



ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Optimizing User Navigation with Pattern based Web Site Restructuring Scheme

Ms. Jissin Mary Kunjukutty¹, Ms. A.Priya, ME,²

II ME (CSE), Dept. of CSE, Sasurie College of Engineering, Vijayamangalam, Tamilnadu, India¹

Assistant Professor, Dept. of CSE, Sasurie College of Engineering, Vijayamangalam, Tamilnadu, India²

Abstract: 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 relink 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.

I. INTRODUCTION

Web Mining, which focuses on automatically discovering information and knowledge through the analysis of Web contents, Web structure and Web usages. Since the Web is huge, heterogeneous and dynamic, automated Web information and knowledge discovery calls for novel technologies and tools, which may take advantage of the state-of-the-art technologies from various areas, including machine learning, data mining, information retrieval, database and natural language processing.

The Web itself and the search engine indices contain information about the documents. Documents have different types of relationships among themselves. Hyperlinks add depth to documents, providing the multi-dimensionality, which characterizes the Web. Documents have an address, a URL, which represents a logical location on a server, which may provide information about the relationship of this document to other on the server. Also, there is a relationship to other documents on the Web unknown to the document, the search engine index may discover such relationships.

Web mining is a huge, interdisciplinary and very dynamic scientific area, converging from several research communities such as database, information retrieval, and artificial intelligence especially from machine learning and natural language processing. The World Wide Web (Web) is a popular and interactive medium, ideal for publishing information. It is huge, diverse and dynamic and thus raises issues of scalability, multimedia and temporal data respectively. This situation,



ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

the users are currently “drowning” in an information overload that expands at a rate that far outpaces human ability to process and exploit it.

Web mining techniques could be used to solve the information overload problems described above, directly or indirectly. Web mining can be broadly defined as the automated discovery and analysis of useful information from the web documents and services using data mining techniques. With the huge amount of information available online, the World Wide Web is a fertile area for data mining research. Web Mining research is at the crossroads of research from several research communities, such as database, information retrieval, and within AI, especially the sub areas of machine learning and natural language processing. Web mining is the use of data mining techniques to automatically discover and extract information from web documents and services. The web is rich with information; gathering and making sense of this data is difficult because publication on the web is largely unorganized.

II. RELATED WORK

The growth of the Internet has led to numerous studies on improving user navigations with the knowledge mined from webserver logs and they can be generally categorized in to web personalization and web transformation approaches. Web personalization is the process of “tailoring” webpages to the needs of specific users using the information of the users’ navigational behavior and profile data. Perkowitz and Etzioni describe an approach that automatically synthesizes index pages which contain links to pages pertaining to particular topics based on the co-occurrence frequency of pages in user traversals, to facilitate user navigation. The methods proposed by Mobasher et al. and Yan et al. create clusters of users profiles from weblogs and then dynamically generate links for users who are classified into different categories based on their access patterns. Nakagawa and Mobasher develop a hybrid personalization system that can dynamically switch between recommendation models based on degree of connectivity and the user’s position in the site. For reviews on web personalization approaches, see [12].

Web transformation, on the other hand, involves changing the structure of a website to facilitate the navigation for a large set of users [8] instead of personalizing pages for individual users. Fu et al. describe an approach to reorganize webpages so as to provide users with their desired information in fewer clicks. However, this approach considers only local structures in a website rather than the site as a whole, so the new structure may not be necessarily optimal. Gupta et al. [9] propose a heuristic method based on simulated annealing to relink webpages to improve navigability. This method makes use of the aggregate user preference data and can be used to improve the link structure in websites for both wired and wireless devices. However, this approach does not yield optimal solutions and takes relatively a long time (10 to 15 hours) to run even for a small website. Lin [10] develops integer programming models to reorganize a website based on the cohesion between pages to reduce information overload and search depth for users. In addition, a two-stage heuristic involving two integer-programming models is developed to reduce the computation time. However, this heuristic still requires very long computation times to solve for the optimal solution, especially when the website contains many links. Besides, the models were tested on randomly generated websites only, so its applicability on real websites remains questionable. To resolve the efficiency problem in [10], Lin and Tseng [8] propose an ant colony system to reorganize website structures. Although their approach is shown to provide solutions in a relatively short computation time, the sizes of the synthetic websites and real website tested in [8] are still relatively small, posing questions on its scalability to large-sized websites.

