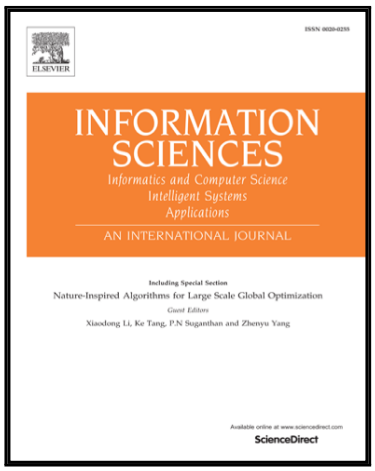

A unified framework of dynamic three-way probabilistic rough sets

Xin Yang, Tianrui Li, Dun Liu, Hongmei Chen, Chuan Luo

PII: S0020-0255(17)30893-9
DOI: 10.1016/j.ins.2017.08.053
Reference: INS 13057

To appear in: *Information Sciences*

Received date: 26 March 2017
Revised date: 11 August 2017
Accepted date: 15 August 2017

Please cite this article as: Xin Yang, Tianrui Li, Dun Liu, Hongmei Chen, Chuan Luo, A unified framework of dynamic three-way probabilistic rough sets, *Information Sciences* (2017), doi: 10.1016/j.ins.2017.08.053

A unified framework of dynamic three-way probabilistic rough sets

Xin Yang[a,b], Tianrui Li[a,*], Dun Liu[c], Hongmei Chen[a], Chuan Luo[d]

[a]*School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China*
[b]*School of Computer Science, Sichuan Technology and Business University, Chengdu 611745, China*
[c]*School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China*
[d]*College of Computer Science, Sichuan University, Chengdu 610065, China*

Abstract

The incremental learning technology has been widely applied in efficient and effective data mining with big data based on granular computing, rough sets and three-way approaches. In real-life applications, the information systems will evolve over time with four levels of variational situations, which can be described by the combination of the variations of attributes, objects, condition attributes values and decision attributes values. Considering updating knowledge with multilevel variations of data, this paper proposes a unified dynamic framework of decision-theoretic rough sets for incrementally updating three-way probabilistic regions, namely, positive region, boundary region and negative region. Through improving the representation of three-way regions based on the well-established Bayesian decision procedure, a novel matrix approach is introduced by the construction of Boolean matrix and specific definition of matrix operation. Subsequently, at the variations of level-1, the fundamental updating propositions, which can induce the corresponding propositions with the variations of level-2, level-3, level-4, respectively, are presented by the matrix updating strategies. Finally, experiments with four incremental algorithms developed for the verification of feasibility and efficiency under multilevel variations of data are conducted by comparison with non-incremental algorithm.

Keywords: Incremental learning, Decision-theoretic rough sets, Three-way decisions, Multilevel variations, Matrix

1. Introduction

With three important perspectives, e.g., philosophy of structured thinking, methodology of structured problem solving and mechanism of structured information processing, granular computing is a multi-view and multi-level of study to deal with uncertain information and concept in human cognition, reasoning and decision-making [39, 40]. As a representative theory of granular computing, rough sets, introduced by Pawlak in 1982 [27] provides a data analysis of approach focusing on knowledge granularity and the approximation of a concept.

With respect to another representation of a pair of lower and upper approximations in rough sets, Yao proposed three-way decisions explained by a trisecting-and-acting framework to interpret three regions [37], namely, three pair-wise disjoint positive, boundary and negative regions. The theory of three-way decisions has been applied in many fields, such as three-way investment decisions [22], three-way group decisions [15], three-way concept lattices [28], three-way cluster [31, 43, 44], three-way recommender [46, 47], three-way multi-class classification [18, 50], sequential three-way decisions [9, 10, 33], etc.

Decision-theoretic rough sets (DTRS) is a probabilistic three-way decisions model. It can systematically calculate the parameters in probabilistic rough sets based on a loss function through the Bayesian decision procedure [36]. Recently, there have many successfully studies in DTRS, such as the extended and modified models, attribute reduction, different areas of application, etc [3, 4, 7, 8, 11, 16, 21, 29, 35]. By considering the minimum expected overall decision risk based on bayesian data analysis, DTRS model divides the universe into three regions, which can be interpreted

*Corresponding author.
Email addresses: yangxin2041@163.com (Xin Yang), trli@swjtu.edu.cn (Tianrui Li), newton83@163.com (Dun Liu), hmchen@swjtu.edu.cn (Hongmei Chen), cluo@scu.edu.cn (Chuan Luo)