

基于粗糙集的科技成果转化风险评价指标选择研究

吕超

(哈尔滨金融高等专科学校 会计系, 哈尔滨 150030)

摘要: 风险是影响科技成果转化成败的一个重要因素, 对新成果转化项目的风险强度进行识别是进行成果转化的前提, 也是转化成功的重要保证。本文利用粗糙集技术对已实施的科技成果转化项目的风险信息进行了分析, 找到其中有意义的依赖关系, 实现了以较少的指标对科技成果转化风险进行科学地分析和测度, 简化了风险强度识别程序。

关键词: 科技成果转化; 粗糙集; 风险评价指标

中图分类号: F123.16 **文献标识码:** A **文章编号:** 1002-980X(2008)02-0031-04

科技成果转化是科技和经济的结合点, 它对一个国家的科技进步和经济发展尤为重要, 各国政府对其都高度重视。然而, 我国科技成果转化工作的实施成功率一直很低, 其中科技成果转化主体不能应对各种突如其来的风险, 从而导致整个转化结果失败是一个重要原因。因此, 对科技成果转化过程中的风险强度进行识别, 以使成果转化主体及早采取对策, 是解决我国科技成果转化率较低的一项重要措施。

1 科技成果转化风险要素

科技成果转化过程一般分 3 个阶段: ①成果的小试、中试阶段; ②成果的工业化生产阶段; ③成果的市场开发、商品化阶段。风险贯穿于科技成果转化过程的 3 个阶段, 正确识别和把握各阶段的风险特点, 是进行风险预警的前提。根据每一阶段的特点, 可将科技成果转化风险分为技术风险、生产风险、市场风险和环境风险这 4 个因素, 再将每一因素细分为若干评价指标, 具体内容如下:

(1) 技术风险(C_1) ①技术的成熟度风险(C_{11})。成熟度风险是实验室阶段科技成果面临的主要风险, 它主要涉及: 科技成果所依据的原理是否合理; 实验的假设条件、实验方案、实验过程和结果是否可靠; 是否具备向产品化阶段转化的可操作性等。②知识产权风险(C_{12})。科技成果转化为实现生产力的过程是将“无形资产”转化为“有形资产”的过程, 该过程涉及成果所有权即知识产权问题, 具体涉及

两类风险: 一是侵权风险, 即在成果转化过程中由于信息不对称引致的非风险企业方的知识产权问题; 二是泄密风险, 它是风险企业方在成果转化过程中泄露技术秘密而引起的风险。

(2) 生产风险(C_2) 生产过程中的风险主要涉及原材料采供风险(C_{21})和生产组织失误风险(C_{22})。原材料的采供情况会影响新产品生产的连续性和产品的质量; 同时, 生产组织管理也会影响新产品的顺利生产, 生产组织不善很可能导致成果转化失败。

(3) 市场风险(C_3) 科技成果转化过程中的市场风险是指新产品、新技术的可行性与市场不匹配而引起的风险。产品化阶段和商品化或产业化阶段的市场风险评价依据如下:

①产品的竞争能力风险(C_{31})。目标产品在功能、特性或价格方面可能会优于现有的产品, 但是高新技术产品常常面临激烈的市场竞争。当确定产品性能指标时, 产品计划可能会低估由于激烈的竞争而造成的水涨船高问题, 而且实验室阶段的性能指标可能与产业化阶段的性能指标有较大偏差。这些都可能影响目标产品的竞争能力, 导致转化过程的市场风险。②产品的市场接受程度(C_{32})。高新技术产品往往是全新产品, 顾客在产品推出后不易及时了解其性能, 常常持观望态度, 从而使企业难以对市场能否接受及有多大容量做出准确判断。评价实验室阶段的科技成果市场风险侧重于预测科技成果的市场前景, 考察其未来产品化和商品化或产业化的市场可行性, 判断该成果转化能否实现预期投资

收稿日期: 2007-07-30

作者简介: 吕超(1976—), 男, 黑龙江哈尔滨人, 哈尔滨工程大学经济管理学院博士研究生, 研究方向: 经济管理与战略研究。

收益率。

(4) 环境风险(c_4) 相对于科技成果转化过程自身而言的各种社会、政治、自然环境所引起的风险,统称为转化过程的环境风险,其基本特征就是这些风险是转化过程的参与者所不可控制的。

