

PARTNER SELECTION IN VIRTUAL ENTERPRISES AN EXPLORATORY APPROACH

José António Crispim,
Universidade do Minho
crispim@eeg.uminho.pt

Jorge Pinho de Sousa
INESC Porto, Faculdade de Engenharia da Universidade do Porto
jsousa@inescporto.pt

Partner selection in virtual enterprises (VE) can be viewed as a multi-criteria decision making problem that involves assessing trade-offs between conflicting tangible and intangible criteria. In general, this is a very complex problem due to the dynamic topology of the network, the large number of alternatives and the different types of criteria. In this paper we propose an iterative and interactive exploratory process to help the decision maker identify the companies that best suit the needs of each particular project. This is achieved by using cluster analysis to distinguish companies according to some selected features. We present an example to illustrate this approach.

1. INTRODUCTION

A virtual enterprise (VE) can be defined as a temporary alliance of independent and geographically dispersed companies set up to share skills or core competencies and resources, in order to respond to business opportunities, the cooperation among the enterprises being supported by computer networks (Camarinha-Matos and Afsarmanesh, 2003). The partner selection problem arises in many other research contexts such as supply chain design, agile manufacturing, network design, dynamic alliances, and innovation management production. In a virtual enterprise (VE) this question is difficult to handle because of the short life-cycles of these organizations (temporary alliances) and because of the lack of formal mechanisms (contracts) to assure participants responsibility. According to Mowshovitz (1994), the functioning of virtual enterprises follows the *switching principle* since connections among members are switched on and off whenever needed. Reactivity and flexibility are the major benefits but, at the same time, the main problems of VEs (Gunasekaran et al. 2008).

The creation of a VE is usually triggered by a market opportunity, giving rise to a “project” that is generally decomposable in relatively independent sub-projects or activities. The work needed to “fulfil” a project involves a set of collaborative activities and the cooperation relationships established can be represented by an activity network. Based on previous experiences the network members can rapidly set up a VE if some organizational structure already exists. This leads to the concept of Virtual Breeding Environment (VBE) which is a long-term networked structure with common operating

principles and infrastructures, common ontologies, and mutual trust (Afsarmanesh & Camarinha-Matos, 2005).

The VE configuration process is a difficult problem due to the complex interactions between the different entities and to the highly dynamic structure of the network (resulting from the connections activation/disactivation), and because the expression of preferences may be based on incomplete or non-available information. To deal with this problem under a multi-criteria perspective, we allow several types of information (numerical, interval, qualitative and binary) in order to facilitate the expression of the stakeholders' preferences or assessments about the potential partners. This is an important requirement in practice as the multiplicity of factors considered when selecting partners for a business opportunity, such as cost, quality, trust and delivery time, cannot be expressed in the same measure or scale (Crispim and Sousa, 2007).

In the selection of partners it is very common to use ranking approaches (see e.g. Gülçin et al. 2007), but according to Munda (2005) rankings are not always trustable, because the results obtained depend for example on the quality of available information, on the set of criteria/indicators used to represent reality, on the direction of each objective/indicator, on the relative importance of these indicators and on the ranking methods adopted. The choice of the method is in fact very important to guarantee consistency between the assumptions considered by the decision maker and the obtained ranking. The quality of the decisions depends crucially on the way the methodology takes into account the various dimensions (social, political, economical, technical, etc.) considered when structuring process. This is the reason why Roy (1996) claimed that what is really important is the decision process and not the final solution.

The companies in a network may be very different from each other, and each company is characterized by a set of attributes that can be large in number. Moreover, these companies may be organised in quite different networks, depending on the particular considered criterion (see figure 1). Collecting and handling the associated data may therefore be a complicated task and require a considerable effort just in structuring the problem.

