

基于粒计算的多源信息融合方法综述

徐伟华, 黄旭东, 蔡 可

(西南大学人工智能学院, 重庆 400715)

摘 要: 多源数据是一种综合多个信息源或数据集的复杂数据类型, 其主要特点是不同的信息源隐含不同的知识结构, 且从不同的角度刻画和描述了样本以及样本之间的关系。如何协同地融合与集成多源数据, 并从不同视角快速地为用户挖掘出整体决策知识, 成为数据科学领域亟待破解的科学问题。经典粗糙集理论、多粒度方法、证据理论和信息熵是常见的、有效的多源信息融合方法, 已取得较为丰硕的成果。本文基于粒计算的角度对多源信息融合工作进行综述研究, 介绍了每种信息融合方法的基本概念以及主要研究思路, 并提出了多源信息融合领域中存在的若干问题, 为该领域的后续研究提供理论参考。

关键词: 粒计算; 粗糙集; 多粒度; 证据理论; 信息熵; 多源信息融合

中图分类号: TP18 **文献标志码:** A

Review of Multi-source Information Fusion Methods Based on Granular Computing

XU Weihua, HUANG Xudong, CAI Ke

(College of Artificial Intelligence, Southwest University, Chongqing 400715, China)

Abstract: Multi-source data is a complex data type that integrates multiple information sources or data sets. Its main feature is that different information sources imply different knowledge structures, and represent and describe samples and relationships between samples from different perspectives. How to fuse and integrate multi-source data cooperatively and how to quickly mine the overall decision-making knowledge for users from different viewpoints have become a scientific problem that needs to be solved urgently in the field of data science. Classical rough set theory, multi-granularity method, evidence theory and information entropy are common and effective multi-source information fusion methods, which have been widely concerned and achieved fruitful results. Therefore, this paper summarizes the work of multi-source information fusion based on granular computing, reviews the basic concepts and main research ideas of each information fusion method, and puts forward some problems in the field of multi-source information fusion. The obtained results can provide a theoretical reference for the follow-up research in this field.

Key words: granular computing; rough set; multi-granularity; evidence theory; information entropy; multi-source information fusion

引言

信息融合作为大数据领域一个重要的研究课题是对多源数据综合进行多层次、多视角信息处理的过程,其主要思想是通过多源数据的检测、关联、组合以及计算等操作来挖掘比单个信息源更全面、更精细的信息^[1]。信息融合最早源于军事领域,随着研究的不断深入,多源信息融合已成功应用于多个军事领域和民用领域,如指挥自动化系统、多目标跟踪与识别、精确制导武器、无线通信、遥感监测及医学诊断等^[2-4]。多源数据综合了多个信息源或数据集的复杂数据,其主要特点是不同的信息源隐含着不同的知识结构,且从不同的角度刻画和描述样本以及样本之间的关系^[5]。在日常生活中经常会面临对多源数据的处理,比如不同医院对同一病人有不同的体检结果;不同气象监测点对同一区域有不同的结果;同一地理空间的时空数据可以通过航空摄影、遥感图像等不同方式获取;监控系统对同一位置的监控数据也可以通过视频监控、红外监控、传感器监控等多种方式获取等。众所周知,数据从不同角度呈现着不同的信息。从多个角度获取的数据可以从多个层次和多个方面来描述事物,并通过这些获取的多源信息对事物有一个更加完整、更加全面的刻画,从而做出更加合理和精准的决策。多源信息融合正是充分利用多渠道、多角度、多层次获取的信息进行数据处理的方法。该方法将多种信息通过特定规则协同融合,对被测对象进行一致性解释进而服务社会^[6]。

粒计算^[7]的核心思想是将信息按照一定的关系划分为不同的“粒”,然后通过处理信息粒的方式获取数据中的信息。近年来,众多学者已将粒计算的思想成功运用到了多源信息融合领域。相较于一般的复杂数据,多源数据中信息或数据之间的关系更加复杂、数据规模更大、获取知识的难度也更大。因此这种信息粒化的思想更适用于复杂数据的知识获取和数据分析,如股票市场分析^[8]、医学诊断^[9]和故障诊断^[10]等领域。粒计算信息融合方法主要利用粗糙集理论(包括经典粗糙集和多粒度粗糙集)、证据理论和信息熵理论等经典粒计算工具和理论。经典粗糙集理论^[11]能够分析隐藏在数据中的信息而不需要任何先验知识,即利用二元关系对论域(研究范围)进行粒度划分,从而通过近似算子的逼近思想从粒度空间中获取有效知识。多粒度粗糙集^[12]则通过多个二元关系来确定论域的粒度空间,并利用逻辑“且”“或”组合运算实现不同粒度空间的知识获取,即融合信息系统中多方面的局部知识获取全局知识。证据理论^[13]是利用辨识框架、基本概率分配(Basic probability assignment, BPA)、信任函数以及似然函数来获取证据,并组合规则融合不同证据实现知识获取。信息熵理论^[14]则主要通过香农熵及其拓展熵获取多源数据中的信息量,并依据特定选取规则实现多源信息融合。

粒计算多源信息融合研究主要包括多源信息融合机制的设计、多源信息融合算法的优化、多源信息知识获取的实现、多源信息规则提取的发展以及多源信息动态知识的更新等。对于多源数据,不同的研究领域采取的研究方法往往都各有特色,其实现的效果也有所不同。本文以多源信息融合为研究对象,以粒计算方法为工具,重点归纳总结当前粒计算领域中已开展的多源信息融合相关工作,并指出具体研究存在的问题,为促进基于粒计算多源信息融合领域的进一步研究与发展提供理论参考。

1 粒计算多源信息融合方法

本节重点对基于粗糙集理论、信息熵、证据理论等粒计算思想方法的多源信息融合理论以及动态多源信息融合研究进行综述。粒计算的具体理论和多源信息系统的详细概念、性质、结论等不再赘述,可以参考文献[15-18]和文献[1, 4-5, 19]详细了解。

1.1 一般粗糙集理论的多源信息融合方法

经典粗糙集理论的核心思想是将一个信息系统的论域划分为不同的颗粒,然后对颗粒进行分析研

究,从而获取有用的知识。假设四元组 (U, A, V, f) 描述一个经典的信息系统,其中 U 为信息系统的论域,代表非空有限的对象集, A 代表论域的属性, $V = \bigcup_{a \in A} V_a$ (V_a 代表属性 a 的范围), f 为一个函数,表示为 $f: U \times A = V (f(x, a) \in V_a, a \in A)$ 。在给定等价关系 R 的情况下,论域 U 被划分为互不相交的等价类 $[x]_R$ 。对于任意 $X \subseteq U$,分别定义 X 的下近似 $\underline{R}X = \{x: [x]_R \subseteq X\}$ 和上近似 $\overline{R}X = \{x: [x]_R \cap X \neq \emptyset\}$,若 $\underline{R}X \neq \overline{R}X$,则认为 X 是粗糙的。由于等价关系是一种较为严苛的关系,因此众多学者在此基础上提出了一系列的拓展模型,如变精度粗糙集模型^[20]、模糊粗糙集模型^[21]和概率粗糙集模型^[22]等,如图1所示。

粗糙集理论提出以来,已受到各个领域的广泛关注,并成功应用于机器学习、决策分析、过程控制、近似推理、模式识别和数据挖掘等领域^[23-26]。一些学者运用粗糙集理论来处理多源数据,并取得了一些较好的成果,主要包括多源数据的Pawlak近似空间^[27-28]、多源数据的决策过程^[29-30]、多源数据的故障诊断^[31]、多源数据的属性选择^[32]以及多源数据信源的可靠性分析^[33]等。除此之外,文献[19]详细论述了粗糙集理论在信息融合领域的属性约简、特征选取以及多源信息融合等。在处理多源数据的一些实际问题中,粗糙集理论也有着较好的表现,如表1所示。文献[34]在无线网络中构建了一种基于粗糙集的信息融合算法用于减少通信数据和能源的消耗,实现了延长无线网络寿命的目的。文献[35]在多源警报数据中利用粗糙集理论对不同警报进行分类,并对警报进行有效聚合从而减少了重复警报。文献[36]利用高斯核、模糊等价关系的拓展和模糊集的包含测度等提出了一种三支决策方法,并成功应用于医学诊断。文献[37]将粗糙集理论与遥感影像结合,融合生成了决策规则。文献[38]构建了基于可

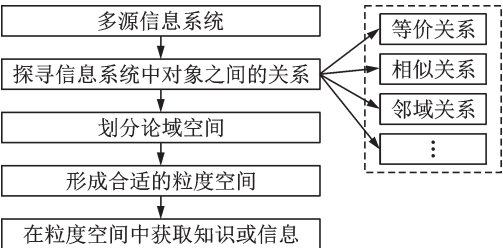


图1 多源信息系统中粗糙集模型框架
Fig.1 Framework of rough set model in multi-source information system

