A STUDY OF MINIMUM SUPPORT VALUE IN ASSOCIATION ANALYSIS-FROM BAYESIAN VIEWPOINT

Yun-Shiow Chen* and Min-Hsun Kuo

Dept. of Industrial Engineering & Management,

Yuan-Ze University, 135 Yuan-Tung Road, Chung-Li 32026, Taiwan, R.O.C.

**** ieyschen@saturn.yzu.edu.tw***

ABSTRACT

Association analysis is used to discover the knowledge and set up rules from a large dataset. The minimum support value in the association analysis is a critical factor to affect the performance of the discovery. Most of the studies in this topic displayed the minimum support value are decided subjectively by the researchers. In this study, we assume that comments for the minimum support value from experts and senior managers follow a priori Beta probability distribution. The Taguchi quality loss function consisting of support value, association rules and noise random variable is established as our loss cost model. A Bayesian estimator of the minimum support value is derived by minimizing the expected loss cost. Statistical properties of the Bayesian estimator are investigated by Monte Carlo simulation. In the model verification, we apply our model into the inventory management of computer assembly and the loss cost function considered includes the experts' experience and knowledge, minimum support value, the space for inventory and the retrenching time for the processing priority of key components. In Bayesian analysis, we derive the Bayesian estimator, as a proper minimum support value and mining the related association rules. We discover that we can modify the experts' opinions by sample analysis.

Key Words: Management Information Systems, Minimum Support Value, Association Rules, Bayesian Estimator, Monte Carlo Simulation.

1. INTRODUCTION

Association analysis, also named market-basket analysis, describes the co-occurrence among data items in a large amount of provided dataset. In the association analysis, the minimum support value directly affects the usefulness of the mining out rules in the association analysis. A larger minimum support value is considered, the less important information is extracted, vice verso, a smaller minimum support prefers, a plenty useless rules are shown. We will misjudge our decision if there are too many noise or useless rules appear. Thus, the determination of threshold value for a minimum support is getting important in mining association rules. Unfortunately, most of the researches focused on developing efficient algorithms in fastening the searching speedy but put little attention on the determination of the threshold value of the supportability.

Bayesian method is different from the traditional statistics in estimating parameter, in this

paper, we imbed the experts' and managers' knowledge into the Bayesian method and set up a prior probability distribution of the minimum support value with which is constituted by all possible minimum support values estimated by experts or experienced managers. We assume that the number of products' types is a factor to affect the decision of the proper values for the minimum support. We combine the prior probability distribution and the working experience to build up the posterior distribution of the minimum support value. We set up a square loss function of support value based on the cost generated from the association rules. From the above set-up, we can derive the Bayesian estimator of the minimum support value by minimizing the posterior risk. The study is organized as follows: in Section 2, literature review of the previous work on association analysis and Bayesian analysis. In Section 3, we introduce our Bayesian model. A numerical example is presented in Section 4. Summary of the work is presented section 5.

2. LITERATURE REVIEW

2.1 Association Analysis

Association analysis is a technology to describe the relationship of data items in a huge collected data. We can find out the focal items and the level of the relationship among these items. Association analysis has been used in various fields like banking, retail business, manufacturing, information business, and so on. Recently, Borchi al et. (2001) applied association analysis to mine out the user's path habit in surfing the web. Lee al et. (2001) improved the query design more effectively by mining out customer's rules that beforehand, and then designing products to satisfy customers' needs and preference closely. Agrawal et al. (1994) proposed "Apriori algorithm" for which is the foundation work of mining rules from the database. Generating frequent itemsets through iterations from the database and generating strong association rules from the frequent itemsets satisfying minimum support and minimum confidence are two main steps in the Apriori algorithm. Long searching time is one major drawback of the Apriori algorithm since for any one of the iteration requires a pass of scanning the database once. In order to reduce the number of scanning of the database in the Apriori algorithm, Chen al et. (2002), Holt al et.(2002) and Yilmaz al et. (2003) proposed alternatives algorithms for mining association rules. Brin al et. (2003), Hong al et. (1999), Hu al et. (2003a) and Rastogi al et. (2001) set up rules with different attribute data like discrete data, continuous data, nominal scale data, and so on. To effectively manage, Goulbourne al et. (2000), Hu al et. (2003b), Li al et. (2002) and Tsay al et. (2004) focused on classifying or clustering the mined out rules and Jan (2000), Liu al et. (1999) and Yun al et. (2003) stressed on dealing from different viewpoints, e.g. taking different minimum support value for each item, or considering uneven weight for each piece. Wang (2001) showed that setting association rules only satisfy the minimum confident value in order to reduce the time being spent on generating a huge of candidate items.

