

# 基于模糊关联聚类的知识网选择方法

杨人子<sup>1,2†</sup>, 严洪森<sup>1,3</sup>

(1. 东南大学 自动化学院, 江苏 南京 210096; 2. 东南大学 数学系, 江苏 南京 210096;

3. 复杂工程系统测量与控制教育部重点实验室, 江苏 南京 210096)

**摘要:** 针对知识化制造系统中相似知识网日益增多和用户需求表达不清晰等导致的知识网选择问题, 提出一种基于模糊关联聚类的知识网选择方法. 综合知识网功能、完善程度和结构等方面构造的相似度具有反映知识网运算规律的特征. 将两两知识网的相似度作为聚类数据, 降低了高维特征空间的维数. 模糊关联矩阵的分解, 获得了知识网-类关系. 目标知识网与类中类隶属度高的知识网的比较缩小了用户选择范围. 最后的实例表明该方法是有有效可行的.

**关键词:** 知识化制造; 知识网; 模糊关联聚类; 相似度; 矩阵分解

**中图分类号:** TP11; TH165 **文献标识码:** A

## Selection of knowledge mesh based on fuzzy relational clustering

YANG Ren-zi<sup>1,2†</sup>, YAN Hong-sen<sup>1,3</sup>

(1. School of Automation, Southeast University, Nanjing Jiangsu 210096, China;

2. Department of Mathematics, Southeast University, Nanjing Jiangsu 210096, China;

3. Key Laboratory of Measurement and Control of CSE, Ministry of Education, Nanjing Jiangsu 210096, China)

**Abstract:** The knowledge mesh's selection method based on fuzzy relation clustering is proposed for the selecting problem caused by the increasing similar knowledge mesh and unclear requirements of user in knowledgeable manufacturing system. The similarity degree which synthesizes the function, perfection degree and structure of knowledge mesh, has the characteristics of reflecting operation laws. The similarity values as cluster data reduce the dimension of high-dimensional feature space. The decomposition of fuzzy relational matrix obtains the groups of knowledge meshes. The comparisons between target knowledge mesh and knowledge mesh with high class membership degree in each class narrow the scope of user selection. The last example shows that the method is effective and feasible.

**Key words:** knowledgeable manufacturing; knowledge network; fuzzy relational clustering; similarity; matrix decomposition

### 1 引言(Introduction)

知识化制造是2000年提出的一种新的制造理念<sup>[1]</sup>. 它将各种先进制造模式表示成知识网(knowledge meshes, KMs)<sup>[2-3]</sup>并纳入知识化制造系统, 这种系统能够随时根据需要选择一种或几种先进制造模式, 通过自适应、自学习、自进化或自重构等特征技术对其进行进化或组合优化, 从而解决现有制造模式中存在的模式单一、缺乏灵活性、不能满足制造企业需求的多样性以及重复研发等问题.

知识化制造系统已经建立了基于知识网和agent网的复杂知识表示方法, 实现了将各种先进制造模式以知识网的形式存储在知识网库中<sup>[2]</sup>; 知识网多重集理论<sup>[2]</sup>实现了知识网的各种组合形式; 知识网

的自动生成方法<sup>[3]</sup>实现了知识网多重集运算表达式的优化. 运用这些技术能够产生满足用户多种需求的新知识网, 但也使得知识网库中知识网之间存在相似或相同的部分, 并且随着各种制造模式的不断涌现和特征技术的完善, 这种相似或相同会越来越多, 用户往往需要在相似知识网之间做出选择. Ding等<sup>[4]</sup>考虑了知识网相似部分的自重构问题, 提出基于近似匹配的知识化制造系统自重构的理论和算法. Yang等<sup>[5]</sup>提出基于信息粒度的知识网模糊分类和检索方法, 避免了自重构中知识网的重复检索和运算. 但这些对知识网相似或相同部分的匹配方法和分类方法, 仅适合能清晰描述其具体需求的高素质用户, 对于中低素质用户, 仍有可能因为表达

不清晰而无法从知识网库中选择知识网. 如果系统能够自动给出一些具有代表性的参考知识网, 就能帮助用户明确需求, 从而快速找到自己需要的知识网. 对于高素质用户, 参考知识网能够提供一定的参考意见, 并缩小选择范围. 因此如何从知识网库中获得类似于模板的参考知识网, 并呈现给用户帮助其选择, 是本文要解决的问题.

聚类是进行数据分析、无监督学习、数据粒化以及信息压缩的重要工具. 聚类算法丰富多彩. 传统的硬聚类如层次聚类、动态聚类, 将每个辨识对象严格地划分为属于某一类. 将模糊概念引入到聚类中, 允许部分隶属所产生的模糊聚类算法, 如模糊 $C$ 均值聚类<sup>[6]</sup>、模糊子空间聚类<sup>[7]</sup>、部分监督聚类<sup>[8]</sup>、模糊关联聚类<sup>[9]</sup>、自适应模糊均值聚类<sup>[10]</sup>等, 使得聚类的结果更符合现实世界的实际情况. 这些算法在智能数据分析、数据挖掘和模式识别等领域<sup>[6-15]</sup>有着广泛的应用. 但在知识化制造系统中运用聚类方法协助用户解决知识网选择问题的文献却极为有限, 仅在文献<sup>[5]</sup>中Yang等提出一种迭代构建聚类中心的知识网模糊聚类方法, 但该方法是在用户需求指导下的知识网特征空间中进行聚类, 而没有用户需求限定的知识网特征空间是一个高维的特征空间, 知识网在该空间上的分布会显得有些稀疏. 模糊关联聚类方法是以样本对的关联数据为聚类数据, 转换后的特征空间降低了维数却增加了聚类样本数, 因此避免了样本的稀疏分布. 不过将其用于知识网聚类需要建立知识网之间的相似度模型. Xue等<sup>[3]</sup>用知识点功能满意度定义相似度, 其完全依赖用户水平而过于主观; Ding等<sup>[4]</sup>用知识点相同功能数目定义相似度, 其受限于知识点功能描述形式且忽略了知识网的其他元素; Yang等<sup>[5]</sup>用匹配度、完善度和复杂度综合定义相似度, 其缺乏对具体元素的测度讨论, 且知识网复杂度解决的是多重数导致的完全相同的知识网对应的多重集未必相同的问题, 并不能体现知识网本身结构的不同. 因此需要在已有的相似度函数的基础之上, 建立更为合理的知识网相似度, 这也是对知识网进行聚类的前提.

本文提出一种基于模糊关联聚类的知识网选择方法. 知识网的相似度是从质、量和知识网的结构等3个方面给出, 匹配度具体为知识网最底层知识点的功能匹配度, 知识网的结构差异由知识网层次数体现. 相似度性质的讨论显示其能够反映知识网在运算过程中的变化特征. 采用模糊关系分解的方法, 构造了样本-类关系. 每个类中类隶属度高的知识网具有参考价值, 可以将其呈现给中低素质用户. 而对于高素质用户, 对用户需求的目標知识网与每个类中类隶属度最高的知识网进行比较则缩小了用户选择范围. 用户只需要对确定类中的知识网进行选择,

而不需要对所有知识网进行选择.

