

Analyzing Actual and Anticipated Usage To Promote Web Navigation Usability

RafatRana S.H. Rizvi, Prof.Ranjit R. Keole

Abstract— Usability is defined as the effectiveness, efficiency, and satisfaction with which specific users can complete specific tasks in a particular environment. Despite the heavy and increasing investments in website design, it is still revealed, however, that finding desired information in a website is not easy and designing effective websites is not a trivial task. Poor website design has been a key element in a number of high profile site failures. It is also found that user having difficulty in locating the targets are very likely to leave a website even if its information is of high quality. A primary cause of poor website design is that the web developers' understanding of how a website should be structured can be considerably different from those of the users. Such differences result in cases where users cannot easily locate the desired information in a website. This paper presents a new method to identify navigation related Web usability problems based on comparing actual and anticipated usage patterns. The actual usage patterns can be extracted from client-side logs routinely recorded for operational websites by first processing the log data to identify users, user sessions, and user task-oriented transactions, and then applying an usage mining algorithm to discover patterns among actual usage paths. The anticipated usage, including information about both the path and time required for user-oriented tasks, is captured by our ideal user interactive path models based on their cognition of user behavior. The comparison is performed for checking results and identifying user navigation difficulties.

Index Terms— Usability, Server Logs, IUIP, Web Navigation, Web analytics..

I. INTRODUCTION

Since the advent of the World Wide Web in 1990, Internet usage worldwide has grown from roughly 2.6 million users (0.05% of the world population) in 1990, to roughly 2.0 billion users (30% of the world population) in 2010 (The World Bank Group, 2012). This trend is expected to continue for the foreseeable future. With this, websites have become increasingly important in the lives of individuals worldwide. With the increasing use of and dependence on websites, it is important that websites be usable (Nielsen, 2003); usability is defined as “[the] extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”.

Manuscript received April, 2016.

Rafat Rana S.H.Rizvi, PG Department of Computer Science And Engineering, SGBAU, HVPM COET., Amravati, India, 9405349147

Ranjit R.Keole, PG Department of Computer Science And Information Technology, SGBAU, HVPM COET, Amravati, India, 9823852893

Good usability is an important factor of successful user experience. Creating good user experience the effectiveness of a website can be increased significantly. Having a good user experience is extremely important to a website because most of the cases the user has many competing sites to choose from. It is easy to change from website to another if the user's experience of the website does not please the visitor. Studies show, that 30 percent of consumers got a negative image from a company based on their poorly designed website. 57 percent of users may change the web-shop or website if they find it complicated to use. If a product is not available at the website 50 percent of users assume that the product is not available at all.

Despite the heavy and increasing investments in website design, it is still revealed, however, that finding desired information in a website is not easy and designing effective websites is not a trivial task. Poor website design has been a key element in a number of high profile site failures. It is also found that user having difficulty in locating the targets are very likely to leave a website even if its information is of high quality. A primary cause of poor website design is that the web developers' understanding of how a website should be structured can be considerably different from those of the users. Such differences result in cases where users cannot easily locate the desired information in a website.

In order to improve web navigation usability and to improve user experience, it is necessary to identify navigation related web usability problems. Usability engineering provides methods for measuring usability and for addressing usability issues. Usability testing is a method used to understand the usability of a website and involves observing users interacting with the website. However, usability testing requires planning, can be expensive, can be time consuming, and yields a poor representation of usage by a small subset of users. Heuristic evaluation by experts are typically used to identify usability issues and to ensure satisfactory usability. However, significant challenges exist, including 1) accuracy of problem identification due to false alarms common in expert evaluation, 2) unrealistic evaluation of usability due to differences between the testing environment and the actual usage environment, and 3) increased cost due to the prolonged evolution and maintenance cycles typical for many Web applications.

