A Framework to improve the Web Performance using Reorganization, Optimized Prediction and Prefetching

K. Shyamala, S. Kalaivani, A. Murugan

Abstract:Improving web performance is becoming more hectic in recent days. This paperelucidates the combination of many ideas to improve web performance and given as a framework. The entire framework depicts various aspects in improving web access performance which includes website reorganization, webpage prediction and prefetching, optimized way of accessing prediction algorithm in webserver and finally improvements in a proxy cache at the time of accessing dynamic content. Each portion of the framework has been successfully proposed and implemented. The various algorithms have been introduced in each portion of the implementation. This research work not only introduced new algorithms but also create scope for further research works in terms of improving web performance.

Keywords: Web performance, Website reorganization, Prediction and prefetching, Proxy cache, Dynamic content.

I. INTRODUCTION

Data mining is nothing but the services which help to mine most relevant information from lakhs of webpages, social networking websites like facebook, twitter, blogging, instant messages etc. Data mining is considered as an important tool inbusiness process and it uses many machine learning algorithms to find exact information from a large database which exists in this modern world. As a result of data increasing, it cannot avoid the contract with trustworthy data mining companies.

Now a day's Data mining performs various tasks in different fields [1]. Web mining is one among the applications of data mining. Web mining is the process of integrating information collected through datamining techniques and methodologies with information collected through the World Wide Web. It identifies or discovers interesting patterns from huge datasets. Web mining is mainly divided into three broad areas (i.e.) web structure, web content and web usage mining.

Since, increasing is the usage of the World Wide Web; web mining has become more popular and important topic in research. Google Analytics is one of the web mining tools which will be helpful to track the information about web traffic. Real-time applications of web mining which includes analysis of the website, analysis of data in website, web user behavior analysis.E-Commerce application plays a significant role in web mining [2].

Web usage mining targets to predict web user behaviors when they communicate with the World Wide Web. Web user navigation can be discovered from reliable secondary data. Web user access pattern can be automatically predicted by using web server log file. The interaction between the web user and websites are stored in the web server log file which will be helpful to identify user behavior, user interest, user activities etc. It containsinformation like How the users access the website?How many paths they cross to achieve their target page? There are many research projects available to analyse the user access pattern for various purpose. Web personalization, website restructuring, security, improvement in web server, business intelligence etc., are the applications of web usage mining.

The main goal of web usage mining is to capture and analysis of user behavior when they interact with the website. Collected or analyzed patterns will be in the form of webpages that are very frequently accessed by web users. The process of web usage mining is associated with three different stages: data collection and data pre-processing, pattern discovery and analysis. Data pre-processing task is to clean the data which will be helpful to extract relevant information from the large data. Pattern discovery is focusedon applying data mining techniques, statistical analysis, and machine learning algorithms to find hidden user behaviour, session identification and so on. The final stage is to discover a pattern from the summarized statistical web resources [3].

To analysis web server log file, there are much software exists which includes splunk, Log DNA, Logz.io,Coralogix, Scalyr, Apache log 4j, Datadog,Graylog, Jaeger, solar winds paper trail, solar winds loggly etc. Among these,splunkEnterprise is highly suggestive to analyse the log file [4].

A. Why website reorganization is needed?

In recent days, most of the businesses have become online, which will increase their brand awareness. This digital world has made websites as interdependent to business. Now a day'scustomer resides in the online space for exorbitant of their lives. Websites playa significant role in many businesses to contact and communicate with their customers.

Manuscript published on 30 September 2019
***** Correspondence Author

K. Shyamala*, Associate Professor, PG & Research Department of Computer Science, Dr.Ambedkar Government Arts College (Autonomous), Affiliated to University of Madras, Chennai, India.

S. Kalaivani, Research Scholar, PG & Research Department of Computer Science, Dr.Ambedkar Government Arts College (Autonomous), Affiliated to University of Madras, Chennai, India.

A. Murugan, Associated Professor, PG & Research Department of Computer Science, Dr.Ambedkar Government Arts College (Autonomous), Affiliated to University of Madras, Chennai, India.

How a good website should be? A good website should be well organized and standardize formatted. It should provide easy navigation. It should be succinct. A website is considered as most important in online business because it is the source of customers and also the merchants in providing goods and services.

