

针对模糊数据近似处理的阴影集研究综述

高满^{1,2,3,4} 张清华^{1,2,3,4} 王国胤^{3,5} 姚一豫⁶

摘要 阴影集 (Shadowed set, SS) 是一种对模糊集进行三支近似处理的不确定性知识发现模型, 其能够对模糊集中具有精确值的不确定性对象进行有效的近似和划分, 从而减少模糊决策过程中不确定性对象的决策划分成本和计算损耗. 首先, 回顾阴影集的发展历程, 并从四个方面介绍其研究现状及内容, 即阴影集的模式构建、理论性质、数据分析以及应用研究. 通过总结分析它们的核心思想、方法体系、相互关系和区别等, 为该领域的后续研究提供借鉴. 随后, 讨论分析阴影集理论与其他不确定性问题处理理论模型的联系, 尤其是阴影集与模糊集、粗糙集和三支决策理论之间的区别、联系以及互补性. 最后, 围绕上述四个研究方面, 对当前若干具有挑战性的研究问题进行分析 and 展望.

关键词 阴影集, 模糊集, 三支决策, 不确定性, 粒计算

引用格式 高满, 张清华, 王国胤, 姚一豫. 针对模糊数据近似处理的阴影集研究综述. 自动化学报, 2024, 50(10): 1906–1927

DOI 10.16383/j.aas.c230808 **CSTR** 32138.14.j.aas.c230808

Review of Shadowed Set for Fuzzy Data Approximation Processing

GAO Man^{1,2,3,4} ZHANG Qing-Hua^{1,2,3,4} WANG Guo-Yin^{3,5} YAO Yi-Yu⁶

Abstract Shadowed set (SS) is a kind of uncertain knowledge discovery model which carries out three-way approximate processing on fuzzy sets. It can effectively approximate and partition the uncertain objects with precise values in fuzzy sets, so as to reduce the decision partitioning cost and calculation loss of uncertain objects in fuzzy decision-making process. Firstly, the development of shadowed set is reviewed, and introduces its research status and content from four aspects: Model construction, theoretical properties, data analysis and application research. By summarizing and analyzing their core ideas, methodological systems, interrelationships, and differences, etc., this paper provides reference for subsequent research in this field. Subsequently, the connection between shadowed set theory and other uncertainty problem handling theoretical models is discussed and analyzed, especially the differences, connections and complementarities between shadowed set and fuzzy sets, rough sets, and three-way decision theories. Finally, based on the above four research aspects, some current challenging research problems are analyzed and prospected.

Key words Shadowed set (SS), fuzzy sets, three-way decision, uncertainty, granular computing

Citation Gao Man, Zhang Qing-Hua, Wang Guo-Yin, Yao Yi-Yu. Review of shadowed set for fuzzy data approximation processing. *Acta Automatica Sinica*, 2024, 50(10): 1906–1927

收稿日期 2024-01-04 录用日期 2024-04-19

Manuscript received January 4, 2024; accepted April 19, 2024

国家重点研发计划 (2021YFF0704101), 国家自然科学基金 (62276038, 62221005), 重庆市创新群体研究项目 (cstc2019jcyjxttX0002), 重庆邮电大学博士人才培养计划 (BYJS202109), 重庆市教委重点合作项目 (HZ2021008) 资助

Supported by National Key Research and Development Program of China (2021YFF0704101), National Natural Science Foundation of China (62276038, 62221005), Foundation for Innovative Research Groups of Natural Science Foundation of Chongqing (cstc2019jcyjxttX0002), Doctoral Talent Training Program of Chongqing University of Posts and Telecommunications (BYJS202109), and Key Cooperation Project of Chongqing Municipal Education Commission (HZ2021008)

本文责任编辑 胡清华

Recommended by Associate Editor HU Qing-Hua

1. 重庆邮电大学计算机科学与技术学院 重庆 400065 中国 2. 重庆邮电大学计算智能重庆市重点实验室 重庆 400065 中国 3. “旅游多源数据感知与决策技术”文化和旅游部重点实验室 重庆 400065 中国 4. 网络空间大数据智能安全教育部重点实验室 重庆 400065 中国 5. 重庆师范大学重庆国家应用数学中心 重庆 401331 中国 6. 里贾纳大学计算机科学系 里贾纳 S4S 0A2 加拿大

1. School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065,

当前智能信息时代, 数据体量高速增长, 数据特性纷繁复杂. 如何从大数据中去粗取精、去伪存真地挖掘出有效的知识是数据挖掘领域所关注的关键问题. 而如何从海量的具有不确定性的数据信息中挖掘和提炼出确定性的知识则是知识发现领域和认知科学领域所关注的关键问题. 因此, 对于不确定性信息的转化和处理的研究就受到了诸多学者的关注. 基于此, 1965 年, Zadeh^[1] 提出模糊集 (Fuzzy sets) 理论, 其能够有效地刻画和处理不确定性信

China 2. Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China 3. “Tourism Multi-source Data Perception and Decision Technology” Key Laboratory of the Ministry of Culture and Tourism, Chongqing 400065, China 4. Key Laboratory of Cyberspace Big Data Intelligent Security, Ministry of Education, Chongqing 400065, China 5. National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing 401331, China 6. Department of Computer Science, University of Regina, Regina S4S 0A2, Canada

息. Zadeh^[1] 认为任一对象总在一定程度上属于某一个集合或目标概念, 也可以在不同程度上属于多个集合或目标概念. 因此, 其给出了“隶属度”的定义, 通过在有限非空论域 $U = \{x_1, x_2, \dots, x_n\}$ 上给定一个映射函数 $\mu_A: U \rightarrow [0, 1], x \rightarrow \mu_A(x)$ 来描述一个对象 x 隶属于某一目标概念的程度, 称 A 为 U 上的一个模糊集, 称 $\mu_A(x)$ 为 A 的隶属度函数, 其表示对象 x 隶属于 A 的程度, 其中 $0 \leq \mu_A(x) \leq 1$. 然而, 精确的隶属度值就像是一把双刃剑. 虽然模糊集能够通过隶属度函数和精确的隶属度值来有效地刻画一个模糊性概念, 但是对于部分隶属度差距非常小的对象, 有时无需进行严格区分, 即对于一般目标概念的刻画和描述并不总是需要完全精确的隶属度, 反而可以通过一些特定值来表示那些非常相似的对象. 这与人类习惯通过已有知识来近似刻画未知事物的认知思想相类似. 此外, 较高的精度需求也伴随着巨大的计算损耗, 这在模糊集数据体量较大时尤其明显, 也极大地影响了模糊集理论的应用^[2].

为解决上述问题, Pedrycz^[2] 提出阴影集 (Shadowed set, SS) 理论: 一种针对模糊信息数据分析的三支近似表达模型. 它可以有效地对模糊集中具有精确值的不确定性对象进行近似表达, 从而减少模糊决策过程中不确定性对象的决策划分成本和计算损耗. 如对于信息检索问题, 可以通过一对给定的近似划分阈值过滤掉大量与检索目标无关的记录, 以快速响应模糊查询的需求^[2]. 而阴影集的核心思想便是通过一对决策阈值 (α, β) ($0 \leq \beta < \alpha \leq 1$) 将模糊集 A 中具有不确定性的对象 x 划分到三个互不相交的区域, 即提升域 ($\mu_A(x) \geq \alpha$)、减小域 ($\mu_A(x) \leq \beta$) 和阴影域 ($\beta < \mu_A(x) < \alpha$). 其中这三个区域分别用常数 1、0 以及区间值 $[0, 1]$ 来表示, $\mu_A(x_i)$ 表示对象 x_i 隶属于模糊集 A , 即目标概念的程度. 阴影集通过这种不太精确的、粗糙的三支近似表达形式替换掉原模糊集中精确值的表达形式, 这有效减少了模糊集中的冗余信息, 即可以理解为一种针对精确值数据信息的粒化处理手段, 这也使得阴影集相比模糊集在处理不确定性信息时更加简洁和高效.

由于阴影集在不确定性信息处理过程中所表现出来的简单性和有效性, 使其吸引了大量学者展开研究, 并将其应用到诸多领域. 对此, 本文通过梳理当前研究现状, 将围绕阴影集所展开的研究工作归纳为四个方面, 如图 1 所示, 即: 1) 阴影集的模式构建; 2) 阴影集的理论性质; 3) 阴影集的数据分析; 4) 阴影集的应用研究. 其中, 在阴影集的模式构建方面, 主要关注: 1) 阴影集的表达形式, 尤其是

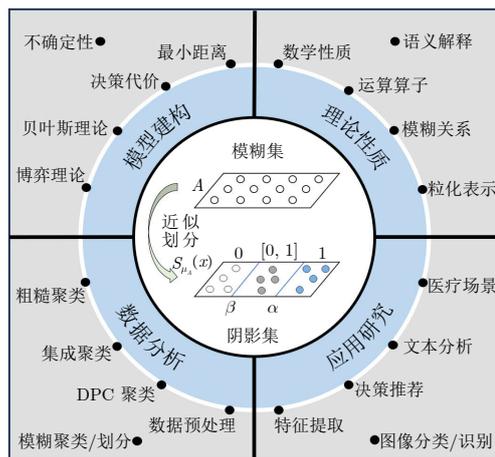


图 1 阴影集研究的四个主要方面

Fig. 1 Four main aspects of shadowed set research

阴影域的表达; 2) 决策阈值对 (α, β) 的求取方法和语义解释. 在阴影集的理论性质方面, 主要关注: 1) 基于阴影集的模糊关系、结构解释、逻辑运算或数学性质等; 2) 基于不同模糊集下的阴影集构建分析. 在阴影集的数据分析方面, 主要关注: 1) 基于阴影集的聚类/分类算法; 2) 基于阴影集的数据预处理. 在阴影集的应用研究方面, 主要关注: 1) 图像分类和识别; 2) 医疗信息处理; 3) 文本分析; 4) 决策推荐.

本文的研究贡献总结如下: 1) 梳理了阴影集模型的历史发展脉络, 总结了不同阴影集模型之间的区别和联系, 介绍了不同模型的核心思想和构建步骤; 2) 按类别和时间对当前阴影集在各个方面的研究现状以及不同研究工作之间的递进关系进行了分析和总结; 3) 分析了阴影集与其他不确定性问题处理理论之间的联系, 尤其是相关联的模糊集、粗糙集以及三支决策理论, 并通过应用实例分析了其相互之间的互补性; 4) 深入分析了当前阴影集研究领域中所存在的挑战性问题, 以及潜在的研究方向和可能的研究思路. 本文的逻辑性研究框图如图 2 所示.

1 研究现状

1.1 阴影集的模式构建研究

1998 年, Pedrycz^[2] 首次提出阴影集理论, 即 Pedrycz 经典阴影集, 并通过构造一对决策划分阈值 (α, β) 将模糊集合近似划分为三个部分. 随后, Pedrycz 和 Vukovich^[3] 讨论了其作为粒计算可能载体的作用, 并分析了其与模糊集之间的关系^[4]. Cattaneo 和 Ciucci^[5-6] 研究了阴影集的代数结构, 并在阴影集表达形式基础上首次提出用常数 0.5 来近似表达阴影域, 即 0.5SS (0.5 shadowed set). 随后, Deng

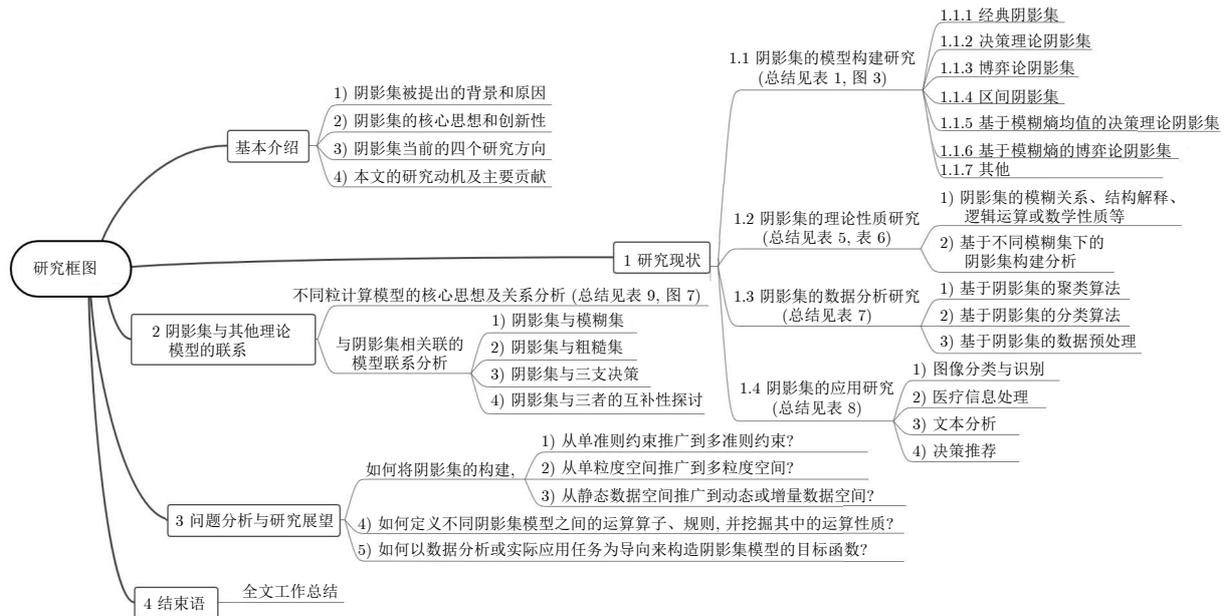


图 2 研究框图

Fig. 2 Research framework diagram

和 Yao^[7-8] 结合贝叶斯理论, 提出基于隶属度均值的决策理论阴影集 (Mean value based decision theoretic shadowed set, MDTSS), 并通过最小化划分动作所造成的决策代价来寻找最优 (α, β) . 此外, Yao 等^[9] 基于阴影集进一步提出针对模糊集的通用三支近似框架, 并总结了计算 (α, β) 的三个准则, 即不确定性不变性、最小距离和最小代价. Zhang 和 Yao^[10-11] 则提出博弈论阴影集 (Game theoretic shadowed set, GTSS), 其结合博弈论的折中原则分析了阴影集不同划分操作所造成的隶属度误差之间的竞争关系, 并通过重复博弈机制来求取最优 (α, β) , 为决策阈值的求取提供了新的语义解释, 并为阴影集的构建提供了一种新颖的视角. Zhang 等^[12] 从熵的角度出发, 提出区间阴影集 (Interval shadowed set, ISS), 其用区间值 $[\beta, \alpha]$ 代替 $[0, 1]$ 来表示阴影域, 并定义一种测量阴影域模糊熵的方法, 通过建立平衡三个区域不确定性的目标函数 (即模糊熵) 来求取构建 ISS 模型的决策阈值. Gao 等^[13] 则从熵的角度出发, 提出模糊熵均值阴影集 (Mean entropy based shadowed set, MESS), 其讨论分析了不同划分操作所产生的模糊熵损失, 并提出相应决策划分规则, 通过规则推导出不涉及人为设定参数的 (α, β) . 随后, 对 MESS 进行了迭代改进, 使其具有更小的熵损失. 受 GTSS 模型的启发, Zhang 等^[14] 和高满^[15] 还从熵的角度出发, 基于博弈论的折中原则分析了阴影集不同划分操作所造成的熵损失, 从博弈对象、策略、收益函数和停止条件四个方面展开讨论, 构建了基于模糊熵的博弈论阴影集 (Game

theoretic shadowed set based on fuzzy entropy, FE-GTSS). 近期, Luo 等^[16] 在 ISS 模型的基础上, 提出一种改进模型, 即 New-ISS. Gao 等^[17] 则在 GTSS 和 FE-GTSS 模型的基础上, 进一步提出基于不确定性和决策代价博弈分析的阴影集 (Shadowed set based on game analysis of uncertainty and decision cost, UC-GTSS).

显然, 这些模型设计了不同的三支近似表达形式或阈值计算方法, 以增强阴影集的语义解释或近似划分能力. 在表达形式方面, 可总结如图 3 所示. 图 3(a) 为 SS 模型, 图 3(b) 为 ISS 和 New-ISS 模型, 其阴影域通过区间值 $[\beta, \alpha]$ 表示. 图 3(c) 代表了 0.5SS、MDTSS、MESS、GTSS、FE-GTSS 和 UC-GTSS 模型, 区别在于阴影域由一个常数值 δ 表示, 而非区间值. 其中在 0.5SS、GTSS 和 UC-GTSS 模型中, $\delta = 0.5$; 在 MDTSS 模型中, δ 通过隶属度均值 $\bar{\delta}$ 表示; 在 MESS 模型中, δ 通过模糊熵均值 δ^* 表示; 而 FE-GTSS 则讨论了 δ 为 0.5 和 δ^* 两种情况下的模型构建. 图 3(d) 是 Yao 等^[9] 所提出的三支近似通用表达框架, 如式 (1) 所示, 三个近似域分别通过三个常数值 $\{p, m, n\}$ 表示, 其具体数值可根据待解决问题和应用场景来设定, 不局限于某一固定值.

$$S_{\mu_A}(x) = \begin{cases} p, & \mu_A(x) \geq \alpha \\ m, & \beta < \mu_A(x) < \alpha \\ n, & \mu_A(x) \leq \beta \end{cases} \quad (1)$$

总结不同阴影集模型的特性如表 1 所示. 下面详细介绍不同阴影集模型的核心思想和构建步骤.

