Supplier Selection in Webinar Supply Chain Using Self-organizing Maps and Data Mining

Roya Hosseini, Nahid Ghassabzadeh Saryazdi, Sayyed Alireza Hashemi Golpayegani
Department of Computer Engineering and IT, Amirkabir University of Technology
Tehran, Iran
{hosseini.ec, nghassabzadeh, sa.hashemi}@aut.ac.ir

Abstract

In today's competitive world, companies are constantly investigating new approaches for reducing costs. To reach this goal, companies need to carefully plan supply chain management processes. Supplier selection is one of the most important strategic issues in supply chain. Supplier selection in supply chain management is a multi criteria decision making process which consists of both quantitative and qualitative criteria. By the rapid growth of outsourcing, supplier selection has gained significant importance. Customer satisfaction, and also satisfying customer requirements and priorities all depends on proper selection of suppliers. On the other hand, the complexity and importance of supplier selection decisions, makes the necessity of having a well-defined method more vivid. There are many methods in the literature which are proposed for supplier selection but each of them has certain limitations. In this paper, as there has been less attention to the importance of ever increasing electronic supply chain, we aim to provide a suppler selection method for one of the most important electronic supply chain, namely webinar. In the proposed method SOM and data mining techniques are deployed for the selection of best suppliers.

1 Introduction

In today's competitive and challenging environment, firms can maintain their competitive advantage by constantly managing organizational resources as well as tightening their industrial relationships. More specifically, companies needs to reorganize the supply chain management strategy (SCMS) by simultaneously integrating the organizational resources, information, and activities. However, SCMS may face multi-dimensional difficulties as it involves numerous organizational functions and resources integration among various departments [1]. On the other hand, the intense competitive markets have forced firms to provide quick and precise responds to customer needs for improving their market situation by gaining customer satisfaction.

In such situations, suppliers' role and their related issues in supply chain management have been of high importance. Supplier selection is a group decision making process which requires participation of different organizational levels and groups. It is quite obvious that each group of decision makers can independently investigate different features of suppliers. Hence, supplier selection has a complex and challenging nature and firms need to consider it as part of their strategic plans.

The rapid development of internet and E-commerce has changed the presentation of most physical services to electronic ones. One of the most widespread services in the world is Webinar. The electronic presentation of this service provides ubiquitous access to knowledge and information.

Actually different suppliers are involved in the supply chain of a typical webinar session.

This research aims to help Webinar providers select the best suppliers by considering pre-defined factors which are critical for providing a high quality webinar. Therefore, a novel method for automatic supplier selection is proposed by using machine learning and data mining techniques.

The organization of this paper is as follows: The following section provides a review of the related works that focus on the supplier selection process. Section 3 provides the definition to the basic concepts. Section 4 presents the proposed methodology. Finally, section 5 concludes the paper.

2 Related Works

Proper supplier selection in supply chain management requires huge range of evaluation criteria which consists of information such as customer concentration, competitive priority, strategic sale, information technology and top management support in the organization. There are different criteria for supplier selection in different conditions.

The problem of supplier performance analysis is a muti-criteria decision making process which involves many factors in the hierarchical structure of the decision analysis system. Therefore, in this paper, we will apply a Multi Criteria Decision Making (MCDM) approach in the context of supplier selection problem. A brief review of the major researches that have been centered on the supplier selection is provided herein.

Analytical Hierarchy Process (AHP) method has been proposed by Narasimhan [2], Hill and Nydick [3], and Barbarosoglu and Yazgac [4] in 1983, 1992, and 1997 respectively. In 1998, Ghodsypour and O'Brien [5] applied AHP method for the multi-criteria decision making problem. In 2001, Boer et al. [6] has reviewed different supplier selection method. In 2001, Lee et al. [7] proposed a method for the selection and management of suppliers. In 2002, Sarkis and Talluri [8] suggested using Analytical Network Process (ANP) for the modeling of supplier selection process. In 2003, Dulmin and Mininno [9] investigated the contribution of a multi-criteria decision aid to supplier selection problems. In 2005, Hwang et al. [10] presented a supplier selection method by means of Fuzzy AHP. The proposed method used AHP to weight different supplier criteria and then by using the Integration Model, results of multiple analyses were added and the best supplier was selected.