表1 基于一般粗糙集的多源融合方法对比

模型	主要工作
基于粗糙集的无线传感器网络 ^[34]	基于聚类的知识约简等数据分析,使得传感器节点仅自动提取和传输高效的最小数据集成为可能,从而产生高质量的信息,支持决策
多源警报分析 ^[35]	基于粗糙集理论,计算了每一类警报的权值。利用特征权重,进行警报相似度计算,以减少来自不同设备的重复警报
图像混合系统 ^[36]	基于粗糙集理论,提出了一种基于高斯核的图像混合信息系统的三支决策方法,该方法可应用到医学诊断
决策应用模型 ^[37]	基于粗糙集的决策级融合算法的思想,利用像素级融合算法对遥感影像进行融合处理,结合融合定量评价以及左右算法,构建策信息表。采用数据分析法决策应用模型进而生成决策规则
可变精度粗糙集 ^[38]	构建了基于可变精度粗糙集模型,通过基于可变精度粗糙集的获取规则方法,得出车辆的行驶模式
多重近似空间 ^[39]	讨论了基于多重近似空间的多传感器数据融合系统的性能评价理论和方法,并在指挥自动化技术系统的背景下建立了信息融合性能指标体系
变精度模糊粗糙集与 Vague 集 ^[40]	提出变精度模糊粗糙集来减少指标,去除冗余和重复指标,保留关键指标。利用改进的模糊集进行了综合评价。对电网企业的协同效应评价指标体系进行筛选

变精度粗糙集模型,并用于驾驶过程获取规则。文献[39]讨论了基于多重近似空间的多传感器数据融合系统的性能评价理论和方法,并在指挥自动化技术系统的背景下建立了信息融合性能指标体系。文献[40]引入变精度模糊粗糙集对电网企业的协同效应评价指标体系进行筛选。

综上,粗糙集理论是一种非常有效的处理不确定数据的理论,主要通过上下近似对偶算子进行逼近刻画,除了可以处理不完整数据、冗余信息压缩和数据关联等之外,也可以高效地处理多源数据的近似空间、属性选择以及不确定决策等问题。

1.2 基于多粒度粗糙集的多源信息融合方法

多粒度粗糙集最初是用于处理具有多粒度结构和多视图的信息系统。多粒度结构通常根据用户的需求或具体问题通过多个关系(如等价关系、容差关系和相似关系等)构建。给定一个信息系统 (U, A, V, f) ,对象子集 $X \subseteq U, R_1, R_2, \dots, R_m$ 是由属性子集 $A_1, A_2, \dots, A_m \subseteq A$ 产生的多个二元关系,引

入支持特征函数 $S_X^{R_i}(x)$, X 关于 $R_1, R_2, \dots, R_m$ 的广义上近似和下近似分别为 $\left(\sum_{i=1}^m R_i(X)\right)^\beta = \left\{x \in U \mid \frac{\sum_{i=1}^m S_X^{R_i}(x)}{m} \geq \beta\right\}, \left(\overline{\sum_{i=1}^m R_i(X)}\right)^\beta = \sim \left(\sum_{i=1}^m R_i(\sim X)\right)^\beta$, 则称 X 在 $R_1, R_2, \dots, R_m$ 下的序对 $\left(\left(\sum_{i=1}^m R_i(\sim X)\right)^\beta, \left(\sum_{i=1}^m R_i(X)\right)^\beta\right)$ 是广义多粒度粗糙集;当 $\beta=0$ 时,序对称乐观多粒度粗糙集;当 $\beta=$

1 时,序对称悲观多粒度粗糙集^[41]。这里只介绍了部分多粒度粗糙集的方法,更多相关理论可以参考文献[12,42-43]。

多粒度粗糙集是实现信息融合(多约束问题)的一种较好方法,其将不同的约束条件灵活处理进而实现信息融合^[44]。由于多粒度粗糙集可以对不同约束条件进行处理数据,因此该理论也能很好地运用在多源信息融合中,实现对多源数据的处理。在多源数据集中,一般将不同的信源看作为不同的粒度,在每个粒度上都有相应的关系或约束条件,通过对多个粒度的处理最终实现多源数据的融合。多源信息系统下多粒度粗糙集的融合框架如图2所示,其中的粒度划分策略(关系、约束问题等)可通过具体问题灵活选取。

已有众多学者对多粒度粗糙集多源信息融合的问题展开了一系列研究,如表2所示,主要有以下3方面:(1)通过多粒度粗糙集及其拓展模型研究多源数据背景下的决策问题^[45-46]; (2)为避免多源数据融合过程中信息丢失问题,研究直接从多源数据中获取信息或知识的问题^[47-48]; (3)与其他理论(证据理论、机器学习方法等)类似,寻求性能和精度更强的融合模型,比如证据理论结合多粒度粗糙集的多源信息融合^[49-51]、结合机器学习相关理论和多粒度粗糙集实现多源信息融合^[52-54]等。除此之外,多粒度粗糙集理论也已经被应用于多源信息融合的目标威胁估计^[55]。

综上,基于多粒度粗糙集的多源信息融合研究,主要通过多粒度的思想从多源信息系统中获取有效信息,对各个粒

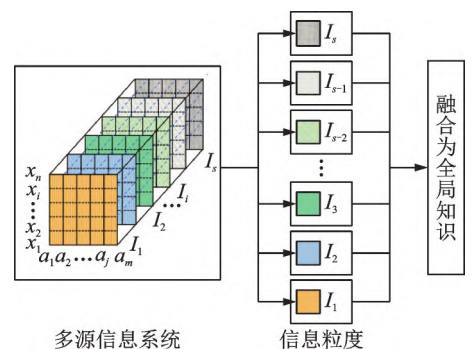


图2 多源信息系统下的多粒度粗糙集融合框架

Fig.2 Multi-granularity rough set fusion framework under multi-source information system

表 2 基于多粒度粗糙集融合方法对比

Table 2 Comparison of fusion methods based on multi-granularity rough sets

融合机制	主要工作
决策问题 ^[46]	利用 4 种类型的 MG-PF-DTRSs 模型及其不确定性测量方法,提出了一种用于多源不完备信息系统决策的方法
直接获取知识 ^[47]	利用多粒度粗糙集的思想,在不丢失信息的情况下直接从多源信息系统中进行知识发现或者规则提取
结合证据理论 ^[49,51]	提出了一种融合 Dempster-Shafer 理论和多颗粒粗糙集理论的两级定量融合方法来处理多源信息系统的 uncertainty 问题,有助于多源信息系统中不确定数据的整合
与机器学习结合 ^[52]	从机器学习的角度出发,提出了一种基于决策树学习分类精度的粒度加权新方法,并构造了加权广义多粒度区间值决策理论粗糙集的基本形式

度进行交叉融合计算,从而解决不同多源信息系统的融合问题。

1.3 基于证据理论的多源信息融合方法

证据理论即 D-S 证据理论^[56-57]作为一种强大的不确定性信息融合工具,不仅能够有效地解决推理过程中证据的不确定性和不精确性,也能够有效地通过不确定信息而做出决策。证据理论主要有辨识框架、基本概率分配、信任函数以及似然函数 4 个基础概念^[58]。

辨识框架 $\Theta = \{H_1, H_2, \dots, H_n\}$ 是由 n 个两两互斥元素 $H_i (1 \leq i \leq n)$ 所组成的有限完备集合,即对象集合全体。将 Θ 中所有子集组成的集合记为幂集 $P(\Theta)$, 且 $P(\Theta) = 2^\Theta$ 。设 A 为 $P(\Theta)$ 的子集, 将集合 A 映射到 $[0, 1]$ 的一个函数 $m(A)$, 且 $m(A)$ 满足 $m(\emptyset) = 0, \sum_{A \subseteq \Theta} m(A) = 1$, 则称 $m(A)$ 为 Mass 函数或 BPA。如果映射 $\text{Bel}: 2^\Theta \rightarrow [0, 1]$ 满足 $\text{Bel}(A) = \sum_{B \subseteq A} m(B), A \in 2^\Theta$, 则称 $\text{Bel}(A)$ 为对命题 A 为真的信任度量, 即命题 A 的信任函数。如果映射 $Pl: 2^\Theta \rightarrow [0, 1]$ 且对所有的 $A \in 2^\Theta$ 有 $Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) = 1 - \text{Bel}(\overline{A})$, 则称 $Pl(A)$ 为似然函数, 即对命题 A 的怀疑程度。

在 D-S 证据理论中, 为了确定多个证据作用下总的信任度, 通常是将两个或多个基本概率分配函数进行合成得到一个基本概率分配函数^[59]。而证据理论算法的核心就是合成规则, 常用的合成规则是 Dempster 合成规则, 其可以通过将两个或多个不同的基本概率分配函数进行正交和运算^[60], 进而融合得到一个 BPA 并将其作为 D-S 证据理论计算的结果输出^[56]。本节只介绍一些证据理论的基本概念和规则, 证据理论更多的性质和结论详见文献[56-58]。

基于证据理论的多源信息融合思想如图 3 所示。首先是从不同的信息源中提取不同证据的 BPA。然后, 通过证据合成规则(一般有很多种合成规则, 常用的有 Dempster 证据合成规则)将不同的 BPA 合成得到一个融合后的 BPA。最后, 通过融合后的 BPA 进行证据决策, 进而完成整个证据理论融合多源信息系统的过程。证据理论具有严密的理论基础, 且具备很强

