

Accepted Manuscript

A Novel Process-based Association Rule Approach through Maximal Frequent Itemsets for Big Data Processing

Zelei Liu, Liang Hu, Chunyi Wu, Yan Ding, Quangang Wen, Jia Zhao



PII: S0167-739X(17)30785-9
DOI: <https://doi.org/10.1016/j.future.2017.08.017>
Reference: FUTURE 3612

To appear in: *Future Generation Computer Systems*

Received date : 27 April 2017
Revised date : 23 June 2017
Accepted date : 8 August 2017

Please cite this article as: Z. Liu, L. Hu, C. Wu, Y. Ding, Q. Wen, J. Zhao, A Novel Process-based Association Rule Approach through Maximal Frequent Itemsets for Big Data Processing, *Future Generation Computer Systems* (2017), <https://doi.org/10.1016/j.future.2017.08.017>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A Novel Process-based Association Rule Approach through Maximal Frequent Itemsets for Big Data Processing

Zelei Liu

College of Computer Science and Technology, Jilin University, Changchun 130012, China

Liang Hu

College of Computer Science and Technology, Jilin University, Changchun 130012, China

Chunyi Wu

College of Computer Science and Technology, Jilin University, Changchun 130012, China
Zhuhai Laboratory of Key Laboratory of Symbol Computation and Knowledge Engineering of
Ministry of Education, Zhuhai College of Jilin University, Zhuhai 519041, China

Yan Ding

College of Computer Science and Technology, Jilin University, Changchun 130012, China
Zhuhai Laboratory of Key Laboratory of Symbol Computation and Knowledge Engineering of
Ministry of Education, Zhuhai College of Jilin University, Zhuhai 519041, China

Quangang Wen

Zhuhai Laboratory of Key Laboratory of Symbol Computation and Knowledge Engineering of
Ministry of Education, Zhuhai College of Jilin University, Zhuhai 519041, China

Jia Zhao*

School of Computer Science and Engineering, ChangChun University of Technology, Changchun
130012, China

High-Assurance Software Laboratory, INESC TEC & University of Minho, Braga, Portugal

Abstract : The maximal frequent itemsets issue in big data processing has become a hot research topic. Most of the previous work on big data processing directly analyzes the data through the existing approaches, which would cause problems of redundant computation, high time complexity, and large storage space. To solve the problems, this paper proposes a Heuristic MapReduce-based Association rule approach through Maximal frequent itemsets mining, HMAM. The main idea is: At first, by directly operating on the transaction database, we allocate transactions to different processing nodes and group all transactions according to dimension. Then, we screen the most frequent transactions from each transaction set using the Bitmap-Sort and obtain best-transaction-set through aggregating all the transaction-elemts of each transaction set. The current candidate maximal frequent itemsets can be acquired by removing sub-transactions in terms of the inclusion relations of the items in best-transaction-set. At the same time, each subset of sub-transactions in the candidate maximal frequent itemsets is discarded from all transaction sets. Then the final candidate maximal frequent itemsets can be obtained by iteration until each transaction set is empty. Finally, we achieve the acquisition of maximal frequent itemsets by employing the minimum support threshold. The experimental results demonstrate that compared with the existing approaches, HMAM significantly avoids producing a large number of candidate itmesets resulting from join operation,

*Corresponding author: zhaiyj049@sina.com

Download English Version:

<https://daneshyari.com/en/article/6873287>

Download Persian Version:

<https://daneshyari.com/article/6873287>

[Daneshyari.com](https://daneshyari.com)