Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 2007 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

2007

Web Structure Reorganization to Improve Web Navigation Efficiency

May Wang
The University of Hong Kong, amaywy@business.hku.hk

Benjamin Yen
The University of Hong Kong, benyen@business.hku.hk

30. Web Structure Reorganization to Improve Web Navigation Efficiency

May Wang
Faculty of Business and Economics,
The University of Hong Kong
amaywy@business.hku.hk

Benjamin Yen
Faculty of Business and Economics,
The University of Hong Kong
benyen@business.hku.hk

Abstract

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.

Keywords: Web Structure, Reorganization, Navigation Efficiency

Introduction

Search, navigation, and transaction are important activities on web use. Navigation, especially, is very important after users identify the corresponding Website using search engines and before make transactions. The study aims to improve Web navigation efficiency by reorganizing Web structure design. Some reorganization completely changes the overall structure and it is very frustrating to the user because they lose their orientation when navigating the Web. Structure stability is taken into consideration to facilitate user orientation. Stability constraint can also alleviate the maintaining problem of Web. This study proposes a mathematical programming method to reorganize Web structure in order to achieve better navigation efficiency. User can specify the requirements and designer can decide on how stable the website structure should be. Navigation efficiency is defined to reflect the average information accumulation performance and an e-banking example is given to illustrate how the method works. This study has the advantage of improving navigation efficiency and of relieving the designer of tedious chore to modify the structure in transformation.

Section 2 reviews the related literature. Definitions and our approach are described in Section 3. An e-banking example is given in section 4 under different user requirements and stability constraints. Conclusions and future research are discussed in Section 5.

Literature Review

Website navigation has been seemed as on of the most important design features across many domains, including finance, e-commerce, entertainment, education, government, and medical (Zhang et al. 2000). The quality of Website design has been studied from both qualitative and quantitative perspective (Guillermo et al. 2006; Luis et al. 2002; Webster et al. 2006). In literature, website design has been approached from different aspects. It has long been deemed as a hypermedia or a database. From engineering point of view, Web engineering and IDEAL take into consideration a whole picture of modeling and enhancement. From human computer interaction's perspective, interface, including graphical design, layout design and usability analysis also play an important role in website design. System design including hardware design, cache scheduling, etc. affect the website performance, too. Besides, structure design, including hyperlinks configuration and information structural design, has a great effect on the website navigation.

Website Evaluation

Evaluation of information systems can be classified as design evaluation and outcome evaluation (Torkzadeh et al. 2002). Website design has been examined from the lens of human-computer interaction which focuses on issues surrounding usability and interface design. Considerable previous research approaches website usability or interface design from a cognitive or empirical perspective. The cumulative findings seemingly point to some high-level design principles or recommendations.

Interface-Based Evaluation

Considerable prior research has taken a cognitive or empirical approach to investigate website usability or interface design. The cumulative findings have pointed to specific design principles or recommendations (Botafogo et al. 1992; Ivory et al. 2002; Nielsen 2000);. Zhang and Dran propose a hygiene and motivator model for website design and evaluation, particularly focusing on considerations pertinent to the interface design (Zhang et al. 2006). Ivory surveys automated website evaluation from the perspective of both research and practice, concluding that the development of automated website evaluation tools and methodologies is still in an infancy stage and requires continued investigative and evaluative efforts. The evaluation and improvement of website design have attracted considerable attention from both researchers and practitioners. The design of a website arguably is more challenging than that pertaining to conventional systems (Robertson et al. 1998). Despite the availability of new modeling tools and development methodologies, website evaluation and enhancement practices have advanced slowly.

Structure-Based Evaluation

Analytical modeling approaches to website design also have been studied. A survey of Web metrics examines fundamental graph characteristics that are relevant to website design and classify a set of essential metrics for quantifying Web graph properties, including page significance, page similarity, search and retrieval, usage characterization, and information theoretic properties for improved Web information access and use (Dhyani et al.). In exploring the Web graph properties, Wu et al. also investigate important Web graph properties by focusing on compactness and stratum which together provide high-level guidance to website designers (Wu et al. 2004). To get the ranking of pages, Brin models website structural designs with graph to get ranking of pages (Brin 1998). Monika and Ricca study the hyperlinks on the Web and compare different ranking strategies (Monika 2005; Ricca 2005). Web mining methods are also applied to investigate the content, structure, and usage. Kumar et al. develop a stochastic model for building a Web graph that consists of edges statistically dependent on each other and allows the creation of new vertices dynamically over time (Broder et al. 2000). Sarukkai analyzes prominent content-access patterns by visitors and develops a Markov Chain model for link prediction and path analysis (Sarukkai 2000). Zhou et al. examine individuals' visiting behaviors and patterns on a website to evaluate the design of its link structures (Zhou et al. 2001).

