website so that imporves can be done.

According to the Web Analytics Association, Web analytics can be defined as the process of measurement, collection, analysis and reporting of web information for purposes of understanding and optimizing web usage [1]. According to Kaushik and Daniel [2] web Analytics is defined as the science and the art of improving websites.

In general terms, web analytics is the process of collecting data about the activities of people accessing your website (visitors) - how they found you, when they visited, what pages they looked at, what they bought or downloaded, and so on - and mining that data for information that can be used to improved [3].

II. RELATED WORK

A. Evolution

In the early 1990s, web analytics is simply saving the internet data as log files. This resulted for the evolution of search engine spiders. Later Javascript tags refined this process. These tags made the viewing of website activities easier. In 2005, Google analytics entered into the battle and it made drilling down into audience behaviour simpler. Figure 1 gives a clear picture about the evolution of web analytics.

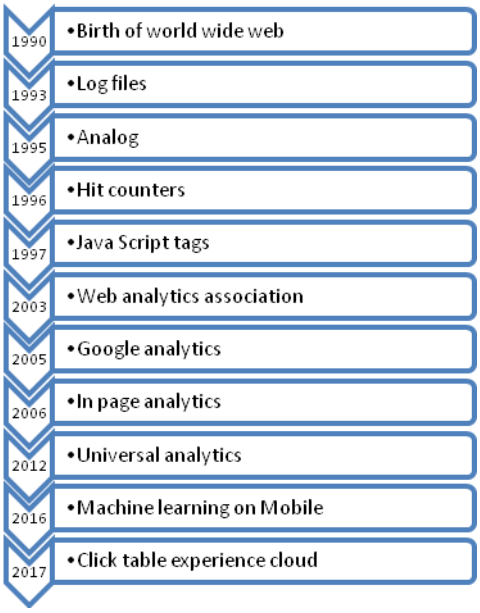

Figure 1. Evolution of web analytics

B. *Technologies*

Web analytics can be categorised into two types; *off-site* and *on-site* web analytics.

- Off-site web analytics refers to measurement of web and analysis of web data without considering whether we own or maintain a website. It includes the measurement of a website's users, visibility, and comments that is occuring on the Internet. Example for an e-commerce website, we can identify the size of the market, identify competitors, determine market peneration for our market, conduct surveys and consult market research sources. Tools like Google insight for search and Compete.com provides off-site web analytics.

- On-site web analytics refers to a most common method to measure a page visitor behavior for our website and performace in a commerical way rather than traditional way. This includes its conversions and drivers. For example, the different landing pages are associated with online purchases. This web data is typically compared against key performance indicators for performance, and which is used to improve a website or marketing based on response of the users. It uses trend analysis software to analyze server logs and tag pages.

  To remain ahead of the race in this competitive digital world, marketers are leveraging upon some web analytics tools. The most widely used on-site web analytics tools are Google Analytics and Adobe Analytics.Even though new tools have come into existence that provide additional information, like heat maps and session replay.

C. *Basic Steps*

Web analytics processes are mainly classified into 4 stages:

- Collection of data: It is the basic step in the web analytics process i.e, collection of data which is an elementary data. Mainly this data includes counts of things.
- Processing of data into information: Whatever the data which is obtained from first stage is translated in terms of ratios which is an information considered as metric and still there may be some counts.
- Developing KPI: It mainly focuses on using the information after processing and combining them with some business strategies, referred to as Key Performance Indicators (KPI). Depending on the organization, KPIs deal with conversion aspects but not always.
- Formulating online strategy: This step mainly deals with the online goals, objectives, and standards for the organization or business. These online strategies are used to increase marketing, making money and saving money.

Additional function available in web analytics process

- Experiments and testings: Changes to web pages are to be identified because it may increase or maximize a statistically tested results of a website which can be done by A/B testing . It is a controlled method which is available in online settings with two variants.

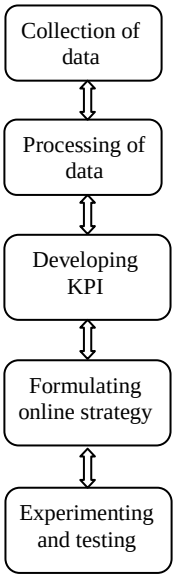

Figure 2. Basic steps in web analytics process

D. Data sources

The primary goal of web analytics is to collect and analyze data related to web traffic (which includes search engines, referring sites direct traffic and others) and usage different patterns. The data mainly comes from four sources:

1. Direct HTTP request data: Data directly comes from HTTP request messages from a HTTP request header.

2. Network level and server generated data associated with HTTP requests: Not only part of an HTTP request it also includes other information like, IP address of a requester.

