# Association Rule Mining Using Multi-objective Evolutionary Algorithms: Strengths and Challenges


Rajul Anand
Student
ABV-IIITM, Gwalior
India
rajulanand@iiitm.ac.in

Abhishek Vaid
Student
ABV-IIITM, Gwalior
India
a.vaid@iiitm.ac.in

Pramod Kumar Singh
Associate Professor
ABV-IIITM, Gwalior
India
pksingh@iiitm.ac.in



***Abstract*— Association rule mining based on support and confidence generates a large number of rules. However, post analysis is required to obtain interesting rules as many of the generated rules are useless. We pose mining association rules as multi-objective optimization problem where objective functions are rule interestingness measures and use NSGA-II, a well known multi-objective evolutionary algorithm (MOEA), to solve the problem. We compare our results vis-à-vis results obtained by a traditional rule mining algorithm – *Apriori* and contrary to the other works reported in the literature clearly highlight the quality of obtained rules and challenges while using MOEAs for mining association rules. Though none of the algorithm emerged as clear winner, some of the rules obtained by MOEA could not be obtained by traditional data mining algorithm. We treat the whole process from data mining perspective and discuss the pitfalls responsible for relatively poor performance of the MOEA which has been shown as a good performer in other paradigms.**



***Key words*— Association Rule Mining, Genetic Algorithms, Interestingness measures, Multi-objective Optimization**


## I. Introduction

The amount of data stored in databases continues to grow fast. Intuitively, this large amount of stored data contains valuable hidden knowledge, which may be used to improve the decision-making process of an organization. Thus, there is a clear need for (semi-)automatic methods for extracting knowledge from data. This need has led to the emergence of a field called data mining and knowledge discovery [1]. Association rule mining is one such data mining task which involves frequent pattern mining.

The aim of frequent pattern mining is to search for recurring relationships in a given data set which enables us to discover various kinds of associations and correlations among different items in data sets. Let us formally define the problem: Let I = $\{i_1, i_2, i_3, ..., i_n\}$ be a set of all items. A k-itemset *l* consists of k items from I, is frequent if *l* occurs in a transaction database D not less than α|D| times. Here α is a user specified parameter called *minimum support* defined later and |D| is total number of tuples in database.

*An association rule is an implication of the form* A → B, where $A \in I$, $B \in I$, and $A \cap B = \Phi$. The rule *A* => *B* holds in the transaction set *D* with support α, where α is the percentage of transactions in |D| that contain *A* U *B*. This is taken to be the probability, *P (A U B) which indicates the probability that a transaction contains the union of set A and set B*. The rule *A* → *B* has confidence β in the transaction set *D*, where β is the percentage of transactions in |*D*| containing *A* that also contains *B*. This is taken to be the conditional probability, P (*B*|*A*). Thus we have

$$support\ (A \rightarrow B) = P(A \cup B)$$
$$confidence\ (A \rightarrow B) = P(B|A)$$

Frequent pattern mining involves finding out frequency of occurence of all kinds of itemsets. A huge number of itemsets are possible even for small number of attributes. To prune these large number of itemsets, Agrawal et al. [6] observed a downward closure property, called *Apriori.* It states that a *k-itemset is frequent only if all of its sub-itemsets are frequent*. This property enables us to prune itemsets which do not satisfy α. We iteratively finds k-itemsets until no more itemsets can be generated.

After the *Apriori* was proposed, a number of methods exploiting the Apriori property and extensions to *Apriori* had been developed. Apriori and all related algorithms based on *Apriori* property generate frequent set by first generating candidate sets. Han et al. [11] developed an algorithm called FP-Growth which mines frequent pattern without candidate set generation. But here also, rule generation uses support and confidence.

### *A. Support-confidence Approach: Pros and Cons*

Rule generation based on support-confidence finds out frequent sets which exceed thresholds namely α and β. Frequently occurring items are preferred using support-confidence approach. Decreasing α and β generates huge itemsets increasing time and memory requirements. Thus α and β helps in managing the rule generation process viable and also removing outliers, noise in the dataset.