## 2 知识网的相似性度量(Similarity measure of KMs)

任何一种先进制造系统都是一种制造模式的具体实现, 都由若干功能相对独立且相互有联系的模块构成. 对任何一个模块的子模块乃至子模块的子模块等, 也都可以这样看待. 将模块视为知识点, 则先进制造模式由知识点和它们之间的联系构成, 形成一个网络结构. 这样, 任何一种先进制造模式均可转化为一种知识网<sup>[2]</sup>. 知识网是由知识点、信息联系、信息流和知识点功能等构成的大集合, 可以简单的记为知识网 $W = \{x_1, x_2, \dots, x_n\}$ .

### 2.1 知识网的匹配度(Matching degree of KMs)

在知识化制造系统中, 每个知识网都一一对应于某种先进制造模式, 用户选择制造模式也就是对知识网进行选择. 用户需求主要为功能需求, 因此对知识网的选择主要是基于功能需求的选择. 讨论知识网的相似性, 也就主要从功能的角度来讨论. 而知识网中父知识点功能必包含子知识点的所有功能, 所以下面讨论的功能是知识网最底层知识点的功能.

**定义 1** 设知识网 $V$ 和 $W$ 的最底层知识点集合分别为 $P_V = \{p_{v_1}, p_{v_2}, \dots, p_{v_m}\}$ 和 $P_W = \{p_{w_1}, p_{w_2}, \dots, p_{w_n}\}$ , 知识点 $p_{v_i}$ 的功能数为 $l_{v_i}$  ( $i = 1, 2, \dots, m$ ), 知识点 $p_{w_j}$ 的功能数为 $l_{w_j}$  ( $j = 1, 2, \dots, n$ ),  $P_W$ 中与知识点 $p_{v_i}$ 具有的共同功能数为 $l_{v_i, W}$ ,  $P_V$ 中与知识点 $p_{w_j}$ 具有的共同功能数为 $l_{w_j, V}$ , 则 $P_V \times P_W$ 上的函数

$$f(P_V, P_W) = \frac{\sum_{i=1}^m l_{v_i, W} + \sum_{j=1}^n l_{w_j, V}}{\sum_{i=1}^m l_{v_i} + \sum_{j=1}^n l_{w_j}}$$

称为知识网 $V$ 和 $W$ 的匹配度.  $\sum_{i=1}^m l_{v_i, W} = \sum_{j=1}^n l_{w_j, V}$ , 也可以化简为

$$f(P_V, P_W) = \frac{2 \sum_{i=1}^m l_{v_i, W}}{\sum_{i=1}^m l_{v_i} + \sum_{j=1}^n l_{w_j}}$$

**性质 1** 知识网 $V$ 和 $W$ 的匹配度具有如下性质:

1)  $0 \leq f(P_V, P_W) \leq 1$ . 当知识网 $V$ 和 $W$ 具有完全相同的功能,  $f(P_V, P_W) = 1$ ; 当知识网 $V$ 和 $W$ 不具有任何相同的功能,  $f(P_V, P_W) = 0$ ; 特别的,  $f(P_V, P_W) = 1$ .

2)  $f(P_V, P_W) = f(P_W, P_V)$ .

3) 当知识网 $V$ 和 $W$ 最底层知识点功能总数一定时, 匹配度随着 $V$ 和 $W$ 相同功能数量的增加而增加.

4) 如果知识网 $V$ 包含知识网 $W$ 的所有功能, 且

$V$ 不同于 $W$ 的那些功能也不同于知识网 $X$ 的任何功能,则 $f(P_V, P_X) \leq f(P_W, P_X)$ .

5) 设 $P_{V+W}$ 为知识网 $V$ 和 $W$ 的并集的最底层知识点集合,如果 $V$ 和 $W$ 不具有相同的功能,且知识网 $X$ 包含 $V$ 的所有功能,则 $f(P_W, P_X) \leq f(P_{V+W}, P_X)$ .

6) 设 $P_{V \cap W}$ 为知识网 $V$ 和 $W$ 的交集的最底层知识点集合,如果知识网 $X$ 既包含 $V$ 的所有功能,也包含 $W$ 的所有功能,则

$$f(P_{V \cap W}, P_X) \leq \min\{f(P_V, P_X), f(P_W, P_X)\}.$$

证 性质1), 2)和3)显然成立,下面证明性质4)–

6). 设知识网 $V, W, X$ 的最底层知识点集合分别为

$$P_V = \{p_{v_1}, p_{v_2}, \dots, p_{v_m}\},$$

$$P_W = \{p_{w_1}, p_{w_2}, \dots, p_{w_n}\},$$

$$P_X = \{p_{x_1}, p_{x_2}, \dots, p_{x_r}\},$$

对应知识点的功能数目分别 $l_{v_i}, l_{w_j}, l_{x_k}$ 并设对应的相同功能的数目分别为 $l_{v_i, X}, l_{w_j, X}, l_{x_k, V}, l_{x_k, W}$ .

性质1中4): 由已知,知识网 $V$ 和 $X$ 的相同功能总数与知识网 $W$ 和 $X$ 的相同功能总数相等,即

$$\sum_i l_{v_i, X} + \sum_k l_{x_k, V} = \sum_j l_{w_j, X} + \sum_k l_{x_k, W},$$

$$\sum_i l_{v_i} \geq \sum_j l_{w_j},$$

$$f(P_V, P_X) = \frac{\sum_i l_{v_i, X} + \sum_k l_{x_k, V}}{\sum_k l_{x_k} + \sum_i l_{v_i}} \leq$$

$$\frac{\sum_j l_{w_j, X} + \sum_k l_{x_k, W}}{\sum_k l_{x_k} + \sum_j l_{w_j}} = f(P_W, P_X).$$