Objectives to reorganize websites:

- ✓ To gain too many users.
- ✓ To improve user experience in website navigation.
- ✓ To avoid difficulty in using the website.
- ✓ To understand the website features easily.

Websites might be wrapped up with enormous information. But from the website, Is it possible for the web user to find their preferable information? Web users can easily identify it within 4 or 5 seconds. Once if they feel confused or complex, they skip the web page or the website and try to navigate another website. Web latency is another indispensable problem in a global network. Latency cannot be controlled completely but it can be minimized. This research work focused to reduce the latency time of web user by reorganizing websites based on frequently accessed web pages.

Methods to enhance website

- ✓ HTTP request should be minimized
- ✓ A response time of server should be reduced
- ✓ Browser caching should be enabled
- ✓ Plugins should be reduced
- ✓ Resources should be minimized

Tools to Analysis of website performance

- ✓ Pingdom page load time
- ✓ Google page speed Insight (it can be used in mobile as well as desktop)
- ✓ Webpage test

Other than website reorganization, most of the researchers focus on web page prediction in directs to reduce user latency time. Prediction is one among the techniques in web usage mining [5]. Prediction is the process of finding web user future request pages based on past user behavior. For this purpose, the web server log file has to be analyzed to find web user behavior. Predicting web user's future needs is based on the past access activities of the web users. Predicting web pages are useful to forecast trends in share market, enhancing website performance etc. predicting and prefetching webpages from the web server to the client browser will help to reduce the web user latency time. This research work focused to reduce web user latency time by proposed prediction algorithms and alsoachieved better accuracy in predicting webpages.

B. Nonetheless, how to improve web performance?

Web performance cannot be improvedcompletely but the steps can be initiated to access the web server add-on algorithms in an optimized way. This research work has focused to improve web performance by running Monte Carlo [14] prediction algorithm in an optimized way and updating modified webpages in the proxy cache through the prefetching engine. The proxy server always acts as an intermediate between the web server and the web client. The proxy cache holds the previously responded pages. When the user request the page proxy server search in its own cache and respond to the user. If the requested page is not found, then it will fetch from the web server and make one copy to its cache and respond to the user. It will be helpful to reduce web user latency time. The web user can access webpages via the proxy server. They are considered as an indirect client or indirect user because it hides the client identification information for security purpose. The person who accesses webpages through webserver is referred as direct webuser or direct client. In this research work add-on programs are included to the web server which will reduce the minimal server burden by running Monte Carlo prediction algorithm in an optimized way. Later prefetching engine is included in the proxy server cache for handling dynamically updated web page content. It requires the following two caches:

- Cache one holds the response webpages for each request.
- The second cache contains the modified or updated web pages.

By using these processesthe web performance is improved and minimal server burden has been reduced.

II. RELATED WORK

This section discusses the existing work of various researchersrelated to improving web performance in various aspects. Authors [6] mainly focused on removing irrelevant information and extract only the related information. Data cleaning algorithm was used to remove *jpg,*gif etc. User identification is done based on three bases. Firstly the users were identified based on IP address which is unique. Secondly, when the user has the same IP address but the user access from a different browser and the different OS is considered as different users. Thirdly, some users may have the same IP address because they may access the web pages through a proxy server. Session identification is identified based on the time in and time out of the website. Path completion is done by the authors. If the requested page by the user should be linked with the previously visited page. Authors have successfully implemented data pre-processing algorithm.

Developing [7] and maintaining a complex website is very difficult for user interaction. Author has proposed an approach to provide preferred web pages to the web users. They have formalized two approaches one is cluster mining and another approach is page gather. Cluster mining deals with less number of well-integrated clusters(which are very similar). This approach is followed by the author instead of considering a large amount of data into different clusters. Page gather is the clustering algorithm which takes inputs as web server log file and it provides content of index pages. An index page is a page which contains a collection of links about a specified topic.

Authors [8] have explored an approach for the adaptive website. They have experimentally analyzed the page Gather's Algorithm output with the frequent item set identified by the Apriori Algorithm. Also the experimental result compared with the web user accessed index page. The adaptive website suggests the contents to the user automatically based on the user access automatically based on the user access pattern. User visited Pages are recorded in the web server log file which will be the main source of information to create a rich and good adaptive website.