3. Application level data sent with HTTP requests: The request is generated and processed by application level programs (such as JavaScript, PHP, and ASP.Net), including session and referrals. These are usually stored by internal logs rather than public web analytics services.

4. External data: It can be combined with on-site data which includes website behavior data described above and interpret web usage. For example, IP addresses are related with Geographic regions and internet service providers (ISP), e-mail open and click- rates, campaign data of a direct mail, sales and lead history, or other data types as needed.

E. Tools

However, the field of web analytics is vast. Numerous tools have been developed with web analytics.Each of these tools deliver different answers and helps in different ways to analyze a particular aspect of web design, content, promotion and strategy. Therefore, which ones of these you pick depends on your need to analyze your website. Tools like Google analytics, Clicky, Mint, Church Analytics, KISSmetrics, Open Web Analytics, Clicktale, CrazyEgg, PIWIK, CloudFlare, Adobe, Urchin are increasingly gaining popularity.

F. Benefits

The role of web analytics has been gaining among the online community. It helps to look at our latest internet site trends and our visitors' or users' preferences in terms of site features. Some examples of the benefits of web analytics.

- Helps to keep an eye on the activities and behaviours of website users.
- Able to see key metrics like unique users, unique sessions, top performing website content, the performance of different traffic sources and much more.
- Gain key insights such as geographical locations and top traffic sources.
- To find number of people visiting using website.
- Segmentation of visitors by new / returning and referral sources.
- Able to identify which pages, the visitors visited most.
- Access the duration of stay of your visitors on pages.
- Able to identify the links clicked by visitors.
- Search terms being used for getting to website.

III. METHODS

In order to collect the data from websites for performing web analytics.In general there are two methods.In general

- Web server log file method
- Page tagging method

A. Web server log file method

The first method, according to Marshall and other experts this method is known as the web server log file method. This process deals with the tracking of files that are commonly stored on a web host server. These files automatically stores visitor behavior (such as site accessed time,visited pages, pages exited and others). Many high lead companies and webmasters use these files to manage storage and bandwidth issues. But, according to Marshall, the log files which are obtained can also be parsed and analyzed by a tool and the data produced by that tools could help web site owners to improve their business.

In order to collect data from websites we can use a third-party data collection server or an in-house data collection server which requires an additional DNS by the user's computer to find the IP address of the data collection server. If there are any delays or failed DNS which results in no data collection.

Web log file method can be implemented with Ajax. As it is popular to obtain solutions or an alternative to the use of an invisible image is to implement a call back to the server from the rendered page. In this case, when the page is rendered on the web browser, a piece of Ajax code would call back to the

server and pass information about the client that can then be aggregated by a web analytics company. This is in some ways flawed by browser restrictions on the servers which can be contacted with XmlHttpRequest objects. Also, this method slightly reduces the traffic levels, as the visitor may stop the page from loading in mid-response before the calling Ajax.

The web server normally produces log files, so the raw data is already available. No changes to the website are required.Logfiles contain information on visits from search engine spiders, which generally do not execute JavaScript on a page and are therefore not recorded by page tagging. Although these should not be reported as part of the human activity, it is useful information for search engine optimization.

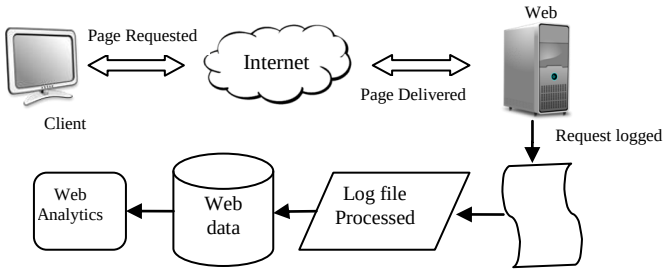

Figure 3. Log File Process

B. Page tagging method

The second method is the page tagging method, also named as the "JavaScript method". This method does not require log files at all. Everything is based on JavaScript code that is included with each web page. The JavaScript sends user activity to a computer that is hosted by the web analytics service provider. The website owner can view the analytics process of a website with a client viewer or web browser.

Counting is important in page tagging. When tha page is opened counting is activated which is given that the web client runs the tag scripts, not requesting it from the server. If a page is cached, it will not be counted by server-based log analysis. Cached pages can account for up to one-third of all page views. If the cached pages are not counted may lead to skew many site metrics. For that reason server-based log analysis is not considered suitable for analysis of human activity on websites.