1.1.1 经典阴影集

定义 1^[1] (模糊集). 设在非空有限集 U 上的映射 $\mu_A : U \rightarrow [0, 1], x \rightarrow \mu_A(x)$, 其中 $U = \{x_1, x_2, \dots, x_n\}$, A 表示 U 上的一个模糊集, $\mu_A(x)$ 表示 A 的隶属度函数, 其意味着集合中对象 x 属于模糊概念 A 的程度, 值越大则隶属程度越高, 反之则越低.

定义 2^[2] (阴影集). 设 U 为非空有限论域, α 和 β 为两个实数且满足不等式 $0 \leq \beta < 0.5 < \alpha \leq 1, \alpha + \beta = 1$, 则定义 U 到集合 $\{0, [0, 1], 1\}$ 的映射函数 $S : U \rightarrow \{0, [0, 1], 1\}$ 为阴影集, 如下所示

$$S_{\mu_A}(x) = \begin{cases} 1, & \mu_A(x) \geq \alpha \\ [0, 1], & \beta < \mu_A(x) < \alpha \\ 0, & \mu_A(x) \leq \beta \end{cases} \quad (2)$$

对决策划分阈值 (α, β) 的求取, Pedrycz^[2] 提出不确定性不变性准则, 即通过平衡提升域 $Ele_{(\alpha, \beta)}(\mu_A)$ 、减小域 $Red_{(\alpha, \beta)}(\mu_A)$ 和阴影域 $Shadow_{(\alpha, \beta)}(\mu_A)$ 之间的不确定性变化 (即模糊熵变化) 来求取和解释

最优 (α, β) , 如图 3(a) 和式 (3) 所示

$$Ele_{(\alpha, \beta)}(\mu_A) + Red_{(\alpha, \beta)}(\mu_A) = Shadow_{(\alpha, \beta)}(\mu_A) \quad (3)$$

随后, 将其转化为最优化问题: 对于给定的任意模糊集, 通过最小化模糊熵损失求取最优 (α, β) , 即式 (4). 其中 $V_{(\alpha, \beta)}(\mu_A)$ 为目标函数, 如式 (5) 所示

$$\arg \min_{(\alpha, \beta)} V_{(\alpha, \beta)}(\mu_A) \quad (4)$$

$$V_{(\alpha, \beta)}(\mu_A) = |Ele_{(\alpha, \beta)}(\mu_A) + Red_{(\alpha, \beta)}(\mu_A) - Shadow_{(\alpha, \beta)}(\mu_A)| \quad (5)$$

基于决策阈值对 (α, β) , 模糊集中的对象 x 将被近似划分到三个子集: 当 $\mu_A(x) \geq \alpha$ 时, x 被划分到 $Ele_{(\alpha, \beta)}(\mu_A)$; 当 $\mu_A(x) \leq \beta$ 时, x 被划分到 $Red_{(\alpha, \beta)}(\mu_A)$; 当 $\beta < \mu_A(x) < \alpha$ 时, x 被划分到 $Shadow_{(\alpha, \beta)}(\mu_A)$.

1.1.2 决策理论阴影集

Cattaneo 和 Ciucci^[5-6] 提出用隶属度常数值 0.5 来表示阴影域, 而不再是式 (2) 中的区间值 $[0, 1]$, 其意味着阴影域中的对象具有最大的不确定性. 随

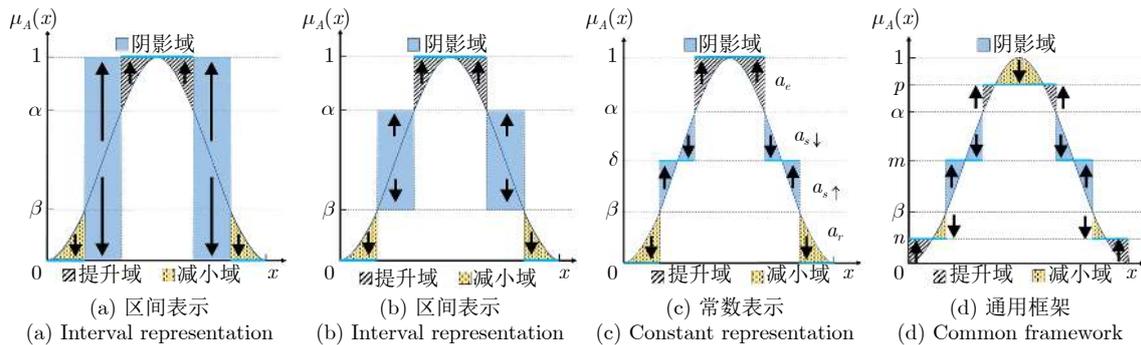


图 3 阴影集的几种三支近似表达形式

Fig. 3 Several three-way approximate expressions of shadowed set

表 1 不同阴影集之间的区别和联系

Table 1 The difference and relation between different shadowed sets

年份	文献	模型	阴影域	构建准则	分析视角	构建方法	人为参数设定	时间复杂度
1998	[2]	SS	$[0, 1]$	不确定性不变性	不确定性损失	最优化目标函数	—	$O(n)$
2003	[5-6]	0.5SS	0.5	不确定性不变性	不确定性损失	最优化目标函数	—	$O(n)$
2013, 2014, 2017	[7-9]	MDTSS	$\bar{\delta}$	最小成本	隶属度误差	决策划分规则推导	$\lambda_e, \lambda_r, \lambda_{s\downarrow}, \lambda_{s\uparrow}$	$O(n)$
2018, 2020	[10-11]	GTSS	0.5	最小成本	隶属度误差	博弈竞争机制	c_E, c_R	$O(n^2)$
2020	[12]	ISS	$[\beta, \alpha]$	不确定性不变性	不确定性损失	最优化目标函数	—	$O(n)$
2020	[13]	MESS	δ^*	不确定性不变性	不确定性损失	决策划分规则推导	—	$O(n)$
2020	[14-15]	FE-GTSS	$0.5, \bar{\delta}$	不确定性不变性	不确定性损失	博弈竞争机制	c_E, c_R	$O(\log_2 n)$
2023	[16]	New-ISS	$[\beta, \alpha]$	不确定性不变性	不确定性损失	最优化目标函数	—	$O(n \log_2 n)$
2023	[17]	UC-GTSS	0.5	最小成本, 不确定性不变性	隶属度误差, 不确定性损失	博弈竞争机制	c_E, c_R	$O(n)$

后, Deng 和 Yao^[7-8] 分析认为用固定值 0.5 并不能总是合理地代表所有对象的不确定性, 因此提出采用模糊集对象的隶属度平均值 $\bar{\delta}$ 来表示阴影域对象的隶属度水平, 其更能反映真实数据的分布情况, 如图 3(c) 和式 (6) 所示, 其中 $S(\mu_A) = \{x \in U | 0 < \mu_A(x) < 1\}$ 表示不确定性对象的集合, $card(\cdot)$ 表示集合的基.

$$\bar{\delta} = \frac{\sum_{x \in S(\mu_A)} \mu_A(x)}{card(S(\mu_A))} \quad (6)$$

随后, Deng 和 Yao^[7-8] 结合贝叶斯决策理论, 以最小化阴影集三个区域中对象的决策划分代价为目标, 即最小代价准则, 求取最优 (α, β) , 提出 MDTSS 模型以及基于阴影集的决策理论三支近似模型. 结合图 3(c) 可知, 对于 $\forall x \in U$, MDTSS 给出了四种划分 $Actions = \{a_e, a_r, a_{s\downarrow}, a_{s\uparrow}\}$, 每种划分 a 均会造成相应决策误差 E_a 和决策代价 λ_a , $a \in Actions$, 总结如表 2 所示. 其中 a_e 和 a_r 分别表示将 $\mu_A(x)$ 提升至 1 和减小至 0 的划分动作, $a_{s\downarrow}$ 和 $a_{s\uparrow}$ 分别表示将 $\mu_A(x)$ 减小至 δ 和提升至 δ 的划分动作, E_a 则表示对象的隶属度变化差值, λ_a 表示该划分动作的代价权重, 其值由专家经验设定.

表 2 MDTSS 中划分所造成的误差和代价

Table 2 Error and cost caused by partitions in MDTSS

划分操作	模糊集	阴影集	决策误差 E_a	决策代价 λ_a
a_e	$\mu_A(x)$	1	$1 - \mu_A(x)$	λ_e
a_r	$\mu_A(x)$	0	$\mu_A(x) - 0$	λ_r
$a_{s\downarrow}$	$\mu_A(x) \geq \delta$	δ	$\mu_A(x) - \delta$	$\lambda_{s\downarrow}$
$a_{s\uparrow}$	$\mu_A(x) < \delta$	δ	$\delta - \mu_A(x)$	$\lambda_{s\uparrow}$

MDTSS 同样将 (α, β) 的求取转换为最优化问题, 即式 (7). 其中 $R(\alpha, \beta)$ 为决策损失函数, 其表示论域 U 中所有对象的决策总损失, 如式 (8) 所示. $R(a|x)$ 表示对象 x 执行划分动作 a 时所造成的损失, 通过决策误差 E_a 和相应决策代价 λ_a 的乘积表示. 进一步通过构建决策划分规则以及规则的约简和推导求取 (α, β) , 如式 (9) 所示, 其中 $\lambda_e, \lambda_r, \lambda_{s\downarrow}, \lambda_{s\uparrow}$ 均大于等于 0 且满足 $\lambda_{s\downarrow} \leq \lambda_r, \lambda_{s\uparrow} \leq \lambda_e, \lambda_e + \lambda_{s\downarrow} \neq 0, \lambda_r + \lambda_{s\uparrow} \neq 0$.

$$\arg \min_{(\alpha, \beta)} R(\alpha, \beta) \quad (7)$$

$$R(\alpha, \beta) = \sum_{x \in U} R(a|x) = \sum_{x \in U} E_a \times \lambda_a = \sum_{x \in U} E_a(\mu_A(x)) \times \lambda_a \quad (8)$$

$$\alpha = \frac{\lambda_e + \delta\lambda_{s\downarrow}}{\lambda_e + \lambda_{s\downarrow}}, \quad \beta = \frac{\delta\lambda_{s\uparrow}}{\lambda_r + \lambda_{s\uparrow}} \quad (9)$$

1.1.3 博弈论阴影集

Zhang 和 Yao^[10-11] 将博弈论与阴影集相结合, 提出 GTSS 模型. 而博弈论的最终目的是尽可能为不同博弈玩家找到对个人而言最有益的行为策略, 并构建一种状态, 在该状态下所有博弈玩家之间的收益都保持均衡或尽可能满足折中的原则^[18-19]. 基于此, Zhang 和 Yao^[10-11] 分析了阴影集中不同划分操作所造成隶属度误差之间的博弈关系, 引入博弈论的折中思想来平衡不同类型隶属度误差, 并通过重复博弈机制来求取最优 (α, β) .

以连续型模糊隶属度函数为例, 如图 4 所示, 在阴影集近似划分过程中, 对象 x 通过 (α, β) 采取提升隶属度和减小隶属度这两种操作. 显然, 将形成两种误差: 提升误差 $TE_{(\alpha, \beta)}^{Mem}(\mu_A)$ (黄色区域) 和减小误差 $TR_{(\alpha, \beta)}^{Mem}(\mu_A)$ (蓝色区域), 分别如式 (10) 和式 (11) 所示, 其中 $R_0(\mu_A), R_\delta(\mu_A), E_1(\mu_A)$ 和 $E_\delta(\mu_A)$ 分别表示对象在不同区域隶属度变化所造成的误差.

$$TE_{(\alpha, \beta)}^{Mem}(\mu_A) = E_1(\mu_A) + E_\delta(\mu_A) \quad (10)$$

$$TR_{(\alpha, \beta)}^{Mem}(\mu_A) = R_0(\mu_A) + R_\delta(\mu_A) \quad (11)$$

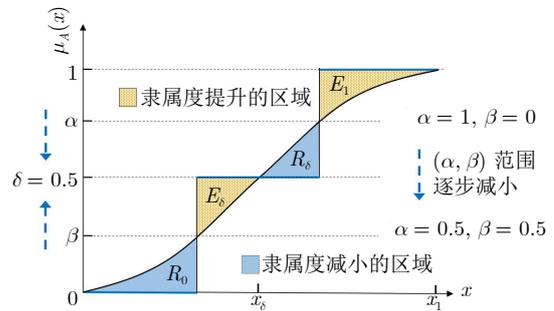


图 4 GTSS 中两种隶属度误差随决策阈值的博弈变化
Fig. 4 The game variation of two kinds of membership errors with decision threshold in GTSS

显然, 由图 4 中黄色区域和蓝色区域可知, 这两种隶属度误差随 (α, β) 的变化始终呈相反的变化趋势, 即一方增大, 另一方必然减小. Zhang 和 Yao^[10-11] 认为这与博弈论的核心思想是相似的: 所有博弈玩家均试图最大化个人收益, 即最小化个人损失, 然而所有玩家的个人收益却会受到其他玩家行为策略的干扰, 形成合作或非合作博弈的局面. 因此, 基于折中原则的考虑, 所有玩家将寻找一种新的策略, 该策略令所有玩家的博弈收益形成平衡状态, 这种平衡是所有玩家的理想状态. 基于该策略, 玩家的个人收益既不是最好的, 也不是最坏的,

但却能够最大化所有玩家的总收益^[18-19]. 综上, GTSS 模型将划分操作所形成的两种隶属度误差分别作为博弈对象, 将 (α, β) 的阈值变化作为博弈策略, 以两种隶属度的总误差最小 (即博弈收益最大) 作为博弈目标, 进行博弈均衡分析, 构建如表 3 所示的博弈收益表, 并通过限定博弈策略的数量来迭代更新博弈收益表, 通过重复博弈的形式来寻找最优的决策划分阈值对 (α, β) . 为易于理解, 相比文献 [10-11], 在表 3 中对符号表示进行了简化.

由表 3 可知, 在 GTSS 中, 博弈对象分别为 $n_1 = TR_{(\alpha, \beta)}^{Mem}(\mu_A)$ 和 $n_2 = TE_{(\alpha, \beta)}^{Mem}(\mu_A)$, 博弈策略集分别为 $S_{n_1} = \{\alpha_1, \alpha_2, \dots, \alpha_p\}$ 和 $S_{n_2} = \{\beta_1, \beta_2, \dots, \beta_q\}$, 其中 $p, q \in \mathbb{N}^+$, $P_{n_i}(\alpha_p, \beta_q)$ 表示 n_i 执行 (α_p, β_q) 时的收益, 如式 (12) 及式 (13) 所示, 其通过一个常量值 C 与隶属度损失的差值来表示. 其中, 常量值 C 及收益函数可根据不同的研究背景和待解决问题来设定^[11]. 在 GTSS 中, 博弈的停止条件 (即确定最优 (α, β) 的条件) 为: 使得 $P_{n_1}(\alpha_p, \beta_q)$ 和 $P_{n_2}(\alpha_p, \beta_q)$ 之和最大的 (α_p, β_q) .

$$P_{n_1}(\alpha_p, \beta_q) = C - TR_{(\alpha_p, \beta_q)}^{Mem}(\mu_A) \quad (12)$$

$$P_{n_2}(\alpha_p, \beta_q) = C - TE_{(\alpha_p, \beta_q)}^{Mem}(\mu_A) \quad (13)$$

1.1.4 区间阴影集

Zhang 等^[12] 首先提出区间模糊熵的计算方法, 随后提出 ISS 模型, 其使用 $[\beta, \alpha]$ 来代表阴影域对象, 替代传统模型的 $[0, 1]$, 如定义 3 及图 3(b) 所示, 并以最小化模糊熵损失为目标构建求取最优 (α, β) 的目标函数.

定义 3^[12] (区间阴影集). 定义 U 为非空有限论域, α 和 β 为两个实数, 并且 $0 \leq \beta < \alpha \leq 1, \alpha + \beta = 1$, 则定义 U 到 $\{0, [\beta, \alpha], 1\}$ 的映射函数, 即 $S^* : U \rightarrow \{0, [\beta, \alpha], 1\}$ 为区间阴影集, 如下所示

$$S_{\mu_A}^*(x) = \begin{cases} 1, & \mu_A(x) \geq \alpha \\ 0, & \mu_A(x) \leq \beta \\ [\beta, \alpha], & \beta < \mu_A(x) < \alpha \end{cases} \quad (14)$$

相比定义 2, Zhang 等^[12] 分析认为区间 $[\beta, \alpha]$ 相比 $[0, 1]$ 能够使得所构建阴影集与模糊集之间具有更小的不确定性损失, 即模糊熵损失, 而更小的模糊熵损失意味着阴影集更有效. 对此, 文献 [2-4] 曾指出: 阴影集本质上是对模糊集的一种符号表示, 即模糊信息的粒化处理, 而基于粒计算思想来看, 可理解为同样的知识在不同粒度层次下的体现, 前者是粗粒度, 后者是细粒度. 综上, 当阴影集与模糊集之间的熵损失越小, 则阴影集对模糊集的近似描述就越好, 即两者间所蕴含知识量和信息量的一致性就保存的越好.