In 2007, Xia and Wu [11] has proposed to improve AHP by rough sets theory and Mixed Integer Programming and then applied the enhanced AHP method for simultaneous selection of both suppliers and demands. Another research in [12] uses Fuzzy extended analytic hierarchy process and simultaneously considers the impacts of risks on supplier selection problem.

In 2008, Ha and Krishnan [13] have used a hybrid model for supplier selection which applied different techniques in the evaluation process for selection of competitive suppliers in supply chain. In this method a score of a supplier was calculated by defining quantitative and qualitative influential factors on the effectiveness of supply chain. Then by using a clustering analysis method a map of suppliers was created which clusters suppliers based on effectiveness of qualitative and quantitative dimensions. Finally, suppliers were ranked based on the key factors.

In 2009, Ming-Lang et al. [1] proposed a new evaluation method which helped experts select optimal suppliers. In this method, a combination of MCDM, ANP, and Choquet Fuzzy Integral were used for selection of suppliers in the supply chain management. Using MCDM analysis, the best supplier was selected according to the set of predefined criteria. ANP enables considering the relationship and feedbacks among criteria and the Choquet Fuzzy Integral removes subjective judge of experts.

In 2009, Hassanzadeh and Razmi [14] proposed a new framework based on firms' strategies for the management of suppliers which involves supplier selection and evaluation. In this method, the best ISP supplier is first selected by quantitative and then by qualitative criteria. Selected suppliers are then evaluated from the customer, effectiveness, and competitive point of view.

In 2010, Bruno et al. [15] investigated different proposed method for the supplier selection and found that most researches applied AHP to the supplier selection process. AHP is one of the most popular methods which have been used by many researchers [2,3,4]. This method is one of the most powerful methods which enable managers to set the priorities of criteria. Although it is easy to comprehend and use AHP, it has some drawbacks such as considering only the one-directional relationships among criteria. In this case, possible relationships among different groups of factors are ignored. To overcome this problem, ANP can be used. ANP uses the interrelationship among criteria. However, it suffers from complexity as the number of factors and relationship increases.

Therefore, in this paper, we will propose a novel method for supplier selection in the webinar supply chain by using machine learning and data mining techniques. The table 1 compares the most recent supplier selection researches with our proposed method. Each row in the table shows the criteria and each column of the table presents a reference article.

3 Basic Concepts

In this section, we will describe the two major techniques that will be used in our proposed method namely, SOM and data mining. Then we will briefly identify the webinar supply chain which is used as our underlying supply chain architecture in the supplier selection process.

3.1 The Self-Organizing Maps (SOMs)

ANN is composed of richly interconnected non-linear nodes that communicate in parallel. The connection weights are alterable, and allow ANN to directly learn from examples without requiring or providing an analytical solution to the problem.

The most popular forms of learning are:

• Supervised learning: Patterns for which both their inputs and outputs that are known are presented to the ANN. Supervised learner predicts the value of the function for any valid input object after training with a number of training examples. ANN employing supervised learning has been widely utilized for the solution of function approximation and classification problems.

• Unsupervised learning: Patterns are presented to the ANN in the form of feature values. It is distinguished from supervised learning by the fact that there is no a priori output. ANN employing unsupervised learning has been successfully employed for data mining and classification tasks. The SOM and adaptive resonance theory (ART) constitutes the most popular exemplar of this class.

Based on the above descriptions, SOM is one of the unsupervised neural network methods that basically entails a non-linear projection of the probability density function of the input data space into a lower dimensional net of units (output space), each represented by a discrete reference weight vector or representative. In this paper, SOM is used to group similar cases from the collected database of past webinar sessions.

An ANN technique has been successfully applied to many applications. However, there is little research that focuses on the SOM approach and empirical investigation of supply chain related topics. Therefore, this paper uses the SOM technique to select the suppliers in the webinar supply chain. The details are provided in section 4.3.

