

新的软集合约简方法在购房决策中的应用

袁红彬, 刘文奇, 石梦婷

(昆明理工大学 理学院, 云南 昆明 650500)

摘要: 为了对软集合各参数进行约简的同时更大程度的保留集合性质, 使约简方法更广泛地运用于决策分析与决策系统, 文中在软集合理论的数据挖掘方法基础上, 从决策系统各因素的重要程度出发, 提出了一种新的约简方法, 得到简化的参数集。文中首先分析了以往软集合的决策规则, 然后定义了参数的影响因子, 并提出了更能反映元素决策等级的加权选择价值, 从而构造了一种新的不可区分关系和决策优先级。最后将该约简算法运用到购房决策系统中, 为决策者提供有价值的信息, 并得到了可行的约简结果。

关键词: 软集合; 参数约简; 等价关系; 数据挖掘

中图分类号: TP301

文献标识码: A

文章编号: 1673-629X(2013)01-0246-03

doi: 10.3969/j.issn.1673-629X.2013.01.060

Application of a New Approach of Soft Set Reduction in House Purchase Decision System

YUAN Hong-bin, LIU Wen-qi, SHI Meng-ting

(School of Science, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: In order to reserve the properties of soft set while reducing the parameters, a new approach of parameter reduction is proposed by using the data mining based on soft set theory, according to the importance degree of each factor in decision making system, which is more widely used in decision analysis systems and obtains a reduction parameter set. Firstly, the former decision rule is analyzed. Next, the impact factor of parameter is defined as well as the weighted choice value that denotes the decision making level for each element. Then, a new indiscernibility relation and decision making priority are presented. Finally, the algorithm for reduction is applied in a house purchase system to show some valuable information and give a feasible result.

Key words: soft set; parameter reduction; equivalence relation; data mining

0 引言

软集合理论^[1]是 Molodtsov 提出的一种新的处理不确定性问题的数学手段, Maji 曾探讨了软集与信息系统之间的关系^[2], 并提出了参数约简^[3]这一概念, Ali^[4]推广了软集合的模糊性, Chen^[5]则将其运用到实际决策中, 马振明和沈伟还将软集与粗糙集相结合, 得到了软粗糙集^[6]的理论和软决策表^[7]的概念。为了对冗余的信息系统进行约简, 学者们还提出了属性约简^[8]、文本分类^[9]、粗糙近似^[10]、多属性决策^[11]等多种处理手段。以商品房选购系统为例, 随着购房者的增多, 房屋特性的信息量巨大, 消费者对住房的选择有

自己的偏好, 如何从庞大数据中提取规则和信息就成了使得研究的问题。这需要对所有的信息进行整合, 筛选出数量较少的, 并在一定程度上能保证房屋挑选规则的若干特性, 这个过程就是参数约简, 约简后的参数属性集可以更加精简的描述购房系统。

1 软集合理论基本概念

定义 1^[2] 设 E 是论域 U 上的一个参数集, $P(U)$ 表示 U 的幂集合且 $A \subseteq E$ 。若 F 表示映射 $F: A \rightarrow P(U)$, 二元组 (F, A) 称为 U 上的一个软集合。因此, 软集合就是一个参数子集族 $(F, A) = \{p_1 = v_1, p_2 = v_2, \dots, p_n = v_n\}$, 其中 p 为参数值, v 为满足参数的集合。

2 软集合的参数与决策分析

2.1 软集合的二元关系表

设论域 $U = \{h_1, \dots, h_6\}$ 表示可供选择的六套住

收稿日期: 2012-05-05; 修回日期: 2012-08-09

基金项目: 科技部科技型中小企业技术创新基金项目 (11C26215 305906)

作者简介: 袁红彬 (1988-), 男, 硕士研究生, 研究方向为数据挖掘、粗糙集理论、模糊数学; 刘文奇, 教授, 硕士生导师, 研究方向为模糊数学、决策分析。