在 ISS 模型中, 模糊集与阴影集之间的不确定性损失通过模糊熵损失来衡量, 后者是 Liu^[20] 所提出的一种描述模糊集合不确定性程度的方法. Zhang 等^[12] 分析发现传统模糊熵度量方法仅针对模糊集有效, 无法直接用于阴影集. 因此, 提出阴影集的模糊熵度量方法, 即针对阴影区间的模糊熵度量方法, 如式 (15) 所示, $\bar{e}_{[a, b]}(x)$ 表示区间 $[a, b]$ 内对象 x 的区间模糊熵, 其中 $[a, b] \subseteq [0, 1], \mu_A(x)$ 等概率地取到 $[a, b]$ 中的任何值.

$$\bar{e}_{[a, b]}(x) = \frac{1}{b-a} \int_a^b \mu_A(x)(1 - \mu_A(x))d\mu_A(x) \quad (15)$$

ISS 模型同样通过平衡阴影集提升域模糊熵 $e_{(\alpha, \beta)}^{Ele}(\mu_A)$ 、减小域模糊熵 $e_{(\alpha, \beta)}^{Red}(\mu_A)$ 以及阴影域模糊熵 $e_{(\alpha, \beta)}^{Shadow}(\mu_A)$ 之间的变化来寻找最优 (α, β) , 如图 3(b) 及式 (16) 所示, 并将其转化为最优化问题, 通过最小化模糊熵损失来求取最优 (α, β) , 如式 (17) 所示. 其中, $E_{(\alpha, \beta)}(\mu_A)$ 为目标函数, 如式 (18) 所示.

$$e_{(\alpha, \beta)}^{Ele}(\mu_A) + e_{(\alpha, \beta)}^{Red}(\mu_A) = e_{(\alpha, \beta)}^{Shadow}(\mu_A) \quad (16)$$

$$\arg \min_{(\alpha, \beta)} E_{(\alpha, \beta)}(\mu_A) \quad (17)$$

$$E_{(\alpha, \beta)}(\mu_A) = \left| e_{(\alpha, \beta)}^{Ele}(\mu_A) + e_{(\alpha, \beta)}^{Red}(\mu_A) - e_{(\alpha, \beta)}^{Shadow}(\mu_A) \right| \quad (18)$$

表 3 博弈机制下阴影集的博弈收益
Table 3 Game payoff of shadowed set under game mechanism

博弈对象	n_2				
	博弈策略	β_1	β_2	\dots	β_q
n_1	α_1	$\langle P_{n_1}(\alpha_1, \beta_1), P_{n_2}(\alpha_1, \beta_1) \rangle$	$\langle P_{n_1}(\alpha_1, \beta_2), P_{n_2}(\alpha_1, \beta_2) \rangle$	\dots	$\langle P_{n_1}(\alpha_1, \beta_q), P_{n_2}(\alpha_1, \beta_q) \rangle$
	α_2	$\langle P_{n_1}(\alpha_2, \beta_1), P_{n_2}(\alpha_2, \beta_1) \rangle$	$\langle P_{n_1}(\alpha_2, \beta_2), P_{n_2}(\alpha_2, \beta_2) \rangle$	\dots	$\langle P_{n_1}(\alpha_2, \beta_q), P_{n_2}(\alpha_2, \beta_q) \rangle$
	\vdots	\vdots	\vdots	\vdots	\vdots
	α_p	$\langle P_{n_1}(\alpha_p, \beta_1), P_{n_2}(\alpha_p, \beta_1) \rangle$	$\langle P_{n_1}(\alpha_p, \beta_2), P_{n_2}(\alpha_p, \beta_2) \rangle$	\dots	$\langle P_{n_1}(\alpha_p, \beta_q), P_{n_2}(\alpha_p, \beta_q) \rangle$

综上, 可求得 ISS 模型的 (α, β) , 进而将模糊集中的对象近似划分到提升域、减小域和阴影域.

1.1.5 基于模糊熵均值的决策理论阴影集

针对 0.5SS 和 MDTSS 模型需要人为设定代价参数所带来的主观性问题以及模型模糊熵损失较大等问题, Gao 等^[13] 提出 MESS 模型, 其从模糊熵的角度探讨了阴影域的表示, 分析了决策划分过程中不同划分行为所造成的熵损失, 在新提出的决策理论划分规则基础上, 推导出 (α, β) , 其无需人为设定代价参数, 避免了主观性, 并通过迭代的形式优化了 (α, β) . 基于定义 2, MESS 模型给出了新的三支近似映射形式, 即 $S^* : U \rightarrow \{0, \delta^*, 1\}$, 其利用全体模糊对象的模糊熵均值 $\bar{e}(A)$ 来表示阴影区域对象的不确定性水平, 即式 (19) (处理连续型数据) 及式 (20) (处理离散型数据), 进而求得阴影域对象的隶属度代表值, 即式 (21), 其中 $0 < \delta_1^* \leq 0.5, 0.5 \leq \delta_2^* < 1$ 且 $\delta_1^* + \delta_2^* = 1$.

$$\bar{e}_{lx}(A) = \frac{\int_{x \in S(\mu_A)} \mu_A(x)(1 - \mu_A(x))dx}{card(S(\mu_A))} \quad (19)$$

$$\bar{e}_{ls}(A) = \frac{\sum_{x \in S(\mu_A)} \mu_A(x)(1 - \mu_A(x))}{card(S(\mu_A))} \quad (20)$$

$$\delta_1^* = \frac{1 - \sqrt{1 - 4\bar{e}(A)}}{2}, \quad \delta_2^* = \frac{1 + \sqrt{1 - 4\bar{e}(A)}}{2} \quad (21)$$

对于 (α, β) 的求取, 受 MDTSS 的启发, MESS 结合贝叶斯决策理论, 并以最小化阴影集三个区域中对象的决策划分代价为目标构建目标函数, 与 MDTSS 所不同的是, MESS 模型的决策代价为模糊熵损失, 而非隶属度损失. 给出阴影集的四中划分操作 $Actions = \{a_e, a_r, a_{s\downarrow}, a_{s\uparrow}\}$ 及其相应的模糊熵损失 $El(a|x), a \in Actions$, 如表 4 所示.

表 4 MESS 中划分所造成的熵损失

Table 4 Entropy loss caused by partitions in MESS

划分动作	模糊集	阴影集	熵损失 $El(a x)$
a_e	$\mu_A(x)$	1	$El(a_e x)$
a_r	$\mu_A(x)$	0	$El(a_r x)$
$a_{s\downarrow}$	$\mu_A(x) \geq \delta^*$	δ^*	$El(a_{s\downarrow} x)$
$a_{s\uparrow}$	$\mu_A(x) < \delta^*$	δ^*	$El(a_{s\uparrow} x)$

由表 4 可知, 在 MESS 中, $\forall x \in U$ 在执行划分操作时所造成的模糊熵损失函数 $El(a|x)$, 如式 (22) 所示, E_b 和 E_e 分别代表执行划分操作前和操作后的模糊熵. 进而构建 U 中全体对象的损失函数 $El(\alpha, \beta)$,

如式 (23) 所示, 将其视为最优化问题, 并进一步构建决策划分规则, 通过规则的约简和推导来求取 (α, β) , 分别如式 (24) 和式 (25) 所示, 其中 $\alpha + \beta = 1$.

$$El(a|x) = |E_e - E_b| \quad (22)$$

$$El(\alpha, \beta) = \sum_{x \in U} El(a|x) = \sum_{x \in U} |E_e - E_b| \quad (23)$$

$$\arg \min_{(\alpha, \beta)} El(\alpha, \beta) \quad (24)$$

$$\alpha = \frac{1 + \sqrt{1 - 2\delta_1^*\delta_2^*}}{2}, \quad \beta = \frac{1 - \sqrt{1 - 2\delta_1^*\delta_2^*}}{2} \quad (25)$$

此外, Gao 等^[13] 分析基于上述步骤所求取的 (α, β) 虽然减小了模糊熵损失, 但不能使 MESS 具有最小的模糊熵损失, 其主要原因是初次求取 $\bar{e}(A)$ 时, $S(\mu_A) = A$, 这其中包含了部分隶属度非常接近 1 或 0 的对象, 即确定性对象, 这使得 $\bar{e}(A)$ 在初次计算时不能充分代表 A 中对象不确定性水平, 从而影响 (α, β) 的计算. 因此, MESS 采用了模型迭代的方法^[13], 利用初次求取的 (α, β) 来约束不确定性对象的集合, 筛选出不确定性较小的对象, 保留不确定性较大的对象, 用作下一次 $\bar{e}(A)$ 的求取, 从而实现最优 (α, β) 的迭代搜索, 减小模型的模糊熵损失.

1.1.6 基于模糊熵的博弈论阴影集

传统 GTSS 通过对隶属度误差之间的博弈分析寻找最优 (α, β) , 然而其缺乏对不确定性变化的考虑, 这使得所构建 GTSS 虽然具有较小的隶属度误差, 但却可能存在较大的不确定性损失. 此外, 在 GTSS 中, 当 (α, β) 的精度要求越高, 则博弈策略的搜索时间越长, 即最优 (α, β) 的搜索效率较低. 因此, Zhang 等^[14] 提出 FE-GTSS, 其遵循文献 [2-4] 提出阴影集时的初衷, 即最小化不确定性损失, 直接基于不确定性的视角, 结合博弈论的折中原理^[18-19], 对阴影集不同划分操作所造成的不确定性损失 (即模糊熵损失) 进行博弈分析. 并从博弈的对象、策略、收益函数和停止条件四个方面讨论了模型的构建步骤. 此外, FE-GTSS 还采用二分法替换传统 GTSS 中的顺序查找法, 从而提高最优 (α, β) 的查找效率. 此外, Zhang 等^[14] 还基于对数据特性的考虑, 对模型进行了扩展分析.

以连续型模糊隶属度函数为例, 如图 5 所示, 在阴影集近似划分过程中, $\forall x$ 通过 (α, β) 采取提升或减小隶属度这两种操作时, 将造成两种类型的模糊熵损失: 模糊熵提升的损失 $TE_{(\alpha, \beta)}^{Ent}(\mu_A)$ (黄色区域) 和模糊熵减小的损失 $TR_{(\alpha, \beta)}^{Ent}(\mu_A)$ (蓝色区域), 分别如式 (26) 和式 (27) 所示, 其中 $R_{e1}(\mu_A), R_{e0}(\mu_A)$,

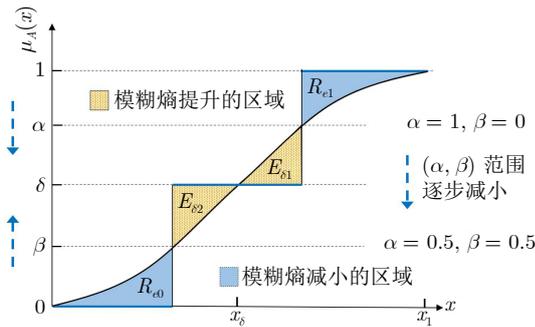


图 5 FE-GTSS 中两种模糊熵损失随决策阈值的博弈变化

Fig. 5 The game variation of two fuzzy entropy losses with decision threshold in FE-GTSS

$E_{\delta 1}(\mu_A)$ 和 $E_{\delta 2}(\mu_A)$ 分别表示对象在不同区域模糊熵变化所造成的损失。

$$TE_{(\alpha, \beta)}^{Ent}(\mu_A) = E_{\delta 1}(\mu_A) + E_{\delta 2}(\mu_A) \quad (26)$$

$$TR_{(\alpha, \beta)}^{Ent}(\mu_A) = R_{e1}(\mu_A) + R_{e0}(\mu_A) \quad (27)$$

显然, 无论 (α, β) 如何变化, 两种模糊熵损失始终呈相反的变化趋势, 即形成非合作型博弈。因此, FE-GTSS 模型将两种模糊熵损失作为相互博弈的对象, 即, $n_1 = TR_{(\alpha, \beta)}^{Ent}(\mu_A)$, $n_2 = TE_{(\alpha, \beta)}^{Ent}(\mu_A)$; 将 (α, β) 的阈值变化作为博弈策略, 即 $S_{n_1} = \{\alpha_1, \alpha_2, \dots, \alpha_p\}$ 和 $S_{n_2} = \{\beta_1, \beta_2, \dots, \beta_q\}$; 将模糊熵的总损失最小 (即博弈收益最大) 作为博弈目标, 博弈收益函数如式 (28) 及式 (29) 所示。并构建如表 3 所示的博弈收益表, 通过博弈均衡分析来搜索最优 (α, β) 。在 FE-GTSS 中, 博弈收益函数的设定以及常数值 C 的设定可根据实际应用场景和待解决问题来设计。博弈的停止条件为: 使得 $P_{n_1}(\alpha_p, \beta_q)$ 和 $P_{n_2}(\alpha_p, \beta_q)$ 之和最大的 (α_p, β_q) 。此时, FE-GTSS 获得最大博弈收益, 即最小的模糊熵损失, 这更好地遵循了不确定性不变性准则。

$$P_{n_1}(\alpha_p, \beta_q) = C - TR_{(\alpha_p, \beta_q)}^{Ent}(\mu_A) \quad (28)$$

$$P_{n_2}(\alpha_p, \beta_q) = C - TE_{(\alpha_p, \beta_q)}^{Ent}(\mu_A) \quad (29)$$

1.1.7 其他

近期, Luo 等^[16] 基于不确定性不变性准则, 在 ISS 模型的基础上提出一种改进模型, 即 New-ISS, 设计一种新的不确定性平衡目标函数以及阈值对快速求取算法, 其目的是使得 New-ISS 模型具有更小的模糊熵损失和时间损耗。Gao 等^[17] 在 GTSS 和 FE-GTSS 模型的基础上, 继续将阴影集与博弈论相结合, 探究了基于多个构建准则约束条件下的阴影集构建方法, 提出基于不确定性和决策代价博弈

分析的阴影集, 即 UC-GTSS 模型。该工作着重讨论了如何通过博弈机制同时满足不确定性不变性准则和最小成本准则, 尝试将阴影集构建从传统的单一准则约束推广到多准则约束。此外, 不同于当前阴影集构建准则都侧重于模型构建过程的语义解释和阈值确定的合理性, 该工作还从划分结果有效性的角度对 UC-GTSS 进行了扩展和讨论, 如近似分类精度与覆盖率之间的博弈, 划分区域的基尼系数之间的博弈等。这些工作丰富了决策阈值的语义解释和合理性, 也增强了阴影集的应用价值, 使其能够满足更多的应用场景。因篇幅限制, 不再详细介绍, 可参考文献 [16–17]。

1.2 阴影集的理论性质研究

鉴于阴影集对不确定性信息近似划分处理所体现出的有效性, 诸多学者围绕其理论性质展开了研究, 可总结如下两方面。

在阴影集的模糊关系、结构解释、逻辑运算或数学性质等方面: 2003 年, Cattaneo 和 Ciucci^[5–6] 研究了阴影集与模糊集之间的关系, 并从代数的角度分析阴影集, 其定义一种 BZMV 代数结构, 使得给定一个模糊集可以返回特定的阴影集, 并对该代数结构进行了理论分析和证明, 并首次提出用精确值 0.5 来近似表达阴影域。文献 [3–4] 以平衡数据的定性布尔描述和定量隶属度描述为目标, 结合粒计算的思想对阴影集进行阐述, 并详细介绍了其概念和被提出的动机, 分析了其基本逻辑运算和应用性, 并在模糊聚类场景中讨论分析了阴影集三个表达值的解释能力。2013 年, 区别于传统区间模糊数近似和梯形模糊数近似算法, Grzegorzewski^[21] 提出结合阴影集的模糊数近似算法, 并讨论其基本性质以及对细粒度计算的益处, 新算法能有效消除描述不精确现象时的过度精度问题, 并分析了新算法相比经典阴影集近似方法的优势。2017 年, Yao 等^[9] 提出一个研究模糊集三支近似的通用框架, 用三值 $\{n, m, p\}$ 的形式替代阴影集 $\{0, [0, 1], 1\}$ 的表达形式, 即式 (1), 并通过回顾现有研究总结了求取和解释阴影集 (α, β) 的三个基本准则: 不确定性不变性, 最小距离, 最小成本, 其中最小距离准则是最小成本准则的一种特例^[22]。文献 [23–24] 在文献 [8–9] 所提出的决策理论阴影集基础上, 采用一个变量值 C 替换常量值 0.5 来表示阴影域, 并通过分析阴影集中的提升和降低操作, 引出四个划分代价参数, 进而建立新的目标函数。详细讨论了 C 值的四种特殊情况及其与 (α, β) 的关系和性质。并通过引入粒子群算法优化目标函数, 以最小化总代价为目标搜索

最优 C 值和相应的 (α, β) . 2019 年, Zhou 等^[22, 25] 详细分析了 Pedrycz^[2] 和 Yao 等^[9] 所构建的三支近似阴影集模型的连续性和凸性, 以及最优解的存在性和唯一性, 并发现即使在相同的优化模型下也会产生不同的近似划分区域, 这需要新的标准来构造阴影集, 因此提出被动和主动约束阴影集的概念, 分析了其数学性质, 并设计一种快速求取最优阈值的算法. 2019 年, Zhao 和 Yao^[26] 提出一种三支模糊划分的概念框架, 将其将阴影集的思想应用于非空模糊集簇, 从而构建该簇的三支近似, 即阴影集簇. 引入了 (α, β) 的三对性质来定义、描述和分类三支模糊划分, 即 α 和 β 的不相交性质、重叠性质和覆盖性质. 此外, 利用两组性质给出了三支模糊划分的一般定义, 并通过考虑所有可能的非等价性质子集得到了 21 种三支模糊划分, 深入研究了其中 3 种, 并分析了它们与划分、模糊划分和区间集簇之间的联系. 2020 年, Ciucci 和 Yao^[27] 作为国际期刊 *Information Sciences* 的客座编辑, 发表特刊介绍了粒计算理论、阴影集理论以及三支决策理论三者之间的关系和协同作用, 并总结了部分代表性研究工作. 2022 年, Boffa 等^[28] 探讨了模糊集上聚合算子与阴影集上相应聚合算子之间的关系, 并重点研究了模糊集近似为阴影集的同态条件. 利用同态条件从所有潜在运算的集合中找到了阴影集上有运算意义的类, 并提出了满足同态条件的一类新聚合算子. 综上所述, 围绕阴影集理论性质的研究可总结如表 5 所示.