An alternative to usability testing and heuristic evaluations is Web Analytics of web server log files (recordings of user interactions on a website). Web Analytics provides a means for understanding website usage and discovering usability insights. Unlike usability testing, log file data is automatically generated by web servers, inexpensive to use, depicts real usage by all website users, and can be used post-hoc. Web Analytics reports on

high-level aggregate usage, which can be used to understand the usability of websites. For example, for an online shopping website, Web Analytics will provide tables and graphs depicting what popular product pages users visited, what the average user clicked on, and common paths through the website.

This paper extract actual user behavior from Web server logs, capture anticipated user behavior with the help of cognitive user models, and perform a comparison between the two. This deviation analysis would help us identify some navigation related usability problems. Correcting these problems would lead to better functional convenience as characterized by both better effectiveness (higher task completion rate) and efficiency (less time for given tasks). This new method would complement traditional usability practices and overcome some of the existing challenges.

II. LITERATURE REVIEW AND RELATED WORK

A. Logs, Usage, Usability

Two types of logs, i.e., server-side logs and client-side logs, are commonly used for Web usage and usability analysis. Server-side logs can be automatically generated by Web servers, with each entry corresponding to a user request.

In 1997, [M.F.Arlitt and C.L. Williamson,] states that by analysing these logs, Web workload was characterized and used to suggest performance enhancements for Internet Web servers. So, they have presents a workload characterization study for Internet Web servers. The workload characterization focuses on the document type distribution, the document size distribution, the document referencing behaviour, and the geographic distribution of server requests. But, they do not provide all of the information that is of interest. For example, the log entries tell only the number of bytes transferred for a document, not its actual size; there is no record of the elapsed time required for a document transfer.

In 1997, [L. Tauscher and S. Greenberg], proposed that users' revisit patterns can be discovered by mining server logs to develop guidelines for browser history mechanism. Through history, a user can return quickly to a previously visited page, possibly reducing the cognitive and physical overhead required to navigate to it from scratch. But in this paper how often user revisit their pages has not been analysed.

In 1999 [TecEd], proposed that Logs can provide insight into real users performing actual tasks in natural working conditions versus in an artificial setting of a lab. Logs also represent the activities of many users over a long period of time versus the small sample of users in a short time span in typical lab testing. So they briefly describes server log file analysis, discusses the requirements for log file data to yield usability data, and presents ways to integrate log file analysis into the usability engineer's toolkit. The drawback of this paper is that it explore only the limitations of log file data for usability analysis but do not provide solutions to overcome limitation.

In 1999, [R Cooley, B. Mobasher, J. Srivastava], discusses that Data preparation techniques and algorithms can be used to process the raw Web server logs, and then mining can be

performed to discover users' visitation patterns for further usability analysis. For example, organizations can mine server-side logs to predict users' behaviour and context to satisfy users' need. So they presents several data preparation techniques in order to identify unique users' and user sessions. It includes data cleaning, user and session identification and path completion. But it has limitations that there is difficulty in identifying users and user sessions, no specific method for solving problems.

In 2001 [C. kallepali and Jeff Tian], proposed that Server-side logs have been used to ensure the quality of web applications. It is also used to construct Web usage models for usage-based Web testing, or to automatically generate test cases accordingly to improve test efficiency. So they have proposed an approach in which statistical testing and reliability analysis can be used to ensure quality of web applications. To support this strategy, they have extract web usage and failure information from web logs. The usage information is used to build models for statistical web testing. The related failure information is used to measure the reliability of web applications and potential effectiveness of statistical web testing. But it has limitations that they have not consider browser rendering problems.

In 2001, [M. Y. Ivory and M. A. Hearst], presents that Client-side logs can capture accurate comprehensive usage data for usability analysis, because they allow low-level user interaction events such as keystrokes and mouse movements to be recorded. For example, using these client-side data, the evaluator can accurately measure time spent on particular tasks or pages as well as study the use of "back" button and user click streams. They have proposed WebVip (visual instrumentor program) and WET (web event logging tool) to capture such client side data. But to capture client side data, evaluator has to add code to a every link on all web pages, so time consuming. And log files analysing is manual.