性质1中5): 由已知,  $P_{V+W} = \{p_{v_1}, p_{v_2}, \dots, p_{v_m}, p_{w_1}, p_{w_2}, \dots, p_{w_n}\}$ ,

$$\sum_i l_{v_i, X} = \sum_k l_{x_k, V} = \sum_i l_{v_i},$$

$$f(P_W, P_X) = \frac{\sum_j l_{w_j, X} + \sum_k l_{x_k, W}}{\sum_k l_{x_k} + \sum_j l_{w_j}} \leq$$

$$\frac{\sum_i l_{v_i, X} + \sum_j l_{w_j, X} + \sum_k l_{x_k, W}}{\sum_i l_{v_i} + \sum_k l_{x_k} + \sum_j l_{w_j}} \leq$$

$$\frac{\sum_i l_{v_i, X} + \sum_j l_{w_j, X} + \sum_k l_{x_k, W} + \sum_k l_{x_k, V}}{\sum_i l_{v_i} + \sum_k l_{x_k} + \sum_j l_{w_j}} =$$

$$f(P_{V+W}, P_X).$$

性质1中6): 设 $P_{V \cap W} = \{p_{s_1}, p_{s_2}, \dots, p_{s_i}\}$ , 有

$$\sum_h l_{s_h} = \sum_h l_{s_h, X} \leq \min\{\sum_i l_{v_i, X}, \sum_j l_{w_j, X}\},$$

$$\sum_i l_{v_i} = \sum_i l_{v_i, X} + \sum_j l_{w_j} = \sum_j l_{w_j, X} + \sum_j l_{w_j},$$

$$f(P_{V \cap W}, P_X) =$$

$$\frac{\sum_h l_{s_h, X} + \sum_k l_{x_k, V \cap W}}{\sum_h l_{s_h} + \sum_k l_{x_k}} \leq$$

$$\frac{\sum_i l_{v_i, X} + \sum_k l_{x_k, V \cap W}}{\sum_i l_{v_i} + \sum_k l_{x_k}} \leq$$

$$\frac{\sum_i l_{v_i, X} + \sum_k l_{x_k, V}}{\sum_i l_{v_i} + \sum_k l_{x_k}} = f(P_V, P_X).$$

同理,  $f(P_{V \cap W}, P_X) \leq f(P_W, P_X)$ . 所以 $f(P_{V \cap W}, P_X) \leq \min\{f(P_V, P_X), f(P_W, P_X)\}$ . 证毕.

性质1说明在功能总数一定的情况下, 匹配度是随着相同功能的数量的增加而增加. 并集的相似度不一定大于子集的相似度, 如性质1里的4); 而在满足性质1里5)的条件下, 并集的相似度可以大于子集的相似度. 同样的, 交集的相似度也不一定小于原集合的相似度, 而是满足性质1里6)的前提下, 交集的相似度才可以小于原集合的相似度. 这些结论说明定义1能够反映知识网间一定程度的相似性.

## 2.2 知识网的功能完善度(Functional perfection degree of KMs)

定义1表现的是知识网功能在量上的相似. 在实际情况中, 当用户需要某些功能, 往往并不仅仅是具有该功能的知识网就一定满足用户的要求, 用户可能希望功能能够完善, 满足以后进一步的发展, 或者也可能为了节约成本, 只需要该功能的基本特征即可. 另外, 笔者知道, 对于知识网中知识点的功能, 虽然功能描述相同, 但其完善程度可能并不一样. 因此用知识网的知识点功能完善度来反映知识网之间质上的不同. 这种完善度不能简单的用完善和不完善、0和1两个数值来定义, 采用一个模糊集合来定义更符合实际意义. 而父知识点功能完善度也可以通过子知识点的功能完善度获得.

**定义2** 设知识网 $W$ 的最底层知识点集合

$$P = \{p_1, p_2, \dots, p_n\},$$

则在 $P$ 上的一个模糊集定义为 $\mu(p_i) : P \rightarrow [0, 1]$ ,  $x_i \in P$ ,  $\mu(p_i)$ 称为知识点 $p_i$ 的功能完善度, 简记为 $\mu_{p_i}$ .  $\mu = (\mu_{p_1}, \mu_{p_2}, \dots, \mu_{p_n})$ 称为知识网 $W$ 的功能完善度.  $\mu_{p_i} = 0$ 表示知识网不具备知识点 $p_i$ , 数值越大说明元素 $p_i$ 的功能越完善.

由定义2可知每个知识网都与一个模糊集 $\mu$ 一一对应, 也可以说一个知识网与一个模糊向量 $\mu$ (完善度)是一一对应的,  $\mu \in [0, 1]^n$ .

### 2.3 知识网的相似度(Similarity degree of KMs)

定义1和定义2从质和量两方面来说明知识网的相似性,但是两个知识网的功能和完善度完全相同,仍不能说明它们是相同的,因为知识网结构会不一样. 知识点所处的层次是一个简单地反映知识网结构的量. 将其引入到知识网的相似度定义中. 综合考虑知识网的质、量和层次结构,给出知识网的相似度定义.

**定义3** 知识网 $V$ 和 $W$ 的相似度为

$$\text{sim}(V, W) = \frac{\sum_{i=1}^m \gamma_{v_i} l_{v_i, W} \kappa_{v_i} + \sum_{j=1}^n \gamma_{w_j} l_{w_j, V} \kappa_{w_j}}{\sum_{i=1}^m \alpha_{v_i} l_{v_i} \mu_{v_i} + \sum_{j=1}^n \beta_{w_j} l_{w_j} \mu_{w_j}}, \quad (1)$$

其中:  $V$ 和 $W$ 的最底层知识点集合分别为

$$P_V = \{p_{v_1}, p_{v_2}, \dots, p_{v_m}\}, P_W = \{p_{w_1}, p_{w_2}, \dots, p_{w_n}\};$$

$\mu_{v_i}, l_{v_i}, \alpha_{v_i}$ 分别为知识点 $p_{v_i}$ 的功能完善度、功能数、所处的知识网层次;  $\mu_{w_j}, l_{w_j}, \beta_{w_j}$ 分别为知识点 $p_{w_j}$ 的功能完善度、功能数、所处的知识网层次;  $l_{v_i, W}$ 为 $P_W$ 中与知识点 $P_{v_i}$ 具有相同功能数,  $\beta_{W(v_i)}, \mu_{W(v_i)}$ 为这些功能相关知识点所处的层次和功能完善度;  $l_{w_j, V}$ 为 $P_V$ 中与知识点 $p_{w_j}$ 具有的相同功能数,  $\alpha_{V(w_j)}, \mu_{V(w_j)}$ 为这些功能相关知识点所处的层次和功能完善度;

$$\gamma_{v_i} = \min\{\beta_{W(v_i)}, \alpha_{v_i}\}, \gamma_{w_j} = \min\{\alpha_{V(w_j)}, \beta_{w_j}\}, \\ \kappa_{v_i} = \min\{\mu_{W(v_i)}, \mu_{v_i}\}, \kappa_{w_j} = \min\{\alpha_{V(w_j)}, \mu_{w_j}\}.$$

这里需要说明的是,知识网的层次数是由根知识点开始依次递增的,即父知识点的层次数小于子知识点的层次数;若 $P_W$ 中与知识点 $P_{v_i}$ 具有相同功能的知识点不唯一,则将 $l_{v_i, W}$ 拆成和式的表达形式,使得每项仅与 $P_W$ 中唯一的知识点对应,再进行计算,如 $P_{v_i}$ 与 $P_W$ 中 $P_{w_1}, P_{w_2}$ 分别具有相同功能,则

$$\gamma_{v_i} l_{v_i, W} \kappa_{v_i} = \gamma_{v_{i1}} l_{v_{i1}, W} \kappa_{v_{i1}} + \gamma_{v_{i2}} l_{v_{i2}, W} \kappa_{v_{i2}},$$

