J Intell Manuf
DOI 10.1007/s10845-012-0684-z

Using data mining synergies for evaluating criteria at pre-qualification stage of supplier selection

Rajeev Jain · A. R. Singh · H. C. Yadav · P. K. Mishra

Received: 5 January 2012 / Accepted: 31 July 2012

Abstract A company must purchase a lot of diverse components and raw material from different upstream suppliers to manufacture or assemble its products. Under this situation the supplier selection has become a critical issue for the purchasing department.The selection of suppliers depends on number of criteria and the challenge is to optimize selection process based on critical criteria and select the best supplier(s). During supplier selection process initial screening of potential suppliers from a large set is vital and the determination of prospective supplier is largely dependent on the criteria chosen of such pre-qualification. In the literature, many judgments based methods are proposed and derived criteria selection from the opinion of either the customers or the experts. All these techniques use the knowledge and experience of the decision makers. These methods inherit certain degree of uncertainty due to complex supply chain structure. The extraction of hidden knowledge is one of the most important tools to address such uncertainty and data mining is one such concept to account for such uncertainty and it has been found applicable in many scenarios. The proposed research aims to introduce a data mining approach, to discover the hidden relationships among the supplier's pre-qualification data with the overall supplier rating that have been derived after observation of previously executed work for a period of time. It provides an overview that how supplier's initial strength influence its final work performance.

R. Jain (✉)
Mechanical Engineering Department, Kalaniketan Polytechnic College, Jabalpur 482001, India
e-mail: rjain@mnnit.ac.in; jainrajeev@rediffmail.com

A. R. Singh · H. C. Yadav · P. K. Mishra
Mechanical Engineering Department, Motilal Nehru National Institute of Technology, Allahabad 211004, India

Keywords Supply chain management (SCM) · Supplier selection · Supplier's pre-qualification · Data mining · i-PM algorithm

Introduction

Supply Chain Management (SCM) has become a critical aspect in today's fiercely competitive business environment. Under the expanded heading of logistics, SCM is now an integral part of company activities covering areas such as purchasing management, transportation management, production management, warehousing management, inventory management, etc (Benyoucef and Tiwari 2010). Strategic collaboration with suppliers and service providers is the key to achieve competitive advantage associated with SCM philosophy. One important aspect in SCM is purchasing and selecting an appropriate supplier that ensures the success of purchasing function. Selecting the suitable supplier has been always a difficult task for the purchasing manager. Supplier selection is an important business decision since a competitive advantage may be gained by the firm with the right suppliers to provide products/services more effectively and efficiently. If the selection process is carried out correctly, a higher quality, longer lasting buyer–supplier relationship is more attainable (Kang et al. 2010). The supplier selection decisions are typically complicated for several reasons such as incomplete information, additional qualitative criteria and imprecise preferences. A range of methods and techniques that may assist the decision-makers in dealing with the increased complexity and importance of their decisions have been offered in the discipline of Operations Management. Examples of such techniques are: multi-criteria decision aid, problem structuring approaches, mathematical programming, and data mining techniques etc.

Published online: 12 August 2012

The problem of supplier selection has been a matter of concern to the researchers since the published work of Dickson (1966). The quantum of work is voluminous that may be obvious from the four articles that reviews the literature on this area only (See Weber et al. 1991; Degraeve et al. 2000; De Boer et al. 2001; Ho et al. 2010). The literature deals with supplier evaluation and selection models, methods and tools.

Supplier selection is defined as the "process of finding the suppliers being able to provide the buyer with the right quality products and/or services at the right price, at the right quantities and at the right time" (Sonmez 2006). The supplier selection process is generally described in the literature consists of five stages: (1) Identification of the need for a new supplier; (2) Identification and elaboration of selection criteria; (3) Initial screening of potential suppliers from a large set; (4) Final supplier selection; and (5) Continuous evaluation and assessment of selected suppliers (De Boer et al. 2001; De Boer and Van der Wegen 2003).

In the supplier selection process, researchers have mainly focused their attention on the choice stage of the process, however, the quality and success of this stage is largely dependent on the quality of a stage prior to choice stage. De Boer et al. (2001) in his work have not only considered the choice stage but have presented a comprehensive framework to cover all stages of supplier selection process.

In the era of globalization of markets buyers are bound to evaluate a large number of suppliers in order to make a comprehensive choice. Mathematical decision methods for supplier evaluation and selection are time consuming and computationally expensive. Owing to these problems it is prudent to prune the supplier search space and focus on the suppliers that are most likely suitable for the organization. The supplier search space can be shortened based on the pre-qualification.

The pre-qualification step can be defined as the process of reducing the set of 'all' suppliers to a smaller set of acceptable suppliers. It's a 'sorting' process rather than a 'ranking' process (De Boer et al. 2001). It consists of two steps. Initially, the set of acceptable suppliers is defined and determined and thereafter the number of suppliers is reduced. The determination of prospective supplier is largely dependent of the criteria chosen of such pre-qualification. There are two basic types of criteria for this process i.e. quantitative and qualitative. Quantitative criteria such as costs etc. are measurable in solid dimensions but qualitative criteria could not match the quality of the design. These criteria may have interdependencies and contradiction with each other that can in turn may increase the complexity in the decision making problem.

Researchers have proposed many methods and tools for selection of suitable criteria such as interpretative structured modeling (ISM), analytical hierarchy process (AHP), analytical network process (ANP) etc. Most of these methods are judgment based and derived criteria selection from the opinion of either the customers or the experts. These techniques use the knowledge and experience of the decision makers and inherit certain degree of uncertainty due to complex supply chain structure.