But using α and β do not always result into interesting rules. For example: for rule *A* → *B*, if the confidence of rule is equal to marginal frequency of itemset B. Then P(*B*/*A*) = P(*B*), meaning A and B are independent. Thus the rule *A* → *B* adds no new information.

There is substantial research done on various interestingness measures for association rules, but integration of these measures during rule generation is minimal. One such initial rule mining was proposed by Agrawal et al. [6].

978-1-4244-5612-3/09/$26.00 ©2009 IEEE

### B. Multi-objective Optimization

When an optimization problem involves more than one objective function, the task of finding one or more optimum solutions is known as *multi-objective optimization*.

The fundamental difference between a single objective and multi-objective optimization task is that the solution in single objective is lone optimum solution, whereas in multi-objective optimization, a number of optimal solutions arise because of trade-offs between conflicting objectives.

### C. Why Genetic Algorithm for Multi-objective optimization

The traditional way to solve multi-objective optimization problems is to follow the preference based approach. In this approach, a preference vector containing relative weight given to each objective is used to create a scalar objective function. Once we have objective function, various single objective optimization algorithms can be used to produce solutions.

Since, single objective optimization algorithms produce one optimal solution. Various runs are required to produce different solutions. Moreover, deciding the weights given to objective functions is an open area which requires sufficient domain knowledge before arriving at decision.

Here Evolutionary Algorithms (EAs) comes to the rescue. Instead of improving single solution during various iterations, that too for a particular preference vector, EAs mimic nature's principles to drive its search towards an optimal solution. Since, in each iteration EAs use a population of solutions, the outcome is a population of solutions. *If an optimization problem has multiple optimal solutions*, an *EA can be used to capture multiple optimal solutions* in its final population.

Since we expect an ideal multi-objective optimization method to give us multiple optimal solutions, EAs seems to be a good alternative.

## II. Association Rule Mining: Multi-objective Optimization Approach

Association rule mining can be considered as multi-objective optimization where support, confidence are optimization objectives. Applying MOEAs (Multi objective Evolutionary Algorithm) will generate rules satisfying above mentioned criteria. However, as discussed in Section, these measures are not good enough to generate so called interesting rules [13], which are generally desired.

Adding new interestingness measures to the rule mining process is normally a post-processing step in KDD methodology. However, the design of MOEAs allows us to introduce new measures in the rule mining process itself. On the other hand,traditional data mining algorithms find frequent sets and then mine rules based on frequent sets. Adding interestingness measures during frequent set mining is not easy because many of the interestingness measures are applicable only to association rules. But in MOEAs, *we directly mine association rules* . Thus, it enables us to integrate various interestingness measures in mining process.

Various interestingness measures can be thought of various objectives of optimization problem, the solution to which is an association rule. The design of MOEAs allows various interestingness measures to be incorporated by simply changing the objective function. Thus we can obtain different association rules based on different interestingness measures.

One major advantage of MOEAs is that rule mining process is guided by interestingness measures . It fulfil our requirement of obtaining rules based on different interestingness measures because practically there is no single measure which outperform other measures in all the domains.

Having said that, there are some challenges and problems in mining rules through MOEA's not yet discussed in detail in previous research on this subject as per our best knowledge. In our result and analysis section, we will discuss some of these aspects based on experimental results.

## III. Related Works

The integration of data mining and EAs is currently an active research area. In recent years, association rule mining using EAs has also garnered interest of researchers.

One of the earlier works utilizing MOEA's to mine association rules was given by Ghosh et al. [2]. They visualized association rule mining as multi-objective optimization and considered three measures: confidence, rule comprehensibility and interestingness measure, latter two being described in [2], as objectives of optimization. However, the approach presented in this work is tested only with the numerical valued attributes. Also, sample of the database was used for finding out support counts of various attributes. Inspiration from MOGA for fitness assignment scheme was used, which use non-domination concept. However, there are some challenges faced which were not highlighted in this work.