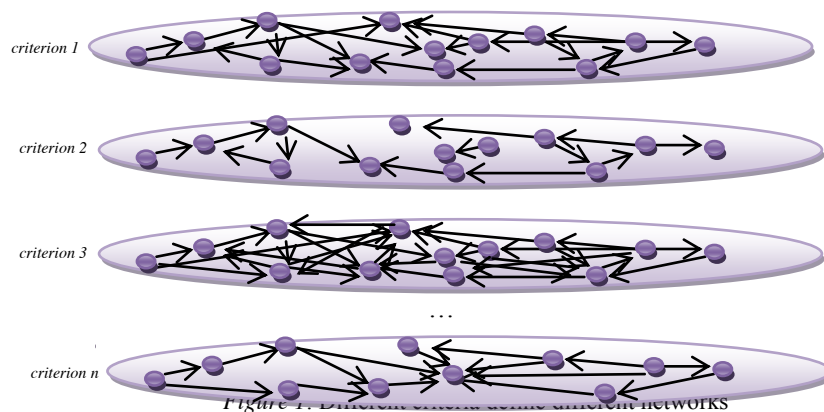


Figure 1. Different criteria define different networks

One company may be more effective, feel more secure or reliant when collaborating with a specific company or group of companies. This requires that the selection of partners follows pre-determined directions of search which demand additional knowledge about the network. In practice, it is often desirable that the companies that will perform a specific project are similar with respect with some issues (for example, organizational

culture or IT usage) and complementary with respect to others (for example, leadership skills, market knowledge or technological strengths). Therefore, we claim that decision support should combine an exploratory process about enterprise's relationships with an algorithm that ranks alternative VE configurations.

The approach taken in this work is quite different from those described in most of the "partner selection" literature in VE (see e.g. Ma et al. (2007) for genetic algorithms, Ng (2007) for linear programming, Li and Liao (2007) for fuzzy theory). Its focus is rather on the *exploratory phase*, performed with the aim of obtaining relevant information about the network. This information is then intended to be used in the subsequent phase of optimally selecting each partner.

The remainder of the paper is organized as follows. In Section 2 the problem is described, in section 3 the exploratory process is presented, in Section 4 an illustrative example is described and finally, in Section 5 some preliminary conclusions are presented.

2. PROBLEM DESCRIPTION

The VE configuration process can be described as follows. Assume a VBE network A representing all potential partners (companies) and their relationships. A specific entity is responsible for the VE formation process (this entity is here referred to as the Decision Maker or DM). Entities and relationships are characterised by a set of m attributes, some assigned to the nodes and some assigned to the edges of the network. These attributes will be used to express the criteria used for evaluating solutions (i.e. VE configurations). The first step in this modelling process is to carefully define what attributes are going to be considered in both subsets.

We consider that the attributes assigned to each node comprise the company's resource availability that will be assessed according to project constraints (e. g. *time windows*, *minimum amount of resources* required) and other company's characteristics relevant to the decision process (e.g. firms' size or financial stability). In the selection process we can consider hard or soft constraints (i.e. we can define strict or loose thresholds for some measures), so that the process does not exclude too many potential candidates in its earliest phases. For example, a good candidate that does not fully respect the time window constraints will still be eligible if there is some slack to re-schedule activities.

The edge attributes include variables that characterise the links between pairs of companies (e.g. assessment of past relationship experiences, distance, level of trust, costs, etc...). We consider that the network is a directed graph because there is the possibility that, for example, the degree of trust between two firms is not reciprocal. From the two sets of attributes (edge and node attributes) it should be possible to derive the organizational characteristics required to achieve the project objectives. The decision maker can give weights to these organizational attributes according to his/hers believes about their relative importance for the project under consideration.

In real-world decision problems we have to handle information that is uncertain, incomplete and/or missing (Li and Lao, 2007). Furthermore, there are many decision situations in which the attributes cannot be assessed precisely in a quantitative form, due to their particular nature (e.g. trust) or because either information is unavailable or the cost of their computation is too high. In these situations an "approximate value" may be acceptable and so the use of a qualitative approach is appropriate (Herrera et al., 2004).

"Linguistic variables" will be used to represent qualitative aspects, with values that are not numbers but words or sentences in a natural language, thus making it easier to express

preferences. *The linguistic term set*, usually called S is a set of linguistic values that are generally ordered and uniformly distributed. For example, a set S of five terms could be defined as follows: $S = \{s_0 = \text{very low}; s_1 = \text{low}; s_2 = \text{medium}; s_3 = \text{high}; s_4 = \text{very high}\}$, in which $s_a < s_b$ if $a < b$. The semantics of the elements in the term set (the meaning of each term set) is given by fuzzy numbers defined on the $[0, 1]$ interval and described by membership functions. For the same attribute, the cardinality of S may vary depending on the decision maker's knowledge about the companies (it may be more detailed in some cases or vaguer in others). Therefore, we allow several types of criteria: numerical, interval, qualitative and binary.