2.2 Bayesian Decision

Recently, knowledge management prevails over the world. In various fields, retaining the experts' knowledge and experience to extract inside knowledge is getting more important. Seidel al et. (2003) and Walley al et. (1996) shows that there are many systems or algorithms generated for learning to keep the experts' and managers' knowledge. The Bayesian method is one popular way to process the experts' knowledge with statistical techiques. The Bayesian statistical model builds up the joint probability function by integrating the information of parameters and the sample in each subjective estimation. Owing to the roughly of the subjective information, we can modify the information by adding the sample information to

increase the accuracy. The previous experience can be applied in the Bayesian method, so we don't need a lot of samples to obtain the valuable inference. The Bayesian method in Hobbs al et. (1997), Hong (1997) and Kamath al et. (2002) were to put on solving the problem of estimating the parameters. Based on its powerful estimating ability, Abdelbar al et. (2002), Baesens al et. (2002) and Lampinen, al et. (2001) joined the Bayesian method and algorithms for two purposes. One is to decide the parameters of algorithms, and the other one is to make the algorithms more powerful.

Previous work put little attention on the determination of the threshold value of the support objectively. Lin al et. (2002), Savinov al et. (2004) and Wang al et. (2003) proposed to set up rule objectively according to the status of the data. In this study, we estimate the minimum support value from cost viewpoint and Bayesian statistics.

3. METHODOLOGY AND APPLICATIONS

In this research, we imbed the experts' and policymakers' knowledge into the association analysis in order to provide more useful and powerful information than the traditional methods do. A suitable minimum support value will make more profit and less loss for the enterprise than the unsuitable ones. Thus, we evaluate the minimum support value from the cost viewpoint.

The Bayesian method is different from the traditional statistics in the definition of parameters. The Bayesian method defines the parameter as a random variable with a probability function, not a constant value. Applying the Bayesian method, therefore, we can put the experts' and managers' opinions into our model. For Bayesian method, we build up the posterior distribution of the parameter by combining the prior probability function from subjective information and sample analysis. Afterward, we integrate the posterior distribution and the loss function to get the Bayesian posterior risk. We can obtain the Bayesian estimator by minimizing the risk, which are the suitable value of the estimator for minimum support of association analysis. The process of the Bayesian model in the association analysis is as follows:

Step1: Set up the prior probability function of the minimum support value

We collect possible values θ_j, which is made from the experts on the most proper minimum support value and the index j means the information of the j^{th} expert or the policymaker. The notation Θ is the set of θ_j, $j = 1,2,\ldots$. The possible value of estimator θ_j is between 0 and 1. The prior probability density function of Θ is denoted as $\pi_\Theta(\bullet)$.

Step 2: Sample analysis

We process *n* independent trials to summarize the number of the products' types. And $\boldsymbol{X}$ is the observed sample vector, also named the random vector of the trials, where $\boldsymbol{X} = (X_1 \quad X_2 \quad \ldots \quad X_n)$. The X_m represents the result from the m^{th} trial, where $1 \le m \le n$. In order to obtain the probability of the $\boldsymbol{X}$ occurring at each possible minimum support value, we proposed the trial is processed in $\Theta = \theta$, the probability function $f_{\boldsymbol{X}}(\boldsymbol{x}, \theta)$ is replaced to be $f_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta)$. It means the conditional probability function of random vector $\boldsymbol{X}$ at given $\Theta = \theta$. We obtain the likelihood function $f_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta)$ based on the following form:

$$f_{X|\Theta}(\boldsymbol{x}|\theta) = f_{X|\Theta}(x_1|\theta) f_{X|\Theta}(x_2|\theta).....f_{X|\Theta}(x_n|\theta) = \prod_{m=1}^{n} f_{X|\Theta}(x_m|\theta) \quad (1)$$

