A Novel Method for Selecting the Supplier Based on Association Rule Mining

Ali Molaali [a], Mohammad Jafar Tarokh [b,*]

[a] *Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*
[b] *Department of industrial engineering, K.N. Toosi University of Technology, Tehran, Iran*

Received 8 October 2015; revised 12 May 2016; accepted 19 August 2016; available online 16 March 2017

Abstract

One of important problems in supply chains management is supplier selection. In a company, there are massive data from various departments so that extracting knowledge from the company's data is too complicated. Many researchers have solved this problem by some methods like fuzzy set theory, goal programming, multi objective programming, the liner programming, mixed integer programming, analytic hierarchy process (AHP), analytic network process model, TOPSIS, etc. Past research gaps are lack of attention to enterprise historical data and extract knowledge from them, review the past performance of suppliers and use effect of the their past performance to their future work. The aim of this paper is to solve supplier selection problem based on historical data by a novel model. The proposed model has tried to uncover hidden relation in massive unstructured industrial data and has used them to extract knowledge for optimizing decision making and predicting in supply chain management by BI tools. The model is based on FP-Growth algorithm integrated with AHP. Moreover, the proposed model is a multi-criteria decision making model (MCDM) with four criteria: quality, priority, delay on delivery and cost that have chosen from literature review. The criteria have been weighed by AHP and finally the model has been validated by industrial group's historical data.

Keywords: *Supply Chain Management, Suppliers Selection Problem, AHP, FP-Growth algorithm, Multi-criteria decision making (MCDM).*

1. Introduction

The importance of supply chain management is significant with the globalization of markets, the rapid development of technology .Due to changes in the business environment and industry, it is impossible to manage industry like before. Decision making is one of important task of each supply chain management that can help CEOs to manage their supply chains. This section has reviewed supply chain management and supplier selection problem:

Supply chain management is increasingly have become important with industry globalization, and the competition between the supply chains will remain likely an important component in the competitive world [2, 3]. A supply chain is a sequences of processes and flows which are at different levels and are combined to fill customer needs for a product [4], it includes the following levels: customers, retailers, general retailers, distributors, manufacturers and component of manufacturers for raw materials [4], through which obtained materials and products are converted and delivered to consumers [1].

* Corresponding author. Email: mjtarokh@kntu.ac.ir

Successful supply chain management requires several decisions related to the information of products and assets, any decision will have to increase surplus is the supply chain [4].

The objective of any supply chain is to maximize the overall value generated [4, 5]. In fact, the main purpose of the supply chain, to satisfy the needs of our customers and making a profit for them [4].

A decision–making process comprising several steps [9], to identity alternative suppliers, manufactures should keep supplier information files and gather information on the past and new suppliers. Supplier information files includes the name of each supplier, a list of materials available from each supplier, the supplier's delivery history, the suppliers quality records, the supplier's overall desirability, and general information concerning the supplier's plant and management [7]. For supplier selection the manager has set of criteria such as quality, technology, and price, to determine how suppler have performed in the previous or current contracts [6].

In supplier selection decisions two issues are of particular significance. One is what criteria should be used, and the other, what methods can be used to compare suppliers [11].

Lima Junior et al [9] has proposed a framework for supplier selection process that consist of four steps: problem definition, formulation of criteria, qualification and final choice. The first step aims at clearly defining the problem at hand, which may mean searching for new suppliers for a completely new product, replacing current suppliers, or choosing suppliers for new products from the existing pool of suppliers. In the next step, the buyer should convert its requirements into decision criteria so as to guide the choices. In the qualification step, the main objective is to reduce the initial set of suppliers by sorting potential suppliers from the initial set of suppliers based on qualifying criteria. The last step aims to rank the potential suppliers so as to make the final choice.

There are a number of management decision concerns to maintain a competitive supply chain: [8]

- Which suppler should be considered for partnering?
- Which supplier should be part of supplier development initiatives?
- Which supplier must be removed from the supply base?
- How can weak suppliers improve their performance?
- How can firm effectively allocate resources to supplier development programs?

2. Literature Review

The number of supplier selection models has been increased exponentially over the past couple decades [8]. There are several supplier selection methods available in the literature. These methods are mostly goal programming (GP), multi objective programming (MOP), the liner programming (LP) or mixed integer programming (MIP), The liner scoring model, AHP (analytic hierarchy process), ANP (analytic network process) model, TOPSIS, Mathematical programming, neural networks, agent based theories, rough set theory, clustering algorithms, genetic algorithms, quality function deployment,… [12, 13, 6, 7]. Here a comprehensive survey has been prepared on supplier selection problem and its solutions that have used in papers in reliable journals and conferences from 2000 to 2016, finally it has categorized and is shown in table 1.