Improvement of Navigation Efficiency

Navigation efficiency is related to system level design, structure level design and page level design. System level characterizes the workload and CPU requirements of the Web application. Page design follows a series of guidelines of user interface to organize appropriate amount of content. Structure level design has been paid attention to in Web usage mining analysis. However, two basic directions, i.e. *customization* and *transformation* exist for improving the browsing efficiency (Perkowitz et al. 2000).

As we can see from last section, transformation has not been paid much attention in order to improve the Website structure. However, Website reorganization is necessary to ensure the logic design as well as to satisfy user's requirements.

Methodology

This study proposes a definition of navigation efficiency. Best, worst and average cases can be defined to assess navigation efficiency. For example, in an optimistic view, efficiency can be modeled as the shortest navigation path to the destination from the home page. In an pessimistic view, the navigation efficiency can be modeled as the longest path from each node to every other node. This paper applies the definition of average information accumulation performance for navigation efficiency. Metric used in this study is described and the correlations among corresponding metrics are illustrated. Several computations are illustrated for representing different possible navigation objectives. Website stability is taken into consideration to facilitate Website maintaining according to user's different requirements. Based on problem properties heuristics are developed to improve the efficiency objective function. Our method can be adaptable to different user defined specific constraints and stability constraints.

Definition of Navigation Efficiency

Navigation is one of very common tasks of web search activity, and information finding task (e.g. looking for product information or the name of a person) accounts for 25% of search activity (McEneaney 2001). In dictionary efficiency is defined as the quality or property of being efficient, or the degree to which this quality is exercised. It can be computed as the ratio of the effective or useful output to the total input in any system or the ratio of the energy delivered by a machine to the energy supplied for its operation. To define navigation efficiency, the average information accumulation in navigation is taken into consideration. The average performance is simply what the performance will most likely be, e.g. the average number of clicks which lead you from one page to another. In the definition of efficiency, the importance of the page, the information volume, and the page loading time / number of clicks are considered. Then related metrics are investigated. Their correlations are studied to get a better insight of the navigation efficiency. Navigation efficiency can be reflected from operation efficiency, or performance efficiency: the time needed to complete the task (Lee et al. 2004). Operating efficiency is the kind of efficiency associated with a group of users when browsing a website. The operating efficiency of a website is defined as a quantification of the amount of data conveyed through users' interactive browsing behavior. Efficiency = Shortest path (shortest path from initial page to target page) / user operating cost (Lee et al. 2004). Overall weighted distance between web pages has also been explored to make preferred pages more accessible using optimization method (Wu et al. 2004). In analytical modeling, efficiency can be transformed to mathematical formulations.

Definition and Notations:

A website consists of Web pages, which connect to each other through hyperlinks. The website can be modeled as a directed graph, G=(V, E). Vertices $V=\{v_1, v_2, \ldots v_n\}$, where v_i

(i=1,2,…n) denotes a page. Edges or arcs E={e_{ij} | the hyperlink from the source page i to the destination page j}. P is the set of ordered pares (i,j) such that there is a path from i to j, where each node is visited once. R is the set of ordered pares (i,j) such that there is a route user navigate from i to j.

Navigation efficiency of one route: Navigation efficiency is the ratio of information accumulated when navigating useful pages and all information accumulated when navigating all pages (including repeatedly visited pages) in one navigation route from i to j. Information accumulation efficiency f_{ij} on one route is defined as a_{ij} / l_{ij}. In this definition, a_{ij} is the information accumulated in necessary navigation to target page j, i.e. information accumulation in the shortest path from i to j. And l_{ij} is the average information accumulation in navigation length of the actual route user taken from i to j, i.e. number of pages (including repeated pages) visited between i and j. If all the information on pages are seemed as of the same importance, and navigation efficiency can be reflected from average path length. The amount of information can be represented by page size. In this study, the time length can be represented by number of links to achieve the target page when downloading time, page size, and viewing time are assumed to be similar among pages.