Author [9] has focused onadaptive websites to improve the website structure. They have proposed a 0-1 programming model, which helps to find the relationship between the web pages and reorganize the websites. The main intention of the author is to reduce website overload.

To overcome the website overload they have reduced the outdegree of a hyperlink from the particular page. By reducing the hyperlink, the author avoided the information overload. Other than the information overload,the author has concentrated on the depth of the page when the user surfs. The accessing path length of the pages has been reduced by providing the shortest path from the home page to the destination pages.

The main goal of the author [9]is achieved by using the spanning tree which will be helpful to reduce the depth of the webpage. The subgraphs which are unreachable in the model is solved by using the transportation problem in the proposed model. He has implemented his model using LINGO 8.0. From the experimental result, it is clear that the model takes lesser time for the computation.Model 2 takes longer time than model 1.TheSub graphs which are obtained from the spanning tree is not only satisfying level constraints but also it is connected fully. From the experimental results, it is clear that by using spanning tree the level constraints and degree constraints are more time-consuming process. So, the author proposed two heuristic approaches. The first stage is to satisfy the level and degree constraints in the spanning tree. Another stage is to satisfy the level and degree constraints are the same which is used in stage one. Their proposed model achieved the optimal results based on the relationship between the webpages. Author has suggested performing some other heuristic algorithms, genetic algorithm for the promising results.

Author has presented a descriptionof the techniques and methods used in web mining. These methods and techniques will be helpful in restructing the website with minimal changes to achieve better web navigation. Paper [10] also includes the recommendation based on the frequency access of the web pages and the user access pattern of the web user. They have achieved an efficient result of better web navigation. Proposed techniques are achieved by three phases. The first phase is to extract the architecture of the website; the Second phase deals with collecting the web user log file. The third phase is to obtain browsing efficiency of the website. Website structure is considered as the graph. Each webpage is considered as a node and every connecting hyperlinks are considered as edges. Their proposed program collects the website structure. In the second phase, data cleaning, user identification, session identification, pattern discovery and pattern Analysis has been done. In the third phase browsing efficiency of a website is measured by calculating the shortest distance from the source to destination divided by operating cost. From the improved web navigation they have restricted the website. This proposed system will be helpful for the web user to achieve better web navigation.

III. PROPOSED WORK

This section deals with a detailed explanation of the research work. The main aim of this research work is to reduce webpage access delay and then to provide desired information as soon as possible to the user and reducing the minimal load for the server. Figure 1 shows the complete framework to minimize the webpage access delay. The work has been explained with four phases as follows,

In the first phase, website reorganization is done to reduce web user latency and to improve website structure.

A. Website Reorganization

The effort started in analyzingWebserver log file, which contains all the details about the web users. It includes accessed date, time, web pages, how long they stayed in the webpages, user agent, and a status code of the particular pages, etc. web server log file contains exorbitant information which will not very relevant to the research. It will not be much efficient when the entire dataset is used to find the user behavior because we may also have irrelevant information in the log file. To get an accurate result, raw data have to be pre-processed first.

Pre-processing is the process of removing irrelevant data from the log file and extracting relevant data. Irrelevant data like error status code, incomplete URL, web robot request etc. these are the irrelevant data which has to be removed from the log file. Data pre-processing algorithm is implemented using SQL server management studio. Web server log file is collected from University of Tokyo [11] and also data pre-processing is done to remove irrelevant data. Log file size is reduced then only relevant data are extracted and stored for further analysis purpose.

Clustering is the process of grouping data objects in such a way that data objects are similar [12]. Clustering is one of the techniques in data mining; it comes under the category of unsupervised learning. The intra-cluster similarity is high only when the data objects are similar in a cluster. Inter clustering similarity is less only when the data objects are not similar in a cluster. In this research work, K-means clustering and farthest first clustering algorithms are used to find frequently accessed webpages from the log file.

After pre-processing, the work started with the K-means clustering algorithm by grouping similar data based on measuring mean distance. Later farthest first clustering approach is used by measuring maximum distance. The execution time of K-means and farthest first clustering approaches are measured using WEKA tool. Among these two clustering approaches, farthest first clustering works better than K-means clustering in finding frequently accessed webpages. Farthest first clustering approach is suitable for website reorganization has been proved by previous work of other researcher [13]. The researchers from various works [13] were proved that farthest first clustering approach is suitable for website reorganization. From the analysis, it has been decided that the webpages can be reorganized according to frequently accessed webpages. It has been implementedby constructing binary search tree and then reorganized using max heap [14].