Table 1: Comparison of the supplier selection methods with our proposed method.

Criteria / Reference	[1]	[12]	[13]	[14]	Our Method
Methodology	Analytic Network Process, Choquet Integral	Fuzzy Extended Analytic Hierarchy Process (FEAHP)	Hybrid Method (AHP, Data Envelopment Analysis, Neural Networks)	Quality Function Deployment, Fuzzy Set Theory	Hybrid Machine Learning-based Method (SOM, Data mining)
Considering Customer Satisfaction	Yes	No	No	Yes	Yes
Considering Risk	Yes	Yes	No	No	No
Considering Performance History Using Data Mining Techniques	No	No	Yes	No	Yes
Focus on Manufacturing Environment(ME)/Service Industry(SI)	ME	ME	ME	SI	SI

3.2 Data Mining (DM)

Data mining known as "knowledge discovery in databases" is the process of discovering meaningful patterns in huge databases. Besides, it is also an application area that can provide significant competitive advantage to make the right decision-making. DM is an explorative and complicated process which consists of multiple iterative steps. Figure 1 shows an overview of the data mining process. In the first step, the main datasets that may be retrieved from operational databases or business transactions is selected.

In the second step, the selected dataset then goes through data cleaning and pre-processing for removing data types with discrepancies and inconsistencies. In the third step, the dataset is examined to identify patterns that express relationships among data by applying some algorithms, such as decision trees, clustering, and Bayesian network and so on. Then the patterns are validated with new datasets or a testing sample. It should be possible to transform the patterns into decision-making rules that are likely to help the administrators to accomplish their objectives. Finally, the steps in the mining process are followed iteratively until meaningful knowledge is extracted.

Although DM has been successfully applied to many domains, there is little research that focuses on the hybrid data mining approach and empirical investigation of supply chain related topics. Therefore, we will use association rules to enhance the accuracy of appropriate suppliers' selection in our supply chain.

3.3 Webinar

Web conference is the web-based and virtual type of conference. By means of an internet connection and a personal computer, each participant can communicate with other participants. This communication is enabled by downloading a program on participants' PCs or by clicking on an invitation e-mail which in turn will trigger a web-based program. The main features of a web conference are as follows: Slide show presentation / Live videos / Web tours / File sharing / Polls and surveys / Screen sharing / and Text chat.
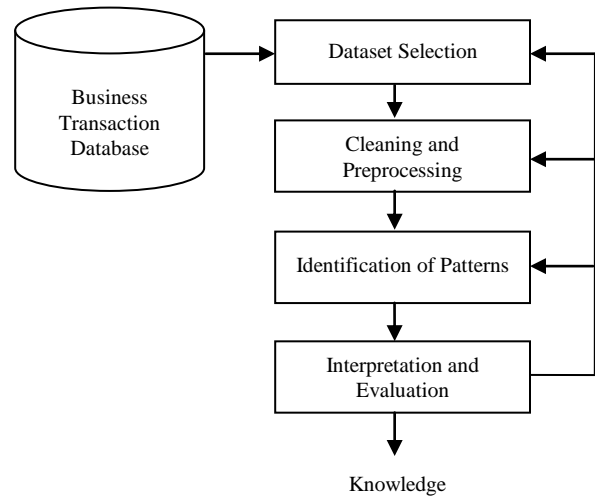

Figure 1: Data mining phases.

Webinar is a new terminology used for description of a specific type of web conference. This type of conference is usually one-way and from the lecturer to the audience. Webinar can also have a collaborative form. In this form, polling and surveying sessions which are held, provides thorough cooperation among presenters and audiences. The presenter may talk on the phone or points to the information on the screen. This way the audience can also answer the questions by phone. New technologies which are used for web conference use VOIP and hence are able to provide a complete web-based communication.