其中:  $\gamma_{v_{i1}} = \min\{\beta_{w_1}, \alpha_{v_i}\}, \gamma_{v_{i2}} = \min\{\beta_{w_2}, \alpha_{v_i}\}, l_{v_{i1}, W}$ 为 $P_{w_1}$ 与知识点 $P_{v_i}$ 具有相同功能数,  $l_{v_{i2}, W}$ 为 $P_{w_2}$ 与知识点 $P_{v_i}$ 具有相同功能数,  $\kappa_{v_{i1}} = \min\{\mu_{w_1}, \mu_{v_i}\}, \kappa_{v_{i2}} = \min\{\mu_{w_2}, \mu_{v_i}\}$ ;  $l_{v_i, W}$ 可以拆开看待,也就是说 $P_V$ 和 $P_W$ 中具有相同功能的知识点可以看成一一对应,因此式(1)分子中

$$\sum_{i=1}^m \gamma_{v_i} l_{v_i, W} \kappa_{v_i} = \sum_{j=1}^n \gamma_{w_j} l_{w_j, V} \kappa_{w_j},$$

则式(1)化简为

$$\text{sim}(V, W) = \frac{2 \sum_{i=1}^m \gamma_{v_i} l_{v_i, W} \kappa_{v_i}}{\sum_{i=1}^m \alpha_{v_i} l_{v_i} \mu_{v_i} + \sum_{j=1}^n \beta_{w_j} l_{w_j} \mu_{w_j}}.$$

知识网间相似度的性质类似于性质1.

**性质2** 1)  $0 \leq \text{sim}(V, W) \leq 1$ . 当知识网 $V$ 和 $W$ 具有完全相同的功能,这些功能所对应的知识点的功能完善度和层次也相同,则 $\text{sim}(V, W) = 1$ ;当知识网 $V$ 和 $W$ 不具有任何相同的功能,  $\text{sim}(V, W) = 0$ ;特别的,  $\text{sim}(V, V) = 1$ .

2)  $\text{sim}(V, W) = \text{sim}(W, V)$ .

3) 如果知识网 $V$ 包括知识网 $W$ 的所有功能,这些功能所对应的最底层知识点的功能完善度和知识网层次相同,且 $V$ 不同于 $W$ 的那些功能也不同于知识网 $X$ 的任何功能,则 $\text{sim}(V, X) \leq \text{sim}(W, X)$ .

4) 如果知识网 $V$ 和知识网 $W$ 不具有相同的功能,且知识网 $X$ 包含 $V$ 的所有功能,对应知识点的功能完善度和知识网层次相同,则

$$\text{sim}(W, X) \leq \text{sim}(V + W, X).$$

5) 如果知识网 $X$ 既包含 $V$ 的所有功能,也包含 $W$ 的所有功能,且对应知识点的功能完善度和知识网层次相同,则

$$\text{sim}(V \cap W, X) \leq \min\{\text{sim}(V, X), \text{sim}(W, X)\}.$$

**证** 类似于性质1的证明,略.

## 3 知识网的模糊关联聚类与选择方法(Fuzzy relational clustering and selection methods of KMs)

### 3.1 模糊关联矩阵的构造和分解(Construction and decomposition of fuzzy relational matrix)

设知识库中有 $V_1, V_2, \dots, V_N$ 共 $N$ 个知识网,两两做比较,由式(1)可获得 $N(N-1)/2$ 个相似度值. 将其组成一个 $N \times N$ 矩阵,称为模糊关联矩阵. 设为 $R = [r_{ij}], i, j = 1, 2, \dots, N$ , 其中 $r_{ij} = \text{sim}(V_i, V_j)$ .  $R$ 具有反身性和对称性,即 $r_{ii} = 1$ 和 $r_{ij} = r_{ji}$ .

现在构造矩阵

$$G = [g_{ij}], i = 1, 2, \dots, N, j = 1, 2, \dots, c, c < N,$$

满足 $R = G \circ G^T$ , 其中:  $G^T$ 表示 $G$ 的转置,  $\circ$ 表示关系合成算子. 即通过对模糊关联矩阵 $R$ 的分解,寻找 $G$ , 获得知识网-类关系, 其中 $c$ 为分类数.

将算子取为模糊集合的 $s-t$ 卷积, 则该问题转化为求 $G = [g_{ij}]$ , 满足 $r_{ij} = \underset{k=1}{\overset{c}{S}}(t(g_{ik}, g_{jk}))$ . 在本文中,

$$t(x, y) = xy, S(x, y) = x + y - xy. \quad (2)$$

要找到一个准确解 $G$ 非常困难, 因此将该问题转化为最优化问题, 即寻找 $G$ 使得 $Q = \|R - G \circ G^T\|^2$ 的值最小, 具体形式为

$$\min_G Q = \min_G \left\{ \sum_{i=1}^N \sum_{j=1}^N [r_{ij} - \underset{k=1}{\overset{c}{S}}(t(g_{ik}, g_{jk}))]^2 \right\}. \quad (3)$$

$G$ 的计算采用基于梯度的方法,由 $G = G - \beta \cdot \nabla_G Q$ 不断更新 $G$ ,其中 $\beta(> 0)$ 代表学习速率,其表达式为

$$g_{ik} = g_{ik} - \beta \cdot \frac{\partial Q}{\partial g_{ik}}. \quad (4)$$

迭代过程从一个随机赋值的模糊关系开始,模糊关系的元素介于0到1之间.梯度的详细计算:

$$\frac{\partial Q}{\partial g_{ab}} = -2 \sum_{i=1}^N \sum_{j=1}^N [r_{ij} - \overset{c}{S}(t(g_{ik}, g_{jk}))] \cdot \frac{\partial Q}{\partial g_{ab}} (\overset{c}{S}(t(g_{ik}, g_{jk}))). \quad (5)$$

式(5)右边和式内部的导数分成4种情况分别考虑:

i) 当 $i \neq a, j \neq a, \overset{c}{S}(t(g_{ik}, g_{jk}))$ 对 $g_{ab}$ 的偏导为0;

ii) 当 $i = a, j \neq a,$

$$\frac{\partial}{\partial g_{ab}} (\overset{c}{S}(t(g_{ik}, g_{jk}))) = \frac{\partial}{\partial g_{ab}} (A + g_{ab}g_{jb} - Ag_{ab}a_{jb}) = g_{jb}(1 - A),$$

其中 $A = \overset{c}{S}_{k=1, k \neq b}(t(g_{ak}, g_{jk}));$

iii) 当 $i \neq a, j = a,$

$$\frac{\partial}{\partial g_{ab}} (\overset{c}{S}(t(g_{ik}, g_{jk}))) = g_{ib}(1 - B),$$

其中 $B = \overset{c}{S}_{k=1, k \neq b}(t(g_{ik}, g_{ak}));$

iv) 当 $i = a, j = a,$

$$\frac{\partial}{\partial g_{ab}} (\overset{c}{S}(t(g_{ik}, g_{jk}))) = 2g_{ab}(1 - C),$$