In 2005, [F. E. Ritter, A. R. Freed, and O. L. Haskett], proposed that Server logs have also been used by organizations to learn about the usability of their products. For example, search queries can be extracted from server logs to discover user information needs for usability task analysis. So they presents a task analysis of user groups and what users look for on university department Web sites. They have developed this through a wide range of analyses, including reviewing existing department Web sites, departmental hardcopy handout materials, search-engine queries, and by interviewing users to see what additional information they require. But it has limitations that the list of user groups and the list of tasks is likely to be difficult to keep in mind, and would be difficult to generate alone in a single setting. There is a cost to testing a Web site using this list. It takes about an hour to work through the list of tasks and search the Web site to determine if the information is available.

In 2011, [T. Carta, F. Paternò, and V. F. D. Santanav], presents a tool that supports remote usability evaluation of Web sites. This work presents Web Usability Probe (WUP), a tool that follows a proxy based architecture, performs remote evaluation, and considers client-side logs as data source. Many problems was analysed like user is not getting accurate information, there was lack of user guidance, etc. But these problems were not solved.

In 2015, Ruili Geng, Jeff Tian, presents a new method to identify navigation related Web usability problems based on comparing actual and anticipated usage patterns. They propose to extract actual user behavior from Web server logs, capture anticipated user behavior with the help of cognitive user models and perform a comparison between the two. But they have not explore additional approaches to discover web usage patterns and related usability problems generalizable to other interesting domains. They have not started exploring deviation calculation and analysis at the trail level, cannot capture comprehensive usage data for usability analysis.

III. PROPOSED METHODOLOGY

Web server logs have been used for usage-based Web testing and quality assurance. They have also been used for understanding user behavior and guiding user interface design. These works are extended in this study to focus on the functional convenience aspect of usability. In particular, this paper focus on identifying navigation related problems as characterized by an inability to complete certain tasks or excessive time to complete them (RQ1). Usability engineers often use server logs to analyze users' behavior and understand how users perform specific tasks to improve their experience.

The proposed method identify navigation related Web usability problems based on comparing actual and anticipated usage patterns. First, method extract actual navigation paths from server logs and discover patterns for some typical events. In parallel, it construct IUIP models for the same events. IUIP models are based on the cognition of user behavior and can represent anticipated paths for specific user-oriented tasks. The result checking employs the mechanism of test oracle. An oracle is generally used to determine whether a test has passed or failed . Here, IUIP models is used as the oracle to identify the usability issues related to the users' actual navigation paths by analyzing the deviations between the two. This deviation analysis would help us identify some navigation related usability problems. Correcting these problems would lead to better functional convenience as characterized by both better effectiveness (higher task completion rate) and efficiency (less time for given tasks). This new method would complement traditional usability practices and overcome some of the existing challenges. Proposed method also anticipate user to access web page currently used by some another user of same category.

Proposed methodology has 2 module:

1) User module who will accessed the site.

2) admin module who will analyse the usage of visitors through server logs to improve web navigation usability.

Flow of user module is shown in fig1.



Figure 1:-user module

Flow of admin module is shown in fig2.

- **Steps:** 1) It extract actual user behaviour from server log.
- **Steps:** 2) compare user behaviour with designer event model
- **Steps:** 3) produce deviation data to indicate usability problems.
- **Steps:** 4) analyse problems and improve it with corrective action

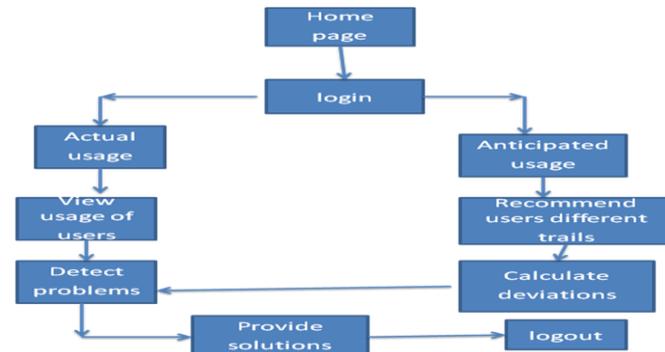


Figure 2; admin module

This paper used the sms.sedna.biz website's server logs as the case study to illustrate our method and its application. The functional convenience aspect of usability for this website is quantified by its task completion rate and time to complete given tasks. To identify problems, logical and temporal deviation was calculated for each pages of website. The results single out these Web pages and their design for further analysis , because such large deviations may be indications of some usability problems. The problems that was identified by our method can then be corrected for improvements in site.

