Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Operational Research 166 (2005) 160–171
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/dsw

Computing, Artificial Intelligence and Information Technology

Cluster analysis based on fuzzy equivalence relation

Gin-Shuh Liang [a,*], Tsung-Yu Chou [b], Tzeu-Chen Han [c]

[a] Department of Shipping and Transportation Management, National Taiwan Ocean University, 2 Pei Ning Road, Keelung 200, Taiwan, ROC
[b] Department of Marketing and Logistics Management, Ling Tung College, Taichung 402, Taiwan, ROC
[c] Department of Shipping and Transportation Management, National Penghu Institute of Technology, Ma-kung, Penghu 880, Taiwan, ROC

Received 2 September 2002; accepted 16 March 2004
Available online 24 June 2004

Abstract

In this paper, a cluster analysis method based on fuzzy equivalence relation is proposed. At first, the distance formula between two trapezoidal fuzzy numbers is used to aggregate subjects' linguistic assessments about attributes ratings to obtain the compatibility relation. Then a fuzzy equivalence relation based on the fuzzy compatibility relation can be constructed. Finally, using a cluster validity index to determine the best number of clusters and taking suitable λ-cut value, the clustering analysis can be effectively implemented. By utilizing this clustering analysis, the subjects' fuzzy assessments with various rating attitudes can be taken into account in the aggregation process to assure more convincing and accurate cluster analysis.

Keywords: Fuzzy sets; Fuzzy cluster analysis; Linguistic value; Equivalence relation; Strategic group

1. Introduction

Cluster analysis is a very useful classification tool. It has been used frequently in product position, strategy formulation, market segmentation studies and business system planning [1–3]. In disjoint cluster analysis, the object, e.g. a brand in product positioning studies or a strategy in business operation studies, is identified as a member of one and only one cluster. It emphasizes the mutually exclusive and exhaustive nature of the cluster development [1,3]. Indeed, objects may well fit into several clusters, e.g. brands can compete against more than one competitive set as well as consumers can belong to more than one segment. To relax the exclusivity constraint of disjoint cluster analysis, Arabie et al. [2] proposed an overlapping

* Corresponding author. Tel.: +886-2-2462-2192; fax: +886-2-2463-1903.
E-mail address: gsliang@mail.ntou.edu.tw (G.-S. Liang).

0377-2217/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2004.03.018

clustering model. Their approach can be used in marketing studies involving products/objects that can belong to more than one group or cluster simultaneously.

To effectively reflect the fuzzy phenomenon existing in clustering analysis, fuzzy set theory [4–6] was utilized to tackle clustering problem [5–15]. Fuzzy clustering methods have been developed widely in the last 20 years. The first fuzzy clustering method was presented by Dunn [5]. Then, Bezdek [6] developed the fuzzy c-means (FCMs) clustering algorithm to classify each object into clusters with different or same degree of membership based on the assumption that the desired cluster number is given and the minimization of an objective function. To overcome the drawback of the FCMs method that only considered the homogeneity within cluster and ignored the well-separated property, Wang et al. [10] proposed a bi-criteria FCMs method to deal with clustering problem. The FCMs-based methods required that the desired numbers must be specified. This is a disadvantage whenever the desired number does not be determined in a clustering problem. Lee [13] proposed a hierarchical clustering algorithm to cluster the business processes identified in the step of business systems planning. And a matching approach was presented to determine the best number of clusters. Klir and Yuan [11] presented a fuzzy equivalent relation-based hierarchical clustering method to deal with the cluster problem in which the desired number of clusters does not be specified. Groenen and Jajuga [15] developed a fuzzy clustering model based on a root of the squared Minkowski distance with includes squared and unsquared Euclidean distance and L_1-distance. This method can be used to reconstruct clusters from error perturbed data. All the methods stated above are based on the concept of accurate measure and crisp evaluation.

In real life, due to the availability and uncertainty of information as well as the vagueness of human feeling and recognition, it is difficult to exactly evaluate and convey the feeling and recognition of objects against the character of various brands/products or the degree of attention customer have to various management strategies, e.g. the loyalty of customers and the attention of customers' satisfaction. Hence, the precision-based cluster analysis may not be practical. Besides, the data evaluation of the objects feeling toward versus brands/products, as well as the degree of attention management strategies are often expressed in linguistic terms, e.g. "very low", "between very poor and poor", "fair", "very high", "about $9000", "approximately between $56,000 and $64,000", etc. It seems that a fuzziness-based method is needed to integrate various linguistic assessments about attributes ratings and construct the fuzzy cluster analysis.

Since the approximate reasoning of fuzzy set theory can also be used to represent linguistic values [16], to effectively handle the ambiguities involved in the evaluation data and the vague property of linguistic expressions, the trapezoidal fuzzy numbers are used to characterize fuzzy measures of quantitative data and denote the approximate reason of linguistic values [16–18]. Furthermore, a cluster method based on fuzzy equivalence relation is presented to tackle the cluster analysis under fuzzy environment.

