International Journal of Production Research
Vol. 48, No. 13, 1 July 2010, 3803–3821

A clustering algorithm for supplier base management

Darshit Parmar[a], Teresa Wu[a]*, Tom Callarman[b], John Fowler[a] and Philip Wolfe[a]

[a]*Department of Industrial Engineering, Arizona State University, Tempe, Arizona 85287-5906, USA;* [b]*China Europe International Business School, 699 Hongfeng Road, Pudong, Shanghai, People's Republic of China, 201206*

(*Received 31 August 2008; final version received 1 April 2009*)

Supplier base management is an important strategy for managing global, customer driven supply chains. Successful supplier base management can effectively handle supply side exceptions, which may have significant business implications. Currently, there is a trend to reduce the size of the supplier base which makes the coordination and interaction among suppliers more effective, less costly and time consuming. The goal of this research is to present a clustering algorithm, named min-min-roughness (MMR) to cluster suppliers into smaller, more manageable groups with similar characteristics. Due to the fact that supplier data are mainly categorical in nature, MMR, based on rough set theory (RST) is developed for categorical data clustering which is also capable of handling the uncertainty during the clustering process. One potential benefit of applying MMR to supplier base management is that more realistic benchmarking can be obtained and the fulfilment operation can be sped up by reducing the number of variables impacting the operations. In addition, the characteristics of each smaller group of suppliers can be summarised and exploited to handle supply side disruptions.

Keywords: supplier base management; rough set theory; data mining

1. Introduction

Supply chain is a commonly heard term in academia as well as industry. It refers to a network of integrated and dependent processes through which specifications are transformed to finished deliverables. In addition to managing the flow of materials, it involves managing the integrated information about product flow, from suppliers to end users, in order to improve customer satisfaction, reduce time to market, and reduce cost in inventories. Supply chain management (SCM) is concerned with managing this complexity. This is not an easy task because the complexity depends on prevailing circumstances characterised by the following market trends: customer behaviour, collaboration, uncertainty, continuously changing business environment, flexibility, globalisation and security (Hicks 2002). During this entire process of transforming specifications to finished deliverables, things often do not go as planned. This is what is commonly referred to as an exception or a disruption in the literature. The implications of exceptions/disruptions can

*Corresponding author. Email: teresa.wu@asu.edu

ISSN 0020–7543 print/ISSN 1366–588X online

DOI: 10.1080/00207540902942891
http://www.informaworld.com

be costly and can lead to significant customer delivery delays. Xu *et al*. (2003) classify exceptions into three categories: input related exceptions, process related exceptions and output related exceptions. Delivery problems from suppliers such as partial delivery and delayed delivery are input related exceptions, disturbances occurring within the manufacturing system are process related exceptions and order changes from customers are identified as output related exceptions. Zsidisin *et al*. (2003) state that lean supply chains, resulting from approaches such as lean systems/just-in-time, six sigma etc., have become fragile and thus more vulnerable to supply side disruptions. Inbound supply disruptions such as unplanned downtime, late deliveries, stock outs/lost sales and delayed launches, generally lead to higher costs, lower revenue, and lower market share. Christopher and Peck (2004) indicate that supply chains are becoming far more vulnerable to external disruptions and while single sourcing may be advantageous from a cost and quality management perspective, it could be dangerous in terms of resilience. They suggest that wherever possible, alternative supply sources should be available. Similarly, Xu *et al*. (2003) propose a business process design for handling rush order exceptions, in which they identify an alternative supplier listing and select one from this listing. As stated by Zsidisin and Ellram (2003), a supplier's failure to deliver goods or services can have a detrimental effect throughout the purchasing firm and subsequently throughout the supply chain, ultimately impacting the customer.

The business implications of supplier exceptions have led to considerable supplier-related research, commonly referred to as supplier base management. Due to the globalisation of trade and the Internet enlarging a purchaser's choice set, supplier base management has become a challenging problem in business and academic research. One solution is to divide all suppliers into smaller sets where each set has a greater degree of similar characteristics. Potential benefits of working with a smaller subset of suppliers include speeding development and fulfilment operations, reducing the number of variables impacting the operation, allowing design engineers to work closely with supplier engineers, etc. In addition, obtaining a reference set of suppliers within the same subset can be effective for supplier development processes. The subset characteristics can also be used to gain insight about the suppliers' performances falling in the same group. Therefore, there is a need to develop a suitable methodology to cluster suppliers. However, much of the data associated with suppliers is categorical which lacks geometric properties upon which the majority of clustering techniques are based. Though there is some recent success in developing clustering algorithms for categorical data, these algorithms do not consider uncertainty during the clustering process. Note uncertainty is an important issue in real world applications since there is often no sharp boundary between clusters. Fuzzy set based algorithms are capable of handling uncertainty yet face other issues such as stability. Thus, the central research issues in this paper are:

What is a suitable algorithm to cluster categorical supplier data for supplier base management?
How can the algorithm help potentially manage the supplier base?

A clustering algorithm, min-min-roughness (MMR), based on rough set theory (RST) (Pawlak 1982) is proposed in this research. There is a three-fold contribution of this research:

(1) a new algorithm is developed for clustering categorical data;
(2) this algorithm is capable of handling uncertainty in the process of clustering;

(3) this algorithm can help managers in obtaining a smaller, more manageable group of suppliers.

This paper is structured as follows: Section 2 presents an overview of standard supplier base management methods that appear in the literature. Section 3 explains the basics of RST followed by a detailed description of the proposed MMR algorithm. We then illustrate MMR using a synthetic dataset and compare the performance of the algorithm with fuzzy set based algorithms using the Zoo dataset (obtained from the UCI Machine Learning Repository; available at http://www.ics.uci.edu/~mlearn/MLRepository.html). In Section 4, we apply the algorithm to cluster supplier data from previous research (Ha and Krishnan 2008) to demonstrate the effectiveness of MMR and summarise the potential managerial insights gained by using MMR. Section 5 provides the conclusion of this work and identifies future research directions.

2. Literature review

Current market trends including globalisation, increased uncertainty, unanticipated customer behaviour, continuously changing business environment, plus the need for flexibility and security, has increased complexities and interdependencies among the various entities of the supply chain. Today, organisations are focusing on core competencies and outsourcing most non-core activities. This has increased the dependence of companies on their suppliers and increased the emphasis on supplier base management. As discussed by Lemke *et al.* (2000):

- Effective supplier management can reduce supply chain costs (Christopher and Peck 2003).
- Involving selected suppliers in new product development can enhance product and process design (Ragatz *et al.* 1997).
- Companies which develop better communication skills with their suppliers achieve better results (Leenders and Fearon 1997).

