基于覆盖的粗糙集模型比较

王丽娟[1,2] 杨习贝[2] 杨静宇[1] 吴 陈[2]

(南京理工大学计算机科学与技术学院 南京 210094)[1]

(江苏科技大学计算机科学与工程学院 镇江 212003)[2]

摘 要 通过分别比较各模型的上近似、下近似以及近似精度,系统地分析了6种基于覆盖的粗糙集模型。得到3条结论:首先,在6种覆盖上近似之间,第2种覆盖上近似最大,并且除了第3种和第4种覆盖上近似之间是不可比较的之外,前5种覆盖上近似集间均有包含关系;其次,两种覆盖下近似间存在包含关系;第三,在6种模型的近似精度之间,第2种模型是最低的,而第5种模型除了和第6种模型不可比较之外其具有最高的近似精度。通过多个实例,验证了所有结论的正确性。这种对不同粗糙集模型的对比研究为深入理解这些模型提供了帮助,并且为不同应用场合模型的选择提供了参考依据。

关键词 覆盖,粗糙集模型,近似精度,模型比较

中图法分类号 TP18 **文献标识码** A

Comparison on Covering-based Rough Set Models

WANG Li-juan[1,2] YANG Xi-bei[2] YANG Jing-yu[1] WU Chen[2]

(School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)[1]

(Department of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)[2]

Abstract By means of comparing the upper approximation sets, the lower approximation sets and accuracy measures of six covering-based rough set models, the six models were systematically studied. Three conclusions were obtained. Firstly, among six covering upper approximation sets, the second one is the largest, and the first five upper approximation sets all have inclusion relationship, except the third and fourth ones. Secondly, there is inclusion relationship between two lower approximation sets. Thirdly, among the accuracy measures of six covering-based rough set models, the second one has the lowest accuracy measure, and the fifth one has the highest accuracy measure, but it has no relationship with the sixth one. By some illustrative examples, all the conclusions gain demonstrations. The comparative study on different rough set models provides a better understanding of these models and some references for model selection in different applications.

Keywords Covering, Rough set models, Accuracy measure, Model comparison

1 引言

近年来,粗糙集理论[1]作为一种有效的知识获取工具受到了人工智能工作者的关注,并成功应用在特征选择[2]、规则获取[3]、不确定推理[4]、决策评估[5]等多个领域。经典粗糙集理论借助于等价关系或划分来进行分类,这在许多应用中要求过于苛刻。为了解决这个问题,人们对等价关系做了一些很有意义的扩展,将其扩展成容差关系[6]、相似关系[7]等多种2元关系。特别是,Zakowski通过将划分扩展为覆盖,提出了基于覆盖的粗糙集模型[8]。随后,这种扩展方式得到了很好的推广。Smanta总结了现有的16种基于覆盖的粗糙集模型[9],并给出了它们的隐含格。W. Zhu系统研究了其中的6种。在文献[10]中,W. Zhu研究了由Zakowski提出的基于覆盖的粗糙集模型,他将其称为第一种覆盖粗糙集模型,讨论了该模型的性质。在文献[11]中,W. Zhu探讨了第二种覆盖粗糙集模型的相关性质,该模型是由J. A. Pomykala首次提出的[12]。第3种覆盖粗糙集模型由E. Tsang在文献[13]中首次提出。W. Zhu在文献[14]中探讨了这3种覆盖粗糙集模型的性质,并给出了3种模型之间的一些联系。随后,W. Zhu提出了第4种覆盖粗糙集模型,并研究了该模型的性质[15]。W. Zhu在文献[16]中给出了第5种覆盖粗糙集模型的定义,并利用拓扑方法对该模型做了系统分析。在文献[17]中,W. Zhu给出了第6种覆盖粗糙集模型的定义,并讨论了6种覆盖粗糙集模型中约简的方法。

到稿日期:2011-06-26 返修日期:2011-11-02 本文受国家自然科学基金(61100116),中国博士后科学基金(20100481149),江苏省自然科学基金(BK2011492),江苏省高校自然科学基金(11KJB520004)资助。

