ISSN 2515-8260 Volume 10, Issue 3, Winter 2023

A CONSOLIDATED WEB EXPERIENCE WITH FEWER PAGES TO NAVIGATE FOR BETTER OUTCOMES

Sai Vamsheedhar Reddy Kuthuru[1], Pallam Ravi[2]
[1]PG Student Department of CSE, Anurag University, Hyderabad, Telangana
[2]Assistant professor, Department of CSE, Anurag University, Hyderabad, Telangana
[1]saivamshi4@gmail.com, [2]Satishpallam@gmail.com

ABSTRACT:

For a long time, it has been difficult to design websites with a clear structure that users can easily navigate. One major cause is the gap that might exist between web developers' and consumers' views on the optimal structure for a website. Although many techniques have been presented for relinking web sites to enhance their navigability by analyzing user navigation data, the totally rearranged new structure may be very unexpected, and the cost of confusing people after the changes has not been quantified. In this article, we will discuss how to tweak a website without completely overhauling it. To enhance user navigation on such a website while making as little changes as possible to the site's present structure, we offer a mathematical programming approach. Extensive testing on a real-world data set that is available to the general public shows that our approach not only improves user navigation dramatically with minor adjustments, while also being efficiently addressed. Further, we have shown that the model scales up well by testing it on extremely large synthetic data sets. We also create two measures for assessment and employ them to measure the enhanced website's effectiveness on a live data set. These evaluations show that the new structure is much easier to navigate for users. Interestingly, we discover that users who are more severely bewildered are now more likely to experience the benefits of the new structure.

Keywords: ***Website design, user navigation, web mining, mathematical programming***.

INTRODUCTION:

As time goes by, the internet will become a more integral aspect of everyone's daily lives. The Internet is often used as a data source. Words, pictures, charts, videos, and other media are all used to convey the same data. The material should be easily accessible, meaning that the website's navigation was straightforward. Website navigation is significantly influenced by the design of the site's architecture. There are two main ways to enhance the user experience while navigating a website: personalization and transformation. Through the process of web customization, consumers may have their web experience tailored to their own needs. Improving web navigation for any and all users is a key part of any overhaul. The two strategies are quite distinct from one another. Personalizing a user's experience on the web requires taking their individual preferences into account. The effects of any modifications performed may vary depending on the user. To have efficient user navigation, one must undergo a web transformation, which requires taking into account aggregate user log data. The process of customizing a website takes a lot of time and processing power. Unlike web personalization, which relies on users' browsing histories, web transformation doesn't ask visitors for any background information. A web personalization strategy is ideal for dynamic websites with frequent data updates, whereas a web transformation strategy is more suited to static websites with infrequent data changes. There have been a number of methods discussed for tailoring the web, and some of them include the automated synthesis of index pages

so order to add connections for frequently visited sites. Another method that has been examined extensively involves clusters of user profiles that are then utilized to construct connections on the fly for each individual user or set of users. We suggest a method that, with little structural modifications to the website, makes it easier to navigate for users. A mathematical model is employed to modify the website's link structure. In order to develop a solution that varies somewhat from the original web structure, you feed the algorithm the web data structure and logs.

LITERATURE SURVEY:

To maximise the user experience while incurring the fewest possible modifications towards the site's infrastructure, Min Chen and Young U. Ryu [1] suggested a mathematical programming paradigm to achieve this goal. The concept worked well for static informative webpages. When website performance was enhanced rather than reorganised, it becomes more amenable to regular maintenance. Problems of enormous scale are quickly resolved using the mathematical programming approach. To help users navigate more easily, based on the co-occurrence among pages in their traversal, Perkowitz and Etzioni [02] present a technique that can dynamically Synthesize index pages. In fact, this is the essence of online customization. Mobasher et al. [3, 4], and Yan et al. [6] suggested a technique to generate clusters with user profiles derived through weblogs as well as to make linkages for users divided into various groups according to their access behaviour followed among them. It's another approach that relies on customising your experience on the web.

The Among the oldest and most important design principles for websites is easy navigation [7], and this holds true across all industries, from business and finance to media and commerce to education and healthcare. Both qualitative and quantitative studies have been conducted on the topic of website design effectiveness [8]. Several theoretical approaches to the study of website design have been published, but the field has often been classified as either hypermedia or perhaps a database. From an engineering point of view, both web engineering as well as IDEAL take into account modelling and subsequent improvement. Graphical aspects, page layout, including usability research and analysis each play crucial roles in human computer communication as they pertain to website design. Website performance is influenced by system design factors including hardware design, cache scheduling, etc. Website usability is impacted by structural design choices like the arrangement of connections and the organisation of content.