The extraction of hidden knowledge is one of the most important tools to address such uncertainty. Ketikidis et al. (2008) have reported that data mining is one such concept to account for such uncertainty and it has been found applicable in many logistics scenarios such as: decision support in production scheduling. Ha and Krishnan (2008) have proposed knowledge extraction technique in the selection of suppliers. The aim of the data mining approach is to mine associated rules that can optimize the desired objectives.

The proposed research aims to introduce a data mining approach, to discover the hidden relationships between the supplier's pre-qualification data with the overall supplier rating that have been derived after observation of previously executed works for a period of time. It provides an overview that how supplier's initial strength influence its final work performance. The approach proposes the critical criteria for selection of supplier at pre-qualification stage in a case study.

Literature review

Procurement is considered to be one of the critical activities in any industrial environment. With the increased level of outsourcing, much attention needs to be paid to the supplier selection and evaluation (Sharma and Ballan 2012). Supplier selection has received an extensive attention by the researchers for last few years as it is vital in success of a supply chain management (SCM) and it deals with evaluation, ranking and selection of the best option from a pool of potential suppliers especially in the presence of conflicting criteria (Shemshadi et al. 2011). Supplier selection decisions are an important component of production and logistics management for most manufacturing firms. Such decisions entail the selection of individual suppliers to employ and the determination of order quantities to be placed with the selected suppliers.

This section presents the insight on the past and recent works in the field of supplier selection decisions and knowledge based approaches. It is categorized into following sub sections:

Supplier selection decisions

From the aspects of supply side, there are a number of key characteristics that should look for when identifying and short listing possible suppliers. Good supplier should be able to demonstrate that they can offer the benefits, such as reliability, quality, value for money, strong service and clear communication, financial security, a partnership approach, etc. (Ozaki et al. 2011). In the literature, supplier selection is characterised as a multiple criteria decision making (MCDM) problem and it is necessary to make a trade-off between conflicting tangible and intangible factors to find the most appropriate supplier. The degree of uncertainty, the number of decision-makers and the nature of the criteria have to be carefully considered to solve this problem (Sevkli 2010). A wide range of methods and tools have been undertaken to provide a solution to the problem.

The supplier selection approaches has been classified in to individual and integrated approaches. Ho et al. (2010) observed that the most popular individual approach had been found to be data envelope analysis (DEA), followed by mathematical programming, AHP, case based reasoning(CBR), ANP, fuzzy set theory, simple multi-attribute rating techniques (SMART), and genetic algorithm (GA). There are various integrated approaches for supplier selection. It was noticed that the integrated AHP approaches are more prevalent. The wide applicability is due to its simplicity, ease of use, and great flexibility. AHP has been integrated with other techniques, including artificial neural network (ANN), bi-negotiation, DEA, fuzzy set theory, goal programming (GP), grey relational analysis, and multi-objective programming. Comparatively, the integrated AHP–GP approach is the most popular. Regarding criteria selection, it was observed that hundreds of criteria were proposed in the articles. The most popular criterion is quality, followed by delivery, price/cost, manufacturing capability, service, management, technology, research and development, finance, flexibility, reputation, relationship, risk, and safety and environment.

There has been considerable amount of work published after 2008. Authors noticed that fuzzy based integrated methods are more prevalent now and fuzzy theory is integrated with other techniques, including AHP, ANP, linear programming (LP) with strengths, weakness, opportunities, threats (SWOT), quality function deployment (QFD), decision making trial and evaluation laboratory (DEMATEL), dempster, Technique For Order preference BY Similarity to Ideal Situation(TOPSIS), DEA, GP, neural network (NN), vlse kriterijumska optimizacija kompromisno resenje (VIKOR), elimination and choice expressing reality (ELECTRE), gray relational analysis (GRA), principal component analysis(PCA), adaptive resonance theory(ART) and mathematical programming. Again, many criteria have been proposed, but there is an increase on emphasis of the criteria like flexibility, environment, risk and corporate social responsibility in this literature.

Pre-qualification is not focused in the literature as a separate stage of supplier selection, there are very few articles found under this heading. De Boer et al. (2001) emphasized that the pre-qualification of supplier is a sorting process than a ranking process, important difference between sorting and ranking is often not explicitly made in the literature. Therefore, the articles discussed by De Boer et al. (2001) under the heading of pre-qualification have originally appeared as supplier selection articles. Many different techniques for pre-qualification have been cited in the review. Some of these techniques are Categorical Methods, Data Envelopment Analysis (DEA), Cluster Analysis and Case-based Reasoning (CBR) systems. Weber and Ellram (1992), Weber and Desai (1996) have developed DEA methods for pre-qualification. Hinkle et al. (1969) and Holt (1998) used cluster analysis for pre-qualification and finally Ng and Skitmore (1995) developed CBR systems for pre-qualification.

Data mining approaches

The proliferation of large masses of data has created many new opportunities for those working in science, engineering and business. The opportunities are offered by the abundance and availability of data and the challenges are posed by the problem of how to organise, retrieve and extract knowledge from this data (Kusiak 2000).