Step 3: Derive the posterior distribution

Integrating steps 1 and 2, we can get the posterior distribution. First, we compute the joint probability distribution $g_{X,\Theta}(\boldsymbol{x},\theta)$, which is defined for $\boldsymbol{\Theta}$ and $\boldsymbol{X}$, the process is as follows:

$$g_{X,\Theta}(\boldsymbol{x},\theta) = f_{X|\boldsymbol{\Theta}}(\boldsymbol{x}|\theta)\pi_{\boldsymbol{\Theta}}(\theta) \qquad \theta \in \Theta, \quad \boldsymbol{x} \in \boldsymbol{X} \quad (2)$$

and we will get the marginal density function of $\boldsymbol{X}$ by integrating the θ of the eqn. (2):

$$g_X(\boldsymbol{x}) = \int_{\Theta} f_{X|\boldsymbol{\Theta}}(\boldsymbol{x}|\theta)\pi_{\boldsymbol{\Theta}}(\theta)d\theta = E_{\pi}(f_{X|\boldsymbol{\Theta}}(\boldsymbol{x}|\boldsymbol{\Theta})) \quad (3)$$

From eqn. (2) and eqn. (3), we can obtain the conditional probability density function of Θ as eqn. (4), under the observation $\boldsymbol{x}$ of $\boldsymbol{X}$.

$$g_{\boldsymbol{\Theta}|X}(\theta|\boldsymbol{x}) = \frac{g_{X,\boldsymbol{\Theta}}(\boldsymbol{x},\theta)}{g_X(\boldsymbol{x})} = \frac{\prod_{m=1}^{n} f_{X|\Theta}(\boldsymbol{x}|\theta)\pi_{\boldsymbol{\Theta}}(\theta)}{\int_{\boldsymbol{\Theta}} \prod_{m=1}^{n} f_{X|\Theta}(\boldsymbol{x}|\theta)\pi_{\boldsymbol{\Theta}}(\theta)d\theta} \quad (4)$$

where $\theta \in \boldsymbol{\Theta}$, $g_X(\boldsymbol{x}) \neq 0$

Step 4: Construct the posterior risk model

The loss function $L(\hat{\boldsymbol{\theta}},\theta)$ is defined for the estimators $\hat{\boldsymbol{\theta}}: \boldsymbol{X} \to \boldsymbol{T}$ of the parameter function $\boldsymbol{C}: \Theta \to \boldsymbol{T}$. We assume that the different losses will be caused based on different support values. The expected loss $B_{\pi|X}(\hat{\boldsymbol{\theta}})$, which is the posterior risk with the posterior distribution $g_{\boldsymbol{\Theta}|X}(\theta|\boldsymbol{x})$ defined on Θ. The expected loss is as follows:

$$B_{\pi|X}(\hat{\boldsymbol{\theta}}) = \int_{\Theta} L(\hat{\boldsymbol{\theta}}(\boldsymbol{x}),\theta) g_{\Theta|X}(\theta|\boldsymbol{x})d\theta \quad (5)$$

The Taguchi quality loss functions is considered as our loss function (figure 1). The perspective point is that the more the estimator closes to the suitable minimum support value, the less the enterprise needs to pay for it, vice versa, when the discrepancy between them is greater, the loss will be larger.

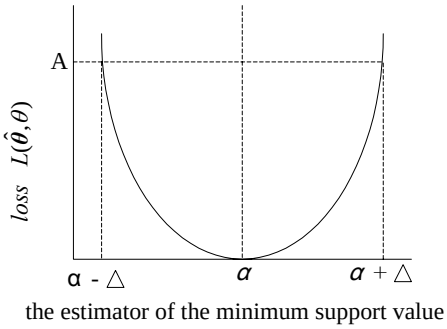


Figure 1 The loss function

In this paper, a squared-error loss function is choiced and showed as eqn. (6).

$$L(\hat{\boldsymbol{\theta}},\theta) = [\hat{\boldsymbol{\theta}}(\boldsymbol{x}) - C(\theta)]^2 \tag{6}$$

And the posterior risk $B_{\pi|X}(\hat{\boldsymbol{\theta}})$ defined for $\hat{\boldsymbol{\theta}}$ will be

$$B_{\pi|X}(\hat{\boldsymbol{\theta}}) = \int_{\Theta} [\hat{\boldsymbol{\theta}}(\boldsymbol{x}) - C(\theta)]^2 g_{\Theta|X}(\theta|\boldsymbol{x}) d\theta \tag{7}$$

