Advanced Engineering Forum *Online: 2011-09-09*
ISSN: 2234-991X, Vol. 1, pp 247-251
doi:10.4028/www.scientific.net/AEF.1.247

Fuzzy Relational Clustering Based on Knowledge Mesh and Its Application

YANG Ren-zi [1,2], YAN Hong-sen [2]

[1] Department of Mathematics, Southeast University, Nanjing 210096, China;

[2] School of Automation, Southeast University, Nanjing 210096, China

yrz@seu.edu.cn

Keywords: knowledge mesh; fuzzy relational clustering; similarity; matrix decomposed

Abstract. A selection method of knowledge meshes based on fuzzy relational clustering is proposed. Considering the perfection degree, the matching degree among knowledge meshes and the level frame of knowledge mesh, the similarity function is defined. Its properties are proved. The similarity values between knowledge meshes are regarded as clustering data. The fuzzy relational matrix is constructed and decomposed. The knowledge meshes with high membership in each class are regarded as referenced knowledge meshes. Or the knowledge meshes in the class are further chosen according to user's needs. Finally the example shows that the method is effective.

Introduction

Knowledgeable manufacturing[1] transform all types of advanced manufacturing modes into corresponding knowledge meshes (KMs)[2] and included them in knowledgeable manufacturing system (KMS), which selects and uses the most appropriate combination of the modes or the best one when necessary. KMS is characterized with self-adaption, self-learning, self-evolution, self-reconfiguration, self-training and self-maintenance, which are named 'six-self-characteristic'. The selection of KM is the first facing problem in the application of most KMS's six-self-characteristic technologies. With the consummate of six-self-characteristic technologies, more and more KMs, which must exist the same or similar parts among these KMs, are stored in KM base. When user search KM in KM base, it is sometimes difficult to choice, especially for the low-quality user who can not clearly describe specific needs. If system can automatically give some representative referenced KMs, it can help user clear needs and quickly find the required KM.

Clustering is an important tool for data analysis, unsupervised learning, data graining and information compression. The fuzzy clustering, such as fuzzy c-means clustering(FCM)[3], supervised fuzzy clustering(SFC)[4], fuzzy relational clustering(FRC)[5], fuzzy kernel clustering with outliers(FKCO)[6], allow partial membership by fuzzy concept and make the result of clustering more conform practical situation. But it isn't reported about using clustering to solve the selection of KMs in KMS. Some literatures have the similar ideas. Shen and Chen[7] discovery the generic model by fuzzy clustering in the level of particular reference model for knowledge management in enterprise modeling. Then constructing the corresponding relationship between the new model object and generic model, and predicting the new mode, solve the retrieval problem. But the clustering objects are reference process models and the clustering data gave by domain experts are the results of comparing each model, which entirely depend on experts' experience. But each KM represents a real manufacturing system which contains a lot of data and the inherent information. It is different from the previous clustering objects. In addition, the clustering methods through comparing the relational degree between the classification samples and standard samples aren't suitable for the clustering of KMs, because the corresponding feature space is difficult to build, and the high-dimensional number and the limited sample set make it difficult to describe the construction of data.

This paper proposes a selection method of KMs based on fuzzy relational clustering whose clustering data are the similarity degree. Construct the sample-class relation by decomposing fuzzy relation. The KMs having high membership in each class have reference value. The comparison of them and target KM narrow the scope of user's choice. User only select the KMs in certain class.

Fuzzy relational clustering based on KM

Similarity Model. Any KM is an abstract mode of advanced manufacturing system in KMS. It is a set including knowledge point(KP), message relationship, function, etc, where the functions of father KP must include the functions of child KP. So the following discusses are about the lowest knowledge points(LKP). And the user's needs are mainly the functional requirements. So the first definition is the functional similarity of the lowest KP.

Definition 1 Suppose the LKP sets of KM V and KM W are $P_V = \{p_{v_1}, p_{v_2}, \cdots, p_{v_m}\}$ and $P_W = \{p_{w_1}, p_{w_2}, \cdots, p_{w_n}\}$. The function number of KP p_{v_i} is l_{v_i} ($i = 1,2,\cdots,m$) and the function number of KP p_{w_j} is l_{w_j} ($j = 1,2,\cdots,n$). The number of the function which is the same as the function of p_{v_i} in P_W is $l_{v_i,W}$ and the number of the function which is the same as the function of p_{w_j} in P_V is $l_{w_j,V}$. Then $f(P_V, P_W) = \dfrac{\sum_{i=1}^{m} l_{v_i,W} + \sum_{j=1}^{n} l_{w_j,V}}{\sum_{i=1}^{m} l_{v_i} + \sum_{j=1}^{n} l_{w_j}}$ on $P_V \times P_W$ is defined as the matching degree of KM V and KM W. In fact $\sum_{i=1}^{m} l_{v_i,W} = \sum_{j=1}^{n} l_{w_j,V}$. It is simply defined as $f(P_V, P_W) = \dfrac{2\sum_{i=1}^{m} l_{v_i,W}}{\sum_{i=1}^{m} l_{v_i} + \sum_{j=1}^{n} l_{w_j}}$.

