Abstract

Advancements in the field of wired and wireless network environments have paved route to the advent of many dynamic distributed computing environments. These environments have diverged computing resources and multiple heterogeneous sources of data. Most mining algorithms are designed to mine rules from monolithic non-distributed databases. Even algorithms exclusively designed to operate on distributed databases normally download the relevant data to a centralized location and then perform the data mining operations. This centralized approach does not work well in many of the distributed, ubiquitous, privacy sensitive data mining applications, which opened a new area of research Distributed Data Mining (DDM) under the data mining domain. Out of various methods employed to mine frequent itemsets, tree based methodology proves some efficiency in distributed environment. So in this paper we study a set of tree based algorithms [DTFIM, PP, LFP and PP] to mine frequent pattern in
Efficient Tree Based Distributed Data Mining Algorithms for mining Frequent Patterns

distributed environment.

Reference


- D. W. Cheung, J. Han, V. T. Ng, and C. Y. Wong. 1996. "Maintenance of Discovered Association Rules in Large Databases: An Incremental Updating Technique". In proceedings of 12th ICDE.

- S. Cong, J. Han, J. Hoefflinger and D. Padua, 2005. "A sampling-based framework for parallel data mining." in PPoPP, pp. 255-65


- David Wai-Lok Cheung, Jiawei Han, Vincent Ng, C. Y. Wong, February 26-March 01. Maintenance of Discovered Association Rules in Large Databases: An Incremental Updating Technique, Proceedings of the Twelfth International Conference on Data Engineering, p.106-114.


- Han, J., Pei, J., and Yin, Y. 2000. Mining frequent patterns without candidate generation. In Proc. 2000 ACMSIGMOD Int. Conf. Management of Data (SIGMOD'00), Dallas, TX, pp. 1–12


- Jian Pei, Jiawei Han, Hongjun Lu, Shojiro Nishio, Shiwei Tang, Dongquing Yang," H-Mine: Hyper-Structure Mining of Frequent Patterns in Large Databases", First IEEE International Conference on Data Mining (ICDM'01)
Efficient Tree Based Distributed Data Mining Algorithms for mining Frequent Patterns


- J. Liu, Y. Pan, K. Wang, and J. Han, 2002. "Mining frequent item sets by opportunistic projection". In SIGKDD.


- Tanbeer SK, Ahmed CF, Jeong B. 2009. "Parallel and Distributed Algorithms for Frequent Pattern Mining in Large Databases" IETE Tech Rev;26:55-65


Index Terms

Computer Science

Data Mining
Efficient Tree Based Distributed Data Mining Algorithms for mining Frequent Patterns

Key words

- Tree based algorithms
- Distributed data mining
- Mining
- Frequent patterns