B2B Supply Chain Performance Enhancement Road Map Using Data Mining Techniques

PONGSAK HOLIMCHAYACHOTIKUL[1]
College of Arts, Media and Technology
Chiang Mai University, Chiang Mai, 50200, THAILAND
Email: holimchayachotikul@hotmail.com[1]

RIDHA DERROUICHE[2]
LIESP Université de Lyon (Lumière Lyon 2)
160 Bd de l'Université 69676 BRON cedex, FRANCE
Email: ridha.derrouiche@gmail.com

KOMGRIT LEKSAKUL[3]
Department of Industrial Engineering
Faculty of Engineering Chiang Mai University, Chiang Mai 50200, THAILAND
Email: komgrit@eng.cmu.ac.th.[3]

GUIDO GUIZZI[4]
Department of Materials Engineering and Operations Management
University of Naples "Federico II"
P.le Tecchio – 80125 Napoli ITALY
Email g.guizzi@unina.it[4]

Abstract: - Presently, modern B2B supply chain management (B2B-SCM) equipped with semi-automated data logging systems accumulate large volumes. However, each SC unit in B2B-SC still individually develops their performance. Besides, their linkages of performance attributes between SC units still lack vital information extraction to improve theirs. Therefore, this paper aims to propose an integrated framework between B2B supply chain (B2B-SC) performance evaluation systems and data mining techniques, developing relationship rules of collaborative performance attribute enhancement. The methodology is as follows. Firstly, B2B-SC performance evaluation questionnaires based on two levels able to characterize collaborative relation between two or more partners in their SC were gathered from the case study companies. The data set of relationships between enterprise and its direct customers of the case study companies in France was used for demonstration. Secondly, data cleaning and preparations for rule extraction were performed on the questionnaire database. The significance of attribute was calculated using attribute ranking algorithms by means of information gain based on ranker search. These results were used to choose the crucial attributes from each micro view. Thirdly, web graph analysis was performed on this data to confirm the strong attribute relationship. Next, association rule was deployed to extract performance attribute relationship rules grounded on support and confidence cross validation method. The quality of each recognized rule is tested and, from numerous rules, only those that are statistically very strong and contain vital information are selected. Last but not least, these rules are interpreted by domain experts and studied by domain engineers to build a collaborative performance attribute enhancement road map. Furthermore, the final rule set of extracted rules contains very interesting information relating to SCs and also point out the critical existing SC attribute improvement. Ultimately, companies in this SC are able to use this framework to design and adjust their units to conform with the exact customer needs.

Key-Words: - Business to Business (B2B), supply chain (SC), Association Rule, Attribute ranking, Data Mining

1 Introduction

Presently supply chain management (SCM) is a crucial constituent of business strategy and operation management according to customers who extremely seek high-quality and low-cost products. The achievement of many organizations count on many aspects; one of them being the performance measurement (PM) system since it can establish understanding, improve competitiveness and mold approach [5]. Over the years, companies have increasingly adopted SCM, which triggers the necessity of the PM of SC. Nonetheless, the current traditional measurement methods have become less appropriate for SCM due to the fact that their scopes are too narrow down to handle the broad extent of measurement activities. The dramatic increase in publications of SCM theories and practices has led to evolution of the SCM itself. However, not much concentration has been devoted to the field of SC performance measurement system [3, 4, 6, 7, 8, 12].

Many researchers have focused on PM framework designs as to exploit these new frameworks in action rather than how to sustain or increase SC performance in the long term, or how to improve and validate the present PM results. Some of the proposed PM methods use either qualitative or quantitative indications, thus the PM assessment might not cover certain other aspects. Most of general PM manners in SC solely offer a guideline on how to manage and control performance based on the previous reported outputs. The general objective of their control mechanism is to analyze and accumulate sequential performance information following both measurement and comparison of the desired performance to indicate performance adjustment activities. Besides, everything depends on things that have already happened. Conversely, none of these means comprise of particular steps for a qualitative reverse-reasoning indicating result which addresses the question of “Is each performance attribute result related with other attribute results?” and “Should we promote, adjust, or leave it with the other suitable attributes in terms of guided performance improvement rule ?’’