There are several remarkable differences between web transformation and personalization approaches. First, transformation approaches create or modify the structure of a website used for all users, while personalization approaches dynamically reconstitute pages for individual users. Hence, there is no predefined/built-in web structure for personalization approaches. Second, in order to understand the preference of individual users, personalization approaches need to collect information associated with these users. This computationally intensive and time-consuming process is not required for transformation approaches. Third, transformation approaches make use of aggregate usage data from weblog files and do not require tracking the past usage for each user while dynamic pages are typically generated based on the users’ traversal



ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

path. Thus, personalization approaches are more suitable for dynamic websites whose contents are more volatile and transformation approaches are more appropriate for websites that have a built-in structure and store relatively static and stable contents. This paper examines the questions of how to improve user navigation in a website with minimal changes to its structure. It complements the literature of transformation approaches that focus on reconstructing the link structure of a website. As a result, our model is suitable for website maintenance and can be applied in a regular manner.

III. WEBSITE STRUCTURE IMPROVEMENT

The advent of the Internet has provided an unprecedented platform for people to acquire knowledge and explore information. There are 1.73 billion Internet users worldwide as of September 2009, an increase of 18 percent since 2008 [1]. The fast-growing number of Internet users also presents huge business opportunities to firms. According to Grau [2], the US retail e-commerce sales totaled \$127.7 billion in 2007 and will reach \$218.4 billion by 2012. In order to satisfy the increasing demands from online customers, firms are heavily investing in the development and maintenance of their websites. InternetRetailer [3] reports that the overall website operations spending increased in 2007, with one-third of site operators hiking spending by at least 11 percent, compared to that in 2006.

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 [5], [6]. Galletta et al. [7] indicate that online sales lag far behind those of brick-and-mortar stores and at least part of the gap might be explained by a major difficulty users encounter when browsing online stores. Palmer highlights that poor website design has been a key element in a number of high profile site failures. McKinney et al. also find that users 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 problem is difficult to avoid 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 judgments. However, the measure of website effectiveness should be the satisfaction of the users rather than that of the developers. Thus, Webpages should be organized in a way that generally matches the user's model of how pages should be organized.

Previous studies on website has focused on a variety of issues, such as understanding web structures finding relevant pages of a given mining informative structure of a news website, and extracting template from webpages [11]. Our work, on the other hand, is closely related to the literature that examines how to improve website navigability through the use of user navigation data. Various works have made an effort to address this question and they can be generally classified into two categories: to facilitate a particular user by dynamically reconstituting pages based on his profile and traversal paths, often referred as personalization, and to modify the site structure to ease the navigation for all users, often referred as transformation.

In this paper, we are concerned primarily with transformation approaches. The literature considering transformations approaches mainly focuses on developing methods to completely reorganize the link structure of a website. Although there are advocates for website reorganization approaches, their drawbacks are obvious. First, since a complete reorganization could radically change the location of familiar items, the new website may disorient users. Second, the reorganized website structure is highly unpredictable, and the cost of disorienting users after the changes remains unanalyzed. This is because a website's structure is typically designed by experts and bears business or organizational logic, but this logic may no longer exist in the new structure when the website is completely reorganized. Besides, no prior studies have assessed the usability of a completely reorganized website, leading to doubts on the applicability of the reorganization approaches. Finally, since website reorganization approaches could dramatically change the current structure, they cannot be frequently performed to improve the navigability.



ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Recognizing the drawbacks of website reorganization approaches, we address the question of how to improve the structure of a website rather than reorganize it substantially. Specifically, we develop a mathematical programming (MP) model that facilitates user navigation on a website with minimal changes to its current structure. Our model is particularly appropriate for informational websites whose contents are static and relatively stable over time. Examples of organizations that have informational websites are universities, tourist attractions, hospitals, federal agencies, and sports organizations. Our model, however, may not be appropriate for websites that purely use dynamic pages or have volatile contents. This is because a steady state might never be reached in user access patterns in such websites, so it may not be possible to use the weblog data to improve the site structure [9].

The number of outward links in a page, i.e., the outdegree, is an important factor in modeling web structure. Prior studies typically model it as hard constraints so that pages in the new structure cannot have more links than a specified out-degree threshold, because having too many links in a page can cause information overload to users and is considered undesirable. For instance, Lin [10] uses 6, 8, and 10 as the out-degree threshold in experiments. This modeling approach, however, enforces severe restrictions on the new structure, as it prohibits pages from having more links than a specified threshold, even if adding these links may greatly facilitate user navigation. Our model formulates the out-degree as a cost term in the objective function to penalize pages that have more links than the threshold, so a page's out-degree may exceed the threshold if the cost of adding such links can be justified.