It is remarkable that definition 1 reflect the quantity of similarity in KM's function. When user needs certain functions, the KM satisfying user's needs is more than just having those functions in practical situation. User may hope to improve functions to meet the further development. Or user only needs the basic functions to save the cost. In addition, we also know that the representation of KP is the same but the corresponding contents in practical system may vary a little. This kind of difference between KMs is taken as the perfection degree of KM for comparison of their "quality". Using a fuzzy set for the definition is a more practical choice. And the perfection degree of father KP's function is gotten by child KP's.

Definition 2 Suppose the LKP's set of KM W is $P = \{p_1, p_2, \cdots, p_n\}$. A fuzzy set on P is defined. $\mu(p_i): P \to [0,1], x_i \in P$. $\mu(p_i)$ is the functional perfection degree of KP p_i and simplified as μ_{p_i}. $\mu = (\mu_{p_1}, \mu_{p_2}, \cdots, \mu_{p_n})$ is the functional perfection degree of KM W. $\mu_{p_i} = 0$ denotes that the element p_i does not exist in KM. The larger the value of p_i, the more perfect its function.

According to definition 2, there is one-to-one correspondence between each KM and its fuzzy set μ. It is assumed that each KM corresponds to one fuzzy vector μ, $\mu \in [0,1]^n$. Definition 1 and definition 2 reflect the similariy from two aspects of functional quality and quantity. Though the matching degree and perfection degree of two KM are completely the same, their corresponding structures of KM may differ. The level number of KP can reflect the KM's structure. Introduce it into the definition of KM's similarity. So taking quality, quantity and structure into account, the similarity degree is defined.

Definition 3 The similarity degree of KM V and W is

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

The LKP of V and W is $P_V = \{p_{v_1}, p_{v_2}, \cdots, p_{v_m}\}$ and $P_W = \{p_{w_1}, p_{w_2}, \cdots, p_{w_n}\}$; μ_{v_i}, l_{v_i}, α_{v_i} is the functional perfection degree, functional number and level number of KP p_{v_i} respectively. $\mu_{w_j}, l_{w_j}, \beta_{w_j}$ is the functional perfection degree, functional number and level number of KP p_{w_j} respectively. $l_{v_i,W}$ is the same functional number as that of KP p_{v_i} in P_W. $\beta_{W(v_i)}, \mu_{W(v_i)}$ is the level number and functional perfection degree of the KPs relating to those functions. $l_{w_j,V}$ is the same functional number as that of KP p_{w_j} in P_V . $\alpha_{V(w_j)}$、$\mu_{V(w_j)}$ is the level number and functional perfection degree of the KPs relating to those functions. $\gamma_{v_i} = \min\{\beta_{W(v_i)}, \alpha_{v_i}\}$, $\gamma_{w_j} = \min\{\alpha_{V(w_j)}, \beta_{w_j}\}$, $\kappa_{v_i} = \min\{\mu_{W(v_i)}, \mu_{v_i}\}$, $\kappa_{w_j} = \min\{\mu_{V(w_j)}, \mu_{w_j}\}$.

It is need to explain that the level number is increased from the root KP, i.e. the level number of father KP is less than that of child KP. If the KP having the same function as KP p_{v_i} 's is not the only KP, divide $l_{v_i,W}$ into sum formula in order to correspond each item with the only KP in P_W . Then operate it. So the KPs having the same function in P_V and P_W can be seen as one-to-one. Then $\sum_{i=1}^{m}\gamma_{v_i}\cdot l_{v_i,W}\cdot\kappa_{v_i} = \sum_{j=1}^{n}\gamma_{w_j}\cdot l_{w_j,V}\cdot\kappa_{w_j}$ and formula (1) is simplified as $sim(V,W) = \dfrac{2\sum_{i=1}^{m}\gamma_{v_i}\cdot l_{v_i,W}\cdot\kappa_{v_i}}{\sum_{i=1}^{m}\alpha_{v_i}l_{v_i}\mu_{v_i} + \sum_{j=1}^{n}\beta_{w_j}l_{w_j}\mu_{w_j}}$.