Hetland et al. [3] had applied the ideas of EAs for mining sequences and time series. Author used modified version of SPEA-2[9], where the selection of individuals for next generation use both archive members and main population. Reason for using this approach was to avoid premature convergence or obtaining solutions which were variations of initial archive. However, the algorithm is applied on spatial data and thus comparison of traditional rule mining techniques is not straightforward because it requires time series analysis.

Aouissi et al. [4] proposed QUANTMINER, a system for mining quantitative association rules. The system is based on a genetic algorithm that dynamically discovers "good" intervals in association rules by optimizing both the support and the confidence. Thus, this approach is more towards what is called optimized rule mining [15]. The power of EAs is put to use to find good intervals which can be promising. But, it faces same issues which traditional mining techniques face i.e. it relies on measures support and confidence only.

Peter et al. [7] specifies that apart from support or confidence as one of the objectives in MOEA's, J-Measure should also be included in the objective functions. Rest all measures are same as in previous studies namely rule comprehensibility, interestingness. Elitism is introduced in the form of archive maintenance inspired from SPEA-2[9].However, fitness calculation is different than SPEA-2 and involves user defined weights of objective functions. Thus, this approach becomes quite similar to traditional preference vector approach. Authors tested their approach on categorical variable only.

A new spatial mining algorithm ARMNGA is proposed by authors in [10] designed using the principles of Genetic Algorithm. It's proposed to be more efficient than Apriori in terms of rule discovery. However, as discussed earlier that only efficiency is not a concern for frequent mining techniques as there exist various algorithms[6,11], some of which are better than Apriori in time efficiency and this approach should have been compared with those algorithms. Moreover, this study used traditional mining measures viz. support and confidence.

Li et al. [14] proposes a new co-evolutionary algorithm to solve multi-objective optimization problem of association rule mining. To enhance the correlation degree and comprehensibility of association rule, two new measures, including statistical correlation and comprehensibility, as objection functions are proposed in this work. Three new co-evolutionary operators are proposed which are: Pareto neighbourhood crossover operator, combination operator and annexing operator. However, no explanation of why these operators are needed and what problems they are intended to solve was given.

## IV. CONTRIBUTION

The earlier works on mining association rules from the databases using EAs gave importance to the core step of mining association rules. The attributes used from the database and encoding scheme for representing the association rules make some assumptions which were not clearly presented. The other steps of the KDD process [1], which are quite important and influence the results especially in the case of numeric attributes, are not dealt earlier.

In this work, we applied the KDD process [1] for mining association rules using MOEAs specifying clearly the underlying assumptions and challenges involved. Algorithm used for frequent pattern mining is Apriori. The pre-processing steps in the KDD process are same for both MOEAs and traditional rule mining. Only, the core Data mining step is changed. Thus comparison of rule generation using MOEAs can be done easily. Using KDD process helped us to handle mixed type of attributes in our database, an issue which was not handled properly in earlier works. Most of works utilizing MOEAs used either all categorical or numerical attributes but never mixed type of attributes.

### *A. KDD Process*

The KDD process can be defined in these iterative and sequential steps namely 1)Data cleaning 2)Data integration 3)Data selection 4)Data transformation 5)Data mining 6)Pattern Evaluation 7)Knowledge Representation.

The steps 1-4 are pre-processing steps where the data is prepared for data mining. The data cleaning steps removes noise and inconsistent data. Data integration involved combination of multiple data sources whereas Data selection involves retrieving task relevant data from the database. Data transformation applied aggregation or summary measures so as to make data suitable for mining operations. Data mining is considered to be the core step of the whole process. Pattern evaluation identifies the truly interesting patterns based on interestingness measures. The last step, Knowledge representation is meant for end user where results are presented in format which can be easily understood using visualization techniques.

