World Academy of Science, Engineering and Technology
International Journal of Industrial and Manufacturing Engineering
Vol:6, No:4, 2012

Fuzzy Hierarchical Clustering Applied for Quality Estimation in Manufacturing System

Y. Q. Lv, C.K.M. Lee

***Abstract*—**This paper develops a quality estimation method with the application of fuzzy hierarchical clustering. Quality estimation is essential to quality control and quality improvement as a precise estimation can promote a right decision-making in order to help better quality control. Normally the quality of finished products in manufacturing system can be differentiated by quality standards. In the real life situation, the collected data may be vague which is not easy to be classified and they are usually represented in term of fuzzy number. To estimate the quality of product presented by fuzzy number is not easy. In this research, the trapezoidal fuzzy numbers are collected in manufacturing process and classify the collected data into different clusters so as to get the estimation. Since normal hierarchical clustering methods can only be applied for real numbers, fuzzy hierarchical clustering is selected to handle this problem based on quality standards.

***Keywords*—**Quality Estimation, Fuzzy Quality Mean, Fuzzy Hierarchical Clustering, Fuzzy Number, Manufacturing system

I. Introduction

QUALITY control and quality monitoring relies on pursuit of a full inter connection of all the activities in a firm. Such integrated cooperation system of all workers, employees from production, through administration, ordering of materials and sale, guarantees full efficiency of operation. And also allows to reach economic benefits and customer's satisfaction [1]. Generally, companies always need to do the quality estimation to determine the quality level of a certain product so the corresponding quality control and improvement activities can be applied accordingly. Currently, most companies rely on the experience and intuition of highly skilled engineers or experts, which is human-interfered and subjective. On the other hand, in the real life situation, the collected data may be vague and not precise. The concept of fuzzy set theory to meet these problems is introduced by Zadeh 1965[2]. Many researchers have applied this concept into different areas [3-6] such as production workflow enhancement, productivity prediction as well as system quality management. In 1978, Dubois and Prade defined any of the fuzzy numbers as a fuzzy subset of the real line [7]. Fuzzy number cannot be easily evaluated and analyzed as real numbers. In this paper, one bottom-top fuzzy clustering method is adopted to classify the sampled fuzzy number data which is fuzzy hierarchical clustering method. Hierarchical clustering is a very systematic algorithm with dendrogram showing the steps of clustering and divergence. Agglomerative nesting (AGNES) is one of the typical hierarchical clustering method using the agglomerative approach.

Y. Q. Lv is with School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore e-mail: lvya0001@e.ntu.edu.sg

AGNES can be described as a bottom-up approach. It starts clustering with a small piece of data object representing an individual cluster, followed by merging a series of clusters together. The cluster will increase continuously until the conditions for termination are met. The fuzzy hierarchical clustering just derives from AGNES. Another important issue in this research is the introduction of quality control means and standard which indicates the quality levels and represents certain linguistic quality estimation. It inspires the authors to apply clustering method based on fuzzy quality mean so as to obtain the quality estimation of products with fuzzy data attributes. This paper is organized as follows: Section 2 introduces the original Agglomerative nesting method. Section 3 depicts the methodology for the new fuzzy hierarchical clustering applied for trapezoidal fuzzy number based on fuzzy dissimilarity. Section 4 is a numerical experiment with the introduced clustering method. And the last section part is the conclusion.

II. Methodology

In this section, fuzzy hierarchical clustering will be applied for quality estimation. Nowadays, any product in manufacturing system gets certain quality standard to verify whether they are up to certain qualification or not. The quality of the products can be differentiated into various levels. Under this concern, quality standards for each level are introduced to help classify the collected data.

i. Fuzzy Hierarchical Clustering of Trapezoidal Fuzzy Number

A. Trapezoidal Fuzzy Number

In general, a generalized fuzzy number A is described as any fuzzy subset of the real line R, whose membership function μ_A satisfies the following conditions [7].