To effectively utilize the limited resource and core competence, seeking out the advantageous position over the competition are very important issue in business management [19]. Strategic group is a group of companies that follow the same strategy to establish a major decision action [20–22]. A decision maker of a company does not pick any type of strategy to learn blindingly. They must understand the core ability and financial ability of their own companies and to follow or to learn the competitor behavior in the same group with similar character. Owing to the change of industrial structure, the type of product mixes changes from large size to precision, small volume and high unit price. Since the request for high technical products logistics is low storage and quick delivery, it assists the fast development of airfreight industry in Taiwan, and cause intense competition in airfreight forwarder market. So, to develop a practical cluster analysis algorithm for discriminating strategy a group analysis is important and helpful for airfreight forwarder decision maker. At the end of this paper, authors will have a empirical example of Taiwan airfreight forwarder for the clustering and analyzing current operation strategies.

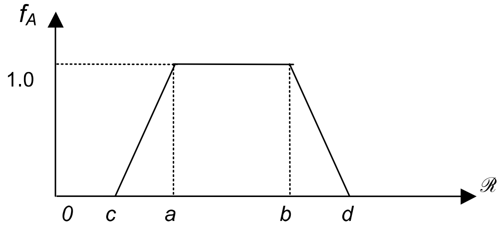

Fig. 1. Membership function of a trapezoidal fuzzy number $A = (c, a, b, d)$.

2. Trapezoidal fuzzy numbers

A fuzzy number [17] A in $\Re$ (real line) is a trapezoidal fuzzy number [18,23], if its membership function $f_A : \Re \to [0, 1]$ is

$$f_A(x) = \begin{cases} (x-c)/(a-c), & c \leqslant x \leqslant a, \\ 1, & a \leqslant x \leqslant b, \\ (x-d)/(b-d), & b \leqslant x \leqslant d, \\ 0, & \text{otherwise}, \end{cases}$$

with $-\infty < c \leqslant a \leqslant b \leqslant d < \infty$. The trapezoidal fuzzy number can be denoted by $A = (c, a, b, d)$ (Fig. 1). The interval $[a, b]$ give the maximal grade of $f_A(x)$, i.e. $f_A(x) = 1$, $x \in [a, b]$, it is the most possible value of evaluation data. The c and d are the lower and upper bounds of the available area for the evaluation data. They are used to reflect the fuzziness of the evaluation data. Here the trapezoidal fuzzy numbers are utilized to denote the approximate reasoning of linguistic values, which are used to convey the subjective evaluation of the subjects.

In a fuzzy decision environment, choosing a preference rating system is an important issue. In general, two preference ratings can be used. They are linguistic values and fuzzy numbers. Based on the practical needs, the decision-makers may apply one or both of them [18]. Herein, the linguistic values [16,18], subjectively represented by the trapezoidal fuzzy number, are utilized to evaluate the preference of objects versus various management strategy attributes. For example, the attention degree is "high" can be represented by (0.6, 0.8, 0.8, 1), "between median and high" can be represented by (0.3, 0.5, 0.8, 1).

3. Distance between two trapezoidal fuzzy numbers

Distance measures, association measures and correlation measures are three primary kinds of methods to characterize the interobjects similarity. By considering the easy implementation and intuition, the distance measures based on trapezoidal fuzzy numbers are used to construct the similarity between two objects. There are three distance formula constructed on trapezoidal fuzzy numbers are suitable for characterizing the interobjects similarity.

Let $A_i = (c_i, a_i, b_i, d_i)$ and $A_k = (c_k, a_k, b_k, d_k)$ be two trapezoidal fuzzy numbers.

In 1997, Heilpern [23] defined the geometrical distance between two trapezoidal fuzzy numbers. Based on this method, the distance between A_i and A_k, denoted by $G_p(A_i, A_k)$, is

$$G_p(A_i, A_k) = \begin{cases} 0.25(|c_i - c_k|^p + |a_i - a_k|^p + |b_i - b_k|^p + |d_i - d_k|^p)^{1/p}, & \text{for } 1 \leqslant p < \infty, \\ \max\{|c_i - c_k|, |a_i - a_k|, |b_i - b_k|, |d_i - d_k|\}, & \text{for } p = \infty. \end{cases}$$

In 1999, Chen and Hsieh [24] studied the Heilpern's geometrical distance based on the geometrical operation of trapezoidal fuzzy numbers, and proposed the modified geometrical distance method. Based on this method, the distance between A_i and A_k, denoted by $d_p(A_i, A_k)$, is

$$d_p(A_i, A_k) = \begin{cases} [0.25(|c_i - c_k|^p + |a_i - a_k|^p + |b_i - b_k|^p + |d_i - d_k|^p)]^{1/p}, & \text{for } 1 \leqslant p < \infty, \\ \max\{|c_i - c_k|, |a_i - a_k|, |b_i - b_k|, |d_i - d_k|\}, & \text{for } p = \infty. \end{cases}$$