表 5 阴影集理论性质的研究总结

Table 5 Research summary of theoretical properties of shadowed set

研究内容	文献
阴影集的代数结构	[5-6]
阴影集与粒计算的关系及被提出的动机	[3-4]
基于阴影集的模糊数近似方法及性质和益处	[21]
阴影集的三支近似通用表达框架及构建准则总结	[9]
阴影集阴影区域表达参数与决策阈值的关系和性质	[23-24]
阴影集的连续性和凸性及最优解的存在性和唯一性	[22, 25]
基于阴影集的模糊划分框架及划分阈值的性质	[26]
阴影集、粒计算及三支决策理论三者间的协同作用	[27]
阴影集的聚合算子及其与模糊集之间的同态条件	[28]

基于不同模糊集下的阴影集构建分析方面: 2013 年, Tahayori 等^[29] 提出一种基于渐进数的模糊集模糊性表示方法, 并证明了模糊测度的不同选择可能导致不同阴影集. 随后, 提出一种针对任意模糊集的阴影集构建公式, 其通过去模糊化处理来

计算从给定模糊集导出阴影集所需的最优阈值. 区别于 Pedrycz^[2] 从三角模糊集和高斯模糊集等连续型论域求出阴影集的最优阈值, Tahayori 等^[29] 提出的方法能够处理离散型论域以及任意隶属函数所描述的模糊集, 并且其同时遵循不确定性不变性原则. 2017 年, Cai 等^[30] 认为在实际应用中具有时变隶属度的动态模糊集更为常见, 因此其探讨了基于动态模糊集的阴影集构建方法, 包括经典阴影集中三划分区域不确定性平衡的方法以及基于贝叶斯理论下最小化决策代价的方法. 2017 年, Bose 和 Mali^[31] 采用模糊集基数的渐进形式来表示模糊集的模糊性, 并证明了两者的关系, 随后通过两个不同且非对称的阈值 α_1 和 α_2 来增强阴影域划分的灵活性, 即 $\alpha_1 < 0.5$, $\alpha_2 \geq 0.5$ 且 $\alpha_1 \neq 1 - \alpha_2$, 最后以最小化模糊性差异为目标, 求取构建阴影集的最优阈值. 2019 年, 文献 [32-33] 以最小化不确定性差异为目标, 提出两种阴影集构建算法: 在文献 [32] 中描述了一种基于不确定性平衡的平均值方法, 阐述了如何从连续模糊隶属函数中求取最优划分阈值; 在文献 [33] 中则分析了在离散型模糊隶属函数下的最优划分阈值求取, 并且给出了量化给定集合不确定性的方法, 同时通过穷举搜索的形式寻找对模糊集近似的最佳对称阈值. 最近, William-West 和 Ibrahim^[34] 还提出一种针对不确定性和确定性平衡考虑的构建准则, 将阴影集的划分域从 3 个推广到 5 个, 讨论了最优阈值的求取, 并分析了其存在性和唯一性. 2020 年, Campagner 等^[35] 基于信息论探讨了将直觉模糊集近似为阴影集的方法, 其目标依旧是使最小化两者间的不确定性差异, 从而保持原始直觉模糊集中通过熵度量所表示的不确定性, 并提出三种构建阴影集的高效优化算法, 分别对应考虑了模糊性和知识的缺失性以及两者的整体考虑. 2021 年, Yang 和 Yao^[36] 研究了从 Atanassov 直觉模糊集构造阴影集的两种可能方案, 方案 1 将隶属度和非隶属度函数视为一对评估函数来构建, 方案 2 将隶属度和非隶属度函数组合成一个评估函数来构建. 前者应用于基于二评估的三支决策模型, 后者应用于基于单一评估的三支决策模型, 此外还引入了切比雪夫距离来解决评估过程中的阈值求取问题. 随后, 在 2022 年, Yao 和 Yang^[37] 基于商空间理论提出粒粗糙集、概率粒粗糙集以及粒阴影集三种三支近似模型, 并详细分析了模型的构造过程和理论性质, 其核心变化是将论域中的划分对象由单个对象改变为具有不可分辨关系的单个等价类. 特别地, 通过对距离函数和划分代价的考虑, 给出了构建粒阴影集时三个区域的决策划分规则. 此外, 还分析证明了使得三个模型近似结果一致的必要条件. 综

上所述, 基于不同模糊集的阴影集构建研究可总结如表 6 所示.

表 6 基于不同模糊集的阴影集构建研究总结
Table 6 Research summary of shadowed set construction based on different fuzzy sets

研究内容	文献
任意隶属度函数下模糊集的阴影集构建	[29]
基于时变隶属度的动态模糊集的阴影集构建	[30]
基于渐进形式表达模糊集模糊性时的阴影集构建	[31]
离散型模糊隶属函数下模糊集的阴影集构建	[32-33]
基于直觉模糊集的阴影集构建	[35]
基于 Atanassov 直觉模糊集的阴影集构建	[36]
基于模糊对象粒化后模糊集的阴影集构建	[37]

1.3 阴影集的数据分析研究

阴影集的三支近似划分操作, 将部分原本具有一定程度不确定性的对象转变成了确定性的对象, 即处在提升域和减小域中分别赋值为 1 和 0 的模糊对象. 从粒计算的视角来看, 这种将海量对象个体粒化处理到三个区域的做法, 在一定程度上减小了数据体量, 提高了数据质量, 也使得阴影集在其他计算和知识推理过程中能够提供更高的效率和更好的可解释性. 因此, 自阴影集被提出后, 诸多学者就将其作为一种有效的数据处理工具, 广泛结合到了不同的数据分析方法中.

基于阴影集的聚类算法方面: 2010 年, Mitra 等^[38] 基于阴影集提出一种结合模糊聚类和粗糙聚类的阴影 C 均值 (Shadowed C-means, SCM) 聚类算法. 该算法能够实现划分阈值参数的自动求取, 减少人为因素的影响, 并且能够利用阴影集三划分的思想约减无效数据, 从而减小计算量, 能够有效处理类簇之间的重叠区域以及类簇边界中的不确定性对象. 2011 年, Zhou 等^[39] 在基于粗糙集的聚类方法中, 分析了近似区域阈值的求取对类簇所形成信息粒描述的有效性, 认为近似阈值的求取至关重要, 因为其反映了数据的结构特性. 因此在基于粗糙集的聚类研究中引入了阴影集, 目的是实现近似阈值的自适应获取, 解决粗糙 C 均值 (Rough C-means, RCM) 聚类算法和粗糙模糊 C 均值 (Rough fuzzy C-means, RFCM) 聚类算法中所存在的对象间缺少全局关系分析的问题, 并证明其有效性. 此外, 在 2018 年, Zhou 等^[40] 还通过集成粗糙集、模糊集、阴影集和多粒度概念等粒计算技术, 提出了基于多粒度近似区域和阴影集的粗糙可能性 C 均值 (Rough possibilistic C-means, MS-RPCM) 聚类算法, 该算

法可充分解决数据中涉及的以及由模型参数产生的不确定性问题, 并且该算法对于处理噪声环境具有较好的鲁棒性. 2011 年, 郭晋华等^[41] 针对软聚类的研究工作, 在 RCM 和 RFCM 的基础上提出基于阴影集的粗糙模糊 C 均值 (Shadowed rough fuzzy C-means, S-RFCM) 聚类算法, 解决了 RCM 和 RFCM 需要人为设定阈值参数来划分不同决策区域所导致的强主观性问题. 而 S-RFCM 可以利用阴影集实现对数据本身特性的考虑, 从而自动化求取决策阈值, 自适应给出决策区域. 相类似地, 2013 年, 汪海良等^[42] 基于经典阴影集, 提出新的粗糙模糊可能性 C 均值 (Shadowed rough fuzzy possibility C-means, S-RFPCM) 聚类算法, 利用阴影集自动化求取传统 RFPCM 算法的决策划分阈值. 随后, Kong 和 Chen 等^[43] 分析 S-RFCM 算法在处理非球形聚类数据时具有较差效果, 因此提出一种新的聚类方案, 其在 S-RFCM 算法基础上融入核方法以及傅立叶随机特征映射, 从而处理非球形数据以及数据中的异常值问题.

2012 年, Falcon 等^[44] 提出一种处理增量数据的阴影集集群聚类算法, 分析了其相比基于 C 均值聚类算法的优势, 并将其应用于国家领土安全环境中风险源的可视化识别场景. 该算法能够对数据进行实时处理, 能很好地处理噪声和异常样本, 并且无需预先获得聚类数量. 2013 年, 李秀馨等^[45] 基于卫星云图聚类分析的应用场景, 将阴影集与模糊 C 均值聚类算法、离子群优化算法相结合, 提出一种改进的聚类算法, 而阴影集的引入则有效去除了类簇中的异常值, 提升了算法有效性. 2015 年, 蒙祖强等^[46] 在 RFCM 算法基础上, 结合混合蛙跳算法及阴影集来处理模糊类簇, 实现类簇划分阈值的自适应求取, 解决了 RFCM 算法的初值敏感、局部最优以及 (α, β) 设定存在主观性的缺陷. 2016 年, Li 等^[47] 基于阴影集提出针对混合属性的聚类边界检测算法, 其利用阴影集对不同类簇进行划分, 根据贡献度更新类簇质心信息, 并不断迭代优化, 最终整合并输出所有类簇基于阴影集划分的边界集. 该算法可有效、准确地检测聚类边界, 还可提取指定聚类的边界点或数据集的整个边界, 并对噪声具有抑制作用. 文献 [48-49] 基于集成学习和阴影集理论, 从三个角度考虑提出三种新颖的模糊聚类算法, 即, 结合阴影集的特征加权考虑, 结合阴影集和粗糙集的特征加权考虑, 以及基于两者的概率不相似函数混合数据考虑. 并在不同的应用场景和现实问题中分析了上述算法的有效性, 如民航机场噪声时间序列的预测、网络入侵的检测以及离群点噪声的检测等.

文献 [50–51] 在密度峰值聚类 (Density peaks clustering, DPC) 算法的基础上引入 ISS 模型, 从而自适应地获取用于设定类簇范围的截断距离, 而非传统的人为主观设定的方法, 同时利用 ISS 对类簇结果进行三支划分, 替换传统二支划分的聚类结果. 该方法能有效减少类簇的错误分类率, 并提升对噪声数据的抗噪性. 2019 年, 雒僖等^[52] 利用阴影集理论对截集式可能性 C 均值聚类算法进行改进, 新算法利用阴影集自动求取划分阈值的优势来为每个类簇确定聚类核, 并实现截集门限的自动化选取. 2020 年, Zhou 等^[53] 结合阴影集提出广义多粒度 RFCM 算法, 很好地实现了数据拓扑结构的捕捉, 并解决了数据中包含的不确定信息, 包括模糊化系数产生的不确定性、边界域和重叠域中产生的不确定性. 此外, Zhou 等^[54] 还分析了构造模糊集三支近似的几种方法, 并概括了三类准则: 基于不确定性平衡的准则, 基于原型的准则, 基于模型分类误差与数据体量之间权衡的准则, 提出一种基于模糊集三支近似的通用无监督学习框架, 在此框架下使用不同数据集对不同构建准则的有效性进行了比较评估, 并给出推荐. 2022 年, William-West 等^[55] 在无监督学习框架下, 对 SCM 聚类算法进行了理论分析, 并定义了三种构建优化准则, 即确定域与不确定域之间的平衡准则、清晰度平衡准则和渐进性平衡准则, 并评估了每种准则的优势.

基于阴影集的分类算法方面: 2020 年, Yue 等^[56] 设计一种新的阴影集来构造不确定性数据分类的阴影邻域. 新模型利用阶跃函数将邻域隶属度映射到三值集合 $\{0, 1, 0.5\}$, 从而将邻域划分为特定的正、负和不确定边界域, 并通过阴影集的最优化原理来寻找阶跃函数中的阈值参数, 设计了相应的三支分类规则, 提出一种基于阴影邻域的三支分类算法. 随后, 文献 [57–58] 进一步提出一种基于半监督阴影集的三支分类方法, 其首先提出一种新的半监督阴影集来构造阴影邻域, 从而实现对含有部分标记的不确定性数据的三支分类, 随后设计了新的阴影集目标函数和分类规则来实现三支分类, 并分析了标记数据对阴影集构造的影响以及新方法的有效性. 文献 [59–62] 基于阴影集理论提出一种多粒度三支聚类集成算法和一种快速 KNN 分类算法. 前者利用阴影集对基聚类进行处理, 从而形成三支聚类结果, 而对于基聚类的构造, 则分别采用模糊 C 均值算法和可能性 C 均值算法进行了探讨; 后者分类算法则利用三支聚类将训练样本进行划分, 从而实现数据样本约减, 优化数据样本质量, 提高分类效率, 而三支聚类的形成则基于阴影集来实现.

基于阴影集的数据预处理方面: 文献 [63–65] 提出一种训练数据筛选方法, 其利用阴影集将训练数据划分成三部分, 将正域和阴影域中的样本数据分别称之为核数据和边界数据, 并验证分析了这两个区域的数据作为训练样本能足够保留典型样本, 减少训练样本数据量并增强数据质量, 有效提高神经网络分类器的性能. 此外, 结合阴影集对训练样本数据筛选的优势, 苏小红等^[66] 也利用阴影集来进行训练样本的筛选, 并提出一种改进的模糊 SVM 数据样本选择算法. 2012 年, 周凡程等^[67] 结合阴影集提出一种提取候选断点集的新算法, 其根据候选断点在单属性上的分布来对数据样本进行分类, 并基于阴影集对决策分类结果进行三支划分, 从而选取候选断点集. 该算法能有效减少候选断点数目, 选出重要断点, 解决了连续属性离散化过程中冗余断点集过多所带来的噪声影响. 2012 年, 王丹等^[68] 提出一种基于阴影集聚类的离群点检测算法. 首先从数据的全局特性和局部特性两个角度出发, 提出一种离群因子的定义, 离群因子值越大表示对象属于离群点的可能性越大, 该定义考虑了离群点的不确定性和欧氏距离. 随后, 融合阴影集及 FCM 算法, 通过对离群因子值的数据处理来筛选离群点, 该算法能够同时实现聚类和离群点的检测.

综上所述, 可知阴影集在数据分析领域得到了广泛应用, 而不同研究内容对于引入阴影集的优势可总结如表 7 所示. 由表 7 可知, 在数据分析领域引入阴影集后的共性优势可总结为: 1) 阈值的自动化求取; 2) 数据体量的有效约减; 3) 异常值或噪声样本的识别; 4) 算法效率或精度的提升. 其中, 优势 1) 的原因是其利用了阴影集不确定性不变性准则求取决策划分阈值, 即式 (3), 这无需人为设定参数, 避免了主观性影响. 优势 2) 和优势 3) 的原因本质上是阴影集对数据三划分处理所造成的. Pedrycz 和 Vukovich^[3] 以聚类过程为例, 分析了阴影集对数据结构的解释和层次描述. 如图 6 所示, 基于阴影集的划分思想, 类簇将被划分成核心域、阴影域和不确定域, 三个区域的数据质量依次递减, 而核心域的数据常被视为典型数据点, 并且用于数据分析的下一阶段, 噪声和异常样本则常出现在阴影域或不确定域, 而后者中通常为高度非典型数据点. 显然, 优势 4) 的原因则是优势 2) 和优势 3) 所导致的必然结果, 因为约减和筛选后的数据体量更小且质量更高, 这将有益于算法性能的提升.