其中 $C = \overset{c}{S}_{k=1, k \neq b}(t(g_{ak}, g_{ak})).$

### 3.2 知识网的选择(Selection method of KMs)

利用第3.1节的方法和最大隶属度原则,知识库中的每个知识网都归属于某一类.将每个类中类隶属度最高的知识网作为参考知识网提供给用户,特别是对于那些不能清晰描述其具体需求的中低素质用户,这些知识网具有很好的参考作用.而对于能够表达具体需求的高素质用户,这些类隶属度最高的知识网并不一定是最佳选择,但是可以缩小选择范围.设用户需求知识网为目标知识网 $\bar{W}$ .将目标知识网与每个类中类隶属度最高的知识网依据式(6)进行相似度计算,数值越大说明目标知识网与该类知识网越接近;再将目标知识网与该类中所有知识网做比较,将相似度数值最大的知识网作为参考知识网.因为用户需求最大化是选择的要求,将知识库中知识网按照用户需求计算相似度,这有别于知识库中知识网的比较,因此式(6)是由式(1)稍作修改获得的公式,其中 $\beta_{W(\bar{w}_i)}, l_{\bar{w}_i, W}, \mu_{W(\bar{w}_i)}$ 分别为知识网 $W$ 中与 $\bar{W}$ 的知识点 $P_{\bar{w}_i}$ 具有相同功能的知

识点的层次、相同功能数和完善度.

$$\text{sim}(\bar{W}, W) = \frac{\sum_{i=1}^m \gamma_{\bar{w}_i} l_{\bar{w}_i, W} \kappa_{\bar{w}_i} + \sum_{j=1}^n \gamma_{w_j} l_{w_j, \bar{W}} \kappa_{w_j}}{\sum_{i=1}^m \alpha_i l_{\bar{w}_i} \mu_{\bar{w}_i} + \sum_{i=1}^m \beta_{W(\bar{w}_i)} l_{\bar{w}_i, W} \mu_{W(\bar{w}_i)}}. \quad (6)$$

### 3.3 算法步骤(Steps of algorithm)

知识网的模糊关联聚类与选择方法的算法步骤可概括为:

**步骤 1** 列出知识库中 $N$ 个知识网最底层知识点的功能完善度、功能数、所处的知识网层次,按照式(1)计算两两知识网的相似度.

**步骤 2** 将相似度数值 $r_{ij} = \text{sim}(V_i, V_j)$ 作为模糊聚类数据,构造模糊关联矩阵

$$R = [r_{ij}], i, j = 1, 2, \dots, N.$$

**步骤 3** 按照式(2)分解矩阵 $R$ ,使得 $R = G \circ G^T$ ,建立形如式(3)的最优化目标函数.

**步骤 4** 采用基于梯度的迭代方法,取定 $\beta$ 和迭代次数,按照式(4)和式(5)计算

$$G = [g_{ij}], i = 1, 2, \dots, N, j = 1, 2, \dots, c,$$

获得知识网的类隶属信息.

**步骤 5** 将类中类隶属度最高的知识网作为参考知识网.或者将用户需求映射为目标知识网,按照式(6)计算目标知识网与每个类中类隶属度最高的知识网的相似度,确定所要选择的类;再按照式(6)计算目标知识网与确定类中知识网的相似度,确定要选择的知识网.

## 4 实例分析(Empirical analysis)

在知识库中有15个知识网 $W_1, W_2, \dots, W_{15}$ ,现在对其进行聚类,选出参考知识网提供给用户.

### 4.1 相似度的计算(Calculation of similarity)

由于相似度的计算需要每个知识网的完整信息,受篇幅所限,这里仅以知识网 $W_1$ 和 $W_2$ 的相似度计算过程为例,设知识网 $W_1$ 和 $W_2$ 为某车身厂简化的管理信息系统,见图1和图2.

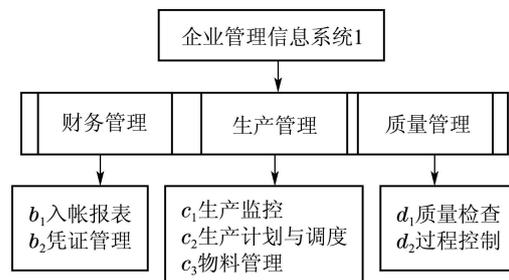


图1 知识网 $W_1$

Fig. 1 Knowledge mesh  $W_1$

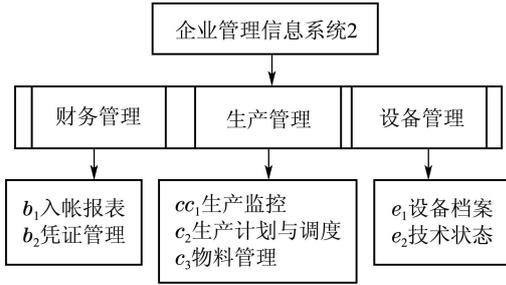


图 2 知识网  $W_2$

Fig. 2 Knowledge mesh  $W_2$

知识网  $W_1$  和  $W_2$  最底层知识点集合分别为

$$P_{W_1} = (b_1, b_2, c_1, c_2, c_3, d_1, d_2),$$

$$P_{W_2} = (b_1, b_2, cc_1, c_2, c_3, e_1, e_2),$$

层次数均为 3, 它们的功能分别是:  $b_1$  为入帐查询、入帐月统计、入帐季统计;  $b_2$  为凭证查询、凭证录入、凭证删除;  $c_1$  为生产完成情况、设备故障情况;  $cc_1$  为生产完成情况、设备故障情况、原料报废情况;  $c_2$  为工序分析、冲压计划报表、冲压加工单;  $c_3$  为零部件需求计划、生产完成情况、原料使用情况;  $d_1$  为检查计划、检查卡片、产品质量、检查信息;  $d_2$  为进货检验、关键工序检验、出厂检验;  $e_1$  为设备卡片、技术资料、附属配件、特种设备;  $e_2$  为设备停用、设备调拨、设备封存、设备闲置. 知识网  $W_1$  和  $W_2$  的功能完善度分别为

$$\mu_{W_1} = (0.9638, 0.8634, 0.4178, 0.5550, 0.6204, 0.2010, 0.1568),$$

$$\mu_{W_2} = (0.8011, 0.8472, 0.3011, 0.4099, 0.4582, 0.1057, 0.1057).$$

由式(1), 知识网  $W_1$  和  $W_2$  相似度为

$$\text{sim}(W_1, W_2) = \frac{C}{A+B} = 0.8150,$$

其中:

$$A = 3 \times 0.8011 + 3 \times 0.8472 + 3 \times 0.3011 + 3 \times 0.4099 + 3 \times 0.4582 + 4 \times 0.1057 + 4 \times 0.1057,$$

$$B = 3 \times 0.9638 + 3 \times 0.8634 + 2 \times 0.4178 + 3 \times 0.5550 + 3 \times 0.6204 + 4 \times 0.2010 + 3 \times 0.1568,$$

$$C = 2 \times 3 \times 0.8011 + 2 \times 2 \times 0.3011 + 2 \times 3 \times 0.4099 + 2 \times 3 \times 0.4582.$$