Navigation efficiency of whole website: Navigation efficiency is defined as average information accumulation efficiency of all routes the user taken to get to destination(s),

i.e. $\sum f_{ij} / |R|$, where $f_{ij} = a_{ij} / l_{ij}$; $a_{ij} = \sum_{i \in \text{shortest path}} c_i x_i$; $l_{ij} = \sum \sum_{i \in \text{actual route}} c_i x_i / |R|$; c_i is the weight of page i, and x_i equals to 1 when it appear in the corresponding route or path.

Sometimes, especially for experienced users, the target pages are already known. So the navigation efficiency can be average information accumulated from home page to a specific page; from all the other pages to one specific page; from home page to all the information (e.g. products or advertisement) pages; or from all the other pages to information pages. For new users, they are unfamiliar with the website. The information accumulation efficiency of one path can be seemed as depth-first-surfing with redundant surfing, or Hamiltonian surfing, where user visit different page only once (Chong et al. 1999). The impact or importance of page can be represented by popularity or visiting frequency, which can be calculated from log file. A lot of factors, e.g. semantic and syntax aspect of website structure will affect efficiency (Fang et al. 2007). In this study, we are concerned with structure oriented improvements and some structure metrics are illustrated here. The ingoing / outgoing degree is the number of edges entering / going out from the node. Depth is the distance from the root, which is the minimum number of links to follow in order to reach the destination. Centrality can be classified as relative out centrality (ROC) and relative in centrality (RIC) (Dhyani et al. 2002). Connectivity is the minimum number of edges that have to be removed from the graph in order to separate two nodes / groups (Ahuja 1993). Accessibility and average connected distance (ACD) are also taken into consideration in past study (Juvina et al. 2006; Yen 2007). Literature found that the *path compactness* will correlate significantly with search performance in a positive fashion, while *path stratum* in an inverse fashion (McEneaney 2001). In his study the correlation between computer experience and navigation performance (hypertext reading) score, however, is found to be weaker than above two path metrics. Stratum and compactness are not independent measures (Botafogo et al. 1992). Table 1 illustrated some of the correlation between metrics according to their definitions.

Table 1: Efficiency Related Web Metrics and Their Correlation

Correlation Table	Depth	Centrality	Connectivity
Depth		No	No
Centrality			Yes
Connectivity			

User Requirements and Stability Constraints

User requirements can be classified as functional requirements and nonfunctional requirements (Robertson et al. 1998). User also cares about different features in different Web application. Web design feature classifications have been studied. Past studies also propose the modeling of user requirements and transformation of requirements into generic design then further into graph problems (Yen 2007).

Stability of Web graph is taken into consideration to avoid that users may lose their orientation and to help designer control the maintaining efforts. A Web graph is stable if the structure does not change much and *structural stability* and *dynamic stability* should be taken into consideration (Bohringer 1990). *Structural stability* is concerned with user specified constraints. *Dynamic stability* if concerned with minimizing the difference between successive web structures (Tamassia et al. 1988). There are three categories for dynamic stability in Web structure reconfiguration. First, there is no structure change at all. Only other constraint, e.g. weigh of the link, of the Web structure will be modified to achieve an improved objective criteria. Second, limited number of links / nodes can be added / deleted to reorganize the structure. Third, the Web graph can be totally restructured. The first and the third are extreme cases. We will focus on the second one in this study. In terms of s*tructural stability* users can specify their requirements of structure constraints. For example, the degree of outgoing degree cannot increase or be limited within some range. Or, the depth of the graph cannot be increased. In terms of *dynamic stability,* we will try to improve the navigation efficiency while taking into consideration minimizing the changes. When limited number of links / nodes can be added / deleted, it becomes the most complex case and with many alternatives. For different category of changes, different scenario can be considered also. For example, to add links, different constraints might be taken into consideration, e.g. maintaining acyclic graph, reversing links allowed, and specific metric requirement. In this study, we only take into consideration the manipulation of links in a directed acyclic graph. Node modification and other extensions will be investigated in future research.

Web Efficiency Modeling and Improvement

We will transform the Web navigation with specific user requirements into a graph problem. In this study, we limit our study scope by bearing in mind two considerations for webmaster and users.

(1) Minimize additional work for webmaster. We will protect the site's original design from destructive changes. Limited number of links can be added or removed, and also limited number of pages can be created or destroyed.
(2) Make the Website transformation for everyone, especially first-time users and casual users. Customization is very useful for experienced users, but does not benefit first-time users (Perkowitz et al. 2000).