1.4 阴影集的应用研究

阴影集的应用研究主要基于其三个优势: 1) 能够实现对不确定性信息的综合考虑, 并给予合理的

表 7 基于阴影集的数据分析研究总结
Table 7 Research summary of data analysis based on shadowed set

类别	研究内容	文献	引入阴影集的优势
聚类	基于阴影集的 C 均值聚类算法 (SCM)	[38]	实现了阈值自动化求取、数据约减、计算效率的提升
	基于多粒度近似区域和阴影集的粗糙可能性 C 均值聚类算法 (MS-RPCM)	[39-40]	实现了阈值自动化求取、数据不确定性问题的解决、噪声抑制能力的增强
	基于阴影集的粗糙模糊 C 均值聚类算法 (S-RFCM)	[41]	实现了阈值自动化求取、异常值的有效识别
	基于阴影集的粗糙模糊可能性 C 均值聚类算法 (S-RFPCM)	[42]	实现了阈值自动化求取、异常值的有效识别
	结合傅立叶随机特征映射和阴影集的粗糙模糊 C 均值聚类算法 (AR-SRFCM)	[43]	实现了非球形数据类簇的有效处理、异常值的有效识别
	基于增量数据的阴影集集群聚类算法 (OSC)	[44]	实现了数据的实时处理、噪声样本的有效识别
	结合阴影集和粒子群优化算法的模糊 C 均值聚类算法 (SP-FCM)	[45]	实现了数据约减、异常值的有效识别、算法效率和精度的提升
	结合阴影集和混合蛙跳算法的粗糙模糊 C 均值聚类算法 (SFLA-SRFCM)	[46]	实现了初值敏感、局部最优和参数设置主观性问题的解决
	基于阴影集的高维混合属性数据集聚类边界检测算法 (CHASM)	[47]	实现了边界检测及噪声抑制能力的增强
	基于阴影集和集成学习的模糊聚类算法及应用场景	[48-49]	实现了阈值自动化求取、异常值的有效识别, 算法鲁棒性增强
	基于区间阴影集的密度峰值聚类算法 (ISS-DPC)	[50-51]	实现了错误分类率的减少、抗噪性的增强
	基于阴影集的截集式可能性 C 均值聚类算法 (C-PCM)	[52]	实现了截集门限的自动化选取
	基于阴影集的广义多粒度粗糙模糊 C 均值聚类算法 (MSRFCM)	[53-54]	实现了数据拓扑结构的捕捉、数据不确定性问题的解决
	基于无监督学习框架下的阴影 C 均值聚类算法 (SCM)	[55]	实现了传统 SCM 算法聚类质量和效率的提升
基于阴影集的多粒度三支聚类集成算法 (MTWCES)	[59-60]	实现了数据不确定性问题的解决、聚类准确率的提升	
分类	基于阴影邻域的三支分类算法 (3WC-SNB)	[56]	实现了不确定性数据的有效分类、分类风险性的降低
	基于半监督阴影集的三支分类算法 (3WC-SSN)	[57-58]	实现了不确定性数据的有效分类、分类风险性的降低
	基于阴影集的快速 KNN 分类算法 (TWC-KNN)	[61-62]	实现了数据约减、数据质量的优化、分类效率的提高
数据预处理	基于阴影集的训练数据选择算法	[63-66]	实现了训练数据质量的提升、训练效率和分类器性能的提升
	基于阴影集的候选断点集提取算法	[67]	实现了候选断点识别效率和准确率的提高
	基于阴影集聚类的离群点检测算法	[68]	实现了聚类离群点检测识别率和识别精度的提高

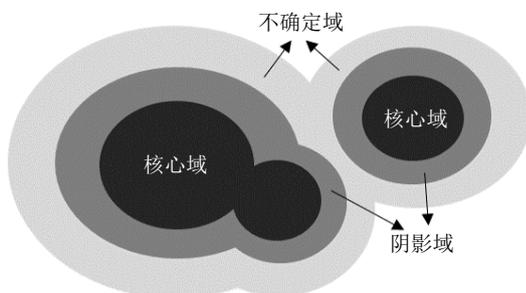


图 6 阴影集在聚类过程中对数据结构的描述
Fig.6 Shadowed set describe the data structure in the clustering process

语义解释; 2) 能够实现模糊信息的自动化三支近似划分处理, 规避人为设定参数所带来的主观性; 3) 作为有效的数据预处理工具, 能够实现数据的约减及质量的优化, 提升处理效率及后续模型性能. 因此, 这也使得阴影集被应用于诸多研究领域, 如图像分类与识别、医疗信息处理、文本分析、决策推荐等.

在图像分类与识别方面: Mitra 和 Kundu^[69] 引入阴影集来处理遥感图像分割过程中像素重叠域的

不确定性问题, 其从像素点聚类的角度入手, 结合 SCM 聚类算法, 通过对边界域之间的不确定性问题建模以及自适应划分来实现图像交叠区域的有效分割, 并通过定性和定量分析证明了该算法在遥感图像和合成图像分割问题上的有效性. Zhang 等^[70] 基于阴影集提出一种图像检索算法, 并开发一种基于显著性检测和边缘检测的阴影集阈值参数自动选择技术. 该算法将图像划分为显著区域、阴影区域和非显著区域, 保留前两个区域来作为检索资源, 以提取图像的有效信息, 忽略无关信息, 有效减少了冗余信息, 降低了信息检索的复杂性, 并提高了图像检索精度. Alavi 和 Kargari^[71] 将阴影集与图像增强领域相结合, 提出一种增强图像灰度对比度的方法. 其通过设计新的隶属度函数来将图像灰度值映射到模糊空间, 并优化图像中的局部参数和全局参数, 从而实现图像对比度的自动提高, 这有益于图像的进一步处理和应用, 如遥感、误差检测和医学图像分析等. 杨晨旭等^[72] 将阴影集应用到人脸图像分类过程中的性别判定问题, 提出一种两阶段的图像分类算法. 在第一阶段通过深度卷积神经网络

络模型分类,并结合阴影集的思想将结果划分到正域、阴影域和负域;在第二阶段将阴影域的不确定性图像再次进行分类,最后融合两个阶段的分类结果,从而实现图像分类精度的提升.相类似地,陈超凡等^[73]将阴影集视为一种狭义的三支决策模型,并基于此构建一种新的两阶段图像分类算法.其首先使用深度学习在第一阶段对图像进行分类,获得隶属度矩阵,并利用阴影集对隶属度矩阵进行划分,获得三支分类结果.随后,在第二阶段利用特征融合方法,结合深度学习和机器学习优势,对上一阶段阴影域中的对象进行二次分类.最后,综合输出两个阶段的分类结果,实现分类精度的提高. Cai 等^[74]分析低质量伪标签的噪声样本会对图像分类产生负面影响,并导致较差的分类结果,因此提出一种基于阴影集的半监督样本选择网络,用于噪声标记图像的分类.该技术将阴影集理论和基于损失相似性的聚类方法相结合,从而实现干净样本的自适应选择,并引入一种高质量伪标签样本选择策略,利用所选择的干净样本和高质量伪标签样本来训练网络,进而完成图像分类.该网络有效减少了训练样本的数量,并提高了图像分类性能.

在医疗信息处理方面: Bose 和 Mali^[75]受阴影集启发,提出一种基于渐进式基数表示的阴影集双聚类模型,并将其应用到基因表达的聚类问题中,从而增强聚类混淆区域基因性质的识别,通过酵母数据、血清数据和小鼠数据三个真实数据集验证了模型在基因表达中的有效性,还结合统计学和生物学证明了阴影集在基因表达问题中的优越性.随后,在 2022 年, Bose 和 Mali^[76]进一步将阴影集与基于乳房 X 光图像的乳腺癌检测场景相结合,将所提出的基于渐进式基数表示的阴影集双聚类模型应用于 X 光图像的对比度增强和微钙化识别问题中,而微钙化的有效识别有助于提高乳腺癌的识别率.其基于阴影集的思想,通过平衡图像不同区域的不确定性,从而将图像划分成“背景、不同密度组织、微钙化”三类.该方法有效增强了整体图像特性,并可以准确地识别“微钙化”,从而提升乳腺癌的识别率.

在文本分析方面: Wang 等^[77]探索了基于阴影集的语言术语建模及其在多属性决策中的应用,其为了有效表达语言变量,开发一种基于数据驱动和语言术语场景的阴影集模型,并将阴影集与勾股模糊集相结合,定义了勾股阴影集,讨论了相关运算规则和基本性质,推导了勾股阴影集的分函数,并在此基础上提出一种新的处理语言术语的多属性决策方法,新方法较好地解决了信息丢失或失真问题,可以提供更合理和准确的决策结果. Li 等^[78]探

究了基于区间数据驱动的阴影集模型构造及其在语言词建模中的应用,针对语言学中常用的模糊集,根据其形状引入四种阴影集,分别命名为常规、左肩、右肩和无核阴影集.随后,利用两种统计方法来确定阴影集的形状和参数,即容差极限法和百分位数法,并提出三个新的评价指标来衡量所构建阴影集的不确定性捕获能力和准确性. Zhang 等^[79]将阴影集与特征提取领域相结合,提出一种基于 TF-IDF 和 GTSS 模型的特征提取方法,其中 TF-IDF 是一种衡量单词与文档关系的常用加权指标,被广泛应用于文本分类或自然语言处理场景下的单词特征提取任务.该方法利用 GTSS 的重复博弈机制,对文档间每个单词 TF-IDF 差异值的重复学习来计算相关单词的划分阈值,从而实现相关单词的特征提取和淘汰.得益于阴影集的优势,该方法能够在保持模型性能的同时,提升处理效率.

在决策推荐方面: He 等^[80]将阴影集与群决策领域结合,提出一种基于阴影集的扩展 TODIM 方法,来解决具有语言信息的大规模群体决策问题,其中 TODIM 是一种经典交互式和多准则决策方法的缩写.新方法可以同时允许决策者以模糊信息的形式表达意见,并考虑了决策者在大规模群体决策问题中的心理行为.其首先基于数据驱动,结合阴影集构建一个用于建模语言术语的码本.随后,提出一种针对阴影集的相似性测度来检测决策者的社会关系.最后,为了提高决策效率,提出一种基于阴影集的聚类模型来对决策者进行划分,并通过装配厂选址实例说明新方法的可行性和合理性. Wu 等^[81]将阴影集与推荐系统相结合,提出一种基于阴影集的推荐算法,实现不确定性信息的有效处理,从而降低决策风险,提高决策质量.其首先将邻域粗糙集引入到推荐系统中以识别活跃用户和相似用户.随后,基于邻域成员关系设计了新的基于不确定性不变性的阴影集,并将用户划分到阴影集的三个区域.最后,对不同区域的用户做出决策,即推荐、延迟推荐和保留推荐.这为不同用户实现了个性化和合理化的推荐,提高了推荐质量.

综上所述,不同应用领域对于引入阴影集的优势可总结如表 8 所示.这些研究成果充分说明,凭借着对模糊信息处理的优势,阴影集可以为解决许多领域的一些算法瓶颈问题提供帮助,而随着阴影集的不断发展和完善,其应用领域也将会更加广泛.

2 阴影集与其他理论模型的联系

事实上,在数据挖掘领域,有众多处理不确定性的理论方法和模型,阴影集理论只是其中一

表 8 基于阴影集的应用研究总结
Table 8 Summary of application research based on shadowed set

类别	研究内容	文献	引入阴影集的优势
图像分类与识别	基于阴影集的遥感图像分割算法	[69]	解决了遥感图像像素交叠区域的不确定性问题, 并提高了遥感图像分割的准确性和稳定性
	基于阴影集的图像信息检索算法	[70]	有效减少了冗余信息, 降低了信息检索的复杂性, 提高了图像检索精度
	基于阴影集的图像对比度增强算法	[71]	优化了图像的局部和全局参数, 实现了图像对比度增强的自动化提升
	基于阴影集和人脸图像的两阶段性别分类算法	[72]	实现了图像数据的多阶段处理, 大大减少了训练时间损耗以及分类的不确定性, 提高了图像的分类精度
	基于阴影集的两阶段图像分类算法	[73]	实现了图像数据的多阶段处理, 大大减少了训练时间损耗以及分类的不确定性, 提高了图像的分类精度
	基于阴影集的半监督样本选择网络在噪声标记图像中的分类算法	[74]	实现了干净样本的自适应选取和高质量伪标签样本的有效选取, 减少了训练样本的数量, 并提高了图像分类性能
医疗信息处理	基于阴影集的基因表达聚类问题及基因性质识别	[75]	增强了聚类混淆区域中基因性质识别的有效性和准确率
	基于阴影集的乳房 X 光图像对比度增强及微钙化识别	[76]	增强了图像不确定性区域的对比度以及微钙化的识别率, 提高了乳腺癌的识别率
文本分析	基于阴影集的语言术语建模及其在多属性决策中的应用	[77]	解决了信息丢失或失真所引起的不确定性问题, 提供了更合理和准确的决策结果
	基于区间数据驱动的阴影集构建及其在语言词建模中的应用	[78]	实现了以区间数据驱动的阴影集模型构建, 为针对语言词问题的建模提供了一种新视角
	基于博弈论阴影集的文本词特征提取算法	[79]	实现了语言词特征的有效提取, 并提高了针对语言词的分类性能和分类效率
决策推荐	基于阴影集的语言信息大规模群体决策算法	[80]	通过阴影集对语言术语进行建模, 实现了对群体决策者意见的综合考虑, 提高了群体决策的效率和有效性
	基于阴影集的用户个性化决策推荐算法	[81]	减少了推荐系统推荐过程中的不确定性, 增强了决策推荐质量, 并能够针对不同用户实现个性化推荐

种, 而随着阴影集理论不断发展, 其与其他方法之间的关系也更加紧密。本节将探讨阴影集理论与其他不确定性问题处理理论模型的联系。

在当前智能信息时代, 数据体量高速增长, 数据特性纷繁复杂。大数据的特性可总结为 4V^[82], 即, Volume, Variety, Velocity, Value。而如何从大数据中去粗取精、去伪存真地挖掘出有效知识则是一个关键问题。传统方法有机器学习、统计学、证据理论、深度学习和归纳学习等。然而, 在将数据转换为知识的过程中, 常常伴随着不确定性问题的存在, 比如: 1) 由于数据的高噪、稀疏和多源异构特性所带来的不确定性; 2) 由于模型设计、方法选择或参数主观设定所带来的不确定性; 3) 由于知识的不完备性或概念漂移所带来的不确定性; 4) 由于知识结果融合、评价标准不一或可解释性差所带来的不确定性。这使得传统数据分析、数据挖掘和处理的方式方法不能够完全适用, 需要新的数据计算范式和模式^[83]。而如何有效解决数据挖掘过程中所存在的不确定性则成为当前大数据领域知识发现的一个关键问题。

在当前人工智能领域中, 粒计算^[84]是用于信息处理的一种新的理念和计算范式, 主要用于处理不确定的、模糊的、不精确的、海量的数据信息, 其基

本思想是利用不同粒化程度或不同粒度上的信息来实现对复杂问题的求解, 并不特指某一具体方法或模型, 而是一种方法论^[85-86]。梁吉业等^[82]曾分析, 从人工智能的视角来看: 粒计算根据实际问题用可行的近似解替换精确解, 从而实现对复杂问题的简化和求解效率的提升, 这是一种模拟人类思考和解决大规模复杂问题的自然模式。而从数据处理与分析的视角来看: 粒计算采用信息粒的形式替换样本个体作为运算基本单元, 这种形式能够将复杂数据进行有效简化, 从而实现计算效率的提高。此外, 王国胤等^[87]曾分析, 知识空间的粒化是粒计算的核心问题之一, 因为知识粒的不确定性将直接对粒计算解决复杂问题的有效性产生影响。因此, 在不同模型下研究数据和知识的不确定性表示与度量问题就受到了诸多学者的关注, 而当前围绕不确定性问题的粒计算理论模型有: 模糊集^[1]、粗糙集^[88]、商空间^[89]、云模型^[90]、阴影集^[2]、三支决策^[91-92]、多尺度粗糙集模型^[93]、形式概念分析^[94]、多粒度认知计算^[86, 95-96]等, 其核心思想总结如表 9 所示。

综上所述, 简要给出不同粒计算理论模型之间的关系图谱, 如图 7 所示。本节将主要关注与阴影集理论相关联的模型, 即讨论阴影集与模糊集、粗糙集以及三支决策理论之间的关系。