Theorem 1 (1) $0 \le sim(V,W) \le 1$. When KM V and W have the identical function, and the functional perfection degree and level number of KP corresponding with those functions are also the same, $sim(V,W) = 1$. When KM V and W don't have any same function, $sim(V,W) = 0$. Especially, $sim(V,V) = 1$. (2) $sim(V,W) = sim(W,V)$. (3) KM V include all function of KM W . The functional perfection degree and level number of LKP corresponding with those functions are the same. And those functions of V which is different from W 's is also different from any function of X . Then $sim(V,X) \le sim(W,X)$. (4) KM V and W don't have the same function. KM X include all function of KM V .The functional perfection degree and level number of KP corresponding with those functions are the same. Then $sim(W,X) \le sim(V+W,X)$. (5) If the functions of KM X include both V 's and W 's, and the functional perfection degree and level number of KP corresponding with those functions are the same, $sim(V \cap W, X) \le \min\{sim(V,X), sim(V,X)\}$.

Proof omitted.

Decomposition of Fuzzy Relational Matrix. Suppose that there are $V_1, V_2, \cdots V_N$ KMs in KM base. Compare in pairwise comparision and get $\frac{N(N-1)}{2}$ similarity values by formula (1).They make up a $N \times N$ matrix which is called fuzzy relational matrix. Let $R = [r_{ij}], i, j = 1,2,\cdots N$,where $r_{ij} = sim(V_i, V_j)$. R has reflexivity and symmetry properties, i.e. $r_{ii} = 1$ and $r_{ij} = r_{ji}$.

Constructing the matrix $G = [g_{ij}], i = 1,2,\cdots N, j = 1,2,\cdots,c, c < N$, satisfies $R = G \circ G^T$, where G^T denotes the transposition of G , and '$\circ$' denotes composition operator of relations. This means that a KM-class relation is gotten by decomposing R . The operator is s-t convolution of fuzzy set. This problem is transformed into finding $G = [g_{ij}]$ to satisfy $r_{ij} = \underset{k=1}{\overset{c}{S}}(t(g_{ik}, g_{jk}))$. Let $t(x,y) = xy$, $S(x,y) = x + y - xy$ in this paper. Find an approximate solution G to minimize Q by form $Q = \left\| R - G \circ G^T \right\|^2$, i.e. $Q = \sum_{i=1}^{N}\sum_{j=1}^{N}[r_{ij} - \underset{k=1}{\overset{c}{S}}(t(g_{ik}, g_{jk}))]^2$. G is calculated by gradient method, $G = G - \beta \cdot \nabla_G Q$, where $\beta (> 0)$ denotes learning rate.

Selection method of KMs

Each KM in KM base is affiliated with a class by the above method and maximum membership principle. The KMs having the highest membership in class are regarded as referenced KM for user. But these KMs having the highest membership, which can reduce the range of choice, is not necessarily the best choice for the high-quality user. Suppose that the target KM $\overline{W}$ satisfy user's needs. Modify formula (1) and get formula (2), where $\beta_{W(\overline{w}_i)}, l_{\overline{w}_i,W}, \mu_{W(\overline{w}_i)}$ is the level number, the same functional number and perfection degree of KP in KM W about the function of KP $p_{\overline{w}_i}$ in KM $\overline{W}$ .According to formula (2), calculate similarity of target KM and KM having the highest membership in each class. The bigger the value is, the more similar the target KM and a class's KM. Then compare all KMs of this class with the target KM. The KM having the biggest simlilarity is referenced KM.

$$sim(\overline{W},W)=\frac{\sum_{i=1}^{m}\gamma_{\overline{w}_i}\cdot l_{\overline{w}_i,W}\cdot\kappa_{\overline{w}_i}+\sum_{j=1}^{n}\gamma_{w_j}\cdot l_{w_j,\overline{W}}\cdot\kappa_{w_j}}{\sum_{i=1}^{m}\alpha_i l_{\overline{w}_i}\mu_{\overline{w}_i}+\sum_{i=1}^{m}\beta_{W(\overline{w}_i)}l_{\overline{w}_i,W}\mu_{W(\overline{w}_i)}} \tag{2}$$

Empirical analysis

Suppose that there are 15 KM $W_1,W_2,\cdots,W_{15}$ in KM base. Now cluster them and select referenced KMs for user. The perfection degrees of 15 KMs are directly given in table 1. These values can be gotten by experts grading method. The financial management of W_1,W_2,W_3,W_8 , the production management of W_4,W_6,W_7,W_{11} , the quality and equipment management of W_9,W_{10},W_{13}, W_{14},W_{15} have higher perfection degrees. But all aspects of W_5 have lower values and those of W_{12} have no prominent values. The data present four obvious classes.