Data mining is a tool which is designed to extract hidden patterns from the given data. It has found to be useful in many fields including medical, defence and crime detection. Data mining is highly capable to tackle tasks such as classification, clustering, association rule discovery, sequence pattern discovery and regression. Data mining is also been made effective in handling large data base as it reduces the search space (Kuri-Morales and Rodriguez-Erazo 2009). Data mining has found a great importance in managing supply chain. Data mining found to been efficient in finding the association rules in the evaluation of supply chain network (Jain et al. 2008). Several works has been contributed in this field. The association clustering method in the joint replenishment policy by Tsai et al. (2009) is an efficient way of cost reduction. As a part of supply chain, distributor selection is explicitly explained by Zhonghai et al. (2011) by introducing the method of rough set theory in it. Most of the data mining approaches embed fuzzy set theory to determine the association rules in the process. Uncertainty and imprecision in the behaviours of data (Lee et al. 2008; Chen and Chang 2008) has been explicitly taken care by the fuzzy set theory. Liao et al. (2008) inquired about the association rules for product line at the customer level. There has been several works in the development of association rules (Wang 2008). Lin et al. (2009) suggested that the supplier selection can be viewed as the problem of mining a large database of shipment. The proposed method incorporates the extended association rule algorithm of data mining with that of set theory to find key suppliers. The study of Chen et al. (2011) proposes an integrated model by combining K-means clustering, feature selection and the decision tree method into a single evaluation model to assess the performance of suppliers and simultaneously tackles the abovementioned shortcomings.

Proposed methodology

In the pre-qualification stage of any supplier selection process, the issue of selection of criteria is crucial. With the proper application of pre-selected criteria the pre-qualification method is rationalized. The purchaser has to decide various criteria and sub-criteria on the basis of nature of product and objective of organization. Once these criteria and sub-criteria are decided, proper weights has been assigned to each of them. Thereafter, the purchasers need to collect data from the supplier firms and assess on quantitative and quantitative criteria. Based on the final score the suppliers are then sorted. Now a day most of the companies used a supplier registration format for collection of data from the supplier and thereafter assessed on the pre-defined criteria scores for supplier qualification. The preliminary selection of supplier is based on certain minimum score achieved. There of two implication of adoption of these methods. First, a pool of supplier has same score but may have different strengths and weaknesses on account of different criteria and second, interdependencies on different criteria can not be assessed at this stage.

By careful consideration of available literature authors decided to choose six main categories of pre-qualification criteria, these are: organizational strength, past experience, performance capabilities, financial soundness, cost and miscellaneous. Table 1 suggest the various criteria and sub-criteria of supplier pre-qualification.

For each qualitative criteria the suppliers are rated by assessors and scores are given on the basis of a pre-defined formula where as the value of quantitative criteria is taken directly. For different criteria the ranges are shown in Table 2

Table 1 Pre-qualification criteria

Main criteria	Sub criteria
Organizational strength	Company size
	Company age
	R & D activities
Past experience	Type of past project completed
	Size of past project completed
Performance capabilities	Contract overruns
	Overall quality
	Responsiveness
Financial soundness	Turnover
	Bank references
Cost	Order change and cancellation cost
	Warranties and claims
Miscellaneous	Labor relationship
	Procedural compliances

Table 2 The symbols of selection criteria parameter and supplier rating

S.N.	Parameter	Range	Symbol
1.	Organizational strength	0–10	A
2.	Past experience	0–20	B
3.	Performance capabilities	0–10	C
4.	Financial soundness	0–100	D
5.	Cost	0–500	E
6.	Miscellaneous	0–10	F
7.	Supplier rating index	0–50	G

First of all a final rating is calculated on the basis of the work performance for a specified period of time. The rating is derived by the assessments of the certain performance criteria. Here quality, cost, delivery and service. The final rating is composed on individual rating considering these criteria. Suggested formula of different parameters is as under:

Quality rating

$$SRQ = \sum_{i=1}^{4} w_i . Q_i \tag{1}$$

where,
Q_1 = Quantity accepted conforming to specification
Q_2 = Quantity accepted with deviation
Q_3 = Quantity accepted after rectification
Q_4 = Quantity rejected
w_i = Weights of $Q_i \forall i \in (1, 4)$
Q = Total quantity offered for inspection$(Q_1 + Q_2 + Q_3 + Q_4)$.

Delivery rating

$$SRD = w_1.(Q_a/Q_c) + w_2.(Q_b/Q_c).(T_c/T_a) \tag{2}$$

where,
Q_a = Quantity supplied on time
Q_c = Quantity ordered
Q_b = Quantity supplied late beyond delivery period
T_c = Delivery period as per supply order in days
T_a = Total time taken to complete the supplies, including late deliveries in number of days
$w_{1,2}$ = Weights of on-time and late deliveries respectively

Price rating

$$SRP = P_L/P_Q \tag{3}$$

where,
P_L = Lowest price quoted by any supplier for a product
P_Q = Price quoted by the supplier

Service rating

Service rating (SRS) is determined as per IS 12040: 2001. The service rating breakup has been presented in Table 3

Composite index for performance or overall rating

$$SRI = W_a.SRQ + W_b.SRD + W_c.SRP + W_d.SRS \quad (4)$$

where,
$W_{a,b,c,d}$ = Weights of parameter of quality, price and service respectively

For a particular item, the average supplier rating, ASR for a supplier over a period of time can be evaluated by applying the formula:

$$ASR = (SRI1 + SRI2 + \cdots + SRIn)/n \quad (5)$$

where

n = the number of orders received by the supplier during the given period for that particular item

The records from supplier database are shown in Table 4 which includes the value of the parameter of different pre-qualification criteria and the overall rating of the suppliers. Table 5 shows the minimum support values of each parameter.