Kumar [12] has integrated fuzzy set theory, analytic hierarchy process and neural network to provide group decision support under achievement. Moghaddam [14] has developed a fuzzy multi-objective mathematical model to identify best supplier and have found the optional number of new and refurbished parts and final products in a reverse logistic network configuration. The objective functions in this study were total profit, total defective parts, total late delivered parts, and economic risk factor. Ayhan & Kilic [15] have used F-AHP (Fuzzy analytical Hierarchy process) and MILP (Mixed Integer Liner programming) to solve problem. Their approach has two stages, in the first stage, the relative weights of each criterion for each type of item is determined via F-AHP technique. In second stage, their outputs are used as input in the MLIP model to determine the suppliers. We chose 4 criteria Price, Quality, Delivery Time Performance and after Sale Performance. Jadidi et al [16] have modeled a supplier selection problem as a multi objective with three minimization objectives: price, reject and lead time, then they for solving problem proposed a multi choice programing (MCGP) approach. Yu and Wong [17] have proposed an agent-based negotiation model to automate the supplier selection process involving a bundle of products with synergy effect, a Mas is established to realize the proposed negotiation model for multi-product supplier selection. Karsak and Dursun [18] have proposed a novel Fuzzy multi-criteria group decision making framework integrating quality function deployment (QFD) and data envelopment analysis (DEA). The Lower and upper bounds of the weights of supplier assessment criteria are identified by adopting fuzzy weighted average (FWA) method. Igoulalene et al [19] have formulated a multi-stakeholder multi-criterion [MSMC) decision making problem and solved using two novel approaches. The first hybrid approach combines the fuzzy consensus-based possibility measure and fuzzy TOPSIS method. The Second hybrid approach combines the fuzzy consensus–based neat OWA and goal programming model. Lirong [20] has used the multi-objective intelligent weighted grey target decision method to dimensionless, processing for supplier evaluation indicators, and then the improved triangle whitening weight function clustering method is applied to evaluate supplier clustering. Their criteria were quality, cost, delivery, cooperation, competitiveness, services support. Sakis and Dhavale [8] have taken a triple- bottom- line (Profit, people and planet) approach. They developed a novel methodological approach based on Bayesian framework and Monte Carlo markov chain (MCMC) simulation to rank and select supplier using specific selection objective. Their criteria were cost variance from expected cost, Acceptance rate for supplier, percent orders before of on due date, energy efficiency of supplier operation, environmental fines, use of green technology and pollution control, contribution to philanthropic causes, social responsibility to community, wages and working condition. Chai and Ngai [21] have proposed novel interval and hesitant fuzzy model. This model shows significant capabilities in handling ambiguous judgment of stake holders and unbiased value preservation of conflicting opinions. Dargi, et al [22] have deployed Nominated Group Technique (NGT) to extract the most important criteria and then used Fuzzy Analytical Network process (FANP) to weight the extracted criteria. Nazari-shirkouhi, et al [23] have developed an interactive fuzzy multi-objective liner programming model to solve the multi-objective supplier selection problem under multi-price and multi-product in the fuzzy environment. Arikan [24] has proposed fuzzy multi-objective programming for multiple sourcing supplier selection problem. Lima Junior et al [9] have proposed competive analysis of use of fuzzy TOPSIS (fuzzy technique for order of preference by similarity to Ideal solution) and fuzzy AHP to aid the supplier selection decision process. Kar [25] has proposed a heuristic model to prove group decision support and predictive decision support by integrating AHP, fuzzy set theory, and mathematical programming theories their Criteria were product, delivery compliance, price technology capability, production capability, financial position, E-transaction capability. Rouyendegh and Saputro [10] have provided fuzzy TOPSIS and multi-choice Goal programming (MCGP) methods. Their criteria were supply capacity, production capacity, response time, production technology, price, warranty, and procedural compliance, and purchase transaction, communication, on time delivery system, quality, completed shipping document, quantity, financial position, location, Reputation, Management and organization. Gholamhossein, et al [26] have constructed the supplier selection index from quality, reputation, technology, delivery, replacement, cost, agility and risk. then use the fuzzy TOPSIS method to select supplier. Jazemi, et al [27] have used coordination between buyers and supplier to solve supplier selection problem. They proposed a fuzzy model, proposed model was "fuzzy multi-objective mixed integer non liner" problem, with three objective. These objectives were cost, quality and timely delivery. A heuristic method is used to gain an efficient solution of the problem. Liu [11] has used an interview scale set of questions, to determine the importance of supplier attributes. They weighted sub criteria and then criteria by knowledge gathered from questionnaire. Finally, they choose these criteria: cost, quality, relationship, delivery, financial and they choose some sub criteria for each criteria. Golmohammadi, et al [6] have developed a decision-making to select suppliers using neural networks (NNs). They have used historical supplier performance data for selection of vendor supplier. The manager’s judgments about suppliers were simulated by using a pair wise comparisons matrix for output estimation in NN. Genetic algorithm (GA) was applied for initial weights. And supplier’s data base information (input) can be updated overtime. Wu [28] has used rough theory to reduce supplier's attribute, then used novel support vector machine to solve problem. Fangqi, et al [29] have evaluated their criteria in three stages: At first, computing value of manufacture criteria and then computing value of supplier's criteria, and finally determining value of finishing the manufacturing task. They solved problem by using Multidimensional vector to find lowest distance, and then use GA to optimize the result. They have four criteria: product, price, processing time on time delivery, product quality. Guosheng & Guohong [30] have used support vector machine (SVM) technique to select supplier then they compared SVM and BPNN. They believed SVM methods are superior to BPNN.

Kong et al [31] have used linguistic value to assess weights for selection of supplier. They describes a fuzzy APH to determine the weighting of subjective judgment. Then the best supplier can be derived with the grey relation model based on the concepts of fuzzy distance to evaluate and select the best supplier. Lee, et al [7] has integrated supplier selection and management system (SSMS) that include, purchasing strategy system, supplier selection system, and supplier management system. The methodology identifies the managerial criteria using information derived from supplier selection process. The weights of criteria are determined by AHP. Their criteria were quality, cost, delivery and services.

In this section, a comprehensive review has conducted on the previous researches. The research gaps are: researchers have not mined historical data to improve decision-making of managers, have not extracted knowledge from them, haven’t reviewed the past performance of suppliers and haven’t used effect of their past performance to their future work.