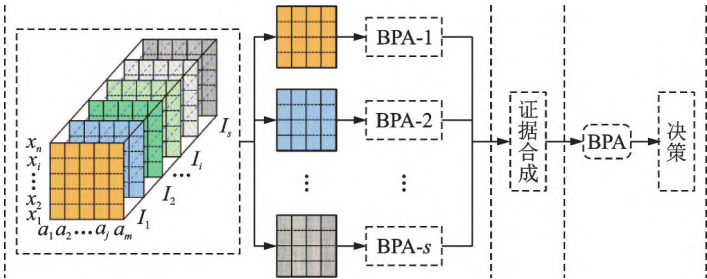


图 3 基于 D-S 证据理论的多源信息系统的融合过程

Fig.3 Fusion process of multi-source information system based on D-S evidence theory

的不确定性描述能力和系统合成推理规则能力,因此证据理论在很多领域都得以成功运用,如工程数据融合^[61]、逻辑推理^[62-63]、目标识别^[64]以及故障诊断^[65-66]等。然而,D-S证据理论在信息融合领域依然存在一些问题有待解决:(1)在不同的应用背景下,BPA函数的合理性问题;(2)在融合冲突的证据时存在证据冲突问题。从Zadeh提出了著名的Zadeh悖论^[67]开始,国内外的学者们对于问题(2)展开了众多研究,本节也将重点介绍关于问题(2)的解决方法。

针对问题(1),在多源信息系统中通常需要将多个信源中的信息转换为证据理论中的多条证据,即给出相应信源的BPA。因此,合理构建生成BPA就极为重要,同时不同的生成BPA方法对融合计算的实时性也有着直接影响。目前,常见的生成BPA的方法主要有2种:一种为Boudraa^[68]和Rombaut^[69]等提出的基于粗糙C-均值分类结果确定像元BPA的方法;另外一种Mandler等^[70]通过计算距离来估计统计分布,并根据统计分布计算类条件概率,然后将其转化为BPA进行组合。同时,近些年如何在多源数据中生成合理的BPA也取得了一些成果,如表3所示,如结合模糊朴素贝叶斯和最近均值分类构建一种数学结构提取多源信息系统中的加权证据^[71]、结合区间数论和K-mean聚类确BPA的方法^[72]、结合样本数据和经验数据之间的距离确定加权BPA的方法^[73]、基于三角模糊数生成广义概率分配的方法^[74]、结合条件属性和决策属性的不同构建改进的基本概率分配^[75]等。

表3 基于证据理论的融合方法对比

Table 3 Comparison of fusion methods based on evidence theory

问题分类	主要工作	参考文献
生成合理的BPA函数	提出了一种基于加权模糊 Dempster-Shafer框架的体系结构,该框架可以调整不同分类方法获得的不一致证据的权重,以实现多模态信息的融合系统	[71]
	通过对从距离直接获得的BPA函数进行加权,提出了权重基本概率分配(Weighted basic probability assignment, WBPA)函数来提高证据组合性能。将WBPA与D-S证据理论相结合传感器采集到的信息进行聚合,解决了故障诊断问题	[73]
	针对多源信息融合问题,本文考虑了从封闭世界到开放世界假设的不确定信息建模情况,研究了不完全信息下基本概率赋值的生成问题。在开放世界假设下,提出了一种基于三角模糊数模型的广义基本概率分配(Generalized basic probability assignment, GBPA)生成方法	[74]
据合成规则的不合理导致高冲突证据合成后产生悖论	考虑质量函数的信任度量和合理性度量来反映不同类型子集的相关性,并构建了一种新的多源数据融合方法	[76]
	基于证据冲突概率加权平均的改进融合算法引入到突水灾害风险预测中,并取得了较好的效果	[77]
出现Zadeh悖论的原因在于证据本身	通过信任 Jensen-Shannon散度和信度熵的方式获取信息量权重实现证据融合	[78]
	结合K-means聚类的方式进行证据内融合和证据间融合,以减少多源证据中冲突的不利影响	[79]
	提出了一种基于加权置信熵和否定BPA的数据融合方法来衡量证据的相对重要性	[80]

针对问题(2),主要基于两种不同的观点。一种观点认为是证据合成规则的不合理导致高冲突证据合成后产生悖论。针对这一观点主要采用对冲突部分证据进行重新定义基本概率分配以实现证据的组合^[81-82]。而对于处理多源数据中的冲突证据则采用了类似的加权方式来实现,如文献[76]考虑质量函数的信任度量和合理性度量来反映不同类型子集的相关性,并构建了一种新的多源数据融合方法,以及文献[77]将基于证据冲突概率加权平均的改进融合算法引入突水灾害风险预测中,并取得了

较好的效果。另一种观点则认为 Dempster 组合规则并不存在问题,而是出现 Zadeh 悖论的原因在于证据本身。针对这一观点主要的方法有考虑集合势的平均算法^[83-84]、算数平均算法^[85]和加权平均算法^[86]等。这些方法同样可以应用于多源信息系统中,如文献[78]通过信任 Jensen-Shannon 散度和信度熵的方式获取信息量权重实现证据融合,文献[79]结合 K-means 聚类的方式进行证据内融合和证据间融合,以减少多源证据中冲突的不利影响,文献[87]使用修正 Minkowski 距离函数和 Betting-commitment 距离函数用于修正冲突证据,以增强证据之间的一致性;文献[80]则提出了一种基于加权置信熵和否定 BPA 的数据融合方法来衡量证据的相对重要性。

综上,D-S 证据理论是处理不确定性信息融合的有效工具,核心思想是利用证据的 BPA 进行证据推理和合成,对人工智能、模式识别和检测诊断等领域的信息融合应用作出了一定的贡献,如多源信息融合在地震烈度判定的应用^[88]、多源信息在时域信息的积累^[89]以及有序信任函数的证据距离度量^[90]等。

1.4 基于信息熵的多源信息融合方法

在信息理论中,信息熵是由 Shannon 等^[91]提出的一种用来量化信息的度量。在信息系统的各个模型和应用中,信息熵仍然是一种很有效的数学工具。

假设一个决策信息系统 $DS = \{U, C \cup D, V, f\}$, $P \subseteq C$ 是条件属性 C 的一个子集,则 P 在论域 U 上的一个划分为 $U/P = \{X_1, X_2, X_3, \dots, X_k\}$ 。论域 U 中的元素是随机分布在 P 中的,关于 P 的信息熵为

$$H(P) = \sum_{i=1}^k P(X_i) \log_2(P(X_i)) \left(\text{其中 } P(X_i) = \frac{|X_i|}{|U|} \right)。$$

当划分中的每一个类 X_i 都是单元集时,信息熵 $H(P)$ 达到最大 $\log_2|U|$;当划分中只有一个类时,即所有论域对象 U 时,该信息熵达到最小值 0,即 $0 \leq H(P) \leq \log_2|U|$ 。类似地,决策属性 D 关于条件属性 P 的条件信息熵为 $H(D|P) =$

$$\sum_{i=1}^m \frac{|X_i|}{|U|} \sum_{j=1}^n \frac{|X_i \cap d_j|}{|X_i|} \log_2 \frac{|X_i \cap d_j|}{|X_i|}。$$

本节只介绍一些信息熵和条件熵的一些基本概念,信息熵和条件熵更多的内容详见文献[41,91]。

信息熵理论结合粒计算的思想取得了一系列的研究成果,如 Düntsch 和 Gediga^[92]定义了用于预测决策属性的粗糙集信息熵和 3 种条件熵;Beaubouef 等^[93]提出了一种基于粗糙熵的粗糙关系数据库不确定性度量方法;Qian 等^[94]提出了一种用于评价信息系统中知识不确定性的组合熵。这些研究都致力于从知识划分能力的角度评估集合的不确定性。同时,Liang 等^[95-96]提出并研究了信息系统中知识的几种度量措施,包括粒度度量、信息熵、粗糙熵和知识粒度等;Xu 等^[97]在序信息系统中引入了知识粒度、知识熵和知识不确定性测度等概念。除此之外,文献[98]开发了一对与分割和近似相关的信息论熵和协熵函数。信息熵结合粒计算理论广泛应用于各个领域,基于信息熵的多源信息融合步骤如图 4 所示。首先,从获取的信息中计算信息熵;其次,通过特定的信息熵选取规则(一般为最小信息选取)选取有效信息;最后通过融合规则实现有效信息的融合,并获得多源信息系统的融合结果。

如表 4 所示,在多源信息系统中,通过信息熵实现多源信息融合的研究主要集中在以下 3 方面:(1)利用条件熵确定每个信息源对于整个系统的重要性,利用最小条件熵的方式实现对多源数据的融合^[99];(2)利用广义信息熵实现对于多源数据的融合,并应用于实际问题,如多传感器决策融合^[100-101]、传感器策略数据处理问题^[102]、融合的指标评价^[103]以及多源数据中的多标签特征选择^[104]等;(3)基于信息熵拓展模型的多源信息融合,如为解决传统电网故障诊断方法造成的误判问题,文献[105]利用贝叶斯网络推导功率元件的故障程度,然后提取各电路故障后电信号的时域奇异谱熵、