王丽娟(1981—),女,博士生,讲师,主要研究方向为智能信息处理、粒计算,E-mail:zjwanglijuan@sina.com;杨习贝(1980—),男,博士后,讲师,主要研究方向为粒计算、智能信息处理;杨静宇(1941—),男,教授,博士生导师,主要研究方向为模式识别、计算机视觉;吴 陈(1962—),男,博士,教授,硕士生导师,主要研究方向为粗糙集理论。

但是，纵观多种基于覆盖的粗糙集模型，发现这些模型提出时作者都没有明确给出提出的原因、背景和应用价值，而且很少涉及这些模型之间优劣的比较。这些基于覆盖的粗糙集模型之间有什么联系和区别？哪种覆盖粗糙集的应用范围最广？哪种模型具有最高近似精度？这些问题是研究覆盖粗糙集模型时必然要解决的问题。

所以，本文以 W. Zhu 系统研究的 6 种基于覆盖的粗糙集模型为例，从粗糙集最根本的上、下近似出发，来比较这 6 种模型，希望能找出这 6 种模型之间的区别和联系，并且找出各模型的应用范围，深入理解和把握这些模型，为选择不同应用场合模型提供帮助。

本文第 2 节介绍 6 种基于覆盖的粗糙集模型；第 3 节首先探讨 6 种覆盖上近似之间的关系，随后比较两种覆盖下近似（第 1—5 共 5 种覆盖粗糙集模型具有相同的下近似集定义），接着给出 6 种覆盖粗糙集模型间近似精度对比表，直观地描述 6 种模型不同的应用范围，并通过多个实例进行验证；最后总结全文。

2 基于覆盖的粗糙集模型

本节主要介绍 6 种基于覆盖的粗糙集模型中的基本概念。

定义 1[17] 令 U 为论域，C 是一簇 U 的子集，并且 C 中不含空集。如果 $\bigcup C=U$，则称 C 是 U 的一个覆盖。如果 P 是 U 的一个覆盖并且 P 中任意两个元素都相交为空，则称 P 为 U 的一个划分。显然，U 的划分必为 U 的覆盖，所以覆盖是划分的一般形式，而划分是一种最简单的覆盖。

定义 2[17] 令 C 是 U 的一个覆盖，$x\in U$，x 的最小描述 $Md_C(x)$ 定义为

$$Md_C(x)=\{K\in C\mid x\in K\wedge(\forall S\in C\wedge x\in S\wedge S\subseteq K\Rightarrow K=S)\}$$

定义 3[17] 令 C 是 U 的一个覆盖，$x\in U$，x 的邻域 $Neighbor_C(x)$ 定义为

$$Neighbor_C(x)=\bigcap\{K\mid x\in K\in C\}$$

定义 4[17] 令 C 是 U 的一个覆盖，$\forall X\subseteq U$，第 1 种基于覆盖 C 的覆盖下近似 $CL_C(X)$ 和覆盖上近似 $FH_C(X)$ 分别定义如下：

$$CL_C(X)=\bigcup\{K\in C\mid K\subseteq X\}$$

$$FH_C(X)=CL_C(X)\bigcup\{Md_C(x)\mid x\in X-CL_C(X)\}$$

定义 5[17] $\forall X\subseteq U$，第 2，3，4，5 种覆盖上近似 $SH_C(X)$，$TH_C(X)$，$RH_C(X)$，$IH_C(X)$ 分别定义如下：

$$SH_C(X)=\bigcup\{K\mid K\in C,K\cap X\neq\emptyset\}$$

$$TH_C(X)=\bigcup\{Md_C(x)\mid x\in X\}$$

$$RH_C(X)=CL_C(X)\bigcup\{K\mid K\cap(X-CL_C(X))\neq\emptyset\}$$

$$IH_C(X)=CL_C(X)\bigcup\{Neighbor_C(x)\mid x\in X-CL_C(X)\}$$

W. Zhu 在文献[11,14-17]中也给出了第 2，3，4，5 种基于覆盖 C 的覆盖下近似的概念，但是它们均等价于 $CL_C(X)$。所以，文中仅仅认为 $CL_C(X)$ 是第 1 种基于覆盖 C 的覆盖下近似。

定义 6[17] $\forall X\subseteq U$，第 6 种基于覆盖 C 的覆盖下近似 $XL_C(X)$ 和上近似 $XH_C(X)$ 分别定义为