下面给出知识网  $W_1, W_2, \dots, W_{15}$  的完善度数值, 见表 1. 这些数值的获得可以采用专家评分的方法, 即请一些有经验的专家对各知识网进行仔细分析、比较和评价后, 分别给出各知识网的每个元素的完善度, 求其平均值作为该元素的完善度. 表中的数值是按照一定规律任意选取的, 如  $W_1, W_2, W_3, W_8$  在财务管理方面、 $W_4, W_6, W_7, W_{11}$  在生产管理方面、 $W_9, W_{10}, W_{13}, W_{14}, W_{15}$  在质量和设备管理方面的完善度值高一些, 而  $W_5$  的各项数值都很低,  $W_{12}$  的各项数值都不突出. 所给的数据呈现 4 个明显的类.

表 1 各知识网的功能完善度

Table 1 The functional perfection degree of knowledge meshes

	$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	$W_8$	$W_9$	$W_{10}$	$W_{11}$	$W_{12}$	$W_{13}$	$W_{14}$	$W_{15}$
入帐报表	0.9638	0.8011	0.8571	0.2461	0.0669	0.1560	0.4579	0.9093	0.2875	0.1875	0.1697	0.6326	0.0000	0.0000	0.2277
凭证管理	0.8634	0.8472	0.8887	0.0000	0.0326	0.3413	0.0000	0.8525	0.2875	0.2875	0.1195	0.8784	0.0000	0.2530	0.0000
生产监控	0.4178	0.3011	0.2087	0.8995	0.0784	0.8277	0.8963	0.2842	0.2625	0.2625	0.8440	0.6530	0.2940	0.3659	0.2124
计划调度	0.5550	0.4099	0.0000	0.9440	0.0530	0.8423	0.8963	0.1540	0.2625	0.3625	0.8940	0.7659	0.3989	0.2274	0.3491
物料管理	0.6204	0.4582	0.3091	0.9940	0.0659	0.9124	0.8205	0.1093	0.3776	0.2364	0.8989	0.6274	0.3684	0.1281	0.1349
计量管理	0.0000	0.0000	0.1255	0.1989	0.0274	0.1491	0.3697	0.2326	0.8413	0.8495	0.3525	0.2697	0.8427	0.8660	0.8511
质量检查	0.2010	0.0000	0.1251	0.2684	0.0281	0.2349	0.3195	0.1784	0.7277	0.7963	0.2842	0.7195	0.8881	0.8277	0.9444
过程控制	0.1568	0.0000	0.2404	0.2427	0.0660	0.2511	0.1440	0.1530	0.7423	0.8963	0.3540	0.6440	0.9017	0.9365	0.9673
设备档案	0.0000	0.1057	0.1912	0.2881	0.0277	0.2444	0.1940	0.1659	0.9124	0.8205	0.0000	0.6940	0.8426	0.9319	0.8020
预修预检	0.0000	0.1057	0.1375	0.1017	0.0365	0.2673	0.1989	0.1274	0.9491	0.9010	0.0000	0.7989	0.8697	0.9285	0.8361

为了简化知识网  $W_1, W_2, \dots, W_{15}$  相似度的计算过程, 首先把所有知识网通过将其原来缺少的底层知识点当成零完善度知识点来归一化成具有相同的底层知识点, 再对式(1)作一些化简和假设, 令  $\gamma_{v_i} = \gamma_{w_i}$ , 即知识网最底层知识点的层次数相同; 令  $l_{v_i} = l_{w_i} = l_{v_i, W}$ , 即匹配度为 1, 则式(1)化简

为

$$\text{sim}(V, W) = \frac{2 \sum_{i=1}^m \gamma_{v_i} \cdot l_{v_i, W} \cdot \kappa_{v_i}}{\sum_{i=1}^m \alpha_{v_i} l_{v_i} \mu_{v_i} + \sum_{j=1}^n \beta_{w_j} l_{w_j} \mu_{w_j}} = \frac{2 \sum_{i=1}^m \kappa_{v_i}}{\sum_{i=1}^m (\mu_{v_i} + \mu_{w_i})}. \quad (7)$$

由式(7)求得15个知识网的相似度构建的模糊关联矩阵(因为对称性只给出矩阵的下半部分),见表2.

表2 知识网的模糊关联矩阵

Table 2 The fuzzy relational matrix of KMs

	$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	$W_8$	$W_9$	$W_{10}$	$W_{11}$	$W_{12}$	$W_{13}$	$W_{14}$	$W_{15}$
$W_1$	1.0000	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$W_2$	0.8278	1.0000	-	-	-	-	-	-	-	-	-	-	-	-	-
$W_3$	0.7346	0.7780	1.0000	-	-	-	-	-	-	-	-	-	-	-	-
$W_4$	0.5519	0.4500	0.4260	1.0000	-	-	-	-	-	-	-	-	-	-	-
$W_5$	0.1835	0.2056	0.2409	0.1928	1.0000	-	-	-	-	-	-	-	-	-	-
$W_6$	0.6117	0.5177	0.5020	0.8826	0.2049	1.0000	-	-	-	-	-	-	-	-	-
$W_7$	0.5934	0.5019	0.4604	0.8882	0.1883	0.8371	1.0000	-	-	-	-	-	-	-	-
$W_8$	0.7605	0.7771	0.8719	0.4331	0.2644	0.4920	0.4967	1.0000	-	-	-	-	-	-	-
$W_9$	0.3893	0.3892	0.4380	0.4573	0.1573	0.5048	0.4858	0.4442	1.0000	-	-	-	-	-	-
$W_{10}$	0.3613	0.3587	0.4007	0.4393	0.1587	0.4990	0.4597	0.4239	0.9367	1.0000	-	-	-	-	-
$W_{11}$	0.5822	0.4199	0.3709	0.8683	0.1902	0.8546	0.8544	0.3955	0.4562	0.4500	1.0000	-	-	-	-
$W_{12}$	0.6590	0.5890	0.5833	0.6243	0.1347	0.6765	0.6612	0.5868	0.7465	0.7266	0.6120	1.0000	-	-	-
$W_{13}$	0.3090	0.3018	0.3151	0.4507	0.1301	0.4585	0.4715	0.3277	0.8894	0.9244	0.4402	0.6927	1.0000	-	-
$W_{14}$	0.2883	0.2640	0.3298	0.3775	0.1398	0.4377	0.3990	0.3841	0.9038	0.9281	0.3905	0.6751	0.9077	1.0000	-
$W_{15}$	0.2816	0.2718	0.3309	0.4257	0.1549	0.4186	0.4469	0.3676	0.8723	0.9271	0.4018	0.6746	0.9257	0.8937	1.0000