In 1991, Kaufmann and Gupta [25] had discussed a distance measure of fuzzy numbers. According to the concept of Dissemblance Index proposed by Kaufmann and Gupta [25], the distance between A_i and A_k, denoted by $D(A_i, A_k)$, can be defined as

$$D(A_i, A_k) = \int_0^1 \left(|a_i^\alpha - a_k^\alpha| + |b_i^\alpha - b_k^\alpha|\right) \mathrm{d}\alpha,$$

where

$$a_i^\alpha = c_i + \alpha(a_i - c_i), \quad b_i^\alpha = d_i + \alpha(b_i - d_i), \quad a_k^\alpha = c_k + \alpha(a_k - c_k), \quad b_k^\alpha = d_k + \alpha(b_k - d_k), \ \ \alpha \in [0, 1].$$

Let

$$h_1 = |c_i - c_k|/(|c_i - c_k| + |a_i - a_k|), \quad h_2 = |d_i - d_k|/(|d_i - d_k| + |b_i - b_k|),$$

then $D(A_i, A_k)$ can be simplified into four cases:

Case 1: If
$((c_i < c_k$ and $a_i < a_k)$ or $(c_i > c_k$ and $a_i > a_k))$ and
$((b_i < b_k$ and $d_i < d_k)$ or $(b_i > b_k$ and $d_i > d_k))$ then
$D(A_i, A_k) = \{|c_i - c_k| + |a_i - a_k| + |d_i - d_k| + |b_i - b_k|\}/2.$

Case 2: If
$((c_k \leqslant c_i$ and $a_i \leqslant a_k)$ or $(c_i \leqslant c_k$ and $a_k \leqslant a_i))$ and
$((b_i < b_k$ and $d_i < d_k)$ or $(b_i > b_k$ and $d_i > d_k))$ then
$D(A_i, A_k) = \{h_1|c_i - c_k| + (1 - h_1)|a_i - a_k| + |d_i - d_k| + |b_i - b_k|\}/2.$

Case 3: If
$((c_i < c_k$ and $a_i < a_k)$ or $(c_i > c_k$ and $a_i > a_k))$ and
$((b_k \leqslant b_i$ and $d_i \leqslant d_k)$ or $(b_i \leqslant b_k$ and $d_i \geqslant d_k))$ then
$D(A_i, A_k) = \{|c_i - c_k| + |a_i - a_k| + h_2|d_i - d_k| + (1 - h_2)|b_i - b_k|\}/2.$

Case 4: If
$((c_k \leqslant c_i$ and $a_i \leqslant a_k)$ or $(c_i \leqslant c_k$ and $a_k \leqslant a_i))$ and
$((b_k \leqslant b_i$ and $d_i \leqslant d_k)$ or $(b_i \leqslant b_k$ and $d_i \geqslant d_k))$ then
$D(A_i, A_k) = \{h_1|c_i - c_k| + (1 - h_1)|a_i - a_k| + h_2|d_i - d_k| + (1 - h_2)|b_i - b_k|\}/2.$

In general, real number t can be considered as the special case of trapezoidal fuzzy number (t, t, t, t). Consequently, trapezoidal fuzzy numbers $A_i = (a_i, a_i, a_i, a_i)$ and $A_k = (a_k, a_k, a_k, a_k)$ can be considered as real numbers a_i and a_k, respectively. Then, the distance of A_i and A_k based on the concept of Dissemblance Index, is equal to $2|a_i - a_k|$, i.e., $D(A_i, A_k) = 2|a_i - a_k|$. Besides, considering parameter $p = 2$, we can obtain Heilpern's geometrical distance $G_2(A_i, A_k) = 0.5|a_i - a_k|$ and Chen's modified geometrical distance $d_2(A_i, A_k) = |a_i - a_k|$. Owing to the Chen's modified geometrical distance with parameter $p = 2$ can meet the concept of the classical distance, we utilize it to solve the distance between two trapezoidal fuzzy numbers in this paper.

4. The construction of a fuzzy equivalence relation

Definition 1. Let R be a fuzzy relation on $X \times Y$, i.e. $R = \{((x, y), f_R(x, y)) | (x, y) \in X \times Y\}$, the λ-cut matrix R_λ is denoted by

$$R_\lambda = \{((x,y), f_{R_\lambda}(x,y)) | f_{R_\lambda}(x,y) = 1, \text{ if } f_R(x,y) \geqslant \lambda;\ f_{R_\lambda}(x,y) = 0, \text{ if } f_R(x,y) < \lambda, (x,y) \in X \times Y, \lambda \in [0,1]\}.$$

Definition 2. Let $R \subset X \times Y$ and $S \subset Y \times Z$ be two fuzzy relations, the max–min composition $R \circ S$ is defined by

$$R \circ S = \{((x,z), \max_y \{\min_{x,z} \{f_R(x,y), f_S(y,z)\}\}) | x \in X, y \in Y, z \in Z\}.$$