Jia-Ching Ying, Chu-Yu Chin, and Vincent S. Tseng [9] offer a model called Ideal-Tree (i.e., Inverted data- basis Expectable Tree) that eliminates the need to search the database. Regarding web navigation mining, another Ideal-Tree Miner method is developed, which uses a variable threshold. Using the found patterns, a navigation prediction model is created. Similarity-based method was also suggested by Dean and Henzinger. It's possible that the algorithm's page source, which includes exclusively of the specified page's siblings, ignores potentially relevant sites that are semantically related. Cocitation degrees are utilized to determine the degree of resemblance between two pages by determining the number of shared parent pages between the two. Relevant pages are those that have a high cogitation index with the specified page. Although the procedure for determining which pages are important is straightforward in the case of deeper relationships, it is inefficient.

ARCHITECTURE:

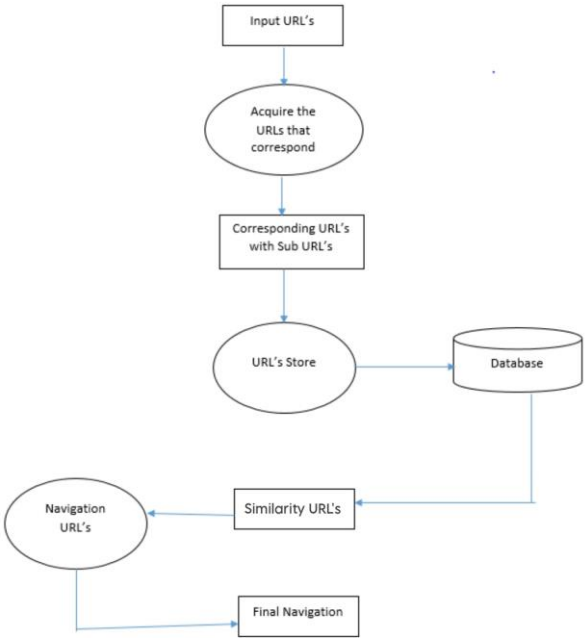

Fig: Architecture system.

PROPOSED APPROACH

Users of something like the new website might feel disoriented due to a thorough restructure that might drastically alter the position of previously accessible features. From the forum's main page to individual threads, there will always be links that visitors may follow implicitly. Costs associated with users being disoriented as a result of the restructured website structure have not been calculated. This happens because, whenever a website is redesigned, the original structure, which was likely developed by professionals and has commercial or organizational logic, could no longer exist. Furthermore, there have been no previous studies to evaluate the usability of either a totally rebuilt website, casting doubt upon that viability of the reorganization strategies. Last but not least, efforts to enhance the website's routing performance through rearrangement strategies are impractical on a regular basis since they risk disrupting the site's existing structure.

Following are the Benefits to Be Obtained from the System Under Consideration:

- Avoids overhead and URL type reorganization problem.
- Provides highly precise index URL.
- No Data loss and Accuracy is maintained.
- Low Cost with time save

Research Methodology

After that, we may apply the following mathematical programming model to solve the issue of how to make a website more user-friendly while making as little modifications as possible to its present structure.

$$\sum \chi_{ij}[1-\lambda_{ij}(1-\in)]+m\sum_{i\varepsilon NE} p_i$$

Improving the website's structure has two parts, and the goal function's job is to decrease both of them.

1) the required quantity of fresh connections (the first summation).

2) The new structure's penalty for pages with an abnormally high number of links (those with more than the out-degree threshold) (the second summation). Our research has shown that consumers may overlook some existing links owing to confusing labelling or layout.

Mini Session and Target Identification

For this reason, I used the page-stay timeout heuristic for identify users' goals and distinguish across sessions. A natural assumption is that people would spend more time upon that designated pages. Numerous studies have revealed that the average amount of time that user spends on a page is indeed a solid absolute indicator of both the page's or document's perceived importance towards the user. The page-stay timeout heuristic and many other time-placed heuristics were frequently employed for session detection in the context of online use mining, and they have shown to be rather challenging with regard to difference of both the threshold values. The setting of page-stay timeout threshold may affect the detection of target pages during micro sessions. We conducted our studies using varying threshold values since it is notoriously difficult to reliably identify micro sessions using unknown user access data.

