## ABSTRACT

Nowadays web usage mining is one of the emerging area and the patterns of user behaviors from web access log are the most important one to be discovered. This is to help designers to develop attractive web sites, create adaptive web site, business and support services, personalization/profiling, network traffic flow analysis, design cache system to increase quality of experience, etc. Most important phase of web usage mining is reconstruction of user sessions by discovering useful patterns by using pattern discovery technique like association rule mining; Apriori algorithm, Frequent Pattern algorithms, etc.

In web usage mining, association rules are used to discover pages often visited together. Association rules are used to find the relationship between attributes from the item set, a set of pages. Rules are applied to discern pages often looked together to reveal associations between groups of users with specific interests and to determine the most preferred. Companies can make use of these patterns to develop cache system and to satisfy their customers. Association rule discovery techniques try to prune the search space according to support for items under consideration. The drawback of association rule mining is too many rules are generated. Many existing algorithms for generating frequent access patterns have less efficient in execution time.

This framework include several stages are data collection, data cleaning, user identification, session identification, association rule mining (build Frequent Pattern-Tree (FP-tree), set information gain, find frequent item set by Frequent Pattern-Growth (FP-Growth)). By improvement Association Rule Mining with information gain, this framework take less execution time 1.34% and generate fewer rules to 40-50% than the previous method.

Keywords: Web Usage Mining, Association Rule Mining, FP-tree, FP-Growth, Information Gain