房,参数集 $E = \{e_1, \dots, e_5\} = \{\text{美观}, \text{价格便宜}, \text{实木结构}, \text{周边绿化良好}, \text{精装修}\}$ 表示购房者关心的几个因素。一个软集合由 $(F, E) = \{\text{价格便宜} = \{h_1, h_2\}, \text{美观} = \{h_1, h_2, h_3\}, \text{实木结构} = \{h_3, h_5\}, \text{周边绿化良好} = \{h_6\}, \text{精装修} = \{h_3, h_4\}\}$ 给定,可以看做描述购房系统中“住房的吸引力”。

软集合 (F, E) 常用二元关系表^[12]表示。关系表的第一行为各参数,表中的对应项为 h_{ij} , 如果 $h_i \in F(e_j)$, 则 $h_{ij} = 1$, 否则 $h_{ij} = 0$, 同时在表尾增加一列选择价值 $c_i = \sum_j h_{ij}$, 选择价值越大表示该对象越符合决策者的预期要求。这时的二元关系表如表 1 所示。

表 1 软集合的二元关系表

U	美观 (e_1)	价格便 宜 (e_2)	实木结 构 (e_3)	周边绿化 良好 (e_4)	精装修 (e_5)	选择价 值 (c_i)
h_1	1	1	0	0	0	2
h_2	1	1	0	0	0	2
h_3	0	1	1	0	1	3
h_4	0	0	0	0	1	5
h_5	0	0	1	0	0	3
h_6	0	0	0	1	0	1

2.2 选择优先级及软划分

由于系统中的参数并不是同等重要的,因此需要在保证软集合某些特征的前提下,删除不重要的参数。这里的特征应包括决策能力及分类能力。以往软集合的不同约简很可能改变其分类能力,因此需要定义一种新的分类,使得约简的同时更好地保证系统特征不变。

定义 2 设 (F, E) 是 U 上的软集合,按选择价值划分论域得到 $U/c_i = \{X_1, \dots, X_n\}$ 。对任意 $X_1, X_2 \in U/c_i$, 若 X_1 中元素的选择价值均大于等于 X_2 , 称 X_1 比 X_2 优先,记为 $X_1 > X_2$ 。并将所有等价类的全排列(如 $X_1 > X_2 > \dots > X_n$)称为软集合的选择优先级。

定义 3 设 (F, E) 是 U 上的软集合,在定义划分之前首先定义任意参数 $e_j \in E$ 的影响因子为 $d_j = \sum_i h_{ij}$, 定义任意元素 $h_i \in U$ 加权选择价值为 $\tau_i = \sum_j h_{ij} d_j$ 。若令等价关系 $E_R = \{c_i, \tau_i\}$, 则不可区分关系 $\text{Ind}(E_R)$ 对 U 的划分 U/E_R 称为 (F, E) 的软划分。

3 参数约简

3.1 软集合参数约简的基本思想及算法实现

设 $\text{Ind}(E_R)$ 是 (F, E) 的软划分,对任意 $e \in E$, 如果 $(F, E - \{e\})$ 的软划分不变,则 e 为 E 中不必要的; 否则, e 为 E 中必要的。对于任意的 $Q \subseteq E$, 如果每个 $e \in Q$ 都是 E 中必要的,且 $\text{Ind}(E_R) = \text{Ind}(Q_R)$, 则称 Q 为 E 的一个约简, (F, Q) 称为约简软集。

参数约简需要逐个分析参数的必要性,若在二元关系表上加一行一列分别表示影响因和加权选择价值,可直观地看出软划分和选择优先级,这样的表称之为二元决策表。若 Q 为 E 的一个约简,显然约简后的软集 (F, Q) 与 (F, E) 有相同的选择优先级。

设 (F, E) 是 U 上的软集合,要找到约简软集 (F, Q) , 可以通过下面几个步骤:

1. 输入软集合 (F, E) 。
2. 生成初始二元决策表,确定初始软划分 U/E_R 和选择优先级。
3. 删除 E 中的一个参数 e ,生成新的二元决策表,并计算新的软划分 $U/(E - e)_R$ 。
4. 比较软划分是否与初始软划分相同;如果相同,则重复步骤 3、4, 如果不同,则保留该参数并重复步骤 3、4;直到遍历了 E 中的所有参数。
5. 得到新参数集 E 即为约简 Q 。