Table 3 Service rating breakup as per IS 12040:2001

Parameter	Maximum score
Cooperativeness and readiness to help in emergencies	30
Readiness to replace rejected material	20
Providing support documents in time	10
Promptness in reply	10
Co-operation in delivering and implementing measures or avoiding recurrence of defects/complaints	30
Total	100

Table 4 Records from the supplier database

Record no.	A	B	C	D	E	F
1	2.5	8	3.5	85	300	7.5
2	4	10	4	95	300	7.5
3	6	15	4.5	65	420	2.5
4	3	11	6.5	70	500	4.5
5	4	11	6	80	350	9
6	6	20	6	90	400	9

Table 5 The minimum support value for each parameter

Parameter	A	B	C	D	E	F	G
Support value	1.8	2.0	1.5	1.9	2.1	2.4	2.2

Objectives

- Identify the critical pre-qualification criteria
- Find the influence of the pre-qualification criteria on final supplier rating

Constraints

- Pre-qualification criteria does not effect each other
- The selected parameters should have a value greater than threshold

Solution approach

To find the solution regarding the identification of critical pre-qualification criteria and their influence on the supplier's final rating, application of i-PM algorithm is proposed.

Notations

Q_{ij}	Observational data item i for jth observation. $i \in (1, n)$, $j \in (1, 10)$
F_{ij}^k	Fuzzy value for Q_{ij} with linguistic value k. $k \in (1, m)$
R_{ij}^k	Linguistic notation (Component. Fuzzy Linguistic notation) of Q_{ij}
R_i^k	Linguistic term after combining all observations.
$Count_i^k$	Fuzzy count for R_i^k
α_i	Minimum support value for Observational data item i
$Max - Count_i^{k'}$	Maximum count for data item i with the fuzzy linguistic value k'
CV_p	Confidence value of rule p.

Application of i-PM algorithm

This process mining technique tries to identify the hidden information among supplier's pre-qualification data and its relationship with final performance rating. The objective of i-PM algorithm is to identify the root cause of failure for supply chain network optimization. This algorithm integrates the fuzzy mining algorithm proposed by Hong et al. (2003) for the identification of fuzzy association rules from quantitative process data. The developed approach has been successfully applied to discover association rules for quality enhancement

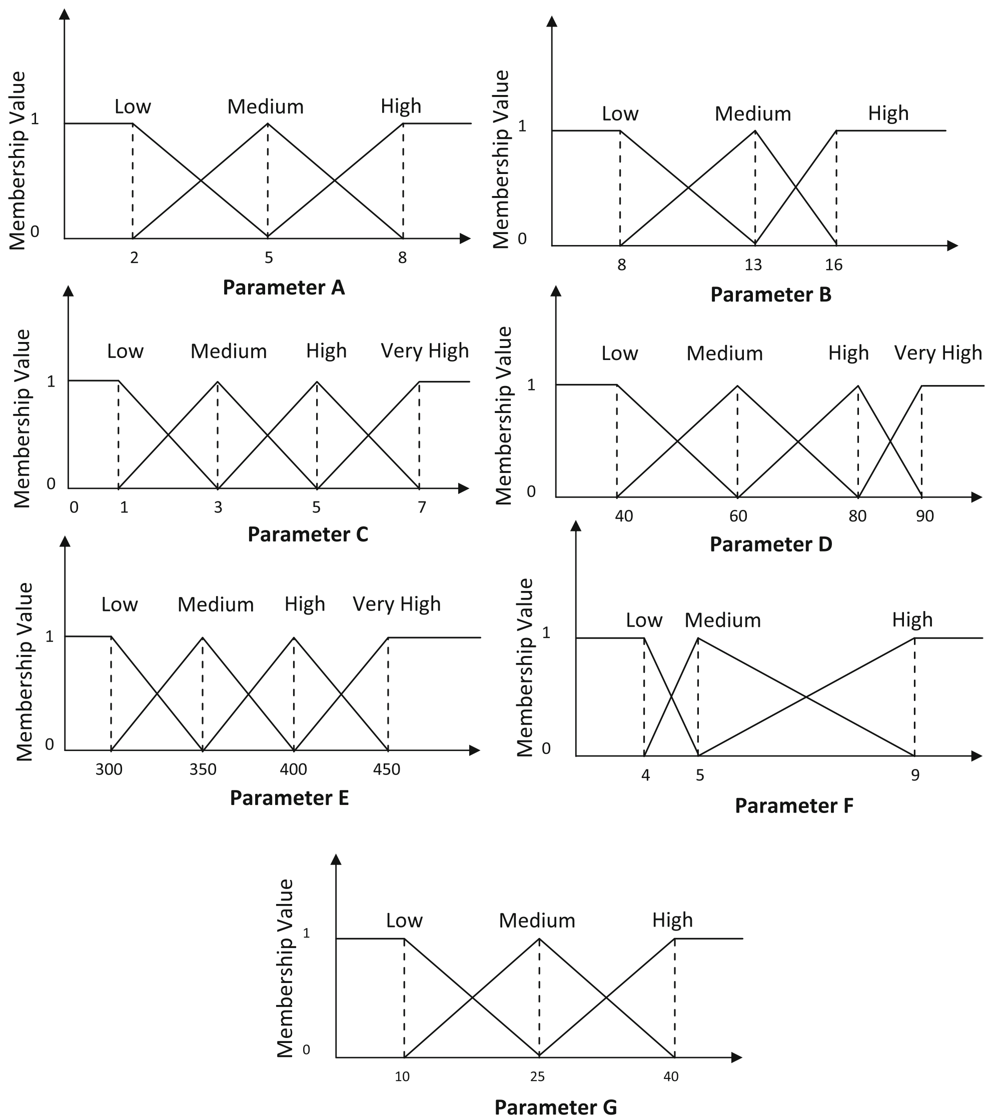

Fig. 1 The membership functions of selection criteria parameters and supplier rating in a case example

in many manufacturing industries (Lau et al. 2009a). This algorithm has been found efficient in identifying the best association rules in works of Lau et al. (2009b). Authors have used the minimum support value to enhance the quality level of the solution. In real applications, these values should be made flexible for making the algorithm adaptive.