Definition 3. A fuzzy relation R on $X \times X$ is called a fuzzy equivalence relation if following three conditions held:

(1) Reflexive, i.e., $f_R(x,x) = 1$, $\forall x \in X$.
(2) Symmetric, i.e., $f_R(x,y) = f_R(y,x)$, $\forall x,y \in X$.
(3) Transitive, i.e., $R^{(2)} = R \circ R \subset R$, or more explicitly.

$$f_R(x,z) \geqslant \max_y \left\{ \min_{x,z} \{f_R(x,y), f_R(y,z)\} \right\},\ \forall x,y,z \in X.$$

Definition 4. A fuzzy relation R on $X \times Y$ is called a fuzzy compatibility relation if it satisfies reflexive and symmetric conditions.

Definition 5. The transitive closure, R_T, of a fuzzy relation R is defined as the relation that is transitive, contains R and has the smallest possible membership grades.

Theorem 1 [11]. *Let R be a fuzzy compatibility relation on a finite universal set X with $|X| = n$, then the max–min transitive closure of R is the relation $R^{(n-1)}$.*

According to Theorem 1, we can get the algorithm to find the transitive closure, $R_\mathrm{T} = R^{(n-1)}$.

Algorithm A. Find the transitive closure R_T of fuzzy compatibility relation.

Step 1: Calculate $R^{(2)}$ if $R^{(2)} \subset R$ or $R^{(2)} = R$, then transitive closure $R_\mathrm{T} = R$ and stop. Otherwise, $k = 2$, go to step 2.

Step 2: If $2^k \geqslant n-1$, then $R_\mathrm{T} = R^{(n-1)}$ and stop. Otherwise, calculate $R^{(2^k)} = R^{(2^{k-1})} \circ R^{(2^{k-1})}$, if $R^{(2^k)} = R^{(2^{k-1})}$, then transitive closure $R_\mathrm{T} = R^{(2^k)}$ and stop. Otherwise, go to step 3.

Step 3: $k = k+1$, go to step 2.

Definition 6 (R_λ classification principle). Let R be a fuzzy compatibility relation on a finite universal set $A = \{A_1, A_2, \ldots, A_n\}$, R_T be the transitive closure of R, and R_λ be the λ-cut matrix of R_T, then x and y belong to the same cluster iff

$$f_{R_\lambda}(x,y) = 1,\ (x,y) \in A \times A,\ \lambda \in [0,1].$$

5. Cluster analysis method based on fuzzy equivalence relation

Suppose there are n objects $A_1, A_2, \ldots, A_i, \ldots, A_n$ and m attributes (factors) $B_1, B_2, \ldots, B_j, \ldots, B_m$. Given a set of $X = \{X_1, X_2, \ldots, X_i, \ldots, X_n\}$, where X_i is a vector, i.e. $X_i = (X_{i1}, X_{i2}, \ldots, X_{ij}, \ldots, X_{im})$, and $X_{ij} = (c_{ij}, a_{ij}, b_{ij}, d_{ij})$, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$, to be the attribute B_j's preference rating of ith object A_i.

The systematic clustering method presented herein can be summarized as follows.

Step 1. Normalize the original attribute preference rating.

In order to ensure the distance summation of different attribute preference ratings can be efficiently tackled. Define the normalized attribute preference rating of X_{ij}, denote it by $X_{ij}^{\sim}$, as follows:

$$X_{ij}^{\sim} = \left((c_{ij} - c_j^*)/t_j^*, (a_{ij} - c_j^*)/t_j^*, (b_{ij} - c_j^*)/t_j^*, (d_{ij} - c_{ij}^*)/t_j^*\right),$$

where

$$t_j^* = d_j^* - c_j^*, \quad d_j^* = \max_i \{d_{ij}\}, \quad c_j^* = \min_i \{c_{ij}\}.$$

Step 2. Find the fuzzy compatibility relation R on X.

According to the distance function between two trapezoidal fuzzy numbers, define the fuzzy compatibility relation R as follows:

$$R(X_i, X_k) = 1 - \eta\left(\sum_{j=1}^{m} d_2\left(X_{ij}^{\sim}, X_{kj}^{\sim}\right)\right),$$

where η is the inverse value of the largest distance in X, i.e.,

$$\eta = \left\{\max_{i,k}\left\{\sum_{j=1}^{m} d_2\left(X_{ij}^{\sim}, X_{kj}^{\sim}\right)\right\}\right\}^{-1}.$$

Step 3. Find the transitive closure R_T by utilizing algorithm A.

Step 4. Find all the feasible clusters by taking suitable λ value.

Based on the fuzzy compatibility relation R stated as above, we can take suitable λ value ($\lambda \in [0, 1]$), variable from 1 to 0, and obtain a series cluster.

Step 5. Determine the best number of clusters and find the objects in each cluster.