$$XL_C(X)=\{x\mid Neighbor_C(x)\subseteq X\}$$

$$XH_C(X)=\{x\}\mid Neighbor_C(x)\cap X\neq\emptyset\}$$

当没有出现歧义时，以上定义中均可省去下标 C，故在下文的讨论中略去。

3 6 种基于覆盖的粗糙集模型间的关系

对任意一个论域 U 上的覆盖 C，W. Zhu 在文献[20]中给出了这 6 种基于覆盖的粗糙集模型，但是并未系统地探讨它们之间的关系，因此本节重点探讨这 6 种模型之间的关系。分 3 步骤进行：首先探讨 6 种覆盖上近似之间的关系，接着探讨两种覆盖下近似之间的关系，最后比较 6 种覆盖粗糙集模型的近似精度的大小。

3.1 6 种覆盖上近似集间的关系

在文献[14]中，W. Zhu 对第 1，2，3 种模型上近似的大小做了比较，得出了以下结论。

命题 1[14] 令 C 是 U 的一个覆盖，对 $\forall X\subseteq U$，满足 $FH(X)\subseteq TH(X)\subseteq SH(X)$。

例 1 给定论域 $U=\{a,b,c,d,e,f\}$，覆盖 $C=\{K_1,K_2,\cdots,K_8\}$，其中 $K_1=\{a\}$，$K_2=\{b\}$，$K_3=\{c\}$，$K_4=\{d,e\}$，$K_5=\{e\}$，$K_6=\{f\}$，$K_7=\{b,f\}$，$K_8=\{e,f\}$。可以得出 $Md(a)=\{\{a\}\}$，$Md(b)=\{\{b\}\}$，$Md(c)=\{\{c\}\}$，$Md(d)=\{\{d,e\}\}$，$Md(e)=\{\{e\}\}$，$Md(f)=\{\{f\}\}$。$Neighbor(a)=\{a\}$，$Neighbor(b)=\{b\}$，$Neighbor(c)=\{c\}$，$Neighbor(d)=\{d,e\}$，$Neighbor(e)=\{e\}$，$Neighbor(f)=\{f\}$。

若假设 $X=\{a,b,c,e\}$，那么 $FH(X)=TH(X)=\{a,b,c,e\}$，$SH(X)=\{a,b,c,d,e,f\}$。所以，命题 1 中的 $FH(X)\subseteq TH(X)\subseteq SH(X)$ 成立。

通过比较这 6 种基于覆盖的上近似之间的关系，同样可以得出其它的一些结论。

命题 2 $IH(X)\subseteq FH(X)$。

证明：由定义 3 知，$Neighbor_C(x)=\bigcap\{K\mid x\in K\in C\}$，因此 $\forall x\in U$，均有 $Neighbor(x)=\bigcap Md(x)$。那么，$\{Neighbor(x)\mid x\in X-CL(X)\}\subseteq\{Md(x)\mid x\in X-CL(X)\}$，$CL(X)\bigcup\{Neighbor(x)\mid x\in X-CL(X)\}\subseteq CL(X)\bigcup\{Md(x)\mid x\in X-CL(X)\}$，即 $IH(X)\subseteq FH(X)$，证毕。

例 2（接例 1） 假设 $X=\{b,d,f\}$，那么 $FH(X)=\{b,d,e,f\}$，$IH(X)=\{b,d,e,f\}$，所以满足 $IH(X)\subseteq FH(X)$。

通过命题 2 及命题 3，显然可以得到如下结论。

推论 1 $IH(X)\subseteq FH(X)\subseteq TH(X)\subseteq SH(X)$。

命题 3 $FH(X)\subseteq RH(X)$。

证明：$FH_C(X)=CL(X)\bigcup\{Md(x)\mid x\in X-CL(X)\}$，$RH_C(X)=CL(X)\bigcup\{K\mid K\cap(X-CL(X))\neq\emptyset\}$。

因此仅需证明 $\bigcup\{Md(x)\mid x\in X-CL(X)\}\subseteq\bigcup\{K\mid K\cap(X-CL(X))\neq\emptyset\}$。