The i-PM algorithm

The steps of i-PM algorithm are given below:

Step 1: Transform the quantitative values of the parameters and overall rating in each record into a fuzzy set using the membership functions given in Fig. 1. Take the parameter A in the fourth record as an example. As shown in Fig. 2, the crisp value "3" of process parameter "A" is converted into the fuzzy set which is calculated as (0.33 of medium + 0.67 of low). This step is repeated for all items in the records and the result is given in Table 6. The converted structure of parameters with fuzzy regions is represented as Convert the Q_{ij} value to F_{ij}^{k} from the respective fuzzy membership function. Put the value of F_{ij}^{k}.

Step 2: Calculate the fuzzy count for each fuzzy region. The fuzzy count for ith data item with kth linguistic value is given by Eq. (6). Insert these values in the list C1 (Table 7).

$$Count_i^k = \sum_{j=1}^{10} F_{ij}^{k} \quad \forall i \in (1, n), \forall k \in (1, m) \qquad (6)$$

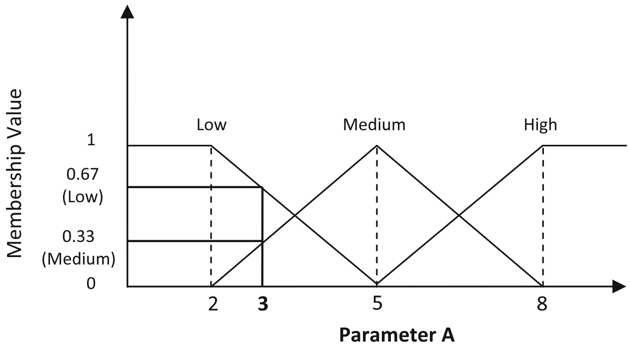

Fig. 2 Fuzzy set conversion of parameter "A"

Step 3: Find the maximum of fuzzy count for a data item i. k' represent the linguistic index which gives the maximum count value. Equation (7) represents the calculation of maximum count value along with the index.

$$Max{-}Count_i^{k'} = \max(Count_i^1, \ldots, Count_i^m) \quad (7)$$

For example if for data item A the L, M, H value are 2.16, 3.18, 0.66 respectively then $Max - Count_1^{k'}$ value will be 3.18 for k = k$'$ = 2. Put the values obtained in the list (Table 8).

Step 4: Compare the $Max - Count_i^{k'}$ values obtained in the step 3 to that of the minimum support value α_i for data item i. If the value of $Max - Count_i^{k'}$ is greater than α_i then put than in the list L1 (Table 9). Suppose that for data item A, a Medium value of 3.18 has been chosen. If its minimum support value is 1.8 (which is lower than the count value) then {A. Medium} with count 3.18 will be put in the list L1.

Step 5: Now combine the data items in the list L1 to get to sets of two data items. All possible combinations from the list L1 should be found out. Like A. Medium and B. Low can be combined as a set and will be represented as {A. Medium, B. Low}. After that their count and the combined minimum support value has to be calculated. Suppose that

Table 6 The fuzzy set transformed from the records

Parameter	Notation	Quantitive values of variables using fuzzy set record number					
		1	2	3	4	5	6
A	Low	0.83	0.33	0	0.67	0.33	0
	Medium	0.17	0.67	0.67	0.33	0.67	0.67
	High	0	0	0.33	0	0	0.33
B	Low	1	0.6	0	0.4	0.4	0
	Medium	0	0.4	0.33	0.6	0.6	0
	High	0	0	0.67	0	0	1
C	Low	0	0	0	0	0	0
	Medium	0.75	0.5	0.25	0	0	0
	High	0.25	0.5	0.75	0.25	0.5	0.5
	Very high	0	0	0	0.75	0.5	0.5
D	Low	0	0	0	0	0	0
	Medium	0	0	0.75	0.5	0	0
	High	0.5	0	0.25	0.5	1	0
	Very high	0.5	1	0	0	0	1
E	Low	0	1	0	0	0	0
	Medium	1	0	0	0	1	0
	High	0	0	0.6	0	0	1
	Very high	0	0	0.4	1	0	0
F	Low	0	0	1	0	0	0
	Medium	0.38	0.38	0	0.5	0	0
	High	0.62	0.62	0	0.5	1	1
G	Low	0	0	0	0	0	0
	Medium	0	0.13	0.67	0	0.19	0
	High	1	0.87	0.33	1	0.81	1

Table 7 The fuzzy counts of itemsets in C1

Parameter items	Count
A. Low	2.16
A. Medium	3.18
A. High	0.66
B. Low	2.40
B. Medium	1.93
B. High	1.67
C. Low	0.00
C. Medium	1.50
C. High	2.75
C. Very high	1.75
D. Low	0.00
D. Medium	1.25
D. High	2.25
D. Very high	2.50
E. Low	1.00
E. Medium	2.00
E. High	1.60
E. Very high	1.40
F. Low	1.00
F. Medium	1.26
F. High	3.74
G. Low	0.00
G. Medium	0.99
G. High	5.01

count value to A. Medium is 3.18 and its minimum support value 1.8. Then for the combined set {A. Medium, B. Low}, the minimum support value is calculated as $\max(\alpha_1, \alpha_2)$ and to calculate the count of that set list in Table 6 must be checked. From the Table 6 the fuzzy values corresponding to A. Medium and B. Low can be extracted. The calculation of the generalised combined count is shown in Eq. (8). Zero Values are not considered in the calculation of minimum.

$$Combined\_Count_{i1,i2}^{k1,k2} = \sum_{j=1}^{10} \min(F_{i1j}^{k1}, F_{i2j}^{k2}) \quad (8)$$

Similarly value of $Combined\_Count_{1,2}^{3,2}$ can be calculated. Let its value comes out to be 1.5. Insert all the values in the list C2 (Table 10).

