Enhancing Web Usability using User Behavior and Cognitive Study

Neha Mahyavanshi
Computer Department
Thakur College of Engineering
and Technology
Mumbai, India

Megharani Patil
Asst Prof.
Computer Department
Thakur College of Engineering
and Technology
Mumbai, India

Veena Kulkarni
Asst Prof.
Computer Department
Thakur College of Engineering
and Technology
Mumbai, India

ABSTRACT
In this era of technology, internet has gained huge popularity and web presence is an important aspect for success of any e-business. This paper presents a new method which focuses on two most important principles of web design i.e. functional convenience and presentational delight for enhancing the web usability. For improving functional convenience the actual usage patterns are extracted from web server logs for an e-commerce website. These logs are firstly preprocessed to detect users, user sessions and user transactions and then an usage mining method is applied to determine actual usage paths through the clickstreams. The actual usage paths will help to identify the usability issues and help suggest the corrective measures to improve it. For improving presentational delight a survey is taken to identify the user's perspective for making web design efficient. This cognitive study of user and actual user behavior help to identify the different usability issues. Based on these issues the different procedures are listed for design improvement.

General Terms
Data mining, Web usage mining, Pattern extraction, Design aspects.

Keywords
User Behavior, Web Usability, User perspective, Cognitive Study, Web Server Log, Usage Mining.

1. INTRODUCTION
Over past few years there has been enormous growth in the use of internet. It provides a platform which works as an international marketplace for various businesses. This has led to more and more businesses growing over internet. The web presence a business can have on internet heavily influence its success or failure. For this reason maintaining the functional convenience and presentational delight of any website has become a necessity. The functional convenience aspect refers to providing effective user interface for easy interaction with ease of navigation. The presentational delight aspect refers to providing attractive website features which has better visual impact on the users.

The proposed system extracts the actual user behavior by capturing the clickstreams obtained through the navigation of user on the website. In addition to this a survey is taken which emphasize on seizing the user perspective regarding the different facets of website design. Both the actual user behavior captured through web server logs and the results of user's perspective survey helps to identify some usability issues related to navigation and website design. Correcting these issues would lead to better functional convenience and presentational delight aspect of web usability.

The rest of this paper is organized as follows: Section II discusses the literature review with respect to the proposed method. Section III illustrates the methodology of the system. Section IV validates the method by providing results and analysis. Conclusions and future scope are presented in Section V.

2. LITERATURE REVIEW
Gupta A and Arora R presented a research work which focuses on web usage mining and specifically keeps tabs on running across the web utilization examples of sites from the server log records using Apriori and FP Tree algorithm [1]. The main drawback of Apriori algorithm is that the candidate set creation is costly, particularly if a large number of patterns or long patterns exist. The main drawback of FP-growth algorithm is that it lacks a good candidate generation method.

Gaol F.L presented web log sequential pattern mining using Apriori-all algorithm, called as Apriori-all Web Log Mining [2]. The experiment conducted is based on the Apriori-all algorithm, which for storing non-sequential data, first stores the original web access sequence database. The main drawback of Apriori-all algorithm is that the candidate set creation is costly, mainly if a large number of patterns and/or long patterns exists.

Yeming Tang, Qiuli Tong and Zhao Du analyzed campus map search log in Tsinghua University on term level, query level and session level [3]. To find out interests of users, generalized sequential pattern algorithm GSP is adopted to mine the frequent sequences of terms. The system performance and user experience is improved analyzing the campus map search log. A lot of time is consumed in candidate sequence generation and scanning database repeatedly hence the GSP algorithm is time inefficient.

Singh A.K, Kumar A and Maurya A.K. determined the empirical comparison of Apriori and FP-growth algorithm for frequent item set sequences for Web Usage data [4]. The analysis shows the comparison of properties like memory size, input data, pre-fetching, scalability and processing efficiency etc. The main drawback of Apriori algorithm is that the candidate set creation is costly, especially if a large number of patterns or long patterns exist. The disadvantage of FP-Growth is that it needs to work out conditional pattern bases and build conditional FP-tree recursively and hence it performs poor for data sets of long patterns.

Kotiyal B., Kumar A. and Pant B. presented a system which emphases on implementing an smart system that can provide

personalized web service for accessing related web pages in order to detect which web pages are more likely to be accessed by the user in future [5].This paper uses two algorithms namely Apriori and Eclat for forecasting the user behavior and also compares its performance with respect to time and space complexity for the filtered data. The results shows that Eclat algorithm serves better for the large databases despite Apriori as it generates less tables and therefore less time it takes to perform the analysis.