表 9 粒计算理论的主要模型及核心思想总结

Table 9 Summary of the main models and core ideas of granular computing theory

年份	文献	理论模型	核心思想
1965	[1]	模糊集	通过使用隶属度函数来对模糊性概念或不确定性信息进行表示与处理, 并进一步通过取值范围在 0 到 1 之间的精确隶属度值来描述一个对象隶属于模糊性概念的程度
1982	[88]	粗糙集	通过利用不可分辨关系构成对象的等价类, 实现对论域空间的近似划分, 并使用具有精确概念的上近似集和下近似集来实现对一个不精确概念的近似表示与度量
1982	[94]	形式概念分析	通过对对象集与属性集之间的某种关联关系, 建立由对象集与属性集对所形成的形式概念层次结构, 并利用哈斯图的形式呈现不同粒概念之间的复杂关系, 实现对知识的挖掘和不确定性推理
1992	[89]	商空间	通过将复杂问题表示成不同粗细的粒度空间, 从而构建多粒度分层递阶空间结构, 通过由粗到细或由细到粗的方式在多粒度空间中进行近似逼近, 将多粒度空间中粒的解组合成原问题的解, 从而获得复杂问题的解
1995	[90]	云模型	通过使用期望、熵和超熵三个参数来对不确定性知识进行描述, 融合人类认知过程中随机与模糊这两种不确定性, 实现知识内涵与外延的相互转换
1998	[2]	阴影集	通过一对决策划分阈值来对模糊集进行近似划分处理, 并通过平衡不同划分区域的不确定性来构建阈值求取的目标函数, 利用这种粗糙的近似表达来替换模糊集过于精确的表达, 从而减少模糊决策过程中不确定性对象的决策划分成本和计算损耗
2009, 2010	[91–92]	三支决策	通过“三分而治”的思想将论域划分为三个子集, 即正域、负域和边界域, 并对不同子集采取不同的决策行为或分治策略, 即接受、拒绝和延迟决策, 进而对相应的行为或策略进行评价和反馈, 实现对复杂问题或信息不充分问题的求解
2011	[93]	多尺度粗糙集模型	通过将属性的取值标记为不同尺度(粒度)层次, 进而讨论知识在不同尺度下的关系及变化规律, 以及在给定尺度好坏评价标准的情况下最优尺度层次的选择方法, 实现在最优尺度空间中对知识的表示和获取
2018, 2019, 2022	[86, 95–96]	多粒度认知计算	通过数据信息驱动, 融合人类“大范围首先”的认知机制(即“由粗粒度到细粒度”的认知过程)以及计算机的信息处理机制(即“由细粒度到粗粒度”的知识挖掘过程), 实现数据和知识双向驱动的空间变换和智能信息处理, 解决认知计算过程中的“数据与知识分离”问题

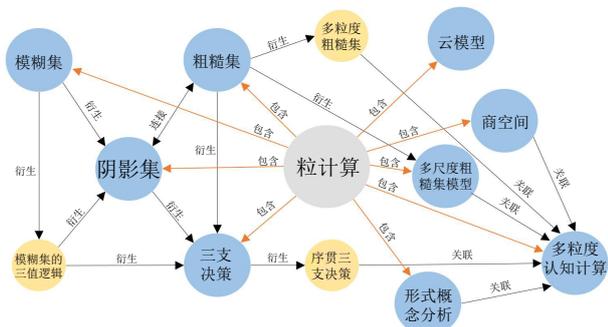


图 7 不同粒计算理论模型之间的简要关系图谱

Fig. 7 Brief relationship graph between different granular computing theory models

阴影集与模糊集: 模糊集被认为是用于捕获、表示和处理模糊问题的载体之一, 正如表 9 中所述, 其通过构造隶属度函数以及精确的隶属度值来描述一个模糊性概念. 然而, 精确的隶属度值就像是一把双刃剑, 对于一般概念的描述并不总是需要完全精确的隶属度, 而且过度的精度也伴随着巨大的计算损耗^[2–4]. 因此, 阴影集通过一种不太精确和粗糙的结构来近似描述模糊集, 其通过提升和减小隶属度的操作, 将原本分布在整个模糊集空间的不确定性缩小到了阴影区域^[2–4], 即区间 $[0, 1]$, 并通过平衡不同区域的不确定性来求取决策划分阈值, 划分思想如图 8 所示.

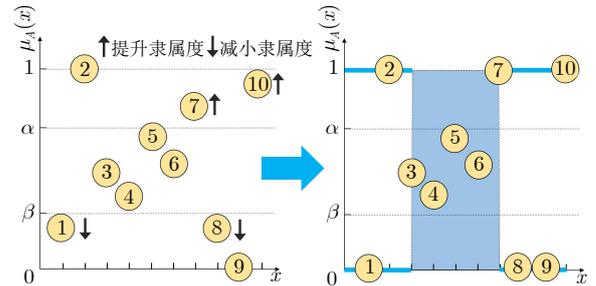


图 8 基于阴影集思想的模糊集近似划分示意图

Fig. 8 Schematic diagram of fuzzy sets approximation partition based on shadowed set idea

事实上, 文献 [3–4] 认为阴影集的这种映射结构也可以被视为模糊集三值逻辑的一种例子, 所不同的是后者缺乏对划分阈值的研究以及合理的语义解释, 而前者则对此做了充分的分析. 此外, Pedrycz^[2] 在首次提出阴影集理论时也曾指出: 可以将阴影集视为一种操作框架, 用于简化模糊集的处理, 增强对模糊集结果的解释. 简而言之, 阴影集是由模糊集诱导的信息粒, 在捕捉模糊集本质的同时, 减少模糊集过精度所带来的计算负担, 在计算和推理中提供了更高的效率和更好的可解释性.

阴影集与粗糙集: 由表 9 可知, 粗糙集通过定义一对上下近似集来实现对不确定性概念或集合的近似表示与度量. 如图 9 所示, 在整个论域空间中,

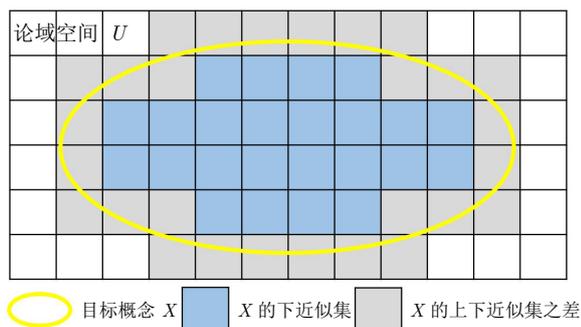


图 9 粗糙集理论模型示意图

Fig.9 Schematic diagram of rough sets theory model

粗糙集下近似集中的等价类, 即蓝色区域 (正域), 被视为完全属于目标概念; 上下近似集差值中的等价类, 即灰色区域 (边界域), 被视为不确定是否属于目标概念; 而上近似集之外的等价类, 即白色区域 (负域), 被视为完全不属于目标概念。

显然, 阴影集和粗糙集从模型思想来讲是相似的, 两者相似之处: 1) 粗糙集正域中等价类对象与阴影集中隶属度为 1 的对象都被视为完全属于目标概念; 2) 粗糙集边界域中等价类对象与阴影集中阴影域的对象都被视为不确定是否属于目标概念; 3) 粗糙集负域中等价类对象与阴影集中隶属度为 0 的对象都被视为完全不属于目标概念。

此外, 文献 [3-4] 指出: 阴影集架起了模糊集与粗糙集之间的桥梁, 减少了两者在概念和算法上的差异, 揭示了它们是如何直接相关的. 而两者也有一些显著区别: 在粗糙集中, 划分对象是基于等价类的形式, 并且等价类对象是固定的, 近似空间被预先定义; 在阴影集中, 划分对象则是模糊数据的形式, 并且模糊对象是动态分配的, 近似空间逐步生成。

阴影集与三支决策: 三支决策是 Yao^[91-92] 在粗糙集的基础上, 结合对损失风险的考虑以及贝叶斯理论所提出的一种处理不确定性决策问题的方法论, 其对粗糙集所形成的三个划分区域赋予了新的语义解释. 刘盾和梁德翠^[97] 分别从宏观和微观的角度将三支决策理论分为广义三支决策和狭义三支决策, 前者注重对概念内涵和外延的诠释, 后者注重三支决策在实际问题中的语义解释. Yao^[98] 进一步指出: 狭义三支决策指代具体的数学模型, 这些模型提供了特定条件下做出三支决策的方法; 而广义三支决策则更多指概念思想的理解, 难以用具体数学模型描述, 其研究重点是三支决策的哲学基础和方法论. 因此, Yao^[99-100] 从宏观角度将三支决策表示为 TAO (Trisecting-acting-outcome) 模型, 其核心思想如图 10 所示, 其将整个论域空间划分为三

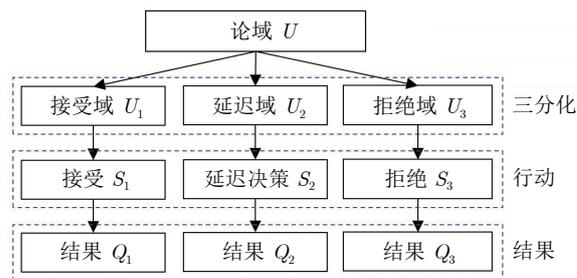


图 10 TAO 模型示意图

Fig.10 Schematic diagram of TAO model

个区域: 接受域、拒绝域和延迟决策域, 并对不同区域采取不同策略, 进而得到不同的分治结果。

显然, 阴影集三支近似划分的思想与三支决策三支划分的思想是一致的, 两者都将论域划分为三个区域, 并针对不同区域采取不同策略, 赋以不同的语义解释. 两者的相似之处: 1) 均利用 (α, β) 对 U 进行三支划分, 形成三个子域; 2) 三支决策接受域中的对象被视为接受该对象属于决策目标类, 这与阴影集中隶属度为 1 的对象含义相同; 3) 同理, 拒绝域中的对象与隶属度为 0 的对象含义相同, 延迟域中的对象与阴影域中的对象含义相同。

此外, 文献 [8-9] 指出: 阴影集是模糊集的一种三支、三值或三区域近似的例子, 其思想与三支决策理论一样. 而文献 [98, 101] 指出: 与粗糙集类似, 三支决策也可以基于区间集或阴影集来实现, 这些集合论模型都是狭义三支决策的体现. 综上所述, 在三支决策理论提出以后, 阴影集可以被视为一种用于处理模糊数据的三支决策模型, 其三支近似的划分思想包含在三支决策理论思想的范畴之内. 换言之, 从宏观角度来看, 阴影集的思想包含在广义三支决策的思想范畴内; 从微观角度来看, 阴影集的数学模型则是狭义三支决策的具体体现。

几种理论模型间的关系及互补性: 首先, 通过篮球运动员挑选这一应用实例来体现上述模型间的关系和互补性. 给定一个信息系统如表 10 所示, 论域 $U = \{x_1, x_2, \dots, x_8\}$, 针对这一复杂目标, 设定条件属性集为 {身高, 力量, 篮球技术}, 决策属性中的 1 表示被选为篮球运动员, 0 表示未被挑选. 显然, 三列属性下的数据格式不同, 身高属性下为模糊数据, 力量和技术属性下为字符型数据. 其中, 模糊数据可通过模糊集来表示, 并通过阴影集来近似处理, 即表 10 第三列, 处理后的属性值可被视为一种字符型数据. 此时, 综合考虑三列属性便可进行论域中对象的等价类划分, 即 $U/C = \{X_1, X_2, \dots, X_7\}$, 其中 $X_1 = \{x_1, x_3\}$, $X_2 = \{x_2\}$, $X_3 = \{x_4\}$, $X_4 = \{x_5\}$, $X_5 = \{x_6\}$, $X_6 = \{x_7\}$, $X_7 = \{x_8\}$. 随

表 10 篮球运动员挑选
Table 10 Selection of basketball players

对象	身高 (m)	身高 ($S_{\mu_A}(x)$)	力量	篮球技术	决策
x_1	1.95	1	大	好	1
x_2	2.00	1	大	一般	1
x_3	2.10	1	大	好	1
x_4	1.98	1	小	一般	0
x_5	1.85	[0, 1]	大	差	0
x_6	1.73	0	中	差	0
x_7	1.88	[0, 1]	中	好	1
x_8	1.70	0	小	好	1

后, 便可通过粗糙集对等价类进行三支决策划分, 而用于决策划分的粗糙集模型有多种, 如概率粗糙集和决策粗糙集, 而三支决策便是由决策粗糙集模型演变而来的, 更详细的介绍可参考文献 [102–104].

综上所述, 给出模糊集、粗糙集、阴影集、三支决策以及粒计算理论之间关系图, 如图 11 所示. 其中, 阴影集通过对模糊集中模糊数据的三支近似处理, 架起了模糊集与粗糙集之间的桥梁, 而近似处理后的数据可支持等价类的形成, 进而支持粗糙集进行决策划分. 三支决策则是粗糙集理论的扩展和外延, 粗糙集理论则是三支决策的特例和内涵, 三支决策的提出源于粗糙集, 但高于粗糙集^[102]. 这几种模型的思想都包含在粒计算方法论的思想范畴内, 即利用不同粒化程度或不同粒度上的信息来实现对复杂问题的求解. 因此, 基于上述几种模型的互补性, 或许可通过集成的架构来处理包含多种数据形式的复杂目标任务, 对此, 本文不做详细探讨.

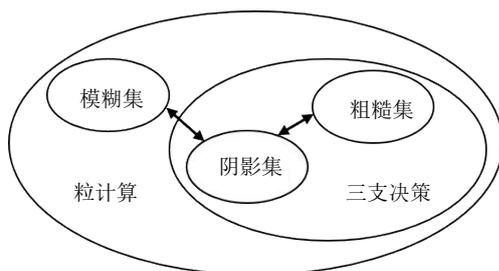


图 11 上述模型之间的关系

Fig. 11 The relationship between the above models

3 问题分析与研究展望

本节围绕阴影集的模式构建、理论性质、数据分析及应用研究四个方面, 对当前若干具有挑战性的研究问题进行总结分析, 便于今后对它们做进一步的研究, 以期推动阴影集理论的不断发展和创新.

1) 多数阴影集模型的构建仅考虑了单一准则

的约束, 缺乏对多准则约束的综合考虑. 如表 1 所示, 当前阴影集的构建主要基于不确定性不变性约束准则和最小成本约束准则^[9], 缺乏对其他类型约束准则的探讨以及多准则综合约束的考虑. 这使得所构建模型容易忽略其他准则, 仅在单一准则目标下达到最优, 但在其他准则下结果可能较差. 这同时也削弱了模型在实际应用场景中的适用性、扩展性和语义解释的合理性. 事实上, 在当前数据分析和知识发现领域的建模过程中, 对多个约束准则的综合考虑需求是非常普遍的. 如, 通常有多个准则来评估分类算法的结果, 即准确度、覆盖率或召回率等. 在分类任务中, 也总是期望每个评价准则都能达到最好, 而不仅只是其中一项. 此外, 同样作为不确定性问题的处理模型, 在粗糙集领域, 已经有针对多准则约束问题的粗糙集模型构建研究, 并在理论和应用方面衍生了诸多相关工作^[105–106]. 因此, 在阴影集模型构建方面, 如何综合考虑和平衡不同评价准则或多个优化目标, 从而将传统阴影集构建从单准则考虑推广到多准则考虑是一个值得研究的问题. 虽然 Gao 等^[17] 在最新的工作中结合博弈论初步探讨了不确定性和决策代价两种评价准则约束下的模型构建, 但对于多准则约束问题的解决方法并不唯一, 比如将阴影集与多目标优化算法、多任务学习或 TOPSIS 多指标融合算法结合等, 这依旧是一个值得探究的问题.

2) 多数阴影集模型的构建仅考虑了单粒度 (单粒层) 空间, 缺乏在多粒度 (多粒层) 空间下的阴影集构建研究. 当前阴影集及各种扩展模型的构建所关注的问题可概括为两个方面: a) 三个区域的表达形式; b) 决策阈值对的获取方法和语义解释. 并且都是基于单一模糊集所展开的, 该模糊集通常由单个属性所构成, 即形成单粒度 (单粒层) 空间. 缺乏对多个属性的综合考虑, 即多粒度 (多粒层) 空间下阴影集的三支近似划分研究. 这使得阴影集无法有效处理复杂目标概念的近似划分问题. 事实上, 在现实决策问题场景下, 一个复杂的目标概念往往需要考虑多种属性、标准或专家意见. 如, 在挑选篮球运动员时, 就需要同时考虑身高、体重、年龄、耐力等多种属性; 在诊断患者是否患有糖尿病时, 不同专家给出的诊断意见也可能并不相同. 而每种属性、标准或专家意见都可构成一个模糊集合, 即单粒度 (单粒层) 空间. 显然, 对于复杂的目标概念或决策问题, 往往需要同时考虑多个模糊集的近似划分结果, 即实现多粒度空间下的阴影集构建. 此外, Qian 等^[107] 曾在多粒度粗糙集的研究中指出: 在多源信息系统中, 每个特征属性都可诱导形成一个粒度

(粒层) 空间, 而从结合粒计算的视角来看, 基于单一属性进行粗糙集的近似划分所形成的是一个单粒度(单粒层)粗糙集. 因此, 同样从粒计算的观点来看, 基于单一模糊集的近似划分所形成的是单粒度(单粒层)阴影集, 而基于多个模糊集的近似划分所形成的则是多粒度(多粒层)阴影集. 因此, 如何将阴影集的构建从单粒度(单粒层)空间推广到多粒度(多粒层)空间, 并进一步研究其在多粒度(多粒层)空间下的表达形式、阈值选择、语义解释和划分结果的融合等, 都是值得探究的问题, 比如将阴影集与多粒度粗糙集或多尺度决策系统相结合等.

