



# A clustering method based on fuzzy equivalence relation for customer relationship management

Yu-Jie Wang

Department of Shipping and Transportation Management, National Penghu University, Penghu 880, Taiwan, ROC

## ARTICLE INFO

### Keywords:

Clustering  
Customer relationship management  
Fuzzy compatible relation  
Fuzzy equivalence relation  
Transitive closure

## ABSTRACT

In real world, customers commonly take relevant attributes into consideration for the selection of products and services. Further, the attribute assessment of a product or service is often presented by a linguistic data sequence. To partition these linguistic data sequences of customers' assessment on a product or service, a proper clustering method is essential and proposed in this paper. In the clustering method, the linguistic data sequences are presented by fuzzy data sequences and a fuzzy compatible relation is first constructed to present the binary relation between two data sequences. Then a fuzzy equivalence relation is derived by max–min transitive closure from the fuzzy compatible relation. Based on the fuzzy equivalence relation, the linguistic data sequences are easily classified into clusters. The clusters representing the selection preferences of different customers on the product or service will be the foundation of developing customer relationship management (CRM).

© 2010 Elsevier Ltd. All rights reserved.

## 1. Introduction

For the selection of an indicated product or service, customers often have to assess related attributes of the product or service. Commonly, the related attributes are so many and their important degrees are varied for different customers. Further, attributes ratings will stand for customers' preferences that are presented with linguistic terms (Delgado, Verdegay, & Vila, 1992; Herrera, Herrera-Viedma, & Verdegay, 1996), such as very poor (VP), poor (P), medium poor (MP), fair (F), medium good (MG), good (G) and very good (VG). Thus, a customer's assessment on several attributes is expressed with a linguistic data sequence and many customers' assessments on the attributes are shown with linguistic data sequences. To discover essential knowledge, linguistic data sequences of customers' preferences have to be processed. In the information procedure, one of the most important technique is clustering (Bandyopadhyay & Maulik, 2002; Deogun, Kratsch, & Steiner, 1997; Dubes & Jain, 1988; Duda & Hart, 1973; Eom, 1999; Hirano, Sun, & Tsumoto, 2004; Kaufman & Rousseeuw, 1990; Khan & Ahmad, 2004; Krishnapuram & Keller, 1993; Kuo, Chang, & Chien, 2004; Kuo, Ho, & Hu, 2002; Kuo, Liao, & Tu, 2005; Lee, 1999, 2001; MacQueen, 1967; Miyamoto, 2003; Paivinen, 2005; Pedrycz & Vukovich, 2002; Ralambondrainy, 1995; Wang & Lee, 2008; Wu & Yang, 2002). For a given product or service, clustering methods can partition linguistic data sequences of customers' assessments into clusters. The clusters respectively

represent preferences of different customer groups in the product or service. Managers will develop related customer relationship management (CRM) and provide necessary assists for different customer groups according to their preferences. Thus, CRM is constructed on a clustering method.

Through the above reason, clustering is a useful technique to develop CRM. In the past, lots of clustering methods (Bandyopadhyay & Maulik, 2002; Deogun et al., 1997; Dubes & Jain, 1988; Duda & Hart, 1973; Eom, 1999; Hirano et al., 2004; Kaufman & Rousseeuw, 1990; Khan & Ahmad, 2004; Krishnapuram & Keller, 1993; Kuo et al., 2004, 2002, 2005; Lee, 1999, 2001; MacQueen, 1967; Miyamoto, 2003; Paivinen, 2005; Pedrycz & Vukovich, 2002; Ralambondrainy, 1995; Wang & Lee, 2008; Wu & Yang, 2002) were proposed. The clustering methods included cluster analysis, discriminant analysis, factor analysis, principal component analysis (Johnson & Wichern, 1992), gray relation analysis (Feng & Wang, 2000), and *K*-means (Bandyopadhyay & Maulik, 2002; Khan & Ahmad, 2004; MacQueen, 1967; Ralambondrainy, 1995; Wu & Yang, 2002), etc. Generally, cluster analysis, discriminant analysis, factor analysis and principal component analysis are often applied in classic statistical problems, such as large sample or long-term data. On the contrary, gray relation analysis and *K*-means are preferred to deal with small sample or short-term data. Thus, an appropriate method of clustering will be selected by the data pattern.