Table. 1. Categorized summery on literature review

Ref. number	algorithms	Weighting algorithms	criteria
6	NN	Genetic Algorithm (GA)	
7	Heuristic solution	Analytical Hierarchy Process (AHP)	Quality, Cost, Delivery, Services
8	Bayesian Framework, Monte Carlo Markov Chain		
9	F-AHP, Fuzzy TOPSIS		
10	Fuzzy TOPSIS,GP		Quality, Cost, Delivery, Financial, Reputation, Technology, Delay, Production Capability, Geographical Location, After Sale Performance, Quantity, Shipping Document, Management & Organization, Communication System, Purchase Transaction, Supply Capacity
11	Questionnaire		Quality, Cost, Delivery, Relationship, Financial
12	Fuzzy set theory ,AHP,NN		
14	Fuzzy set theory		
15	F-AHP, MILP		Quality, Cost, Delivery, Services, Reputation, Production , Technical Capability, Geographical , Location
16	Multi choice programming		Cost, Delay
17	Agent-based Model		
18	Fuzzy set theory ,QFD ,DEA	Fuzzy weighted Average (FWA)	
19	Fuzzy-CPM, Fuzzy-CNO, Fuzzy TOPSIS		
20	Triangle whitening weight function		Quality, Cost, Delivery, Cooperation, Competiveness
21	Interval and Hesitant Fuzzy		
22	Nominated Group Technique	Fuzzy Analytical Network Process(F-ANP)	Quality, Cost, Delivery, Services, Reputation, Production , Technical Capability, Geographical , Location
23	F-MOLP		
24	F-MOLP		
25	Fuzzy set theory ,AHP, Mathematical Programming		Cost, Delivery, Financial, Technology, Production Capability, E-transaction Capability
26	Fuzzy TOPSIS		Quality , Cost, Delivery, Cooperation, Competiveness
27	Fuzzy set theory, Integer non Liner		Quality, Cost, Delivery, Financial, Reputation, Technology, Delay, Production Capability, Geographical Location, After Sale, Performance, Quantity, Shipping Document, Management & Organization, Communication System, Purchase Transaction, Supply Capacity
28	Rough Theory, Support Vector Machine (SVM)		
29	Multi-Dimensional Vector, Genetic Algorithm (GA)		Quality, Cost, Delivery, Delay
30	Support Vector Machine (SVM), BPNN		
31	Grey Relation Model	Fuzzy Analytical Hierarchy Process (F-AHP)	

3. Proposed Model

The paper has proposed new multi-criteria decision-making model based on mining association rules to solve the supplier selection problem, then for weighting the criteria, analytical hierarchical process (AHP) have been used. The proposed model consists of four stages are shown in figure 1:

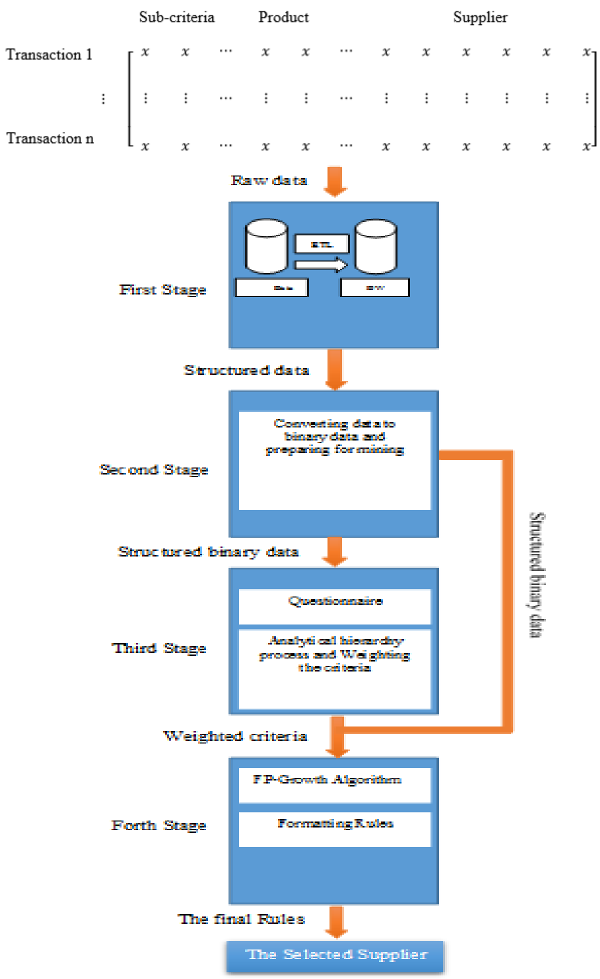

Fig. 1. Schematic view of proposed model

3.1. The First Stage

Before exploring the association rules, data must be prepared by a novel method for mining. In this stage, the data should be extracted from data sources and stored in a single database, then data should be cleaned up and the extraction, transformation and loading (ETL) operations are done on them.

3.2. The Second Stage

Selected ARM algorithm just can accept binary data as input, while historical data usually are nominal or numerical. This part of paper has proposed a solution for converting the data into binary data. To do this, a matrix has been used, where the rows indicate transaction and columns indicate suppliers, products and criteria. The below procedures is used to fill the matrix:

1. For Element, a_{ij} where "i" represents the number of rows or transaction and "j" is the number of columns or suppliers, products and criteria. (Initial value of i is equal to 1).
2. First j=1, then (m=1, m<=s, m++):

$a_{ij=1}$,else $a_{ij=0}$

(Where supplier "m" is processing while "m" is the code of suppliers, "s" is the number of suppliers)

3. First j=m+1 , then(n=1 , n<=k ,n++):

$a_{ij=1}$,else $a_{ij=0}$

(Where product "n" is processing while "n" is the code of products, "k" is the number of products)

4. First j=m+n+1 ,then (p=1, p<=c, p++):

$a_{ij=1}$,else $a_{ij=0}$

(Where criteria "p" is processing while "p" is the code of criteria, "c" is the number of criteria)

5. i=i+1, then go 1
6. Go 1

After running the above procedure, all data is converted to binary data. Overview of the matrix shown in Figure 2:

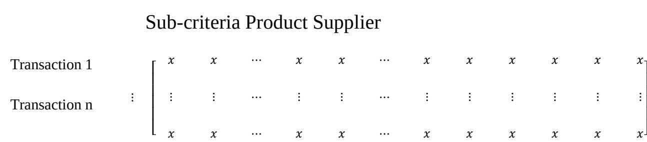

Fig. 2. Convert matrix

3.3. The Third Stage

According to literature review, a number of criteria and sub-criteria have been chosen that are important in decision-making. To obtain the priority of criteria and sub-criteria to each other, a questionnaire has been used, and it was provided to the expert of manufacturer Company for the survey on the importance of the criteria. This questionnaire compared criteria and sub-criteria to pair on a scale of 1 to 9 (Figure 3). Respondent must determine the priority rate of a criteria than other one by ranking. First, all the main criteria, and then the sub-criteria compared to pair separately. Finally, the results have been used for weighting by AHP. In this step, the relative weight of each sub-criteria should be multiplied by the weight of higher criteria (that have gathered from AHP) to obtain the final weight. With this step for each sub-criteria, the final weight is achieved:

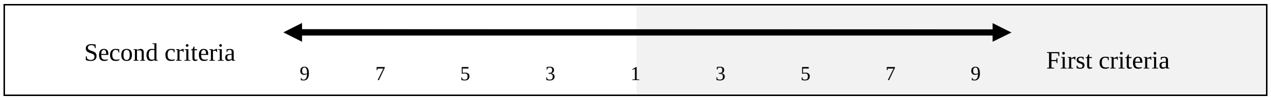

Fig. 3. Compare scale

$$Final\ W_{SC} = \prod (W_{SC} W_C) \quad (1)$$

(Where: Final W_{SC}= final weight of sub-criterion, W_C= the weight of higher criterion, W_{SC}= weight of sub-criterion)

Thomas has shown that if inconsistency ratio less than 0.10 represent acceptable comparison consistency and otherwise the comparisons should be revised [32, 33].

3.4. The Fourth Stage

This stage uses output of the second and third stages as input. The fourth stage is divided into two parts: FP-growth algorithm and formatting rules.

The first part uses output of the second stage, structured binary data, as input and deals to explore the association rules by FP-Growth algorithm. We are considered MST=40% and MCT=60%. The output of this part is unformatted rules that have minimum of MST and MCT.

According to association rules mining, If I = {i_1, i_2, ..., i_d} is considered as a set of data and T = {t_1, t_2, ..., t_n} be a set of transactions, While each transaction t_i should be a subset of I set items. The general overview of association rule is same as below form:

$$X \Rightarrow Y \quad (2)$$

(Where: $\epsilon I, Y \epsilon I\ AND\ X, Y \neq \emptyset\ AND\ X \cap Y = \emptyset$)

The Strength of an association rule could be measured by using two amounts of support and confidence. The support determines how times are both items X and Y comes together, while the confidence determines how many repetitive items in Y are involved in the transactions of X. Support threshold of an item set (X) in a transaction (T), which is mentioned as supp (X), represents the number of transactions that includes item set X:

$$\text{supp}(X) = \{ti | X \subseteq ti, ti \in T\} \quad (3)$$

The mathematical definition of these two values for the **X ⇒ Y** rule is as follows [34]:

$$\text{supp}(X \Rightarrow Y) = P(X \cup Y) \quad (4)$$

$$\text{conf}(X \Rightarrow Y) = P(Y|X) = \frac{\text{supp}(X \cup Y)}{\text{supp}(X)} \quad (5)$$

As stated earlier, FP-Growth algorithm's output is a set of rules, however, not all them are applicable. So, here two parts of antecedent and consequent have been set based on the weight of the AHP and then rules have been classified. As shown in Figure 5, the antecedent part includes the supplier, products and all of sub-criteria, with the exception of the most scored sub-criteria placed in consequent part. Now all outputs rules of FP-Growth algorithm that fit in this format will be accepted and will be sent to final part of model.

Finally, operation of third and fourth stages have formulated as below to calculate highest score of each supplier that they can gain:

$$S_{score} = \max\left(\prod F_c \left(\sum_{i}^{n} E_i W_{ci}\right)\right) \quad (6)$$

(Where Sscore =the highest score of the supplier for each product, Fc =confidence index, i=sub-criteria counter, n=number of sub-criteria, Ei=sub-criteria existence index, Wci=sub-criteria weight)

4. Model Implementation

To implement model, the historical data of a refrigerator manufacturer have been used. This company collected data from different departments during 2007 to 2014 and stored in a database. Data related to supply items were about 1 million transactions.

Table. 2. Selected criteria and sub-criteria and ranges

Man criteria	**Delay in delivery**		**Cost**		**priority**		**Quality**	
Sub-criteria	D_1	Without delay	C_1	-20% to -5%	Pr_1	Less than 3 days	QC_OK	Quality Approved
	D_2	One to five days.	C_2	-5% to +5%	Pr_2	3 to 14 days	QC_NOK	Quality NOT Approved
	D_3	5 to 10 days	C_3	+5% to +20%	Pr_3	14 to 30 days		
	D_4	10 to 15 days			Pr_4	30 to 60 days		
	D_5	15 to 30 days			Pr_5	More than 60 days		

According to the reviewed literature and company data, 10 suppliers, 12 products and four criteria including quality, priority, cost and delay in delivery have been selected. Then, to initialize main criteria, some sub-criteria have been selected (including intervals or numerical ranges for each main criteria) that is shown in the table 2.

After choosing criteria, data is ready to enter the FP-Growth algorithm, so mentioned procedure of previous chapter have been used to convert data in a matrix. Some symbols have been replaced with supplier's names (from S_1 to S_10) and also product's names (from Part_1 to Part_12).