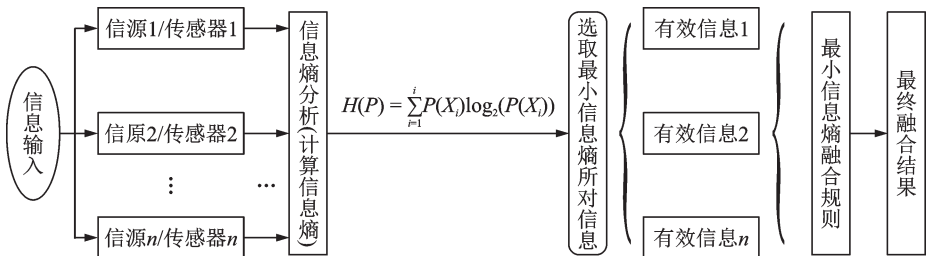


图4 信息熵的多源信息融合策略

Fig.4 Multi-source information fusion strategy of information entropy

表4 基于信息熵的信息融合方法

Table 4 Comparison of fusion methods based on information entropy

模型	主要工作
条件熵 ^[99]	利用条件熵确定每个信息源对于整个系统的重要性,利用最小条件熵的方式实现对多源数据的融合
广义信息熵 ^[100]	提出了一种新的基于广义信息熵的信息填充技术来解决决策融合中的不完备性问题
基于信息熵拓展模型 ^[105]	利用贝叶斯网络推导功率元件的故障程度,然后提取各电路故障后电信号的时域奇异谱熵、频域功率谱熵和小波包能量谱熵,并将这3个特征量作为电源的故障支持度成分结合多源信息融合理论提出了一种新的诊断方法

频域功率谱熵和小波包能量谱熵,并将这3个特征量作为电源的故障支持度成分结合多源信息融合理论提出了一种新的诊断方法。

综上,基于信息熵的信息融合方法重点是利用信息熵来衡量各个信源的某种信息量,从而借助信息熵的重要性质解决多源数据中信息不对称问题,进而实现多源数据信息或规则的获取及融合。

1.5 基于粒计算理论的动态多源信息融合方法

数据具有一种特殊的性质——动态性,即数据由于各种原因会发生变化。在生活中这种特殊情况普遍存在,例如气象站收集的天气数据总是随着时间变化而不断变化,导航软件的路线信息也会随着路况、时间等不断发生变化。如果想要更好地研究多源数据的动态性,就需要对数据进行准确且高效地知识发现和更新。近些年,如何设计基于多源动态数据的计算模型和高效的融合算法逐渐成为数据处理领域的研究热点。在单源动态数据中,增量学习方法^[106-107]是一种有效的方法,它可以随着数据收集量的增加而逐渐获得数据中的信息或知识。通过增量学习的方式,根据数据的添加和删除,逐步实现对知识的修改、增强、更新和维护。增量学习方法不仅能使新信息和旧信息有效融合,还能降低对存储和计算时间的要求^[5]。在多源动态数据环境中,增量学习依然可以有效地实现知识获取,但相对于单源数据多源数据的结构更复杂,需要考虑的因素更多。在多源数据的环境中,数据的变化可能是由单一结构改变造成的,也可能是由多结构改变造成的;例如多源信息系统中单一结构变化有属性(特征)、对象(样本)、信源以及特征值等随时间分别变化,多结构变化中又有多种因素同时变化的多源动态信息系统。

针对多源信息系统中结构的变化(图5)可以将多源动态信息融合可以细分为以下5种变化形式:

(1)对象变化的多源动态信息融合。对象变化主要表现为对象数量的增加或减少。目前,基于对象变化的融合机制主要集中在对象变化下的规则提取^[108]、对象变化下的近似更新^[109-110]、对象变化下的属性约简^[111-112]以及对象变化下的特征选择^[113-114]等。另外,文献[115]在多源不完备混合信息系统中添

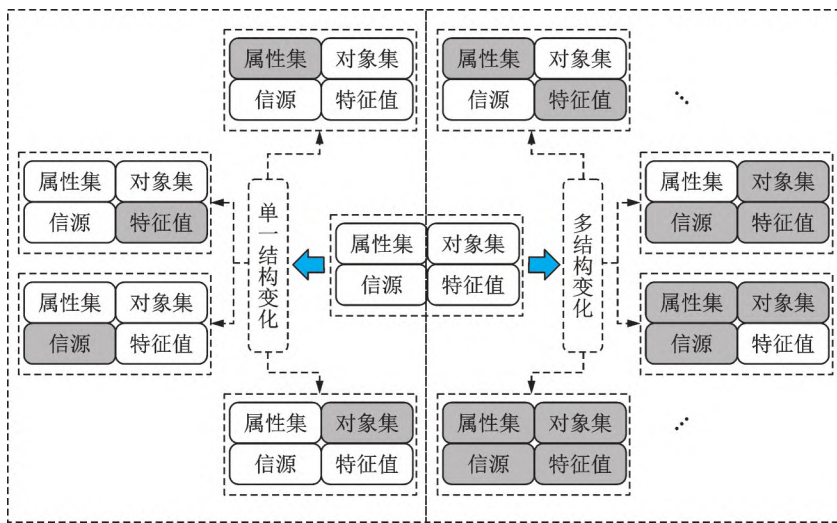


图5 多源信息系统中的结构变化

Fig.5 Structural changes in multi-source information systems

加对象和删除对象时设计了一种基于条件熵信息融合的增量学习方法。

(2)属性集变化的多源动态信息融合。属性集变化通常是指多源信息系统中的特征属性发生变化,例如增加、减少以及子集重组等。在属性集发生变化时,信息粒度、概念近似以及信息系统的属性约简都会发生改变。因此,将变化的信息融合到变化前的信息中有助于实现对知识或信息的维护。目前已有部分学者研究了属性集变化时的融合机制,如基于属性变化的属性约简^[116-117]和近似更新^[118-119]。同时,在其他领域也有类似应用,如文献[120]利用熵的融合理论对P2P网络信任关系的多个互补决策因素进行权重分配,提出了一种动态信任模型。

(3)信源变化的多源动态信息融合。信源变化是指信源的增加或减少或某些信源的数据发生变化等。信源的变化总是伴随着数据维度的变化,以及数据量的大规模增加。因此将变化的信源信息通过一定的方式融合进原始的多源信息系统中有着重要的意义。现有对于信源变化的多源数据研究主要集中于不同的研究背景下进行,如多源区间值数据下的模糊信息粒动态更新机制^[121]、冲突信源下的信源权重动态更新机^[122]、疲劳驾驶识别^[123]、多源数据的特征选择^[124]、交叉路口智能寻路预测^[125]以及更新物理网络的非结构化文本^[126]等。

(4)属性值变化的动态信息融合。属性值变化主要是指信息系统中的具体数据值发生变化而引起多源信息融合的结果产生偏差的一种现象。对于这一现象通常需要对融合结果实现动态更新来减少数据变化对融合结果的影响。一些学者在属性值变化的情况下针对属性约简^[127-128]、近似更新^[129]等进行了一系列的研究。此外,在应用领域也取得了相关成果,如针对海量数据,文献[130]利用D-S证据理论和多源信息融合决策方法建立了分布式、面向对象、面向互联网的动态管理实时数据库。

(5)多结构变化的多源动态信息融合。多结构变化在信息系统中通常是指属性集、对象集、信源以及特征值不再单独变化的情况,其内部的变化机制则更加复杂。对于多结构变化现有的动态融合方法主要集中在对象和属性同时变化^[131-133],对象和属性值同时变化^[134],属性和属性值同时变化^[118],对象、属性和属性值同时变化^[32]。相对而言,目前关于对象、属性、属性值结合信源变化的研究还相对较少。

上述多源动态信息融合不同方法对比如表5所示。在动态多源信息系统中,借用粒计算方法实现对于动态多源数据的融合主要需要考虑的是多源信息系统的变化情形。只有针对相应的结构变化做出动态更新才能更好地实现多源动态信息系统的融合。

表5 多源动态信息融合的不同方法对比

Table 5 Comparison of different dynamic multi-source information methods

动态类型	主要工作
对象变化 ^[115]	在多源不完备混合信息系统中添加对象和删除对象时设计了一种基于条件熵信息融合的增量学习方法
属性集变化 ^[120]	利用熵的融合理论对 P2P 网络信任关系的多个互补决策因素进行权重分配,提出了一种动态信任模型
信息源变化 ^[121-126]	多源区间值数据下的模糊信息粒动态更新机制 ^[121] ,冲突信源下的信源权重动态更新机制 ^[122] ,疲劳驾驶识别 ^[123] ,多源数据的特征选择 ^[124] ,交叉路口智能寻路预测 ^[125] ,更新物理网络的非结构化文本 ^[126]
属性值变化 ^[130]	利用 D-S 证据理论和多源信息融合决策方法建立了分布式、面向对象、面向互联网的动态管理实时数据库
多结构变化 ^[32,118,131-134]	不再局限于单一的维度变化,而是多维度同时发生变化。如:对象和属性同时变化 ^[131-133] ,对象和属性值同时变化 ^[134] ,属性和属性值同时变化 ^[118] ,对象、属性和属性值同时变化 ^[32]

2 基于粒计算的多源信息融合的若干问题

本节对基于经典粗糙集理论、多粒度粗糙集、证据理论、信息熵多源信息融合方法和基于粒计算的动态多源信息融合的研究中存在的问题进行归纳。

(1) 基于一般粗糙集模型的多源信息融合。该方法的重点在于建立多源信息系统的 Pawlak 近似空间^[27]。不管是多源数据中的决策过程^[29]、故障诊断^[31]还是属性选择^[32]都需要确定多源数据的近似空间从发,然后通过不同的关系划分为不同的“颗粒”,进而从不同的粒度中获取知识。此外,粗糙集理论的核心在于如何构建信息系统中的关系^[18],然而不同的信息系统在求解问题时所采用的关系往往不同。这就会引出一个问题:在不同的研究背景下无法使用相同的关系来求解问题时,如何确定在信息系统中所使用关系的适用性或有效性变成了该模型下信息融合的难点。同时,如何利用多源信息系统内部的粒度粗细变化提高多源信息融合的性能也自然成了该模型下信息融合的另一难点。

