Reducing User Latency in Web Prefetching Using Integrated Techniques

Naveed Ahmad[1], Owais Malik[2], Mahmood ul Hassan[3], Muhammad Shuaib Qureshi[4], Asim Munir[5]

[1,5]*Department of Computer Science, International Islamic University, Islamabad, Pakistan*

[2]*School of Electrical Engineering and Computer Science NUST, Islamabad, Pakistan*

[3]*Department of Computer Science, Al Jouf University, Tabarjal, Saudi Arabia*

[4]*Ministry of Interior, Islamabad, Pakistan*

[1,5]{naveed.mscs407, asim}@iiu.edu.pk   [2]owais.malik@seecs.edu.pk   [3]mahmood.mscs@yahoo.com

4qureshi.shuaib@gmail.com

Abstract—**Web caching and Web Prefetching are the areas for the research in Web Mining. Web Prefetching improves the performance of the Web Caching techniques due to prediction of the user pages in advance before the user requests. Both techniques provide the web pages local to the user; they provide the resources of web for user's ease and access. Web caching is limited due to its size. Web prefetching is the process of accessing the web objects before the user's request. Whenever a client requests before accessing the web page a prediction is made for accessing that web page. All the web objects are brought from server to the client. The access to the web objects are on the basis of the data prefetched from the server. This research focused on when a user requests for a web page, how to improve the overall performance of web prefetching mechanism? The proposed mechanism provides the pages locally available to a user or group of users by utilizing bandwidth of the network. The server contains an algorithm for the prediction of web pages and the prediction of a web page is based on counting the number of times a page is accessed by a user from each cluster.**

Keywords: **Prefetching;Prediction,User Latency**

I. INTRODUCTION

Web Caching is one of the techniques used for the client, server and proxies to help users in fast access to the web documents. Access is fast because the cache keeps the most recently used documents in the cache. Cache can be deployed to client, server or proxy depending on the purpose for which it is used. Web cache reduced latency because cache keeps the data/web pages local to the users. If the documents are present at the client's cache, then the documents will be accessed faster and this will reduce the network traffic of overall system [1].

The cache has the limited size if we consider the web pages that are recently stored in the cache. The repeated access to same pages will be performed on cache. Prefetching is classified into server initiated prefetching, client initiated prefetching, proxy initiated prefetching and hierarchical prefetching [2].

Our work is based on some concepts presented in the following text. A single user is a person that accessed the file from one or more files through the browser. A web page is the collection of all the files that are seen from the user's screen at any point and contained several files at any time for a user. The web page contains frames, graphics and media. The clicks stream for a user is a series of pages that a user takes from a site. When a client requests for web page, then proposed framework preprocesses the requests of user and provides the pages local to the user for web prefetching. The web page usage information of server is considered in our framework at server side. The main contribution of this research is to enhance the performance of web prefetching mechanism and improve the efficiency of web resource's usage. The paper is organized into following sections:

Related work, proposed framework, experiment for the performance evaluation and conclusion.

II. RELATED WORK

In [3] the method of prediction of the web pages is made at the proxy server. The proxy contains a framework known as prediction engine. In the prediction engine, proxy server log are maintained. The proxy server log records all the user/clients that send request to the server that are kept on the web log which used to form mineable warehouse. The warehouse is used to track the user's activity. In this way predictive web prefetching is defined at the proxy server for   access time optimization of user. The   predicted web pages from the server are loaded into the proxy server while user is busy in performing other tasks using the data mining rules (Association rules).

To improve the mechanism of accessing the web document with accuracy, web logging is used. This method has improved the network performance using prefetching mechanism. In web log mechanism server keeps the web logs of its each client. Log contains the probability of each web object to prefetch. On the basis of weblog, a prediction is made for the future request documents. The prediction about the web object/ web page is based on the current and future requests to the same web object-web page. Prediction about web page is made on the basis of access frequency of user to access that page in [4].

In [5], a mechanism is used for web prefetching that is referred to optimizing the web server using the page rank prefetching for clustered accesses. In the mechanism the rank page algorithm calculates the importance of a page or document to be prefetched. The importance of page to be prefetched depends upon the next page and previous