In general, supplier base management practices are classified into three categories:

(1) supply base rationalisation,
(2) supplier development and
(3) supplier evaluation.

Supply base rationalisation includes identification and elimination of suppliers that are not capable of meeting the company's needs. The outcome of this strategy is a pool of suppliers that are potentially capable of meeting the purchasing organisation's need for products and services. To be a source of competitive advantage, suppliers' performance must be managed and developed to meet the needs of the buying firm (Krause *et al.* 1998).

Supplier development refers to any activity that a buyer undertakes to improve a supplier's performance and/or capacities to meet the buyer's short-term or long-term supply needs (Krause and Handfield 1999). Techniques for supplier development include Kaisen breakthroughs, process mapping, inventory reductions, training, total preventive maintenance, etc. These techniques are complemented by the use of award programmes and increased business for the best suppliers, which serve as incentives for improved performance. Other techniques include introducing competition for the company's business and taking business away from poor performers (Krause and Handfield 1999).

Supplier evaluation includes all efforts expended by companies in evaluating their suppliers using various supplier selection models and techniques. Some basic supplier selection techniques include:

- data envelopment analysis (Charnes *et al*. 1978, Talluri 2000),
- analytic hierarchy process (Satty 1980, Barbarosoglu and Yazgas 1997, Nydick and Hill 1992),
- multi-objective programming (Weber and Ellram 1993),
- total cost of ownership (Ellram 1995),
- statistical analysis (Mummalaneni *et al*. 1996, Choi and Hartley 1996),
- outranking methods (de Boer *et al*. 1998),
- interpretative structural modelling (Mandal and Deshmukh 1994),
- discrete choice analysis (Verma and Pullman 1998), and
- case-based reasoning (Choy *et al*. 2002).

Many extensions of the above-mentioned techniques exist in the literature. For instance, Chan (2003) proposes an interactive selection model to systemise the determination of buyer–supplier relationships and formation of selection criteria, before the implementation of the analytic hierarchy process (AHP). Chan and Chan (2004) illustrate a model which adopts AHP and quality management system principles in the development of the supplier selection model. Humphreys *et al*. (2003) use the case-based reasoning approach to evaluate supplier environmental management performance. They propose a knowledge-based system which integrates environmental factors into the supplier selection process. Jain *et al*. (2004) use an evolutionary fuzzy based approach to evaluate supplier performance.

Note that the three supplier base management practices described above are not independent of each other. Many organisations attempt to apply an effective supplier evaluation system to reduce their supplier base and effectively manage relationships with the suppliers (Tully 1995). While promising, the globalised economy significantly increases the supplier set. Given the size of the larger supplier set and thus diversified performance of the suppliers, the effectiveness of the existing methods is questionable. Lemke *et al*. (2000) point out that if the supplier base is too large, coordination and interaction becomes costly, time consuming and inefficient. One solution is to group the suppliers into smaller segments with similar characteristics. As stated by Talluri and Narasimhan (2004), grouping suppliers based on performance will provide useful insights to management in identifying benchmarks for ineffective suppliers and assist in decisions relating to supplier development. Unfortunately, the importance of segmenting suppliers is simply realised and research on suitable methodologies to cluster suppliers into smaller groups is less explored. Among the few documented efforts, we note that Holt (1998) explains the use of cluster analysis for pre-qualifying all suppliers. Hong *et al*. (2005) and Ha and Krishnan (2008) explore the use of self organising maps (SOMs) for supplier clustering. In this paper, a data mining clustering algorithm based on RST is proposed to segment a large number of suppliers into smaller groups which share similar characteristics. This algorithm is explained in the following section.

3. Min-min-roughness (MMR) algorithm

3.1 *Nomenclature*

U	Universe or the set of all objects $(x_1, x_2, \ldots)$.
X	Subset of the set of all objects ($X \subset U$).
x_i	Object belonging to the subset of the set of all objects, $x_i \in X$.
A	The set of all attributes (features or variables).
a_i	Attribute belonging to the set of all attributes, $a_i \in A$.
$V(a_i)$	Set of valid values of attribute a_i (or called domain of a_i).
B	Non-empty subset of A ($B \subseteq A$).
$\overline{X_B}$	Upper approximation of X with respect to B.
$\underline{X_B}$	Lower approximation of X with respect to B.
$R_{ai}(X)$	Roughness with respect to $\{a_i\}$.
MR (a_i)	Minimum roughness of attribute a_i.
$Rough_{aj}(a_i)$	Mean roughness on attribute a_i with respect to $\{a_j\}$.
MMR	Minimum of MR of all attributes.
Ind (B)	Indiscernibility relation.
$[x_i]_{Int(B)}$	Equivalence class of x_i in relation $Ind(B)$, also known as elementary set in B.

3.2 *Rough set theory*

Rough set theory (RST) was introduced by Pawlak (1982) and was developed to classify uncertain or incomplete information. It complements fuzzy set theory (Dubois and Prade 1990). The concept of rough sets assumes that objects are characterised by their features. For instance, objects can be described with geometric features such as length, width, height. Objects described by identical data are indiscernible. In RST, a set of all similar objects is called elementary, and it makes up a fundamental atom of knowledge (Pawlak 1982). Any union of elementary sets is referred to as a precise set as opposed to an imprecise (rough) set. As a result, each rough set has boundary-line objects. For example, some objects cannot be classified for sure as members of the set or its complement. In other words, when the available knowledge is employed, boundary-line cases cannot be properly classified. Therefore, rough sets can be considered as uncertain.

In RST, let U be the set of all objects, A be the set of all attributes, B be a non-empty subset of A, $Ind(B)$ is a relation on U. Given two objects, x_i, $x_j \in U$, they are indiscernible by the set of attributes B in A, if and only if $a(x_i) = a(x_j)$ for every $a \in B$. That is, $(x_i, x_j) \in Ind(B)$ if and only if $\forall a \in B$, where $B \subseteq A$ and $a(x_i) = a(x_j)$, $Ind(B)$ is called indiscernibility relation. Given $Ind(B)$, the set of objects x_i having the same values for the set of attributes in B consists of an elementary set, represented by $[x_i]_{Ind(B)}$. For a set of objects X, the lower approximation is defined as the union of all the elementary sets which are contained in X, and the upper approximation is defined as the union of the elementary sets which have a non-empty intersection with X, that is,

The lower approximation of X is presented by

$$\underline{X_B} = \cup \{ x_i | [x_i]_{Ind(B)} \subseteq X \},$$

and the upper approximation of X is presented by

$$\overline{X_B} = \{x_i \in U | [x_i]_{Ind(B)} \cap X \neq 0\}.$$

The ratio of the cardinality of the lower approximation and the cardinality of the upper approximation is defined as the accuracy of estimation, which is a measure of roughness. The roughness is presented as:

$$R_B(X) = 1 - \frac{|\underline{X_B}|}{|\overline{X_B}|}$$

If $R_B(X) = 0$, X is *crisp* with respect to B, in other words, X is *precise* with respect to B. If $R_B(X) < 1$, X is *rough* with respect to B, that is, B is *vague* with respect to X. The concept of roughness enables RST the ability to deal with uncertainty since the calculation of the lower approximation and the upper approximation gives different objects a different degree of belonging to the group.