The network is formed by a set of n companies (nodes) connected with each other, capable of performing activities and of providing a finite amount of resources that are available over specific intervals of time. We also assume that project P involves k activities that demand a specific amount Q of resources and have to be performed within a given interval of time S . These activities have some precedence relationships and therefore form an activity network.

The partner selection problem consists in choosing the best group of companies to perform all k activities of project P taking into account a set of evaluation criteria based on the m attributes established for the network. In this problem, the alternatives correspond to groups of enterprises that have the resources and skills needed to carry out the project. Given the multi-criteria nature of the problem, there is generally no "optimal" alternative, and a good "trade-off" solution must therefore be identified.

3. EXPLORATORY PHASE

The exploratory process takes place at an initial phase of the whole decision support process. This phase enables the DM to test various scenarios in which the companies are grouped in different ways, and/or different levels of importance are assigned to the different criteria.

In this work we use cluster analysis to somehow better structure the interesting knowledge about the network. This knowledge can be used to create or avoid some alternatives (potential groups of firms that have the resources and skills needed to carry out the project) or to confine the search to a given cluster of companies.

This approach has been applied in the supply chain area, namely as follows: Ha and Krishnan (2008) to compose a portfolio of suppliers using quantitative criteria; Bottani e Rizzi (2008) to reduce the problem dimension by grouping suppliers based on their similarities (even if they cannot control the number of elements in the clusters); and Sarkar and Mohapatra (2006) to group the selection criteria into long- and short-term categories.

A VE implies the existence of cooperation at several levels, such as R&D, production, marketing or distribution. This requirement can lead to the selection of companies belonging to the same cluster (for example, group of companies with similar (high) technical skills) or to the selection of companies belonging to different clusters according to the specific activities to be allocated (for example, for activities related to distribution choose the companies in the cluster that are stronger at this function).

In spite of the additional computational effort required by this interactive learning process when compared with a less guided search (which may be significant if the network size and/or the number of criteria considered is high), it adds the possibility of identifying different solutions, possibly closer to the DM ideals.

3.1 Criteria selection - dimensions

In order to simplify the decision process, we use in this work the concept of “dimensions”. Here a “dimension” is viewed as a small set of criteria, considered together as a way to obtain a simpler representation of all network characteristics. Attribute selection is an important issue in the VE configuration process as the selected attributes should be able to explain the data, provide significant information and be not redundant.

This process of finding out which attributes should be kept (i.e. identifying attributes that are relevant to the decision making process) often provides valuable structural information and is therefore important in its own right. Moreover, if we take into account the dynamic nature of the network, relevant attributes for one project (VE) may be inappropriate for another.

It is important to notice that only some of the criteria are useful to characterize the enterprise for each dimension (e.g. financial stability), so one key task of the decision maker is to carefully define what criteria are going to be considered (e.g. ROE, Debt/Assets, Cash Flow, etc.). Moreover, these criteria have to be statistically analysed before they can be considered suitable for inclusion. For example, it would be incorrect to consider criteria that are highly correlated.

3.2 Clustering

Cluster analysis (CA) is a popular data mining technique (see e.g. Olafsson et al., 2008) that involves the partitioning of a set of objects into a set of mutually exclusive clusters such that the similarity between the observations within each cluster (i.e. subset) is high, while the similarity between the observations from the different clusters is low. In our case, this technique is useful to determine clusters of companies that can be viewed as related with the others, according to specific dimensions.

Clustering may be categorized in various ways such as hierarchical or partitional, deterministic or probabilistic, hard or fuzzy. The general approaches to clustering are: hierarchical clustering and partitional clustering (e.g. Samoilenko and Osei-Bryson, 2008). Hierarchical clustering forms clusters through agglomerative or divisive methods. The agglomerative method assumes that each data point is its own cluster, and with each step of the clustering process, these clusters are combined to form larger clusters, which may be combined to form a single cluster. The divisive method of the hierarchical clustering, on the other hand, starts with one single cluster containing all data points within the sample and proceeds to divide it into smaller dissimilar clusters.

In partitional clustering, the k-means procedure (Kim and Ahn, 2008) is a simple way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. The centroid of a cluster is the average point in the multidimensional space defined by the criteria, i.e., the cluster’s center of gravity. These centroids should be placed as much as possible far away from each other.