A cluster validity index L, modified the compactness and separation validity function [8], can be used to determine the best number of clusters. Define

$$L = \frac{T}{n \times d_{\min}^2},$$

where

$$T = \sum_{r=1}^{h} \sum_{i=1}^{n} u_{ir} d_2^2\left(X_i^{\sim}, V_r^{\sim}\right),$$

$$d_{\min}^2 = \min_{q,r} d_2^2\left(V_q^{\sim}, V_r^{\sim}\right),$$

$$u_{ir} = \begin{cases} 1, & \text{if } A_i \in C_r, \\ 0, & \text{if } A_i \notin C_r, \end{cases}$$

$$d_2^2\left(X_i^{\sim}, V_r^{\sim}\right) = \sum_{j=1}^{m} d_2^2\left(X_{ij}^{\sim}, V_{rj}^{\sim}\right),$$

$$d_2^2\left(V_q^{\sim}, V_r^{\sim}\right) = \sum_{j=1}^{m} d_2^2\left(V_{qj}^{\sim}, V_{rj}^{\sim}\right),$$

A_i: the ith object, h: the number of clusters, n: the number of objects, m: the number of attributes, $X_i^{\sim}$: the normalized attribute preference rating vector of object A_i versus all attributes, i.e. $X_i^{\sim} = (X_{i1}^{\sim}, \ldots, X_{ij}^{\sim}, \ldots, X_{im}^{\sim})$, $V_r^{\sim}$: the normalized centroid vector of cluster C_r, i.e.

$$V_r^{\sim} = \left(V_{r1}^{\sim}, \ldots, V_{rj}^{\sim}, \ldots, V_{rm}^{\sim}\right),$$

C_r: the rth cluster, $d_2^2(X_{ij}^{\sim}, V_{rj}^{\sim})$: the Chen's modified geometrical distance square between $X_{ij}^{\sim}$ and $V_{rj}^{\sim}$ with parameter $p = 2$, $d_2^2(V_{qj}^{\sim}, V_{rj}^{\sim})$: the Chen's modified geometrical distance square between $V_{qj}^{\sim}$ and $V_{rj}^{\sim}$ with parameter $p = 2$.

Since, the more separate the clusters, the larger $d_{\min}^2$, and the smaller L. A smaller L indicates a clustering in which all the clusters are more compact and separate to each other. Thus, the smallest L indeed indicates a valid optimal clustering. After calculating the L value of various clusters obtained by step 4, the best number of clusters and the objects belong to each cluster can be obtained.

6. Empirical example

Ones can obtain strategy group by strategy factors, and strategy can be evaluated by strategy factors. There are different strategic factors to consider in different industries. Thus, we can understand the character and resource configuration of airfreight forwarder business by describing of strategic factors. Further, we could discriminate one or more strategies from airfreight industry and to comprehend the competitive situation deeply. Strategy factors were presented by scholars, such as innovation, difference, sustainable competitive advantage, scope of products, constructing and strengthening network relation, cost controlling [19]. In this paper, we gather 28 strategic criteria (Table 1) from scholars, experts and proprietors. We select 30 companies of airfreight forwarder in Taiwan by random selection to obtain the strategic factors. Using Statistical Analysis System (SAS), we obtain seven factors (Table 1). The proposed fuzzy cluster analysis method will be utilized to determine the relevantly strategic group.

The decision-makers may tackle preference rating system by adopting one of various rating scales assumed in the literature [18,26,27] or may develop their own rating scales system by using trapezoidal fuzzy number to show the individual conception of the linguistic variable "attention degree". As to how many elements of linguistic rating scale set should be included; the values 5, 7 and 9 are suitable [28]. According to the preference ratings proposed by Liang and Wang [18], it is suggested that the decision-makers utilize the linguistic rating set $W = \{\text{VL}, \text{B.VL \& L}, \text{L}, \text{B.L \& M}, \text{M}, \text{B.M \& H}, \text{H}, \text{B.H \& VH}, \text{VH}\}$, where VL = Very Low, B.VL & L = Between Very Low and Low, L = Low, B.L & M = Between Low and Medium, M = Medium, B.M & H = Between Medium and High, H = High, B.H & VH = Between High and Very High, VH = Very High, to assess the attention degree of subjects companies under each of the management strategies. Herein, they are subjectively defined as: VL = (0,0,0,0.2), B.VL & L = (0,0,0.2,0.4), L = (0,0.2, 0.2,0.4), B.L & M = (0,0.2,0.5,0.7), M = (0.3,0.5,0.5,0.7), B.M & H = (0.3,0.5,0.8,1), H = (0.6,0.8,0.8,1), B.H & VH = (0.6,0.8,1,1), and VH = (0.8,1,1,1). Since the calculating process is complex for 30 samples, we give a simple example to spread out authors' fuzzy equivalence clustering process.

The decision-makers utilize the linguistic rating as above and obtain the evaluation results as Table 2.