Step 6: The values obtained in C2 have to be compared with its combined minimum support value (Table 11) and the one qualify that test will be put in the list L2 (Table 12).

Table 8 The maximum fuzzy counts and the corresponding fuzzy regions of the itemsets in C1

Parameter items	Count
A. Medium	3.18
B. Low	2.40
C. High	2.75
D. Very high	2.50
E. Medium	2.00
F. High	3.74
G. High	5.01

Table 9 Items in large 1-itemsets L1

Parameter items	Count
A. Medium	3.18
B. Low	2.40
C. High	2.75
D. Very high	2.50
F. High	3.74
G. High	5.01

Table 10 Items in C2 and their fuzzy counts

Parameter Items	Count
{A. Medium, B. Low}	1.50
{A. Medium, C. High}	2.59
{A. Medium, D. Very high}	1.51
{A. Medium, F. High}	2.46
{A. Medium, G. High}	2.84
{B. Low, C. High}	1.40
{B. Low, D. Very high}	1.10
{B. Low, F. High}	2.02
{B. Low, G. High}	2.40
{C. High, D. Very high}	1.25
{C. High, F. High}	2.00
{C. High, G. High}	2.33
{D. Very high, F. High}	2.12
{D. Very high, G. High}	2.37
{F. High, G. High}	3.55

Step 7: Now again we will group the data items in the groups of 3 and the count (Table 13) and minimum support values (Table 14) has to be find out.

Step 8: The process will keep on repeating from step 5 to Step 6 till the list becomes null.

Step 9: The final list obtained will have the entries like {A. Medium, C. High, G. High} (Table 15). Now the rules have to be developed using this set. The fuzzy rules can be of any form made of combinations of different elements from the set like If

Table 11 The supports of the 2-itemsets

Parameter items	Count
{A. Medium, B. Low}	2.0
{A. Medium, C. High}	1.8
{A. Medium, D. Very high}	1.9
{A. Medium, F. High}	2.4
{A. Medium, G. High}	2.2
{B. Low, C. High}	2.0
{B. Low, D. Very high}	2.0
{B. Low, F. High}	2.4
{B. Low, G. High}	2.2
{C. High, D. Very high}	1.9
{C. High, F. High}	2.4
{C. High, G. High}	2.2
{D. Very high, F. High}	2.4
{D. Very high, G. High}	2.2
{F. High, G. High}	2.4

Table 12 Items in large 2-itemsets L2

Parameter items	Count
{A. Medium, C. High}	2.59
{A. Medium, F. High}	2.46
{A. Medium, G. High}	2.84
{B. Low, G. High}	2.40
{C. High, G. High}	2.33
{D. Very high, G. High}	2.37
{F. High, G. High}	3.55

Table 13 Items in C3 and their fuzzy counts

Parameter items	Count
{A. Medium, C. High, G. High }	2.25
{A. Medium, G. High, F. High }	2.46

Table 14 The support of the 3-itemsets

Parameter items	Count
{A. Medium, C. High, G. High }	2.20
{A. Medium, G. High, F. High }	2.40

{A. Medium, C. High} then {G. High}. In order to optimise the rules confidence values have been used as proposed by Lau et al. (2009a), depicted in Eq. (9).

$$CV_p = \frac{Count(\{A.Medium, C.High, G.High\})}{Count(\{A.Medium, C.High\})} \tag{9}$$

Table 15 Items in large 3-itemsets L3

Parameter items	Count
{A. Medium, C. High, G. High }	2.25
{A. Medium, F. High, G. High }	2.46

Table 16 All possible association rules and their corresponding confidence values

Association rule	Confidence value
{A. Medium, C. High} then {G. High}	0.87
{A. Medium, G. High} then {C. High}	0.79
{C. High, G. High } then {A. Medium}	0.97
{A. Medium} then {C. High, G. High}	0.71
{C. High} then {A. Medium, G. High}	0.82
{G. High} then {A. Medium, C. High}	0.45
{A. Medium, F. High} then {G. High}	1.00
{A. Medium, G. High} then {F. High}	0.87
{F. High, G. High } then {A. Medium}	0.69
{A. Medium} then {F. High, G. High}	0.77
{F. High} then {A. Medium, G. High}	0.66
{G. High} then {A. Medium, F. High}	0.49
{A. Medium} then {C. High}	0.81
{C. High} then {A. Medium}	0.94
{A. Medium} then {F. High}	0.77
{F. High} then {A. Medium}	0.66
{A. Medium} then {G. High}	0.89
{G. High} then {A. Medium}	0.57
{B. Low} then {G. High}	1.00
{G. High} then {B. Low}	0.48
{C. High} then {G. High}	0.85
{G. High} then {C. High}	0.47
{D. Very high} then {G. High}	0.95
{G. High} then {D. Very high}	0.47
{F. High} then {G. High}	0.95
{G. High} then {F. High}	0.71

Here CV_p represents the confidence value of rule p. All the rules along with the confidence value are listed in the Table 16.