The next step is to take each point belonging to a given data set and associate it to the nearest centroid. After all points have been grouped, new centroids are re-calculated and the points are grouped again. This process is repeated until centroids do not change. The k-means algorithm aims at minimizing an objective function, in this case the euclidian distance between each data point and the cluster centre.

Thus, k-means clustering will produce k different clusters of greatest possible distinction. In our work, since we want to explore the data and we do not know the number of clusters in advance, we will use hierarchical clustering.

4. ILLUSTRATIVE EXAMPLE

Assume we would like to form a VE to perform a project decomposed in 5 activities.

Consider a network with 100 candidates (companies) characterized by 20 attributes (10 nodes attributes and 10 edge attributes) expressed in four different types of information: numerical, percentage, binary and linguistic (table 1). Some attributes are chosen to define clusters of candidates according to several dimensions such as organizational culture, management capability, financial stability or market knowledge. It is reasonable to assume that the group of companies that will perform the project will match better together if they have similar cultures, even if we do not have preferences for a specific culture. On the other hand, the VE may have a better performance if with respect to other characteristic (e.g. leadership, managerial competencies) its companies are complementary.

Table 1. Criteria characteristics

criteria	type	edge attribute	cardinality (for linguistic)	Organizational culture	Competencies
c1	linguistic	yes	5	-	-
c2	number	yes	-	-	-
c3	linguistic	no	7	✓	-
c4	number	no	-	-	-
c5	number	no	-	✓	-
c6	percentage	yes	-	-	✓
c7	linguistic	yes	5	-	-
c8	linguistic	no	5	✓	-
c9	percentage	no	-	-	-
c10	binary	no	-	-	-
c11	linguistic	yes	7	-	✓
c12	number	no	-	✓	-
c13	number	no	-	-	✓
c14	linguistic	yes	5	✓	-
c15	linguistic	yes	3	-	-
c16	number	no	-	-	✓
c17	binary	yes	-	-	-
c18	linguistic	no	7	-	✓
c19	binary	yes	-	-	-
c20	linguistic	no	7	-	-

Notes:c3: attitude towards uncertainty/risk

c5: power distance (# of hierarchical levels from top to bottom of organization)

c8: individualism vs collectivism

c12: age of the organization (years)

c14: masculinity vs. femininity

c6: market entrance capability

c11: managerial skills

c13: productivity

c16: cost (per unit)

c18: technical expertise

In our example we will sequentially use two illustrative dimensions - organizational culture and competencies - following the postulation in section 3. Figures have been randomly generated and the algorithm was implemented in C++ with the use of the SPSS software to perform cluster analysis.

The DM maker will carry out a two steps analysis: first, he/she will partition the companies into groups with similar organizational cultures, and then he/she will distinguish the companies selected in the previous step according to their competencies.

Taking a set of variables based on the Hofstede (2002) framework to define organizational culture (attitude towards uncertainty/risk, masculinity⁴ vs. femininity⁵, individualism vs collectivism, small⁶ vs. large⁷ power distance) and the age of the organization, we have obtained the clusters presented in figure 2 and in table 2.

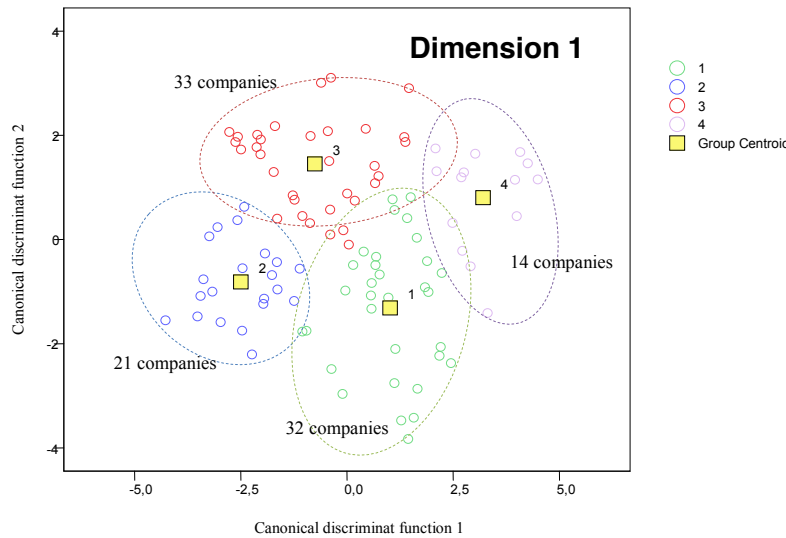


Figure 2. Clusters formation of Dimension 1

The centroid of a cluster is the average point in the multidimensional space defined by the criteria, i.e., the cluster’s center of gravity.