Modeling and Evaluation

In this study we want to improve web navigation efficiency while taking into consideration of website structure stability. Past studies also propose the modeling of user requirements and transformation of requirements into generic design then further into graph problems (Yen

2007). In the case where the constraints are not consistent, some constraints can be deactivated if they are assigned lower priority. Some studies suggests binary search through inconsistent set of constraints and all constraints causing an inconsistency are deactivated (Bohringer 1990). In our study, only consistent constraints are considered. To model optimistic, pessimistic, or average scenarios in navigation, information accumulation along shortest, longest, or average path length can be considered. When we would like to find the worst case in the navigation, a longest path of navigation may be taken into consideration. However, longest path problem has been proved to be NP and quite hard to be solved (Ahuja 1993; Garey 1979). Table 2 illustrates the four scenarios in navigation with target(s), i.e. Single-source and Single-destination (SS), Single-source and Multiple-destination (SM), Multiple-source and Single-destination (MS), and Multiple-source and Multiple-destination (MM). Two scenarios for no target navigation have also been modeled as well established problems.

Table 2: Efficiency Evaluation Model for Based on Optimistic Scenario

<table>
<tr><th>Experienced User Model</th><th>Complexity</th><th>Possible Solution</th><th>New User Model</th><th>Complexity</th><th>Possible Solution</th></tr>
<tr><td>SS</td><td>P</td><td rowspan="4">(Ahuja 1993; Garey 1979)</td><td rowspan="2">Depth-first navigation</td><td rowspan="2">P</td><td rowspan="4">(Chong et al. 1999)</td></tr>
<tr><td>SM</td><td>NP</td></tr>
<tr><td>MS</td><td>P</td><td rowspan="2">Hamiltonian surfing</td><td rowspan="2">NP</td></tr>
<tr><td>MM</td><td>NP</td></tr>
</table>

Efficiency Improvement

To improve efficiency of whole website navigation, we need to look into the efficiency of each route. By identifying the inefficiency routes, we may improve the overall efficiency while not decreasing the efficiency on other routes. To improve the efficiency of one route, we need to look at the average information accumulation along the route between two nodes of the route. Therefore, we need to add a link to create a path that reduces the route length. We may add the link which can reduce repeated path, trim down unnecessary / repeated information accumulation, and maintain important information. Therefore, heuristic may be developed to identify these kinds of links.

The information accumulated on shortest path will not change by adding a link. Literature shows that adding hotlinks (shortcut) may reduce average number of searching steps (Prosenjit et al. 2003). However, as number of link increases, the complexity of graph increases and the search time may not decrease by adding shortcuts (hotlink).

Adding, deleting, and hybrid methods can be applied to achieve structure with better average navigation efficiency. Due to page limit, we only consider adding a link. In this study, a branch-and-bound method implicitly enumerates all the feasible solutions, relaxing the integer constraints. The method can be very time-consuming, but is in principle capable of yielding an exactly optimal solution (Bertsekas 1998). The main idea of the branch-and-bound is to save computation by discarding the nodes / subsets of the tree that have no chance of containing an optimal solution. The branch-and-bound solution of a problem may be accelerated by adding side constraints to make the feasible set small without integer constraint (Bertsekas 1998).

Table 3: Efficiency Improvement Based on Branch-and-Bound

Initial Feasible Solution	Bounding Function	Live Solution Selecting Strategy	Branching Rule	Reference
Using current website structure if feasible or generate a new one using heuristic	Relaxing subtour constraint and use shortest path as lower bound	Best first search/ breadth first depth first search based on decreasing impact	Start from high impact neighbor or trade-off between impact and path length	(Beezer et al. 2000; Brusco 2005)

Examples and Result Analysis

Based on stability constraints, examples for navigation efficiency are given. Besides system updates and improvements, modifications in a HCI perspective also help improve the factors that contribute to navigation efficiency. To improve the navigation efficiency while modifying limited number of links (including adding and deleting limited number of links) with node degree constraints is a much more complicated case as the constraints vary in a lot of forms and some of the problem is NP hard.