Recently, the RST has been applied to a variety of problems and some of the applications are summarised in Table 1. The majority of research has focused on supervised learning (also known as classification), that is, the prior group membership is known. Results usually generated are rules for group membership (Pawlak 1982). There also exist some attempts to introduce rough set into the clustering for managing imprecise concepts. Lingras and West (2004) propose a modified K-means algorithm to create intervals of clusters based on RST. Voges *et al.* (2002a, b) propose a technique called rough clustering, which is a simple extension of RST, and apply it to the problem of market segmentation. However, the majority of research in exploring RST for clustering aims to handle numerical data sets, where distance can be easily derived from the data set. Other than the vagueness introduced by the physical measurement, linguistic vagueness in terms of characterisation of members of the same class in different ways, which usually is represented as categorical data, has also been studied. For instance, Mitra *et al.* (2003) suggest the membership of each object is initialised with fuzziness, that is, the membership is represented by three fuzzy linguistic sets: low, medium and high. RST is applied to help faster convergence and avoidance of the local minima problem thereby enhancing the performance of a well known clustering algorithm termed expectation maximisation (EM) algorithm. Greco *et al.* (2001) apply a dominance-based RST approach to generate linguistic rules that represent a decision-maker's preference based on the classification of a test dataset. Yet, the focus has been either on how to effectively extract knowledge/rules from existing data by using RST concept to represent the vagueness of membership, or on how to incorporate the semantic data which represents the human-comprehensible concepts into the decision models. Please note the supplier base management fundamentally differs from the problems reviewed above, that is,

(1) it is an unsupervised learning problem,
(2) the supplier database is categorical in nature,
(3) the vagueness of membership is implicitly embedded in the data.

Therefore, we propose a new RST-based clustering algorithm, aimed at reducing the vagueness in the clustering process, in this research.

Table 1. RST applications.

Applications	Researchers	Description
Medical decision making	Kusiak *et al*. (2000)	Analysis of large datasets to identify the key factors in a medical dataset
Customer retention	Kowalczyk and Slisser (1997)	Development of rough data model involving clustering and a linear ordering on clusters to analyse customer retention related data
Neural networks	Szczuka and Slezak (1997)	Rough set methods to construct a network with parameters corresponding to decision classes
Data filtering	Skowron (1994)	Rough set approach to searching for new classifiers (features) for data filtering
Intelligent image filtering	Wojcik (1993)	Effective removal of noise and enhancing edges to represent a class of novel, high-quality image filters.
Risk management	Dimitras *et al*. (1999)	Rough set based approach to rule extraction to discriminate between healthy and failing firms for risk management
Software engineering	Ruhe (1996)	Finding the interaction between goal-oriented measurements and rough sets based analysis in the context of experimental software engineering
Web mining	Lingras and West (2004)	Extraction of structure from a dataset containing the characteristics of users' information.
Marketing data examination	Voges *et al*. (2002a, b)	Used for market segmentation

3.3 *Min-min-roughness*

The first attempt to use the RST for clustering categorical data was conducted by Mazlack *et al*. (2000). A total roughness is created to determine the crispness of a partition. The attribute with the lowest total roughness value might be selected for partition. However, for partitioning, the method starts with binary valued attributes and uses the total roughness criterion only for multi-valued attributes. This suffers from a considerable handicap, for instance, the partitioning is done on a binary attribute even though the total roughness for a multi-valued attribute is lower. Therefore, we advance the roughness and total roughness concept and propose a new measurement, called min-min-roughness (MMR), to cluster the objects (such as suppliers) on all attributes.

As defined by Mazlack *et al.* (2000), given any attribute $a_i \in A$, $V(a_i)$ refers to the set of values of attribute a_i, X is a subset of objects having one specific value, α, of attribute a_i, that is, $X(a_i = \alpha)$, $\underline{X_{a_j}(a_i = \alpha)}$ refers to the lower approximation, and $\overline{X_{a_j}(a_i = \alpha)}$ refers to the upper approximation with respect to $\{a_j\}$, then $R_{aj}(X)$ is defined as the *roughness* of X with respect to attribute a_j, that is

$$R_{a_j}(X|a_i = \alpha) = 1 - \frac{\left|\underline{X_{a_j}(a_i = \alpha)}\right|}{\left|\overline{X_{a_j}(a_i = \alpha)}\right|},$$

where a_i, $a_j \in A$ and $a_i \neq a_j$.

Let $|V(a_i)|$ be the number of values of attributes a_i, the mean roughness on attribute a_i with respect to $\{a_j\}$ is defined as

$$Rough(a_i)_{(aj)} = \frac{R_{a_j}(X|a_i = \alpha_1) + \cdots R_{a_j}(X|a_i = \alpha_{|V(a_i)|})}{|V(a_i)|} \tag{1}$$

In MMR, given n attributes, min-roughness of attribute a_i ($a_i \in A$) refers to the minimum of the mean roughness, that is,

$$\mathrm{MR}(a_\mathrm{i}) = \min\big(Rough_{a_1}(a_i), \ldots Rough_{a_1}(a_i)\ldots\big), \quad \text{where } a_i, a_j \in A,\ a_i \neq a_j,\ 1 \leq i,\ j \leq n \tag{2}$$

The minimum of the min-roughness of the n attributes is calculated as min-min-roughness, that is,

$$\mathrm{MMR} = m(\mathrm{MR}(a_1), \ldots \mathrm{MR}(a_\mathrm{i}), \ldots) \text{ where } a_\mathrm{i} \in \mathrm{A}, \quad i \text{ goes from 1 to cardinality}(A) \tag{3}$$

Based on the MMR, the algorithm (as shown in Table 2) iteratively divides the group of objects aiming to achieve better clustering crispness. The algorithm takes the number of clusters, k, defined by the user, as one input and will terminate when the number k is reached.