当 $x\in K\cap(X-CL(X))$ 时，$x\in K\wedge x\in(X-CL(X))$。

既然 $Md(x)$ 是包含 x 的最小的集合，那么

$\bigcup\{Md(x)\mid x\in X-CL(X)\}\subseteq\bigcup\{K\mid x\in K\wedge x\in(X-CL(X))\}\subseteq\bigcup\{K\mid K\cap(X-CL(X))\neq\emptyset\}$

因此，$FH(X)\subseteq RH(X)$。证毕。

例 3（接例 1） 假设 $X=\{a,b,c,e\}$，那么 $FH(X)=\{a,b,c,e\}$，$RH(X)=\{a,b,c,e\}$，所以满足 $FH(X)\subseteq RH(X)$。

命题 4 $RH(X)\subseteq SH(X)$。

证明：$RH(X)=CL(X)\bigcup\{K\mid K\cap(X-CL(X))\neq\emptyset\}\subseteq CL(X)\bigcup\{K\mid K\cap X\neq\emptyset\}=\bigcup\{K\mid K\cap X\neq\emptyset\}=SH(X)$。

因此，$RH(X)\subseteq SH(X)$。证毕。

例 4(接例 1)　假设 $X=\{a,c,d,e,f\}$，那么 $RH(X)=\{a,c,d,e,f\}$，$SH(X)=\{a,b,c,d,e,f\}$，所以满足 $RH(X)\subseteq SH(X)$。

由命题 2—命题 4，很容易得到如下结论：

推论 2　$IH(X)\subseteq FH(X)\subseteq RH(X)\subseteq SH(X)$。

命题 5　$TH(X)$ 和 $RH(X)$ 是不可比较的。

例 5　(1)给定论域 $U=\{a,b,c,d,e\}$，覆盖 $C=\{\{a\},\{b\},\{c,d\},\{c,d,e\}\}$。

当 $X=\{a,b,c\}$ 时，$TH(X)=\{a,b,c,d\}$，$RH(X)=\{a,b,c,d,e\}$，可得 $TH(X)\subset RH(X)$；

当 $X=\{b,c,d\}$ 时，$TH(X)=\{b,c,d\}$，$RH(X)=\{b,c,d\}$，可得 $TH(X)=RH(X)$。

(2)给定论域 $U=\{a,b,c,d\}$，覆盖 $C=\{\{a\},\{a,b\},\{c\},\{b,d\}\}$。

当 $X=\{a,b\}$ 时，$TH(X)=\{a,b,d\}$，$RH(X)=\{a,b\}$，可得 $TH(X)\supset RH(X)$；

当 $X=\{c,d\}$ 时，$TH(X)=\{b,c,d\}$，$RH(X)=\{b,c,d\}$，可得 $TH(X)=RH(X)$。

可见，在情况(1)，$TH(X)\subseteq RH(X)$；在情况(2)，$TH(X)\supseteq RH(X)$。因此可以说明，在 $TH(X)$ 和 $RH(X)$ 之间是不可以比较的。

下面探讨第 6 种覆盖上近似集和前 5 种覆盖上近似集之间的关系。

命题 6　$XH(X)\subseteq SH(X)$。

证明：由定义 3 $Neighbor(x)=\bigcap\{K|x\in K\in C\}\subseteq K$，那么 $Neighbor(x)\bigcap X\neq\emptyset\Rightarrow K\bigcap X\neq\emptyset$。

因为 $XH(X)=\{x|Neighbor(x)\bigcap X\neq\emptyset\}$，$SH(X)=\bigcup\{K|K\in C,K\bigcap X\neq\emptyset\}$，

所以，$XH(X)\subseteq SH(X)$。证毕。

例 6(接例 1)　若假设 $X=\{a,b,c,e\}$，那么 $XH(X)=\{a,b,c,d,e\}$，$SH(X)=\{a,b,c,d,e,f\}$，所以满足命题 6 中的 $XH(X)\subseteq SH(X)$。

命题 7　$XH(X)$ 和 $FH(X)$，$TH(X)$，$RH(X)$ 或者 $IH(X)$ 都是不可以比较的。

例 7　给定论域 $U=\{a,b,c,d,e,f\}$ 以及覆盖 $C=\{\{a\},\{b\},\{c\},\{d\},\{e\},\{b,f\},\{b,d\},\{c,d\}\}$。