3) 阴影集模型的研究大多是针对静态数据考虑的, 缺乏动态或增量数据背景下的阴影集构建研究. 当前阴影集模型主要研究对单个模糊集中不确定性对象的近似划分问题, 并且大多数只讨论了模糊对象是静态的情况, 其中模糊集可被视为是多源信息系统的某一系列特征属性. 然而, 由当前智能信息时代所反映出的大数据 4V 特性^[82]可知, 数据信息体量浩大且增长快速. 因此, 复杂动态的多源信息系统则更为常见, 并且现实场景中复杂目标概念的决策近似划分需求也更为常见. 显然, 这就需要在阴影集的构建过程中, 结合对动态或增量数据背景及需求的考虑, 以及对多列特征属性的综合考虑. 而对于信息系统的动态或增量情况大致可分为三种^[108]: a) 对象的添加和删除; b) 特征属性的添加和删除; c) 特征属性值的更新和修改. 虽然 Cai 等^[30]探讨了基于动态模糊集的阴影集构建方法, 但其主要是对具有时变隶属度模糊集的考虑, 即模糊集的隶属度随时间变化的场景下阴影集的构建, 可以视为是对情况 c) 的考虑, 而对前两种情况的研究依旧缺乏. 因此, 如何在动态或增量数据背景下实现单粒度(单粒层)或多粒度(多粒层)阴影集的构建, 从而将阴影集从静态数据空间的考虑推广到动态或增量数据空间的考虑, 实现动态数据的知识挖掘就是一个值得探究的问题. 比如将阴影与序贯三支决策的思想相结合, 通过属性的动态添加来考虑模糊对象的逐步划分.

4) 针对不同阴影集模型之间运算算子、规则及性质的研究还比较少, 并且缺乏针对不同模糊现象下不同运算需求的考虑. 由第 2 节分析可知, 阴影集有效地将模糊集和粗糙集联系起来. 而从完善模型理论性质的角度来看, 对集合之间运算规则、方法设计及规律性质的研究是非常有必要的, 许多学者对此也做了相关研究. Pawlak^[88]提出并研究了粗糙集理论及其基本运算算子的设计. Zhang 等^[109]进一步研究了粗糙集近似集的基本运算算子及其不

确定性变化规律. Zadeh^[1]提出并研究了模糊集的基本运算算子及性质. Pedrycz^[2]提出并定义了阴影集的基本运算算子, 即模糊集的近似集. 最近, Boffa 等^[28]以代数系统的同构方程为载体, 讨论和分析了阴影集与模糊集运算算子之间的关系. 此外, 由第 1.1 节可知, 随着阴影集理论的不断发展和完善, 已经衍生出诸多扩展阴影集模型, 对于阴影域的表达形式也有诸多改进, 如图 3 所示. 然而, 在当前对阴影集运算算子的研究中, 只分析了经典阴影集的基本算子, 即阴影域为区间值 $[0, 1]$ 的情况, 缺乏对不同阴影集之间运算算子、规则及性质的研究. 此外, 在模糊数学领域, 为了使模糊集能够对不同模糊现象进行合理描述, 在“析取合取”经典算子的基础上, 也衍生出了诸多模糊运算算子, 如代数和与积、有界和与积、Hamacher 算子和 Yager 算子等. 这些算子优缺点及特性迥异, 可根据具体问题和需求选择使用. 然而, 当前对于阴影集运算算子的设计仅考虑了集合运算中经典的“析取合取”形式, 缺乏对不同模糊现象下不同运算需求的考虑. 因此, 如何定义不同阴影集模型之间的运算算子、运算规则, 并挖掘其中的运算性质就成为了一个值得探究的问题, 而参考当前模糊集运算算子来设计不同类型的阴影集运算算子或许是可行的方案. 这也将有助于探究不同阴影集模型之间的特性和优势以及结合的可能性, 更好地实现不确定性信息的处理和知识挖掘.

5) 基于阴影集理论的数据分析及应用研究中, 还缺乏对不同扩展模型应用效果的探讨, 并且所构建的阴影集模型缺乏对任务需求的考虑. 由第 1.3 节和第 1.4 节可知, 阴影集被广泛应用于不同研究领域, 这得益于经典阴影集模型构建的简单性和对不确定性信息处理的有效性. 然而, 这些研究工作大多是结合经典阴影集模型所展开的, 缺乏对各种扩展模型的考虑. 由第 1.1 节可知, 在经典阴影集模型被提出后, 衍生出了各种改进模型和针对不同需求考虑的扩展模型. 显然, 其构造方法、模型优势及特性迥异, 并且不同模型对于数据分析和应用场景的影响也显著不同. 因此, 不同阴影集模型在数据分析和现实应用场景下的有效性是一个值得探究的问题. 此外, 当前阴影集模型的优劣评价主要是从不确定性信息处理的角度出发, 考虑了 Yao 等^[9]所总结的三个准则. 然而, 这些评估准则更侧重于考虑模型构建和决策阈值的语义合理性解释, 缺乏对决策划分结果有效性评估的考虑. 这使得阴影集在模糊信息处理的体系下能获得有效的结果和解释, 但在其他研究任务体系下则可能未取得最好的结果. 因此, 如何以数据分析或实际应用任务为导

向来构造阴影集模型的目标函数也是一个值得探究的问题. 比如, 基于聚类或分类的任务场景下, 以分类精度或准确度为目标来构建阴影集; 基于图像识别的任务场景下, 以识别准确率为目标来构建阴影集等.

4 结束语

阴影集在不确定性信息处理过程中所表现出的简单性和有效性使其受到了大量学者的关注. 本文从四个方面对阴影集的研究工作进行了总结和分析: 在模型构建方面, 介绍了不同阴影集的核心思想、方法体系及模型差异, 如表 1 所示; 在理论性质方面, 介绍了基于阴影集的模糊关系、结构解释、逻辑运算或数学性质等研究, 以及基于不同模糊集的阴影集构建研究, 如表 5 和表 6 所示; 在数据分析方面, 介绍了基于阴影集的聚类、分类算法以及数据预处理研究, 如表 7 所示; 在应用研究方面, 介绍了基于阴影集的图像分类与识别、医疗信息处理、文本分析和决策推荐研究, 如表 8 所示. 此外, 还讨论分析了阴影集与粗糙集、模糊集和三支决策等处理不确定性问题的理论模型之间的联系、区别及互补性. 最后, 分析了当前所存在的若干挑战性问题, 即, 如何将阴影集的构建考虑从单准则约束推广到多准则约束; 如何从单粒度 (单粒层) 空间推广到多粒度 (多粒层) 空间; 如何从静态数据空间推广到动态或增量数据空间; 如何定义不同阴影集模型之间的运算算子、运算规则, 并挖掘其中的运算性质; 如何结合模糊集运算算子的考虑设计不同类型的阴影集运算算子; 如何以数据分析或实际应用任务为导向来构造阴影集模型的目标函数. 本文工作将为关注粒计算、不确定性问题处理和知识发现领域的研究者提供一些理论参考, 并促进阴影集的不断发展和推广应用.

References

- Zadeh L A. Fuzzy sets. *Information and Control*, 1965, **8**(3): 338–353
- Pedrycz W. Shadowed sets: Representing and processing fuzzy sets. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 1998, **28**(1): 103–109
- Pedrycz W, Vukovich G. Granular computing with shadowed sets. *International Journal of Intelligent Systems*, 2002, **17**(2): 173–197
- Pedrycz W. From fuzzy sets to shadowed sets: Interpretation and computing. *International Journal of Intelligent Systems*, 2009, **24**(1): 48–61
- Cattaneo G, Ciucci D. Shadowed sets and related algebraic structures. *Fundamenta Informaticae*, 2003, **55**(3–4): 255–284
- Cattaneo G, Ciucci D. An algebraic approach to shadowed sets. *Electronic Notes in Theoretical Computer Science*, 2003, **82**(4): 64–75
- Deng X F, Yao Y Y. Mean-value-based decision-theoretic shadowed sets. In: Proceedings of the Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS). Edmonton, Canada: IEEE, 2013. 1382–1387
- Deng X F, Yao Y Y. Decision-theoretic three-way approximations of fuzzy sets. *Information Sciences*, 2014, **279**: 702–715
- Yao Y Y, Wang S, Deng X F. Constructing shadowed sets and three-way approximations of fuzzy sets. *Information Sciences*, 2017, **412–413**: 132–153
- Zhang Y, Yao J T. Determining strategies in game-theoretic shadowed sets. In: Proceedings of the 17th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems. Cádiz, Spain: Springer, 2018. 736–747
- Zhang Y, Yao J T. Game theoretic approach to shadowed sets: A three-way tradeoff perspective. *Information Sciences*, 2020, **507**: 540–552
- Zhang Q H, Chen Y H, Yang J, Wang G Y. Fuzzy entropy: A more comprehensible perspective for interval shadowed sets of fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 2020, **28**(11): 3008–3022
- Gao M, Zhang Q H, Zhao F, Wang G Y. Mean-entropy-based shadowed sets: A novel three-way approximation of fuzzy sets. *International Journal of Approximate Reasoning*, 2020, **120**: 102–124
- Zhang Q H, Gao M, Zhao F, Wang G Y. Fuzzy-entropy-based game theoretic shadowed sets: A novel game perspective from uncertainty. *IEEE Transactions on Fuzzy Systems*, 2020, **30**(3): 597–609
- Gao Man. Research on Shadowed Sets Model Based on Uncertainty Analysis [Master thesis], Chongqing University of Posts and Telecommunications, China, 2020. (高满. 基于不确定性分析的阴影集模型研究 [硕士学位论文], 重庆邮电大学, 中国, 2020.)
- Luo Z Q, Hu J, Zhang Q H, Wang G Y. Induction of interval shadowed sets from the perspective of maintaining fuzziness. *International Journal of Approximate Reasoning*, 2023, **153**: 219–238
- Gao M, Zhang Q H, Zhao F, Wu C Y, Wang G Y, Xia D Y. Constructing shadowed set based on game analysis of uncertainty and decision cost. *Applied Soft Computing*, 2023, **147**: Article No. 110762
- Vincent T L. Game theory as a design tool. *Journal of Mechanical Design*, 1983, **105**(2): 165–170
- Rao S S. Game theory approach for multiobjective structural optimization. *Computers & Structures*, 1987, **25**(1): 119–127
- Liu X C. Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems*, 1992, **52**(3): 305–318
- Grzegorzewski P. Fuzzy number approximation via shadowed sets. *Information Sciences*, 2013, **225**: 35–46
- Zhou J, Miao D Q, Gao C, Lai Z H, Yue X D. Constrained three-way approximations of fuzzy sets: From the perspective of minimal distance. *Information Sciences*, 2019, **502**: 247–267
- Zhang Q H, Xia D Y, Liu K X, Wang G Y. A general model of decision-theoretic three-way approximations of fuzzy sets based on a heuristic algorithm. *Information Sciences*, 2020, **507**: 522–539
- Xia De-You. Research on Three-way Decision-theoretic With Rough Set Models and Its Application [Master thesis], Chongqing University of Posts and Telecommunications, China, 2018. (夏德友. 三支决策粗糙集模型及其应用研究 [硕士学位论文], 重庆邮电大学, 中国, 2018.)
- Zhou J, Gao C, Pedrycz W, Lai Z H, Yue X D. Constrained

- shadowed sets and fast optimization algorithm. *International Journal of Intelligent Systems*, 2019, **34**(10): 2655–2675
- 26 Zhao X R, Yao Y Y. Three-way fuzzy partitions defined by shadowed sets. *Information Sciences*, 2019, **497**: 23–37
- 27 Ciucci D, Yao Y Y. Synergy of granular computing, shadowed sets, and three-way decisions. *Information Sciences*, 2020, **508**: 422–425
- 28 Boffa S, Campagner A, Ciucci D, Yao Y Y. Aggregation operators on shadowed sets. *Information Sciences*, 2022, **595**: 313–333
- 29 Tahayori H, Sadeghian A, Pedrycz W. Induction of shadowed sets based on the gradual grade of fuzziness. *IEEE Transactions on Fuzzy Systems*, 2013, **21**(5): 937–949
- 30 Cai M J, Li Q G, Lang G M. Shadowed sets of dynamic fuzzy sets. *Granular Computing*, 2017, **2**(2): 85–94
- 31 Bose A, Mali K. A two threshold model for shadowed set with gradual representation of cardinality. In: Proceedings of the 14th IEEE India Council International Conference (INDICON). Roorkee, India: IEEE, 2017. 1–6
- 32 Ibrahim A M, William-West T O. Induction of shadowed sets from fuzzy sets. *Granular Computing*, 2019, **4**(1): 27–38
- 33 William-West T O, Ibrahim A M, Kana A F D. Shadowed set approximation of fuzzy sets based on nearest quota of fuzziness. *Annals of Fuzzy Mathematics and Informatics*, 2019, **17**(2): 133–145
- 34 William-West T O, Ibrahim M A. Trade-off principle for standard shadowed sets and its generalization to five-regions. *Fuzzy Sets and Systems*, 2023, **461**: Article No. 108373
- 35 Campagner A, Dorigatti V, Ciucci D. Entropy-based shadowed set approximation of intuitionistic fuzzy sets. *International Journal of Intelligent Systems*, 2020, **35**(12): 2117–2139
- 36 Yang J L, Yao Y Y. A three-way decision based construction of shadowed sets from Atanassov intuitionistic fuzzy sets. *Information Sciences*, 2021, **577**: 1–21
- 37 Yao Y Y, Yang J L. Granular rough sets and granular shadowed sets: Three-way approximations in Pawlak approximation spaces. *International Journal of Approximate Reasoning*, 2022, **142**: 231–247
- 38 Mitra S, Pedrycz W, Barman B. Shadowed c-means: Integrating fuzzy and rough clustering. *Pattern Recognition*, 2010, **43**(4): 1282–1291
- 39 Zhou J, Pedrycz W, Miao D Q. Shadowed sets in the characterization of rough-fuzzy clustering. *Pattern Recognition*, 2011, **44**(8): 1738–1749
- 40 Zhou J, Lai Z H, Gao C, Miao D Q, Yue X D. Rough possibilistic C-means clustering based on multigranulation approximation regions and shadowed sets. *Knowledge-based Systems*, 2018, **160**: 144–166
- 41 Guo Jin-Hua, Miao Duo-Qian, Zhou Jie. Shadowed sets based threshold selection in rough clustering. *Computer Science*, 2011, **38**(10): 209–210
(郭晋华, 苗夺谦, 周杰. 基于阴影集的粗糙聚类阈值选择. 计算机科学, 2011, **38**(10): 209–210)
- 42 Wang Hai-Liang, She Kun, Zhou Ming-Tian. Shadowed sets-based rough fuzzy possibilistic C-means clustering. *Computer Science*, 2013, **40**(1): 191–194
(汪海良, 余莹, 周明天. 基于阴影集的粗糙模糊可能性 C 均值聚类算法. 计算机科学, 2013, **40**(1): 191–194)
- 43 Kong L N, Chen L. Shadowed set-based rough-fuzzy clustering using random feature mapping. In: Proceedings of the International Conference on Security, Pattern Analysis, and Cybernetics (SPAC). Shenzhen, China: IEEE, 2017. 400–405
- 44 Falcon R, Nayak A, Abielmona R. An online shadowed clustering algorithm applied to risk visualization in territorial security. In: Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defence Applications. Ottawa, Canada: IEEE, 2012. 1–8
- 45 Li Xiu-Xin, Wang Jing-Dong, Xu Ye-Ye, Wen Jia-Wang. Satellite image clustering research based on improved FCM algorithm. *Infrared Technology*, 2013, **35**(3): 150–154
(李秀馨, 王敬东, 徐烨晔, 温家旺. 基于改进 FCM 算法的卫星云图聚类方法研究. 红外技术, 2013, **35**(3): 150–154)
- 46 Meng Zu-Qiang, Hu Yu-Lan, Jiang Liang, Chang Hong-Yan. Shuffled frog leaping algorithm and shadowed sets based rough fuzzy clustering algorithm. *Control and Decision*, 2015, **30**(10): 1766–1772
(蒙祖强, 胡玉兰, 蒋亮, 常红岩. 基于混合蛙跳与阴影集优化的粗糙模糊聚类算法. 控制与决策, 2015, **30**(10): 1766–1772)
- 47 Li X L, Geng P, Qiu B Z. A cluster boundary detection algorithm based on shadowed set. *Intelligent Data Analysis*, 2016, **20**(1): 29–45
- 48 Wang Li-Na. Research on Fuzzy Clustering Algorithms Based on Shadowed Sets and Rough Sets and Their Applications [Ph. D. dissertation], Nanjing University of Aeronautics and Astronautics, China, 2016.
(王丽娜. 基于阴影集和粗糙集的模糊聚类算法研究与应用 [博士学位论文], 南京航空航天大学, 中国, 2016.)
- 49 Wang Li-Na, Wang Jian-Dong, Li Tao, Ye Feng. Cluster's feature weighting fuzzy clustering algorithm integrating rough sets and shadowed sets. *Systems Engineering and Electronics*, 2013, **35**(8): 1769–1776
(王丽娜, 王建东, 李涛, 叶枫. 集成粗糙集和阴影集的簇特征加权模糊聚类算法. 系统工程与电子技术, 2013, **35**(8): 1769–1776)
- 50 Chen Yu-Hong, Zhang Qing-Hua, Yang Jie. Density peak clustering algorithm based on interval shadowed sets. *Pattern Recognition and Artificial Intelligence*, 2019, **32**(6): 531–544
(陈玉洪, 张清华, 杨洁. 基于区间阴影集的密度峰值聚类算法. 模式识别与人工智能, 2019, **32**(6): 531–544)
- 51 Chen Yu-Hong. Interval Shadowed Set Model Based on Fuzzy Entropy and Its Application [Master thesis], Chongqing University of Posts and Telecommunications, China, 2019.
(陈玉洪. 基于模糊熵的区间阴影集模型及其应用 [硕士学位论文], 重庆邮电大学, 中国, 2019.)
- 52 Luo Xi, Fan Jiu-Lun, Yu Hai-Yan, Liang Dan. Selection of cut-set threshold for cutset-type possibilistic C-means clustering based on shadowed set. *Computer Science*, 2019, **46**(8): 249–254
(雒维, 范九伦, 于海燕, 梁丹. 基于阴影集的截集式可能性 C-均值聚类截集阈值的选取. 计算机科学, 2019, **46**(8): 249–254)
- 53 Zhou J, Lai Z H, Miao D Q, Gao C, Yue X D. Multigranulation rough-fuzzy clustering based on shadowed sets. *Information Sciences*, 2020, **507**: 553–573
- 54 Zhou J, Pedrycz W, Gao C, Lai Z H, Yue X D. Principles for constructing three-way approximations of fuzzy sets: A comparative evaluation based on unsupervised learning. *Fuzzy Sets and Systems*, 2021, **413**: 74–98
- 55 William-West T, Kana A F D, Ibrahim M A. Shadowed-set-based three-way clustering methods: An investigation of new optimization-based principles. *Information Sciences*, 2022, **591**: 1–24
- 56 Yue X D, Zhou J, Yao Y Y, Miao D Q. Shadowed neighborhoods based on fuzzy rough transformation for three-way classification. *IEEE Transactions on Fuzzy Systems*, 2020, **28**(5): 978–991
- 57 Yue X D, Liu S W, Qian Q, Miao D Q, Gao C. Semi-supervised shadowed sets for three-way classification on partial labeled data. *Information Sciences*, 2022, **607**: 1372–1390
- 58 Liu Si-Wen. Research on Classification Method With Uncertain Shadowed Prototypes [Master thesis], Shanghai University, China, 2022.
(刘思雯. 基于不确定阴影原型的分类方法研究 [硕士学位论文], 上海大学, 中国, 2022.)