2 科技成果转化风险指标选择方法

2.1 方法选择

科技成果转化风险指标的设置虽涵盖了技术风险、生产风险、市场风险和环境风险,但并不是所有的指标都对成果转化项目的整体风险强度构成决定性的影响。因此,利用模糊综合评判法、灰色预测法等方法对科技成果转化项目风险强度进行识别,难免会因冗余指标主观赋值过大而使项目风险强度评价失真。而利用粗糙集技术对已知项目的风险信息进行分析,找到其中有意义的依赖关系,是简化风险强度识别程序、提高识别准确性的重要课题。

2.2 粗糙集基本原理

粗糙集(rough set)理论^[1]是由波兰数学家 Pawlak 教授于 1982 年提出的一种研究模糊性、不完整性、不确定性的数据分析工具,它能够有效地分析和处理具有不精确、不一致和不完整等特点的各种不完备信息,并从中发现隐含的知识,揭示潜在的规律^[2,3]。粗糙集理论具有两个显著的特点:其一,它仅利用数据本身所提供的信息,无需提供所需处理的数据集合之外的任何先验信息,因此与其他不确定推理理论相比更具客观性;其二,它具备从大量数据中求取最小不变集合(称为核)与求解最小规则集(称为约简)的能力,这一特性有助于简化冗余属性和属性值,提取有用的特征信息。

在利用粗糙集技术对科技成果转化项目风险指标进行约减之前,先引入以下定义^[4,5]:

定义 1: 设 $U \neq \phi, U = (x_1, x_2, \dots, x_n)$ 是我们感兴趣的对象组成的集合,称为论域。

定义 2: 任何子集 $X \subseteq U$, 称为 U 上的一个概念或范畴, U 中的任何概念簇称为关于 U 的抽象知识,简称知识。

定义 3: 设 R 是 U 上的一个等价关系,则给定论域 U 和等价关系 R ,在 R 下对 U 的划分称为知识,由它确定的等价类记为 U/R 。 U 上的一簇划分(对 U 的分类)称为关于 U 的知识库,一个知识库可用 $K = (U, R)$ 表示,其中 U 是非空有限集, R 是 U 上的一簇等价关系。

定义 4: 设 $S = \{U, C, d\}$ 为一完全信息决策表,

其中 U 为论域, C 为条件属性集, d 为决策属性, c_i 为 C 中的一个条件属性, $(X_1, X_2, \dots, X_m)$ 为 c_i 在论域 U 上形成的一个划分, $(Y_1, Y_2, \dots, Y_n)$ 为决策属性 d 在论域 U 上形成的一个划分。

定义 5: 设 U 是给定论域, $X \subseteq U$, R 是 U 上的等价关系簇。当 X 能表达成属性子集 B 所确定的 U 上的基本集的并时,称 X 是 B 可定义的,也称为 B 精确集,否则称 X 是 B 不可定义的或称为 B 粗糙集。包含于 X 中的最大可定义集和包含 X 的最小可定义集,都是根据 B 能够确定的,前者称为 X 的下近似集,记为 $B_-(X)$,后者称为 X 的上近似集,记为 $B_+(X)$ 。

定义 6: 对信息系统 $S = \{U, R, V, f\}$, 设 $r_0 \in R$, 如果 $IND(R - \{r_0\}) = IND(R)$, 则称属性 r_0 在 R 中是冗余的, r_0 为冗余属性;否则,称 r_0 在 R 中是绝对必要的。如果每个属性 $r \in R$ 在 R 中都是绝对必要的,则称属性集 R 是独立的;否则,称 R 是可约简的。 R 中所有绝对必要属性组成的集合称为 R 的属性核,记作 $CORE(R)$ 。

定义 7: 对于决策表 $S = \{U, R, V, f\}$, $C \cup D = R$, 条件属性 $c_i \in C (i = 1, 2, \dots, n)$, c_i 对于决策属性的重要性程度定义为 k_i , 如下:

$$k_i = r_c(D) - r_{c_i}(D) \\ = \frac{\text{card}(\text{POS}_C(D)) - \text{card}(\text{POS}_{\{C/c_i\}}(D))}{\text{card}(U)}。$$