A sample transaction is explained in figure 6. In this transaction, supplier 5th (S_5) arranged to procurement the 8th product (Part_8) with first priority (Pr_1), means delivery below three days. This product must be delivered with first delay (D_1), means without delay and third cost (C_3) means +5% to +20% above the cost stated by the producer, ultimately the products are quality verified at the factory (QC_OK).

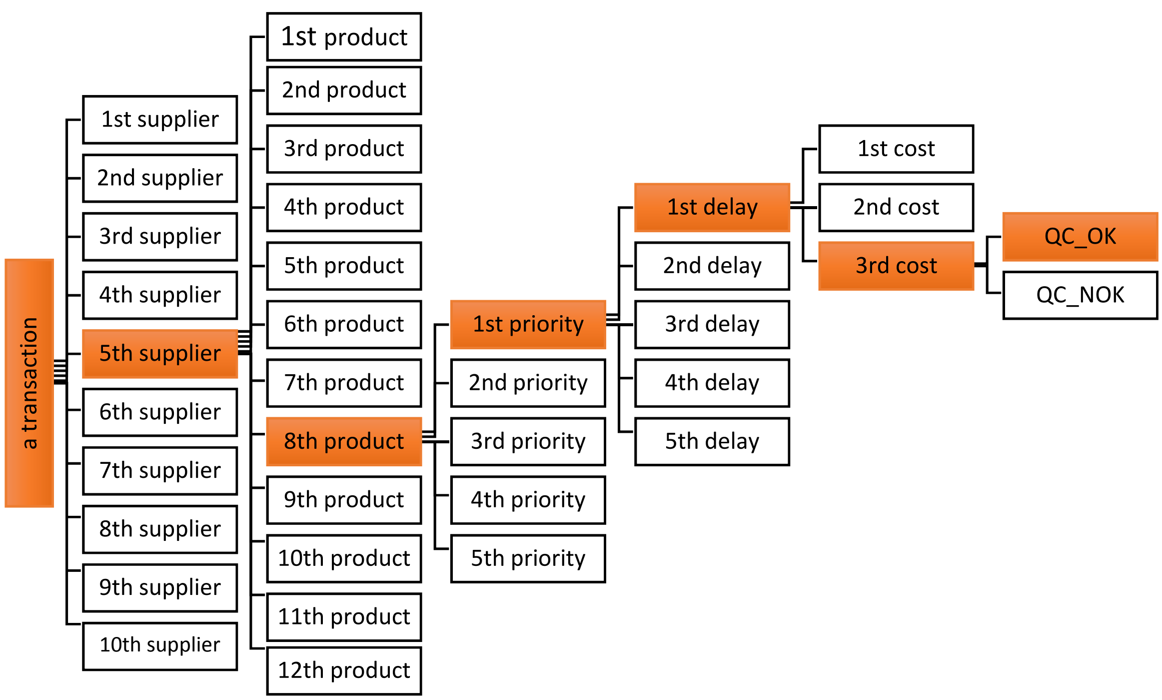

Table. 3. Criteria weighting result (software output)

Criteria	**Quality**		**Delay**					**Cost**			**Priority**				
Weight	0.056		0.279					0.139			0.525				
Subcriteria	**Qc_nok**	**Qc_ok**	**D_1**	**D_2**	**D_3**	**D_4**	**D_5**	**C_1**	**C_2**	**C_3**	**Pr_1**	**Pr_2**	**Pr_3**	**Pr_4**	**Pr_5**
weight	0.1	0.9	0.513	0.261	0.129	0.063	0.033	0.637	0.258	0.105	0.513	0.261	0.129	0.063	0.033

A questionnaire was prepared for paired comparison of criteria and provided to 10 experts of the company (purchasing managers and sales managers). They were asked to compare the main criteria and sub-criteria separately, two by two and ranked them according to preference of any criteria than other between 1 and 9, then the results were averaged and given value among 1, 3, 5, 7 and 9.

Now, binary data is ready to enter to the association rule mining, so, FP-Growth algorithm has used for this purpose. The output of this stage is a series of rules. To ensure about the produced rules, MCT and MST have chosen 60% and 40%, but unfortunately all rules cannot be used. To weight the criteria and sub-criteria, questionnaires and AHP were used. AHP has been implemented by using ExpertChoice11 software. Inconsistency ratio of paired comparisons of main criteria, priority criteria, delay criteria, cost criteria, quality criteria and hierarchy were 00.526, 0.0535, 0.0535, 0.0367 and 0.0523. Weighting results are summed up in table 3.

Now, the final weight of sub-criteria for the ranking has been computed due to the formula (1). The final weight of sub-criteria is shown in table 4.

To implement FP-Growth algorithm, RapidMiner6 software has been used. A format has been designed for rules in regard to the obtained weights from formula (1). As it is observed, the most preferred criteria are first priority (the priority lower than three days). So it is placed in antecedent part and all the other criteria in the consequent part of rules. The determined format for rules should be established same as figure 7.
The generated rules have been ordered by the product, then supplier's score has been calculated for each product according to the obtained rules and calculation of formula (6). The results of this calculation are presented in table 5.