Scenarios and Modeling

An e-banking scenario is illustrated as example. Internet users grow fast since e-banking launched. At the same time, banks have learned more about the benefits of e-banking and therefore are able to provide innovative services to their customers. Several studies have shown e-banking customers are very valuable because they make fewer service calls and are less likely to switch to a different provide. The ultimate success of e-banking services demands effective website designs and continuous improvement. In the following example, we illustrate the use of the proposed framework to evaluate an e-banking website and specific recommendations for service improvement. To make this example simple, the fictitious website is consists of two basic functions. One is account balance checking and the other is to apply for visa card. Node 1 is the home page of the website. Node 2 is a login page, which is mandatory before checking user's account at node 4. Node 8 is the detail transaction information of the account. Node 7 is the contact information of the bank. Node 3, 5, and 6 are fund products information. Node 9 is an advertisement page that appears after user finishes one function. We transform the scenario to a graph problem with a single objective. Multiple objectives are not considered in this study. The goal is to increase the navigation efficiency from homepage to advertisement page (node 9), so that the user can be kept in the website for a long time. Page size, accessing time, and page significance are known. The x_{ij} denotes the existence of edge e_{ij}. We assume that the structural stability constraint required by users is to constraint the outgoing degree to avoid overwhelming information load. In the following we will discuss the improvement of efficiency in different category of dynamic stabilities, i.e. stable website and no change on structure, limited changes, and unlimited changes.

Extreme Cases

There are some special web structures that no more links can be added, e.g. complete graph. The efficiency of such graph may be improved by adding new nodes by creating an index page or splitting current page. In this study we will not discuss the situation of adding nodes.

Example of Adding Links Using Branch and Bound

In this example, limited number of links can be added or deleted. Moreover, only sub-structure is allowed to be modified instead of whole structure. Link locality has been noticed and studied by several parties. Research implies that links within a site tend to be correlated to pages that are "nearby" in some measure (Eiron et al. 2003; Eiron et al. 2004; Watts et al. 1998). Therefore, sub-graph modification is taken into consideration in this study. In the scenario the objective is to achieve the maximum average efficiency while limiting the degree and depth of the node. Navigation efficiency can be reflected from minimum average path length problem (Beezer et al. 2000; Broder et al. 2000). When we take into consideration of link impact and path length, the following heuristic is developed for adding links to improve navigation efficiency from homepage to one destination node. If user set a maximum outgoing degree of 4, then no links is allowed to be introduced from homepage. If maximum number of added links is constrained, only those along paths with best trade off between impacts and path length will be chosen. Adding more links may not increase efficiency as the some length is increasing as adding more links.

Table 4: Components in Heuristic Development for Adding Links

Stop Criteria	No new path can be found, whose efficiency is higher than existing one. Level of links and paths are updated after each reconfiguration.
Adding limited number of links	Add the link which can reduce repeated path, trim down unnecessary / repeated information accumulation, and maintain important information.
Constraint of depth	Links can only be added not to increase depth.
Constraint of degree	Start from the node which has the lowest outgoing degree. Links can only be added not to increase outgoing degree.
Constraint of sub-graph	Links can only be added from one node to adjacent node without violating other constraints.

The navigation efficiency improvement problem is formulated as

$$Max\ |N| / \sum_{\substack{i \in leave \\ i \in P}} p_i a_i \quad \text{or} \quad Min \sum_{\substack{i \in leave \\ i \in P}} p_i a_i \tag{1}$$

Subject to

$$\sum_{v_j \in A(v_i)} x_{ij} \le d_i, i = 1,\ldots,n, \tag{2}$$

$$z_{ij} \le x_{ij}, \forall (i,j) \in E, \tag{3}$$

$$\sum_{v_j \in B(v_i)} z_{ij} = 1, j = 2,\ldots,n, \tag{4}$$

$$\lambda_j + m(1 - z_{ij}) \ge \lambda_i + 1, \forall (i,j) \in E, \tag{5}$$

$$\lambda_i \le l, \forall v_i \in V, \tag{6}$$

$$x_{ij}, z_{ij} = 0,1; \lambda_i \ge 0, \tag{7}$$

The d_i denotes the maximal allowable outgoing degree for nodes v_i and $a(v_i)$ denotes the set of succeeding nodes of v_i. λ_i denotes the level of v_i and l determine the depth. Constraint (2) is to limit the outgoing degree of the node. Constraint (3) (4) is to maintain a connected structure. Constraint (5) is to ensure that level of parent node is smaller than that of child node. Constraint (7) is to constraint the depth of the whole graph. To maximize the navigation efficiency from homepage to advertisement (node 9), it is obvious that links may be added to improve the overall performance. Branch-and-bound method is applied here to guarantee the optimal solution of the improvement method. J_0=and Ø, J_1= Ø are disjoint subsets of link set.