On the business side, most of the present performance development in Business-to-Business (B2B-SC) collaboration consumes much more time and cost on the grounds that each SC unit in B2B-SC database still individually develops their performance. Besides, their linkage of leading and lagging indicators between SC units still lack information extraction to improve their collaboration even though they share some SC information based on business contract and context. Consequently, each SC unit does not have the appropriate collaborative PM roadmap development together. This leads to the one question which is “How can we extract the interesting association rules for collaborative PM roadmap enhancement from the relationship between leading and lagging attributes among SC units?”

In this paper, a novel alternative by integrated (B2B-SC) performance evaluation system together with data mining technique were developed to provide the performance attribute association rules for SC performance improvement. This paper is organized as follows. In section 2, the literature review is briefly described. In section 3, association rule algorithm is shortly explained. In section 4, the methodology of the proposed approach is shown. Results and discussion of the proposed framework are shown in section 5. Finally, conclusions are provided in section 6.

2 Literature Review

A reliable PM is beneficial in evaluating the SCM effectiveness and efficiency. Supposing we have a good performance measurement; we can profoundly understand the current performance status of the SC network easily so as to effectively recognize our strengths, weaknesses, threats, and opportunities. However, it has been a challenge to establish an “appropriate” collaborative network for the SC network. Kittelson and Associates, 2003 [2] point out that PM, among collaborative SC network, which is crucial for management. There have been many attempts to apply and explore AI and data mining techniques to make up for the typical techniques in optimizing the PM in SCM with a better development roadmap. M. F. Joanna et.al 2007[10] have rendered that data mining is a tool for efficiency of supply chain integration to analyze many quality perspectives. Mostly, many publications have concentrated on the overall relationship through the chain; these are slightly focused on the linkage relationship in each link in the chain. These result in the lack of the long term relationship in terms of trust and commitment to push up their performance together. Recently, a combination of fuzzy theory has been extensively employed to handle and optimize supply chain configurations in many aspects. For example, Bevilacqua et.al 2006[11] has employed A fuzzy-QFD approach to supplier selection and Jain et.al 2007 [16] has then proposed supplier selection using fuzzy association rules mining approach. On the process view, H.C.W. Lau et al., 2009 [9] has developed a process mining system for supporting knowledge discovery in a supply chain network using fuzzy association rules to fine tune the configuration of process parameters. While Huang et al., 2008 [13] has developed a fuzzy neural network optimized by particle swarm optimization to model solve the problem of demand uncertainty in SC. Moreover, the number of publications in operational optimization in terms of scheduling, routing, and inventory using Genetic Algorithms (GA) have increased according to its performance. In addition, Almejalli, K et.al 2008 [1] has applied fuzzy neural network and GA for real time identification of road traffic control measures. The more complicated both methodology and tools used in the PM systems; the more complex implementation consumed the time and acquired the well results. Furthermore, It seems to be better way for the academic research, but it might not be suitable for the SC domain managers from the various fields. To simplify, modify and develop tools and effective frameworks, related with the general knowledge of SC domain managers is the appropriated way for academic and business. At this point, we found the gap of the integration between PM model, focused on B2B-SCM and SC performance rule development based on data mining techniques. They should be able to develop going hand in hand to establish the collaborative performance attribute enhancement rules for B2B-SCM using the techniques from the data mining combined with PM model.

3 Association Rule

One of eminent data mining is very powerful to mine the hidden vital information in a mountain of database by discovering frequent itemsets and only examining rules that are made up of frequent itemsets [15] ; It is called association rules. Apriori -association rule induction acquires interesting correlation relationships or/and associations among large sets of data items. Apriori handles the items and itemsets that make up transactions. Items are flag-type conditions that indicate the presence or absence of a particular thing in a specific transaction. An itemset is a set of items which may or may not incline to co-occur within transactions. Most of association rules are employed in many domains such as sales promotion planning; market basket analysis and production attribute relationship etc.

Processes were divided in two main tasks. Firstly, it indicates frequent itemsets in the interested data, and then it produces rules from the frequent itemsets table. To choose attractive rules from the result of Apriori process, the finest-known triggers are minimum thresholds on confidence and support. The former is the support $\mathrm{supp}(M)$ of an itemset, M is defined as the ratio of transactions in the data set, contained the itemset. For instance, the itemset {yoghurt,cookies} has a support of 2/5= 0.4 because it takes place in 40% of all transactions (2 out of 5 transactions). The latter is the confidence of a rule. It is defined conf $(M \rightarrow N)$ $= \mathrm{supp}(M \cup N) / \mathrm{supp}(M)$ For instance, the rule $\{yoghurt, cookie\} \rightarrow \{icecream\}$ has a confidence of 0.3/0.4 = 0.75 in the record set, which denotes that for 75% of the transactions, included yoghurt and cookies the rule is right. Confidence can be interpreted as an estimate of the probability P(N|M). Moreover, the result can be interpreted following this format : If Antecedent is/are are happed together with support percentage of the probability, then Consequence is/are happen with confidence percentage of the probability.

