

Frequent Pattern Extraction from Web Usage data using Hash based Association Rule

Thet Htoo Swe, Nilar Aye

*University of Computer Studies, Yangon
thethtooswe.ucsy@gmail.com*

Abstract

World Wide Web is a huge repository of web pages and links. Users' accesses are recorded in web logs. Because of the tremendous usage of web, the web log files are growing at a faster rate and the size is becoming huge. Web data mining is the application of data mining techniques in web data. Web Usage Mining applies mining techniques in log data to extract the behavior of users which is used in various applications like personalized services, adaptive web sites, customer profiling, prefetching, creating attractive web sites etc., This paper presents web usage mining where frequent patterns of accessed patterns are generated by using hashing approach. The main problem in generating association rules by Apriori algorithm is processing time. In this system, Apriori algorithm is improved by Hashing approach. Hashing is very efficient for the generation of candidate large itemsets. In addition, hashing employs effective pruning techniques to reduce the transaction database size. Generation of smaller candidate sets by hashing enables to effectively trim the transaction database at a much earlier stage of the iterations, thereby reducing the computational cost for later iterations significantly.