FACILITATING EFFECTIVE USER NAVIGATION THROUGH WEBSITE STRUCTURE IMPROVEMENT

Mr. Suraj Rajaram Nalawade
M.Tech (CSE), Dept.of CSE MLR Institute technology, Dundigal Hyderabad Telangana-500043

Poreddy Dayaker
Assistant Professor, Dept.of CSE MLR Institute technology, Dundigal Hyderabad Telangana-500043

ABSTRACT

Designing well-structured websites to facilitate effective user navigation has long been a challenge, one of the reason is user behaviour is keep changing and web developer or designer not think according to user’s behaviour, so to improve user navigability by reorganizing website can be done by web transformation. This paper discusses how to improve a website without introducing substantial changes. In this paper we proposed a mathematical model to improve the user navigation on website. In addition, we define two evaluation metrics and use them to assess the performance of the improved website using the real data set. Evaluation results confirm that the user navigation on the improved structure is indeed greatly enhanced.

Index Terms: Website Design, User Navigation, Web Mining, Mathematical Model.

I. INTRODUCTION

There are millions of user for website since it is large source of information, web site also contain many links and pages every user require different pages at same time or same user may access different pages at different time. As user increases over www we need to make web intelligent, we concern here about intelligent website. To make web site intelligent we must know what is content of website, which are users and how website structured all this known as web mining.

Web design encompasses many different skills and disciplines in the production and maintenance of websites. The different areas of web design include web graphic design; interface design; authoring, including standardized code and proprietary software; user experience design; and search engine optimization. Often many individuals will work in teams covering different aspects of the design process, although some designers will cover them all. The term web design is normally
used to describe the design process relating to the front-end (client side) design of a website including writing mark up. Web design partially overlaps web engineering in the broader scope of web development. Web designers are expected to have an awareness of usability and if their role involves creating markup then they are also expected to be up to date with web accessibility guidelines.

Previous studies on website has discusses on a variety of issues, such as, extracting template from WebPages, mining informative structure of a news website, finding relevant pages of a given page, and understanding web structures. On the other hand, our work is closely related examines how to improve website navigability through the use of user navigation data. Various works have made an effort to address this question and generally it can be classified into two categories: first is to facilitate a particular user by dynamically reconstituting pages based on his profile and traversal paths, often referred as personalization, and second is to modify the site structure to ease the navigation for all users, often referred as transformation.

We perform experiments on a data set which is collected from a real websites. The results of these experiments indicate that our model can significantly improve the site structure with only few changes. Besides all this, the optimal solutions of the mathematical model are effectively obtained, suggesting that our model impractical to real-world websites. We also tested our model with synthetic data sets that are larger than the real data set. The solution times are remarkably low for all cases tested, ranging from fraction of second to up to 34 seconds. The solution times are shown to increase reasonably with the size of the website, indicating that the proposed MP model can be easily scaled to a large extent.

Motivation for choosing web structure effective user navigation through website structure improvement is: since web site is big source of information, but users mostly browsing useless page which irritates user and user lost interest from searching data over website. 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; however, the measure of website effectiveness should be the satisfaction of the users rather than that of the developers. Thus, Web pages should be organized in a way that generally matches the user’s model of how pages should be organized.

II. LITERATURE SURVEY

The purpose of this review is to report, evaluate, and discuss the findings from research. A particular focus of this review is to facilitating effective user navigation through website structure improvement.

- May Wang, Benjamin Yen, The study aims to improve Web navigation efficiency by reorganizing Web structure. Navigation efficiency is defined mathematically for both navigation with / without target destination pages, e.g. for experienced and new users. To help experienced users not to lose their orientation, structure stability is taken into consideration. Stability constraint can also help website designers control the maintaining effort of Web. This study proposes a mathematical programming method to reorganize Web structure in order to achieve better navigation efficiency. Designer can specify the user requirements and how stable the website structure should be. An e-banking example is given to illustrate how the method works in scenarios where user surfs with target destination. This study has the advantage of assessing and improving navigation efficiency and of relieving the designer of tedious chore to modify the structure in transformation.

- Devenish Dyane, NG EeeKeong and Sourav S Bhowmick, The unabated growth and increasing significance of the World Wide Web has resulted in a flurry of research activity to improve its capacity for serving information more effectively. But at the
Heart of these efforts lie implicit assumptions about “quality” and “usefulness” of Web resources and services. This observation points towards measurements and models that quantify various attributes of web sites. The science of measuring all aspects of information, especially its storage and retrieval or Informetrics has interested information scientists for decades before the existence of the Web. Is Web Informetrics any different, or is it just an application of classical Informetrics to a new medium? In this paper, we examine this issue by classifying and discussing a wide ranging set of Web metrics. We present the origins, measurement functions, formulations and comparisons of well-known Web metrics for quantifying Web graph properties, web page significance, web page similarity, search and retrieval, usage characterization and information theoretic properties. We also discuss how these metrics can be applied for improving Web information access and use.

- Ramakrishnan Srikant and Yinghui Yang, Many websites have a hierarchical organization of content. This organization may be quite different from the organization expected by visitors to the website. In particular, it is often unclear where a specific document is located. In this paper, we propose an algorithm to automatically find pages in a website whose location is different from where visitors expect to find them. The key insight is that visitors will backtrack if they do not find the information where they expect it: the point from where they backtrack is the expected location for the page. We present an algorithm for discovering such expected locations that can handle page caching by the browser. Expected locations with a significant number of hits are then presented to the website administrator. We also present algorithms for selecting expected locations (for adding navigation links) to optimize the benefit to the website or the visitor. We ran our algorithm on the Wharton business school website and found that even on this small website, there were many pages with expected locations different from their actual location.