(2) 基于多粒度粗糙集的多源信息融合。该方法与一般粗糙集理论的不同在于其通过将多源信息系统划分为多个粒度(多子集),并从这个多个粒度中获取信息或知识。从前面的分析来看,不管是在多源数据的决策问题^[45-46]还是为了避免信息丢失直接获取多源数据中的信息^[47-48],多粒度粗糙集理论实现多源信息融合的特点是首先确定研究对象所蕴涵的关系,然后利用不同的多粒度粗糙集模型实现多源信息的融合。利用多粒度粗糙集理论实现多源信息融合都集中于求解在近似空间中研究对象的上下近似。然而,随着多源信息融合领域的进一步发展,越来越多的具体应用场景需要融合,多粒度粗糙集理论面临着不少挑战性的问题,如现有多粒度粗糙集模型或多粒度的思想能否应用于承载不同需求的多源数据,或者如何融合多需求信息等。显然,将多粒度粗糙集模型或多粒度思想融合于多源数据对应的不同具体需求,将是一个非常前景的研究方向。

(3) 基于证据理论的多源信息融合方法。证据理论实现对多源信息融合的核心在于 BPA 的构建、证据合成规则或证据融合规则的选取以及证据可信度的确定。由上节可知 BPA 可以由概率知识^[68-69]、数据的统计分布^[70]等构建,除此之外还可以结合其他方法实现对其的构建^[71-75];证据合成规则选取则基于经典的 D-S 组合规则进行改进^[81-82]或 D-S 组合规则结合其他方法构建^[76-77];对于证据可信度则主要通过获取信源权重的方式对证据可信度进行判断^[78-80,87]。同时,由于对某一事物认知逐渐加深时获取新的知识对于事物认知提升的幅度也在逐渐减少,故当从多源数据中获取证据越来越多时具体问题

的求解精度也逐渐下降。从多源数据中获取更多的证据也意味着需要投入更多的成本实现证据确定,若新证据对问题求解精度提升不大,则该证据实际上是一种无效的证据。因此,研究多源数据中证据的数量和求解精度之间的关系,这也是值得深思的问题。

(4)基于信息熵的多源信息融合。该方法通过信息熵反映多源数据中的信息,并利用多源数据的信息熵实现多源信息融合。在融合过程中会根据某些特定选取规则来选取重要的信息,如选取最小条件熵^[99]。此外也有采用其他拓展的信息熵衡量信息系统中的信息,如粗糙熵^[93]、组合熵^[94]等。使用信息熵实现多源信息融合通常是利用信息的大小衡量数据的重要性,保留重要的数据进行融合。虽然该方法可以保证融合结果能够获取较好的有用信息,但是信息的损失对最终融合效果会产生比较大的影响。因此,如何实现信息熵融合的信息能够更好地代表整个多源信息系统,如何量化选取出的信息对于整个系统的代表程度,都是需要解决的问题。

(5)基于粒计算理论的动态多源信息融合。对于实现多源数据动态信息融合的核心在于如何从更新的动态信息获取有效知识。对于动态多源信息融合的研究主要集中在对象变化^[108-115]、属性变化^[116-120]、信源变化^[121-126]、属性值变化^[127-129]以及多结构变化^[131-133]的动态多源信息融合。由于动态多源数据时刻都有着不同的变化,需要对变化的数据进行准确高效的知识发现和更新,所以处理动态多源信息融合首要需要解决的问题是能否实现多源动态数据的实时计算高效融合。虽然增量学习^[106-107]是一种动态获取知识的有效方法,但多数的增量学习方法很少过滤处理动态数据,这就导致信息系统中会存在大量噪声数据或干扰数据。因此,研究有效过滤动态多源数据理论与方法,也是实现对多源动态数据实时计算和信息高效融合的另一重要课题。

综上,粒计算融合理论方法目前在多源信息融合的研究中仍然面临一定的挑战。同时,随着多源信息融合研究的不断深入还会涌现新的问题,有的问题甚至会覆盖或包含原有的问题,所以上述所提到的若干问题对于现阶段的研究具有一定的理论意义。

3 结束语

本文重点综述了基于经典粗糙集理论、多粒度方法、证据理论和信息熵等粒计算多源信息融合理论和方法,介绍了每一类多源信息融合方法的核心思想及相关研究,并提出了若干问题,为今后多源信息融合的应用研究提供了理论参考。现实应用中,根据不同的需求,信息融合的出发点可能不同,不同的融合方法可根据自身的具体问题采取不同的融合策略,也可以采取多种方法同时结合的方式实现多源信息融合。需要指出的是,除了本文介绍的多源信息融合方法之外,基于粒计算实现多源信息融合的方法还有很多,还存在本文没有总结到的理论和方法,希望本文能够起到引玉之砖的作用,以进一步促进信息融合领域的研究进展,推动粒计算在信息融合领域的理论和应用研究。

参考文献:

- [1] 潘泉. 多源信息融合理论及应用[M]. 北京:清华大学出版社, 2013.
PAN Quan. Multi-source information fusion theory and its applications[M]. Beijing: Tsinghua University Press, 2013.
- [2] 韩崇昭, 朱洪艳, 段战胜. 多源信息融合[M]. 北京: 清华大学出版社, 2006.
HAN Chongzhao, ZHU Hongyan, DUAN Zhansheng. Multi-source information fusion[M]. Beijing: Tsinghua University Press, 2006.
- [3] HALL D L. Mathematical techniques in multisensor data fusion[M]. Fitchburg: Artech House, 2004.
- [4] 何友, 王国宏, 关欣. 信息融合理论及应用[M]. 北京: 电子工业出版社, 2010.
HE You, WANG Guohong, GUAN Xin. Information fusion theory with applications[M]. Beijing: Electronic Industry Press, 2010.
- [5] ZHANG Pengfei, LI Tianrui, WANG Guoqiang, et al. Multi-source information fusion based on rough set theory: A review [J]. Information Fusion, 2021, 68: 85-117.

- [6] 李洋,赵鸣,徐梦瑶,等.多源信息融合技术研究综述[J].智能计算机与应用,2019,9(5): 186-189.
LI Yang, ZHAO Ming, XU Mengyao, et al. A survey of research on multi-source information fusion technology[J]. Intelligent Computer and Applications, 2019, 9(5): 186-189.
- [7] ZADEH L A. Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems[J]. Soft Computing, 1998, 2(1): 23-25.
- [8] GOLAN R H, ZIARKO W. A methodology for stock market analysis utilizing rough set theory[C]//Proceedings of 1995 Conference on Computational Intelligence for Financial Engineering (CIFEr).New York: IEEE, 1995: 32-40.
- [9] LI Zhaowen, ZHANG Pengfei, XIE Ningxin, et al. A novel three-way decision method in a hybrid information system with images and its application in medical diagnosis[J]. Engineering Applications of Artificial Intelligence, 2020, 92: 103651.
- [10] SHI Wengang, WANG Rixin, HUANG Wenhui. Application of rough set theory to fault diagnosis of check valves in reciprocating pumps[C]//Proceedings of Canadian Conference on Electrical and Computer Engineering 2001. Toronto, Canada: IEEE, 2001: 1247-1250.
- [11] PAWLAK Z. Rough sets: Theoretical aspects of reasoning about data[M]. Berlin: Springer Science & Business Media, 1991.
- [12] QIAN Yuhua, LI Shunong, LIANG Jiye, et al. Pessimistic rough set based decisions: A multigranulation fusion strategy[J]. Information Sciences, 2014, 264: 196-210.
- [13] SHAFER G. A mathematical theory of evidence[M]. New Jersey: Princeton University Press, 1976.
- [14] SHANNON C E. A mathematical theory of communication[J]. The Bell System Technical Journal, 1948, 27(3): 379-423.
- [15] YAO Jintao, VASILAKOS A V, PEDRYCZ W. Granular computing: Perspectives and challenges[J]. IEEE Transactions on Cybernetics, 2013, 43(6): 1977-1989.
- [16] 王国胤,张清华,马希骛,等.知识不确定性问题的粒计算模型[J].软件学报,2011,22(4): 676-694.
WANG Guoyin, ZHANG Qinghua, MA Xiao, et al. Granular computing models for knowledge uncertainty[J]. Journal of Software, 2011, 22(4): 676-694.
- [17] 苗夺谦,徐菲菲,姚一豫,等.粒计算的集合论描述[J].计算机学报,2012,35(2): 2351-2363.
MIAO Duoqian, XU Feifei, YAO Yiyu, et al. Set-theoretic formulation of granular computing[J]. Chinese Journal of Computers, 2012, 35(2): 2351-2363.
- [18] 陈德刚,徐伟华,李金海,等.粒计算基础教程[M].北京:科学出版社,2019.
CHEN Degang, XU Weihua, LI Jinhai, et al. Elements of granular computing[M]. Beijing: Science Press, 2019.
- [19] WEI Wei, LIANG Jiye. Information fusion in rough set theory: An overview[J]. Information Fusion, 2019, 48: 107-118.
- [20] SALIDO J M F, MURAKAMI S. Rough set analysis of a general type of fuzzy data using transitive aggregations of fuzzy similarity relations[J]. Fuzzy Sets and Systems, 2003, 139(3): 635-660.
- [21] MIESZKOWICZ-ROLKA A, ROLKA L. Variable precision fuzzy rough sets[C]//Proceedings of Transactions on Rough Sets I. Glasgow, UK: Springer, 2004: 144-160.
- [22] YAO Yiyu. Three-way decisions with probabilistic rough sets[J]. Information Sciences, 2010, 180(3): 341-353.
- [23] PAWLAK Z, SOWINSKI R. Rough set approach to multi-attribute decision analysis[J]. European Journal of Operational Research, 1994, 72(3): 443-459.
- [24] HAN J, CAI Y, CERONE N. Data-driven discovery of quantitative rules in relational databases[J]. IEEE Transactions on Knowledge and Data Engineering, 1993, 5(1): 29-40.
- [25] CHAN C C. A rough set approach to attribute generalization in data mining[J]. Information Sciences, 1998, 107(1/2/3/4): 169-176.
- [26] FAN Tuanfang, LIU Duenren, TZENG G H. Rough set-based logics for multicriteria decision analysis[J]. European Journal of Operational Research, 2007, 182(1): 340-355.
- [27] KHAN M A, BANERJEE M. Multiple-source approximation systems: Membership functions and indiscernibility[C]//Proceedings of International Conference on Rough Sets and Knowledge Technology. Berlin, Heidelberg: Springer, 2008: 80-87.
- [28] HUANG Yanyong, LI Tianrui, LUO Chuan, et al. Dynamic maintenance of rough approximations in multi-source hybrid information systems[J]. Information Sciences, 2020, 530: 108-127.
- [29] SANG Binbin, GUO Yanting, SHI Derong, et al. Decision-theoretic rough set model of multi-source decision systems[J]. International Journal of Machine Learning and Cybernetics, 2018, 9(11): 1941-1954.
- [30] CHEN Xiuwei, XU Weihua. Double-quantitative multigranulation rough fuzzy set based on logical operations in multi-source