Table. 4. The final weight of sub-criteria

Sub-criteria	1st priority	2nd priority	3rd priority	4th priority	5th priority	1st delay	2nd delay	3rd delay	4th delay	5th delay	1st cost	2nd cost	3rd cost	1st quality	2nd quality
Symbol	Pr_1	Pr_2	Pr_3	Pr_4	Pr_5	D_1	D_2	D_3	D_4	D_5	C_1	C_2	C_3	QC_OK	QC_NOK
Final weight	0.269	0.137	0.067	0.033	0.017	0.143	0.072	0.035	0.017	0.009	0.088	0.035	0.014	0.050	0.005

[Supplier]	[Product]	[Delay criteria]	[Quality criteria]	[Cost criteria]	→	[Less than three days priority]

Fig. 7. Determined format for rules

Table. 5. Selected supplier for each product that has gain highest score

Product	Selected supplier	Score
1	First	0.11625
2	Second	0.077444
3	Second	0.118672
4	First	0.10841
5	Third	0.10883
6	7th	0.110059
7	9th	0.110714
8	Second	0.112107
9	Third	0.11288
10	Third	0.107308
11	10th	0.116812
12	5th	0.1176

5. Model Validation

First, the validation of the past researchers' model has been reviewed in table 6.

Table. 6. Review on evaluation and validation methods of previous models

Ref. number	Evaluation and validation method
22	By a questionnaire (Check List) contains 66 questions on 7 sections for each criteria, the supplier was evaluated, but not compared the validation.
10	Data of a factory includes 17 criteria has used, but not compared to validation.
25	the test data is used for evaluation
17	there is no evaluation and validation
8	there is no evaluation and validation
28	A numeric sample was designed to evaluate, but not validation.
35	Use two data sets, one for evaluation the model and the other for model validation test.
18	The hospital data was used for evaluation but not validation.
14	A numerical example is used that the results show the effective function of model.
15	To evaluate the model , data of an oil motor producer has used, but not validation
6	Validation of model is done by asking experts
7	The data sample of a Korean air conditioner manufacturer is used, the t-test was used for validity measurement.
11	A questionnaire was used to evaluate the model.
31	A numerical model has been used for evaluation.
20	To evaluate the model , a numerical model of aircraft parts company has been used
29	To evaluate the model, a numerical model of plastic injection company has been used. To measure the validity of model, the calculation of data distances and data received from supplier have been used.

As it was seen how researchers assess the validity in past, on the one hand it is impossible to compare results with past researchers' results because of the lack of access to researchers' data and the full version of their model. On the other hand, it is impossible to assess the exact validity of supplier selection compared to other model because who knows which is the best performance, Golmohammadi et al have mentioned: “To compare the performance of the proposed model, a basis is needed to determine the deviation of the results from the best results. One way is to compare the results with other decision- making models for a system, but no one can evaluate the results except the expert of that system. Furthermore, evaluation via comparison with other systems may require cumbersome data reformatting. The best judgment for the resulting solution can be obtained from the experts of the system [6] ". So it has intended to use the current company data to obtain the results deviation with the best possible results. Actually proposed model has compared to real company performance to know which one in more optimized than other one.

We offer in two parts to measure validity and reliability of the proposed model and its results: validation using the algorithm validation factors and calculate the normalized score for each supplier criterion of system data and comparison with results of model.

Due to the large volume of data and calculations and limitations of research, we just intent to prove two above parts for first product.

5.1. The Algorithms Validation Factor

Many algorithms use some values to validate their output as validation factor and then define the acceptable range for these factors. MCT and MST have been set 60% and 40% to validate produced rules, all laws that are less than the mentioned amounts were unacceptable and were not used.

In addition, AHP is used inconsistency ratio to validate, Saati is the inventor of the AHP, all results with inconsistency ratio less than 0.1 is acceptable. As shown before, all of our results are acceptable.

5.2. Calculate Normalized Criteria Weight

This part examines the optimization of proposed model by using calculating the deviation between the results of model and result of normalized criteria weight. For this purpose, the normalized weight of each criterion has been calculated for each supplier who buys first product and has been able to satisfy the first priority. Finally, all obtained scores by suppliers that have been extracted from the past transaction of company were compared with the scores obtained from proposed model and it has expressed if the experts of company use proposed model, how much improvement was seen in the results of supply?. Calculation of supplier score has been divided to 3 part (based on number of criteria), supplier cost score, supplier quality score and supplier delay on delivery score .in each part it has calculated normalized criteria weight as supplier score.

$$S_C = \sum_{i=1}^{number\ of\ sub-criteria} \left(\frac{(\prod C_{Cni} W_{Ci})}{n_{Ci}} \right) \quad (7)$$

(Where SC: cost criteria score, CCni: number of nth supplier transactions that include ith cost sub-criteria, WCi: ith cost sub-criteria weight, nCi: number of transaction that include ith cost criteria (all of supplier))

$$S_Q = \sum_{i=1}^{number\ of\ sub-criteria} \left(\frac{(\prod C_{Qni} W_{Qi})}{n_{Qi}} \right) \quad (8)$$

(Where SC: quality criteria score, CQni: number of nth supplier transactions that include ith quality sub-criteria, WQi: ith quality sub-criteria weight, nQi: number of transaction that include ith quality criteria (all of supplier))

$$S_D = \sum_{i=1}^{number\ of\ sub-citeria} \left(\frac{(\prod C_{Dni} W_{Di})}{n_{Di}} \right) \quad (9)$$

(Where SD: delay criteria score, CDni: number of nth supplier transactions that include ith delay sub-criteria, WDi: ith delay sub-criteria weight, nDi: number of transaction that include ith delay criteria (all of supplier))

6. Discussion

This paper has suggested a novel model based on association rule mining and AHP to choose the most optimized supplier. The results of the formulas (7), (8) and (9) on historical data (product 1) are given in table 7.Then, result of normalization of criteria weights (from historical data) and proposed model's results have compared (for product number 1).

Table. 7. Result of normalization of criteria weights (product 1)

	Quality (S_Q)	**Delay (S_D)**	**Cost (S_C)**
First supplier	0.022263	0.024775	0.021129
Second supplier	0.00444	0.005543	0.00456
5th supplier	0.01821	0.019912	0.016775
6th supplier	0.00257	0.002973	0.002288
7th supplier	0.00002	0.000058	0.000005
10th supplier	0.00203	0.002246	0.001995
Our selected supplier	0.05	0.017	0.088

As mentioned in figure 8, which is displayed on the graph, the f optimization of model selected supplier's gathered score for product number 1 than selected supplier from normalized score (historical data) is obvious. That indicating improving results of the proposed model.