- 59 Jiang Chun-Mao, Zhao Shu-Bao. Multi-granulation three-way clustering ensemble based on shadowed sets. *Acta Electronica Sinica*, 2021, **49**(8): 1524–1532
(姜春茂, 赵书宝. 基于阴影集的多粒度三支聚类集成. 电子学报, 2021, **49**(8): 1524–1532)
- 60 Jiang C M, Li Z C, Yao J T. A shadowed set-based three-way clustering ensemble approach. *International Journal of Machine Learning and Cybernetics*, 2022, **13**(9): 2545–2558
- 61 Zhao Shu-Bao, Jiang Chun-Mao. Fast KNN classification algorithm based on three-way clustering. *Journal of Chinese Computer Systems*, 2021, **42**(9): 1845–1851
(赵书宝, 姜春茂. 一种基于三支聚类的快速 KNN 算法. 小型微型计算机系统, 2021, **42**(9): 1845–1851)
- 62 Zhao Shu-Bao. Clustering Ensemble and Classification Under the Perspective of Three-way Decision [Master thesis], Harbin Normal University, China, 2021.
(赵书宝. 三支决策视角下的聚类集成与分类研究 [硕士学位论文], 哈尔滨师范大学, 中国, 2021.)
- 63 Zhou Y, Su H B, Zhang H T. A novel data selection method based on shadowed sets. *Procedia Engineering*, 2011, **15**: 1410–1415
- 64 Zhou Yu, Qian Xu, Wang Zi-Qiang. Performance improvement of extension neural network using data selection method based on shadowed sets. *Journal of Beijing University of Technology*, 2013, **39**(3): 430–437
(周玉, 钱旭, 王自强. 基于阴影集数据选择的可拓神经网络性能改进. 北京工业大学学报, 2013, **39**(3): 430–437)
- 65 Zhou Yu, Zhu An-Fu, Zhou Lin, Qian Xu. Sample data selection method for neural network classifiers. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 2012, **40**(6): 39–43
(周玉, 朱安福, 周林, 钱旭. 一种神经网络分类器样本数据选择方法. 华中科技大学学报(自然科学版), 2012, **40**(6): 39–43)
- 66 Su Xiao-Hong, Zhao Ling-Ling, Xie Lin, Ma Pei-Jun. Shadowed sets-based sample selection method for fuzzy support vector machine. *Journal of Harbin Institute of Technology*, 2012, **44**(9): 78–84
(苏小红, 赵玲玲, 谢琳, 马培军. 阴影集的模糊支持向量机样本选择方法. 哈尔滨工业大学学报, 2012, **44**(9): 78–84)
- 67 Zhou Fan-Cheng, Wu Meng-Da, Wang Dan. Discretization of continuous attributes based on shadowed sets. *Fuzzy Systems and Mathematics*, 2012, **26**(2): 120–128
(周凡程, 吴孟达, 王丹. 基于 Shadowed Sets 的连续属性离散化. 模糊系统与数学, 2012, **26**(2): 120–128)
- 68 Wang Dan, Mao Zi-Yang, Wu Meng-Da. Outlier detection algorithm on shadowed sets clustering. *Journal of Frontiers of Computer Science and Technology*, 2012, **6**(11): 985–993
(王丹, 毛紫阳, 吴孟达. 融合 Shadowed Sets 聚类的离群点检测算法. 计算机科学与探索, 2012, **6**(11): 985–993)
- 69 Mitra S, Kundu P P. Satellite image segmentation with shadowed C-means. *Information Sciences*, 2011, **181**(17): 3601–3613
- 70 Zhang H Y, Zhang T, Pedrycz W, Zhao C R, Miao D Q. Improved adaptive image retrieval with the use of shadowed sets. *Pattern Recognition*, 2019, **90**: 390–403
- 71 Alavi M, Kargari M. A novel method for contrast enhancement of gray scale image based on shadowed sets. In: Proceedings of the 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS). Mashhad, Iran: IEEE, 2020. 1–7
- 72 Yang Chen-Xu, Cai Ke-Can, Zhang Hong-Yun, Miao Duo-Qian. Facial image based two-level model for gender classification. *Journal of Frontiers of Computer Science and Technology*, 2021, **15**(3): 524–532
(杨晨旭, 蔡克参, 张红云, 苗夺谦. 基于人脸图像的二阶段性别分类算法. 计算机科学与探索, 2021, **15**(3): 524–532)
- 73 Chen Chao-Fan, Zhang Hong-Yun, Cai Ke-Can, Miao Duo-Qian. Two-stage image classification method based on three-way decisions. *Pattern Recognition and Artificial Intelligence*, 2021, **34**(8): 768–776
(陈超凡, 张红云, 蔡克参, 苗夺谦. 基于三支决策的二阶段图像分类方法. 模式识别与人工智能, 2021, **34**(8): 768–776)
- 74 Cai K C, Zhang H Y, Pedrycz W, Miao D Q. SSS-Net: A shadowed-sets-based semi-supervised sample selection network for classification on noise labeled images. *Knowledge-based Systems*, 2023, **276**: Article No. 110732
- 75 Bose A, Mali K. Gradual representation of shadowed set for clustering gene expression data. *Applied Soft Computing*, 2019, **83**: Article No. 105614
- 76 Bose A, Mali K. Mammogram image enhancement using a two-threshold model of shadowed set with gradual representation of cardinality. In: Proceedings of the IEEE Region 10 Symposium (TENSymp). Mumbai, India: IEEE, 2022. 1–6
- 77 Wang H D, He S F, Pan X H, Li C D. Shadowed sets-based linguistic term modeling and its application in multi-attribute decision-making. *Symmetry*, 2018, **10**(12): Article No. 688
- 78 Li C D, Yi J Q, Wang H K, Zhang G Q, Li J Q. Interval data driven construction of shadowed sets with application to linguistic word modelling. *Information Sciences*, 2020, **507**: 503–521
- 79 Zhang Y, Zhou Y, Yao J T. Feature extraction with TF-IDF and game-theoretic shadowed sets. In: Proceedings of the 18th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems. Lisbon, Portugal: Springer, 2020. 722–733
- 80 He S F, Pan X H, Wang Y M. A shadowed set-based TODIM method and its application to large-scale group decision making. *Information Sciences*, 2021, **544**: 135–154
- 81 Wu C Y, Zhang Q H, Zhao F, Cheng Y L, Wang G Y. Three-way recommendation model based on shadowed set with uncertainty invariance. *International Journal of Approximate Reasoning*, 2021, **135**: 53–70
- 82 Liang Ji-Ye, Qian Yu-Hua, Li De-Yu, Hu Qing-Hua. Theory and method of granular computing for big data mining. *Scientia Sinica Informationis*, 2015, **45**(11): 1355–1369
(梁吉业, 钱宇华, 李德玉, 胡清华. 大数据挖掘的粒计算理论与方法. 中国科学: 信息科学, 2015, **45**(11): 1355–1369)
- 83 Xu Ji, Wang Guo-Yin, Yu Hong. Review of big data processing based on granular computing. *Chinese Journal of Computers*, 2015, **38**(8): 1497–1517
(徐计, 王国胤, 于洪. 基于粒计算的大数据处理. 计算机学报, 2015, **38**(8): 1497–1517)
- 84 Zadeh L A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets and Systems*, 1997, **90**(2): 111–127
- 85 Yao J T, Vasilakos A V, Pedrycz W. Granular computing: Perspectives and challenges. *IEEE Transactions on Cybernetics*, 2013, **43**(6): 1977–1989
- 86 Wang Guo-Yin, Li Shuai, Yang Jie. A multi-granularity cognitive computing model bidirectionally driven by knowledge and data. *Journal of Northwest University (Natural Science Edition)*, 2018, **48**(4): 488–500
(王国胤, 李帅, 杨洁. 知识与数据双向驱动的多粒度认知计算. 西北大学学报(自然科学版), 2018, **48**(4): 488–500)
- 87 Wang Guo-Yin, Zhang Qing-Hua, Ma Xi-Ao, Yang Qing-Shan. Granular computing models for knowledge uncertainty. *Journal of Software*, 2011, **22**(4): 676–694
(王国胤, 张清华, 马希骛, 杨青山. 知识不确定性问题的粒计算模型. 软件学报, 2011, **22**(4): 676–694)
- 88 Pawlak Z. Rough sets. *International Journal of Computer & Information Sciences*, 1982, **11**(5): 341–356
- 89 Zhang B, Zhang L. *Theory and Applications of Problem Solving*. New York: Elsevier Science Inc., 1992.
- 90 Li De-Yi, Meng Hai-Jun, Shi Xue-Mei. Membership clouds and membership cloud generators. *Journal of Computer Research*

- and Development, 1995, **32**(6): 15–20
(李德毅, 孟海军, 史雪梅. 隶属云和隶属云发生器. 计算机研究与发展, 1995, **32**(6): 15–20)
- 91 Yao Y Y. Three-way decision: An interpretation of rules in rough set theory. In: Proceedings of the 4th International Conference Rough Sets and Knowledge Technology. Gold Coast, Australia: Springer, 2009. 642–649
- 92 Yao Y Y. Three-way decisions with probabilistic rough sets. *Information Sciences*, 2010, **180**(3): 341–353
- 93 Wu W Z, Leung Y. Theory and applications of granular labelled partitions in multi-scale decision tables. *Information Sciences*, 2011, **181**(18): 3878–3897
- 94 Wille R. Restructuring lattice theory: An approach based on hierarchies of concepts. In: Proceedings of the NATO Advanced Study Institute. Banff, Canada: Springer, 1982. 445–470
- 95 Wang Guo-Yin, Yu Hong. Multi-granularity cognitive computing—A new model for big data intelligent computing. *Frontiers of Data & Computing*, 2019, **1**(2): 75–85
(王国胤, 于洪. 多粒度认知计算——一种大数据智能计算的新模型. 数据与计算发展前沿, 2019, **1**(2): 75–85)
- 96 Wang Guo-Yin, Fu Shun, Yang Jie, Guo Yi-Ke. A review of research on multi-granularity cognition based intelligent computing. *Chinese Journal of Computers*, 2022, **45**(6): 1161–1175
(王国胤, 傅顺, 杨洁, 郭毅可. 基于多粒度认知的智能计算研究. 计算机学报, 2022, **45**(6): 1161–1175)
- 97 Liu Dun, Liang De-Cui. Generalized three-way decisions and special three-way decisions. *Journal of Frontiers of Computer Science and Technology*, 2017, **11**(3): 502–510
(刘盾, 梁德翠. 广义三支决策与狭义三支决策. 计算机科学与探索, 2017, **11**(3): 502–510)
- 98 Yao Y Y. The geometry of three-way decision. *Applied Intelligence*, 2021, **51**(9): 6298–6325
- 99 Yao Y Y. An outline of a theory of three-way decisions. In: Proceedings of the 8th International Conference on Rough Sets and Current Trends in Computing. Chengdu, China: Springer, 2012. 1–17
- 100 Yao Y Y. Three-way decisions and cognitive computing. *Cognitive Computation*, 2016, **8**(4): 543–554
- 101 Yao Y Y. The Dao of three-way decision and three-world thinking. *International Journal of Approximate Reasoning*, 2023, **162**: Article No. 109032
- 102 Liu Dun, Li Tian-Rui, Yang Xin, Liang De-Cui. Three-way decisions: Research perspectives for rough sets and granular computing. *CAAI Transactions on Intelligent Systems*, 2019, **14**(6): 1111–1120
(刘盾, 李天瑞, 杨新, 梁德翠. 三支决策——基于粗糙集与粒计算研究视角. 智能系统学报, 2019, **14**(6): 1111–1120)
- 103 Liu D, Yang X, Li T R. Three-way decisions: Beyond rough sets and granular computing. *International Journal of Machine Learning and Cybernetics*, 2020, **11**(5): 989–1002
- 104 Liu Dun, Ye Xiao-Qing, Li Tian-Rui. Three-way decision based on the interpretability research perspective. *Journal of Northwest University (Natural Science Edition)*, 2023, **53**(6): 991–1003
(刘盾, 叶晓庆, 李天瑞. 三支决策——基于可解释研究视角. 西北大学学报(自然科学版), 2023, **53**(6): 991–1003)
- 105 Azam N, Yao J T. Multiple criteria decision analysis with game-theoretic rough sets. In: Proceedings of the 7th International Conference on Rough Sets and Knowledge Technology. Chengdu, China: Springer, 2012. 399–408
- 106 Zhang Y, Yao J T. Multi-criteria based three-way classifications with game-theoretic rough sets. In: Proceedings of the 23rd International Symposium on Foundations of Intelligent Systems. Warsaw, Poland: Springer, 2017. 550–559
- 107 Qian Y H, Liang J Y, Yao Y Y, Dang C Y. MGRS: A multi-granulation rough set. *Information Sciences*, 2010, **180**(6): 949–970
- 108 Wang Jun-Yu, Li Yan, Li Li-Hong. Summary and prospect of the research on dynamic three-way decisions. *Digital Technology & Application*, 2021, **39**(9): 174–176
(王君宇, 李言, 李丽红. 动态三支决策研究综述及展望. 数字技术与应用, 2021, **39**(9): 174–176)
- 109 Zhang Q H, Liu K X, Feng L. Research on rough equality and rough inclusion of sets in multi-granulation spaces. *Journal of Intelligent & Fuzzy Systems*, 2019, **36**(3): 2793–2806



高满 重庆邮电大学计算机科学与技术学院博士研究生. 主要研究方向为粒计算, 三支决策和不确定性信息处理.

E-mail: d200201005@stu.cqupt.edu.cn
(GAO Man Ph.D. candidate at the School of Computer Science and Technology, Chongqing University of Posts and Telecommunications. His research interest covers granular computing, three-way decision, and uncertainty information processing.)



张清华 重庆邮电大学计算机科学与技术学院教授. 主要研究方向为粒计算, 不确定性信息处理和大数据智能. 本文通信作者.

E-mail: zhangqh@cqupt.edu.cn
(ZHANG Qing-Hua Professor at the School of Computer Science and Technology, Chongqing University of Posts and Telecommunications. His research interest covers granular computing, uncertainty information processing, and big data intelligence. Corresponding author of this paper.)



王国胤 重庆师范大学重庆国家应用数学中心教授. 主要研究方向为粒计算, 认知计算和智能信息处理.

E-mail: wanggy@cqupt.edu.cn
(WANG Guo-Yin Professor at the National Center for Applied Mathematics in Chongqing, Chongqing Normal University. His research interest covers granular computing, cognitive computing, and intelligent information processing.)



姚一豫 里贾纳大学计算机科学系教授. 主要研究方向为三支决策, 粒计算和不确定性信息处理.

E-mail: Yiyu.Yao@uregina.ca
(YAO Yi-Yu Professor in the Department of Computer Science, University of Regina. His research interest covers three-way decision, granular computing, and uncertainty information processing.)