Commonly, most of the previous clustering methods are often utilized in crisp values, no matter how data belong to large sample, small sample, long-term, or short-term. However, a clustering method in this paper has to partition linguistic data sequences into clusters to present the preferences of different customer groups.

E-mail addresses: [knight\\_edu@yahoo.com.tw](mailto:knight_edu@yahoo.com.tw), [knight@axp1.stm.ntou.edu.tw](mailto:knight@axp1.stm.ntou.edu.tw)

That is to say, the clustering method will partition data under impression, subjectivity and vagueness. Based on the concept, the method is proposed and expressed below. First, linguistic terms are transferred into fuzzy numbers in the clustering method. Then the method will construct a fuzzy binary relation between any two fuzzy data sequences. The fuzzy binary relation is rooted in the similarity of two data sequences and satisfies compatible relation. However, the elements of different clusters may overlap as the fuzzy binary relation merely satisfies compatible relation for partitioning. To solve the overlap problem, some additional mechanisms (Lee, 1999, 2001; Wang & Lee, 2008) are essential on the compatible relation. For instance, Wang and Lee (2008) proposed additional mechanisms to resolve the ties, but the mechanisms may be complex and difficult in computation. Generally, a simpler mechanism is max-min transitive closure (Lee, 1999, 2001), being one of most popular methods, which is often utilized on compatible relation. In this paper, we utilize the max-min transitive closure as the additional mechanism to transform fuzzy compatible relation into fuzzy equivalence relation. Then fuzzy data sequences are partitioned into clusters by the equivalence relation. After clustering, the intra-relations relations will be high, whereas the inter-cluster relations are low.

In short, fuzzy data sequences are from linguistic data sequences and partitioned into clusters by fuzzy equivalence relation for an indicated product or service. One cluster stands for a customer group having similar preferences to the product or service on attributes. After grasping these attribute preferences of different clusters, CRM will be developed on the varied attribute preferences.

For the sake of clarity, mathematical preliminaries are presented in Section 2. The fuzzy compatible relation and equivalence relation are proposed in Section 3. Finally, an empirical example of CRM concerning the application of credit cards is illustrated in Section 4.

**2. Mathematical preliminaries**

In this section, some mathematical theories are stated below. First, we review some basic notions of fuzzy sets and fuzzy numbers (Zadeh, 1965; Zimmermann, 1987, 1991).

**Definition 2.1.** Let  $U$  be a universe set. A fuzzy set  $A$  of  $U$  is defined with a membership function  $\mu_A(x) \rightarrow [0, 1]$ , where  $\mu_A(x), \forall x \in U$ , indicates the degree of  $x$  in  $A$ .

**Definition 2.2.** A fuzzy set  $A$  is normal iff  $\sup_{x \in U} \mu_A(x) = 1$ , where  $U$  is the universe set.

**Definition 2.3.** A fuzzy set  $A$  of the universe set  $U$  is convex iff  $\mu_A(\lambda x + (1 - \lambda)y) \geq (\mu_A(x) \wedge \mu_A(y)), \forall x, y \in U, \forall \lambda \in [0, 1]$ , where  $\wedge$  denotes the minimum operator.

**Definition 2.4.**  $A$  is a fuzzy number iff  $A$  is both normal and convex on  $U$ .

**Definition 2.5.** A triangular fuzzy number  $A$  is a fuzzy number with piecewise linear membership function  $\mu_A$  defined by

$$\mu_A(x) = \begin{cases} (x - a_l)/(a_h - a_l), & a_l \leq x < a_h, \\ 1, & x = a_h, \\ (a_r - x)/(a_r - a_h), & a_h < x \leq a_r, \\ 0, & \text{otherwise,} \end{cases}$$

that can be denoted as a triplet  $(a_l, a_h, a_r)$  (see Fig. 1).