Khodambashi and Ø Nytrø presented a usability evaluation of published clinical guidelines (GLs) on the web and their web presentation [6]. Different methods were used such as interview, survey, system usability scale questionnaire, etc. From the step-by-step synthesis method and analysis of different results it was found that users are mostly concerned about various features like the amount of text and scrolling, font size of content and controls, presence of not more than one navigation bar and no redirection to other websites. The limitation of this paper is that the analysis is based on the survey of 14 users only.

Usability test is an essential process in the human computer interface design. It is a process which includes assembly of usability aspects of interface and evaluation and refining of those aspects. Najmeh Ghasemifard, Mahboubeh Shamsi, Abol Reza Rasouli Kenar and Vahid Ahmadi has presented the usability test methods in the human computer interface design, analysed them and stated the comparison of these methods [7]. The main drawback of this paper is that the results of the comparison did not evaluate the best method for usability testing.

The significance of end user documentation is well known for improving usability and operability of software applications. D. Amalfitano, A. R. Fasolino and P. Tramontana presented a semi-automatic approach for user documentation generation of Web 2.0 applications [8]. The method exploits dynamic analysis techniques for capturing the user behaviour of a web application and, hence, producing end user documentation amenable with known standards and guidelines for software user documentation. The limitation of this paper is that to verify the validity of the proposed approach experiments are needed which can prove the effectiveness and usefulness of the obtained documentation results.

The objective of any e-commerce website is to achieve high level of user participation to gain success. R Safavi presented a study for evaluating the different parameters which are important in the GUI and interface design. And which also affects the usability of the system from the customers' point of view [9].

3. METHODOLOGY

A new method is proposed to identify the navigation related usability problems from the actual usage patterns extracted from the server logs and cognitive study targeting user's perspective which identify the different design issues through survey. The architecture of the system for identifying usability problems is shown in Fig. 1.

The proposed system includes three main modules:

1. Usage Pattern Extraction
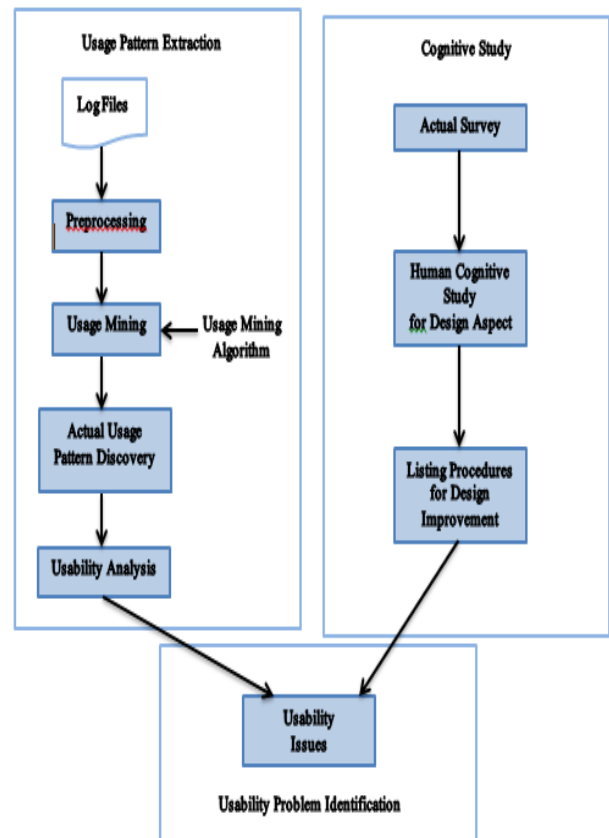2. Cognitive Study
3. Usability Problem Identification.

Fig 1: Architecture of the system for identifying usability issues

3.1 Usage Pattern Extraction

The web server logs provides input source of data to the system. The raw data is preprocessed to identify required information, remove extraneous fields and identify the unique users based on the combination of username and IP field. The activity record of each user is represented separately as a single visit to the website. A new usage mining method of Improved prefixspan algorithm is applied to discover the interesting frequent patterns from the actual usage paths of the users. This algorithm is based on the concept of Divide and Conquer method. It works in three main steps: 1) Scan the database to find all 1-itemset sequential patterns. 2) Get the projected databases by dividing the search space. 3) Find subsets of sequential patterns which can be mined by constructing projected databases. This algorithm uses certain specific sequential patterns pruning and duplicated projection which reduces the number and size of projected databases. This will improve the mining efficiency of the algorithm compared to prefixspan algorithm by reducing the runtime of scanning projected databases. The frequent patterns discovered from the actual usage paths are used to identify some common problems of the users while interacting with the website.

3.2 Cognitive Study

The Cognitive Study module is used to enhance the web usability by concentrating on design aspect of the website. The survey is conducted which focuses on different parameters like color, size, grouping, positions and presentation of information on the user interface. Based on the results of the survey procedures are listed for website design improvement.

