GUJARAT TECHNOLOGICAL UNIVERSITY

(Established by Government of Gujarat under Gujarat Act No. : 20 of 2007)

ગુજરાત ટેકનોલોજિકલ યુનિવર્સિટી

(ગુજરાત સરકારના ગુજરાત અધિનિયમ ક્રમાંક :૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Abstract of the Thesis

Name: PATEL MAHENDRAKUMAR NAROTTAMDAS

Enrolment No: 189999913028

Faculty: Engineering

Discipline/Branch: Computer/IT Engineering

Title of the Thesis: An Adjacency Matrix based Multiple Fuzzy Frequent
Itemsets Mining

Abstract

Discovering helpful information from transactions is becoming an important research issue. Several frequent itemsets mining algorithms are proposed for association rule mining, which handle only binary datasets. These methods concentrate on an item's presence or absence in a dataset. However, in some situations in real life, it is crucial to consider the quantity of items. A fuzzy technique is used to handle quantitative datasets and to generate meaningful representations of the dataset. Thus several algorithms were developed to discover fuzzy frequent itemsets from quantitative transactions. Most of them merely take the linguistics term with the highest cardinality into account. As a result, the number of original elements and fuzzy regions processed is equal. On the other hand, decision-making can be made more successful when an item has several fuzzy zones.

Existing approaches scan the database more than once, and the high number of join counts (candidate itemsets) required thus degrade the algorithm's performance by increasing execution time.

In this research, we proposed AMFFI (Adjacency matrix based multiple fuzzy frequent itemsets mining) and MFFPA-2 (multiple fuzzy frequent itemsets mining using adjacency matrix with type-2 membership function) using an Adjacency matrix and Fuzzy-Tid-list structures to discover multiple fuzzy frequent itemsets (MFFI) that scan the database only once.

AMFFI is proposed for mining MFFI from quantitative transactions. AMFFI technique uses a type-1 membership function to transform quantitative datasets into fuzzy linguistics terms. An efficient search space exploration strategy is proposed to find the occurrence of two fuzzy linguistic terms together immediately from the adjacency matrix to minimize the join counts and speed up the discovery of MFFI.

The proposed MFFPA-2 uses a type-2 membership function to transform quantitative databases into fuzzy linguistics terms. The type 2 Fuzzy Set could be useful for providing more reliable and agile decision-making by considering many uncertainty possibilities and considering more complex relationships between variables.

Extensive experiments have been conducted to verify efficiency regarding runtime, memory usage, and join counts with different min support thresholds. Experimental results demonstrate that the designed approaches AMFFI and MFFPA-2 achieved superior performance compared to cutting-edge techniques. The AMFFI improves execution time by 8% to 81% and node join count by 93% to 99%. The MFFPA-2 improves execution time by 38% to 75% and node join count by 93% to 99%.

This PhD Thesis would be useful for:

The thesis will helpful to the professionals working in industries where data mining techniques are applied, such as retail, healthcare, finance, etc., can benefit from the practical implications discussed in the thesis. They can potentially implement the proposed algorithms to extract valuable insights from their data. Overall, a PhD thesis on fuzzy frequent itemsets mining contributes to both academic knowledge advancement and practical applications in various domains where data analysis is essential.

List of Publication(s):

1) Patel, Mahendra Narottamdas, Sanjay M. Shah, and Suresh B. Patel. "An Adjacency matrix-based Multiple Fuzzy Frequent Itemsets mining (AMFFI) technique." International Journal of Intelligent Systems and Applications in Engineering 10, no. 1 (2022): pp. 69–74. (SCOPUS Approved, ISSN: 2147–6799).

2) Patel, Mahendra Narottamdas, Sanjay M. Shah, and Suresh B. Patel. "An Efficient (MFFPA-2) Multiple Fuzzy Frequent Patterns Mining with Adjacency Matrix and Type-2 Member Function." In International Conference on Advances in Computing and Data Sciences, pp. 502-515. Cham: Springer Nature Switzerland, 2023. (ISSN: 1865-0929)