具体算法实现如下:
S1 set $H = \begin{matrix} \text{ } & \text{ } & \text{ } & \text{ } & \text{ } & \text{ } \\ \text{ } & \text{ } & \text{ } & \text{ } & \text{ } & \text{ } \end{matrix} m \times n$;
for $i = 1, 2, \dots, m$; for $j = 1, 2, \dots, n$; if $(h_i \in F(e_j))$
 $h_{ij} = 1$; else $h_{ij} = 0$

S2 define function partition {array[]}
set $X_k = A = \emptyset, \text{par} = \{\emptyset\}$;
for $k = 1, 2, \dots, \max$; for $j = 2, 3, \dots, m$; if $(c_1 = c_j \tau_1 = \tau_i)$
{ $X_k = X_k \cup \{h_i\}$; $A = A \setminus X_k$; $\text{par} = \text{par} \cup X_k$ } return
 par ;
S3 for $i = 1, 2, \dots, m$; for $j = 1, 2, \dots, n$; $\text{par} = \text{partition}\{E\}$;
S4 for $j = 1, 2, \dots, n$;
{ $\text{temp} = \text{partition}\{E \setminus \{e_j\}\}$; if $(\text{temp} = \text{par})$ $E = E \setminus \{e_j\}$; else continue}
S5 $Q = E$

3.2 参数约简的结果

按上述步骤对文中第二部分的例子进行约简,首先生成 (F, E) 的初始二元决策表(见表 2),可以得到初始软划分形成的等价类 $U/E_R = \{(\{h_1, h_2\}, \{h_3\}, \{h_4, h_5\}, \{h_6\})\}$, 初始选择优先级 $\{h_3\} > \{h_1, h_2\} > \{h_4, h_5\} > \{h_6\}$ 。再检验参数“美观”是否必要,删除 e_1 并刷新表格可得 $U/(E - \{e_1\})_R = \{(\{h_1, h_2\}, \{h_3\}, \{h_4, h_5\}, \{h_6\})\} = U/E_R$, 因此 e_1 是 E 中不必要的。

依次检验所有参数最终得到最简二元决策表(见表 3),即 e_2, e_3, e_5 是必要的, e_1, e_4 是不必要的, $Q = \{e_2, e_3, e_5\}$ 是 E 的一个约简,可以验证软划分和选择优先级均不变。也就是说“价格便宜”、“实木结构”、“精装修”是所有购房者考虑的因素中不可缺少的三项,而“美观”、“周边绿化良好”则是可以删除的因素,

这样的约简就足够将所有住房划分等级,并保证 house 3 的吸引力最大,为最优选择,house 1 和 house 2 的吸引力其次,为次优选择,house 4 和 house 5 的吸引力再次,而 house 6 的吸引力最差,不建议选择。

表 2 软集合的初始二元决策表

U	美观 (e_1)	价格便宜 (e_2)	实木结构 (e_3)	周边绿化良 好 (e_4)	精装修 (e_5)	选择价值 (c_i)	加权选择 价值 (τ_i)
h_1	1	1	0	0	0	2	5
h_2	1	1	0	0	0	2	5
h_3	0	1	1	0	1	3	7
h_4	0	0	0	0	1	1	2
h_5	0	0	1	0	0	1	2
h_6	0	0	0	1	0	1	1
d_j	2	3	2	1	2		

表 3 最简的二元决策表

U	价格便宜 (e_2)	实木结构 (e_3)	精装修 (e_5)	选择价值 (c_i)	加权选择价值 (τ_i)
h_1	1	0	0	1	3
h_2	1	0	0	1	3
h_3	1	1	1	3	7
h_4	0	0	1	1	2
h_5	0	1	0	1	2
h_6	0	0	0	0	0
d_j	3	2	2		

4 结束语

文章通过定义影响因子和加权选择价值,直观地刻画了购房决策系统中每个元素与参数之间的关系。在此基础上进行约简,既保证了理论上的软集合分类能力不变,又保证了实际的待购房屋选择优先级不变。

通过实例可以验证,该约简方法可以讲购房系统的数据进行有效约简,为购房者提供更简洁的房屋信息,有一定的实用价值。同时,文中所述方法具有一般性,也适用于其他商品采购系统的属性约简。

参考文献:

[1] Molodtsov D. Soft set theory—first results[J]. Computers and Mathematics with Applications,1999,37(4/5):19–31.

