عنوان مقاله:

A New Technique to Improve the Performance of Distributed Association Rules Mining

محل انتشار:

دوازدهمین کنفرانس سالانه انجمن کامپیوتر ایران (سال: 1385)

تعداد صفحات اصل مقاله: 7

نویسندگان:

Mahmood Deypir - *Department of Computer Science and Engineering, School of Engineering, Shiraz University, Shiraz, Iran.*

Mohammad Hadi Sadreddini - *Department of Computer Science and Engineering, School of Engineering, Shiraz University, Shiraz, Iran.*

خلاصه مقاله:

Mining association rules in distributed environments is one of the most important problems in the field of knowledge discovery and parallel and distributed computing. Communication and computation are two important factors in distributed mining of association rules. Current proposed distributed association rules mining algorithms treat all types of frequent itemsets as being the same, while there are different types of itemsets in distributed databases, e.g., derivable and non-derivable. In this study a new technique is developed to reduce communication and computation by exploiting derivability of itemsets in distributed data. In this technique derivable frequent itemsets are mined without any communication and I/O costs. This approach can be utilized in every distributed association rules mining algorithm. Experimental evaluations on real-life datasets show the effectiveness of our technique in terms of communication and run time.

کلمات کلیدی:

Distributed association rules mining, data mining, Non-derivable itemsets, Distributed deduction rules

لینک ثابت مقاله در پایگاه سیویلیکا:

https://civilica.com/doc/44520