Table 1 The perfection degree 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}
Financial management	Enter-in-ledger report	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
	Credentials anagement	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
Production management	Production monitoring	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
	Planning scheduling	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
	Materials management	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
Quality management	Measure management	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
	Quality inspection	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
	Process control	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
Equipment management	Equipment file	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
	Technique state	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

Simplify the calculation process of fuzzy relational matrix because of the lack of space. Transform all KMs into KMs having the same LKPs by regarding the lacking LKPs as KPs of zero perfection degree. Then make some simplifications and assumptions for formula (1). Let $\gamma_{v_i}=\gamma_{w_i}$, i.e, the level number of LKP is the same. And Let $l_{v_i}=l_{w_i}=l_{v_i,W}$, i.e. the matching degree is 1. Formula (1) is simplified as formula (3). Get the fuzzy relational matrix of 15 KMs by formula (3) in table 2. (Only give the lower part of fuzzy relation matrix because of symmetry.)

$$sim(V,W)=\frac{2\sum_{i=1}^{m}\gamma_{v_i}\cdot l_{v_i,W}\cdot\kappa_{v_i}}{\sum_{i=1}^{m}\alpha_{v_i}l_{v_i}\mu_{v_i}+\sum_{j=1}^{n}\beta_{w_j}l_{w_j}\mu_{w_j}}=\frac{2\sum_{i=1}^{m}\kappa_{v_i}}{\sum_{i=1}^{m}(\mu_{v_i}+\mu_{w_i})}, \tag{3}$$

Decompose the fuzzy relational matrix. The initial value of G is the random assignment of the matrix. Let $c=3$, β =0.07. Iterate 500 times. The result of clustering is given in table 3. It is shown that the result of clustering exactly matches the characteristic of table 1.Regard KMs having higher membership as the referenced KM, such as W_4,W_8,W_{10} . If user is the high-quality user, his target KM is $\overline{W}$, whose level number is the same as that of existing KMs' LKP. Then formula (2) is simplified as $sim(\overline{W},W)=2\sum_{i=1}^{m}\kappa_{\overline{w}_i}\Big/\sum_{i=1}^{m}(\mu_{\overline{w}_i}+\mu_{W(\overline{w}_i)})$. Suppose that the perfection degree of $\overline{W}$ for each function in table 1 is {0.2000,0.4000,0.9500,0.9500,0.9000,0.0000, 0.0000,0.0000,0.0000}. Then

Table 2 The fuzzy relation 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

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}
The first class	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
The second class	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
The third class	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

$sim(\overline{W},W_4)=0.9080, sim(\overline{W},W_8)=0.4020$ and $sim(\overline{W},W_{10})=0.5643$. According to the principle of maximum membership, $\overline{W}$ belongs to the first class. Calculate the similarity degrees of $\overline{W}$ and them. And $sim(\overline{W},W_4)=0.9080$, $sim(\overline{W},W_6)=0.9467$, $sim(\overline{W},W_7)=0.8694$, $sim(\overline{W},W_{11})=0.9251$. So W_6 is the best selection for user.

Conclusion

With the promotion of KMS technologies, the selection of KM must be faced for user. How to select some representative referenced KMs for user by the convenient and quick way, some try are made in this paper. Forming a reasonable similarity model is the precondition of clustering for KMs. The similarity model from the quality, quantity and structure aspects can completely reflect on the similarity among KMs. It makes the original clustering space convert to that constituted by similarity. It overcomes the sparse distribution of high-dimensional sample sets. The KMs having high membership help user narrow the choice scope. It makes the relations of KMs more clear. It also helps user clear needs and make the right choice.

References

[1] H. S. Yan, F. Liu. Knowledgeable manufacturing system——a new kind of advanced manufacturing system. Computer Integrated Manufacturing Systems, 7(2001) 7-11.

[2] H. S. Yan. A new complicated-knowledge representation approach based on knowledge meshes. IEEE Transactions on Knowledge and Data Engineering, 18(2006) 47-62.

[3] J. C. Bezdek. Pattern recognition with fuzzy objective function algorithm. New York, Plenum Press, 1982.

[4] J. Abonyi, F. Szeifert. Supervised fuzzy clustering for the identification of fuzzy classifiers. Pattern Recognition Letters, 24(2003) 2195-2207.

[5] T. A. Runkler , J. C. Bezdek. Web mining with relational clustering. Journal of Approximate Reasoning, 32(2003) 217-236.

[6] H. B. Shen, S. T. Wang, X. J. Wu. Fuzzy kernel clustering with outliers. Journal of Software, 15(2004) 1021-1029.

[7] H. Shen, Y. L. Chen. Applying fuzzy neural network classifier to knowledge management in enterprise modeling. Systems Engineering Theory & Practice, 11(2003) 1-7.