### *B. Multi Objective Optimization based Rule Mining*

For multi-objective rule mining, NSGA-II has been a good choice for obtaining Pareto-optimal solutions. The reason for using NSGA-II is that it is an elitist MOEA which produces more Pareto-optimal solutions while maintaining diversity. SPEA-2[9] is another elitist MOEA with nearly comparable results as of NSGA-II.

Multi-objective rule mining involves representing the chromosomes as the possible rules for which a particular encoding/decoding scheme is required. We used modified Michigan approach where each rule is represented by single chromosome. The encoding of the rules depends on the data attributes a lot. The assumptions behind rule encodings are

1. Attributes are categorical in nature and take fixed set of values
2. Number of attributes are fixed in the database, i.e., each tuple in whole database contains exactly k attributes and value of k is fixed for one database

But, when total number of attributes in dataset increases and only fraction of attributes appear in each tuple, then representing the rule using the modified Michigan approach is very inefficient in both time and memory because of very large chromosomes. Thus, in our study, we will be limited to datasets where number of attributes in each tuple is fixed. However, attributes can be categorical, numerical or mixed.

The chromosomes will be binary where each attribute contains two parts. First part contains tag bit represented by two bits which helps us to determine one of the three states

- An attribute is present in antecedent of rule
- An attribute is present in consequent of rule
- An attribute is not present in the rule

Since, we have to represent 3 possible states using 2 bits, one of the three states will be represented twice. To remove any bias regarding whether attribute should be present in antecedent/consequent part of the rule, we represent absence of attribute by two states. Also keeping in mind the nature of evolutionary operator mutation, we set the tag bits as follow:

- 10 for representing attribute in antecedent of rule
- 11 for representing attribute in consequent of rule
- 00 and 01 for representing absence of attribute

The tag bit encoding ensures single bit mutation to work effectively. Mutation of tag bits for absence of attributes gives equal probability for an attribute to be in antecedent or consequent, i.e., single bit mutation of 00 can lead to 10 and 01 and similarly for 01, it is 11 and 00. Also, single bit mutation for 10, 11 have similar properties.

The second part of an attribute in chromosome contains encoding for representing the M states of a categorical variable. For numerical variables, we adopt the strategy of discretization because after discretization, each range of a numerical attribute can be represented by a state.

For our study, we kept discretization simple with equi-width binning with bin size k . Simply, if a variable $p$ is bounded by $p_{min}$ and $p_{max,}$ then bin width = ($p_{max}$ – $p_{min}$ / k). Then bin boundaries are $p_{min}$ + i*(bin width) where i = 1, 2, .., k-1. Here k=10 for our study. The reason for choosing k=10 arise from the fact that one of the study [5] reported using k=10 give marginally less performance than advance discretization measures. Moreover, this type of discretization is unsupervised as it makes no use of interval class information and thus can be applied easily.

So, we have attributes which can take both M states for categorical variable and k for numerical variables. We represent in the binary form the number of these states, rather than the actual value of attribute.

The measures used in our study are interestingness, comprehensibility [2], support, confidence, cosine measure [1] and a new measure *attributes_freq* whose purpose will be revealed later on. It can be described simply as:

*attributes_freq = number of attributes in rule / total number of attributes in dataset*

We took these measures 3 at a time and obtain the results. The reason for taking 3 measures as objective functions was two-folded. One, the MOEAs work very well for objective functions ≤ 3. Second, sometimes few measures are correlated and thus one of them can suffice instead of taking all of them.

### C. *Methodology*

Steps 1-4 in KDD process are same for the both the task viz. Association Rule Mining using Apriori and Multi-objective rule mining. Step 1 was used minimally as data set was standard, still for the missing values, we filled the missing value with the mode of the particular attribute. In case of numerical variables, the average of the particular attribute is replaced with missing value.
Step 2 was not performed as dataset was obtained from single source repository UCI. Step 3 involves removing the irrelevant attributes. All the attributes which were tuple identifier, unique serial number, names were removed from consideration. All numerical attributes were preserved. Among categorical attributes, those specifying class information about tuple are removed as they are only needed in classification.