IV. CONCLUSION

This paper have developed a new method for the identification and improvement of navigation-related Web usability problems by checking extracted usage patterns against cognitive user models. The method can identify areas with usability issues to help improve the usability of Web systems. Server logs in our method represent real users' operations in natural working conditions, and our IUIP models injected with human behavior cognition represent part of cognitive experts' work.

Some of research topics are as following:

-Presenting some strategies to further expand our usability research to cover more usability aspects to improve Web users' overall satisfaction.;

-Introducing some methods to explore additional approaches to discover Web usage patterns and related usability problems generalizable to other interesting domains;

ACKNOWLEDGMENT

My thanks to the Guide, Prof. R.R.Keole and Principal Dr.A.B.Marathe, who provided me constructive and positive feedback during the preparation of this paper.

REFERENCES

- [1] M. F. Arlitt and C. L. Williamson, "Internet Web servers: Workload characterization and performance implications," *IEEE/ACM Trans. Netw.*, vol. 5, no. 5, pp. 631–645, Oct. 1997.
- [2] L. Tauscher and S. Greenberg, "Revisitation patterns in World Wide Web navigation," in *Proc. ACM SIGCHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, 1997, pp. 399–406.
- [3] Tec-Ed, "Assessing web site usability from server log files," White Paper, Tec-Ed, 1999.
- [4] F. E. Ritter, A. R. Freed, and O. L. Haskett, "Discovering user information needs: The case of university department web sites," *ACM Interactions*, vol. 12, no. 5, pp. 19–27, 2005.
- [5] T. Carta, F. Paternò, and V. F. D. Santana, "Web usability probe: A tool for supporting remote usability evaluation of web sites," in *Human-Computer Interaction—INTERACT 2011*. New York, NY, USA: Springer, 2011, pp. 349–357.
- [6] T. Arce, P. E. Román, J. D. Velásquez, and V. Parada, "Identifying web sessions with simulated annealing," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1593–1600, 2014.
- [7] Ruili Geng, Jeff Tian, "Improving Web Navigation Usability by Comparing Actual and Anticipated Usage" *IEEE Transactions On Human-Machine Systems*, Vol. 45, No. 1, February 2015.
- [8] R. Cooley, B. Mobasher, and J. Srivastava, "Data preparation for mining World Wide Web browsing patterns," *Knowl. Inf. Syst.*, vol. 1, no. 1, pp. 5–32, 1999.
- [9] J. Nielsen, *Usability Engineering*. San Mateo, CA, USA: Morgan Kaufmann, 1993.
- [10] T. Tullis and B. Albert, *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics (Interactive Technologies)*. San Mateo, CA, USA: Morgan Kaufmann, 2008.
- [11] Ruili Geng, and Jeff Tian, "Improving Web Navigation Usability by Comparing Actual and Anticipated Usage", *IEEE Transactions On Human-Machine Systems*, Vol. 45, No. 1, February 2015

RafatRana S.H. Rizvi, received the B.E.degree in Computer Science and Engineering from H.V.P.M's College Of Engineering And Technology, Amravati in 2014. She is currently pursuing Master's Degree in Computer Science and Engineering from H.V.P.M's College of Engineering And Technology, Amravati.

Prof. Ranjit R.Keole, received the B.E.and M.E degree in Computer Science from Prof. Ram Megha Institute of Technology, Badnera in 1992 and 2008, respectively. His field of specialisation is web Mining. He is currently working as Associate Professor at H.V.P.M's college of Engineering and Technology, Amravati.