其中, $\text{card}(U)$ 表示集合 U 中元素的数量; $\text{POS}_{\{C/c_i\}}(D)$ 称为 D 的相对于 $\{C/c_i\}$ 的正域,即 U 中所有根据属性集 $\{C/c_i\}$ 进行划分后,仍可准确地划分到 D 的等价类中的对象集合; $\text{card}(\text{POS}_{\{C/c_i\}}(D))$ 表示集合 $\text{POS}_{\{C/c_i\}}(D)$ 中元素的个数。 k_i 越大,表明当从条件属性中去掉属性 c_i 后再对论域中对象进行分类时,分类 U/D 的正域所受的影响越大,则条件属性 c_i 即第 i 个指标对决策结果越重要;如果 $k_i = 0$,说明该指标对决策结果影响很小,可直接约简,不必赋予权重和评分。

决策表的属性约减就是化简决策表中的条件属性。化简后的决策表具有化简前的决策表的功能,但是化简后的决策表具有更少的条件属性,因此同样的决策表可以基于更少量的条件属性^[6,7]。

3 科技成果转化风险评价指标选择实例分析

科技成果转化风险强度在转化前的估计带有主观性,而在实施科技成果转化项目后再进行风险强

度大小的反思则更具有客观性。因此,本文选择 16 项已实施的国防领域科技成果转化项目作为样本,将风险评价指标作为条件属性,将科技成果转化风险总体强度作为决策属性,运用粗糙集理论对科技成果转化风险评价指标选择问题进行探讨。

3.1 风险值的确定原则

运用粗糙集理论进行对象评价、处理决策表时,要求决策表中的值用离散(如整型、字符串型、枚举型)数据表达。

在评价某项科技成果转化风险时,风险强度的大小可按照日常人们认识风险的习惯分为 3 类:弱风险、中等风险、强风险。为使指标值离散化,我们采用分级赋分的方式,即用 0、1、2 分别表示弱风险、中等风险和强风险。作为训练样本的 16 项成果转化项目的风险指标值经成果转化主体评价后的分值分布如表 1 所示。

表 1 16 项科技成果转化项目风险识别信息表

序号	C11	C12	C21	C22	C31	C32	C4	d
1	0	1	1	0	0	1	0	0
2	0	1	2	1	1	1	0	1
3	0	1	1	2	0	1	0	1
4	1	1	1	0	0	1	0	0
5	1	2	1	1	0	0	0	2
6	0	2	1	0	0	1	0	1
7	0	1	2	0	0	1	0	2
8	1	2	0	1	0	1	0	1
9	0	1	1	1	0	1	0	1
10	0	0	2	1	0	1	0	1
11	1	0	1	1	0	1	0	2
12	0	0	1	2	0	1	0	1
13	0	1	2	1	0	1	1	1
14	1	2	1	1	0	1	1	2
15	0	1	2	1	2	1	1	2
16	1	2	1	1	0	2	1	1

3.2 基于粗糙集的属性约减

令 $Q =$ 决策属性集 $= \{d\}$, $P =$ 条件属性全集 $= \{C_{11}, C_{12}, C_{21}, C_{22}, C_{31}, C_{32}, C_{41}, C_{42}\}$, 则

$IND(P) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}\},$
 $IND(Q) = \{\{1,4\}, \{2,3,6,8,9,10,12,13,16\}, \{5,7,11,14,15\}\},$
 $POS_P(Q) = U。$

因此,论域 U 是 P 上相对于 Q 一致的,这说明该决策表是完全确定的决策表,决策表中不包含不一致信息(样本)。

$IND(P\{C_{11}\}) = \{\{1,4\}, \{2\}, \{3\}, \{5\}, \{6\},$

$\{7\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}\};$

$IND(P\{C_{12}\}) = \{\{1,6\}, \{2\}, \{3,12\}, \{4\}, \{5\}, \{7\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}\};$

$IND(P\{C_{21}\}) = \{\{1,7\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}\};$

$IND(P\{C_{22}\}) = \{\{1,3,9\}, \{2\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}\};$

$IND(P\{C_{31}\}) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{10\}, \{11\}, \{12\}, \{13,15\}, \{14\}, \{16\}\};$

$IND(P\{C_{32}\}) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14,16\}, \{15\}\};$