Step 4 involves converting numerical attributes to categorical attributes described in previous section. Step 5 for obtaining association rules Apriori is used as frequent mining algorithm. Steps 6 and 7 are omitted here as these are post-processing steps and are a separate area of concern.

Step 5 for multi-objective rule mining used NSGA-II [8] and objective functions were measures described earlier taking 3 at a time. One of advantages of multi-objective rule mining is clearly indicated here that Step 6 is inherently contained in the process of rule mining using MOEAs and can be changed easily to obtain results as per our requirements. The facility of obtaining rules based on interestingness measures is not available in traditional data mining algorithms. It is a post-processing step which is done once we obtain the rules.

## V. EXPERIMENTS

The Original Wisconsin Breast Cancer Database was obtained from UCI repository [12]. There were total of 11 attributes and after attribute selection, 9 attributes were selected for mining rules based on techniques described in previous section. All attributes were categorical in nature and value of M is 10 for all attributes. There are 699 instances in the dataset. The attributes in dataset has M=10 state. The rule can be deciphered like this:

a1-b1 ^ a2-b2 ^ .... => ak-bk ^ ar-br.....

Here, a, a2, ..., ak, ar represents attributes present in dataset and b1,b2, ..., bk, br represents the corresponding state of that attribute. For example: 1-5 represents 1st attribute 5th state. 3-9 represent 3rd attribute 9th state and so on. In dataset, total attributes after selection are 9 and each of 9 attributes can take states from 1 to 10.

The results for association rule mining for breast cancer dataset was obtained using settings for minimum support set to be 0.85% and minimum confidence set to 10%. Such low settings were used to obtain exhaustively the number of rules generated. Total numbers of rules generated were 9844.Some of the rules are shown below.

- 1-5 ^ 6-1 ^ 7-3 ^ 8-1 ^ 9-1=> 2-1 ^ 3-1 ^ 4-1 ^ 5-2
  confidence: 44.4444%
- 1-3 ^ 5-2 ^ 6-1 ^ 7-3 ^ 8-1 ^ 9-1 => 2-1 ^ 3-1
  confidence: 80%
- 4-1 ^ 2-4 ^ 6-1 ^ 7-3 ^ 8-1 ^ 9-1 => 2-1
  confidence: 100%

. Results for multi-objective rule mining using NSGA-II:

1. Using Interestingness, Confidence and Comprehensibility
   a. *1-1^ 5-2 ^ 6-3 ^ 8-1 ^ 9-1 => 3-1*
      support: 0.286% confidence: 100%
   b. *3-1 ^ 7-1 ^ 9-1 => 8-1*
      support:16.595% confidence: 97.479%
2. Using Interestingness, Confidence and attributes_freq
   a. *1-3 ^ 3-1 ^ 7-1 ^ 8-1 ^ 9-1 => 5-2 ^ 6-1*
      support: 2.288% confidence: 76.190%
   b. *2-7 ^ 8-1 ^ 9-1 => 1-8 ^ 5-4*
      support : 0.143% confidence : 50%
3. Using Interestingness, Cosine and attributes_freq
   a. 1-5 ^ 3-1 ^ 6-1 => 2-1 ^ 7-1 ^ 9-1
      support: 2.145% confidence: 31.25%
   b. 6-1 ^ 9-1 => 7-1
      support:18.598% confidence: 32.098%
4. Using Interestingness, Support and attributes_freq
   *a. 2-5 ^ 8-1 => 1-9 ^ 5-2*
      support: 0.286% confidence: 40%
   *b. 1-5 ^ 2-1 ^ 9-1 => 6-1*
      support: 8.011% confidence: 90.322%

The mutation probability was taken to be *1/ (chromosome length)* and crossover probability was taken to be 0.8. The rules were obtained till 50$^{th}$ generation. We observed that premature convergence was obtained after running NSGA-II for more than 50 generations.

## VI. RESULT ANALYSIS

Results for association rules using MOEAs fall short of our expectations. Nevertheless, some important observations from the results are described below.