4 Methodology

The methodology of the Apriori association rule induction for B2B-SC performance improvement road map development is shown in Figure. 1. The data set of relationship between enterprise and its direct customers, from R. Derrouiche et al., 2008 PM model and questionnaire [14] , was used for demonstration. Before association rule performing, the attribute selection was used to select the essential attributes which can give the information and much more detail of relationship among attribute.

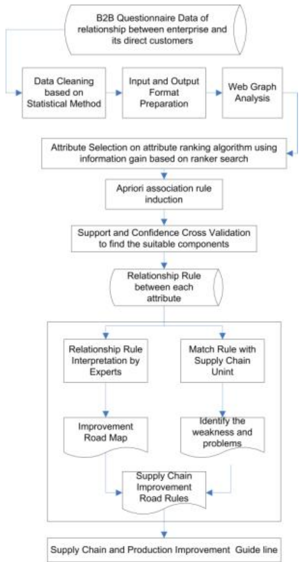

Figure. 1. The methodology of Apriori -association rule induction for B2B supply chain performance improvement road map development

In this research, the attribute selection method based on attribute ranking algorithm using information gain based on ranker search was performed for the two types of relationships. Since many of the data fields are categorical, a web graph was performed to map associations between different categories. Last but not least,, the Apriori association rule induction was performed based on the two datasets. One was the best relationship and the other was the bad relationship data set to develop the supply chain performance improvement road map and identify the top weaknesses and problems between enterprise and its direct customers. The support and confidence cross validation is performed to find the suitable number of rules with the domain experts. Finally, the results were interpreted by the domain expert to develop the collaborative performance attribute enhancement rules.

5 Result and Discussion

After The data set of relationship between enterprise and its direct customer questionnaire gathering following R. Derrouiche et al., 2008 's Framework[14] in Figure. 2. to analyze a dyadic relation and to evaluate its performance, the attribute ranking algorithm using information gain based on ranker search was calculated for the two types of relationships. In addition, R.

Derrouiche et al., 2008's questionnaire[14] was able to characterize collaborative relation between two or more partners in a supply chain, evaluating their related performances. The former level is the common perspective as follows: relation climate, relation structure, IT-used and relation lifecycle and the later level consists of the perceived satisfaction of the relation and its perceived effectiveness. These represent the macro view of model. For example, macro view of relation climate has the main six micro views and each micro view has also at two sub-micro views.

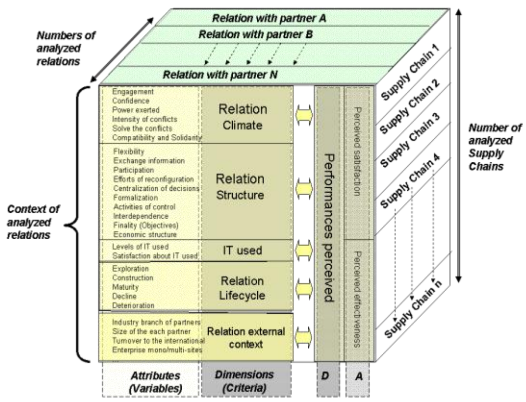

Figure. 2. Framework to analysis a dyadic relation and to evaluate its performan [14]

The PM framework of R. Derrouiche et al., 2008[14] has many sub-micro view attributes; therefore , the vital sub-micro attributes have to be selected for the Apriori association rule induction process. In this research, the sub-micro attributes which have an information gain proportion greater than 70 percentages of the impact ranking or the top of the 20 percentages of the ranking in their micro view, were chosen as the input of Apriori association rule induction process. The results of the attribute ranking are shown in Figure. 3. It is found that the top sub-micro views, impact on the two relationship types are centralization of decision, perceived effectiveness and Finance, control and information exchange. For instance, the one of three of sub-micro of engagement micro view, in relation climate view because this attribute pass the top of the 20 percentages of the ranking in their micro view.