The implementation starts with K-means with the K-value as 10 (i.e.) number of clusters considered as 10. The binary search tree is constructed and then the search cost is measured then Max heap is constructed based on the frequently accessed webpage. The property of max heap tree is to bring the maximum value to the root node. So, in this research work max heap tree is considered to bring frequently accessed pages to the root node. Hence, the search cost is reduced when compared with the existing organization of a website.

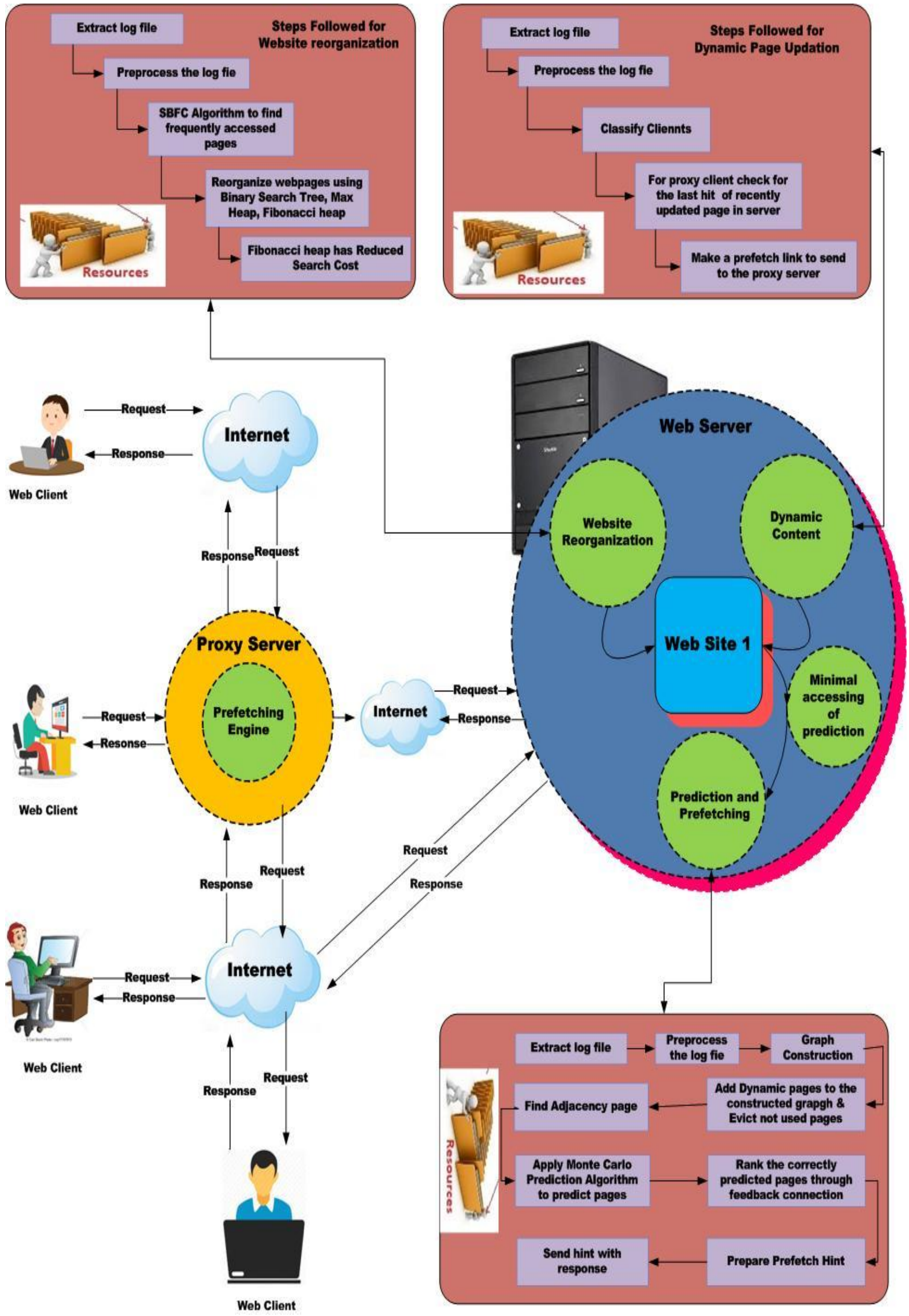
Extract log file
Steps Followed for Website reorganization
Preprocess the log fie
SBFC Algorithm to find frequently accessed pages
Reorganize webpages using Binary Search Tree, Max Heap, Fibonacci heap
Fibonacci heap has Reduced Search Cost
Resources
Extract log file
Steps Followed for Dynamic Page Updation
Preprocess the log fie
Classify Clinents
For proxy client check for the last hit of recently updated page in server
Make a prefetch link to send to the proxy server
Resources
Request
Response
Internet
Web Client
Response
Request
Web Server
Website Reorganization
Dynamic Content
Web Site 1
Minimal accessing of prediction
Prediction and Prefetching
Proxy Server
Prefetching Engine
Request
Resonse
Web Client
Internet
Request
Response
Response
Request
Request
Response
Request
Response
Internet
Web Client
Response
Request
Web Client
Extract log file
Preprocess the log fie
Graph Construction
Find Adjacency page
Add Dynamic pages to the constructed grapgh & Evict not used pages
Apply Monte Carlo Prediction Algorithm to predict pages
Rank the correctly predicted pages through feedback connection
Send hint with response
Prepare Prefetch Hint
Resources

