Conclusion and Scope for Future Enhancement

CHAPTER 5

CONCLUSION AND SCOPE FOR FUTURE ENHANCEMENT

In this thesis work a distinct method of analyzing the web log data to improve web page navigation pattern prediction is suggested. The huge amount of data present in the internet and the user's difficulty to search various pages are addressed in this work. The data set is preprocessed to obtain the required data for effective mining purpose by removing the noisy data. The DBscan algorithm helps in the process of determining the links between pages which are related to information, the time stamp helps in deviating the pattern of the pages. The FP-growth algorithm uses the FP-tree structure to attain a divide-and conquer to break down the mining problem into a set of smaller problems. The frequent patterns which are commonly referred by the user help for prediction of web pages. The algorithm Flame which is a sequential pattern mining algorithm can be used for navigational pattern prediction. The result shows that the final flame process achieves better result in terms of session identified and patterns identified. The Flame algorithm also outperforms the DBscan and FP-growth algorithm.

Future Enhancement

- The prediction accuracy can be improved by implementing some effective clustering process to deal with different log file format.
- Map reduce function can be implemented in the preprocessing when the large session file needs to be get processed.
- Flame algorithm can be further enhanced to perform incremental online prediction of navigational pattern.