Abstract

Pattern mining from the web log data leads to discovery of usage patterns of the user who navigate the web. Patterns which appear frequently in the web log data are item-sets and sequences. In this paper, a novel algorithm Intelligent Generalized Sequential pattern (IGSP) is designed which shows better results than the Generalized Sequential Pattern (GSP) algorithm. Experiment is conducted with respect to running time and number of patterns discovered from the log data and results has shown that IGSP outperforms the well-known algorithms (GSP) algorithm.

References

Sequential Pattern Discovery from Web Log Data

- J. Pei, J. Han, B. Mortazavi-Asl, and H. Zhu, "Mining access patterns efficiently from web logs", in PADKK: Proceedings of the 4th Pacific-Asia Conference on Knowledge Discovery and Data Mining, Current Issues and New Applications London, UK: Springer-Verlag, 2000, pp. 396-407
- S. Chakrabarti, "Data Mining for hypertext: A tutorial survey", SIGKDD Explorations; Newsletter of the special Interest Group (SIG) on Knowledge Discovery and Data Mining, ACM, Vol. 1, No. 2, pp. 1-11, 2000.

Index Terms

Computer Science
Data Mining

Keywords
Web Usage Mining Sequential Patterns Web Log