Fig 1:Framework to minimize the web page access delay

B. Proposed and implemented SBFC Algorithm

Instead of using existing algorithms, the effort started to introduce a new algorithm called Split Based Frequently Count Algorithm which is used to find frequently accessed webpages. The SBFC algorithm was proposed and implemented in previous work [15] and the results were compared with theWeka tool. The work shows that the time taken by the weka tool and SBFC proposed algorithm isthe same. Webserver log file taken for the experiment is from the US Securities and Exchange Commission, the Division of Economic and Risk Analysis (DERA). The proposed algorithm is successfully implemented in java with the Net Beans IDE. The working processes of SBFC areto extract the log files and split the unique pages from the URL and keep counting the repeated occurrences of the same page in order to find the frequency count of the page.

From the result of an algorithm,the top 10 frequently accessed web pages were considered for the experimental analysis. Binary search tree was constructed for the frequently accessed web pages then Max heap tree was constructed for the frequently accessed web pages. In this phase the Fibonacci heap was introduced which has the collection of heap ordered tree. It contains two properties one is Fibonacci min heap and the other one is Fibonacci max heap. This research works with the Fibonacci max heap such that each and every node will be considered as root node such that the time complexity for insertion of the node is O (1). So, it will be suitable to add dynamic pages in the root node. In the work [15],Fibonacci max heap is constructed for the top 10 frequently accessed web pages. Then the search cost is measured for all the three structures. Among these three structures Fibonacci max heap gives better performance and reduced search cost.

Hence, it is proved that the Fibonacci max heap is suitable for website reorganization. Web users' interest may vary to time to time. So in this digital world there is phenomenal growth of the dynamic website. Website reorganization is suitable for the dynamic website (it changes seasonally or it changes according to the user interest).

C. Prediction and Prefetching

The second phase deals with prediction and prefetching.This phase consists of two portions one is Monte Carlo Prediction(MCP) with prefetching and other one is EnhancedMonte Carlo Prediction (EMCP) with prefetching. Prediction can be done by collecting log file from the web server and then identifying how users navigated in the website [16]. The web pages which are accessed frequently in the same sequence by the different web users will be considered as the future request of other users. Anticipating user future request pages is calleda prediction. Prefetching is the process of preloading the predicted pages into user browser cache to reduce latency. Before user requesting the page, the prediction algorithm predicts the web pages and preload in the user browser cache. So, the user latency time can be reduced.

In this research work, the prediction was implemented by constructing a complete graph called User Navigation Graph Construction (UNGC) according to the pre-processed log file entries. The working process of graph construction is identifying the unique user and unique pages from the log file then the sequence of access is analyzed to identify the adjacency vertex. Here the unigram sequence has been considered to predict the future request page. The work in [16] shows the table of content which contains the sequence of pages accessed by various users. There may be n number of users for example: A user can access page 1 to page 2 and page 2 to page 7 etc. The graph is constructed based on the reference string. Here the term reference string is considered as sequence of pages accessed by the users. This proposed UNGC algorithm is implemented successfully. Thecomplete graphpaves the way to anticipate the future request page of a user.