One of the interesting observations was rule length. The measures used earlier namely interestingness, comprehensibility and support/confidence obtain rules where total number of attributes in rule are few. One of the measures comprehensibility is designed in such a way that long rules are prohibited. Thus, we included another measure called cosine instead of comprehensibility. But, our average rule length didn't change much. Although, cosine measure is influenced only by support of items in rule, there is no indication regarding rule length.

The theoretical justification for rule length being small in comparison to total number of attributes can be analyzed easily. As the number of attributes in rule increases, the probability of occurrence of all such items in terms of frequent set decreases. Moreover, even if such items occur together, measures for such rules like support, confidence etc is low and they are easily dominated by shorter rules having high interestingness, confidence or support. Thus to compensate for rule length being small, we devised a naïve measure called *attributes_freq* which promotes rules having more attributes. In our dataset, upto 9 frequent set while using Apriori were obtained covering all the attributes although using minimal settings. Therefore, we expected some rules from MOEAs from higher frequent sets like 7, 8 or 9 in our case. Even after introducing measure *attributes_freq*, only marginal improvement in results was observed. Thus, we are losing out on rules generated using higher frequent sets using MOEAs despite giving preference to rule length by using it as one of the measures.

Now, we turn our attention to another important observation. Number of rules obtained after application of MOEA. Ghosh et al. [2] took inspiration from MOGA for obtaining rules. Simple archive strategy was used so as to obtain good number of rules because final population may or may not contain better rules generated during intermediate generations. We used NSGA-II for our purpose and similar problem was also encountered by us, despite NSGA-II being elitist MOEA and maintains non-dominated solutions in next generation. We can say that more refinement is needed in parameters affecting MOEAs to solve the problem of number of rules.

Two major problems were encountered while obtaining rules using MOEA. One of the problems is of invalid rules. After the application of evolutionary operators, invalid rules are generated. Rules are termed invalid if either of the antecedent or consequent side of the rule is empty. For example, the following rule: " $1-4 \wedge 2-7 \wedge 8-2 \rightarrow$ " is invalid because consequent part is empty and it can happen vice-versa.

Second problem is of non-existent rules. Basically, the rule is derived from a frequent set. But, the *non-existent rules refer to a combination of items which never occurred together in database*. This scenario occurs most of the times in short datasets like in our case. This issue is resolved to some extent in large datasets with few attributes. However, if dataset size increases more horizontally than vertically, the problem remains as it is. By this statement, we mean that if database size increase due to more number of attributes per tuple rather than increasing the tuples for same attribute size. The chance of non-existent rule increases because number of combinations increases exponentially due to new attributes.

These two problems are reason for reduced performance of elitist MOEAs like NSGA-II to obtain more number of good rules. Due to the presence of invalid or non-existent rules, they are dominated by either weak solutions or non-dominated solutions found so far. Since, invalid or non-existent rules do not decrease in new generations because they are obtained after applying evolutionary operators like crossover, mutation. The search space tends to get stuck at local Pareto optimal solutions found so far or converge to very few Global Pareto optimal solutions if found.

Although, we did found some of the rules which would not be generated by Apriori as they would be pruned earlier due to support constraint, the number of such rules is very less.

Another good observation is that by changing measure like comprehensibility, there is marginal improvement in obtained results. This is because comprehensibility exists for non-existent rules and even for some of the invalid rules.

One of the interesting applications we found was by using combination of measures like interestingness and confidence (leaving aside comprehensibility) was outlier analysis. Attributes which occur rarely but consistently have high interestingness and also high confidence. In most of the cases, these are outliers which occur once or twice in a database. By storing the support of such rules, outliers can be easily identified as they have high values of interestingness and confidence. Rules having high interestingness, confidence and also marginally good support are the ones we really required as they are normally pruned while using traditional data mining algorithms. But due to the problems discussed above, number of such rules is very less.