J_0 is the subset of feasible solutions such that link (u,v) existing in the solution. With constraint on outgoing degree, no link can be added from homepage, therefore the structure with link (1,9) is not feasible. We identify the high impact node 3, 4, 5 and 7. However node 3, 5 and 7 are not on the same branch of node 9, which is defined as a sub-graph, therefore no links will be added. Without stability constraint on outgoing degree and sub-graph, only trade-off between cumulated impact and length will be considered. We can start from the node which has the lowest outgoing degree. Based on all these constraints we can add links to achieve a minimum average search steps using Branch and Bound method. Figure 1 illustrates the result with / without stability constraints on outgoing degrees, sub-graph constraints, and maximum number of adding links.

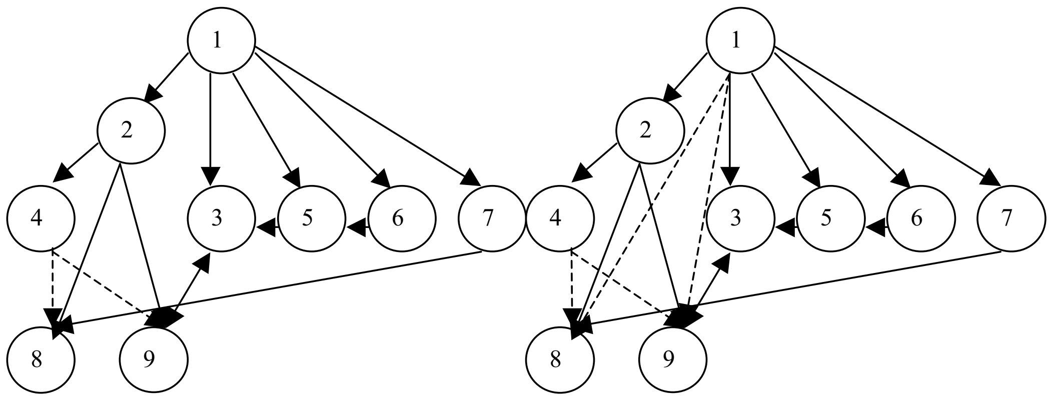

Figure 1: The Structure after Adding Links with / without Stability Constraints (degree≦5)

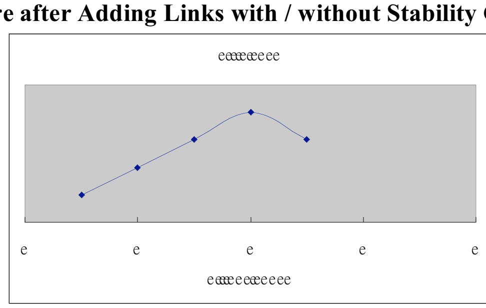

Figure 2: The Correlation between Efficiency and Number of Links Added

Figure 2 illustrates that as number of link increases, the complexity of graph increases and the search time may not decrease by adding shortcuts (hotlink). Some improvements are significant but will saturate or even decrease as after adding a certain number of links.

Discussions

The trade-off between stability and navigation efficiency is necessary for achieving desirable performance while confirming to requirements. However the problem of improve navigation is not an easy problem. Some instances of addling links have been proved to be NP-hard and the trade-off between stability and efficiency is necessary in order to achieve a good design.

Proof of Problem Complexity

The problem of minimizing average search length is a very complicated problem. Hotlink (shortcut) assignment problem in some instance has been proved to be NP-hard in the shortest path formulation (Bose et al. 2000; Czyzowicz et al. 2003). Even adding one optimal link is NP-hard for arbitrary directed acyclic graph with uniform distributions (Czyzowicz et al. 2003).

Structural Stability and Efficiency

We investigate the outgoing degree constraint of structural stability and its correlation with navigation efficiency. The efficiency value might saturate as the outgoing degree increase and number of paths increase. The outgoing degree of web graph is defined as the maximum degree of all nodes. And the depth is the maximum depth of all nodes.

Though branch-and-bound method ensures the optimal result by implicitly enumerating all feasible solutions, significance is still need to be investigated in different scenario, e.g. different website structures (applications), and different requirements.

Conclusions

This study aims to improve Web navigation efficiency by reorganizing Web structure. To avoid users to lose their orientation, structure stability is taken into consideration. This study proposes a mathematical programming method to reorganize Web structure in order to achieve better navigation efficiency. User can specify the requirements and how stable the website structure should be. Navigation efficiency is defined systematically and an e-banking example is given to illustrate how the method works. This study has the advantage of improving navigation efficiency mathematically and relieving the designer of tedious chore to modify the structure in transformation. If some constraint, e.g. structural stability constraint, is relaxed, there is still room for improvement. Some improvements are significant but will saturate or even decrease as after adding a certain number of links.