(1) μ_A is a continuous mapping from R to the closed interval $[0,1]$,
(2) $\mu_A(x) = 0, -\infty < x \le c$,
(3) $\mu_A(x) = L(x)$ is strictly increasing on $[c, a]$,
(4) $\mu_A(x) = w, a \le x \le b$,
(5) $\mu_A(x) = R(x)$ is strictly decreasing on $[b, d]$,
(6) $\mu_A(x) = 0, d \le x < \infty$,

where $0 < w \le 1$, a, b, c, and d are real numbers. We denote this type of generalized fuzzy number as $A = (c, a, b, d; w)_{LR}$. When $w = 1$, we denote this type of generalized fuzzy number as $A = (c, a, b, d)_{LR}$. When L(x) and R(x) are straight line, Then A is a trapezoidal fuzzy number, we denote it as(c, a, b, d):

$$\begin{cases} \mu_A(x) = 0, \quad x \le c \\ \mu_A(x) = L(x) = w\left(\frac{x-c}{a-c}\right)^k, \quad c < x \le a \\ \mu_A(x) = w, \quad a < x \le b \\ \mu_A(x) = R(x) = w\left(\frac{x-d}{b-d}\right)^k, \quad b < x \le d \\ \mu_A(x) = 0, \quad x > d \end{cases} \tag{1}$$

B. Graded Mean Integration Representation (GMIR)

The fuzzy numbers, with given fuzzy set and corresponding fuzzy membership grade, are not easy to compare based on the fuzzy form. Hence, it's necessary to defuzzify the fuzzy numbers under comparison. The concept of maximizing set to defuzzify fuzzy numbers has been introduced by Jain early in the 1970s [8], there are many other defuzzification methods such as weighted fuzzy mean, mean of maximum etc [9].

In this research, Graded Mean Integration Representation [10] is selected to represent fuzzy numbers because of its simplicity and accuracy. The GMIR can be described as follow [11]:

Based on section A, suppose L^{-1} and R^{-1} are inverse functions of functions L and R, respectively, and the graded mean h-level value of generalized fuzzy number $A = (c, a, b, d; w)_{LR}$ is $h[L^{-1}(h) + R^{-1}(h)]/2$ as Figure 2. Then the graded mean integration representation of generalized fuzzy number based on the integral value of graded mean h-levels is

$$P(A) = \frac{\int_0^w h\left(\frac{L^{-1}(h)+R^{-1}(h)}{2}\right)dh}{\int_0^w h dh}, \tag{2}$$

where h is between 0 and w, $0 < w \le 1$.

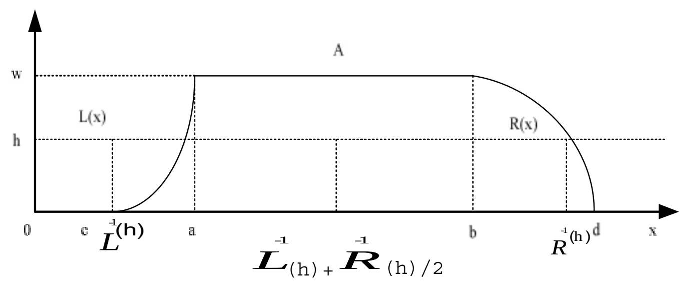


Fig. 2 The graded mean integration of a generalized fuzzy number A= (c, a, b, d; w)LR

Suppose $A = (c, a, b, d; w)$ is a trapezoidal fuzzy number, based on equation (1), to simplify the calculation, let k=1: Since,

$$L(x) = w\left(\frac{x-c}{a-c}\right), \qquad c < x \le a,$$

and

$$R(x) = w\left(\frac{x-d}{b-d}\right), \qquad b < x \le d,$$

then

$$L^{-1}(h) = c + \frac{(a-c)h}{w}, 0 \le h \le w,$$

$$R^{-1}(h) = d + \frac{(d-b)h}{w}, 0 \le h \le w,$$

and

$$\frac{L^{-1}(h) + R^{-1}(h)}{2} = \frac{c + d + (a - c - d + b)h/w}{2}$$

The graded mean integration representation of A is

$$P(A) = \frac{\int_0^w h\left(\frac{L^{-1}(h) + R^{-1}(h)}{2}\right)dh}{\int_0^w h dh} = \frac{c + 2a + 2b + d}{6}.$$