The proper value for the minimum support value as the expected loss being least is the value of $\hat{\boldsymbol{\theta}}$ which minimizes the quadratic loss function. After that, we differentiate eqn(7) and get the Bayesian estimator $\hat{\boldsymbol{\theta}}^*(\boldsymbol{x})$ of a suitable minimum support value as follows:

$$\begin{aligned} \hat{\boldsymbol{\theta}}^*(\boldsymbol{x}) &= \int_{\theta} C(\theta) g_{\theta|X}(\theta|\boldsymbol{x}) = E(C(\theta)|\boldsymbol{X} = \boldsymbol{x}) \\ &= \frac{\int_{\Theta} C(\theta) f_{X|\Theta}(\boldsymbol{x}|\theta) \pi_{\Theta}(\theta) d\theta}{\int_{\Theta} f_{X|\Theta}(\boldsymbol{x}|\theta) \pi_{\Theta}(\theta) d\theta} \end{aligned} \tag{8}$$

4. MODEL VERIFICATION

In this section, we apply our model structure into the computer assembly area and try to locate key components most customers preferring. We process these key components in prior to reduce the customers' waiting time, hence, a proper minimum support value can make the process effectively. For the performance of the model, we collect the possible support values decided by experts and analyze the sample to summarize the variety of products. By combining these two parts, we will obtain the posterior distribution of the minimum support value. We adopt BSM algorithm to mine out the association rules at the given possible support value. In this work, our cost includes the space for inventory and the retrenching time for processing the key components. Parameters affected by the association rules are considered in the cost loss function.

Because there are few source to get the real data and relative information to verify our thought, we just set up some fake terms to verify this method.

The simulated data is used to verify our approach. In this study, we consider five different categories including PC board (3 brands), CPU (4 brands), RAM (2 brands), Hard Disk (4 brands) and CD-Rom (3 brands) for the components in a computer.

The time intervals of the customers giving orders are proposed to be exponential distributed and the quantities of the computer in each order follow the poisson distribution. Based on the foregoing, we generated 300 orders including 3082 sets of computers in total.

For these orders, we assume the possible minimum support values decided by the experts follow the beta distribution with parameter *a* and *b.* And the probability density function is shown as follows:

$$\pi(\theta) = \begin{cases} \dfrac{\theta^{a-1}(1-\theta)^{b-1}}{\beta(a,b)} & , \ 0 \le \theta \le 1 \\ 0 & , otherwise \end{cases} \tag{9}$$

and

$$\beta(a,b) = \int_0^1 \theta^{a-1}(1-\theta)^{b-1} d\theta = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{10}$$

We believe the suitable minimum support value will be small when there are more variety of the products, vice versa, the value is large as the products being less variety. Therefore, we adopt geometric trials to see the variety of the products in the studied period. If the variety of products is high then the probability of picking the right products tends to be small, on the contrary, the probability will be high for a less variety. Based on the products from these orders, we repeat geometric trials *m* times and obtain the random vector $\boldsymbol{X}$. The random variable X_j is the number of the needed samples for the first time to pick the right product for the j^{th} trial.

In order to estimate a suitable minimum support value, we observe the probabilities of the sample occurring at each possible support value and get the likelihood function shown as follows:

$$f_{X|\Theta}(\boldsymbol{x}|\theta) = \prod_{m=1}^{n} \theta(1-\theta)^{x_m - 1} = \theta^n (1-\theta)^{\sum_{m=1}^{n} x_m - n} \tag{11}$$

and

$$1 \le x_m < \infty, \ 0 \le \theta \le 1$$

Integrating eqns(9) and (11), we have the joint probability function $g_{X,\Theta}(\boldsymbol{x},\theta)$ defined for Θ and $\boldsymbol{X}$. And the process is described as follows.