Second, we review the compatible relation and equivalence relation (Epp, 1990; Lee, 1999, 2001) on fuzzy numbers.

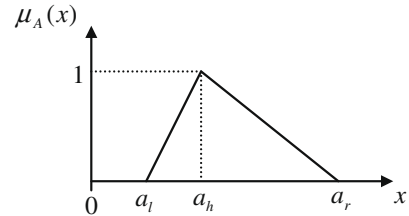


Fig. 1. The membership function of a triangular fuzzy number  $A$ .

**Definition 2.6.** A fuzzy binary relation  $R$  on  $X \times Y$  is defined as the set of ordered pairs:

$$R = \{((x, y), r(x, y)) | (x, y) \in X \times Y\},$$

where  $r$  is a function that maps  $X \times Y \rightarrow [0, 1]$ .

In particular, the relation  $R$  is called a fuzzy binary relation on  $X$  as  $X = Y$ .

**Definition 2.7.** Let  $R$  be a fuzzy binary relation on  $S$ , where  $S$  is a set composed of fuzzy data sequences. The following conditions may hold for  $R$ .

1.  $R$  is reflexive, if  $r(x, x) = 1, \forall x \in S$ .
2.  $R$  is symmetric, if  $r(x, y) = r(y, x), \forall x, y \in S$ .
3.  $R$  is transitive, if  $r(x, y) \geq \max_{y \in S} \min(r(x, y), r(y, z)), \forall x, y, z \in S$ .

If  $R$  satisfies reflexive and symmetric laws,  $R$  is a fuzzy compatible relation on  $S$ . If  $R$  is reflexive, symmetric, and transitive,  $R$  will be a fuzzy equivalence relation on  $S$  (Lee, 1999).

**Definition 2.8.** Let  $R_\lambda$  be a binary relation on  $S$ . Define

$$R_\lambda = \{(x, y) | r(x, y) \geq \lambda, \forall x, y \in S\},$$

where  $0 \leq \lambda \leq 1$ .

**Lemma 2.1.** Let  $R$  be a fuzzy equivalence relation on  $S$ . Then  $R_\lambda$  is an equivalence relation on  $S$  (Epp, 1990).

**Proof.**  $R_\lambda$  is an equivalence relation on  $S$ , because it satisfies the three following conditions:

1. Reflexive:  $r(x, x) = 1 \geq \lambda, \forall x \in S$ .
2. Symmetric: If  $(x, y) \in R_\lambda$ , then  $(y, x)$  is also in  $R_\lambda$  for  $r(y, x) = r(x, y) \geq \lambda, \forall x, y \in S$ .
3. Transitive: Suppose both  $(x, y)$  and  $(y, z)$  are in  $R_\lambda$ , then  $(x, z)$  is also in  $R_\lambda$  for  $r(x, z) \geq \max_{y \in S} \min(r(x, y), r(y, z)) \geq \lambda, \forall x, y, z \in S$ .  $\square$

Commonly, a partition is based on an equivalence relation (Lee, 1999, 2001).

**Definition 2.9.** An  $n \times n$  fuzzy relation matrix  $A = [r_a(i, j)]_{n \times n}$  is a matrix, where  $0 \leq r_a(i, j) \leq 1$  indicates the relation between two sequences  $i$  and  $j$ , and  $0 \leq i, j \leq n$ . It is reflexive, if  $r_a(i, i) = 1$  for  $0 \leq i \leq n$ . A matrix  $A = [r_a(i, j)]_{n \times n}$  is symmetric, if  $r_a(i, j) = r_a(j, i)$ . The matrix  $A$  is called max-min transitive, if  $r_a(i, j) \geq \max_k \min(r_a(i, k), r_a(k, j))$ .

**Definition 2.10.** A fuzzy compatible matrix is a fuzzy relation matrix that is both reflexive and symmetric.

**Definition 2.11.** A fuzzy compatible matrix which satisfies max-min transitive is called a fuzzy equivalence matrix.

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات