Abstract

The mining of rare itemsets involves finding rarely occurring items. It is difficult to mine rare itemsets with a single minimum support (minsup) constraint because low minsup can result in generating too many rules in which some of them can be uninteresting [3]. In the literature [4, 5], "multiple minsup framework" was proposed to efficiently discover rare itemsets. However, that model still extracts uninteresting rules if the items' frequencies in a dataset vary widely. In this paper, we are using the notion of "item-to-pattern difference" and multiple minsup based FP-growth-like approach proposed in [6] to efficiently discover rare itemsets in the distributed environment. To discover global rare itemsets in distributed environment, information regarding itemsets of local sites is collected in the form of MIS-tree at one site; that is, each site sends its local MIS-tree to a single site where a global MIS-tree will be constructed from all the MIS-trees received from all the sites. This global MIS-tree is mined to generate global rare itemsets. Experimental results show that this approach is efficient in terms of communication bandwidth consumed.

Liu, B., Hsu, W., Ma, Y. 1999: Mining Association Rules with Multiple Minimum Supports. In: ACM Special Interest Group on Knowledge Discovery and Data Mining Explorations, pp. 337–341

R. Uday Kiran, P Krishna Reddy. 2010: Mining Rare Association Rules in the Datasets with Widely Varying Items Frequencies The 15th International Conference on Database Systems for Advanced Applications Tsukuba, Japan, April 1-4, 2010


J. Han, J. Pei, Y. Yin. 2000: Mining frequent patterns without candidate generation, Proceedings 2000 ACM-SIGMOD International Conference on Management of Data (SIGMOD’00), Dallas, TX, USA,

Taweechai Ouypornkochagorn Kitsana Waiyamai Apriori_MSG-P 2011: A Statistic-Based Multiple Minimum Support Approach to Mine Rare Association Rules. Proceedings of the Third International Conference on Knowledge and Smart Technologies

Jutamas Tempaiboolkul. 2013: Mining rare association rules in a distributed environment using multiple minimum supports IEEE

Frequent Itemset Mining Repository, http://fimi. cs. helsinki. fi/data/

Index Terms

Computer Science

Information Sciences

Keywords

Association rules multiple minimum supports MIS-tree rare itemsets