A Novel Cache Replacement Policy for Web Proxy Caching System Using Web Usage Mining

V. Sathiyamoorthi, Sona College of Technology, Salem, India

ABSTRACT

Network congestion remains one of the main barriers to the continuing success of the internet and Web based services. In this background, proxy caching is one of the most successful solutions for civilizing the performance of Web since it reduce network traffic, Web server load and improves user perceived response time. Here, the most popular Web objects that are likely to be revisited in the near future are stored in the proxy server thereby it improves the Web response time and saves network bandwidth. The main component of Web caching is it cache replacement policy. It plays a key role in replacing existing objects when there is no room for new one especially when cache is full. Moreover, the conventional replacement policies are used in Web caching environments which provide poor network performance. These policies are suitable for memory caching since it involves fixed sized objects. But, Web caching which involves objects of varying size and hence there is a need for an efficient policy that works better in Web cache environment. Moreover, most of the existing Web caching policies have considered few factors and ignored the factors that have impact on the efficiency of Web proxy caching. Hence, it is decided to propose a novel policy for Web cache environment. The proposed policy includes size, cost, frequency, ageing, time of entry into the cache and popularity of Web objects in cache removal policy. It uses the Web usage mining as a technique to improve Web caching policy. Also, empirical analyses shows that proposed policy performs better than existing policies in terms of various performance metrics such as hit rate and byte hit rate.

KEYWORDS


1. INTRODUCTION

Web proxy caching plays a key role in improving Web performance by keeping Web objects that are likely to be used in the near future in the proxy server which is closer to the end user (Ali, 2011). It helps in reducing user perceived latency, network bandwidth utilization, and alleviating loads on the Web servers. Thus, it improves the efficiency and scalability of Web based system (Sathiyamoorthi, 2013).

Since the space apportioned to a cache is limited, the space must be utilized judiciously. Therefore, an intelligent mechanism is required to manage Web objects present in the web cache system. The web cache replacement policy is the core or heart of Web caching system since it plays a critical role in replacing existing web objects especially when cache is full and new one arrives (Kumar, 2009;
Arlitt, 2000). Therefore, the design of efficient cache replacement policy is extremely important and crucial for caching techniques. In other words, there is a question of which objects should be cached and which Web objects should be replaced from the cache to make the best use of available cache space.

As discussed earlier, Web caching is one of the most successful solutions for improving the Web based systems performance. Typically, a Web cache is located in a browser or proxy server, and/or origin server (Kumar, 2008). The browser cache is located on the client’s machine. The user can notice the setting of any modern browser. At the origin server, Web pages can be stored in a server-side cache for reducing the redundant computations and the server loads. Figure 1 shows the architecture of proxy based Web caching system. In this, client side caching only deals with single user whereas proxy caching deals with multiple user’s interest on multiple web servers. Web server caching deals with multiple users’ interest on a single web server.

The most commonly used Web caching methods are not efficient enough and may suffer from a cache pollution problem since they consider just one factor and ignore other factors that may have an impact on the efficiency of Web proxy caching (Koskela, 2003; Romana, 2011). Cache pollution means that a cache contains objects that are not frequently visited, thus reduces the effective cache size and effects on performance of the Web proxy caching negatively.

Many Web proxy caching policies have attempted to combine some factors which can influence the performance of Web proxy caching for making decisions about caching. However, this is not an easy task because one factor in a particular environment may be more important in other environments. Also, the difficulty in determining which ideal Web object will be revisited is still a major challenge faced by the existing web caching system. This is solved by determining page popularity. This article is organized as follows: section 2 deals with related work on web caching, section 3 describes the proposed system methodology and section 4 discusses the results and discussions on proposed system and then finally conclusion and future enhancement.

2. RELATED WORK

Most Web proxy servers are still based on traditional caching policies. These conventional policies are suitable for traditional caching like CPU caches and virtual memory systems, but they are not efficient in Web caching area. This is because they only consider one factor in order to make cache decisions and ignore the other factors that have impact on the efficiency of Web proxy caching.
Web Graph Clustering for Displays and Navigation of Cyberspace
www.igi-global.com/chapter/web-graph-clustering-displays-navigation/31142?camid=4v1a

Uses, Limitations, and Trends in Web Analytics
www.igi-global.com/chapter/uses-limitations-trends-web-analytics/21999?camid=4v1a

Extracting Usage Patterns from Power Usage Data of Homes' Appliances in Smart Home using Big Data Platform