3.4 *MMR example*

Let us use a simple example (Table 3) to illustrate the MMR algorithm. In this example, there are eight objects ($|U| = 8$) and three attributes ($|A| = 3$). Our interest is to create clusters of similar objects. As seen from the dataset, variables can be multi-valued. That is, the domain of an attribute can contain more than two distinct values.

Step 1: *Initialise*: First, the dataset is loaded. Second, the desired number of clusters is input by the user. This is subjective and is pre-decided based either on user requirements or domain knowledge.

Step 2: *Split current dataset (procedure MMRMain)*:

Step 2.1: *Calculate mean roughness*: The mean roughness for each attribute a_i ($i = 1, \ldots, 3$) is calculated using Equation (1). For example, for attribute a_1, there are two elementary sets: 'High' (1, 2, 3, 4, 5) and 'Low' (6, 7, 8). We now calculate the mean roughness of these two sets with respect to all other attributes. Let us calculate the mean

Table 2. MMR algorithm.

Set $U = \{x_1, x_2, \ldots\}$//Load dataset
Set $k =$ *desired number of clusters specified by the user*//Input desired number of clusters
Set CNC =1//Set Current Number of Clusters to be 1
Set Clustering [] = $\emptyset$
Procedure **MMR (*U*, *k*)**
Begin
    Cur_Dataset = U//Set Current Dataset to be the entire dataset as no partition has occurred.
    Do until CNC > k//If the Current Number of Clusters is less than the desired, algorithm continues
        Call **MMRMain (Cur_Dataset)**
        Cur_Dataset = DetermineNextSplittingDataset (CNC)
    Loop
End
Procedure **MMRMain (Cur_Dataset)**
    For each $a_i \in A$ ($i = 1$ to n, where n is the number of attributes in A)
        //Calculates the Min-Roughness for each attribute
        Determine $\left[x_j\right]_{Ind(a_i)}$//Calculates the elementary sets for attribute a_i
        For each $a_j \in A$($j = 1$ to n, where n is the number of attributes in A, $j \neq i$)
        //Calculates the Roughness with respect to the remaining attributes
            Calculate $Rough_{a_j}(a_i)$ (see Equation (1))
        Next
        Min-Roughness $(a_i) = \text{Min}\ (Rough_{a_j}(a_i))$ (see Equation (2))
    Next
    Set min-min-roughness = min (min-roughness (a_i)) (see Equation (3))
    Determine splitting attribute a_i corresponding to the min-min-roughness
        //Identifying attribute for splitting
    Determine splitting point on attribute a_i
    //All possible splitting points on a_i are evaluated using the min-min-roughness approach
    Do binary split on the attribute a_i on the splitting point
    Set the split datasets to Clustering (CNC) and Clustering (CNC + 1)
    CNC = CNC + 1
End
Procedure **DetermineNextSplittingDataset (CNC)**
    Set $i = 1$
    Do until $i >$ CNC
        Size (i) = Count (Clustering (i))
        $i = i + 1$
    Loop
    Find max (size (i))//Determine the cluster with the largest number of elements.
    Return (Clustering (i)) corresponding to max (size (i))
End

Table 3. Example dataset.

Objects	a_1	a_2	a_3
x_1	High	Hard	One
x_2	High	Hard	Two
x_3	High	Medium	Three
x_4	High	Soft	Four
x_5	High	Soft	One
x_6	Low	Hard	Two
x_7	Low	Medium	Three
x_8	Low	Soft	Four

Table 4. Min-min-roughness calculation.

Attributes	Mean roughness	Min roughness
a_1	$R_{a_2}(a_1) = 1$ $R_{a_3}(a_1) = 0.875$	0.875
a_2	$R_{a_1}(a_2) = 1$ $R_{a_3}(a_2) = 0.333$	0.333
a_3	$R_{a_1}(a_3) = 1$ $R_{a_2}(a_3) = 0.750$	0.750
		MMR: 0.333 (a_2)

roughness of a_1 with respect to a_3 as an example. There are four elementary sets for a_3: 'One' (1, 5), 'Two' (2, 6), 'Three' (3, 7) and 'Four' (4, 8). The lower approximation of 'Low' (6, 7, 8) with respect to a_3 is empty, that is, $\left|\underline{X}_{a_3}(a_1 = \text{``Low''})\right| = 0$, thus, the roughness is

$$R_{a_3}(X|a_1 = \text{``Low''}) = 1 - \frac{\left|\underline{X}_{a_3}(a_1 = \text{``Low''})\right|}{\left|\overline{X}_{a_3}(a_1 = \text{``Low''})\right|} = 1.$$

The lower approximation of 'High' $(1, 2, 3, 4, 5)$ with respect to a_3 is (1,5) (elementary set 'One' with cardinality as 2) and the upper approximation is (1,2,3,4,5,6,7,8) (sum of all four elementary sets with cardinality as 8), that is, $\left|\underline{X}_{a_3}(a_1 = \text{``High''})\right| = 2$, and $\left|\overline{X}_{a_3}(a_1 = \text{``High''})\right| = 8$, thus, the roughness is $R_{a_3}(X|a_1 = \text{``High''}) = 1 - \frac{2}{8} = 0.75$, The mean roughness of a_1 with respect to a_3, $R_{a_3}(a_1)$ is 0.875. Similar calculations are performed for all the attributes.

Step 2.2: *Determine splitting attribute*: The partitioning attribute is determined based on the MMR. Given the mean roughness of each attribute, MR and MMR are calculated based on the attribute used to determine MMR. As shown in Table 4, attribute a_2, having the smallest MMR, is chosen to partition the dataset. Note that ties can be broken by randomly selecting any tied attribute.

Step 2.3: *Determine Splitting point*: Binary splitting is chosen as it prevents the fragmentation of data too quickly which might leave insufficient data at the next level (Hastie *et al.* 2001). In the situation where the splitting attribute has binary value, splitting is straightforward. Given the splitting attribute having multiple values ($|V(a_i)| > 2$), there are $C^2_{|V(a_i)|}$options of splitting points. For each option, MMR is calculated to determine the splitting point. In this example, attribute a_2 is selected as the partitioning attribute. There are three possible split points on attribute a_2:

Split point 1 → 'Hard' vs. 'Medium-Soft' → (1, 2, 6) and (3, 4, 5, 7, 8)
Split point 2 → 'Medium' vs. 'Hard-Soft' → (3, 7) and (1, 2, 4, 5, 6, 8)
Split point 3 → 'Soft' vs. 'Hard-Medium' → (4, 5, 8) and (1, 2, 3, 6, 7)

For each possible split point, mean roughness, min-roughness followed by MMR are calculated. As shown in Table 5, Split point 2 is chosen for splitting as it has MMR.