By using steps 1 and 2 of proposed fuzzy cluster analysis method, the fuzzy compatibility relation can be obtained as follows:

$$R = \begin{bmatrix} 1 & 0 & 0.415 & 0.242 & 0.679 \\ 0 & 1 & 0.389 & 0.080 & 0.108 \\ 0.415 & 0.389 & 1 & 0.050 & 0.292 \\ 0.242 & 0.080 & 0.050 & 1 & 0.590 \\ 0.679 & 0.108 & 0.292 & 0.590 & 1 \end{bmatrix}.$$

Table 1
The strategic factors and criteria of airfreight forwarder in Taiwan

Factors	Criteria
Core ability	Ability of EDI handling
	Work attitude and profession knowledge of sales
	Ability of new customer creating
	Ability of coordination and problem solving
Organization management	Ability of world wild organization network management
	Convenience of cargo tracking
	Operation process is simple and definite
	Definite business object
Pricing	Discount
	The rationality of service charge
	The site of company
	Sales ability
	Effective advertisement and promotion
	Ability of process improving
	The completeness of indemnification
Competitive forces	The reputation of company
	The ability of freight safekeeping
	Holding high loyal customers
	New route developing
	Ability of market demand forecasting
Finance	Ability of capital raising
	Ability of delivery on time
Different advantage	Completeness of route
	Completeness of service (land transportation, apply to customs)
	Ability of accident and customer complain handling
Information technology	Degree of IT manpower
	Degree of IT equipment investing
	Degree of IT phase into operation flow

Table 2
The evaluation results of five companies

Company	Factor						
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
C_1	M	H	H	B.H & VH	VH	L	B.M & H
C_2	H	B.L & M	M	B.M & H	H	B.M & H	VL
C_3	H	H	B.M & H	H	H	VH	B.M & H
C_4	VL	M	H	B.VL &L	H	B.L & M	M
C_5	L	M	B.H & VH	H	B.H & VH	B.VL & L	B.M & H

Then, by the step 3, the transitive closure of compatibility relation R_T can be found, i.e.,

$$R_T = \begin{bmatrix} 1 & 0.389 & 0.415 & 0.590 & 0.679 \\ 0.389 & 1 & 0.389 & 0.389 & 0.389 \\ 0.415 & 0.389 & 1 & 0.415 & 0.415 \\ 0.590 & 0.389 & 0.415 & 1 & 0.590 \\ 0.679 & 0.389 & 0.415 & 0.590 & 1 \end{bmatrix}.$$

Based on the elements of R_T matrix, we can find five intervals, $[0, 0.389]$, $(0.389, 0.415]$, $(0.415, 0.590]$, $(0.590, 0.679]$, and $(0.679, 1]$, which are suitable to generate clusters. Finally, by taking suitable $\lambda \in [0, 1]$

and according to the R_λ classification principle, the relevantly strategic clusters can be obtained. For example, by taking $\lambda \in (0.590, 0.679]$, then R_λ is

$$R_\lambda = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and the relevantly strategic clusters are $\{C_1, C_5\}$, $\{C_2\}$, $\{C_3\}$ and $\{C_4\}$.

After calculating the simple sample, the clustering results of 30 companies will be shown as following.

At first, we try to know the respectable degree of strategic factors of airfreight forwarder in Taiwan. And another target is to make cluster analysis for airfreight forwarder industry.

From the results of examination, we can obtain the order of respectable strategies in Taiwan's airfreight forwarder industry. The most respectable strategy of operation used by airfreight forwarder in Taiwan is differential advantage. The others are organization management, information technology etc. The results are shown as Fig. 2. From the outcome, we know the trend of airfreight forwarder industry is to supply integrated service for customers. Maybe the forwarders can provide special third party logistics service. The purpose of this action, don't doubt, is to make customer have diversity of opinions from their service.

Before the clustering analysis, we want to integrate the evaluate results from our survey, and the outcome list as Appendix A. After transferring the interval ratings to linguistic value, we will use the way of fuzzy equivalence clustering analysis that is presented in this paper. By taking suitable $\lambda \in [0, 1]$ and making the validity analysis, four kinds of cluster numbers and their L values can be obtained. The results are shown in Table 3. As Table 3 shows, the L value was determined to be the least while divided into five clusters. The best fuzzy clustering number was, therefore, five. That is, there are five

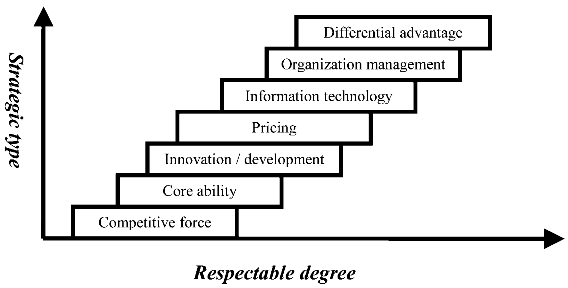

Fig. 2. Respectable degree of airfreight forwarder in Taiwan.

