Abstract

The long-term success of the World Wide Web depends on fast response time. People use the Web to access information from remote sites, but do not like to wait long for their results. The rapid growth in the amount of information and the number of users has lead to difficulty in providing effective response time for the web users and this increased web latency; resulting in decreased web performance. Although several proposals have been made for reducing this
latency, like it can be improved by caching, the benefit of using it is rather limited owing to filling the cache with documents without any prior knowledge. Predictive caching becomes an attractive solution wherein the forthcoming page likely to be requested soon are predicted based on user access logs information and pre-fetched, while the user is browsing the current display pages. As web page prediction gained its importance, this paper proposes a bracing approach for increasing web server performance by analyzing user behavior, in this pre-fetching and prediction is done by pre-processing the user access log and integrating the three techniques i.e. Clustering, Markov model and association rules which achieve better web page access prediction accuracy. This work also overcomes the limitation of path completion i.e. by extracting web site structure paths are completed, which helps in better prediction, decreasing access time of user and improving web performance.

Reference

- Thesis (_PhD/Research)
- Liu, B., Hsu, W., and Ma, Y., “Integrating Classification and Association Mining”,Proc. of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), 1998.
- Cadez I., Heckerman D., Meek C. Symth P., and Whire S., “Visualization of Navigation Patterns on a website using Model Based Clustering”, March, 2002

**Index Terms**

Computer Science  
Web Server

**Key words**

Web log file  
pattern discovery  
Petri nets

prefetching  
prediction  
Association rules  
Markov models  
clustering