3.3 Usability Problem Identification

This module takes in the frequent patterns obtained from the actual usage paths of the website users and the results obtained from the survey conducted. These data is analyzed to identify some usability problems of the users while interacting with the website and some procedures are listed for improving the website usability with respect to functional convenience and presentational delight aspects.

4. RESULTS AND ANALYSIS

This section describes certain results obtained after applying the algorithm on the website. To illustrate the working of the system an electronics website named elexforyou was developed and published which was used by different group of users. This website alike other e-commerce websites provided registration, selection and purchasing of various electronics products. The web server log data of this website was collected after it was made operational and used by different users. The frequent patterns were categorized into different types based on tasks such as selection of first item, selection of second item and change of selection. The usage patterns extracted were analyzed to identify some usability issues. Fig. 2 gives an example of a navigation problem identified during change in selection of an item. The user registers, does login, selects a category and selects an item from that category. For changing the selection of the item the user has to go back to state S3, from state S6, followed by states S4, S5, and S6. Hence this navigation problem increases the effort required by the user to perform change in selection of an item. Likewise, various such issues were identified related to different tasks which can be improved by redesigning the website for reducing the user effort and thus improving the web usability.
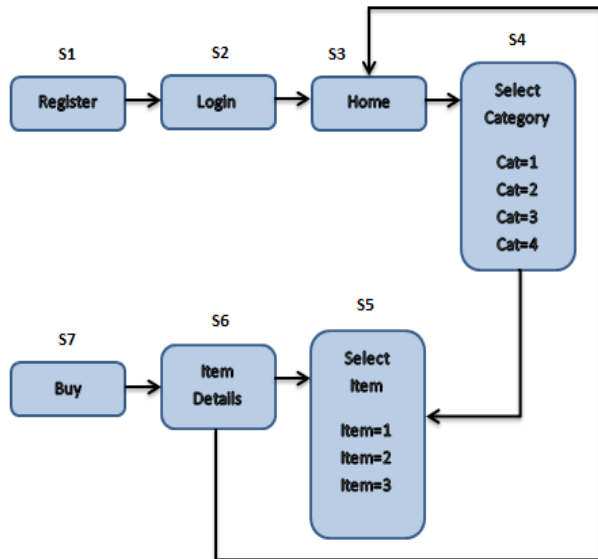

Fig 2: Example of a navigation problem

The website1 and website2, original electronics website elexforyou and redesigned website respectively, were used to evaluate the usability improvement. All the usability problems in website 1 identified by our method were fixed in website 2. The effort required in terms of number of steps required to perform the task and the task success rate were used to examine the usability improvement.

Table 1. Usability improvement between website1 and website2 based on Average Effort

Task	Average Effort(Number of Steps)		
	Website 1	Website 2	Improvement
1	4	3	1
2	7	4	3
3	7	5	2
Overall	6	4	2

Table 2. Usability improvement between website1 and website2 based on Task Success Rate

Task	Task Success Rate(%)		
	Website 1	Website 2	Improvement
1	36/41 = 87.80%	32/34 = 94.11%	6.31%
2	27/32 = 84.37%	35/38 = 92.10%	7.73%
3	19/25 = 76%	24/27 = 88.88%	12.88%
Overall	82.72%	91.69%	8.97%

Table 1 shows the usability improvement between website1 and website2 in terms of average effort required for a task. The average number of steps for each task decreased from 6 to 4 with an improvement of 2 steps. Table 2 shows the usability improvement between website1 and website2 in terms of the task success rate. The average improvement of task success rate is 8.97%.

Some of the website design issues were also found by analyzing the results of the survey conducted on the website. The survey questions were formed based on the different parameters like size, color, grouping, positions and presentation of information on the user interface. The survey conducted gives an understanding of the overall impression of design on the users, which can help improve the web usability based on these factors.

The survey questions are as follows:

1. How visually appealing is our website?

2. How easy was it to find what you were looking on our website?

3. Do you agree or disagree with the following statement: "The position of different controls on the screen is proper".

4. How would you rate our website in terms of the ease of navigation?

5. Do you agree or disagree with the following statement: "The overall design of the website is simple and familiar".

6. How would you rate our website in terms of organization of information?

7. Is the website looking attractive and its appearance is not hindering the user experience?

8. Is the information presented in appropriate size and color?

9. Is the grouping of different chunks of information proper or will you prefer it separated?

Table 3. shows the usability improvement between website1 and website2 based on the analysis of the survey. The average rating value between 1 to 5 is used for examining the usability improvement.

Table 3. Usability improvement between website1 and website2 based on Analysis of Survey

Design Aspect	Average Rating Value (1 to 5)	
	Website 1	Website 2
Q1	2.69	3.41
Q2	1.98	4.03
Q3	3.37	4.15
Q4	1.87	3.5
Q5	3.45	4.23
Q6	2.81	3.85
Q7	3.53	3.91
Q8	2.73	4.17
Q9	4.73	4.83

Based on the cognitive study of the survey results user's perspective related to different design aspects of the website was identified and used for listing the improvement procedures.