Table 5. Min-min roughness to decide the split point for attribute a_2.

Splits	Mean roughness	Min roughness
Hard' vs. 'Medium-Soft'	$R_{a_1}(a_2) = 1$ $R_{a_3}(a_2) = 0.417$	0.417
'Medium' vs. 'Hard-Soft'	$R_{a_1}(a_2) = 1$ $R_{a_3}(a_2) = 0$	0
'Soft' vs. 'Hard-Medium'	$R_{a_1}(a_2) = 1$ $R_{a_3}(a_2) = 0.417$	0.417
		MMR: 0 ('Medium' vs. 'Hard-Soft')

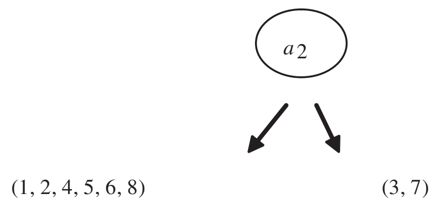

Figure 1. Partition after first iteration.

Step 2.4: *Partition dataset*: The original dataset is divided into two sets based on the split point on the split attribute. In this example, the dataset is divided into two on attribute a_2 on medium and hard-soft values. The partition at this stage can be represented as a binary tree and is shown in Figure 1.

The numbers in the parenthesis at each of the child nodes correspond to the objects in the original dataset. Set (1, 2, 4, 5, 6, 8) corresponds to all objects having either 'Hard' or 'Soft' as a value for attribute a_2 and set (3, 7) corresponds to all objects having 'Medium' as a value for attribute a_2.

Step 3: *Select dataset for further splitting* (*procedure DetermineNextSplittingDataset*): Choose the largest dataset after partitioning and then repeat the process in Step 2 until the number of clusters reaches the pre-defined number.

3.5 *Implementation and comparison experiments*

MMR is implemented in VB. Net and tested on the Zoo dataset obtained from the UCI Machine Learning Repository. MMR, based on RST, is similar to algorithms based on fuzzy set theory (FST) in that both RST and FST handle uncertainty in the data and both allow membership of an object in more than one cluster. We therefore decide to compare MMR with the algorithms based on fuzzy set theory namely, *K*- modes, fuzzy *K*-modes and fuzzy centroids (Kim *et al*. 2004) using overall purity of clustering (Overall_PoC), similar measure as in Kim *et al*. (2004) and Guha *et al*. (2000). PoC is defined as the

Table 6. Applying MMR on the Zoo dataset ($k=7$).

Cluster	1	2	3	4	5	6	7	PoC (%)
1	39	0	0	0	0	0	0	100
2	0	20	0	0	0	0	0	100
3	0	0	1	0	1	0	0	50
4	0	0	1	13	0	0	0	93
5	0	0	3	0	3	0	0	50
6	2	0	0	0	0	6	0	75
7	0	0	0	0	0	2	10	83
Overall_PoC = 91%								

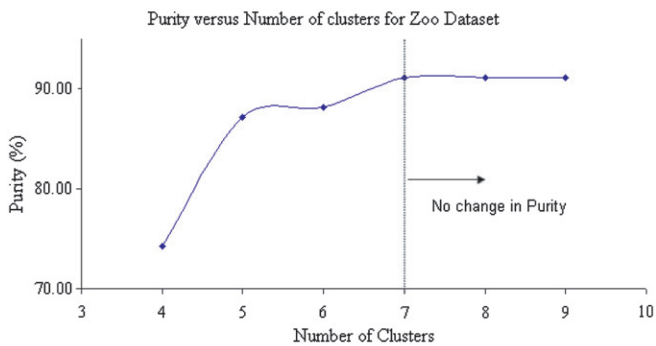

Figure 2. Purity versus number of clusters for Zoo dataset.

number of objects being correctly clustered in the group over the number of all objects in the group. Since Zoo dataset is a benchmark dataset provided by UCI machine learning repository, the correct classification of the dataset is a known fact. In Table 6, the columns represent the correct classification based on the original Zoo dataset. The rows represent the classification obtained from MMR. Let us take cluster 4 from Table 6 as an example, the given dataset should have 13 objects belonging to cluster 4, however, MMR identifies 14 objects in cluster 4, which indicates a misclassification of 1 object. Given cluster i, we define

$$PoC(i) = \frac{\textit{the \# of objects being correctly classifed to ith cluster}}{\textit{the \# of objects in ith cluster}}\%$$

We conclude PoC of cluster 4 is 92%. Considering the size of each cluster varies, overall_PoC is derived based on the classification of all objects instead of the average of PoC(i), that is:

$$Overall\_POC = \frac{\sum_{i=1}^{\#\text{of clusters}} \text{the \# of objects being correctly classified to the ith cluster}}{\text{total \# of objects}}\%$$

The Zoo dataset has 101 objects, where each object represents information of an animal in terms of 18 categorical attributes. As shown in Figure 2, several experiments are conducted to test the change of PoC on the number of clusters. It is interesting to note that

Table 7. Comparison between MMR with algorithms based on fuzzy set theory (on Overall_PoC).

Dataset	*K*-modes	Fuzzy *K*-modes	Fuzzy centroids	MMR
Zoo	60%	64%	75%	91%

Overall_PoC increases to more than 90% when the number of clusters changes from four to seven and keeps constant thereafter. In addition, the algorithms based on FST all generate seven clusters for the Zoo dataset (Kim *et al*. 2004). Therefore, for comparison purposes, the MMR experiment with seven set as the number of clusters is summarised in Table 6.

As shown in Table 6, clusters 1 and 2 achieve 100% PoC as all 39, 20 objects are correctly identified for the specific clusters, respectively. For clusters 3 to 7, 1 out of 2, 13 out of 14, 3 out of 6, 6 out of 8 and 10 out of 12 are correctly classified to the specific clusters. The PoCs are 50%, 93%, 50%, 75% and 83% respectively. The overall_PoC is 91%. Table 7 summarises the results of comparison of MMR with *K*-modes, fuzzy *K*-modes and fuzzy centroid.

Clearly, MMR outperforms the algorithms based on FST in terms of purity of clustering. In addition, one issue with the algorithms based on FST is that a control parameter, m, presented as the power of the membership degree of the data to the specific cluster, is required to control the fuzziness of membership of each object. These algorithms need to be run at different values of this control parameter, and at each value of the control parameter, the algorithms need to be run multiple times to achieve a stable solution. Such computational efforts to adjust the parameter are saved in MMR.