After construction of UNGC, the effort started to design an algorithm to make a decision to choose a page from a number of anticipated pages. Here the decision-making technique Monte Carlo was applied for prediction which is called Monte Carlo Prediction (MCP)algorithm. Monte Carlo search is the heuristic search and it is used for decision making in game theory and in machine learning applications. Monte Carlo search chooses most winning probability as a next move is game theory. In this research work, most frequently accessed page was considered as winning probability. Prediction algorithm looks for UNGC to find its adjacency node from the current node. Then it will apply the Monte Carlo [16] on the adjacency nodes. After calculating the Monte Carlo prediction value for each adjacency node,the winning node may have maximum value to move further. Top two maximum value holding pages are considered as future request page from that current node. Four datasets are considered for the experimental analysis [16].The successful implementation has been demonstrated in [16] and it achieved up to 71% of better accuracy for the considered datasets. To improve the accuracy Monte Carlo prediction algorithm has been enhanced with rank based prediction using the feedback process.

D. Rank based Feedback Process

A step forward the Monte Carlo Prediction algorithm was enhanced by introducing rank through a feedback process. Here the purpose of using the feedback process is after the anticipation of a page we have to check whether the user requests the predicted page. The algorithm EMCP (Enhanced Monte Carlo Prediction)was proposed and implemented successfully [17]. The algorithm EMCP used to rank the pages that are predicted correctly. EMCP algorithm does two things; one is to Rank the correctly predicted pages and the other one is to include dynamic pages (newly added web pages) in the constructed graph. The constructed graph will get updated periodicallyand update the rank of a page whenever the correct prediction occurs. The EMCP algorithm describes that the feedback process is the task of comparing the previous and current request of the user to analyze whether the predicted page has been requested by the user. The four same datasets which were used in the MCP algorithmwere consideredfor the EMCP algorithm. Accuracy gets increasedup to 75% when rank based Monte Carlo algorithm is used. This phase paves the way to reduce user latency.

Prefetching was implemented with the two portions of the prediction algorithm. The prefetching process is done by sending the hint of the predicted pages to the user, which makes the preload in the user browser to reduce the user waiting time. The complete architecture of prediction with prefetching has been shown in [16, 17].

The third and fourth phase demonstrate [18] the improvement of web performance by updating updated page to the proxies and optimized way of accessing the EMCP algorithm in the server.

E. Improving web performance

When the usage of internet is increased,the web performance gets decreased. Web performance cannot be improved completely but it can be improved little by introducing various efficient algorithms.

Updating dynamic content in the proxy cache

The proxy cache contains all the responded pages, if the user requests the same page then the page will be taken from the proxy cache then response will be sent to the user. But the drawback is proxy cache contains the old copy of the responded page even it has been updated in the webserver. To overcome this drawback, the new algorithm has been proposed [18] to checks whether the page is modified in webserver.

If the page is modified in the webserver, then the algorithm should check whether the modified page was accessed very recently byan indirect (client access through proxy) client. If it is true then the modified webpage link will be sent asa hint to the proxy to prefetch the page through a prefetching engine.The work of the prefetch engine is to sense the hints from the server response and preload the pages to its cache. In this work,the proxy cache has been categorized into two types [18] cache-1 for holding the server response pages and cache-2 holds the preloaded pages.

The final phase demonstrates the optimized way of running the EMCP algorithm [18].

The optimized way of accessing EMCP algorithm

In the second phase the prediction and prefetching process have been successfully implemented to anticipate and preload the pages to reduce user latency. The problem here is whenever the user requests the page, EMCP will be accessed repeatedly to predict the future request page. This creates the anonymous access of prediction algorithm even for the continuous request of the same page. If the different types of clients like a direct or indirect client request the same page again it creates the anonymous access to the prediction algorithms. Here the optimized running of the prediction algorithm [18] has been proposed tominimize the anonymous access of prediction algorithm.The working process of the optimized way of accessing prediction algorithm is once the prediction is processed then the predicted page hint can be taken into temporary storage for a specific period of time. If the same request has been given to the server then the server can use that temporary storage to send the hint of a predicted page. This process will be efficient only when the predicted page can be maintained with the specific time limit because the prediction may get varies when a number of request and responses increases. This optimized way of accessing EMCP may reduce the minimal burden of a web server.