C. Fuzzy Dissimilarity

The distance of two real numbers is a new real number, while the distance of two fuzzy numbers should be another fuzzy set. In this paper, the fuzzy dissimilarity is defined as defuzzified fuzzy distance as follow [14]:

Definition 1: Let $A = (a_1, a_2, a_3, a_4)$, $B = (b_1, b_2, b_3, b_4)$ be two trapezoidal fuzzy numbers, and their Graded mean integration representation are P(A), P(B) respectively[12].
Assume
$s_i = (a_i - P(A) + b_i - P(B))/2$, i= 1, 2, 3, 4;
$c_i = |P(A) - P(B)| + s_i$, i=1, 2, 3, 4;
then the fuzzy distance of A,B is $C = (c_1, c_2, c_3, c_4)$
then the fuzzy dissimilarity of A, B $d_\wedge(A, B)$ is the Graded Mean integration representation of the new fuzzy number P(C).

Example:
There are 2 trapezoidal fuzzy numbers X (less or greater than 5)=(3, 4, 6, 7), Y (less or greater than 8) =(6,7, 9, 10)
then
s1=-2, s2=-1., s3=1. s4=2. Therefore.
c1=2,c2=3, c3=5, c4=6; So the fuzzy distance of A, B is C=(2, 3, 5, 6), We can say C as greater or less than 4.
$d_\wedge(X, Y)=3$

D. Fuzzy Hierarchical Algorithm

The generalized fuzzy hierarchical algorithm for trapezoidal fuzzy number can be summarized as follow:

1. Initially, each cluster contains one fuzzy number;
2. Calculate the fuzzy dissimilarity between each fuzzy number and create the dissimilarity matrix for fuzzy numbers;
3. Two "most similar" clusters are chosen at each step
4. The dissimilarity between two clusters can be estimated in various ways: single-linkage, complete-linkage, and average-linkage;
5. The steps 2 and 3 will be repeated until every object is in one cluster.

ii. Clustering into different quality levels based on quality mean standards

A. Fuzzy quality Mean

In manufacturing system production lines, the products parameters are inspected based on quality standards to exclude those unqualified items. As mentioned, the data represented the products may be not accurate but in fuzzy way, so Quality Mean Standard is introduced in term of Trapezoidal fuzzy number.

Definition 2: Fuzzy quality Mean is defined as the statement of the mean value for a certain quality level, represented in MQ. In this research the MQ is fuzzy number, so a linguistic label can be given for different MQ for easy delivery of information.

B. Fuzzy quality Mean help clustering products

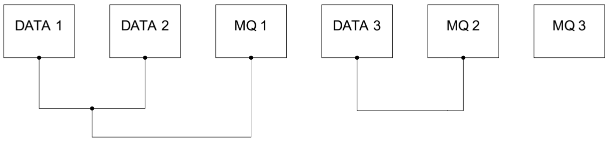
Assume for a certain product, there are three quality levels, and the quality mean standards for each level are MQ1, MQ2 and MQ3 with the linguistic labels are "good", "medium" and "bad" respectively.

Now a sample data representing this product is captured in term of trapezoidal fuzzy number. Based on fuzzy hierarchical clustering, the collected data can be classified based on its natural attributes but referring to the real condition, it's necessary to figure out which quality level the product belongs to.

Example: We randomly choose 3 same products P1 P2 P3, and the parameter attributes representing each product are DATA1, DATA2, and DATA3. The target is to estimate the quality of the products.