- decision systems[J]. International Journal of Machine Learning and Cybernetics, 2021, 13: 1021-1048.
- [31] 杨广, 吴晓平, 宋业新, 等. 基于粗糙集理论的多源信息融合故障诊断方法[J]. 系统工程与电子技术, 2009(8): 2013-2019.
YANG Guang, WU Xiaoping, SONG Yexin, et al. Multi-sensor information fusion fault diagnosis method based on rough set theory[J]. Systems Engineering and Electronics, 2009(8): 2013-2019.
- [32] ZHANG Pengfei, LI Tianrui, YUAN Zhong, et al. A data-level fusion model for unsupervised attribute selection in multi-source homogeneous data[J]. Information Fusion, 2022, 80: 87-103.
- [33] XU Weihua, YU Jianhang. A novel approach to information fusion in multi-source datasets: A granular computing viewpoint[J]. Information Sciences, 2017, 378: 410-423.
- [34] LI Tongying, FEI Minrui. Information fusion in wireless sensor network based on rough set[C]//Proceedings of ICNIDC. Beijing: IEEE, 2009: 129-134.
- [35] YAO Yiyang, WANG Zhiqiang, GAN Chun, et al. Multi-source alert data understanding for security semantic discovery based on rough set theory[J]. Neurocomputing, 2016, 208: 39-45.
- [36] WANG Pei, ZHANG Pengfei, LI Zhaowen. A three-way decision method based on Gaussian kernel in a hybrid information system with images: An application in medical diagnosis[J]. Applied Soft Computing, 2019, 77: 734-749.
- [37] 董广军, 张永生, 戴晨光, 等. 基于粗糙集的多源信息融合处理技术[J]. 仪器仪表学报, 2005, 26(S1): 570-571.
DONG Guangjun, ZHANG Yongsheng, DAI Chenguang, et al. The processing of information fusion based on rough set theory[J]. Institute of Surveying and Mapping, 2005, 26(S1): 570-571.
- [38] 梁燕飞, 何汉武, 郑德涛, 等. 基于可变精度粗糙集的驾驶员多源信息融合研究[J]. 计算机工程与应用, 2008, 44(10): 192-194.
LIANG Yanfei, HE Hanwu, ZHENG Detao, et al. Research on driver multi-information fusion based on variable precision rough set[J]. Computer Engineering and Applications, 2008, 44(10): 192-194.
- [39] LIU Yunxiang, CHEN Yan, YUAN Xinxin. Study of information fusion evaluation methodology based on multiple approximation space[C]//Proceedings of 2010 2nd International Symposium on Information Engineering and Electronic Commerce. Qingdao, China: IEEE, 2010: 1-4.
- [40] XU Xiaomin, WANG Qiang, NIU Dongxiao, et al. Synergistic effect evaluation of main and auxiliary industry of power grid based on the information fusion technology from the perspective of sustainable development of enterprises[J]. Sustainability, 2018. DOI: <https://doi.org/10.3390/su10020457>.
- [41] 徐伟华. 序信息系统与粗糙集[M]. 北京: 科学出版社, 2013.
XU Weihua. Ordered information system and rough set[M]. Beijing: Science Press, 2013.
- [42] QIAN Yuhua, LIANG Jiye. Rough set method based on multi-granulations[C]//Proceedings of 2006 5th IEEE International Conference on Cognitive Informatics. Beijing, China: IEEE, 2006, 1: 297-304.
- [43] QIAN Y H, LIANG J Y, DANG C Y. MGRS in incomplete information systems[C]//Proceeding of 2007 IEEE International Conference on Granular Computing (GRC 2007). [S.l.]: IEEE, 2007: 163-163.
- [44] YANG Xibei, QI Yunsong, SONG Xiaoning, et al. Test cost sensitive multigranulation rough set: Model and minimal cost selection[J]. Information Sciences, 2013, 250: 184-199.
- [45] SUN Bingzhen, MA Weimin, CHEN Xiangtang, et al. Heterogeneous multigranulation fuzzy rough set-based multiple attribute group decision making with heterogeneous preference information[J]. Computers & Industrial Engineering, 2018, 122: 24-38.
- [46] MANDAL P, RANADIVE A S. Multi-granulation Pythagorean fuzzy decision-theoretic rough sets based on inclusion measure and their application in incomplete multi-source information systems[J]. Complex & Intelligent Systems, 2019, 5(2): 145-163.
- [47] YANG Lei, XU Weihua, ZHANG Xiaoyan, et al. Multi-granulation method for information fusion in multi-source decision information system[J]. International Journal of Approximate Reasoning, 2020, 122: 47-65.
- [48] SANG Binbin, YANG Lei, CHEN Hongmei, et al. Generalized multi-granulation double-quantitative decision-theoretic rough set of multi-source information system[J]. International Journal of Approximate Reasoning, 2019, 115: 157-179.
- [49] LIN Guoping, LIANG Jiye, QIAN Yuhua. An information fusion approach by combining multigranulation rough sets and evidence theory[J]. Information Sciences, 2015, 314: 184-199.
- [50] 韩瑞. 面向不完备多源信息系统的多粒度融合方法研究[D]. 太原: 山西大学, 2014.
HAN Rui. Research on incomplete multi-source information system for multi-granularity fusion[D]. Taiyuan: Shanxi University, 2014.