We perform extensive experiments on a data set collected from a real website. The results indicate that our model can significantly improve the site structure with only few changes. Besides, the optimal solutions of the MP model are effectively obtained, suggesting that our model is practical to real-world websites. We also test our model with synthetic data sets that are considerably larger than the real data set and other data sets tested in previous studies addressing website reorganization problem. The solution times are remarkably low for all cases tested, ranging from a fraction of second to up to 34 seconds. Moreover, 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.

To assess the user navigation on the improved website, we partition the entire real data set into training and testing sets. We use the training data to generate improved structures which are evaluated on the testing data using simulations to approximate the real usage [4]. We define two metrics and use them to assess whether user navigation is indeed enhanced on the improved structure. Particularly, the first metric measures whether the average user navigation is facilitated in the improved website, and the second metric measures how many users can benefit from the improved structure. Evaluation results confirm that user navigation on the improved website is greatly enhanced.

In summary, this paper makes the following contributions. First, we explore the problem of improving user navigation on a website with minimal changes to the current structure, an important question that has never been examined in the literature. We show that our MP model not only successfully accomplishes the task but also generates the optimal solutions surprisingly fast. The experiments on synthetic data indicate that our model also scales up very well. Second, we model the out-degree as a cost term in the objective function instead of as hard constraints. This allows a page to have more links than the out-degree threshold if the cost is reasonable and hence offers a good balance between minimizing changes to a website and reducing information overload to users. Third, we propose two evaluation metrics and use them to assess the improved structure to confirm the validity of our model. The evaluation procedure developed in this paper provides a framework for evaluating website structures in similar studies.

IV. WEBSITE RESTRUCTURING MODEL

Our problem can be regarded as a special graph optimization problem. We model a website as a directed graph, with nodes representing pages and arcs representing links. Let N be the set of all webpages and λ_{ij} , where $i, j \in N$, denote page connectivity in the current structure, with $\lambda_{ij} = 1$ indicating page i has a link to page j , and $\lambda_{ij} = 0$ otherwise. The current out-degree for page i is denoted by $W_i = \sum_{j \in N} \lambda_{ij}$.

**International Journal of Innovative Research in Computer and Communication Engineering**

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)**Organized by****Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014**

From the log files, we obtain the set T of all mini sessions. For a mini session $S \in T$, we denote $\text{tgt}(S)$ the target page of S . Let $L_m(S)$ be the length of S , i.e., the number of paths in S , and $L_p(k, S) L_p(k, S)$, for $1 \leq k \leq L_m(S)$, be the length of the k th path in S , i.e., the number of pages in the k th path of S . We further define $\text{docno}(r, k, S)$, for $1 \leq k \leq L_m(S)$ and $1 \leq r \leq L_p(k, S)$, as the r th page visited in the k th path in S . Take the mini session S , it follows that $L_m(S) = 3$, $L_p(1, S) = 3$, and $\text{docno}(1, 1, S) = A$, as this mini session has three paths and the first path has three pages (A, D, and H) in which page A is the first page. We define $E = \{(i, j) : i, j \in N \text{ and } \exists S \in T \text{ such that } i \in S \text{ and } j = \text{tgt}(S)\} \text{ and } N_E = \{i : (i, j) \in E\}$. In essence, E is the set of candidate links that can be selected to improve the site structure to help users reach their targets faster. Our problem is to determine whether to establish a link from i to j for $(i, j) \in E$. Let $x_{ij} \in \{0, 1\}$ denote the decision variable such that $x_{ij} = 1$ indicates establishing the link.

As explained earlier, Webmasters can set a goal for user navigation for each target page, which is denoted by b_j and is termed the path threshold for page j . Given a mini session S with target page j and a path threshold b_j , we can determine whether the user navigation goal is achieved in S by comparing the length of S , i.e., $L_m(S)$, with path threshold (b_j) for the target page of S . If the length of S is larger than b_j , it indicates the user navigation in S is “below” the goal. Then, we need to alter the site structure to improve the user navigation in S to meet the goal. Otherwise, no improvement is needed for S .