OBSERVATIONS ON REDUCTION OF PROBLEM SIZE

Formulation includes E binary variables indicating the number of candidate connections and TR restrictions representing the number of applicable mini sessions. However, it turns that within the context with our issue, the formulation may be simplified to a substantially smaller one that could be swiftly handled, despite the fact that in actuality overall size of such a website as well as the number of micro sessions collected from server logs might be extremely high. Regarding the scope of the issue, we provide a number of comments. Collectively, these findings provide light regarding why the issue size in proposed formulation may be drastically decreased, which in turn helps to explain the quick solution times seen in our studies. As we will see, our formulation actually makes allowances for a smaller issue size.

RELEVANT MINI SESSIONS

Mini sessions are only important if their duration exceeds the route threshold. Thus, my MP model eliminates many unnecessary mini sessions (T^I) since only important mini sessions should be evaluated for improvement. In these other words, define $T^I = T \backslash T^R$, every micro session cannot be included in our formulation since user navigation within S already satisfies the aim (set as path threshold).

Path threshold may affect relevant mini sessions, as illustrated later. Enhancing the route threshold reduces meaningful mini sessions, whereas reducing it increases them. For the actual data set utilized in the studies, increasing the route threshold (b) from 3 to 5 decreases the number of relevant mini sessions from several thousand down a few hundred. Also when b = 1, many irrelevant mini sessions may be avoided.

RELEVANT CANDIDATE LINKS

As possible linkages among all pages of a website using node set N. This decision problem may examine every connection from E0 without preprocessing, yielding N X N linkages (variables). Even a tiny website might have a huge quantity. E^I links cannot increase user navigation intuitively. Candidate links (E) may be gathered from mini sessions to assist users browse. Thus, in the best solution x(i,j) = 0. Thus, no candidate links don't really aid user navigation and shouldn't be considered.

	n_1	n_2	n_3	n_4	n_5	n_6
n_1	0	0	0	1	1	0
n_2	1	0	0	1	1	1
n_3	1	1	0	0	0	0
n_4	1	1	1	0	1	0
n_5	1	1	1	1	0	1
n_6	0	1	1	1	0	0

Fig: The Connectivity matrix

Two considerations make many candidate connections irrelevant to the judgement. First, assuming path thresholds, designate the E^{RM} and E^{IM} candidate linkages for relevant and non-relevant mini sessions, respectively. Thus, $E^{IM}\backslash E^{RM}$ candidate links are really for irrelevant mini sessions that don't require improvement and may be ignored. Second, E^{RM} may not consider all potential linkages. For such a mini session having path threshold b, a link was significant to S if adding/improving link helps the user during S reach the destination in b pathways, i.e., accomplish the user navigation objective in S. Thus, potential links from sites seen during the both route or previously are significant to the choice. S cannot meet the user navigational objective by picking the other potential links.

COMPARISON WITH A GREEDY ALGORITHM

Many methods have been offered to reorganize a website, but none have optimally eliminated connections to minimize web page and structure complexity. Thus, we cannot compare earlier research since their objective functions or restrictions employ different measurements. We use the efficient greedy dynamic least-frequency (DLF) technique to compare the MP model. DLF was two-step. First, no traversed links were removed. No traversed links may be deleted without impacting user sessions. Reduces out degrees on relevant and irrelevant nodes. Thus, certain important nodes might well be improved with just a lower out degree meeting the threshold, making them irrelevant and eliminated from consideration inside the following stage. After this phase, relevant nodes and superfluous linkages remain nontrivial. Thus, we must proceed by deleting links from the visited relevant links collection.

Inside the second stage of the DLF method, a loop analyses the session frequency of every traversed relevant connection and picks the link only with lowest session frequency should delete. When one link is withdrawn, it really is added into set E^I and its impacted sessions to set T^I: Those

ISSN 2515-8260 Volume 10, Issue 3, Winter 2023

two sets of links and sessions were unsuitable for future inclusion to avoid "double counting" in later cycles.