Step 10: In order to further optimise the rules, after putting the constraints a limit on the confidence value must be imposed. In order to further optimise the rules we some put the constraints. Thereafter a limit on the confidence value must be imposed. A value of $CV \geq 0.80$ is accepted. This will curtail the number of rules to less than 5 or 6. Depending on the criteria and highest CV the rules must be selected. Table 17 gives the rules with confidence values equal to or larger than the predefined confidence threshold.

Step 11: Keep the rules with pre-qualification criteria parameters in the consequent part only and put constraints

Table 17 The association rules with confidence values ≥0.80

Association rule	Confidence value
{A. Medium, C. High} then {G. High}	0.87
{C. High, G. High } then {A. Medium}	0.97
{C. High} then {A.Medium, G. High}	0.82
{A. Medium, F. High} then {G. High}	1.00
{A. Medium, G. High} then {F. High}	0.87
{A. Medium} then {C. High}	0.81
{C. High} then {A. Medium}	0.94
{A. Medium} then {G. High}	0.89
{B. Low} then {G. High}	1.00
{C. High} then {G. High}	0.85
{D. Very high} then {G. High}	0.95
{F. High} then {G. High}	0.95

Table 18 The association rules which satisfy the relevant constraints

Association rule	Confidence value
{A. Medium, C. High} then {G. High}	0.87
{A. Medium, F. High} then {G. High}	1.00
{A. Medium} then {G. High}	0.89
{C. High} then {G. High}	0.85
{F. High} then {G. High}	0.95

if any. Now output the rules to users as interesting association rules as presented in Table 18.

For the given example, the rules indicate that three pre-qualification criteria (Organizational strength, Performance capabilities and Miscellaneous) are the important factors which affect the final supplier rating. Association rules suggest that organizational strength and performance capabilities of the firms have impacted supplier's final work performance. For this particular case the criteria like work experience and cost does not have any influence on final supplier rating. This methodology can be applied to supplier selection process of a supply chain network to extract some actionable knowledge. By introducing the methodology, the analysis engineers in the supply chain network can anticipate work performance at pre-qualification stage itself. Moreover, the measurement of the improvements of supplier work performance reflects the feasibility and utility of the application of the i-PM method to the supply chain network.

Discussion

In this research, the i-PM algorithm is used and the focus is on discovering the correlations among supplier's pre-qualification data. This is done so as to support knowledge discovery. This algorithm is designed to be applied in the cases-data stored in a data warehouse. If the enterprise does not collect data regularly or if the number of records of stored data is not enough, this algorithm is inappropriate and may reduce the reliability and quality of mining rule sets. This may limit the application of proposed methodology in practice (Lau et al. 2009a).

The fuzzy sets derived from knowledge acquisition used in the proposed methodology can have a variety of shapes including triangle, trapezoid, parabola, etc. In order to provide adequate representation of expert knowledge and simplify the process of computation, a triangular shape is used to represent different fuzzy sets for the processes within the supply chain. The framework of i-PM algorithm can be modified in the future by allowing different shapes of fuzzy sets to be computed. Further more, the determination of membership functions and fuzzy sets by relying on the knowledge of experts is quite time-consuming and subjective (Lau et al. 2009a). It is suggested that some AI techniques with self-learning ability be incorporated to derive the fuzzy sets automatically.

Conclusion

A methodology for supporting knowledge discovery in a supply chain network is proposed and developed. The methodology has been evaluated in a case study and the algorithm shows its potential to figure out the primary criteria that have an influence on the supplier's final work performance in a supply chain. In present case, three pre-qualification criteria (Organizational strength, Performance capabilities and Miscellaneous) have been identified as the important factors that have impacted the overall rating of the supplier after final work performance. It is noticed that some association rules are relatively difficult to understand. There is a need to integrate such association rules with practical situation for in-depth analysis or resort to other techniques to examine or even exclude it. This is a universal challenge for the techniques of mining association rules to provide effective rules and avoid dubious rules.

Furthermore, the advantage of the i-PM algorithm is that it can be used and applied in an enterprise where its centralized data warehouse is huge. Adequate sample data can ensure the quality of association rules. However, the disadvantage is that it may take a large amount of computational time to identify the significant association rules. The membership functions of different parameters along the logistics workflow are always subjective (Lau et al. 2009a). Selection of suitable fuzzy sets, like most "fuzzy" quantifier (Wang 2008) is one suggestion for the possible future work. In commercial applications, the developed methodology could be based on the integration of the i-PM algorithm with the right fuzzy sets. Therefore, another future research is necessary to validate the performance of the system in a variety of applications through extensive practical evaluations.

References

Benyoucef, L., & Tiwari, M. K. (2010). Editorial note for the special issue on 'Advanced metaheuristics for integrated supply chain management'. *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-010-0456-6.

Chen, S. H., & Chang, S. M. (2008). Optimization of fuzzy production inventory model with unrepairable defective products. *International Journal of Production Economics, 113*, 887–894.

Chen, Y.-S., Cheng, C.-H., & Lai, C.-J. (2011). Extracting performance rules of suppliers in the manufacturing industry: An empirical study. *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-011-0530-8.

De Boer, L., & Van der Wegen, L. L. M. (2003). Practice and promise of formal supplier selection: A study of four empirical cases. *Journal of Purchasing & Supply Management, 9*, 109–118.