4. Applying MMR to cluster suppliers

The above has described the basis for MMR with comparison results presented. In this section, the application of MMR to supplier base management is demonstrated using the supplier dataset adopted from Ha and Krishnan (2008). The data is from an automobile company which mainly manufactures automatic transmission parts. Ha and Krishnan (2008) apply SOM to segment 27 suppliers into six clusters based on seven quantitative characteristics:

(1) Quality system outcome (QSO).
(2) Claims (CL).
(3) Quality improvement (QI).
(4) Response to claims (RC).
(5) On-time delivery (OD).
(6) Internal audit (IA).
(7) Data administration (DA).

Though the characteristics should be categorical in nature, numerical data is presented in Ha and Krishnan (2008) for ease of SOM application. The original numerical dataset is shown in Table 8 and the SOM clustering results are summarised in Table 9.

To apply MMR, we first discretise the data using an existing BN software package (available at http://www.cs.ualberta.ca/~jcheng/download.htm) which uses unsupervised discretisation algorithm to divide the attribute range into a constant number of intervals

Table 8. Supplier dataset adopted from Ha and Krishnan (2008).

Supplier	QSO	CL	QI	RC	OD	IA	DA
C1	4.5	5	9.5	3	3	3	27.5
C2	4	1	7	2	2	2.5	26
C3	4.5	5	8	1	1	2	22.6
C4	5	5	9	1	1	5	28.5
C5	4.3	5	7.7	3	3	4	25.4
C6	2.5	5	8	1	1	3	23.7
C7	5	1	10	2	2	4	24.5
C8	4	1	8	2	2	2.5	27.9
C9	5	5	10	3	3	1.5	27.5
C10	4	5	10	3	3	4.5	27.5
C11	5	5	10	3	3	3.5	27.5
C12	5	1	10	2	2	2.5	22.5
C13	3	5	10	3	3	3	22.5
C14	5	5	10	3	3	4	30
C15	5	5	10	3	3	4	28
C16	5	5	10	3	3	3	28
C17	4.5	5	8	1	1	3.5	22
C18	3.5	5	7.5	2.5	2.5	3.5	24
C19	4	5	8	2	2	3	22.5
C20	4	5	8.5	1	1	3.5	26.3
C21	4	5	7	3	3	4.5	26.5
C22	4	1	8	2	2	4	24.5
C23	5	1	10	2	2	4	24.5
C24	5	5	10	3	3	4.5	29
C25	5	5	10	3	3	4	28
C26	4	5	10	3	3	5	29.5
C27	4	5	9	3	3	4	29

Table 9. Supplier clusters from SOM.

Cluster	Member suppliers
1	C3, C4, C6, C17, C20
2	C2, C7, C8, C12, C22, C23
3	C5, C18, C21
4	C1, C13, C27
5	C9, C10, C11, C14, C15, C16, C24, C25, C26
6	C19

containing an equal number of the attribute values. The results after discretisation are summarised in Table 10.

MMR is now applied on this discretised dataset and six clusters are obtained so that the results can be compared with the supplier clusters obtained using SOM.

In the earlier comparison study on the Zoo dataset, overall_PoC can be derived since the correct clustering results are known. In this experiment, we use similar measures as that from Ha and Krishnan (2008), that is, the mean and standard deviation of each cluster on

Table 10. Discretised dataset.

Supplier	QSO (0, 1)	CL (1, 5)	QI (0, 1)	RC (0, 1)	OD (0, 1)	IA (0, 1, 2)	DA (0, 1, 2, 3, 4)
C1	0	5	0	0	0	0	2
C2	1	1	1	1	1	0	1
C3	0	5	1	1	1	0	0
C4	0	5	0	1	1	2	3
C5	0	5	1	0	0	1	1
C6	1	5	1	1	1	0	0
C7	0	1	0	1	1	1	1
C8	1	1	1	1	1	0	3
C9	0	5	0	0	0	0	2
C10	1	5	0	0	0	2	2
C11	0	5	0	0	0	1	2
C12	0	1	0	1	1	0	0
C13	1	5	0	0	0	0	0
C14	0	5	0	0	0	1	4
C15	0	5	0	0	0	1	3
C16	0	5	0	0	0	0	3
C17	0	5	1	1	1	1	0
C18	1	5	1	0	0	1	1
C19	1	5	1	1	1	0	0
C20	1	5	0	1	1	1	2
C21	1	5	1	0	0	2	2
C22	1	1	1	1	1	1	1
C23	0	1	0	1	1	1	1
C24	0	5	0	0	0	2	3
C25	0	5	0	0	0	1	3
C26	1	5	0	0	0	2	4
C27	1	5	0	0	0	1	3

Table 11. Supplier clusters from MMR on discretised dataset.

Cluster	Member suppliers
1	C3, C4, C6, C17, C19, C20
2	C2, C7, C8, C12, C22, C23
3	C5, C18, C21
4	C10, C13, C26, C27
5	C1, C9, C11, C14, C15, C16, C25
6	C24

the seven characteristics to evaluate the degree of homogeneous of MMR comparing to SOM (Table 12). From Table 9 and Table 11, we conclude, the six clusters identified by MMR and SOM have 81.5% match. Given both MMR and SOM have one cluster with only one supplier, we further study the standard deviations of the remaining five clusters on the seven characteristics. Compared with SOM (as shown in Table 12), nine out of 35 MMR results have smaller standard deviations, four out of 35 MMR results have larger

Table 12. Comparing study results.