It is very important that the DM describes each cluster carefully in order to verify if the results are valid: cluster 1 includes companies which are neutral towards uncertainty/risk, have in average 6 hierarchical levels, have a individualist culture, are relatively old (approximately 15 years in average) and are neutral in relation to masculinity/femininity. The same kind of analysis must be performed regarding the other clusters.

Table 2. Clusters data of Dimension 1

criteria	Cluster			
	1	2	3	4

⁴ Based on traditional male values (e.g. competitiveness, assertiveness, ambition)

⁵ Based on traditional female values (e.g. relationships orientated)

⁶ People relate to one another as equals regardless of formal positions

⁷ There is a formal hierarchy accepted by all

<i>attitude toward uncertainty/risk</i>	neutral	neutral	keen	keen
<i>power distance</i>	6	6	2	2
<i>individualism vs collectivism</i>	individualist	neutral	collectivist	neutral
<i>age of the organization (years)</i>	14,69	17,57	6,76	5
<i>masculinity vs. femininity</i>	neutral	feminine	neutral	masculine

Notes: a) attitude toward uncertainty/risk = {extremely adverse, very adverse, adverse, neutral, keen, very keen, totally keen}
 b) power distance = {9, 8, 7, 6, 5, 4, 3, 2, 1}
 c) individualism vs collectivism = {very individualist, individualist, neutral, collectivist, very collectivist}
 d) masculinity vs. femininity = {very masculine, masculine, neutral, feminine, very feminine}

The DM may (or may not) prefer one of these clusters. Let us assume, for the purpose of this example, that the DM thinks the organizational culture represented by cluster 1 suits the project better. In this case, companies belonging to the other clusters will be excluded from subsequent analysis.

In the next step, he/she divides the 32 companies from cluster 1 according to their competencies (see the resulting clusters in figure 3 and table 3).

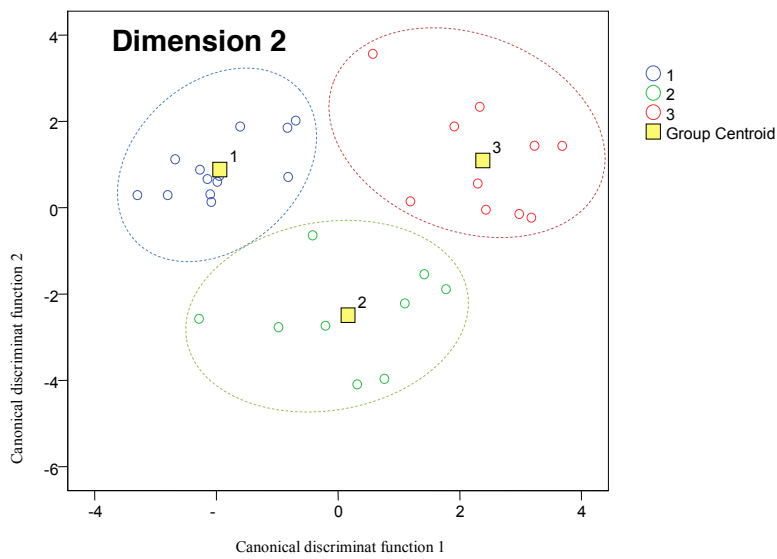


Figure 3. Clusters formation of Dimension 2

Table 3. Clusters data of Dimension 2

criteria	Cluster		
	1	2	3

<i>market entrance capability</i>	39%	35%	74%
<i>managerial skills</i>	positive	neutral	positive
<i>productivity</i>	57,77	31,89	39,80
<i>cost (per unit)</i>	6,46	7,31	7,56
<i>technical expertise</i>	large	large	large

Notes: a) managerial skills = {none, very negative, negative, neutral, positive, very positive, total}
 b) technical expertise = {none, very small, small, neutral, large, very large, total}

In this dimension the DM is looking for complementary competencies, so he/she will choose companies from cluster 1 to perform production tasks, and companies from cluster 3 to perform marketing and managerial activities. In a real situation, involving more companies, the DM may use optimization or a multicriteria ranking algorithm to select the best companies from each cluster (Crispim and Sousa, 2007).