(1)若 $X=\{b,c,e,f\}$ 时，$FH(X)=TH(X)=RH(X)=IH(X)=XH(X)=\{b,c,e,f\}$。

(2)若 $X=\{a,c,d,e,f\}$ 时，$FH(X)=TH(X)=RH(X)=IH(X)=\{a,b,c,d,e,f\}$，$XH(X)=\{a,c,d,e,f\}$。可见，$XH(X)\subseteq FH(X)=TH(X)=RH(X)=IH(X)$。

(3)若 $X=\{b,d,e\}$ 时，$FH(X)=TH(X)=RH(X)=IH(X)=\{b,d,e\}$，$XH(X)=\{a,b,d,e,f\}$。可见，$FH(X)=TH(X)=RH(X)=IH(X)\subseteq XH(X)$。

故，通过例 7 中的 3 种情况，可以验证命题 7 的正确性。

3.2　两种覆盖下近似集间的关系

以上 7 个命题，揭示了 6 种覆盖上近似集之间的关系。同样可以得到两种覆盖下近似集之间的关系。

命题 8　$CL(X)\subseteq XL(X)$。

证明：由定义 4 知，$CL(X)=\bigcup\{K\in C|K\subseteq X\}$，对 $\forall x\in CL(X)$，假设覆盖 $C=\{K_1,K_2,\cdots,K_m\}$ 并且 $x\in K_i\wedge x\in K_j\wedge x\in K_l\wedge\cdots$，那么 $K_i\subseteq X\wedge K_j\subseteq X\wedge K_l\subseteq X\wedge\cdots$，从而 $K_i\bigcap K_j\bigcap K_l\bigcap\cdots\subseteq X$。

因此，$\bigcap\{K|x\in K\in C\}\subseteq X$，即 $Neighbor(x)\subseteq X$，$x\in XL(X)$。

所以，$CL(X)\subseteq XL(X)$ 成立。证毕。

例 8(接例 1)　若 $X=\{a,b,c,e\}$，那么可以得出 $CL(X)=\{a,b,c,e\}$，$XL(X)=\{a,b,c,e\}$，所以满足命题 8 中的 $CL(X)\subseteq XL(X)$。

3.3　6 种覆盖粗糙集模型近似精度的关系

在粗糙集理论中，近似精度是一个衡量粗糙集模型适应范围的重要度量。下面将近似精度的概念引入基于覆盖的粗糙集模型中，给出 6 种覆盖粗糙集模型近似精度的概念。

定义 7　令 C 是 U 的一个覆盖，$\forall X\subseteq U$ 且 $X\neq\emptyset$，第 1 种覆盖粗糙集模型的近似精度 $\alpha_1(X)$ 定义为

$$\alpha_1(X)=cardCL(X)/cardFH(X)$$

类似地，可以给出第 2,3,4,5,6 种覆盖粗糙集模型的近似精度 $\alpha_2(X)$，$\alpha_3(X)$，$\alpha_4(X)$，$\alpha_5(X)$ 和 $\alpha_6(X)$ 的定义：

$$\alpha_2(X)=cardCL(X)/cardSH(X)$$

$$\alpha_3(X)=cardCL(X)/cardTH(X)$$

$$\alpha_4(X)=cardCL(X)/cardRH(X)$$

$$\alpha_5(X)=cardCL(X)/cardIH(X)$$

$$\alpha_6(X)=cardXL(X)/cardXH(X)$$

通过计算，可以得到以下的大小关系：

①因 $IH(X)\subseteq FH(X)\subseteq TH(X)\subseteq SH(X)$，且这几个模型具有相同的下近似集，故 $\alpha_5(X)\geqslant\alpha_1(X)\geqslant\alpha_3(X)\geqslant\alpha_2(X)$。

②因 $IH(X)\subseteq FH(X)\subseteq RH(X)\subseteq SH(X)$，且这几个模型具有相同的下近似集，故 $\alpha_5(X)\geqslant\alpha_1(X)\geqslant\alpha_4(X)\geqslant\alpha_2(X)$。

