

A Survey on Enhancing User Navigation Using Web Transformation

Madhuri B. Aswale¹, Gandhali. S. Gurjar²

¹Pune University, SAE, Kondhwa, Pune, Maharashtra, India

²Assistant Professor, SAE, Kondhwa, Pune, Maharashtra, India

Abstract: *Currently there are billions of internet users worldwide. A primary function of Web sites is to serve necessary information to the users. But Web sites are not always as successful or as usable as they could be. Discovery of required information in a website is a difficult task. Web site effectiveness depends on need of users. Web site users can encounter various problems when trying to acquire knowledge from a Web site and trying to use a Web site's functionalities. It is important to improve the navigation of website because they are mostly use for collecting information and for business purpose. Thus, it is important for the provider of a Web site to ensure that it is developed in a way that it should be useful for users. Designing a website which efficiently and effectively retrieves required Web pages on the Web is becoming a challenge. A main primary cause is that web developers may have understanding of how a web site should be design can be different from that of users this cause will lead to the problem for user to locate their intended information, however, the measure of website effectiveness should be the satisfaction of the users rather than that of the developers. Various techniques have been proposed to solve the problems.*

Keyword: Web mining, website design, user navigation, navigation technique

1. Introduction

Today internet is widely used by people around the world. Huge amount of information available on internet. It connect peoples to share information. Social site like Facebook connect individuals to each other. This tremendous increase in the use of internet leads the researcher, service provider to increase the functionality and usability of this source of information. In order to fulfill needs of users, providers investing heavily in maintenance and development in their websites. Discovery of required information in a website are a difficult task. If the user is not able to reach their intended information, then the user can exit the website even containing more quality information.

The process of obtaining useful information from World Wide Web is known as web mining. Web mining is a huge area and web mining techniques are used to solve problems related to information retrieval. It uses various data mining algorithms to extract information from websites. There are different types web mining: web content mining, usage mining and structure mining.

1.1 Web Content Mining

Extract information from web content. It is also known as text mining.

1.2 Web Structure Mining

Extract information from web structure.

1.3 Web Usage mining:

Extracting and analyzing information from web server logs. Data mining algorithms are used to evaluate the information. The main cause of poor website design is that the web

developers understanding is different from that of users[2]. Previous studies focus on various approaches related to website design to find the desired information like understanding web structure[5], finding the best relevant web page[6]. There are two ways to improve website navigation: Web personalization and Web Transformation. Web personalization deals with individual user and web logs are created by user's activity on website while Web transformation approach deals with structure of website which is applicable for all users[3]. In this paper a survey is based on transformation approach, which mainly focuses on reorganizing the complete website structure. But this results into difficulty for users. So improving website structure rather than reorganize is more beneficial.

2. Related Work

The fast growing number of internet users leads to numerous study on improving user navigation based on webserver logs. They are classified into web personalization and web transformation. There are several differences in web transformation and personalization. Transformation approaches modify the structure of a website which is applicable for all users, while personalization approaches applicable for individual users. To understand the user preferences, personalization approaches need to collect information associated with individual users. This is time consuming process not required in transformation.

Transformation approaches use weblog files and do not need to collect individual user information. Thus, personalization approaches are more suitable for dynamic websites and transformation approaches are more appropriate for static websites. Web personalization [19] is the process of finding relevant information from webpages according to particular user's need, by using user's profile data. Perkowitz and Etzioni [3] describe this Web personalization approach in

which index page contain links to particular pages based on the frequency of user traversal. However the method proposed by Mobasher et al. [15],[16] and Yan et al. [18] consist of weblogs which are retrieve using clustering. These clusters of user profile dynamically generate links for users according to their access pattern.

Nakagawa and Mobasher [19] presents an idea in which degree of connectivity is calculated to find out the user location. It is a hybrid personalization system. Bamshad Mobasher¹, Honghua Dai, Tao Luo, Miki Nakagawa develop two techniques based on clustering, to generate collective profiles that can be used for recommendation system to improve web personalization by using user profile. B. Mobasher, R.Cooley, J. Srivastava describe an approach where near about all web mining techniques and activities are used for web personalization. They propose automatic and dynamic web personalization. B.Mobasher, R. Cooley and J.srivastava have developed an approach where they collect information about users offline task and mining data. They have proposed a technique to gather user profiles using association rule discovery and uses based clustering.