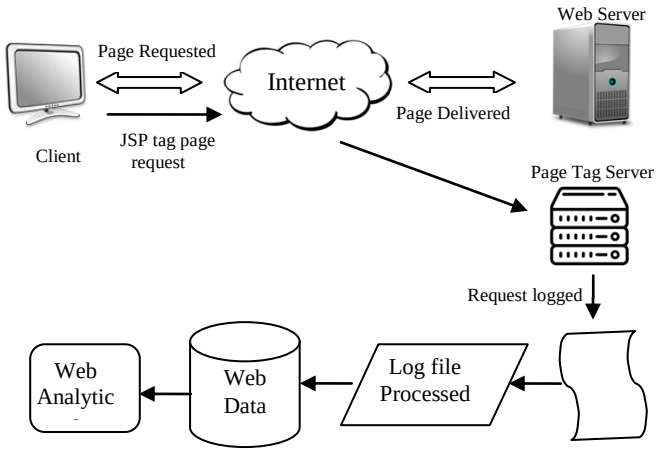

Figure 4. page tagging Process

Data is gathered via a component called tag in the page, usually written in JavaScript, though Java can be used, and increasingly Flash is used. Ajax can also be used in conjunction with a server-side scripting language (such as PHP) to manipulate and (usually) store it in a database, basically enabling complete control over how the data is represented.

Table 1. Comparison of methods

Measure	**Page tagging**	**Log file method**
Tracking of bandwidth	Yes	No
Accurately measure traffic across domains	Easy	Difficult
Expensive	less	
Caching effects the data		Yes
Reprocess data	Difficult	Easy
Tracking of scripting events	Easy	Difficult
Single page tracking	Accurate	Not accurate
Trace sessions across multiple domains	Easy	Difficult
Viewing of extensive information	Yes	No
Anchor URL Tracking	Yes	No
Information about unique visitors	Yes	No
IT overhead for custom setup	Low	High
use existing historical data logs	No	Yes
Search engine spiders	Ignores	
Cookie-based session IDs	Yes	No
Device tracing	Available	Not available
Data transparency	Yes	No
Historical data access	No	Yes
Download Tracking	No	Yes
Security	No	Yes

Few parameters of table1 are described as:

Event tracking:- Is tracking of events , user interaction like page views and downloads. Page tagging can perform event tracking completely but log file method only tracks the downloads.

Single Page Tracking:- JavaScript Tracking can track single pages accurately. In Web Log Analytics, It may be

challenging to achieve with Log Analytics, as the data provided by logs may be hard for the marketer or analyst to read. They will only see the URLs of API calls, which may not mean anything.

Device Tracking:-In JavaScript Tracking deals with tracking page titles, browser resolutions, and installed plugins. Web Log Analytics – not available as the relevant data is not available in the server http log files.

Unique Visitor Tracking :- In JavaScript Tracking uses both first-party cookies and browser fingerprint in order to accurately identify unique visitors. Log file analytics method cannot accurately calculate returning visitors, visitor recent visit, number of visits of a user to conversion and other metrics.

Data Transparency :- In JavaScript Tracking it provides clean and reliable data. Log analytics analyzes all http server calls. There is quite a bit of work required to filter out URLs that are not relevant (like calls from monitoring systems or single page app API calls).

Tag Implementation:- In JavaScript Tracking , every page of the website code must be placed in order to implement tag. Web Log Analytics does not require tracking code because java script includes tag.

Bandwidth Analytics:- In JavaScript Tracking-Bandwidth usage cannot be tracked, because the information is not provided by the web server to the browser client. Web Log Analytics can also be used for bandwidth analytics of how much bandwidth each website generates.

Historical Data Access:- In JavaScript Tracking can access to data collected only from the moment of implementation of our analytics software. Web Log Analytics – we can process historical data even though any analytics sysatem is not available previously.

Download Tracking:- In JavaScript Tracking downloads and other events from the website may not be registered if the user has a slow connection or JavaScript disabled. Web Log Analytics- We can track all downloads based on the information from the http server.

Security:- In JavaScript Tracking anyone who access server can essentially inject any code into your website, capture data from users that leave it on the website, etc. Moreover, JS Tracking requires adding external files into the website. This may not be feasible for many organizations due to data security considerations. In Web Log Analytics - Analytics server can be separated from the website for better security. Banks, healthcare and other data-sensitive industries take advantage of this capability. Various methods of security can be provided by using Internet Analysis tools. Some of them are brower history captures, web historian, web page saver, my last search etc., incorporates the procedures of monitoring and identifies users offline activities [6][4].

According to John Marshall, "The log file method tends to be less accurate than the JavaScript method. But, since the JavaScript method usually involves paying monthly fees to an analytics company, the log file method is typically less expensive. With the log file method, a website operator typically incurs a one-time software expense, but he can then analyze the log files at no additional cost."