③因 $XH(X)\subseteq SH(X)$，$CL(X)\subseteq XL(X)$，那么 $cardHX(X)\leqslant cardSX(X)$，$cardCL(X)\leqslant cardXL(X)$，故 $\alpha_6(X)\geqslant\alpha_2(X)$。

④因 $TH(X)$ 和 $RH(X)$ 之间不可比较，故不能得出 $\alpha_3(X)$ 和 $\alpha_4(X)$ 之间的大小关系，记为“无”。

⑤因 $XH(X)$ 与 $FH(X)$，$TH(X)$，$RH(X)$ 或者 $IH(X)$ 都是不可以比较的，故不能得出 $\alpha_6(X)$ 与 $\alpha_1(X)$，$\alpha_3(X)$，$\alpha_4(X)$ 或者 $\alpha_5(X)$ 之间的大小关系，记为“无”。

从而可以得到 6 种覆盖粗糙集模型近似精度间的大小关系，如表 1 所列。

表 1　近似精度间关系表

关系	$\alpha_1(X)$	$\alpha_2(X)$	$\alpha_3(X)$	$\alpha_4(X)$	$\alpha_5(X)$	$\alpha_6(X)$
$\alpha_1(X)$	=	≥	≥	≥	≤	无
$\alpha_2(X)$	≤	≥	≤	≤	≤	≤
$\alpha_3(X)$	≤	≥	=	无	≤	无
$\alpha_4(X)$	≤	≥	无	=	≤	无
$\alpha_5(X)$	≥	≥	≥	≥	=	无
$\alpha_6(X)$	无	≥	无	无	无	=

由表 1，可以得到如下一些结论：

① 第 2 种覆盖粗糙集模型具有最低的近似精度，可见该模型仅适用于数据中含有大量噪声、对精度要求较低的应用；

② 除第 6 种覆盖粗糙集模型外，第 5 种模型是剩下 5 种覆盖粗糙集模型中精度最高的，该模型可以应用于精度要求较高的场合；

③ 在对于精度有一定要求但不是特别严格的情况下，第

（下转第 236 页）

是求一个取样点集 S，要求 S 包含客户点集 C 的每个最优划分子集中至少一个点，根据 S 再进一步确定一个设施子集 F_s，利用反向贪心算法从 F_s 中找出 k 个点服务 C。本文在求解取样点集时基于均衡限制参数 α，算法近似性能比的期望值依赖于 α。α 越小，则 $|S|$ 越大，算法近似性能比期望值越高；α 越大，则 $|S|$ 越小，算法近似性能比的期望值越小，因此该算法适用于各划分子集大小相差不太大的情况。

参考文献

[1] Balinski M L. On finding integer solutions to linear programs [C]// Proceedings of IBM scientific computing symposium on combinatorial problems. 1966:225-248

[2] Hochbaum D S. Heuristics for the fixed cost median problem [J]. Mathematical Programming, 1982, 22: 148-162

[3] Lin J H, Vitter J S. ε-approximation s with minimum constraint violation[C]// Proceedings of stoc'92. 1992: 771-782

[4] Arora S, Raghavan P, Rao S. Approximation schemes for Euclidean k-median and related problems [C] // Proceedings of stoc' 98. 1998: 106-113

[5] Charikar M, et al. A constant approximation algorithm for the k-median problem[C]// Proceedings of stoc'1999. Atlanta GA USA, 1999

[6] Jain K, Vazirani V V. Primal-dual approximation algorithms for metric facility location and k-median problems[C]// Proceedings of focs'99. 1999: 2-13

[7] Charikar M, Guha S. Improved combinatorial algorithms for the facility location and k-median problems [C] // Proceedings of focs'99. 1999: 1-10

[8] Arya V, et al. Local search heurictics for k-median and facility location problems [C] // Proceedings of stoc' 2001. Hersonissons, Crete, Greece, 2001

[9] 潘锐，朱大铭，等. k-median 近似计算复杂度与局部搜索近似算法分析[J]. 软件学报，2005，16(3)：392-399

[10] 肖进杰，范辉，等. 贪心算法求解 k-median 问题[J]. 计算机工程与应用，2006，42(3)：57-58

[11] 肖进杰，谢青松. k-median 问题贪心近似算法的分析与实验[J]. 计算机工程，2008，34(22)：213-215

[12] Chrobak M, Kenyon C, Young N. The reverse greedy algorithm for the metric k-median problem [J]. Information Processing Letters, 2006, 97: 68-72

