



SDI Review Form 1.6

Journal Name:	Journal of Advances in Mathematics and Computer Science
Manuscript Number:	Ms_JAMCS_39635
Title of the Manuscript:	Improved FTWeightedHashT Apriori Algorithm on Big Data using Hadoop-MapReduce Model
Type of the Article	

General guideline for Peer Review process:

This journal's peer review policy states that **NO** manuscript should be rejected only on the basis of '**lack of Novelty**', provided the manuscript is scientifically robust and technically sound. To know the complete guideline for Peer Review process, reviewers are requested to visit this link:

(<http://www.sciencedomain.org/page.php?id=sdi-general-editorial-policy#Peer-Review-Guideline>)



SDI Review Form 1.6

PART 1: Review Comments

	Reviewer's comment	Author's comment (if agreed with reviewer, correct the manuscript and highlight that part in the manuscript. It is mandatory that authors should write his/her feedback here)
Compulsory REVISION comments	Abstract must be simplified, starting from "Mining frequent itemset on big datasets..." and removing technical data and future work. Lines 38-40: ARM is the process of finding the CO-OCURRENCES between items in a given dataset and describing the association relationships among different ITEMS. Neither "correlations" nor "attributes". These inconsistencies often appear in the manuscript. Line 47: About 50MB and 9900ms to analyze what dataset? Are those results collected from [6] or obtained from authors? Line 80: Is "Frist" a technical word that I do not know? Lines 89-90: Results and conclusions are not part of this section. Line 92: I can not understand the description of the proposed algorithm. I could not reproduce it with the explanations of this paragraph and the revision of figures 1 and 2. Line 102: From this line, "item" is confused with "itemset". The item filtering process is elementary while filtering itemsets is highly complex. Line 130: I don't know what indirect relationship between items in a dataset is. The pseudo code is not well indented and uses variables and functions unknown to the reader: m, F_i, F_i, F_{ii}, Indexof(), Cal_Weight(), T... Three small datasets are mentioned but only results are given for one of them, we do not know which one. The authors do not know the number of readings made with HashTableTrie. Are the rest of the data in table 1 from [5] and [6] or are they prepared by the authors? If the algorithms are not executed under the same conditions, comparisons cannot be made. They conclude "...it is made sure that the presented work is the best in term of time and space..." but only show the times and scans used in their experiments, performed with an unknown small dataset. It would be necessary to test larger and denser datasets to arrive at a conclusion in this regard, and to reflect the memory usage data.	
Minor REVISION comments	There are many versions of the datasets used in this article, with big differences between them. The exact origin of the datasets should be reflected.	
Optional/General comments	There are many spelling errors, punctuation, sentence construction... Paragraphs are too long. Authors should review the manuscript in depth.	

Reviewer Details:

Name:	Enrique Lazcorreta Puigmartí
Department, University & Country	Institute Center of Operations Research, Miguel Hernández University of Elche, Spain