Research Contribution and Managerial Implications

We proposed a quantitative method to improve navigation efficiency systematically. From our problem solving, we identify some potential important research areas that has not been explored or paid attention to. For the website designer in the company, it is very important to evaluate the efficiency quantitatively and to improve the IS design continually. Moreover, to improve a website should be put on schedule after the assessment if the result is not satisfactory.

Limitations and Future Directions

This study adopts the approaches for measuring the performance using average-case analysis, which has some major draw backs. First, the analysis depends crucially on the probability distribution chosen to represent the problem instance. Second, the appropriate probability distribution, which is got from usage data, is difficult to be obtained or determined in practice. Third, the analysis may require quite intricate mathematics even for assessing simple algorithm (Ahuja 1993). Moreover, mathematical proof is expected to illustrate the feasibility of heuristic to improve the navigation efficiency. In particular, following extensions can be taken into consideration. First, multiple goals can be considered in the future. Second, pages can be merged or spitted in order to modify the structure. And adding, deleting, and modifying links should be considered at the same time. Moreover, similar methodology can be applied in Web page layout. Finally, simulation can be used for the evaluation to justify the significant level of the improvement empirically. The discussion of models of with specific property (e.g. monotonic property), different web structure, and sensitivity analysis are expected also.

References

Ahuja, R.K. *Network flows : theory, algorithms, and applications* Prentice Hall, Englewood Cliffs, N.J. :, 1993, pp. xv, 846 p. :.

Beezer, R.A., Riegsecker, J.E., and Smith, B.A. "Using minimum degree to bound average distance," *Discrete Mathematics* (226:1-3) 2000, pp 365-371.

Bertsekas, D.P. *Network optimization : continuous and discrete models* Athena Scientific, Belmont, Mass. :, 1998, pp. xiii, 593 p. :.

Bohringer, K.F. "Using Constraints to Achieve Stability in Automatic Graph Layout Algorithms.," Proceedings of the ACM CHI 90 Human Factors in Computing Systems Conference, Seattle, Washington,USA, 1990, pp. 43-51.

Bose, P., Jurek Czyzowicz, Leszek Gasieniec, Evangelos Kranakis, Danny Krizanc, Andrzej Pelc, and Martin, M.V. "Strategies for Hotlink Assignments " The Eleventh Annual International Symposium on Algorithms and Computation, 2000, pp. 23-34.

Botafogo, R.A., Rivlin, E., and Shneiderman, B. "Structural analysis of hypertexts: Identifying hierarchies and useful metrics," *ACM Transactions on Information Systems* (10:2) 1992, pp 142-180.

Brin, S. "The anatomy of a large-scale hypertextual Web search engine 1," *Computer Networks* (30:1-7) 1998, pp 107-117.

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., and Wiener, J. "Graph structure in the Web," *Computer Networks* (33:1), 15 May 2000 through 19 May 2000 2000, pp 309-320.

Brusco, M.J. *Branch-and-bound applications in combinatorial data analysis* Springer, New York :, 2005, pp. xii, 221 p. ;.

Chong, C.W., Ramachandran, V., and Eswaran, C. "Web navigation efficiency analysis," Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, IEEE, Tokyo, Jpn, 1999.

Czyzowicz, J., Kranakis, E., Krizanc, D., Pelc, A., and Martin, M.V. *Journal of Web Engineering* (1:2) 2003, pp 93-127.

Dhyani, D., Ng, W.K., and Bhowmick, S.S. "A survey of Web metrics," *ACM Computing Surveys* (34:4) 2002, pp 469-503.

Eiron, N., and McCurley, K.S. "Locality, Hierarchy, and Bidirectionality in the Web " Workshop on Algorithms and Models for the Web Graph, Budapest, 2003.

Eiron, N., McCurley, K.S., and Tomlin, J.A. "Ranking the Web frontier," Thirteenth International World Wide Web Conference Proceedings, WWW2004, New York, NY, 2004, pp. 309-318.

Fang, X., and Holsapple, C.W. "An empirical study of web site navigation structures' impacts on web site usability," *Decision Support Systems* (43:2) 2007, pp 476-491.

Garey, M.R. *Computers and intractability : a guide to the theory of NP-completeness* W. H. Freeman, New York :, 1979, pp. x, 338 p. :.