IV. CONCLUSION

The proposed and implemented algorithms show the improved web performance by reducing user latency time and optimized way of accessing add-on algorithms in the server. The complete frame work shows the four portion of the work. The first portion of the work shows the implementation of different algorithms to reorganize the website in order to reduce the search cost. Split-Based Frequency Count Algorithm (SBFC) with Fibonacci Max Heap shows reduced search cost in website reorganization. The second portion of work elucidates prediction and prefetching to reduce user latency time. The successful implementation of Enhanced Monte Carlo Prediction (EMCP) algorithm with feedback ranking process shows the improved performance. The third portion shows that the dynamic content from the server to the proxy server also updated through the prefetching engine. The fourth portion shows the successful implementation of accessing the EMCP algorithm in an optimized way to reduce the minimal server load. This research work provides improved web performance by reducing user latency and also reducing the minimal load of a server.

REFERENCES

1. https://www.quora.com/What-are-the-uses-of-data-mining
2. https://www.quora.com/What-are-the-real-time-applications-of-web-mining
3. http://facweb.cs.depaul.edu/mobasher/classes/ect584/Lectures/12-web-usage-mining.pdf
4. https://www.g2crowd.com/categories/log-analysis#highest_rated
5. https://www.quora.com/What-are-some-not-so-obvious-applications-of-link-prediction
6. Mary, S. Prince, and E. Baburaj. "An efficient approach to perform pre-processing." *International Journal of Computer Scienceand Engineering* 4.5 (2013).
7. Perkowitz, Mike, and Oren Etzioni. "Adaptive web sites: Automatically synthesizing web pages." *AAAI/IAAI*. 1998.
8. Perkowitz, Mike, and Oren Etzioni. "Towards adaptive web sites: Conceptual framework and case study." *Artificial intelligence* 118.1-2 (2000): 245-275.
9. Lin, Chang-Chun. "Optimal Web site reorganization considering information overload and search depth." *European Journal of Operational Research* 173.3 (2006): 839-848.
10. Sona, Joy Shalom, and Asha Ambhaikar. "A reconciling website system to enhance efficiency with web mining techniques." *International Journal of Scientific and Engineering Research* 3.2 (2014): 498-500.
11. S. Kalaivani and K.Shyamala, "A Novel Technique Pre-Process Web Log Data Using SQL Server Management Studio", *International Journal of Advanced Engineering, Management and Science*, Vol. 2 (7), PP 2454-1311, 2016.
12. S. Kalaivani and K.Shyamala, "Clustering of Web users Behavior based on the Session Identification through Web Server Log File", *International Journal of Control Theory and Applications*, Vol. 10(23), PP 7-16, 2017. **SCOPUS Indexed.**
13. Deepshree.et.al, "Farthest First Clustering in Links Reorganization," *International Journal of Web & Semantic Technology (IJWesT)*.Vol.5(3). July 2014
14. K. Shyamala and S.Kalaivani, "An Effective Webpage Reorganization through Heap Tree and Farthest first Clustering Approach", *In Proceedings of the IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017)*, Vol.4 , PP 263-266, Saveetha Engineering College, Chennai, **IEEE-CATALOG NUMBER**: 978-1-5386-0814-5, 21st and 22nd September 2017.
15. K.Shyamala and S.Kalaivani,"Website Restructuring using Fibonacci Heap and Split Based Frequency Count", *International Journal of Applied Engineering Research*, ISSN 0973-4562 Vol. 13(22), PP 15470-15476, 2018.
16. K.Shyamala and S.Kalaivani, "Application of Monte Carlo Search for Performance Improvement of Webpage Prediction", *International Journal of Engineering and Technology (UAE)*, Vol.(3.4), PP 133-137, 2018. **SCOPUS Indexed.**
17. K. Shyamala and S. Kalaivani, "Enhanced webpage prediction using Rank based Feedback Process", Accepted to publish in *Springer - Lecture Notes in Computational Science and Engineering*.
18. K. Shyamala and S. Kalaivani, "Improvement of Web Performance using Optimized Prediction Algorithm and Dynamic webpage content updation in Proxy Cache." Accepted to publish in *Springer – Lecture Notes on Data Engineering and Communications Technologies*