Another drawback of using multi-objective rule mining is multiple scanning of database. In the whole algorithm, the most time consuming step is objective function calculation. For large databases, multiple scanning of database is usually avoided. Thus, for large databases, multi-objective rule mining needs to use only sample of databases for objective function calculation to make it a viable option.

## VII. CONCLUSION AND FUTURE WORK

Comparing the results obtained using NSGA-II and Apriori algorithm. We can clearly see that some of the rules obtained using NSGA-II cannot be obtained using traditional data mining methods like Apriori, FP-tree because support for such rules is very low. Fore ex: rules like *2-7 ^ 8-1 ^ 9-1 => 1-8 ^ 5-4* having support 0.143%.

However, the numbers of rules which are interesting and cannot be generated using traditional data mining algorithms are not significant enough. Various problems plague the multi-objective rule mining, but further studies and potential remedies can improve the results. One such area is dominance criteria. Strict dominance criteria used in our algorithm and previous works fail to produce good number of results. Improving the dominance criteria might favor the MOEAs to generate significant number of good rules normally not obtained using traditional algorithms.

## VIII. REFERENCES


[1] Jiawei Han and Micheline Kamber. *Data Mining: Concepts and Techniques* 2 Edition, Morgan Kaufmann

[2] Ashish Ghosh and Bhabesh Nath. *Multi-objective Rule Mining Using Genetic Algorithms*. Elsevier Information Sciences 163 (2004) 123–133

[3] Magnus Lie Hetland And P ˚ Al Sætrom. *Evolutionary Rule Mining in Time Series Databases*. Machine Learning, 58, 107–125, 2005, Springer 2005

[4] Ansaf Salleb-Aouissi, Christel Vrain, Cyril Nortet. *QuantMiner: A Genetic Algorithm for Mining Quantitative Association Rules.* International Joint Conference on Artificial Intelligence 2007, Pg 1035-1040

[5] Pfahringer B. *Supervised and Unsupervised Discretization of Continuous Features*. In Proc. 12th Int. Conf. Machine Learning, 456-463. 1995.

[6] Rakesh Agrawal, R. Srikant, *Fast algorithms for mining association rules*, in: Proceedings of the 20th International Conference on Very Large Databases, Santiago, Chile, 1994.

[7] Peter P. Wakabi-Waiswa, Venansius Barymureaba , *Extraction of Interesting Association Rules Using Genetic Algorithms*, Int'l Journal of Computing and ICT Research, Vol 2 No 1, June 2008

[8] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. *A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II.* IEEE Transactions on Evolutionary Computation, 6(2):182–197, 2002

[9] Eckart Zitzler, Marco Laumanns, and Lothar Thiele, *SPEA2: Improving the Strength Pareto Evolutionary Algorithm*, TIK Report 103, May 2001

[10] Shangping Dai, Li Gao, Qiang Zhu and Changwu Zhu. *A Novel Genetic Algorithm Based on Image Databases for Mining Association Rules*. In 6th IEEE/ACIS International Conference on Computer and Information Science, 2007

[11] J. Han, J. Pei and Y. Yin. *Mining Frequent Patterns without Candidate Generation*. In Proc. 2000 ACM-SIGMOD Int. Conf. Management of Data(SIGMOD'00), pages 1-12, Dallas, Tx, May 2000

[12] Breast Cancer Wisconsin (Original) Data Set http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original)

[13] Roberto J. Bayardo Jr., Rakesh Agrawal. *Mining the Most Interesting Rules*. In Proc. of the Fifth ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining, 145-154, 1999.

[14] Xiang Yang-Li and Jian Hiu. *Association Rules Mining Using Multi-objective Coevolutionary Algorithm*. International Conference of Computational Intelligence and Security Workshop 2007.

[15] Rajeev Rastogi, Kyoseuk Shim. *Mining Optimized Association Rules* with *Categorical and Numerical Attributes.* IEEE Transactions on Knowledge and Data Engineering, January/February 2002 (vol. 14 no. 1) 29-50