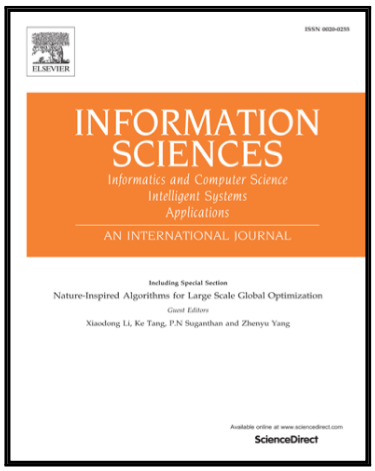
Accepted Manuscript



Effective Lossless Condensed Representation and Discovery of Spatial Co-location Patterns

Lizhen Wang , Xuguang Bao , Hongmei Chen , Longbing Cao

PII: S0020-0255(18)30014-8
DOI: 10.1016/j.ins.2018.01.011
Reference: INS 13365

To appear in: *Information Sciences*

Received date: 24 April 2017
Revised date: 19 October 2017
Accepted date: 7 January 2018

Please cite this article as: Lizhen Wang , Xuguang Bao , Hongmei Chen , Longbing Cao , Effective Lossless Condensed Representation and Discovery of Spatial Co-location Patterns, *Information Sciences* (2018), doi: 10.1016/j.ins.2018.01.011

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.

ACCEPTED MANUSCRIPT

Effective Lossless Condensed Representation and Discovery of Spatial Co-location Patterns

Lizhen Wang[a], Xuguang Bao[a], Hongmei Chen[a*], Longbing Cao[b]
[a]School of Information Science and Engineering, Yunnan University, Kunming, China
[b]Advanced Analytics Institute, University of Technology Sydney, Sydney, Australia

ABSTRACT: A spatial co-location pattern is a set of spatial features frequently co-occuring in nearby geographic spaces. Similar to *closed frequent itemset mining*, *closed co-location pattern (CCP) mining* was proposed for losslessly condensing large collections of prevalent co-location patterns. However, the state-of-the-art condensation methods in mining CCP are inspired by closed frequent itemset mining and do not consider the intrinsic characteristics of spatial co-locations, e.g., the participation index and ratio in spatial feature interactions, thus causing serious containment issues in CCP mining. In this paper, we propose a novel *lossless condensed representation* of prevalent co-location patterns, *Super Participation Index-closed* (*SPI-closed*) *co-location*. An efficient *SPI-closed Miner* is also proposed to effectively capture the nature of spatial co-location patterns, alongside the development of three additional pruning strategies to make the SPI-closed Miner efficient. This method captures richer feature interactions in spatial co-locations and solves the containment issues in existing CCP methods. A performance evaluation conducted on both synthetic and real-life data sets shows that SPI-closed Miner reduces the number of CCPs by up to 50%, and runs much faster than the baseline CCP mining algorithm described in the literature.

Keywords: Spatial data mining, Spatial co-location patterns, SPI-closed co-location patterns, Lossless condensed representation.

* Corresponding author.
E-mail addresses: lzhwang@ynu.edu.cn (L. Wang), bbaaooxx@163.com (X. Bao), hmchen@ynu.edu.cn (H. Chen), longbing.cao@uts.edu.au (L. Cao)

ACCEPTED MANUSCRIPT

Download English Version:

https://daneshyari.com/en/article/6856627

Download Persian Version:

https://daneshyari.com/article/6856627

Daneshyari.com