Intuitively, given path thresholds, we can determine which mini sessions need to be improved and hence are relevant to our decision (termed relevant mini sessions). The irrelevant mini session are not considered in our model. We denote the set of relevant mini sessions by $T^R \subseteq T$. For a mini session $S \in T^R$, it is said to be improved if the website is altered in a way such that the user could reach the target within the associated path threshold after changes are made to the site structure.

We define parameters $a_{ijk,r}^S$ to be 1 if $\text{docno}(r, k, S) = I$ and $\text{tgt}(S) = j$, and 0 otherwise. In other words, $a_{ijk,r}^S = 1$ if and only if page i is the r th visited page in the k th path in S and page j is the target of S . Further, we define variable C_{kr}^S which will be set to one if the solution indicates establishing a link from the r th page in the k th path in S to the target page of S , i.e., $\text{tgt}(S)$, and 0 otherwise. As will be explained later, the use of $a_{ijk,r}^S$ is to build connections between variables x_{ij} and C_{kr}^S , where the first variable uses global indices and the latter is defined using local indices.

Similar to prior studies, appropriate out-degree thresholds can be specified for webpages. We denote C_i the outdegree threshold for page i . Nevertheless, our model “penalizes” a page if its out-degree is larger than the threshold instead of modeling the threshold as a hard constraint. In effect, out-degree C_i indicates the maximum number of links that page i can have without being penalized. Let p_i be the number of links exceeding the out-degree threshold C_i for page i in the improved structure. Depending on the application of our model, different weights of penalties can be imposed on pages whose out-degree exceeds the respective out-degree threshold. We denote the weight by m and term it the multiplier for the penalty term.

V. ISSUES ON WEBSITE RESTRUCTURING PROCESS

User navigation data is used to relink 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 following issues are identified from the current website restructuring methods.

- User access pattern are not considered
- Out degree threshold is not optimized



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

- User target identification accuracy is low
- User interest levels are not considered

VI. ASSOCIATION RULE MINING PROCESS

A number of data mining algorithms have been introduced to the community that perform summarization of the data, classification of data with respect to a target attribute, deviation detection and other forms of data characterization and interpretation. One popular summarization and pattern extraction algorithm is the association rule algorithm, which identifies correlations between items in transactional databases.

Given a set of transactions, each described by an unordered set of items, an association rule $X - Y$ may be discovered in the data, where X and Y are conjunctions of items. The intuitive meaning of such a rule is that transactions in the database, which contain the items in X, tend to also contain the items in Y. An example of such a rule might be many observed customers who purchase tires and auto accessories also buy some automotive services. In this case, X = {tires, auto accessories} and Y = {automotive services}. Two numbers are associated with each rule that indicates the support and confidence of the rule. The support of the rule X - Y represents the percentage of transactions from the original database that contain both X and Y. The confidence of rule X-Y represents the percentage of transactions containing items in X that also contain items in Y. Applications of association rule mining include cross marketing, attached mailing, catalog design and customer segmentation. An association rule discovery algorithm searches the space of all possible patterns for rules that meet the user-specified support and confidence thresholds. The problem of discovering association rules can be divided into two steps:

Find all item sets whose support is greater than the specified threshold. Item sets with minimum support are called frequent item sets. Generate association rules from the frequent item sets. To do this, consider all partitioning of the item set into rule left-hand and right-hand sides. Confidence of a candidate rule X-Y is calculated as support (XY) / support (X). All rules that meet the confidence threshold are reported as discoveries of the algorithm.

L1: = {frequent 1-itemsets};
k:= 2; // k represents the pass number
While (Lk-1)

Ck = New candidates of size k generated from Lk-1
For all transactions t $\square \square$ DIcrement count of all candidates in Ck
That are contained in t
Lk = All candidates in Ck with minimum support
k = k+1

Report Uk Lk as the discovered frequent item sets

The first pass of the algorithm calculates single item frequencies to determine the frequent 1-itemsets. Each subsequent pass k discovers frequent item sets of size k . To do this, the frequent item sets L_{k-1} found in the previous iteration are joined to generate the candidate item sets C_k . Next, the support for candidates in C_k is calculated through one sweep of the transaction list. From L_{k-1} , the set of all frequent ($k-1$) item sets; the set of candidate k -item sets is created. The intuition behind this Apriori candidate generation procedure is that if an item set X has minimum support, so do all the subsets of X . Thus new item sets are created from ($k-1$) item sets p and q by listing p.item1, p.item2, p.item ($k-1$), q.item ($k-1$). Items p and q are selected if items 1 through $k-2$ are equivalent for p and q and item $k-1$ is not equivalent. Once candidates are generated, items etcs are removed from consideration if any ($k-1$) subset of the candidate is not in L_{k-1} .