Table 3
The L value

λ interval	Number of clusters	L value
(0.4, 0.5)	2	1.565
(0.5, 0.6)	3	1.292
(0.6, 0.7)	5	0.099
(0.7, 0.8)	9	0.111

operating strategic groups from airfreight forwarders in Taiwan. The character of each group describes as follows.

Group 1: *Performance upgrade strategy*

Compared with other groups, the objects in group 1 pay more attention to each strategic item compared to the others. Especially, the strategies of 'organization management' and 'information technology' are most important for them. About the financial performance of group 1, we get the trend of profit is on the downside. To judge the reason, we find that the main line of the objects of group 1 is Japan and Korea. After Asia finance storm, the import cargo quantity of Japan and Korea area is withered. And it causes the financial performance inferior to other groups.

Group 2: *Core resource strategy*

Many objects belong to group 2. We can judge that the operation strategy of airfreight forwarders in Taiwan is learning mutually from the phenomenon. In this group, the main strategies are 'core ability' and 'differential advantage'. Pricing and information technology are less important for them. So, we can suggest that the management of group 2 is mature. The samples age is older than others, and the market sharing is keeping stable. For this reason, they are not focused on price competition.

Group 3: *Pricing strategy*

In this group, pricing is the main strategy. The strategies of 'Competitive weapon' and 'Innovation/development' are not important for them. Financial performance in this group is not perfect. Thus, we can suggest that the strategy of price war is not a good policy in the Taiwan's airfreight market.

Group 4: *Complex strategy*

The character of this group is that there is not a particular strategy of operation. The managers think that all of the strategic items are equal important. And the financial performance of this group is not successful in the airfreight forwarder industry.

Group 5: *Differential advantage strategy*

Differential advantage is the principal operating strategy for this group. The main differential strategy is risk reduction for customers. Base on the research of Global Trade [29], the safety of cargo is the minimum requirement of customers. So, this group is focus on cargo tracking and security of consignment. Using this strategy, a young company will increase their profit keep growing.

7. Conclusion

This paper presents a new cluster analysis method based on fuzzy equivalence relation to solve the cluster analysis problem. This method can be used for fuzzy data and does not specify any desired number of clusters. Besides, a cluster validity index based on notions of maximum cohesion and minimum coupling is proposed as guidance for determining the best number of clusters.

To efficiently grip the ambiguity exists in available information as well as the essential fuzziness in human judgment and preference, the linguistic values, subjectively represented by trapezoidal fuzzy numbers, are used to evaluate the importance of factors.

From this research, we obtained the order of respectable strategies in Taiwan's airfreight forwarder industry. Such as differential advantage, organization management, information technology etc. Using the fuzzy clustering analysis method, we get five strategic groups in this industry. This result also fits the current industry. Thus, results presented the proposed fuzzy cluster analysis method are practical and useful.

Appendix A

The evaluate results of 30 companies

Company	Factor						
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
C_1	H	H	H	B.H & VH	VH	B.M & H	B.M & H
C_2	H	H	B.M & H	B.M & H	H	B.M & H	B.M & H
C_3	H	H	B.M & H	H	H	B.M & H	B.M & H
C_4	H	B.H & VH	H	B.H & VH	H	B.H & VH	B.M & H
C_5	H	B.H & VH	H	B.H & VH	B.H & VH	B.H & VH	B.M & H
C_6	H	H	B.M & H	M	B.M & H	B.M & H	B.M & H
C_7	H	H	M	B.H & VH	B.M & H	B.M & H	M
C_8	B.M & H	H	M	B.M & H	B.M & H	H	B.L & M
C_9	B.M & H	H	H	H	B.H & VH	B.M & H	M
C_{10}	H	VH	H	B.M & H	B.H & VH	H	B.M & H
C_{11}	M	H	M	H	B.M & H	B.H & VH	B.H & VH
C_{12}	VH	VH	H	B.H & VH	VH	H	M
C_{13}	B.M & H	H	B.M & H	H	B.H & VH	B.M & H	B.M & H
C_{14}	H	H	B.M & H	B.H & VH	B.H & VH	H	M
C_{15}	H	H	H	H	B.H & VH	VH	M
C_{16}	H	H	H	B.H & VH	B.H & VH	B.H & VH	B.H & VH
C_{17}	B.M & H	H	M	H	B.H & VH	B.M & H	B.M & H
C_{18}	M	H	B.M & H	B.H & VH	B.H & VH	M	M
C_{19}	B.M & H	H	B.M & H	B.M & H	VH	B.M & H	B.H & VH
C_{20}	B.M & H	M	M	H	B.H & VH	B.M & H	B.L & M
C_{21}	B.H & VH	VH	B.H & VH	B.H & VH	B.H & VH	B.H & VH	VH
C_{22}	H	B.H & VH	B.M & H	B.H & VH	B.H & VH	H	M
C_{23}	B.H & VH	VH	H	B.H & VH	H	B.M & H	B.M & H
C_{24}	H	B.M & H	M	M	B.H & VH	M	M
C_{25}	VH	VH	H	B.H & VH	B.H & VH	B.H & VH	B.M & H
C_{26}	H	M	H	B.H & VH	B.H & VH	B.H & VH	L
C_{27}	B.M & H	B.M & H	H	H	B.H & VH	H	B.M & H
C_{28}	B.H & VH	B.H & VH	B.M & H	H	B.H & VH	B.M & H	B.H & VH
C_{29}	H	B.M & H	M	H	B.M & H	B.M & H	L
C_{30}	H	B.H & VH	H	B.H & VH	H	H	M

