

Intelligent cooperative web caching policies for media objects based on J48 decision tree and naïve Bayes supervised machine learning algorithms in structured peer-to-peer systems

ABSTRACT

Web caching plays a key role in delivering web items to end users in World Wide Web (WWW). On the other hand, cache size is considered as a limitation of web caching. Furthermore, retrieving the same media object from the origin server many times consumes the network bandwidth. Furthermore, full caching for media objects is not a practical solution and consumes cache storage in keeping few media objects because of its limited capacity. Moreover, traditional web caching policies such as Least Recently Used (LRU), Least Frequently Used (LFU), and Greedy Dual Size (GDS) suffer from caching pollution (i.e. media objects that are stored in the cache are not frequently visited which negatively affects on the performance of web proxy caching). In this work, intelligent cooperative web caching approaches based on J48 decision tree and Naïve Bayes (NB) supervised machine learning algorithms are presented. The proposed approaches take the advantages of structured peer-to-peer systems where the contents of peers' caches are shared using Distributed Hash Table (DHT) in order to enhance the performance of the web caching policy. The performance of the proposed approaches is evaluated by running a trace-driven simulation on a dataset that is collected from IRCache network. The results demonstrate that the new proposed policies improve the performance of traditional web caching policies that are LRU, LFU, and GDS in terms of Hit Ratio (HR) and Byte Hit Ratio (BHR). Moreover, the results are compared to the most relevant and state-of-the-art web proxy caching policies.

Keyword: Web caching; Machine learning algorithms; Peer-to-peer systems