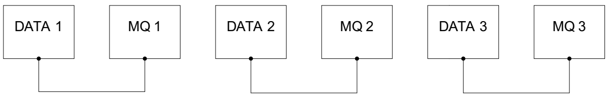
What we need to do is to regard different MQ and DATA as different cluster initially. Then use fuzzy hierarchical clustering to classify them. It's not difficult to know there are several cases can be obtained (some typical case shown as follow):

Case 1



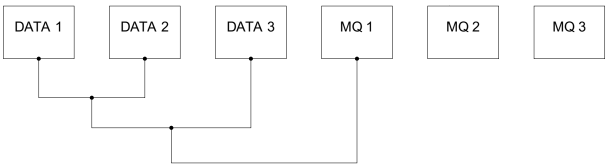
After this clustering, it can easily estimate that P1 and P2 are with "good" quality, while P3 is only in "medium" quality.

Case 2



After this clustering, it can easily estimate that P1 is with "good" quality, while P2 is only in "medium" quality and P3 "bad". So P3 needs rework.

Case 3



After this clustering, it can easily estimate that P1 P2 P3 are all with "good" quality, which is the best case for manufacturing process.

Based on the dendrogram of case 1 to 3 shown above, it's can be found the clustering for quality estimation is different from theoretical fuzzy clustering that all the object will merge into one cluster at the end. There is one important constraint for quality estimation with fuzzy hierarchical clustering MQs can never merge into one cluster because it may combine cluster 1 and cluster 2 to one cluster and the boundary of two clusters may not be clear after combination (Fig. 3 shows the constraint.).

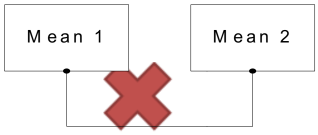


Fig. 3 Constraint of merging rules

III. Numerical Experiment

There are 3 same type products with the key parameter in term of trapezoidal fuzzy numbers PA (less or greater than 3)=(1, 2, 4, 5), PB (less or greater than 7)=(5, 6, 8, 9), PC (less or greater than 4)= (2, 3, 5, 6), For this product the quality levels are divided in level two with the given quality mean standards are MQ1 (less or greater than 5)=(3, 4, 6, 7), MQ2 (less or greater than 8) =(6,7, 9, 10), now use the proposed fuzzy hierarchical method to do the clustering:

i. Get the fuzzy distance matrix

TABLE I
Fuzzy Distance Matrix of 5 Trapezoidal Fuzzy Numbers

	PA(1,2,4,5)	PB(5,6,8,9)	PC(2,3,5,6)	MQ1(3,4,6,7)	MQ2(6,7,9,10)
PA(1,2,4,5)	0				
PB(5,6,8,9)	(2,3,5,6)	0			
PC(2,3,5,6)	(-1,0,2,3)	(1,2,4,5)	0		
MQ1(3,4,6,7)	(0,1,3,4)	(0,1,3,4)	(-1,0,2,3)	0	
MQ2(6,7,9,10)	(3,4,6,7)	(-1,0,2,3)	(2,3,5,6)	(1,2,4,5)	0

ii. Get the fuzzy dissimilarity matrix

TABLE II
Fuzzy Dissimilarity Matrix in term of Graded Mean Integration Representation

	PA	PB	PC	MQ1	MQ2
PA	0				
PB	4	0			
PC	1	3	0		
MQ1	2	2	1	0	
MQ2	5	1	4	3	0

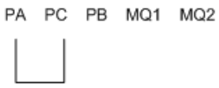
iii. Clustering the data based on dissimilarity matrix with dendrogram

From Fig. 3, it can estimate that PA and PC belong to quality level 1 and PB belongs to quality level 2. This hierarchical dendrogram clearly shows the steps of clustering or divergence, and it terminates when MQ1 and MQ2 will merge into one cluster which is strictly forbidden by this method constrain.