References

[1] M. Aldenderfer, R. Blashfield, Cluster Analysis, Beverly Hill Sage Publications, 1984.
[2] P. Arabie, J.D. Carroll, W. Desarbo, J. Wind, Overlapping clustering: A new method for product positioning, Journal of Marketing Research 18 (1981) 310–317.
[3] G.N. Punj, W.S. David, Cluster analysis in marketing research: Review and suggestion for application, Journal of Marketing Research 20 (1983) 135–148.
[4] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–353.
[5] J.C. Dunn, Will-separated clusters and the optimal fuzzy partition, Journal of Cybernetics 4 (1974) 95–104.

[6] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum, New York, 1981.
[7] H. Hruschka, Market definition and segmentation using fuzzy clustering methods, International Journal of Research in Marketing 3 (1) (1986) 17–34.
[8] X.L. Xie, G. Beni, A validity measure for fuzzy clustering, IEEE Transportation of Pattern Analysis and Machine Intelligence 13 (1991) 841–847.
[9] J.C. Bezdek, S. Pal, Fuzzy Models for Pattern Recognition: Methods That Search for Structures in Data, IEEE Press, Piscataway, NJ, 1992.
[10] H.F. Wang, C. Wang, G.W. Wu, Bi-criteria fuzzy c-means analysis, Fuzzy Sets and System 64 (1994) 311–319.
[11] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic Theory and Application, Prentice Hall PTR, Upper Saddle River, NJ, 1995.
[12] H. Wang, P.B. McCauley, Fuzzy clustering analysis and multi-factorial evaluation for students' imaginative power in physics problem solving, Fuzzy Sets and Systems 78 (1996) 95–105.
[13] H.S. Lee, Automatic clustering of business process in business systems planning, European Journal of Operational Research 114 (1999) 354–362.
[14] O. Muammer, User segmentation of online music service using fuzzy clustering, Omega 29 (2001) 193–206.
[15] P.J.F. Groenen, K. Jajuga, Fuzzy clustering with squared Minkowski distances, Fuzzy Sets and Systems 120 (2) (2001) 227–237.
[16] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Parts I, II and III, Information Science 8 (1975) 199–249, 301–357; 9 (1976) 43–80.
[17] D. Dubois, H. Prade, Operation on fuzzy numbers, International Journal of Systems Science 9 (1978) 613–626.
[18] G.S. Liang, M.J. Wang, A fuzzy multi-criteria decision-making method for facility site selection, International Journal of Product Research 29 (11) (1991) 2313–2330.
[19] D.A. Aaker, Strategic Market Management, John Wiley & Sons, New York, 1992.
[20] M. Hunt, Competition in the major home appliance industry, 1960–1970, Ph.D. dissertation, Graduate School of Business Administration, Harvard University, 1972.
[21] M.E. Porter, Competitive Strategy: Techniques for Analyzing Industries and Competitors, Free Press, New York, 1980.
[22] K.O. Cool, Strategic Group Formation and Strategic Group Shifts: A Longitudinal Analysis of the U.S. Pharmaceutical Industry, 1963–1982, Doctorial Dissertation, Graduate School of Business Administration, Purdue University, 1985.
[23] S. Heilpern, Representation and application of fuzzy numbers, Fuzzy Sets and Systems 91 (1997) 259–268.
[24] S.H. Chen, C.H. Hsieh, A model and algorithm of fuzzy product positioning, Information Sciences 121 (1999) 61–82.
[25] A. Kaufmann, M.M. Gupta, Introduction to Fuzzy Arithmetic Theory and Application, Van Nostrand Reinhold, New York, 1991.
[26] T.S. Liou, M.J.J. Wang, Subjective assessment of mental workload—a fuzzy linguistic multi-criteria approach, Fuzzy Sets and Systems 62 (1994) 155–165.
[27] S.H. Gbyym, A semi-linguistic fuzzy approach to multifactor decision-making: Application to aggregation of experts' judgments, Annals of Nuclear Energy 26 (1999) 1097–1112.
[28] J.M. Booker, M.A. Mayer, Common problem in the elicitation and analysis of expert opinion affecting probabilistic safety assessments, presented in CSNI Workshop on PSA Application and Limitation (1990) 353–368, Santa Fe, NM, September 4–6.
[29] Global Trade, Pilot Air Freight Defines Quality Transport Service 108, 1988, p. 20.