The optimization percentage of supplier score is shown in table 8. As mentioned, quality criteria, cost criteria and delay on delivery criteria have achieved +224.58%, 416.48% and -31.38%. Finally, by multiplying the criteria weight in optimization percentage and averaging of them, the proposed model provide 20.57% better results.

Table. 8. Optimization percentage

	Quality	**Delay**	**Cost**
Our selected supplier	0.05	0.017	0.088
Historical data best supplier	0.022263	0.024775	0.021129
Optimization percentage	224.58	-31.38	416.48

7. Conclusion

In the recent years, many researchers have solved supplier selection problem using different approaches. This paper has reviewed these approaches and has found research gaps that showed they have not investigated some methods like historical data mining, haven't extracted knowledge, so it has proposed a novel model to solve problem. The proposed model is based on association rule mining and AHP. It has used heuristic method to ready data for mining, then has used FP-Growth algorithm to extract knowledge, and has used AHP to weight criteria, finally model has been validated by an industrial company's data, we have compared our result with historical result, as seen, the results has indicated the improvement of proposed model's output compared to the historical data's output.

This project has faced some problems like unavailability of historical data, data inconsistency, high volume of computing, lack of research about implementation of this method, lake of method for validating SSP models and etc.

To continue or improve research, following topics have proposed:

Using more efficient algorithms to mining for the association rules, using algorithms that aren't independent to the questionnaire for weighting the criteria, reduce the size of paired comparisons in the AHP, using this method at other levels of the supply chain models, add criteria such as quantity, risk, agility, supplier's features, new services, electronics, communications, and cooperation, providing an optimal method for validation of SSP models.

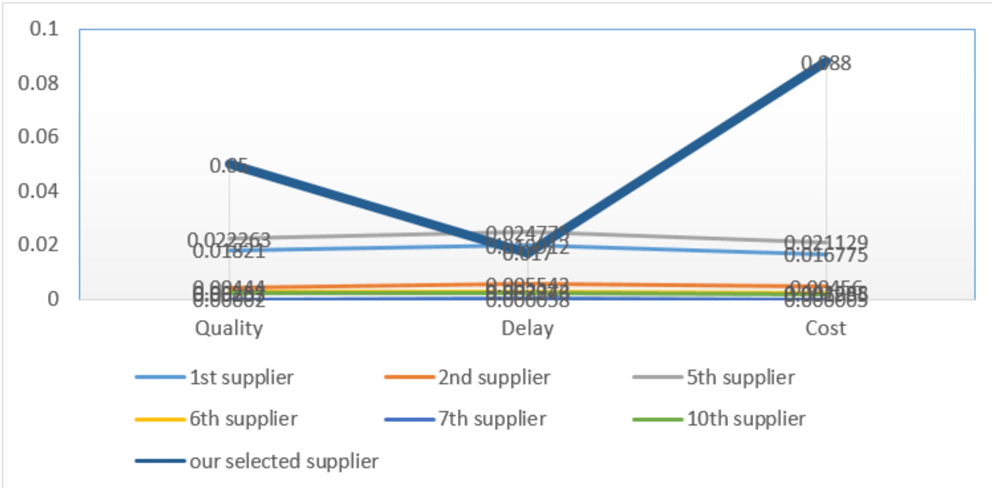

Fig. 8. Normalized score and our model's result comparison (for product 1)

References

[1] Li, Y., "Adaptive multi agent modeling in intelligent supply chain management", 2008 ISECS International Colloquium on Computing, Communication, Control, and Management ,Volume 3 , pp: 330 – 333 (2008).

[2] Samuel, K. E.; Goury, M.L.; Gunasekaran, A.; Spalanzani, A., "Knowledge management in S C: An empirical study from France", The Journal of Strategic Information Systems archive ,Volume 20, Issue 3, pp: 283-306 ,September (2011).

[3] Chen, C.; Kim, J.K., "Optimization for intelligent operation of supply chains", Chemical Engineering Research and Design, Volume 85, Issue 12, pp: 1611–1629, (2007).

[4] Chopra, S.; Meindl, P., "Supply Chain Management Strategy, Planning, and Operation", Prentice Hall, 5th edition, ISBN-13: 978-0132743952, February (2012).

[5] William, E.; Hoover, J.R.; Eloranta, E.; Holmstrom, J.; Huhunen, K., " Managing the Demand-Supply Chain: Value Innovations for Customer Satisfaction: Value Innovations for Supplier Excellence" , John Wiley & Sons , 1st edition , ISBN-13: 978-0471384991, (2001).

[6] Golmohammadi, D.; Creese, R.C.; Valian, H.; Kolassa, J., "supplier selection based on a neural network model using genetic algorithm", IEEE Transactions on Neural Networks, Volume 20, Issue 9, (2009).

[7] Lee, E.-K.; Ha, S.; Kim, S.-K., "supplier selection and management system considering relation-ships in supply chain management", IEEE Transactions on Engineering Management, Volume 48, Issue 3, (2001).

[8] Sarkis, J.; Dhavale, D. G., "supplier selection for sustainable operations: A Triple-Bottom-Line approach using a Bayesian framework ", International Journal of Production Economics, Volume 166, pp: 177–191, August (2015).

[9] Lima Junior, F. R., et al,"A comparison between fuzzy AHP and fuzzy TOPSIS methods to supplier selection", Applied Soft Computing Journal, Volume 21, pp: 194–209, August (2014).