$IND(P\{C_4\}) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}\}。$

从而,

$POS_{(P\{C_{11}\})}(Q) = U = POS_P(Q);$
 $POS_{(P\{C_{12}\})}(Q) = \{2,3,4,5,7,8,9,10,11,12,13,14,15,16\};$
 $POS_{(P\{C_{21}\})}(Q) = \{2,3,4,5,6,8,9,10,11,12,13,14,15,16\};$
 $POS_{(P\{C_{22}\})}(Q) = \{2,4,5,6,7,8,10,11,12,13,14,15,16\};$
 $POS_{(P\{C_{31}\})}(Q) = \{1,2,3,4,5,6,7,8,9,10,11,12,14,16\};$
 $POS_{(P\{C_{32}\})}(Q) = \{1,2,3,4,5,6,7,8,9,10,11,12,13,15\};$
 $POS_{(P\{C_4\})}(Q) = U = POS_P(Q)。$

由此可知,条件属性 C_{11} 、 C_4 是相对于决策属性 d 可省略的,但不一定可同时省略。而条件属性 C_{12} 、 C_{21} 、 C_{22} 、 C_{31} 、 C_{32} 是相对于决策属性 d 不可省略的,因此 $CORE_Q(P) = \{C_{12}, C_{21}, C_{22}, C_{31}, C_{32}\}$ 。进一步:

$IND(P\{C_{11}, C_4\}) = \{\{1,4\}, \{2\}, \{3\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}\},$

$POS_{(P\{C_{11}, C_4\})}(Q) = U = POS(Q),$
故属性 C_{11} 是条件属性集 $P\{C_4\}$ 相对于决策属性 d 可省略的,属性 C_4 是条件属性集 $P\{C_{11}\}$ 相对于决策属性 d 可省略的。因此, $CORE_Q(P) = \{C_{12}, C_{21}, C_{22}, C_{31}, C_{32}\}。$

4 分析及总结

本文选取 16 项已实施的国防领域科技成果转化

化项目,运用粗糙集理论对其风险评价指标进行了筛选。从筛选结果来看,本文所选取的风险评价指标比较符合实际情况。技术市场上的非国防科技成果项目更注重市场前景,而对于国防科技成果而言,其技术成熟度是至关重要的。另外,国防科技成果的转化主体一般是国家,转化的直接目的是满足国防建设的需要,而非市场经济利益的驱动,因此政府部门会为国防科技成果的转化开辟有利途径,降低转化过程中的经济、政策等风险。

从以上分析来看,基于粗糙集的科技成果转化风险评价指标筛选方法具有较好的应用效果,选择的样本较好地验证了方法的正确性。要说明的是,本文筛选出的指标并非适用于所有类型的科技成果转化风险评价,针对不同类型的科技成果转化项目,应该选择不同类型的成果样本进行指标筛选,并要进行验证。只有得到的结果较好地模拟了实际情况,才能将其应用于实践,这需要引起注意,但这并

不影响本方法的实用性。

参考文献

- [1] PAWLAK Z. Rough sets [J]. International Journal of Computer and Information Science, 1982:11(5):341-356.
- [2] 王宗军,李红侠,邓晓岚.粗糙集理论的最新应用[J].统计与决策,2005,41(1):64-68.
- [3] 孙国梓吴志军等.基于粗糙集的供应商绿色评价体系研究[J].计算机工程与应用,2004:40(32):190-192.
- [4] ZIARKO W. Variable precision rough set model[J]. Journal of Computer and System Sciences, 1993,46(1):39-59.
- [5] 苗夺谦,王珏.基于粗糙集的多变量决策树的构造方法[J].软件学报,1997(6):46-48.
- [6] TSUMOTO S. Mining diagnostic rules from clinical databases using rough sets and medical diagnostic model [J]. Information Sciences, 2004,162(2):65-80.
- [7] LI R P, WANG Z O. Mining classification rules using rough sets and neural networks[J]. European Journal of Operational Research, 2004,157(2):439-448.

Research on Risk Assessment Index of Technological Achievement Transformation Based on Rough Set Theory

Lv Chao

(Accounting Department, Harbin Finance College, Harbin 150030, China)

Abstract: Risk is an important factor influencing the result of the technological achievement transformation. And risk recognition is the premise and the important guarantee for successful technological achievement transformation. Based on rough set theory, this paper analyzes the risks of technological achievement transformation projects implemented, finds out the significant dependence relations and achieves the objective of scientifically analyzing and measuring the risks of technological achievement transformation projects by less indexes.

Key words: technological achievement transformation; rough set; risk assessment index