[13] Motawani R, Raghavan P. Randomized Algorithms. Cambridge University Press[M]. Cambridge, UK, 1995

[14] Chen K. A constant factor approximation algorithm for k-median clustering with outliers[C] // Proceedings of the 19th Annual ACM-SIAMSymposium on Discrete Algorithms. 2008

(上接第 231 页)

1,3,4,6 这 4 种模型都可以适用。

结束语 本文系统地分析、比较了 6 种主要基于覆盖的粗糙集模型，找到了各模型之间的区别和联系。各模型之间要么明确可比，要么明确不可比。可比的给出了定理证明，不可比的举出了反例。通过比较多种模型之间关系，可以加深对这些模型的理解，使得读者方便地在不同应用中根据需要选择不同的模型。

今后将在本文工作的基础上，对于可比较的模型研究其相等的充要条件，对于不可比较的模型研究其可比必须满足的条件；也可以从最小覆盖粗糙集模型和最大覆盖粗糙集模型的角度出发，构造出满足需要的覆盖粗糙集模型。

参考文献

[1] Pawlak Z. Rough sets[J]. International Journal of Computer and Information Science, 1982(11): 341-356

[2] 邓维斌，王国胤，洪智勇. 基于粗糙集的加权朴素贝叶斯邮件过滤方法[J]. 计算机科学，2011，38(2)：218-221

[3] Qian Y H, Liang J Y, Dang C Y. Converse approximation and rule extraction from decision tables in rough set theory [J]. Computers & Mathematics with Applications, 2008, 55 (8): 1754-1765

[4] Shi Z H, Gong Z T. The further investigation of covering-based rough sets: Uncertainty characterization, similarity measure and generalized models [J]. Information Science, 2010, 180 (19): 3745-3763

[5] Dembczyński K, Greco S, Słowiński R. Rough set approach to multiple criteria classification with imprecise evaluations and assignments [J]. European Journal of Operation Research, 2009, 198(2): 626-636

[6] Kryszkiewicz M. Rough set approach to incomplete information systems [J]. Information Science, 1998, 112: 39-49

[7] Stefanowski J, Tsoukias A. On the extention of rough sets under incomplete information [C] // 7th International Workshop on New Directions in Rough Sets, Data Mining, and Granular-soft Computing. 1998: 422-450

[8] Zakowski W. Approximations in the space(u, π)[J]. Demonstration Mathematics, 1983, 16: 761-769

[9] Samanta P, Chakraborty M K. Covering based approaches to rough sets and implication lattices [C]// RSFD GrC. 2009: 127-134

[10] Zhu W, Wang F Y. Properties of the First Type of Covering-based Rough Sets [C]// Proceedings of DM Workshop 06 (ICDM 06). Hong Kong, China, December 2006: 407-411

[11] Zhu W. Properties of the second type of covering-based rough sets [C]// Workshop Proceedings of GrC&BI 06 (IEEE WI 06). Hong Kong, China, December 2006: 494-497

[12] Pomykala J A. Approximation operations in approximation space [J]. Bulletin of the Polish Academy of Sciences, 1987, 35 (9/10): 653-662

[13] Tsang E, Cheng D, Lee J, et al. On the upper approximations of covering generalized rough sets [C]// Proceedings of the 3rd International Conference Machine Learning and Cybernetics. 2004: 4200-4203

[14] Zhu W, Wang F Y. On three types of covering rough sets[J]. IEEE Transactions on Knowledge and Data Engineering, 2007, 19(8): 1131-1144

[15] Zhu W, Wang F Y. A new type of covering rough sets [C]// IEEE IS 2006. London, September 2006: 444-449

[16] Zhu W. Topological approaches to covering rough sets[J]. Information Science, 2007, 177 (6): 1499-1508

[17] Zhu W. Relationship between generalized rough sets based on binary relation and covering [J]. Information Science, 2009, 179 (1): 210-225