Use Web Log Analytics when: Organization handling great amounts of sensitive data and we need to track activity of a particular server ,if we want to analyze historical data from your log files and our website contains too many pages and we don't want to place JS tags on every single one of them.

C. Hybrid method

Hybrid method can be as combination of both logfiles and page tagging .Some companies produce solutions with hybrid method and can analyze both kinds. By using a hybrid method,more accurate statistics can be obtained. An early hybrid solution was produced in 1998 by Rufus Evison.

IV. CONCLUSION

Web analytics always keep an eye on the activities and behaviours of website users. It brings a set of comprehensive requirements, improved metrics and optimization practices. It monitors the impact of marketing, design, or product changes. Web analytics can give ample of insights connected to any specific area, regardless of the size of the organization. These areas can be online presence, reach to social media, experience of the user, online reputation, web traffic and so many. It can be used by any company, as it detects problems in the early stage.

REFERENCES

[1] WAA Standards Committee. "*Web analytics definitions*." Washington DC: Web Analytics Association (2008).

[2] D. Waisberg and A. Kaushik, "*Web Analytics 2.0: Empowering Customer Centricity*," The original Search Engine Marketing Journal 2 , vol. 2, no. 1, pp. 5– 11, 2009.

[3] Peterson, Eric. "What Is Web Analytics." *Xml.com*. N.p., 2017. Web. 17 Oct. 2017.

[4] Prasanthi, B. V., Prathyusha Kanakam, and S. Mahaboob Hussain. "*Cyber Forensic Science to Diagnose Digital Crimes-A study."*International Journal of Scientific Research in Network Security and communication (IJSRNSC),2017

[5] Jeremy Aube, Director of Engineering. "*Collecting Web Data: A Look At Web Analytics Methodology - ROI Revolution.*" *ROI Revolution*. N.p., 2017. Web. 17 Oct. 2017.

[6] B.V.Prasanthi ,"*Cyber Forensics Tools : A Review*" International Journal of Engineering Trends and Technology (IJETT),vol 41 no 5(2016).

[7] *"A Brief History Of Web Analytics & DCX Insights." Clicktale*. N.p., 2017. Web. 17 Oct. 2017.

[8] "Web Analytics Basics | Usability.Gov." *Usability.gov*. N.p., 2017. Web. 17 Oct. 2017.

[9] Bonthu, S., S S R Murthy, Y., & Srilakshmi, M. (2014). *Building an Object Cloud Storage Service System using OpenStack Swift*. International Journal Of Computer Applications, 102(10), 39-42. http://dx.doi.org/10.5120/17854-8827

[10] T. Stewart, "*Websites – Quality and Usability*", Behaviour & Information Technology , vol. 31, no. 7, pp. 645-646, 2012.

[11] Sridevi Bonthu et al, (2017). *Automation of Pre-processing of Students Data*. Indian Journal of Computer Science and Engineering (IJCSE), 8(3), 241-245. ISSN : 0976-5166 Vol. 8 No. 3 Jun-Jul 2017.

[12] Mentes, S. Ahmet, and Aykut H. Turan. "*Assessing the usability of university websites: an empirical study on Namik Kemal University.*" TOJET: The Turkish Online Journal of Educational Technology 1 1 .3 (2012).

Authors Profile

Mrs. U.Padma Jyothi pursed Bachelor of Technolgy in Computer Science & Engineering from JNTUK affiliated college, Kakinada in year 2012 and and Master of Technology in Computer Science & Engineering from JNTUK affiliated college, Kakinada in year 2014. She is currently working as Assistant Professor in Department of Computer Science, in Vishnu Institute of Technology since 2015. Her main research work focuses on Big Data, data and web analytics. She has 2 years of teaching experience.

Sridevi Bonthu received her B.Tech. in Computer Science and Engineering from Acharya Nagarjuna University, India and her MTech in Computer Science from JNTUK, Kakinada, India. She is currently a faculty in Computer Science and Engineering department at Vishnu Institute of Technology, Bhimavaram, Andhra Pradesh, India. Her primary research interests include web application development, Big Data, data and web analytics.

Mrs. B.V.Prasanthi pursed Bachelor of Technolgy in Computer Science & Engineering from Kakatiya University, Kothagudem in 2009 and Master of Technology in Computer Science & Engineering from JNTUH affiliated college,Hyderabad in year 2014. She is currently working as Assistant Professor in Department of Computer Science, in Vishnu Institute of Technology since 2014. Her main research work focuses on Information Forensics, Network Security and Biometrics. She has 3 years of teaching experience.