4.2 知识网的聚类与选择(Clustering and selection of KMs)

在Intel酷睿双核2.0G, 内存2G计算机上, MATLAB7.5.0环境下编程, 对模糊关联矩阵进行分解.  $G$ 的初值为随机赋值的矩阵, 计算式(4)-(5), 取  $c = 3, \beta = 0.07$ , 迭代500次, 聚类结果见表3.  $W_4, W_6, W_7, W_{11}$ 属于第1类,  $W_9, W_{10}, W_{13}, W_{14}, W_{15}$ 属于第2类,  $W_1, W_2, W_3, W_8$ 属于第3类. 这些知识网关于归属类都具有较高的类隶属度(大于0.9), 说明  $c = 3$  是一个合理的分类数. 比较表1, 第1类知识网在生产管理方面具有较高的功能完善度, 而在其他方面功能完善度偏低; 第2类知识网在质量、设备方面具有较高的功能完善度; 第3类知识网在财务管理方面具有较高的功能完善度. 可以看到知识网的分类结果与表1的完善度数值特点完全吻合, 说明文中的相似度函数能够合理地反映知识网间的相似程度, 模糊关联聚类能够把具有相似功能特点的知识网归类在一起. 另外,  $W_5$ 关于3个类的类隶属度都很低, 这是因为它的各项完善度数值都很低, 比较后将其划分为第3类.  $W_{12}$ 被划分为第2类, 但它对3个类的类隶属度相差不是很大, 因为它的各项功能完善度数

值并不突出. 由表3, 不需要用户的任何参与, 知识网库中的知识网进行了自动归类, 此时可以调出每个类中类隶属度高的知识网作为参考知识网, 如  $W_4, W_8, W_{10}$ , 协助用户进行知识网的选择. 若用户为能够表达具体需求的高素质用户, 其目标知识网为  $\bar{W}$ , 并且  $\bar{W}$  与已有知识网最底层知识点的层次数相同, 匹配度为1, 则式(6)化简为

$$\text{sim}(\bar{W}, W) = \frac{2 \sum_{i=1}^m \kappa_{\bar{w}_i}}{\sum_{i=1}^m (\mu_{\bar{w}_i} + \mu_{W(\bar{w}_i)})}$$

设  $\bar{W}$  针对表1中各功能的功能完善度为

$$(0.2000, 0.4000, 0.9500, 0.9500, 0.9000, 0.0000, 0.0000, 0.0000, 0.0000),$$

则  $\text{sim}(\bar{W}, W_4) = 0.9080, \text{sim}(\bar{W}, W_8) = 0.4020, \text{sim}(\bar{W}, W_{10}) = 0.5643$ . 根据最大类隶属度原则,  $\bar{W}$  属于第1类. 第1类的知识网包括  $W_4, W_6, W_7, W_{11}$ , 分别计算  $\bar{W}$  与它们的相似度.

$$\begin{aligned} \text{sim}(\bar{W}, W_4) &= 0.9080, \text{sim}(\bar{W}, W_6) = 0.9467, \\ \text{sim}(\bar{W}, W_7) &= 0.8694, \text{sim}(\bar{W}, W_{11}) = 0.9251. \end{aligned}$$

可见  $W_6$  为最符合用户需求的知识网.

表3 知识网的类隶属关系

Table 3 The membership of KMs' class

	$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	$W_8$	$W_9$	$W_{10}$	$W_{11}$	$W_{12}$	$W_{13}$	$W_{14}$	$W_{15}$
第1类	0.4054	0.1398	0.0000	0.9602	0.0642	0.9351	0.9390	0.0124	0.2573	0.2421	0.9487	0.5506	0.2844	0.1884	0.2249
第2类	0.2853	0.3095	0.3799	0.3635	0.2058	0.4042	0.3911	0.4065	0.9775	1.0000	0.3642	0.7411	0.9832	0.9835	0.9795
第3类	0.9010	0.9327	0.9440	0.4770	0.3670	0.5493	0.5279	0.9488	0.2719	0.2234	0.4442	0.5839	0.1195	0.1340	0.1300

### 4.3 算法比较(Algorithm comparison)

利用模糊C均值方法(fuzzy C-means, FCM)<sup>[6]</sup>对表1知识网数据进行聚类, 模糊化因子 $m = 2$ , 分类数为3, 最大迭代次数为100, 迭代终止条件为隶属度最小变化量 $1e - 5$ , 隶属度见表4. 比较表3与表4, 聚类结果完全相同, 这是因为表1数据是一组具有明显结构特征的数据. 但知识网是制造系统的抽象模型, 它拥有相应制造系统的所有知识. 随

着知识网库中知识网的增多, 整个知识网集合的结构会越来越复杂, 知识网特征空间的维数也会越来越高. 在这种情况下, 本文的方法就会显得更为有效. 表5与表6为表1扩充为 $20 \times 15$ 任意数据后, 利用本文方法和FCM分别获得的隶属度值, 很明显表5的数据更为有效. 因此构造知识网相似性度量的模糊关联聚类方法较FCM法更能够解决知识网高维和集合结构复杂的问题.

表4 FCM方法的隶属度

Table 4 Membership degree of FCM

	$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	$W_8$	$W_9$	$W_{10}$	$W_{11}$	$W_{12}$	$W_{13}$	$W_{14}$	$W_{15}$
第1类	0.1343	0.0467	0.0387	0.9581	0.3152	0.9254	0.9281	0.0406	0.0272	0.0053	0.9331	0.3245	0.0265	0.0206	0.0255
第2类	0.0456	0.0187	0.0269	0.0169	0.1915	0.0268	0.0276	0.0280	0.9477	0.9899	0.0255	0.3711	0.9548	0.9614	0.9528
第3类	0.8201	0.9346	0.9344	0.0249	0.4933	0.0478	0.0443	0.9315	0.0251	0.0048	0.0414	0.3044	0.0187	0.0181	0.0217

表5 本文方法的隶属度

Table 5 Membership degree of this paper's method

	$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	$W_8$	$W_9$	$W_{10}$	$W_{11}$	$W_{12}$	$W_{13}$	$W_{14}$	$W_{15}$
第1类	0.9337	0.5921	0.2497	0.6715	0.4644	0.3919	0.8457	0.6095	0.8856	0.0000	0.5651	0.8487	0.5602	0.8827	0.3440
第2类	0.5661	0.9094	0.9558	0.6860	1.0000	0.7283	0.4636	0.8636	0.5019	0.6645	0.7770	0.7440	0.9511	0.7300	0.5968
第3类	0.6548	0.4941	0.7161	0.8243	0.5366	0.8494	0.9121	0.5989	0.8008	0.9773	0.8144	0.5399	0.4432	0.7161	0.9304