4 Methodology

4.1 Webinar Supply Chain Model

A typical supply chain generally consists of five major layers, namely manufacturers, distributors, retailers, and customers. All of these layers except retailers are considered in the webinar supply chain model. Figure 2 depicts the
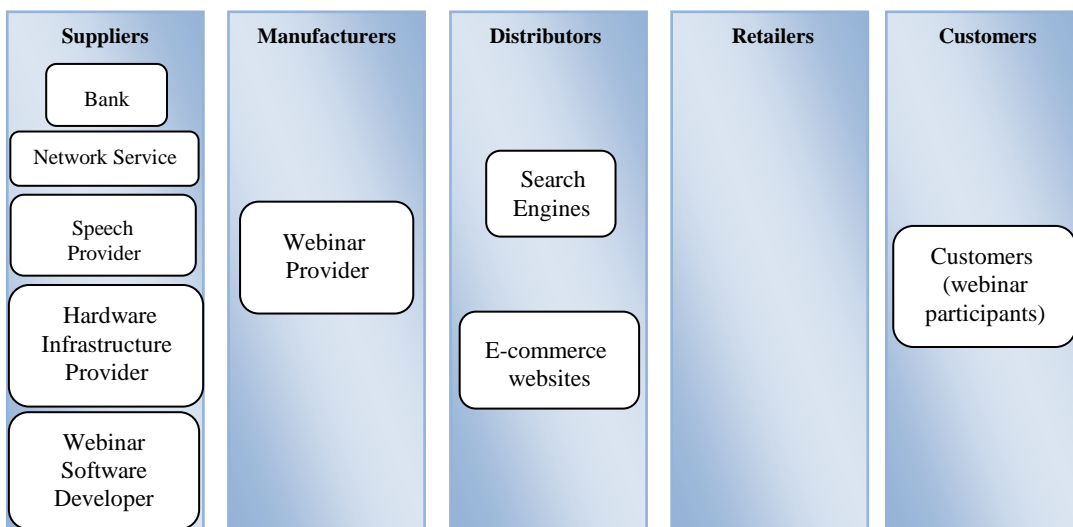

Figure 2: Layers of Webinar Supply Chain.

Table 2: Description of layers for the Webinar supply chain.

Row	Organization name	Organization Role	Description
1	Bank	Supplier	Used in case of electronic payment.
2	Network Service	Supplier	Provides network services for Webinar.
3	Speech Provider	Supplier	Presents speech about the selected topic.
4	Hardware Infrastructure Provider	Supplier	Provides required hardware infrastructure for the webinar sessions.
5	Webinar Software Developer	Supplier	Responsible for designing required webinar software.
6	Webinar Provider	Manufacturer	The sole manufacturer of the chain who delivers the final service of the chain i.e. webinar sessions.
7	Search Engines	Distributors	Used for indexing, and finding webinar sessions.
8	E-commerce websites	Distributors	Served as an advertiser of webinar sessions of a webinar provider.
9	Customers	Customers	The webinar participants.
10	Supply chain management	Supply chain management	Manages the whole supply chain.

different parts of this chain in each layer. Details of layers are shown in table 2.

4.2 Data Collection and Preprocessing

In this section, the data for each factors shown in table 3, should be collected for 600 webinar sessions which are provided by a specific provider. Description of each factor is shown in table 3.

For the purpose of evaluating customer satisfaction, a questionnaire shown in table 4 is presented to each customer. Each of these questions are scored from 1 to 5 (1: Very Bad, 2: Bad, 3: Normal, 4: Good, 5: Very Good). The satisfaction degree for each customer is calculated by averaging the score of each question.

After collecting the data set, it should be filtered to find sessions in which the value of costs and customer satisfaction degree are less and more than a threshold value respectively. To obtain an advisable webinar session, customer satisfaction must at least be of value 3 (Normal) and total cost must be less than a certain value which is determined periodically regarding the specific conditions.

As it is obvious from table 3, some of these factors are quantitative and others are qualitative. Hence, all these data should be normalized for using in the clustering algorithms. After filtering the initial webinar sessions, all the values of data are normalized by (1) as follows:

$$x_t' \equiv \frac{x_t - \overline{x_t}}{\sigma_t} \qquad (1)$$

where x_t', x_t, $\overline{x_t}$ is the normalized, real, and the average value of the t-th factor respectively. σ_t shows the standard deviation of the t-th factor.