Guillermo, J.C., and Luis, A.O. "Assessing quality in use in a consistent way," in: *Proceedings of the 6th international conference on Web engineering*, ACM Press, Palo Alto, California, USA, 2006.

Ivory, M.Y., and Hearst, M.A. "Improving web site design," *IEEE Internet Computing* (6:2) 2002, pp 56-63.

Juvina, I., and van Oostendorp, H. "Individual differences and behavioral metrics involved in modeling web navigation," *Universal Access in the Information Society* (4:3) 2006, pp 258-269.

Lee, J.H., and Shiu, W.K. "An adaptive website system to improve efficiency with web mining techniques," *Advanced Engineering Informatics* (18:3) 2004, pp 129-142.

Luis, O., and Gustavo, R. "Measuring Web Application Quality with WebQEM," *IEEE MultiMedia* (9:4) 2002, pp 20-29.

McEneaney, J.E. "Graphic and numerical methods to assess navigation in hypertext," *International Journal of Human Computer Studies* (55:5) 2001, pp 761-786.

Monika, H. "Hyperlink analysis on the world wide web," in: *Proceedings of the sixteenth ACM conference on Hypertext and hypermedia*, ACM Press, Salzburg, Austria, 2005.

Nielsen, J. *Designing web usability* New Riders, Indianapolis, Ind. :, 2000, pp. xiii, 419 p. :.

Perkowitz, M., and Etzioni, O. "Towards adaptive Web sites: conceptual framework and case study," *Artificial Intelligence* (118:1-2) 2000, pp 245-275.

Prosenjit, B., Danny, K., Stefan, L., and Pat, M. "Asymmetric Communication Protocols via Hotlink Assignments," *Theory of Computing Systems* (36:6) 2003, p 655.

Ricca, F. "Hyperlinks analysis in multilingual web applications," Proceedings - Seventh IEEE International Symposium on Web Site Evolution, WSE 2005, Budapest, 2005, pp. 57-62.

Robertson, G., Czerwinski, M., Larson, K., Robbins, D.C., Thiel, D., and van Dantzich, M. "Data Mountain: Using spatial memory for document management," UIST (User Interface Software and Technology): Proceedings of the ACM Symposium, Association for Computing Machinery, San Francisco, CA, USA, 1998, pp. 153-162.

Sarukkai, R.R. "Link prediction and path analysis using Markov chains," *Computer Networks* (33:1), 15 May 2000 through 19 May 2000 2000, pp 377-386.

Tamassia, R., Di Battista, G., and Batini, C. "AUTOMATIC GRAPH DRAWING AND READABILITY OF DIAGRAMS," *IEEE Transactions on Systems, Man and Cybernetics* (18:1) 1988, pp 61-79.

Torkzadeh, G., and Dhillon, G. "Measuring Factors that Influence the Success of Internet Commerce," *INFORMATION SYSTEMS RESEARCH* (13:2), June 1, 2002 2002, pp 187-204.

Watts, D.J., and Strogatz, S.H. "Collective dynamics of 'small-world9 networks," *Nature* (393:6684) 1998, pp 440-442.

Webster, J., and Ahuja, J.S. "Enhancing the design of Web navigation systems: The influence of user disorientation on engagement and performance," *MIS Quarterly: Management Information Systems* (30:3) 2006, pp 661-678.

Wu, Y., Wang, Y., and Dai, W. "Quantitative Analysis of Website Based on Web Graph Theory," Proceedings the Eighth Pacific Asia Conference on Information Systems(PACIS 2004),, July 8-11, Shanghai, China. , 2004.

Yen, B.P.C. "The design and evaluation of accessibility on web navigation," *Decision Support Systems* (42:4) 2007, pp 2219-2235.

Zhang, H., and Jiao, L. "Usage of user navigation pattern clustering in web personalized recommendation," *Jisuanji Gongcheng/Computer Engineering* (32:15) 2006, pp 64-66.

Zhang, P., Dran, G.v., Blake, P., and Pipithsuksunt, V. "A Comparison of the Most Important Website Features in Different Domains: An Empirical Study of User Perceptions," Proceedings of Americas Conference on Information Systems (AMCIS'2000), Long Beach, CA, 2000, pp. 1367-1372.

Zhou, B., Chen, J., Shi, J., Zhang, H., and Wu, Q. "Website link structure evaluation and improvement based on user visiting patterns," Proceedings of the ACM Conference on Hypertext, 2001, pp. 241-242.