[10] Daneshvar Rouyendegh, B.; Eko Saputro, T., "supplier selection using integrated fuzzy TOPSIS and MCGP: a case study", 5th world conference on educational science WCES, Volume 116, pp: 3957-3970, February (2014).

[11] Liu, Y. N., "A case study of evaluating supplier's selection criteria in a steel bars manufacture ", IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp: 994 - 998 (2010).

[12] Kar, A.K., "A hybrid Group decision support system for supplier selection using analytic hierarchy process, fuzzy set theory and neural network", Journal of Computational Science, Volume 6, pp: 23–33, January (2015).

[13] Chai, .J; Liu, J.N.K; Ngi, E.W.T,"Application of decision making techniques in supplier selection: A systematic review of literature", Expert Systems with Applications Journal, Volume 40, Issue 10, pp: 3872–3885, August (2013).

[14] Moghaddam, K.S, "Fuzzy multi-objective model for supplier selection and over allocation in reverse logistics systems under supply and demand uncertainty", Expert Systems with Applications ,Volume 42, Issues 15–16, pp: 6237–6254 ,September (2015).

[15] Ajhan, M.B; Kilic, H.S, "A two stage approach for supplier selection problem in multi-item, multi supplier environment

with quantity discounts", Computers & Industrial Engineering, Volume 85, pp: 1–12 July (2015).

[16] Jadidi, O.; Cavalieri, S.; Zolfaghari, S., "An Improved Multi-choice god programming approach for supplier selection problem", Applied Mathematical Modelling Journal, Volume 39, Issue 14, pp: 4213–4222, July (2015).

[17] Yu, C.; wong, T.N.,"An agent-based negotiation model for supplier selection of multiple products with synergy effect", Expert Systems with Journal, Volume 42, Issue 1, pp: 223–237, January (2015).

[18] Ertugrul Karsak, E.; Dursun, M., "An integrated supplier selection methodology incorporating QFD and DEA with imprecise data", Expert Systems with Applications Journal,Volume 41, Issue 16, pp: 6995–7004 ,November (2014).

[19] Igoulalene, I.; Benyoucef, L.; Tiavari, M.K., "Novel fuzzy hybrid multi-criteria group decision making approaches for the strategic supplier selection problem" , Expert Systems with Applications Journal, Volume 42, Issue 7, pp; 3342–3356, May (2015).

[20] Lirong, J.; Huanhuan, Z.; Jong, L.; Jianjun, Z., "Civil aircraft suppliers selection based on grey target and grey cluster decision method " , 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), PP: 1691-1696, Oct (2014).

[21] Chai, J.; Ngai, E.W.T., "Multi-perspective strategic supplier selection in uncertain environments", International Journal of Production Economics, Volume 166, pp: 215–225, August (2015).

[22] Dargi, A.; Anjomshoae, A.; Rahiminezhad Galankashi M.; Memari, A.; Md.Tap, M. B., "Supplier selection: A Fuzzy-ANP approach", 2nd International Conference on Information Technology and Quantitative Management, Volume 31, pp: 691-700, (2014).

[23] Nazari-Shirkouhi; Shakouri, S.; Javadi, H.; Keramati, B., "A supplier selection and other allocation problem using a two-phase fuzzy multi-objective liner programming", Applied Mathematical Modelling, Volume 37, Issue 22, pp: 9308–9323, November (2013).

[24] Arilcan, F., "A fuzzy solution approach for multi objective supplier selection", FUZZYSS11: 2nd International Fuzzy Systems Symposium, Volume 40, Issue 3, pp: 947–952, February (2013).

[25] Kar, A.K., "Revisiting the supplier selection problem: An integrated approach for group decision support", Expert Systems with Applications, Volume 41, Issue 6, pp: 2762–2771, May (2014).

[26] Mehralian, Gh.; Rajabzadeh Gatari, A.; Morakabati, M., "Developing a suitable model for supplier selection based on supply chain risks : An empirical study from Iranian pharmaceutical" ,Iranian Journal of pharmaceutical research, Volume 11, Issue 1 , (2012).

[27] Jazemi, R.; Ghodsypour, S.H.; Gheidar-Kheljani, J., "Considering supply chain benefit in supplier selection problem by using information sharing benefits", IEEE transactions on industrial information, Volume 7, Issue 3, (2011).

[28] Wu, Sh., "A new method of selecting supplier", International joint conference on artificial intelligence, pp: 370-372, April (2009).

[29] Fraggi , Ch.; Hsaiao, W.; Feifan, Y., "A multi-criteria interval optimization model for manufacturing supplier selection using genetic algorithm", first international workshop on education technology and computer science, Volume 03, pp: 757-761, (2009).

[30] Guoshang, H.; Guohang, Z. "comparison on neural networks and support vector machines in supplier selection", Journal of systems engineering and electronics, Volume 19, Issue 2, pp: 316-320, (2008).

[31] Kong, F.; Zahng, Z.; liu, J., "Selection of suppliers based on fuzzy multi-criteria decision making", Fifth international conference on fuzzy systems and knowledge discovery, Volume 5, pp: 198-202, Oct (2008).

[32] Saaty, T. L., "Decision making with the analytic hierarchy process", Int. Journal Services Sciences, Volume 1, Issue 1, 2008.

[33] S.H Ghodsypour," analytical hierarchy process ", Amirkabir University Publication, ISBN 978-964-463-056-9, (In Persian, 2013).

[34] Tan, P.-N.; Steinbach, M.; Kumour, V., "Introduction To Data Mining", Pearson, 1ST edition, May (2005).

[35] Tsoulfas, G.T.; Pappis, C. P., "A model for supply chains environmental performance analysis and decision making", Journal of Cleaner Production, Volume 16, Issue 15, pp: 1647–1657, October (2008).