Clustering the Products based on Quality mean standards

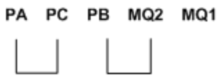
Corresponding dendrogram drawn:

	PA	PB	PC	MQ1	MQ2
PA	0				
PB	4	0			
PC	1	3	0		
MQ1	2	2	1	0	
MQ2	5	1	4	3	0



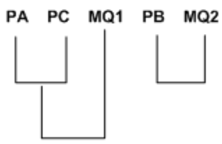
1. Merge of PA and PC to cluster (PAPC)

	(PAPC)	PB	MQ1	MQ2
(PAPC)	0			
PB	3	0		
MQ1	1	2	0	
MQ2	4	1	3	0



2. Merge of B and E to cluster (PBMQ2)

	(PAPC)	(PBMQ2)	MQ1
(PAPC)	0		
(PBMQ2)	3	0	
MQ1	1	2	0



3. Merge of (PAPC) and MQ1 to cluster (PAPCMQ1)

Fig. 3 Clustering of fuzzy numbers based on fuzzy dissimilarity

IV. CONCLUSION

This paper proposes a systematic method for applying fuzzy hierarchical clustering method to do quality estimation based on fuzzy quality mean. This clustering method derives from normal hierarchical clustering agglomerative nesting, with the fuzzy dissimilarity replacing of the normal one. With the proposed method, the quality estimation with fuzzy data can be done without relying on experienced experts or highly skilled engineer. What's more, the hierarchical clustering can clearly show the steps how data merges into clusters and easily extract the results on the quality level of products. In this paper the definition of fuzzy dissimilarity based on fuzzy distance has been given. The sample fuzzy number is trapezoidal fuzzy number in this paper. The defuzzified representation way is Grades Mean Integration Representation which is the most simple and precise way. The quality standards are introduced with clustering the collected data and terminates once the fuzzy quality means will merge into one cluster. In the numerical experiment, it's clearly shown the process that how fuzzy hierarchical clustering can be applied for quality estimation.

Future work of this research will focus on improving the hierarchical clustering algorithm and apply the concept with the practical cases so as to verify the feasibility of the proposed methodology.

REFERENCES

[1] Dudek-Burlikowska, M. (2005). Quality estimation of process with usage control charts type XR and quality capability of process Cp, Cpk Journal of materials processing technology 162: 736-743.
[2] L. A. Zadeh, "Fuzzy set", *Information and Control 8: 338-353*, 1965.
[3] Ho, G. T. S., H. C. W. Lau, et al. (2006). "An intelligent production workflow mining system for continual quality enhancement." The International Journal of Advanced Manufacturing Technology 28(7): 792-809.
[4] Lau, H. C. W., E. N. M. Cheng, et al. (2008). "A fuzzy logic approach to forecast energy consumption change in a manufacturing system." Expert Systems with Applications 34(3): 1813-1824.
[5] Lau, H. C. W., G. T. S. Ho, et al. (2009). "Development of an intelligent quality management system using fuzzy association rules." Expert Systems with Applications 36(2, Part 1): 1801-1815.
[6] Yaqiong, L., L. K. Man, et al. "Fuzzy theory applied in quality management of distributed manufacturing system: A literature review and classification." Engineering Applications of Artificial Intelligence.
[7] Dubois, D. and H. Prade (1978). "Operations on fuzzy numbers." International Journal of Systems Science 9(6): 613-626.
[8] S. Heilpern, Representation and application of fuzzy numbers, Fuzzy Sets andSystems 91 (1997)259-268。
[9] Defuzzification: criteria and classification, from the journal Fuzzy Sets and Systems, Van Leekwijck and Kerre, Vol. 108 (1999), pp. 159-178
[10] Chen, S. H., S. T. Wang, et al. (2006). "Some Properties of Graded Mean Integration Representation of LR Type Fuzzy Numbers." Tamsui Oxford Journal of Mathematical Sciences 22(2): 185.
[11] Chen, S. H. and C. C. Wang (2006). Fuzzy distance of trapezoidal fuzzy numbers. Proceedings of the 9th Joint Conference on Information Sciences,.
[12] Lv, Y. Q. and C. K. M. Lee Fuzzy hierarchical clustering based on fuzzy dissimilarity. Industrial Engineering and Engineering Management (IEEM), 2011 IEEE International Conference on, IEEE.

Lv Yaqiong is now pursuing her PhD degree in division of system engineering and engineering management, school of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. She got her B.eng in Huazhong University of Science and Technology, China, 2008, majoring in industrial engineering.