5. CONCLUSION
A new method is developed which contribute towards enhancing the web usability by identifying the navigation related web usability problems and website design problems with respect to user's perspective. This combination approach of user behavior and cognitive study has better results since it represents the real users' activities and perspective in natural working environment. The result shows that web usability enhancement is achieved with respect to some usability issues identified based on actual user paths and user perspective for different design aspects. The future scope of the proposed idea lies in the incorporation of the validation studies with large-scale Web applications. Also additional approaches can be explored to discover Web usage patterns and related usability problems generalizable to other interesting domains. The usability research can also be expanded to cover more usability aspects to improve Web users' overall satisfaction.

6. REFERENCES
[1] A. Gupta, R. Arora, R. Sikarwar and N. Saxena, "Web usage mining using improved Frequent Pattern Tree algorithms, "International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), Ghaziabad, 2014, pp. 573-578.

[2] F. l. Gaol, "Exploring the Pattern of Habits of Users Using Web log Squential Pattern, "Second International Conference on Advances in Computing, Control, and Telecommunication Technologies, Jakarta, 2010, pp. 161-163.

[3] Y. Tang, Q. Tong and Z. Du, "Mining frequent sequential patterns and association rules on campus map system, "2nd International Conference on Systems and Informatics (ICSAI 2014), Shanghai, 2014, pp. 954-958.

[4] A. K. Singh, A. Kumar and A. K. Maurya, "An empirical analysis and comparison of apriori and FP- growth algorithm for frequent pattern mining, "IEEE International Conference on Advanced Communications, Control and Computing Technologies, Ramanathapuram, 2014, pp. 1599-1602.

[5] B. Kotiyal, A. Kumar, B. Pant, R. H. Goudar, S. Chauhan and S. Junee, "User behavior analysis in web log through comparative study of Eclat and Apriori,"7th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, Tamil Nadu, India, 2013, pp. 421-426.

[6] S. Khodambashi and Ø Nytrø, "Usability Evaluation of Published Clinical Guidelines on the Web: A Case Study, "IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS), Dublin, 2016, pp. 140-145.

[7] Najmeh Ghasemifard, Mahboubeh Shamsi, Abol Reza Rasouli Kenar and Vahid Ahmadi, "A New View at Usability Test Methods of Interfaces for Human Computer Interaction," Global Journal of Computer Science and Technology: A Hardware and Computation, vol 15, Issue 1, Version 1.0, 2015.

[8] D. Amalfitano, A. R. Fasolino and P. Tramontana, "Using dynamic analysis for generating end user documentation for Web 2.0 applications, "13th IEEE International Symposium on Web Systems Evolution (WSE), Williamsburg, VI, 2011, pp. 11-20.

[9] R. Safavi, "Interface Design Issues to Enhance Usability of E-commerce Websites and Systems, "International Conference on Computer Technology and Development, Kota Kinabalu, 2009, pp. 277-281.

[10] E. Perakakis and G. Ghinea, "Responsive web design for the internet connected TV: The answer to more smart TV content?," IEEE 5th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), Berlin, 2015, pp. 38-42.

[11] Peishan Tsai, "A survey of empirical usability evaluation methods", GSLIS Independent Study, 2006.

[12] Julie M. Rinder, "The Importance of Website Usability Testing," Applied Information Management, University of Oregon, July 2012.

[13] J. Gao, "Research on teaching design of cartography teaching website of quality courses, "IEEE International Symposium on IT in Medicine and Education, Cuangzhou, 2011, pp. 400-402.

[14] A. Gregoriades and O. Vozniuk, "Combining physiological with cognitive measures of online users to evaluate a physicians' review website, "12th International Joint Conference on e-Business and Telecommunications (ICETE), France, 2015, pp. 152-159.

[15] S. M. R. S. Beheshti and P. Liatsis, "CAPTCHA Usability and Performance, How to Measure the Usability Level of Human Interactive Applications Quantitatively and Qualitatively?, "International Conference on Developments of E-Systems Engineering (DeSE), Duai, 2015, pp. 131-136.

[16] B. Lee, P. Isenberg, N. H. Riche and S. Carpendale, "Beyond Mouse and Keyboard: Expanding Design Considerations for Information Visualization Interactions," in IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 12, pp. 2689-2698, Dec. 2012.

[17] S. B. Singh and H. K. Singh, "An Efficient Approach for Improving Website Design," Fifth International Conference on Communication Systems and Network Technologies (CSNT), Gwalior, 2015, pp. 945-949.