Web transformation is the process of changing structure of website to improve navigation for all users [22] rather than personalizing webpages for individual users. Fu et al. [26] proposed a method to reorganize Webpages to provide required information to users in few clicks. But, this method considers only local structures in a website rather than the whole site, so the new structure may not an optimal solution. Gupta et al. [12] describe an approach based on finding maximum links from user preference data in websites for both wired and wireless devices. However, this approach takes relatively a long time to run even for a small website. Lin [13] proposed integer programming models that uses inter relationship between the web pages to reorganize a website. In addition, this model used to reduce the computation time. However, it is applicable to small websites only, for large websites this model still requires very long computation time.

Lin and Tseng [22] propose an ant colony system to improve website structures. Their approach provide solutions in a less time. It is also useful for small websites. It not appropriate for large websites. W. Yan, M. Jacobsen, H. Garcia-Molina and U. Dayal develop an approach in which users are classified based on pages that user visits. Using web pages logs, groups are made on the basis of users which visit similar pages. This data is used to improve navigation. According to users group they recommend links dynamically. Dean and Henzinger proposed algorithm to find relevant pages from page similarities. Page source consist of siblings and they are compare to parent pages when visited by users. According to this concept they find the relevant pages. Zhai and Liu describe an algorithm to extract a template using structural as well as visual layout information. Min Chen and Young U. Ryu [1] develops a mathematical programming model to enhance the navigation. This model uses web log files to do the changes in structure. Mini sessions are calculated and backtracking algorithm is used to find backtrack pages with the help of

mini sessions. This model improve navigation with minimum alteration in current website structure.

3. Issues in Website Restructuring

This survey based on the literature which improves user navigation by examining issue. Previous work on website structure improvement i.e. Web transformation approach having some issues. So there is need to work on these. Some algorithms or techniques are applicable to only small websites. Some are time consuming. Method describe in [1] do not consider user access pattern due to which accuracy of finding target is low. Efficiency is the most important issue while restructuring. Web transformation approach can be further improved by using various data mining algorithms for accuracy. This is important when websites are regularly improved.

4. Conclusion

In this paper we have presented a literature survey on Enhancing user navigation by using web transformation approach. Web transformations approaches mainly focus on developing methods to create or modify the link structure of a website. Various web transformation techniques are listed and current issues also examined in this survey.

References

- [1] Min Chen and Young U. Ryu Facilitating Effective User Navigation Through Website Structure Improvement"IEEE Transaction on knowledge and Data Engineering, Vol. 25, no. 3, March2013
- [2] T. Nakayama, H. Kato, And Y. Yamane, "Discovering the gap between Web Site Designers Expectations and Users Behavior," Computer Network, vol. 33, pp. 811-812,2000.
- [3] M. Perkowitz and O. Etzioni, "Towards Adaptive Web Sites: Conceptual Framework and Case Study," Artificial Intelligence, vol. 118, pp. 245-275, 2000.
- [4] J. Lazar, User-Centered Web Development. Jones and Bartlett Publishers, 2001.
- [5] Y. Yang, Y. Cao, Z. Nie, J. Zhou, and J. Wen, "Closing the Loop in Webpage Understanding," IEEE Trans. Knowledge and Data Eng.vol. 22, no. 5, pp. 639-650, May 2010.
- [6] J. Hou and Y. Zhang, "Effectively Finding Relevant Web Pages from Linkage Information," IEEE Trans.Knowledge and Data Eng., vol. 15, no. 4, pp. 940-951, July/Aug. 2003
- [7] H. Kao, J. Ho, and M. Chen, "WISDOM: Web Intrapage Informative Structure Mining Based on Document Object Model," IEEE Trans. Knowledge and Data Eng., vol. 17, no. 5, pp. 614-627, May 2005.
- [8] H. Kao, S. Lin, J. Ho, and M. Chen, "Mining Web Informative Structures and Contents Based on EntropyAnalysis," IEEE Trans.Knowledge and Data Eng., vol. 16, no. 1, pp. 41-55, Jan. 2004.
- [9] C. Kim and K. Shim, "TEXT: Automatic Template Extraction from Heterogeneous Web Pages," IEEE