- Mingjun Li, Mingxin Zhang, JinlongZheng and Ying Lu, The improved algorithm selects expected locations (for adding navigation links) in backtracks set at the point of the earlier and the less backtracks, which avoids effectively negative impact to the accuracy of the overall analysis by the long access sequence. The experimental results show that the improved algorithm can find expected pages effectively, thus can achieve the target of adjustment and reorganization of website.

- Ms. Jissin Mary Kunjukutty and Ms. A. Priya, Web mining techniques are used to analyze web resource details. Content mining, structure mining and usage mining are the main types of web mining. Web page contents are analyzed in the content mining process. Structure mining technique is used to analyze the web site and page layouts. User access details are analyzed using usage mining methods. Web site structures are altered to improve the user navigations. Web personalization method reconstructs the page links with reference to the traversal path and profile of a particular user. Transformation mechanism is applied to modify the site structure for all users. User navigation data is used to link web pages to improve navigability. The out degree refers the number of outward links in a page. Out degree threshold is used to control the number of links in a page to minimize information overload in a page. Targeted pages are identified with page-stay time information. Mini sessions are identified with processed logs and path threshold information. Mathematical programming model is used to improve the user navigation on a website with minimum alteration in the current structure. Backtracking algorithm is used to estimate backtracking pages from mini sessions. Average user navigation and benefited user count metrics are used to evaluate the navigation performance. The web site restructuring scheme is enhanced with frequent pattern mining mechanism. Dynamic out degree threshold estimation model is adapted for the system. Target page identification process is performed with sequential patterns. Relative link information is used for the navigation pattern analysis.
III. IMPLEMENTATION DETAILS

3.1 Existing Work

A most important cause of poor website design is that the web developers perceptive of how a website should be structured and can be considerably different from those of the users. Such differences result in cases where users cannot easily find the desired information in a website. This issue is difficult to handle because when creating a website, web developers may not have a clear understanding of users’ preferences and can only organize pages based on their own ideas.

Existing System Algorithm:

In an existing system k-means algorithm is used for effective user navigation through website structure improvement.

Input: set of $k$ means $m_1^{(1)}, M_k^{(1)}$

Assignment step: Assign each observation to the cluster whose mean yields the least within-cluster sum of squares (WCSS). Since the sum of squares is the squared Euclidean, this is intuitively the "nearest" mean. (Mathematically, this means partitioning the observations according to the Voronoi diagram generated by the means).

$$S_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\|^2 \leq \|x_p - m_j^{(t)}\|^2 \ \forall j, 1 \leq j \leq k\},$$

Where each $x_p$ is assigned to exactly one, even if it could be is assigned to two or more of them.

Update step: Calculate the new means to be the centroids of the observations in the new clusters.

$$m^{(t+1)}_i = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j$$

Since the arithmetic mean is a least-squares estimator, this also minimizes the within-cluster sum of squares (WCSS) objective.

3.2 Propose Work

In this project we are presenting and extending the Cure clustering algorithm for user navigation through website structure improvement. The current method is dealing with user navigation through website structure. This approach delivers the efficiency as well as effectiveness of proposed methods for improvement of website. However this method is suffered from limitations like while creating website web developers don’t have clear understanding of clients’ requirement. Thus in this project our main aim is to present approaches to overcome the limitations. In this existing method we will add new algorithm which will efficiently do the improved user navigation through website structure. For this purpose we are using CURE clustering algorithm.

Algorithm

Input: Real websites dataset
- Random sampling: To handle large data sets, we do random sampling and draw a sample data set. Generally the random sample fits in main memory. Also because of the random sampling there is a tradeoff between accuracy and efficiency.
• Partitioning for speed up: The basic idea is to partition the sample space into $p$ partitions. Each partition contains $n/p$ elements. Then in the first pass partially cluster each partition until the final number of clusters reduces to $n/pq$ for some constant $q \geq 1$. Then run a second clustering pass on $n/q$ partial clusters for all the partitions. For the second pass we only store the representative points since the merge procedure only requires representative points of previous clusters before computing the new representative points for the merged cluster. The advantage of partitioning the input is that we can reduce the execution times.

• Labeling data on disk: Since we only have representative points for $k$ clusters, the remaining data points should also be assigned to the clusters. For this a fraction of randomly selected representative points for each of the $k$ clusters is chosen and data point is assigned to the cluster containing the representative point closest to it.

Output: Set of web links that needs to be redesign and relink.

IV. RESULTS

4.1 Input Dataset
Datasets of real websites.

4.2 Hardware and Software Used

Hardware Configuration
- Processor - PentiumIV 2.6 GHz
- RAM - 512 mb dd ram
- Monitor - 15” color
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard

Software Configuration
- Operating System - Windows XP/7
- Programming Language - Java
- Database - MySQL
- Tool - Net beans

4.4 Results of Practical Work
V. CONCLUSION

This paper presents a comprehensive study for improvement of user navigation through website structure using CURE algorithm. We use this algorithm to improve the navigation effectiveness of a website while minimizing changes to its current structure. The tests on a real websites dataset showed that CURE algorithm could provide significant improvements to user navigation by adding only few new links. Optimal solutions were quickly obtained, suggesting that the CURE algorithm is very effective to real world websites datasets.

REFERENCES