	Characteristics of each cluster							
	QSO	CL	QI	RC	OD	IA	DA	
	Average	Average	Average	Average	Average	Average	Average	
Cluster	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Member suppliers
1 (SOM)	4.10	5.00	8.30	1.00	1.00	3.40	24.62	C3, C4, C6,
	0.96	0.00	0.45	0.00	0.00	1.08	2.72	C17, C20
1	4.08	5.00	8.25	1.17	1.17	3.33	24.27	C3, C4, C6,
	0.86	0.00	0.42	0.41	0.41	0.98	2.59	C17, C20, C19
2 (SOM)	4.50	1.00	8.83	2.00	2.00	3.25	24.98	C2, C7, C8,
	0.55	0.00	1.33	0.00	0.00	0.82	1.81	C12, C22, C23
2	4.50	1.00	8.83	2.00	2.00	3.25	24.98	C2, C7, C8,
	0.55	0.00	1.33	0.00	0.00	0.82	1.81	C12, C22, C23
3 (SOM)	3.93	5.00	7.40	2.83	2.83	4.00	25.30	C5, C18, C21
	0.40	0.00	0.36	0.29	0.29	0.50	1.25	
3	3.93	5.00	7.40	2.83	2.83	4.00	25.30	C5, C18, C21
	0.40	0.00	0.36	0.29	0.29	0.50	1.25	
4 (SOM)	3.83	5.00	9.50	3.00	3.00	3.33	26.33	C1, C13, C27
	0.76	0.00	0.50	0.00	0.00	0.58	3.40	
4	3.75	5.00	9.75	3.00	3.00	4.13	27.13	C10, C13, C26,
	0.50	0.00	0.50	0.00	0.00	0.85	3.20	C27
5 (SOM)	4.78	5.00	10.00	3.00	3.00	3.78	28.33	C9, C10, C11, C14,
	0.44	0.00	0.00	0.00	0.00	1.03	0.94	C15, C16, C24, C25, C26
5	4.93	5.00	9.93	3.00	3.00	3.29	28.07	C9, C11, C14,
	0.19	0.00	0.19	0.00	0.00	0.91	0.89	C15, C16, C25, C1
6 (SOM)	4.00	5.00	8.00	2.00	2.00	3.00	22.50	C19
6	5.00	5.00	10.00	3.00	3.00	4.50	29.00	C24

standard deviations (highlighted) and 22 out of 35 MMR results have the same standard deviations. For the four under-performed clusters, we further carefully examine the data from Table 8 (original dataset) and Table 10 (discretised dataset), we conclude that larger standard deviations are mainly introduced by discretised errors instead of the misclassification from the MMR algorithm. For example, the difference between cluster 1 from MMR and SOM is C19. In the original dataset (Table 8), C19 has 2 for RC and 2 for OD while all other suppliers have 1 for RC and 1 for OD. However, after the discretisation, all the suppliers in MMR cluster 1 have the same values for both RC and OD. Therefore, we conclude MMR could generate comparable homogenous clusters to SOM, such performance can be further strengthened if the discretisation errors can be reduced which will be our immediate next step. However, training SOM is relatively computationally expensive as it requires the adjustment of several parameters such as the size of the neighbourhood, the learning rate while MMR simply sets six as the number of clusters to achieve comparable results.

Given the clustering results, some managerial insights can be gained by summarising the characteristics of each cluster. For instance, cluster 1 suppliers deliver auto parts on-time, but they receive many claims from the purchaser and respond promptly to resolve them. Cluster 2 obtains less number of claims and deals with them quickly (above average). Cluster 3 is weak in improving quality. Cluster 4 receives many claims and has low scores in response to the claims, on-time delivery and quality system outcome. Cluster 5 has better quality outcome however underperforms in response to claims and on-time delivery. If the purchaser needs supply sources that perform well in terms of responsiveness, it can select a supplier from either cluster 1 or cluster 2.

5. Conclusion

Considering the complex nature of supply chain systems, special attention should be paid to supplier base management to better handle supplier delivery exceptions. Managing a massive supplier base is a challenging task. One of the key trends observed in recent years is for manufacturers to reduce their total number of suppliers. There is a need for an algorithm that can be used to cluster suppliers into smaller, more manageable subsets. However, little supplier clustering work has been done, especially, when categorical data are involved and uncertainty of belonging to a cluster is considered. This paper presents an algorithm, named MMR, to deal with uncertainty in the process of clustering categorical supplier data. MMR is then applied to clustering suppliers. The benefits of group homogeneous suppliers include:

(1) in the case of multiple sourcing, it provides the manager a reasonable size of suppliers within a different group and the characteristics of each group can be summarised;
(2) benchmarking within the same group instead of all suppliers has the potential to help the underperforming supplier locate more realistic benchmarks thus make supplier development more effective.

We will further assess this potential benefit in the future research. While the research presented is potentially quite useful, there are a number of issues that remain to be addressed. First, MMR is developed mainly for categorical data. We will extend MMR for mixed datasets containing both numerical including interval data and categorical data. Second, we will test this algorithm using a less subjective stopping criterion rather than a pre-defined number of clusters. Third, we will utilise real industry data to test the algorithm. We will also study various discretisation methods and the impact of the discretisation on the clustering performance.

References

Barbarosoglu, G. and Yazgas, T., 1997. Application of the analytic hierarchy process to the supplier selection problem. *Production and Inventory Management Journal*, 38 (1), 14–21.

Chan, F., 2003. Interactive selection model for supplier selection process: an analytic hierarchy process approach. *International Journal of Production Research*, 41 (15), 3549–3579.

Chan, F. and Chan, H., 2004. Development of the supplier selection model – a case study in the advanced technology industry. *Proceedings of the Institution o f Mechanical Engineers, Part B, Journal of Engineering Manufacture*, 218 (12), 1807–1824.

Charnes, A., Cooper, W., and Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.

Choi, T. and Hartley, J., 1996. An exploration of supplier selection practices across the supply chain. *Journal of Operations Management*, 14, 333–343.

Choy, K., Lee, W., and Lo, V., 2002. An intelligent supplier management tool for benchmarking suppliers in outsource manufacturing. *Expert Systems Application*, 22 (3), 213–224.

Christopher, M. and Peck, H., 2004. The five principles of supply chain resilience. *Logistics Europe*, 12 (1), 16–21.

Christopher, M. and Peck, H., 2003. *Marketing logistics*. Oxford, England: Butterworth-Heinemann.

de Boer, L., van der Wegen, L., and Telgen, J., 1998. Outranking methods in support of supplier selection. *European Journal of Purchasing & Supply Management*, 4 (2), 109–118.

Dimitras, A.I., *et al.*, 1999. Business failure prediction using rough sets. *European Journal of Operational Research*, 114 (2), 263–280.

Dubois, D. and Prade, H., 1990. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17, 191–209.

Ellram, L., 1995. Total cost of ownership: an analysis approach for purchasing. *International Journal of Physical Distribution & Logistics Management*, 25 (8), 4–23.

Greco, S., Matarazzo, B., and Slowinski, R., 2001. Rough set theory for multicriteria decision analysis. *European Journal of Operational Research*, 129, 1–47.

Guha, S., Rastogi, R., and Shim, K., 2000. ROCK: a robust clustering algorithm for categorical attributes. *Information Systems*, 25, 345–366.

Ha, S. and Krishnan, R., 2008. A hybrid approach to supplier selection for the maintenance of a competitive supply chain. *Expert Systems with Applications*, 34 (2), 1303–1311.

Hastie, T., Tibshirani, R., and Friedman, J., 2001. *The elements of statistical learning – data mining, interference, and prediction*. New York: Springer-Verlag.