- ty, 2014.
- [51] CHE Xiaoya, MI Jusheng, CHEN Degang. Information fusion and numerical characterization of a multi-source information system[J]. *Knowledge-Based Systems*, 2018, 145: 121-133.
 - [52] GUO Yanting, TSANG E C C, XU Weihua, et al. Adaptive weighted generalized multi-granulation interval-valued decision-theoretic rough sets[J]. *Knowledge-Based Systems*, 2020, 187: 104804.
 - [53] YANG Lei, ZHANG Xiaoyan, XU Weihua, et al. Multi-granulation rough sets and uncertainty measurement for multi-source fuzzy information system[J]. *International Journal of Fuzzy Systems*, 2019, 21(6): 1919-1937.
 - [54] YIN Tao, MAO Xiaojuan, ZHANG Ying, et al. Decision-theoretic rough set: A fusion strategy[J]. *IEEE Access*, 2020, 8: 221027-221038.
 - [55] 庄颖, 刘文奇, 范敏, 等. 集值信息系统上的多粒度优势关系与信息融合[J]. *模式识别与人工智能*, 2015, 28(8): 741-749.
ZHUANG Ying, LIU Wenqi, FAN min, et al. Multi-granulation dominance relation and information fusion based on set-valued information system[J]. *Pattern Recognition and Artificial Intelligence*, 2015, 28(8): 741-749.
 - [56] DEMPSTER A P. Upper and lower probabilities induced by a multivalued mapping[C]//*Proceedings of Classic Works of the Dempster-Shafer Theory of Belief Functions*. Berlin, Heidelberg: Springer, 2008: 57-72.
 - [57] SHAFER G. A mathematical theory of evidence[M]. Princeton: Princeton University Press, 1976.
 - [58] 易革军. 基于D-S证据理论的信息融合方法研究与应用[D]. 长沙: 湖南大学, 2015.
YI Gejun. Research and application of information fusion based on D-S evidence theory[D]. Changsha: Hunan University, 2015.
 - [59] 李海生. 基于证据理论的分类方法研究[D]. 广州: 华南理工大学, 2013.
LI Haisheng. Research of classification methods based on evidence theory[D]. Guangzhou: South China University of Technology, 2013.
 - [60] HANSEN L K, SALAMON P. Neural network ensembles[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2002, 12(10): 993-1001.
 - [61] 桂劲松, 陈志刚, 邓晓衡, 等. 基于D-S证据理论的网格服务行为信任模型[J]. *计算机工程与应用*, 2007, 43(2): 25-28.
GUI Jinsong, CHEN Zhigang, DENG Xiaoheng, et al. Behavior trust model for grid services based on D-S evidence theory[J]. *Computer Engineering and Applications*, 2007, 43(2): 25-28.
 - [62] SMETS P. Analyzing the combination of conflicting belief functions[J]. *Information Fusion*, 2007, 8(4): 387-412.
 - [63] 邓勇, 韩德强. 广义证据理论中的基本概率指派生成方法[J]. *西安交通大学学报*, 2011, 45(2): 35-38.
DENG Yong, HAN Deqiang. Methods to determine generalized basic probability assignment in generalized evidence theory[J]. *Journal of Xi'an Jiaotong University*, 2011, 45(2): 35-38.
 - [64] 邓勇, 朱振福, 钟山. 基于证据理论的模糊信息融合及其在目标识别中的应用[J]. *航空学报*, 2005, 26(6): 754-758.
DENG Yong, ZHU Zhenfu, ZHONG Shan. Fuzzy information fusion based on evidence theory and its application in target recognition[J]. *Acta Aeronautica et Astronautica Sinica*, 2005, 26(6): 754-758.
 - [65] 苏晓燕, 邓勇, 吴英, 等. 基于改进D-S组合规则的故障模式分类[J]. *振动、测试与诊断*, 2011, 31(2): 144-149.
SU Xiaoyan, DENG Yong, WU Ying, et al. Fault pattern classification using modified Dempster-Shafer (D-S) combination rule[J]. *Journal of Vibration Measurement & Diagnosis*, 2011, 31(2): 144-149.
 - [66] YANG Guang, YU Shuofeng, LU Shan, et al. The comprehensive diagnostic method combining rough sets and evidence theory[J]. *Applied Mathematics and Nonlinear Sciences*, 2021, 6(2): 171-180.
 - [67] ZADEH L A. Review of a mathematical theory of evidence[J]. *AI Magazine*, 1984, 5(3): 81.
 - [68] BOUDRAA A O, BENTABET L, SALZENSTEIN F. Dempster-Shafer's basic probability assignment based on fuzzy membership functions[J]. *ELCVIA: Electronic Letters on Computer Vision and Image Analysis*, 2004, 4(1): 1-10.
 - [69] ROMBAUT M, ZHU Yuemin. Study of Dempster-Shafer theory for image segmentation applications[J]. *Image and Vision Computing*, 2002, 20(1): 15-23.
 - [70] MANDLER E, SCHÜMANN J. Combining the classification results of independent classifiers based on the Dempster/Shافر theory of evidence[C]//*Proceedings of Machine Intelligence and Pattern Recognition*. North-Holland: [s.n.], 1988, 7: 381-393.
 - [71] LIU Yuting, PAL N R, MARATHE A R, et al. Weighted fuzzy Dempster-Shafer framework for multimodal information integration[J]. *IEEE Transactions on Fuzzy Systems*, 2017, 26(1): 338-352.
 - [72] QIN Bowen, XIAO Fuyuan. An improved method to determine basic probability assignment with interval number and its application in classification[J]. *International Journal of Distributed Sensor Networks*, 2019. DOI: <https://doi.org/10.1177/>

- 1550147718820542.
- [73] ZHANG Hepeng, DENG Yong. Weighted belief function of sensor data fusion in engine fault diagnosis[J]. *Soft Computing*, 2020, 24(3): 2329-2339.
- [74] TANG Yongchuan, WU Dongdong, LIU Zijing. A new approach for generation of generalized basic probability assignment in the evidence theory[J]. *Pattern Analysis and Applications*, 2021, 24(3): 1007-1023.
- [75] XU Yaoyu, LI Yuan, WANG Yijing, et al. Integrated decision-making method for power transformer fault diagnosis via rough set and DS evidence theories[J]. *IET Generation, Transmission & Distribution*, 2020, 14(24): 5774-5781.
- [76] WANG Hongfei, DENG Xinyang, JIANG Wen, et al. A new belief divergence measure for Dempster-Shafer theory based on belief and plausibility function and its application in multi-source data fusion[J]. *Engineering Applications of Artificial Intelligence*, 2021, 97: 104030.
- [77] LI Shucan, LIU Cong, ZHOU Zongqing, et al. Multi-sources information fusion analysis of water inrush disaster in tunnels based on improved theory of evidence[J]. *Tunnelling and Underground Space Technology*, 2021, 113: 103948.
- [78] WANG Zhe, XIAO Fuyuan. An improved multi-source data fusion method based on the belief entropy and divergence measure [J]. *Entropy*, 2019. DOI:<https://doi.org/10.3390/e21060611>.
- [79] JIANG Wen, WEI Boya, XIE Chunhe, et al. An evidential sensor fusion method in fault diagnosis[J]. *Advances in Mechanical Engineering*, 2016. DOI:<https://doi.org/10.1177/1687814016641820>.
- [80] CHEN Yong, TANG Yongchuan, LEI Yan. An improved data fusion method based on weighted belief entropy considering the negation of basic probability assignment[J]. *Journal of Mathematics*, 2020(1): 1-11.
- [81] YAGER R R. On the Dempster-Shafer framework and new combination rules[J]. *Information Sciences*, 1987, 41(2): 93-137.
- [82] LEFEVRE E, COLOT O, VANNOORENBERGHE P. Belief function combination and conflict management[J]. *Information Fusion*, 2002, 3(2): 149-162.
- [83] 何兵, 胡红丽. 一种修正的 DS 证据融合策略[J]. *航空学报*, 2003, 24(6): 559-562.
HE Bing, HU Hongli. A modified DS evidence combination strategy[J]. *Acta Aeronautica et Astronautica Sinica*, 2003, 24(6): 559-562.
- [84] 何兵, 胡红丽. 一种分级的 DS 证据合成策略[J]. *计算机工程与应用*, 2004, 40(10): 88-90, 98.
HE Bing, HU Hongli. Multi-level DS evidence combination strategy[J]. *Computer Engineering and Applications*, 2004, 40(10): 88-90, 98.
- [85] MURPHY C K. Combining belief functions when evidence conflicts[J]. *Decision Support Systems*, 2000, 29(1): 1-9.
- [86] DENG Yong, SHI Wenkang, ZHU Zhenfu, et al. Combining belief functions based on distance of evidence[J]. *Decision Support Systems*, 2004, 38(3): 489-493.
- [87] CHEN Jie, YE Fang, JIANG Tao, et al. Conflicting information fusion based on an improved DS combination method[J]. *Symmetry*, 2017. DOI:<https://doi.org/10.3390/sym9110278>.
- [88] 李芋均, 郭红梅, 黄丁发, 等. 基于区间证据理论的多源灾情信息融合及烈度判定研究[J]. *计算机应用研究*, 2020, 37(S2): 92-94.
LI Yujun, GUO Hongmei, HUANG Dingfa, et al. Information fusion and intensity determination of multi-source earthquake disaster based on interval evidence theory[J]. *Application Research of Computers*, 2020, 37(S2): 92-94.
- [89] 安春莲, 黄静, 吴耀云. 基于证据理论的多源信息融合模型[J]. *电子信息对抗技术*, 2017, 32(1): 23-26.
AN Chunlian, HUANG Jing, WU Yaoyun. A multi-source information fusion model based on evidence theory[J]. *Electronic Information Warfare Technology*, 2017, 32(1): 23-26.
- [90] 程翠萍. 面向多源信息融合的证据距离算法研究[D]. 重庆: 西南大学, 2021.
CHENG Cuiping. Research on evidence distance algorithm and application for multi-source information fusion[D]. Chongqing: Southwest University, 2021.
- [91] SHANNON C E, WEAVER W. The mathematical theory of communication[J]. *Bell Labs Technical Journal*, 1950, 3(9): 31-32.
- [92] DÜNTSCH I, GEDIGA G. Uncertainty measures of rough set prediction[J]. *Artificial Intelligence*, 1998, 106(1): 109-137.
- [93] BEAUBOUF T, PETRY F E, ARORA G. Information-theoretic measures of uncertainty for rough sets and rough relational databases[J]. *Information Sciences*, 1998, 109(1/2/3/4): 185-195.
- [94] QIAN Yuhua, LIANG Jiye. Combination entropy and combination granulation in rough set theory[J]. *International Journal of*