Comparing As previously, for such a fair comparison, they omit the portion of the result attributable to the deletion of no traversed connections since this step is common for both with us model as well as the DLF method and will not affect their real performances. Therefore, we again examine just the net excessive links deleted, i.e., overall traversed link removal result. In those other words, however many superfluous links each approach removes after removing all no traversed connections. While in column "Performance imp. using MP model (%)," the result is calculated:

$$\frac{\text{Net excessive links removed by MP model --- by DLF}}{\text{Net excessive links removed by DLF}}$$

The values reported indicate the number of additional excessive links removed by our MP model over the DLF algorithm (%) under different session impact thresholds (σ) and outdegree thresholds (d). .

Overall, the result shows our method greatly outperforms the greedy algorithm in all categories and the margins are over 50% in many cases

EVALUATION

We execute approximations of its real-world use in addition to doing large-scale computer studies on actual and synthetic data sets that determine whether the enhanced structure's navigational efficacy is, in fact, strengthening. More particular, we divide the whole dataset together into training set (covering the first three months) and just a testing set (the rest of the data) (last month). And used the training data, we develop the enhanced structure, and afterwards evaluate based on the testing data utilizing two metrics: overall average number of pathways per mini session, and indeed the proportion of mini meetings that amplify to a determined threshold. The very first statistic determines whether the new structure helps users reach their goals more quickly than the old one does on average, while the second metric determines if indeed the site's structural changes are likely to pique the attention of users who are having trouble navigating the site. The very first measure has a three-stage evaluation process, which is outlined below:

- To get the new set connections and links to really be improved, apply this MP model towards the training data.
- Mini sessions with two or more upgradeable paths should be identified from test data, together with the length of those paths and indeed the candidate connections that may be utilized to implement those upgrades.
- The findings from either the training data should be compared to the connections acquired during Step 1 to see whether certain candidate links collected in Step 2 match.
- If the answer to this question is affirmative, then perhaps the upgraded website's source node session duration information should be obtained, and it should be assumed that users would negotiate either new link or perhaps the strengthened connection in the better structure.

ISSN 2515-8260 Volume 10, Issue 3, Winter 2023

EXPERIMENTAL RESULTS:

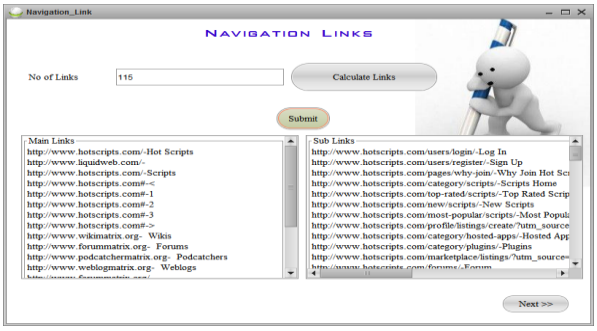

Fig: Navigation Links

EACH URL EDGE IDENTIFICATION

url	indexs
http://www.hotscripts.com	http://www.hotscripts.com/most-popular/scripts/-...
http://www.hotscripts.com	http://www.hotscripts.com/profile/listings/create/?...
http://www.hotscripts.com	http://www.hotscripts.com/category/hosted-apps/-...
http://www.hotscripts.com	http://www.hotscripts.com/category/plugins/-Plugins
http://www.hotscripts.com	http://www.hotscripts.com/marketplace/listings/?u...
http://www.hotscripts.com	http://www.hotscripts.com/forums/-Forum
http://www.hotscripts.com	http://www.hotscripts.com/blog/-Blog
http://www.hotscripts.com	http://www.hotscripts.com/pages/advertisers/-Adv...
http://www.hotscripts.com	http://www.hotscripts.com/contact/-Contact Us

url	indexs
http://www.forummatrix.org	http://www.forummatrix.org/show/JForum-
http://www.forummatrix.org	http://www.forummatrix.org/show/Jitbit-AspNetF...
http://www.forummatrix.org	http://www.forummatrix.org/show/Jive-Forums-
http://www.forummatrix.org	http://www.forummatrix.org/show/Lokboard-
http://www.forummatrix.org	http://www.forummatrix.org/show/MercuryBoard-
http://www.forummatrix.org	http://www.forummatrix.org/show/MetaForum-
http://www.forummatrix.org	http://www.forummatrix.org/show/miniBB-
http://www.forummatrix.org	http://www.forummatrix.org/show/Monkey-Boards-