表6 FCM方法的隶属度

Table 6 Membership degree of FCM

	$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	$W_8$	$W_9$	$W_{10}$	$W_{11}$	$W_{12}$	$W_{13}$	$W_{14}$	$W_{15}$
第1类	0.3353	0.3325	0.3318	0.3327	0.3325	0.3338	0.3343	0.3333	0.3350	0.3323	0.3324	0.3346	0.3328	0.3350	0.3329
第2类	0.3329	0.3334	0.3336	0.3335	0.3335	0.3333	0.3332	0.3333	0.3330	0.3336	0.3335	0.3330	0.3334	0.3329	0.3336
第3类	0.3319	0.3340	0.3346	0.3338	0.3340	0.3329	0.3325	0.3334	0.3320	0.3341	0.3341	0.3324	0.3339	0.3321	0.3336

## 5 结论(Conclusions)

随着知识化制造技术的推广, 知识网的选择是用户必须面临的问题. 如何以方便快捷的方式将具有代表性的参考知识网推荐给用户, 本文进行了一些尝试. 针对知识网特征空间高维和集合结构复杂的特点, 提出基于模糊关联聚类的知识网选择方法. 从质、量和结构层次3个方面综合构造两两知识网的相似度, 具有反映知识网运算规律的特征, 并将原来的高维聚类空间转换到由相似度构成的聚类空间, 克服了样本集稀疏分布的问题. 聚类结果中类隶属度高的知识网为低素质用户提供了参考依据, 同时也为高素质用户缩小了选择范围. 这种通过挖掘知识网间关系去帮助用户选择知识网的方法为知识网选择问题提供了一种新思路, 它使得知识网间的关系变得清晰, 并辅助用户理解知识网和明确需求. 但文中方法会造成相似度矩阵中存在大量重复数据, 需要对矩阵

进行化简, 以及分类数的确定和聚类有效性, 都将做进一步的研究.

## 参考文献(References):

- [1] 严洪森, 刘飞. 知识化制造系统—新一代先进制造系统[J]. 计算机集成制造系统, 2001, 7(8): 7-11.  
(YAN Hongsen, LIU Fei. Knowledgeable manufacturing system — a new kind of advanced manufacturing system [J]. *Computer Integrated Manufacturing Systems*, 2001, 7(8): 7-11.)
- [2] YAN H S. A new complicated-knowledge representation approach based on knowledge meshes [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2006, 18(1): 47-62.
- [3] 薛朝改, 严洪森. 基于用户功能需求的知识网的自动生成研究[J]. 控制与决策, 2005, 20(9): 996-1001.  
(XUE Chaogai, YAN Hongsen. Research of automatic construction of the knowledge mesh based on the user's functional requirement [J]. *Control and Decision*, 2005, 20(9): 996-1001.)
- [4] 丁雪峰, 严洪森, 薛朝改. 基于近似匹配的知识化制造系统自重构研究[J]. 控制与决策, 2008, 23(1): 70-74.  
(DING Xuefeng, YAN Hongsen, XUE Chaogai. Self-reconfiguration of knowledgeable manufacturing system based on approximate match [J]. *Control and Decision*, 2008, 23(1): 70-74.)

- [5] 杨人子, 严洪森. 基于信息粒度的知识网的模糊分类与检索方法 [J]. 自动化学报, 2011, 37(5): 584 – 595.  
(YANG Renzi, YAN Hongsen. The method of fuzzy classification and searching for knowledge meshes based on information granularity [J]. *Acta Automatica Sinica*, 2011, 37(5): 584 – 595.)
- [6] BEZDEK J C. *Pattern Recognition with Fuzzy Objective Function Algorithm* [M]. New York: Plenum Press, 1981.
- [7] GAN G, WU J. A convergence theorem for the fuzzy subspace clustering (FSC) algorithm [J]. *Pattern Recognition*, 2008, 41(6): 1939 – 1947.
- [8] ABONYI J, SZEIFERT F. Supervised fuzzy clustering for the identification of fuzzy classifiers [J]. *Pattern Recognition Letters*, 2003, 24(14): 2195 – 2207.
- [9] RUNKLER T A, BEZDEK J C. Web mining with relational clustering [J]. *International Journal of Approximate Reasoning*, 2003, 32(2): 217 – 236
- [10] 唐成龙, 王石刚. 基于数据间内在关联性的自适应模糊聚类模型 [J]. 自动化学报, 2010, 36(11): 1544 – 1556.  
(TANG Chenglong, WANG Shigang. Adaptive fuzzy clustering model based on internal connectivity of all data points [J]. *Acta Automatica Sinica*, 2010, 36(11): 1544 – 1556.)
- [11] 潘天红, 李少远. 基于模糊聚类的PWA系统的模型辨识 [J]. 自动化学报, 2007, 33(3): 327 – 330.  
(PAN Tianhong, LI Shaoyuan. Model identification of piecewise affine (PWA) systems based on fuzzy cluster [J]. *Acta Automatica Sinica*, 2007, 33(3): 327 – 330.)
- [12] 贾文君, 柴天佑. 稀土串级萃取分离过程元素组分含量的多模型软测量 [J]. 控制理论与应用, 2007, 24(4): 569 – 580.  
(JIA Wenjun, CHAI Tianyou. Soft-sensor of element component content based on multiple models for the rare earth cascade extraction process [J]. *Control Theory & Applications*, 2007, 24(4): 569 – 580.)
- [13] 吴敏, 丁雷, 曹卫华, 等. 铅锌烧结过程烧穿点的集成预测模型 [J]. 控制理论与应用, 2009, 26(7): 739 – 744.  
(WU Min, DING Lei, CAO Weihua, et al. An integrated prediction model for burn-through-point in lead-zinc sintering process [J]. *Control Theory & Applications*, 2009, 26(7): 739 – 744.)
- [14] 徐超, 张培林, 任国全, 等. 基于改进半监督模糊C-均值聚类的发动机磨损故障诊断 [J]. 机械工程学报, 2011, 47(17): 55 – 60.  
(XU Chao, ZHANG Peilin, REN Guoquan, et al. Engine wear fault diagnosis based on improved semi-supervised fuzzy C-means clustering [J]. *Journal of Mechanical Engineering*, 2011, 47(17): 55 – 60.)
- [15] 周敬泉, 颜春兰, 袁鹏, 等. 用模糊聚类法研究中药成分特征谱 [J]. 控制理论与应用, 2004, 21(4): 569 – 573.  
(ZHOU Jingquan, YAN Chunlan, YUAN Peng, et al. Research and analysis of componential characteristic spectrum of Chinese traditional medicine with fuzzy clustering [J]. *Control Theory & Applications*, 2004, 21(4): 569 – 573.)

### 作者简介:

**杨人子** (1977–), 女, 博士, 讲师, 研究领域为知识化制造系统,

E-mail: yrz@seu.edu.cn;

**严洪森** (1957–), 男, 博士, 教授, 博士生导师, 研究领域为知识化制造系统、计算机集成制造系统和生产计划与调度, E-mail: hsyang@seu.edu.cn.