- Uncertainty, Fuzziness and Knowledge-Based Systems, 2008, 16(2): 179-193.
- [95] LIANG Jiye, SHI Zhongzhi. The information entropy, rough entropy and knowledge granulation in rough set theory[J]. International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, 2004, 12(1): 37-46.
- [96] LIANG Jiye, SHI Zhongzhi. Information entropy, rough entropy and knowledge granulation in incomplete information systems [J]. International Journal of General Systems, 2006, 35(6): 641-654.
- [97] XU Weihua, ZHANG Xiaoyan, ZHANG Wenxiu. Knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems[J]. Applied Soft Computing, 2009, 9(4): 1244-1251.
- [98] ZHU Ping, WEN Qiaoyan. Information-theoretic measures associated with rough set approximations[J]. Information Sciences, 2012, 212: 33-43.
- [99] XU Weihua, LI Mengmeng, WANG Xizhao. Information fusion based on information entropy in fuzzy multi-source incomplete information system[J]. International Journal of Fuzzy Systems, 2017, 19(4): 1200-1216.
- [100] HAN Shan, CHEN Lin, ZHANG Zhi, et al. A new information filling technique based on generalized information entropy[J]. International Journal of Computers Communications & Control, 2014, 9(2): 172-186.
- [101] 万树平, 潘鹏. 基于信息熵的多传感器数据的融合方法[J]. 传感器与微系统, 2008, 27(5): 64-68.
WAN Shuping, PAN Peng. Method of fusion for multi-sensor data based on entropy[J]. Transducer and Microsystem Technologies, 2008, 27(5): 64-68.
- [102] 高崎, 曹琳, 熊学军. 基于信息熵的溶解氧传感器数据融合处理方法[J]. 山东科技大学学报(自然科学版), 2019, 38(6): 67-73, 80.
GAO Hao, CAO Lin, XIONG Xuejun. Data fusion algorithm of dissolved oxygen sensor based on information entropy[J]. Journal of Shandong University of Science and Technology (Natural Science), 2019, 38(6): 67-73, 80.
- [103] CHEN Y, ORADY E. An entropy-based index evaluation scheme for multiple sensor fusion in classification process[C]// Proceedings of International Mechanical Engineering Congress and Exposition. [S.l.]: American Society of Mechanical Engineers, 1997: 3-8.
- [104] QIAN Wenbin, YU Sudan, YANG Jun, et al. Multi-label feature selection based on information entropy fusion in multi-source decision system[J]. Evolutionary Intelligence, 2020, 13(2): 255-268.
- [105] ZENG Xin, XIONG Xingzhong, LUO Zhongqiang. Grid fault diagnosis based on information entropy and multi-source information fusion[J]. International Journal of Electronics and Telecommunications, 2021, 67(2): 143-148.
- [106] LI Tianrui, RUAN D, GEERT W, et al. A rough sets based characteristic relation approach for dynamic attribute generalization in data mining[J]. Knowledge-Based Systems, 2007, 20(5): 485-494.
- [107] CHEN Hongmei, LI Tianrui, LUO Chuan, et al. A rough set-based method for updating decision rules on attribute values' coarsening and refining[J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(12): 2886-2899.
- [108] HU Chengxiang, ZHANG Li. A dynamic framework for updating neighborhood multigranulation approximations with the variation of objects[J]. Information Sciences, 2020, 519: 382-406.
- [109] CHEN H M, LI T R, RUAN D, et al. A rough-set-based incremental approach for updating approximations under dynamic maintenance environments[J]. IEEE Transactions on Knowledge and Data Engineering, 2011, 25(2): 274-284..
- [110] GE Hao, YANG Chuanjian, XU Yi. Incremental updating three-way regions with variations of objects and attributes in incomplete neighborhood systems[J]. Information Sciences, 2022, 584: 479-502.
- [111] GUO Yanting, TSANG E C C, HU Meng, et al. Incremental updating approximations for double-quantitative decision-theoretic rough sets with the variation of objects[J]. Knowledge-Based Systems, 2020, 189: 105082.
- [112] YANG Yanyan, CHEN Degang, WANG Hui. Active sample selection based incremental algorithm for attribute reduction with rough sets[J]. IEEE Transactions on Fuzzy Systems, 2016, 25(4): 825-838.
- [113] LIANG Jiye, WANG Feng, DANG Chuangyin, et al. A group incremental approach to feature selection applying rough set technique[J]. IEEE Transactions on Knowledge and Data Engineering, 2012, 26(2): 294-308.
- [114] YANG Yanyan, CHEN Degang, WANG Hui, et al. Incremental perspective for feature selection based on fuzzy rough sets[J]. IEEE Transactions on Fuzzy Systems, 2017, 26(3): 1257-1273.
- [115] HUANG Qianqian, LI Tianrui, YANG Xin, et al. Dynamical information fusion of multisource incomplete hybrid information systems based on conditional entropy[C]//Proceedings of International Conference on Intelligent Systems and Knowledge Engineering (ISKE). Dalian, China: IEEE, 2019: 1-8.

- [116] WANG Feng, LIANG Jiye, QIAN Yuhua. Attribute reduction: A dimension incremental strategy[J]. Knowledge-Based Systems, 2013, 39: 95-108.
- [117] SHU Wenhao, SHEN Hong. Updating attribute reduction in incomplete decision systems with the variation of attribute set[J]. International Journal of Approximate Reasoning, 2014, 55(3): 867-884.
- [118] WANG Shu, LI Tianrui, LUO Chuan, et al. A novel approach for efficient updating approximations in dynamic ordered information systems[J]. Information Sciences, 2020, 507: 197-219.
- [119] LIU Dun, LI Tianrui, ZHANG Junbo. Incremental updating approximations in probabilistic rough sets under the variation of attributes[J]. Knowledge-Based Systems, 2015, 73: 81-96.
- [120] LI Xiaoyong, ZHOU Feng, YANG Xudong. Developing dynamic P2P trust model using theory of entropy-based multi-source information fusion[J]. International Journal of Innovative Computing, Information and Control, 2011, 7(2): 777-790.
- [121] HUANG Yanyong, LI Tianrui, LUO Chuan, et al. Dynamic fusion of multisource interval-valued data by fuzzy granulation[J]. IEEE Transactions on Fuzzy Systems, 2018, 26(6): 3403-3417.
- [122] LI Yuting, XIAO Fuyuan. A novel dynamic weight allocation method for multisource information fusion[J]. International Journal of Intelligent Systems, 2021, 36(2): 736-756.
- [123] SUN Wei, ZHANG Xiaorui, PEETA S, et al. A self-adaptive dynamic recognition model for fatigue driving based on multi-source information and two levels of fusion[J]. Sensors, 2015, 15(9): 24191-24213.
- [124] HUANG Yanyong, GUO Kejun, YI Xiuwen, et al. Matrix representation of the conditional entropy for incremental feature selection on multi-source data[J]. Information Sciences, 2022, 591: 263-286.
- [125] DUBEY R K, SOHN S S, HOELSCHER C, et al. Fusion-based way finding prediction model for multiple information sources[C]//Proceedings of International Conference on Information Fusion (FUSION). Ottawa, Canada: IEEE, 2019: 1-8.
- [126] KEONG C H, HU Z X, BENG L H. Fusion of simplified entity networks from unstructured text[C]//Proceedings of International Conference on Information Fusion. Chicago, USA: IEEE, 2011: 1-7.
- [127] SHU Wenhao, SHEN Hong. Incremental feature selection based on rough set in dynamic incomplete data[J]. Pattern Recognition, 2014, 47(12): 3890-3906.
- [128] WEI Wei, WU Xiaoying, LIANG Jiye, et al. Discernibility matrix based incremental attribute reduction for dynamic data[J]. Knowledge-Based Systems, 2018, 140: 142-157.
- [129] ZENG Anping, LI Tianrui, HU Jie, et al. Dynamical updating fuzzy rough approximations for hybrid data under the variation of attribute values[J]. Information Sciences, 2017, 378: 363-388.
- [130] XU Lijuan, ZHANG Lihong, DU Zhenhua. Coastal ecological environment monitoring and protection system based on multisource information fusion decision[J]. Journal of Sensors, 2021, 2021: 1-15.
- [131] CHEN Hongmei, LI Tianrui, LUO Chuan, et al. A decision-theoretic rough set approach for dynamic data mining[J]. IEEE Transactions on Fuzzy Systems, 2015, 23(6): 1958-1970.
- [132] HUANG Yanyong, LI Tianrui, LUO Chuan, et al. Matrix-based dynamic updating rough fuzzy approximations for data mining [J]. Knowledge-Based Systems, 2017, 119: 273-283.
- [133] YANG Xin, LI Tianrui, LIU Dun, et al. A unified framework of dynamic three-way probabilistic rough sets[J]. Information Sciences, 2017, 420: 126-147.
- [134] WANG Shu, LI Tianrui, LUO Chuan, et al. Domain-wise approaches for updating approximations with multi-dimensional variation of ordered information systems[J]. Information Sciences, 2019, 478: 100-124.

作者简介:



徐伟华(1979-),通信作者,男,教授,博士生导师,研究方向:确定性人工智能、智能计算、粒计算、数据挖掘、知识发现、机器学习、认知计算、不确定性处理等,E-mail:chxuwh@gmail.com。



黄旭东(1996-),男,硕士研究生,研究方向:粒计算、多源信息融合、数据挖掘等,E-mail:hxdml@gmail.com。



蔡可(1997-),女,硕士研究生,研究方向:粒计算、多源信息融合等,E-mail:2279732003@qq.com。

(编辑:刘彦东)