Since many of the data fields are categorical, you can also try plotting a web graph which maps associations between different categories. Next, web graph analysis, as in Figure. 4, of each relationship type were drawn based on the frequency of the significant attribute; these results are used to confirm the strong relationship among each attribute level.

Finally, the association rules were performed based on the two datasets. One was the best relationship and the other was the bad relationship data set to develop the supply chain performance improvement road map and identify the top weakness and problems between enterprise and its direct customers.

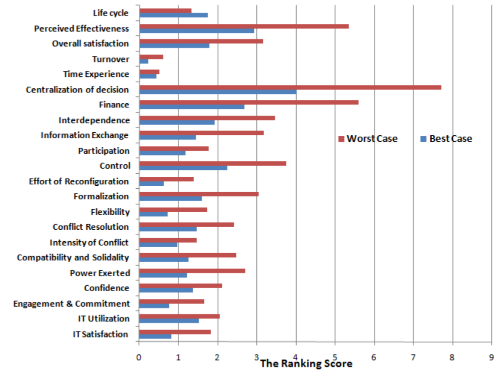

Figure.3. The aggregated sub-micro results of the attribute ranking algorithm using information gain ,based on ranker search for macro view

From support and confidence cross validation in this research, we found that the suitable value support and confidence was 40 percentages and 80 percentages, respectively. As a result, the association model provided that the interested rules have merely extracted following 40 percentages of minimum support and 80 percentages of minimum confidence structure component of Apriori -association rule induction

We noticed that most of the best relation type merely started in the maturity phase of product life cycle. Between the two partners, they take effort to develop the cooperation climate in the long-term orientation with a very high participation. Besides, It can improve satisfaction with their partner.[Rule 23,24,41-43] It is seemly founded on the strong commitment and high confidence degree. The more the partners make a loyal, trust and honest relationship to collaborate to sustain their strong relationship; the more they know the profound information of each other as well. This relationship shows that both partners take much more effort and strategies to balance their power and authority on their collaborations [Rule 6]

At this point, we found some conflicts from the extended relation based on their commercial information; nevertheless, these conflicts could be rapidly resolved using the right information when they flexibly managed their supply chain transactions through their smart technology information system.

From the performance evaluation database, we can extract the performance perspective rules such as if utili1cp = B and plm_cp = C and engag2cp = A are happed together with a 41.176 percentage of probability, then utili3cp = B is also happed with a 85.714 percentage of probability. To sum up, we can interpret that when both enterprise and its direct customers apply

IT used with RCAP in the maturity phase of the product life cycle management, they can work faster in the long-term.

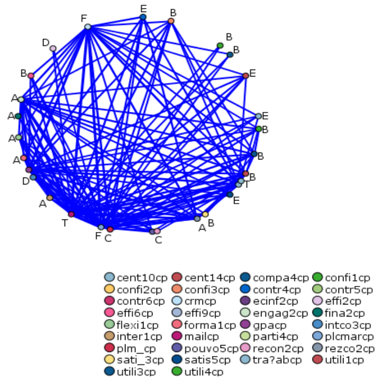

Figure. 4. web graph analysis of best relationship type.

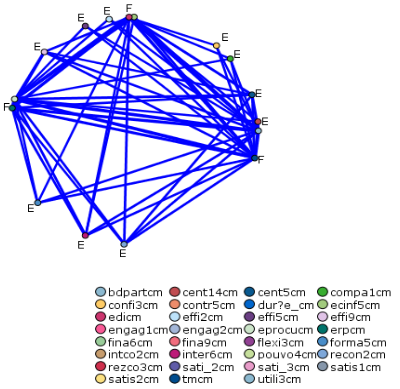

Figure. 5. web graph analysis of bad relationship type

As a result of association rule in Table 1, we can draw the supply chain performance improvement road map between enterprise and its direct customers as following

1. The great trust is the important role to establish and sustain their relationship such as performance report providing, information sharing and staff train exchange. [Rule13,14,36]

2. IT used is crucial to the drive of the extended relationship because it makes fast responsiveness to deliver the right data and information in business transactions.[Rule1,2,4,5,9,10,16,19,20,22,26,28,31,32 ,34,35,37,38,39,40,42]

3. The convenience is suitable way to reduce their partner's opposition in the maturity phase of the product life cycle. Moreover, both enterprise and its direct customers are enable to adjust their performance to reach a common benefit easily[Rule 3,6,7,27]