Hicks, M., 2002. When the chain snaps. *eWeek Enterprise News & Reviews*, February 18.

Holt, G.D., 1998. Which contractor selection methodology? *International Journal of Project Management*, 16 (3), 153–164.

Hong, G.H., *et al.*, 2005. An effective supplier selection method for constructing a competitive supply relationship. *Expert Systems with Applications*, 28, 629–639.

Humphreys, P., McIvor, R., and Chan, F., 2003. Using case-based reasoning to evaluate supplier environmental management performance. *Expert Systems with Applications*, 25, 141–153.

Jain, V., Tiwari, M., and Chan, F., 2004. Evaluation of the supplier performance using an evolutionary fuzzy-based approach. *Journal of Manufacturing Technology Management*, 15 (8), 735.

Kim, D., Lee, K., and Lee, D., 2004. Fuzzy clustering of categorical data using fuzzy centroids. *Pattern Recognition Letters*, 25, 1263–1271.

Kowalczyk, W. and Slisser, F., 1997. Analyzing customer retention with rough data models. *Proceedings of the 1st European symposium on principles of data mining and knowledge discovery* (*PKDD*–1997), Trondheim, Norway. Lecture Notes in AI 1263. London: Springer-Verlag, 4–13.

Krause, D. and Handfield, R., 1999. *Developing a world-class supply base* [online]. Center for Advanced Purchasing Studies. Available from: http://www.capsresearch.org/Login.aspx?ReturnURL=publications/pdfs-protected/krause1999.pdf [Accessed May 2009].

Krause, D., Handfield, R., and Scannel, T., 1998. An empirical investigation of supplier development: reactive and strategic processes. *Journal of Operations Management*, 17, 39–58.

Kusiak, A., *et al.*, 2000. Autonomous decision-making: a data mining approach. *IEEE Transactions on Information Technology in Biomedicine*, 4 (4), 274–284.

Leenders, M. and Fearon, H., 1997. *Purchasing and supply management*. 11th ed. London, England: Irwin.

Lemke, F., *et al.*, 2000. Supplier base management: experiences from the UK and Germany. *International Journal of Logistics Management*, 11 (2), 45–48.

Lingras, P. and West, C., 2004. Interval set clustering of web users with rough K-means. *Journal of Intelligent Information System*, 23 (1), 5–16.

Mandal, A. and Deshmukh, S., 1994. Vendor selection using interpretative structural modeling (ISM). *International Journal of Operations & Production Management*, 14 (6), 52–59.

Mazlack, L., *et al.*, 2000. A rough set approach in choosing partitioning attributes. *Proceedings of the ISCA 13th international conference* (*CAINE*-2000), November 2000, Honolulu, Hawaii, USA. IEEE Computer Society, 1–6.

Mitra, P., Pal, S.K., and Siddiqi, M.A., 2003. Non-convex clustering using expectation maximisation algorithm with rough set initialisation. *Pattern Recognition Letters*, 24, 863–873.

Mummalaneni, V., Dubas, K., and Chao, C., 1996. Chinese purchasing managers preferences and trade-offs in supplier selection and performance evaluation. *Industrial Marketing Management*, 25 (2), 115–124.

Nydick, R. and Hill, R., 1992. Using the analytic hierarchy process to structure the supplier selection procedure. *International Journal of Purchasing and Materials Management*, 28 (2), 31–36.

Pawlak, Z., 1982. Rough sets. *International Journal of Computer and Information Sciences*, 11, 341–356.

Ragatz, G., Hanfield, R., and Scannel, T., 1997. Success factors for integrating suppliers into new product development. *Journal of Product Innovation Management*, 14 (3), 190–202.

Ruhe, G., 1996. Qualitative analysis of software engineering data using rough sets. *Proceedings of the 4th international workshop on rough sets, fuzzy sets, and machine discovery*, Tokyo, Japan. London: Springer-Verlag, 292–299.

Satty, T., 1980. *The analytical hierarchical process*. New York: McGraw-Hill.

Skowron, A., 1994. Data filtration: a rough set approach. *Proceedings of the international workshop on rough sets and knowledge discovery*, Banff, Alberta, Canada. London: Springer-Verlag, 108–118.

Szczuka, M. and Slezak, D., 1997. Hyperplane-based neural networks for real-valued decision tables. *Proceedings of the RSSC97*, March 1997, Raleigh, NC. New York: Elsevier Science Publishing Inc., 265–268.

Talluri, R., 2000. Data envelopment analysis: models and extensions. *Decision Line, Production/Operations Management*, 31, 8–11.

Talluri, S. and Narasimhan, R., 2004. A methodology for strategic sourcing. *European Journal of Operational Research*, 154, 236–250.

Tully, S., 1995. Purchasings' new muscle. *Fortune*, 20, 76.

Verma, R. and Pullman, M., 1998. An analysis of the supplier selection process. *Omega: The International Journal of Management Science*, 26 (6), 739–750.

Voges, K., Pope, N., and Brown, M., 2002b. Cluster analysis of marketing data examining on-line shopping orientation: a comparison of K-means, rough clustering approaches, *In*: H.A. Abbas, R.A. Sarker and C.S. Newton, eds. *Heuristics and optimisation for knowledge discovery*. Pennsylvania, USA: Idea Group Publishing, 217–225.

Voges, K., Pope, N., and Brown, M., 2002a. Cluster analysis of marketing data: a comparison of K-means, rough set, and rough genetic approaches, *In*: H.A. Abbas, R.A. Sarker and C.S. Newton, eds. *Heuristics and optimisation for knowledge discovery*. Pennsylvania, USA: Idea Group Publishing, 208–216.

Weber, C. and Ellram, L., 1993. Supplier selection using multi-objectivity programming: a decision support system approach. *International Journal of Physical Distribution and Logistics Management*, 23 (2), 3–14.

Wojcik, Z., 1993. Rough sets for intelligent image processing. *Proceedings of the international workshop on rough sets and knowledge discovery*, Banff, Alberta, Canada. London: Springer-Verlag, 399–410.

Xu, H., Besant, B., and Ristic, M., 2003. System for enhancing supply chain agility through exception handling. *International Journal of Production Research*, 41 (6), 1099–1114.

Zsidisin, G. and Ellram, L., 2003. An agency theory investigation of supply risk management. *Journal of Supply Chain Management*, 39 (3), 15–29.

Zsidisin, G., Ragatz, G., and Melnyk, S., 2003. *Effective practices for business continuity planning in purchasing and supply management. Department of Marketing and Supply Chain Management, The Eli Broad Graduate School of Management*. East Lansing, Michigan: Michigan State University.