5. CONCLUSIONS

The selection of partners is a critical issue in the formation of a virtual enterprise, the basic problem consisting in choosing the entities to be involved in an emergent business opportunity, according to their attributes and interactions. This paper tries to emphasise the need to obtain knowledge about the network before starting to search best candidates. This exploratory phase demands some interactivity with the DM and can be enhanced by using of cluster analysis. The potential of this approach was demonstrated by a small illustrative example. This example shows the feasibility of the method that allows directional searches in order to identify different types of solutions, hopefully closer to the DM's expectations or to the project specific requirements.

6. REFERENCES

1. Afsarmanesh, H, Camarinha-Matos LM. (2005). A framework for management of virtual organization breeding environments. In: Camarinha-Matos L M, Afsarmanesh H and Ortiz A, eds. IFIP International Federation for Information Processing. Collaborative Networks and their Breeding Environments. Springer, Boston. 2005; 207: 35-48.
2. Bottani, E and Rizzi, A. An adapted multi-criteria approach to suppliers and products selection—An application oriented to lead-time reduction. *International Journal of Production Economics*, 2008; 111 (2): 763-781.
3. Camarinha-Matos LM Afsarmanesh H. Elements of a base VE infrastructure. *Computers in Industry* 2003; 51: 139-163.
4. Crispim, J and Sousa, JP. Multiple criteria partner selection in virtual enterprises. In Camarinha-Matos LM, Afsarmanesh H, Novais P and Analide C, eds. IFIP International Federation for Information Processing. *Establishing the Foundation of Collaborative Networks*. Springer, Boston. 2007; 243: 197-206.
5. Gülçin B, Orhan F and Erdal N. Selection of the strategic alliance partner in logistics value chain. *International Journal of Production Economics*, 2007, DOI:10.1016/j.ijpe.2007.01.016.
6. Gunasekaran A, Lai K-H and Cheng E. Responsive supply chain: A competitive strategy in a networked economy. *Omega*, 2008; 36:4: 549-564.
7. Ha SH and Krishnan R. A hybrid approach to supplier selection for the maintenance of a competitive supply chain. *Expert Systems with Applications*, 2008; 34(2): 1303-131.
8. Herrera F, Martinez, L and Sanchez PJ. Decision Aiding Managing non-homogeneous information in group decision making. *European Journal of Operational Research*, 2004; 166 (1): 115-132
9. Hofstede, Geert. *Culture's Consequences, Comparing Values, Behaviors, Institutions, and Organizations cross Nations*, Sage Publications; Second Edition 2003.
10. Li Y and Liao X.. Decision support for risk analysis on dynamic alliance. *Decision Support Systems*, 2007; 42: 2043- 2059.
11. Ma, X, Han, J, Wei, Z. and Wang, Y. An improved adaptive Genetic Algorithm in optimization of Partner Selection. *Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, SNPDC 2007*. Volume 3, 30 July 2007- 1 Aug. 2007: 455 - 460.
12. Mowshovitz, A. Virtual organizations: A vision of management in the information age, *Information Society* 1994, 10: 267-288.
13. Munda G. Multiple Criteria Decision Analysis and Sustainable Development. In Figueira, J, Greco, S, Ehrgott, M. eds. *Multiple Criteria Decision Analysis: State of the Art Surveys*, 2005, New York, USA, Springer 78: 953-989.
14. Ng W.L. An efficient and simple model for multiple criteria supplier selection problem. *European Journal of Operational Research*, 2007; DOI:10.1016/j.ejor.2007.01.018.
15. Roy, B. *Multicriteria Methodology for Decision Aiding*. Netherlands, Kluwer Academic Publishers, 1996.
16. Samoilenko, S Osei-Bryson K-M. Increasing the discriminatory power of DEA in the presence of the sample heterogeneity with cluster analysis and decision trees. *Expert Systems with Applications*, 2008; 34 (2): 1568-1581.
17. Sarkar, A and Mohapatra, PKJ. Evaluation of supplier capability and performance: A method for supply base reduction. *Journal of Purchasing and Supply Management*, 2006; 12(3): 148-163.
18. Olafsson S, Li X, Wu S. Operations research and data mining. *European Journal of Operational Research*, 2008; 187(3): 1429-1448.