VII.PATTERN BASED WEB SITE RESTRUCTURING SCHEME

Web site restructuring is performed with links and access log details. Apriori algorithm is used for the usage pattern analysis. Page domains are used for the reconstruction process. The system is divided into five major modules. They



ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

are log optimizer, site structure analysis, statistical model based restructuring, pattern identification and pattern based restructuring.

Log optimizer module is designed to perform access log preprocessing operations. Web page links are analyzed under the site structure analysis module. Statistical model based restructuring module is designed to rearrange the web site links. The pattern identification module is designed to fetch access patterns. Pattern based restructuring module is designed to perform site rearrangement using patterns.

7.1. Log Optimizer

The web page requests are maintained under the access log files. Data populate process transfers the log data into database. Redundant page requests are removed from the log files. The page requests are grouped into sessions.

7.2. Site Structure Analysis

Link information are collected for each page in the web site. Indegree is calculated with the inlink count for the web page. Outlink count is used to estimate the outdegree value for the web page. Indegree and outdegree values are used in the site restructuring process.

7.3. Statistical Model Based Restructuring

Mini session identification is carried out with the access log details. Back tracking algorithm is used to estimate back tracking pages. Outdegree threshold is used to rearrange the links in a web page. Outdegree and request frequency for the pages are used in the restructuring process.

7.4. Pattern Identification

Association rule mining method is used to identify the user access patterns. Apriori algorithm is used in the pattern extraction process. Support and confidence values are estimated from the session details. Minimum support and minimum confidence values are used for the pattern identification process.

7.5. Pattern Based Restructuring

Web site restructuring is performed with usage patterns. Dynamic outdegree threshold identification scheme is used to estimate the outlink levels. Search domains are used to identify the user interest levels. Access patterns, outdegree threshold and stay time details are used in the restructuring process.

VIII. CONCLUSION AND FUTURE WORK

Web usage mining techniques are used to analyze the user search behaviors. Web page links are used to manage the user navigations in the web site. Web page links are restructured to improve the user navigations. Pattern based model is used to improve the web site restructuring process. The system requires minimum site structure changes. High scalability is supported by the system. Web site navigability is improved by the pattern based model. Cost minimized model is used in the website restructuring process. The system can be enhanced with the following features. The system can be integrated with web personalization problem. The system can be connected with web proxy caching schemes. The prediction scheme can be adopted to predict dynamic pages with query values. Dynamic page based site restructuring can be adapted with the system.

REFERENCES

- [1] Pingdom, "Internet 2009 in Numbers," <http://royal.pingdom.com/2010/01/22/internet-2009-in-numbers/>, 2010.
- [2] J. Grau, "US Retail e-Commerce: Slower But Still Steady Growth," 2008.
- [3] Internetretailer, "Web Tech Spending Static-But High-for the Busiest E-Commerce Sites," <http://www.internetretailer.com/dailyNews.asp?id=23440>, 2007.
- [4] Min Chen and Young U. Ryu, "Facilitating Effective User Navigation through Website Structure Improvement", IEEE Transactions on Knowledge and Data Engineering, Vol. 25, No. 3, March 2013
- [5] X. Fang and C. Holsapple, "An Empirical Study of Web Site Navigation Structures' Impacts on Web Site Usability," Decision Support Systems, vol. 43, no. 2, pp. 476-491, 2007.
- [6] J. Lazar, Web Usability: A User-Centered Design Approach. Addison Wesley, 2006.



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

- [7] D.F. Galletta, R. Henry, S. McCoy, and P. Polak, "When the Wait Isn't So Bad: The Interacting Effects of Website Delay, Familiarity, and Breadth," *Information Systems Research*, vol. 17, no. 1, pp. 20-37, 2006.
- [8] C.C. Lin and L. Tseng, "Website Reorganization Using an Ant Colony System," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7598-7605, 2010.
- [9] R. Gupta, A. Bagchi, and S. Sarkar, "Improving Linkage of Web Pages," *INFORMS J. Computing*, vol. 19, no. 1, pp. 127-136, 2007.
- [10] 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.
- [11] 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.
- [12] 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.