Fig: Edge Identification

1	0	1	0	0	0	0	0	0
0	1	0	1	0	0	0	0	0
0	0	1	0	1	0	0	0	0
0	0	0	1	0	1	0	0	0
0	0	0	0	1	0	1	0	0
0	0	0	0	0	1	0	1	0
0	0	0	0	0	0	1	0	1
0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1

Fig: Confusion Matrix

CONCLUSION:

This study addresses a significant gap in the literature by proposing a mathematical programming model for increasing the efficiency of website navigation with little disruption to the site's existing structure. Our concept works well for ever-constant informative websites. Because it enhances rather than restructures a website, it may be used for routine upkeep. The results of our model's testing on a real website shown that with very few more connections, the site's navigation may be vastly improved for users. Rapid optimization results are promising for the model's applicability to real-world websites. We have also validated the MP model using many simulated data sets that really are substantially bigger than the biggest data set addressed in previous research and the actual data set itself. It was found that the MP model scales up quite effectively, with optimum solutions being found for large-sized issues in a matter of seconds on such a desktop PC.

We've established two criteria and put them to good use in a simulated evaluation of the enhanced website to verify our model's effectiveness. Our tests demonstrated that the revised layouts significantly simplified site navigation. In addition, we discovered an encouraging finding suggesting that visitors who are more disoriented those with a greater likelihood of giving up on the website are more likely to profit from the revamped layout than those who are less confused. The experiments also showed that lowering the route threshold might improve the findings, but would greatly increase the number of new linkages. Therefore, while defining acceptable route thresholds, Webmasters have to carefully weigh the tradeoff among desired improvements towards the user navigation and indeed the amount of new links required to achieve the job. We likened our model to a heuristic because no previous research has looked at the same goal as we have. Compared towards the heuristic, my model only required a fraction of the number of new linkages to obtain the same or better results.

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. M. Perkowitz and O. Etzioni, "Towards Adaptive Web Sites: Conceptual Framework and Case Study," Artificial Intelligence, vol. 118, pp. 245-275, 2000.

3. M. Eirinaki and M. Vazirgiannis, "Web Mining for Web Personalization," ACM Trans. Internet Technology, vol. 3no. 1,pp. 1-27, 2003.

4. 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.

5. B. Mobasher, R. Cooley, and J. Srivastava, "Automatic Personalization Based on Web Usage Mining," Comm. ACM, vol. 43, no. 8,pp. 142-151, 2000.

6. B. Mobasher, R. Cooley, and J. Srivastava, "Creatin Adaptive Web Sites through UsageBased Clustering of URLs," Proc. Workshop Knowledge and Data Eng. xchange, 1999.

7. 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.

8. 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.

9. Jia-Ching Ying, Chu-Yu Chin,Vincent S. Tseng "Mining Web Navigation Patterns with Dynamic Thresholds for Navigation.

10. Pingdom, "Internet 2009 in Numbers," http://royal.pingdom. Com/2010/01/22/internet-2009-in-numbers/, 2010.

11. J. Grau, "US Retail e-Commerce: Slower But Still Steady Growth," http://www.emarketer.com/Report.aspx?code=emarketer_ 2000492, 2008.

12. Interne retailer, "Web Tech Spending Static-But High-for the Busiest E-Commerce Sites," http://www.internetretailer.com/ daily News.asp?id = 23440, 2007.

13. D. Dhyani, W.K. Ng, and S.S. Bhowmick, "A Survey of Web Metrics," ACM Computing Surveys, vol. 34, no. 4, pp. 469-503, 2002.

14. 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.

15. J. Lazar, Web Usability: A User-Centered Design Approach. Addison Wesley, 2006.

16. 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.

17. J. Palmer, "Web Site Usability, Design, and Performance Metrics," Information Systems Research, vol. 13, no. 2, pp. 151-167, 2002.

18. 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

19. Reddy, G. Raghupal, and G. Radha Devi. "Security Privacy Content and Impact of Trust in Social Networks." International Journal of Advanced Information Science and Technology (IJAIST), vol. 6, issue 11, pp. 394-398.

20. G. Raghupal Reddy, and A. Shravan Kumar "An Automated Research For Analysing And Visualising The Data Using Machine Learning Analytics" Journal Of Critical Reviews, Vol. 06, Issue 05, pp. 470-475.