4. IT used with RCAP is linkage between their common goal and information flow for long-term commitment collaboration. [Rule 8,11,12,30,33]

5. Trust breaking with partners has to be avoid or resolved quickly. It directly impacts other invoked partner's collaborative degree. [Rule 15]

Table 1 The result of association rules based on best relationship

Antecedent	Consequent	Support %	Confidence %
utili4cp = B	utili3cp = B	41.176	100.000
utili4cp = B	engag2cp = A	41.176	100.000
contr4cp = E	rezco2cp = A	47.059	100.000
utili4cp = B and utili3cp = B	engag2cp = A	41.176	100.000
utili4cp = B and engag2cp = A	utili3cp = B	41.176	100.000
contr4cp = E and plm_cp = C	rezco2cp = A	41.176	100.000
plm_cp = C and rezco2cp = A	contr4cp = E	41.176	100.000
utili1cp = B and plm_cp = C	engag2cp = A	41.176	100.000
utili1cp = B	utili3cp = B	52.941	88.889
utili3cp = B	utili1cp = B	52.941	88.889
utili1cp = B	engag2cp = A	52.941	88.889
utili3cp = B	engag2cp = A	52.941	88.889
contr4cp = E	plm_cp = C	47.059	87.500
confi3cp = B	plm_cp = C	47.059	87.500
inter1cp = A	engag2cp = A	47.059	87.500
utili3cp = B and engag2cp = A	utili4cp = B	47.059	87.500
contr4cp = E and rezco2cp = A	plm_cp = C	47.059	87.500
utili1cp = B and utili3cp = B	engag2cp = A	47.059	87.500
utili1cp = B and engag2cp = A	utili3cp = B	47.059	87.500
utili3cp = B and engag2cp = A	utili1cp = B	47.059	87.500
utili1cp = B and engag2cp = A	plm_cp = C	47.059	87.500
plm_cp = C and engag2cp = A	utili1cp = B	47.059	87.500
cent14cp = E	plm_cp = C	41.176	85.714
cent14cp = E	engag2cp = A	41.176	85.714
effi6cp = B	plm_cp = C	41.176	85.714
utili4cp = B	utili1cp = B	41.176	85.714
pouvo5cp = C	flexi1cp = A	41.176	85.714
pouvo5cp = C	utili1cp = B	41.176	85.714
satis5cp = E	plm_cp = C	41.176	85.714
satis5cp = E	engag2cp = A	41.176	85.714
sati_3cp = B	utili1cp = B	41.176	85.714
sati_3cp = B	utili3cp = B	41.176	85.714
sati_3cp = B	engag2cp = A	41.176	85.714
utili4cp = B and utili3cp = B	utili1cp = B	41.176	85.714
utili4cp = B and engag2cp = A	utili1cp = B	41.176	85.714
confi3cp = B and plm_cp = C	engag2cp = A	41.176	85.714
utili1cp = B and plm_cp = C	utili3cp = B	41.176	85.714
rezco2cp = A and engag2cp = A	utili1cp = B	41.176	85.714
utili4cp = B and utili3cp = B and engag2cp = A	utili1cp = B	41.176	85.714
utili1cp = B and utili3cp = B and engag2cp = A	utili4cp = B	41.176	85.714
utili1cp = B and utili3cp = B and engag2cp = A	plm_cp = C	41.176	85.714
utili1cp = B and plm_cp = C and engag2cp = A	utili3cp = B	41.176	85.714
plm_cp = C	engag2cp = A	58.824	80.000

On the other hand, Apriori-association rule induction from the result of association rules based on worst relationship is to identify the significant problems in the relationship between enterprise and its direct customers. Thus, we can draw the supply chain performance improvement road map as follows

1. The stock planning problem in terms of financial perspective mostly comes from the lack of partner's convenience, ability to deliver, information sharing, data accuracy and product life cycle planning. [Rule 1,2,3,5,7,12,16,17]

2. The major consequences on the activity of partners come from misunderstanding and fuzzy agreement in their contracts. It leads to the recessed real-time delivery performance [Rule 4,11,22]

3. The stock planning performance depends on the consistency of partner's operations. [Rule 6,8,9]

4. The less fast-track changing and information sharing, the less trust and satisfaction to operate the collaborative transactions [Rule 13,15,18]

5. The lack of data accuracy and planning between units can generate the problem of the trust degrees in the long-term.[Rule 19,20,21].