$$g_{X,\Theta}(\boldsymbol{x},\theta) = f_{X|\Theta}(\boldsymbol{x}|\theta)\pi_{\Theta}(\theta) = \theta^{n}(1-\theta)^{\sum_{m=1}^{n} x_m - n}\left(\frac{\theta^{a-1}(1-\theta)^{b-1}}{\beta(a,b)}\right)$$
$$= \frac{\theta^{n+a-1}(1-\theta)^{\left(\sum_{m=1}^{n} x_m - n + b - 1\right)}}{\beta(a,b)} \tag{12}$$

and $\theta \in \Theta$, $\boldsymbol{x} \in \boldsymbol{X}$

We will get the marginal density function of $\boldsymbol{X}$ by integrating the θ of the eqn.(12):

$$g_X(\boldsymbol{x}) = \int_0^1 f_{X|\Theta}(\boldsymbol{x}|\theta)\pi_{\Theta}(\theta)d\theta$$
$$= \int_0^1 \frac{\theta^{n+a-1}(1-\theta)^{\left(\sum_{m=1}^{n} x_m - n + b - 1\right)}}{\beta(a,b)}d\theta \tag{13}$$

From eqns (12) and (13), we can obtain the conditional probability density function of Θ as eqn. (14), if we have the observation $\boldsymbol{x}$ of $\boldsymbol{X}$:

$$g_{\Theta|X}(\theta|\boldsymbol{x}) = \frac{g_{X,\Theta}(\boldsymbol{x},\theta)}{g_X(\boldsymbol{x})} = \frac{\frac{\theta^{n+a-1}(1-\theta)^{\left(\sum_{m=1}^{n} x_m - n + b - 1\right)}}{\beta(a,b)}}{\int_0^1 \frac{\theta^{n+a-1}(1-\theta)^{\left(\sum_{m=1}^{n} x_m - n + b - 1\right)}}{\beta(a,b)}d\theta} \tag{14}$$

and, $\theta \in \Theta$, $g_X(\boldsymbol{x}) \neq 0$

For simplicity, here, we only estimate two factors of the cost function. One is the space of the inventory, we assume that inventory cost is the cubic of the number of all items in a rule. The other is the retrenching time as processing the priority items in a rule. We assume that we can save two units of time if there are two components in one rule. And eqn.(6) is used in this model, we get the loss function as eqns.(15).

$$C(\theta) = \sum_{i=1}(k_1(M_i^{\theta})^3 - k_2 M_i^{\theta})Q_i^{\theta} \tag{15}$$

and

i : the index of the mined-out association rules

k_1 : the constant term for estimating the space of inventory

k_2 : the constant term for estimating the retrenching time

M_i^{θ} : the number of items in the i^{th} association rule with support value θ

Q_i^{θ} : the number of the itemsets in the i^{th} association rule with support value θ for all orders.

And the posterior risk of $\hat{\boldsymbol{\theta}}$ is

$$B_{\pi|X}(\hat{\boldsymbol{\theta}}) = \int_{\Theta} (\hat{\boldsymbol{\theta}} - C(\theta))^2 g_{\Theta|X}(\theta|\boldsymbol{x}) d\theta \tag{16}$$

For the simulation purpose, we suppose that parameters in the beta distribution of prior probability distribution of minimum support value are $a = 2$, $b = 3$. Owing to the nonlinear relationship between the support value and its cost, it's difficult to find the solution analytically. Therefore, we adapt the simulation method to approximate a suitable minimum support value. The steps of the simulation process are as follows:

Step 1: Get a possible support value θ from the beta distribution, and generate a set sample by geometric trials.

Step 2: Repeat step 1 for 100 times, we can build the posterior distribution by combining the prior distribution and sample information.

Step 3: Get a possible support value θ from the posterior distribution and estimate the cost caused by θ.

Step 4: Evaluate the mean value of these costs and find the corresponding support value.

Step 5: Repeat step 3 and step 4 for 30 times from those data sets, find the minimum cost and the corresponding support value, as the proper minimum support value.

From the foregoing, an estimated suitable minimum support value 0.2045 is obtained, and the related association rules. The association rules include 15 1-itemset (one item in a rule) and 10 2-itemsets (two items in a rule). We can depict the relationship between the support value and its cost from figure 1 when $k_1 = 1$ and $k_2 = 1$. From fig 2, we discover that there are many mean minimum support values clustered within the region between 0.1725 and 0.1775. From their cost values, we find that the highest drop occurred between 0.1761 (cost value being 1690.8) and 0.1762 (cost value being 1113.6). Also no points fall between 0.1775 and 0.1887, one possible reason is that the optimal minimum support value is determined from the cost. When the cost is less than 1113.6, the minimum support value tends to near 0.1887, and tends to close 0.1775 for the cost larger than 1113.6.

