Cover page

Proceedings of the 5th International Conference on Engineering and MIS 2019

L.N.Gumilyov Eurasian National University, Astana (Nur Sultan), Kazakhstan

ICEMIS'19

Conference Chair: Prof. Y. B. Sydykov

TPC Chair: Prof. Shadi A. Aljawarneh

Organization Chairs: Prof. Sharipbay Altynbek & Dr. Gulmira Bekmanova

Sponsor: International Association of Researchers (IARES Inc., Canada)

The 5th ACM International Conference on Engineering and MIS 2019 Proceedings Series

ACM ISBN: 978-1-4503-7212-1

ACM Press

Content page

First page

TREE BASED DATA FUSION APPROACH FOR MINING TEMPORAL PATTERNS

Vangipuram Radhakrishna
Information Technology Dept
VNR VJIET
HYDERABAD, INDIA
radhakrishna_v@vnrvjiet.in

Shadi Aljawarneh
Software Engg Dept
JUST
Jordan
saaljawarneh@just.edu.jo

P.V.Kumar
Computer Science and Engg Dept,
Acharya Institute of Technology
Bangalore, INDIA
pvkumar58@gmail.com

Vinjamuri Janaki
Computer Science and Engg Dept
Vaagdevi College of Engineering
Warangal, INDIA
janakicse@yahoo.com

Aravind Cheruvu
Information Technology Dept
VNR VJIET
HYDERABAD, INDIA
aravindcheruvu1995@gmail.com

ABSTRACT

Discovering time profiled temporal patterns from time stamped transaction datasets is addressed in our previous research works which includes proposing new support estimation techniques, similarity measures for computing similarity between temporal patterns. This paper proposes a novel approach for discovering temporal pattern by introducing the concept of data fusion w.r.t the temporal pattern tree. The tree is generated for each timeslot and then the trees obtained for individual timeslots are merged or fused to get the overall tree for the entire dataset. The concept of tree based data fusion helps to prune elements efficiently and well ahead during pattern mining process. A pruning function is also introduced in this paper to prune invalid temporal patterns.

General Terms

Novel approach, Computational Complexity

Keywords

Z-score, support, temporal pattern, similarity computation

1. INTRODUCTION

Similarity based temporal pattern mining aims at mining for similarity profiled temporal patterns that essentially satisfy a finite set of similarity constraints guided by the user. Most of the present contributions that are addressed for temporal pattern mining include methods and algorithms for discovering temporal patterns that are frequent within a interested time span, calendar temporal patterns and rules, weighted temporal pattern, cohesive temporal patterns, sequential patterns which essentially do not involve any similarity constraint.

ICTMIDW, June 6–8, 2019, Astana, Kazakhstan.

ACM ISBN 978-1-4503-7212-1/19/06…$15.00.
https://doi.org/10.1145/3330431.3330463

All the previous and present works related to temporal pattern mining did not consider deriving the temporal patterns which are similarity profiled (or) time profiled w.r.t an explicitly mentioned reference (which is usually a multi-dimensional vector sequence of interested prevalence values) that satisfy user guided set of similarity constraints. Another limitation is that the temporal patterns discovered using such methods is based on simple interest measure. By simple interest measure, we mean to say that the support and confidence values are single dimension values.

Extracting the time profiled temporal patterns of user interest from time-stamped transaction datasets that satisfy certain guided similarity condition and other subset constraints from the time-stamped transaction datasets have various important challenges to be addressed. The challenges are as follows:

i) The first challenge is obtaining the multi-dimensional support time sequences of temporal patterns (or temporal itemsets) with minimum data scan operations and the similarity between multi-dimensional temporal pattern support sequence and reference sequence.

ii) The second challenging issue is finding the precise similarity between temporal patterns and eliminate the dissimilar temporal pattern effectively and efficiently. Elimination of dissimilar temporal pattern at an early stage in the discovery process facilitates the algorithm to consume minimal computation time and also optimizes the required computational space.

iii) Current methods of time profiled temporal pattern mining are mostly tree-tree based strategies and these methods (SPAMINE, SEQUENTIAL, G-SPAMINE, and MASTER) require the dataset to be unconditionally retained in the main memory till completion of the algorithm. Tree structures that are designed to store the temporal dataset in compressed representation can be helpful in reducing the total computation time and computation space required. So, the third challenge is to devise methods for pattern mining by proposing compressed tree based approaches for similarity profiled pattern mining in time stamped temporal datasets.