4.3 Clustering Webinar Sessions

For clustering webinar session, a specific type of artificial neural network (ANN), called SOM is used. The inputs of this network are webinar sessions. Each session is shown as a seven dimensional vector $X = (x_1, x_2, ..., x_7)$. Each x_i has the value of the i-th factor in table 3. Therefore, SOM network has seven inputs. After clustering, a specific label will be assigned to each session. This label represents the cluster of the session. Each label is in the form of an integer in the range of $[1, n]$ where n shows the total number of clusters. Different size of maps can be used to train network and the result with

higher clustering precision is used in the next step. After training the SOM network, the outputs are sessions that are assigned to different clusters. The sessions in the same cluster have the most degree of similarities with each other. Figure 3 shows the structure of the SOM network for clustering webinar sessions.

4.4 Rule Extraction

After clustering webinar sessions, Association Rule Mining is used for finding specific patterns in each cluster in order to select the best supplier. An example of an extracted rule is as follows:

(Webinar Subject = Cloud Computing) and (Number of Customers = 2000) and (Geographical Distribution of Customers = Normal) and (Maximum Required Bandwidth = 50 MB) and (Desired Software Features = PowerPoint Slides with Video Clips) and (Customer Satisfaction = Good) and (Total Cost = 1000000) → (Speech provider = Queen Inc.)

4.5 Supplier Selection for a New Webinar Session

In the Webinar supply chain speech provider is selected separately for each webinar session while other suppliers are selected periodically. Regarding this fact, after the above steps, the best supplier for a new webinar session can be selected from the data of the new webinar session. Data of a new webinar session will be given to the SOM. The SOM will assign a cluster to the session. After this step, the best supplier will be selected by using data mining rules of the given cluster extracted in section 4.4.
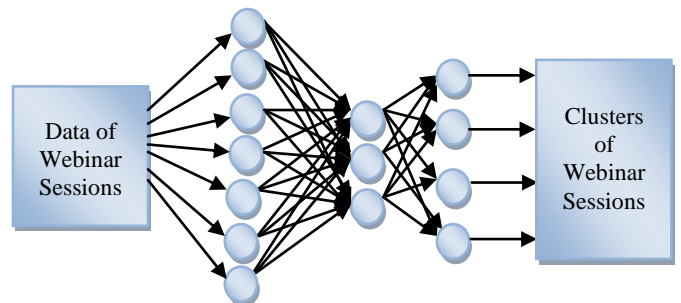

Figure 3: Structure of SOM network for clustering webinar sessions.

Table 3: Description of the main criteria of a webinar session.

Row	Criteria	Description
1	Webinar Subject	8 general subjects are considered: E-commerce, Information Security, Computer Networks, Artificial Intelligence, Multimedia Systems, and Information Systems Management, Cloud Computing, and Software Engineering which are all scored from 1 to 8 respectively.
2	Number of Customers	Less than 200: score 1, between 200 to 500: score 2, between 500 to 1000: score 3, between 1000 to 5000: score 4, and more than 5000: score 5.
3	Geographical Distribution of Customers	Is considered as Low, Normal, and High and score of 1 to 3 are assigned to each one respectively.
4	Maximum Required Bandwidth	In kilo byte per second.
5	Desired Software Features	Determines the features of the webinar software.
6	Customer Satisfaction	Is determined from the answers of the customer to the questionnaire and has the score of 1 to 5 (Very Bad, Bad, Normal, Good, Very Good).
7	Total Cost	Is the total cost from the beginning of the webinar session until its end.

Table 4: Customer satisfaction questionnaire.

Question / Option	Very Bad	Bad	Normal	Good	Very Good
How was the scientific quality of webinar? Was the webinar subject useful for you? Was the webinar cost rational for you? How would you evaluate the speech provider? How would you evaluate presenting other subjects related to the subject of this webinar session? How was the reception quality of the webinar? (receiving webinar continuously without interruptions)					

5 Conclusions

Supplier selection is a group decision making process which requires participation of different organizational levels and groups and hence is a critical and challenging process in the supply chain management. On the other hand, by the growth of internet and information technology, electronic services have gained more importance. Webinar is among the most popular electronic services which are used in our today's competitive environment. Hence, in this paper, we proposed a novel supplier selection method using preprocessed data. The data will be then clustered by using SOM. Next by using data mining techniques, supplier selection rules are extracted from each of the created clusters. Finally, supplier selection for a new webinar session will be done by first finding the cluster of the webinar session and then applying the supplier selection rules of that cluster to the webinar session data.