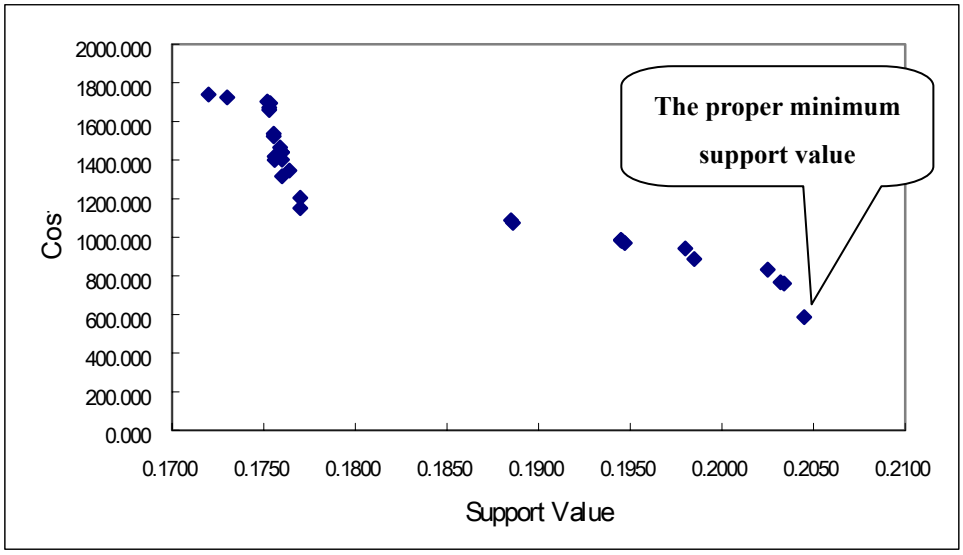


Fig. 2 The relationship between the support value and cost when $k_1 = 1$ and $k_2 = 1$

5. CONCLUSION

In this work, we treat the experts' knowledge and experience and the enterprise's work loss as the prior information in our Bayesian model and then determine a suitable minimum support value. First of all, we assume the beta distribution as the prior probability function for the possible value of the minimum support, which are generated from the experts. Based on the prior probability function, we summarize the products types for each possible support value and compute the conditional probability function. Secondly, we integrate these two parts to build up the posterior distribution. Next, we combine the posterior distribution and the loss function to establish the posterior risk model. Finally, we obtain the Bayesian estimator of the minimum support value by minimizing the posterior risk.

It is difficult to quantitate the customers' service level or satisfactory, thus, we focus on the inventory cost and exclude the profit from saving time after receiving the order. Generally, determining the suitable minimum support value is still difficult for the experts in various fields. We propose a navel approach to process this problem in this study.

ACKNOWLEDGMENTS

Part of this work was supported by the National Science Council of Taiwan under grant No. NSC92-2213-E-155-032.

REFERENCES

Abdelbar, Ashraf, Sahar, M. and Tagliarini, Gene A. (2002), A hybridization of Bayesian and neural learning, *Neurocomputing*, **48**, 443-453.

Agrawal, R. and Srikant, R. (1994). Fast algorithms for mining association rules. *Proceedings of the 20th VLDB Conference*, Santiage-Chile, 478-499.

Baesens, Bart, Viaene, Stijn, Poel, Dirk Van den, Vanthienen, Jan and Dedene, Guido (2002), Bayesian neural network learning for repeat purchase modeling in direct marketing, *European Journal of Operational Research*, **138**, 191-211.

Bonchi, F., Giannotti, F., Gozzi, C., Manco, G. and Nanni, M. (2001), Web log data warehousing and mining for intelligent web caching, *Data and Knowledge Engineering*, **39.2**, 165-189.

Brin, S., Rastogi, R. and Shim, Kyuseok (2003), Mining optimized gain rules for numeric attributes, Knowledge and Data Engineering, *IEEE Transaction* on, **15.2**, 324-338.

Chen, Guoqing, Wei, Qiang, Liu, De and Wets, Geert (2002), Simple association rules(SAR) and the SAR-based rule discovery, *Computer and Industrial Engineering*, **43.4**, 721-733.