6 Conclusion

This paper has described the integrated application between B2B supply chain (B2B-SC) performance evaluation systems and data mining techniques, developing relationship rules of the collaborative performance attribute enhancement road map. After the proposed methodology implementation, the result is extended to the B2B-SC improvement road map guidelines to promote the right direction for performance attribute improvement, to resolve the conflict among the attributes and eliminate the weakness problem issues. It is useful and easy for the B2B-SC invokers to implement because the results are shown in the simplified graphical diagram and the vital rules. This work is constructed by a qualitative and very subjective approach on how managers consider the role of each attribute and its impact on the performance. There are many concerning points from the association rule process such as the key success factor of this process relies on the quality of the database, data preparation, and data cleaning. The fuzzy theory can be applied to handle the subjective attributes before the association rule process; it might increase accuracy input values related to the real B2B-SC environment. To continue this work, it leads to four main directions. The first is we are now extending the present PM including quantitative attributes.

Acknowledgments:
The financial is supported by Erasmus Mundus, External Cooperative Window Project (**eLINK** ***e***ast-***w***est ***L***ink for ***I***nnovation, ***N***etworking and ***K***nowledge exchange) and College of Arts, Media and Technology Chiang Mai University, THAILAND

References:

[1] K. Almejalli, Dahal, K., Hossain, M.A. , Real time identification of road traffic control measures . *Advances in Computational Intelligence in Transport, Logistics, and Supply Chain Management*. vol. 144: Springer Berlin / Heidelberg, 2008.

[2] K. a. Associates, A Guidebook for Developing a Transit Performance-measurement System. Washington, DC., 2003.

[3] B. M. Beamon, Measuring supply chain performance, *International Journal of Operations and Production Management*, vol. 19, pp. 275-292, 1999.

[4] F. T. S. Chan and H. J. Qi, An innovative performance measurement method for supply chain management, *Supply Chain Management*, vol. 8, pp. 209-223, 2003.

[5] S. E. Fawcett and M. B. Cooper, Logistics Performance Measurement and Customer Success, *Industrial Marketing Management*, vol. 27, pp. 341-357, 1998.

[6] P. Folan and J. Browne, A review of performance measurement: Towards performance management,*Computers in Industry*, vol. 56, pp. 663-680, 2005.

[7] A. Gunasekaran, C. Patel, and R. E. McGaughey, A framework for supply chain performance measurement, *International Journal of Production Economics*, vol. 87, pp. 333-347, 2004.

[8] A. Gunasekaran, C. Patel, and E. Tirtiroglu, Performance measures and metrics in a supply chain environment, *International Journal of Operations and Production Management*, vol. 21, pp. 71-87, 2001.

[9] H.C.W. Lau, Y. Zhao, N.S.H. Chung, Development of a process mining system for supporting knowledge discovery in a supply chain network, *International Journal of Production Economics*, 2009.

[10] M. F. Joanna Oleskow, Paulina Golinska and Katarzyna Maruszewska, Data Mining as a Suitable Tool for Efficient Supply Chain Integration Springer Berlin Heidelberg, 2007.

[11] M. Bevilacquaa, A fuzzy-QFD approach to supplier selection, *Journal of Purchasing and Supply Management* vol. 12, pp. 14-27, 2006.

[12] J. H. Park, J. K. Lee, and J. S. Yoo, A framework for designing the balanced supply chain scorecard, *European Journal of Information Systems*, vol. 14, pp. 335-346, 2005.

[13] Y. Huang Z. Qiu and Q.Liu, chain network design based on fuzzy neural network and PSO, *IEEE International Conference on Automation and Logistics (ICAL 2008) Qingdao, China*, 2008.

[14] R. Derrouiche, A. Bouras, G. Neubert, M. Savino, A dashboard to analyze relations in enterprise networks, *International Conference on Advances in Production Management Systems (APMS'08) Finland*, 2009.

[15] R. Agrawal and R. Srikant. Fast Algorithms for Mining Association Rules in Large Databases, *Proceedings of the 20th International Conference on Very Large Data Bases*, p.487-499, September 12-15, 1994

[16] V. Jain, S. Wadhwa, and S. G. Deshmukh Supplier selection using fuzzy association rules mining approach, *International Journal of Production Research*, vol. 45, pp. 1323 – 1353, 2007.