- Trans. Knowledge and Data Eng., vol. 23, no. 4, pp. 612-626, Apr. 2011.
- [10] V. McKinney, K. Yoon, and F. Zahedi, "The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach," *Information Systems Research*, vol. 13, no. 3, pp. 296-315, 2002
- [11] M. Kilfoil et al., "Toward an Adaptive Web: The State of the Art and Science," *Proc. Comm. Network and Services Research Conf.*, pp. 119-130, 2003.
- [12] R. Gupta, A. Bagchi, and S. Sarkar, "Improving Linkage of Web Pages," *INFORMS J. Computing*, vol. 19, no. 1, pp. 127-136, 2007
- [13] C.C. Lin, "Optimal Web Site Reorganization Considering Information Overload and Search Depth," *European J. Operational Research*, vol. 173, no. 3, pp. 839-848, 2006.
- [14] M. Eirinaki and M. Vazirgiannis, "Web Mining for Web Personalization," *ACM Trans. Internet Technology*, vol. 3, no. 1, pp. 1-27, 2003.
- [15] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, "Discovery and Evaluation of Aggregate Usage Profiles for Web Personalization," *Data Mining and Knowledge Discovery*, vol. 6, no. 1, pp. 61-82, 2002.
- [16] B. Mobasher, R. Cooley, and J. Srivastava, "Automatic Personalization Based on Web Usage Mining," *Comm. ACM*, vol. 43, no. 8, pp. 142-151, 2000.
- [17] B. Mobasher, R. Cooley, and J. Srivastava, "Creating Adaptive Web Sites through Usage-Based Clustering of URLs," *Proc. Workshop Knowledge and Data Eng. Exchange*, 1999.
- [18] W. Yan, M. Jacobsen, H. Garcia-Molina, and U. Dayal, "From User Access Patterns to Dynamic Hypertext Linking," *Computer Networks and ISDN Systems*, vol. 28, nos. 7-11, pp. 1007-1014, May 1996.
- [19] M. Nakagawa and B. Mobasher, "A Hybrid Web Personalization Model Based on Site Connectivity," *Proc. Web Knowledge Discovery Data Mining Workshop*, pp. 59-70, 2003.
- [20] M. Eirinaki and M. Vazirgiannis, "Web Mining for Web Personalization," *ACM Trans. Internet Technology*, vol. 3, no. 1, pp. 1-27, 2003.
- [21] B. Mobasher, "Data Mining for Personalization," *The Adaptive Web: Methods and Strategies of Web Personalization*, A. Kobsa, W. Nejdl, P. Brusilovsky, eds., vol. 4321, pp. 90-135, Springer-Verlag, 2007.
- [22] C. Lin and L. Tseng, "Website Reorganization Using an Ant Colony System," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7598-7605, 2010.
- [23] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, (2002) "Discovery and Evaluation of Aggregate Usage Profiles for Web Personalization," *Data Mining and Knowledge Discovery*, vol. 6, no. 1, pp. 61-82.
- [24] B. Mobasher, R. Cooley, and J. Srivastava, (2000), "Automatic Personalization Based on Web Usage Mining," *Comm. ACM*, vol. 43, no. 8, pp. 142-151.
- [25] B. Mobasher, R. Cooley and J. Srivastava, (1999) "Creating Adaptive web Sites through Usage-Based clustering of URL," *Proc. Workshop Knowledge and Data Eng. Exchange*.
- [26] Y. Fu, M. Y. Shih, M. Creado, and C. Ju, "Reorganizing Web Sites Based on User Access Patterns", vol. 11, no. 1, pp. 39-53, 2002.

Author Profile

Ms. Madhuri B. Aswale is Student of ME Computer Engineering, SAE, Kondhwa, Pune, Pune University. India.

Ms. Gandhali S. Gurjar is Assistant Professor in Computer Engineering Department, SAE, Kondhwa, Pune, Pune University. India