References

1. Ming-Lang, T., Chiang, J.H. and Lan, L.W., "Selection of Optimal Supplier in Supply Chain Management Strategy with Analytic Network Process and Choquet Integral", *Computers & Industrial Engineering*, Vol. 57, No. 1 (2009), pp. 330–340.

2. Narasimhan, R. "An Analytical Approach to Supplier Selection", *Journal of Purchasing and Materials Management*, Vol. 19, No. 4 (1983), pp. 27-32.

3. Hill, R.P. and Nydick, R.J., "Using the Analytic Hierarchy Process to Structure the Supplier Selection Procedure", *International Journal of Purchasing and Materials Management*, Vol. 28, No. 2 (1992), pp. 31-36.

4. Barbarosoglu, G. and Yazgac, T., "An Application of the Analytic Hierarchy Process to the Supplier Selection Problem," *Production and Inventory Management Journal*, Vol. 38, No. 1 (1997), pp.14-21.

5. Ghodsypour, S. H. and O'Brien, C., "A Decision Support System for Supplier Selection Using an Integrated Analytic Hierarchy Process and Linear Programming", *International Journal of Production Economics*, Vol. 56-57, No. 1 (1998), pp. 199-212.

6. Boer, L., Labro, E. and Morlacchi P., "A review of methods supporting supplier selection", *European Journal of Purchasing & Supply Management*, Vol. 7, No. 2 (2001), pp. 75-89.

7. Lee, E.-K., Ha, S. and Kim, S.-K., "Supplier Selection and Management System Considering Relationships in Supply Chain Management", *IEEE Transactions On Engineering Management*, Vol. 48, No. 3 (2001), pp. 307-318.

8. Sarkis, J. and Talluri, S., "A Model for Strategic Supplier Selection", *The Journal of Supply Chain Management*, Vol. 38, No. 1 (2002), pp. 18-28.

9. Dulmin, R. and Mininno, V., "Supplier Selection Using a Multi-Criteria Decision Aid Method", *Journal of Purchasing & Supply Management*, Vol. 9, No. 4 (2003), pp. 177-187.

10. Hwang, H.-S., Moon, C., Chuang, C.-L. and Goan, M.-J., "Supplier Selection and Planning Model Using AHP", *International Journal of the Information Systems for Logistics and Management (IJISLM)*, Vol. 1, No. 1 (2005), pp. 47-53.

11. Xia, W. and Wu, Z., "Supplier Selection with Multiple Criteria in Volume Discount Environments", *Omega*, Vol. 35, No. 5 (2007), pp. 494–504.

12. Chan, F. T. S. and Kumar, N., "Global Supplier Development Considering Risk Factors Using Fuzzy Extended AHP-based approach", *Omega*, Vol. 35, No. 4 (2007), pp. 417–431.

13. Ha, S.H. and Krishnan, R., "A Hybrid Approach to Supplier Selection for the Maintenance of a Competitive

Supply Chain", *Expert Systems with Application*, Vol. 34, No. 2 (2008), pp. 1303–1311.

14. Hassanzadeh Amin, S. and Razmi, J., "An Iintegrated Fuzzy Model for Supplier Management: A Case Study of ISP Selection and Evaluation", *Expert Systems with Applications*, Vol. 36, No. 4 (2009), pp. 8639–8648.

15. Bruno, G., Esposito, E., Genovese, A. and Passaro, R., "The Analytical Hierarchy Process in the Supplier Selection Problem", Retrieved November 24, 2010, from *http://www.superdecisions.com/~saaty/ISAHP2009/Final_Papers/28_GenoveseA_SupplierSelectionMethodAnalysis_REV_FIN.pdf.*