[2] Maji P K,Biswas B. Soft Set Theory[J]. Computers and Mathematics with Applications,2003,45(4/5):555–562.

[3] Maji P K,Roy A R. An Application of Soft Sets in a Decision Making Problem[J]. Computers and Mathematics with Applications,2002,44(8/9):1077–1083.

[4] Ali M I. A note on soft sets,rough soft sets and fuzzy soft sets [J]. Applied Soft Computing,2011(11):3329–3332.

[5] Chen D G. The Parameterization Reduction of Soft Sets and Its Applications[J]. Computers and Mathematics with Applications,2005,49(5/6):757–763.

[6] 马振明. Pawlak 近似空间中软集的软上(下)近似[J]. 计算机工程与应用,2011,47(18):60–61.

[7] 沈 玮,赵佳宝. 一种新的启发式粗集决策表属性约简算法[J]. 计算机技术与发展,2012,22(10):16–20.

[8] 汪 蓉,陈黎伟,吴 涛. 一种改进的属性约简算法[J]. 计算机技术与发展,2008,18(2):146–148.

[9] 洪智勇,秦克云. 基于模糊软集合理论的文本分类方法[J]. 计算机工程,2010,36(13):90–92.

[10] 张文修,吴伟志. 粗糙集理论与方法[M]. 北京:科学出版社,2001.

[11] 付 清,张小红. 一种应用软粗糙模糊集的多属性决策方法[J]. 计算机工程与应用,2011,47(30):39–43.

[12] 胡逢彬,桂现才. 决策表属性约简的相对信息量表示[J]. 计算机技术与发展,2006,16(7):39–41.

(上接第 245 页)

统的设计与实现[J]. 计算机技术与发展,2008,18(6):236–239.

[4] 李先军,刘 波,余 丹,等. 一种基于 AJAX 技术的 B/S 与 C/S 混合构架模式[J]. 计算机应用,2009(4):1135–1138.

[5] 余名高,吴海林,熊童满,等. AJAX 在 Struts 中的应用[J]. 计算机技术与发展,2007,17(10):69–72.

[6] 张小兵. 人力资源管理系统与组织绩效关系研究评述与展望[J]. 软科学,2009(3):106–110.

[7] 熊 文,熊淑华,孙 旭,等. Ajax 技术在 Web2.0 网站设计中的应用研究[J]. 计算机技术与发展,2012,22(3):145–148.

[8] 荣 耀,李 昕. 企业级 AJAX 框架设计与实现[J]. 南京

师范大学学报(工程技术版),2007(3):64–69.

[9] 宋瑜辉,马光思. 基于两种 Web 开发模式构建 AJAX 应用的研究[J]. 计算机技术与发展,2009,19(5):86–88.

[10] Liu C. Data flow analysis and testing of JSP-based Web applications[J]. Information and Software Technology,2006,48(12):1137–1147.

[11] 杨智勇,赵瑞锋,何 盼,等. 基于 AJAX 的自定义报表中间件平台的研究与实现[J]. 计算机工程与设计,2009(3):713–715.

[12] Yang J,Liao Z,Liu F. The impact of Ajax on network performance[J]. The Journal of China Universities of Posts and Telecommunications,2007,14(Sup):32–34.

新的软集合约简方法在购房决策中的应用

作者: [袁红彬](#), [刘文奇](#), [石梦婷](#)
作者单位: [昆明理工大学 理学院, 云南 昆明 650500](#)
刊名: [计算机技术与发展](#)
英文刊名: [Computer Technology and Development](#)
年, 卷(期): 2013(1)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wjfz201301062.aspx