Abstract

The tremendous growth in the World Wide Web has led to the user perceived latency when requesting for resources from the web servers. Millions of users are connected to the web server for different needs. To improve the performance of the servers, caching is used where the frequently accessed pages are stored in proxy server caches. Pre-fetching of web pages is the new research area which when used with caching greatly increases the performance. In this paper, a better algorithm for predicting the web pages is proposed. Clustering of web users according to their location using K-Means clustering is done and then each cluster is mined using FP-Growth algorithm to find the association rules and predict the pages to be pre-fetched for storing in cache.

References

- Web Usage Mining: Discovery and Applicaions of Usage Patterns from Web Data, Jaideep Srivastava, Robert Cooleyzy, Mukund Deshpande, Pang-Ning Tan, ACM SIGKDD, Jan 2000.
- K. Chinen and S. Yamaguchi. An Interactive Pre-fetching Proxy Server for
Enhanced Integrated Approach to Predict Web User’s Future Requests using K-Means and FP-Growth


Index Terms

Computer Science
Information Sciences
Enhanced Integrated Approach to Predict Web User’s Future Requests using K-Means and FP-Growth

**Keywords**
- Web Usage Mining
- Apriori
- FP-Growth algorithm
- K-Means clustering