Goulbourne, G., Coenen, F. and Leng, P. (2000), Algorithms for computing association rules using a partial-support tree, *Knowledge-Based Systems*, **13**, 141-149.

Hobbs, Benjamin F. (1997), Bayesian methods for analysing climate change and water resource uncertainties, *Journal of Environment Management*, **49.1**, 53-72.

Holt, John D, and Chung, Soon M. (2002), Mining association rules using inverted hashing and pruning, *Information Processing Letters*, **83.4**, 211-220.

Hong, Tzung-Pei, Kuo, Chan-Sheng and Chi, Sheng-Chai (1999), Mining association rules from quantitative data, *Intelligent Data Analysis*, **3.5**, 363-376.

Hong, Xu (1997), Valuation-based systems for decision analysis using belief functions, *Decision Support Systems*, **20.2**, 165-184.

Hu, Yi-Chung, Chen, Ruey-Shun and Tzeng, Gwo-Hshiung (2003), Discovery fuzzy association rules using fuzzy partition methods, *Knowledge-Based Systems*, **16.3**, 137-147.

Hu, Yi-Chung, Chen, Ruey-Shun Chen, and Tzeng, Gwo-Hshiung (2003), Mining fuzzy association rules for classification problems, *Computers and Industrial Engineering*, **43.4**, 735-750.

Jan, Kuang Tai (2000), A study of mining multidimensional association rules based on dynamic minimum supports. M.S. Thesis, Department of Business Management , Soochow University, Taiwan.

Kamath, K. Rajashree, and Pakkala, T.P.M. (2002), A Bayesian approach to a dynamic inventory model under an unknown demand distribution, *Computers & Operations Research*, **29**, 403-422.

Lampinen, Jouko, and Vehtari, Aki (2001), Bayesian approach for neural network—review and case studies, *Neural Networks*, **14**, 257-274.

Lee, C.-H., Kim, Y.-H. and Rhee, P.-K. (2001), Web personalization expert with combining collabrative filtering and association rule mining technique, *Expert Systems with Applications*, **21.3**, 131-137.

Li, Jiuyong, Shen, Hong and Topor, Rodney (2002), Mining the optimal class association rule set, *Knowledge-Based System*, **15**, **7**, 399-405.

Lin, Weiyang, Alvarez, Sergio A. and Ruiz, Carolina (2002), Efficient adaptive-support association rule mining for recommender systems, *Data Mining and Knowledge Discovery*, **6**, 83-105.

Liu., Bing, and Hsu, Wynne (1999). Mining association rules with multiple minimum supports. *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining ACM*, San Diego, CA. 337-340, 1999.

Rastogi, Rajeev, and Shim, Kyuseok (2001), Mining optimized support rules for numeric attributes, *Information Systems*, **26.6**, 425-444.

Savinov, Alexandr (2004), Mining dependence rules by finding largest itemset support quota, *ACM Symposium on Applied computing*, 525-529.

Seidel, M., Breslin, C., Christley, R. M., Gettinby, G., Reid, S. W. J. and Revie, C. W. (2003), Comparing diagnoses from expert systems and human experts, *Agricultural Systems*, **76**, 527-538.

Tsay, Yuh-Jiuan, and Chang-Chien, Ya-Wen (2004), An efficient cluster and decomposition algorithm for mining association rules, *Information Sciences*, **160**, 161-171.

Walley, Peter (1996), Measures of uncertainty in expert systems, *Artificial Intelligence*, **83**, 1-58.

Wang, Ke, He, Yu and Cheung, David W. (2001), Mining confident rules without support

requirement, *ACM*, 89-96.

Wang, Shyue-Liang, Wang, Mei-Hwa, Lin, Wen-Yuen and Hong, Tzung-Pei (2003), Efficient generation of adaptive-support association rules, *IEEE*, 894-899.

Yilmaz, E., Triantaphyllou E., Chen, J. and Liao, T.W. (2003), A heuristic for mining association rules in polynomial time, *Mathematical and Computer Modelling*, **37**, 219-233.

Yun, Hyunyoon, Ha, Danshim, Hwang, Buhyun and Ryu, Keun Ho (2003), Mining association rules on significant rare data using relative support, *The Journal of Systems and Software*, **67**, 181-191.