De Boer, L., Labro, E., & Morlacchi, P. (2001). A review of methods supporting supplier selection. *European Journal of Purchasing & Supply Management, 7*, 75–89.

Degraeve, Z., Labro, E., & Roodhooft, F. (2000). An evaluation of supplier selection methods from a total cost of ownership perspective. *European Journal of Operational Research, 125*(1), 34–58.

Dickson, G. W. (1966). An analysis of vendor selection systems and decisions. *Journal of Purchasing, 2*(1), 5–17.

Ha, S. H., & Krishnan, R. (2008). A hybrid approach to supplier selection for the maintenance of a competitive supply chain. *Expert System with Application, 34*, 1303–1311.

Hinkle, C. L., Robinson, J. P., & Green, P. E. (1969). Vendor evaluation using cluster analysis. *Journal of Purchasing, 5*(3), 49–58.

Ho, W., Xu, X., & Dey Prasanta, K. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of Operational Research, 202*, 16–24.

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

Hong, T.-P., Lin, K.-Y., & Wang, S.-L. (2003). Fuzzy data mining for interesting generalized association rules. *Journal of Fuzzy Sets and Systems, 138*, 255–269.

Jain, V. B., Benyoucef, L., & Deshmukh, S. G. (2008). A new approach for evaluating agility in supply chains using fuzzy association rules mining. *Engineering Applications of Artificial Intelligence, 21*, 367–385.

Kang, H.-Y., Lee, A. H. I., & Yang, C.-Y. (2010). A fuzzy ANP model for supplier selection as applied to IC packaging. *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-010-0448-6.

Ketikidis, P. H., Koh, S. C. L., Dimitriadis, N., Gunasekaran, A., & Kehajova, M. (2008). The use of information systems for logistics and supply chain management in South East Europe: Current status and future direction. *Omega, 36*, 592–599.

Kuri-Morales, A. F., & Rodriguez-Erazo, F. (2009). A search space reduction methodology for data mining in large databases. *Engineering Applications of Artificial Intelligence, 22*(1), 57–65.

Kusiak, A. (2000). *Computational intelligence in design and manufacturing. Industrial Engineering/Mechanical Engineering*. Newyork: Wiley, Intersciences.

Lau, H. C. W., Ho, G. T. S, Zhao, Y., & Chung, N. S. H. (2009a). Development of a process mining system for supporting knowledge discovery in a supply chain network. *International Journal of Production Economics, 122*, 176–187.

Lau, H. C. W., Ho, G. T. S., Chu, K. F., Ho, W., & Lee, C. K. M. (2009b). Development of an intelligent quality management system using fuzzy association rules. *Expert Systems with Application, 36*, 1801–1815.

Lee, Y.-C., Hong, T.-P., & Wang, T.-C. (2008). Multi-level fuzzy mining with multiple minimum supports. *Expert Systems with Applications, 34*, 458–468.

Liao, S.-H., Chen, C.-M., & Wu, C.-H. (2008). Mining customer knowledge for product line and brand extension in retailing. *Expert Systems with Applications, 34*, 1763–1776.

Lin, R.-H., Chuang, C.-L., Liou James, J. H., & Wua, G.-D. (2009). An integrated method for finding key suppliers in SCM. *Expert Systems with Applications, 36*, 6461–6465.

Ng, S. T., & Skitmore, R. M. (1995). CP-DSS: Decision support system for contractor prequalification. *Civil Engineering Systems: Decision Making Problem Solving, 12*(2), 133–160.

Ozaki, T., Lo, M.-C., Kinoshita, E., & Tzeng, G.-H. (2011). Decision-making for the best selection of suppliers by using minor ANP. *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-011-0563-z.

Sevkli, M. (2010). An application of the fuzzy ELECTRE method for supplier selection. *International Journal of Production Research, 48*(12), 3393–3405.

Sharma, S., & Ballan, S. (2012). An integrative supplier selection model using Taguchi loss function, TOPSIS and multicriteria goal programming. *Journal of Intelligent Manufacturing*. doi:10.1007/s10845-012-0640-y.

Shemshadi, A., Toreihi, M., Shirazi, H., & Tarokh, M. J. (2011). Supplier selection based on supplier risk: An ANP and fuzzy TOPSIS approach. *The Journal of Mathematics and Computer Science, 2*(1), 111–121.

Sonmez, M. (2006). A review and critique of supplier selection process and practices. Loughborough University Business School occasional papers series, No. 2006:1.

Tsai, C. Y., Tsai, C. Y., & Huang, P. W. (2009). An association clustering algorithm for can-order policies in the joint replenishment problem. *International Journal of Production Economics, 117*, 30–41.

Wang, F. H. (2008). On discovery of soft associations with "most" fuzzy quantifier for item promotion applications. *Information Sciences, 178*, 1848–1876.

Weber, C. A., Current, J. R., & Benton, W. C. (1991). Vendor selection criteria and methods. *European Journal of Operational Research, 50*(1), 2–18.

Weber, C. A., & Desai, A. (1996). Determination of paths to vendor market efficiency using parallel co-ordinates representation: A negotiation tool for buyers. *European Journal of Operational Research, 90*, 142–155.

Weber, C. A., & Ellram, L. M. (1992). Supplier selection using multi-objective programming: A decision support system approach. *International Journal of Physical Distribution & Logistics Management, 23*(2), 3–14.

Zou, Z., Tseng, T.-L. B., Sohn, H., Song, G., & Gutierrez, R. (2011). A rough set based approach to distributor selection in supply chain management. *Expert Systems with Applications, 38*, 106–115.