Ant Colony Optimization to Discover the Concealed Pattern in the Recruitment Process of an Industry

N. Sivaram
Research Scholar, Department of CSE
National Engineering College
Kovilpatti, India
Sivaram.Natarajan@airtelmail.in

K. Ramar
Department of CSE
National Engineering College
Kovilpatti, India
kramar_nec@rediffmail.com

M. Janaki Meena
Department of CSE
PSG College of Technology
Coimbatore, India
mjanakimeena@yahoo.com

***Abstract*— Recruitment of the most appropriate employees and their retention are the immense challenges for the HR department of most of the industries. Every year IT companies recruit fresh graduates through their campus selection programs. Usually industries examine the skills of the candidate by conducting tests, group discussion and number of interviews. This process requires enormous amount of effort and investment. During each phase of the recruitment process, candidates are filtered based on some performance criteria. The recruitment process of an industry differs each year based on their requirement traits and the process and criteria changes among the industries. This research focuses on investigating the underlying criteria and tries to capitalize on the existing patterns, to minimize the effort made during the recruitment process. Knowledge about the recruitment process was collected from the domain experts and decision trees were constructed with it to identify superior selection criteria. Most of the machine learning algorithms including decision trees have tainted performance in high dimensional feature space and substantiate significant increase in performance with selected features. In this paper, a novel technique based on Ant Colony Optimization is proposed to identify the attributes that impacts the selection process. The proposed ACO technique assigns heuristic information for the attributes based on the estimated conditional probabilities. Experiments were carried out using the dataset collected from an industry to identify the feature sets that give greater accuracy. Decision trees constructed using the C4.5 algorithm with the set of attributes that influence the recruitment process were used to extract feasible rules after making discussions with the domain experts.**

Keywords- Machine learning, feature selection, Ant colony optimization, C4.5 algorithm, decision trees

I. INTRODUCTION

Human resource is one of the back bones for industries to maintain their competitive advantages in the knowledge economy. Selecting fresh people with high talent and potential retention is challenging and daunting task faced by any HR department. The demand is increasing year over year and the supply has also been on the raise. But, the problem is with the quality in the supply for its recruitment. Different companies are fighting it out by taking the first slot, opting for dual placements, addressing the gaps by faculty development programs, etc. Campus recruitment is the predominant mode of recruitment for fresh talented graduates. Because of the inconsistency in the quality of the students produced by different universities and the type of skill set they acquire during their program, selecting the right candidate among those who graduate becomes a herculean task. This involves lot of effort by the recruiting team and money spent for the process is phenomenal. One of the mechanisms used by the industries is to conduct tests and group discussions during the filtration process. The selection process uses different criteria which include the average of their semester marks, marks obtained in the aptitude, programming and technical tests conducted by the company, group discussion, technical and HR interviews. These criteria are common for all the students, but the skill level of the students vary since they are from different disciplines and backgrounds. The time taken and expenditure for conducting group discussions and interviews consumes more than 90% of the total effort of the recruitment process. It has been observed that 1 among 120 students who apply get selected and the ratio of number of candidates selected against the number of candidates interviewed after tests is approximately 1:20. Reducing these ratios will immensely help the industries to save the effort. This research focuses on determining a set of selection criteria to be applied to filter candidates based on their background and academic data.

Decision tree is a simple data mining approach used to establish the hidden knowledge in the data for classification and prediction. They have the advantage of easy interpretation and understanding for the decision makers to compare with their domain knowledge for validation and justify their decisions [7]. Knowledge required for the analysis was collected from an industry and attributes associated with the recruitment process were identified. There were totally forty-four attributes. The structured dataset was split into training and testing set and decision trees were constructed with training data and tested without overfitting. Experiments were carried out with different subset of attributes and it was observed that the accuracy of the classifier increases with decision trees constructed using a few subsets of the domain attributes. Hence a feature selection algorithm based on ACO is proposed to identify the right set of attributes that impact the selection process and rules were extracted from decision trees constructed with those attributes.

In machine learning and statistics, feature selection, is the technique of selecting a subset of relevant features for building robust learning models. Ant Colony Optimization is a metaheuristic approach that could be applied for feature selection. Characteristics of the ACO include positive feedback, distributed computation and constructive greedy heuristic. Ant Colony Optimization is a technique inspired by the observation of real ants in their search for shortest paths to food source. The ants communicate through a chemical aromatic substance called as pheromones [1], [6]. In this model, subset selection problem is reformulated as a search problem in a connected graph.

The rest of this paper is organized as follows: Section 2 presents a brief overview of the methodologies used in the proposed algorithm; Section 3 describes the proposed algorithm and Section 4 describes the dataset and about the experimentations carried out and results.

II. BACKGROUND KNOWLEDGE

In this section, the general procedure of recruitment process, the process of feature selection, decision tree algorithm and Ant colony optimization algorithm are explained.

A. Recruitment Process

Selecting the right persons for the right job is the most important challenge in the human resource management. The various selection methods include analysis of application form, self-assessment, telephone screening, tests depending on the requirement of the industry (such as aptitude, technical, programming, personality, interest test, etc.) [13], [14]. Generally industries use a combination of the selection methods, based on their job nature, cost, time, accuracy, culture and acceptability. According to Lewis, there are three aspects of selection criteria. They are organizational criteria, functional/departmental criteria and individual job criteria. Finally, the recruitment committee must consider the adaptation of the job, departmental and organizational characteristics to the applicant's characteristics [12], [8]. Hence the recruitment committee designs each level of the recruitment process to reflect their needs.

The recruitment process in the campus interviews of an IT industry includes filtering based on their semester marks in the graduation, marks obtained in the aptitude, technical and programming tests conducted by the industry, grade obtained during group discussion and interviews. The company prepares a set of questions to test, if the candidate is really capable of applying what he/she learnt in his/her course of study. The questions also map to the expectations and job description for which he is recruited. To check the presentation, communication and behavioral skills of the candidate a group discussion is conducted.

B. General Procedure for Feature Selection

Attributes of the training data are the nodes in the search space and the search is made to determine the optimal subset. The procedure of feature selection consists of four steps subset generation, subset evaluation, stopping criteria and result validation [2], [3].

Subset Generation

Subset generation is a process of heuristic search in the search space, to identify a subset S. The disposition of this process is determined by two issues: the starting point of the search procedure and search strategy. Search procedure may start with an empty set and add features consecutively (forward approach) or start with all the features in the subset and eliminate one at a time in each step (backward approach) or start with an empty set and a full set, add and remove feature simultaneously in each step. Search may also begin with a randomly selected subset in order to avoid getting trapped in the local optima. Existing search strategies include complete search, sequential search and random search [2]. Generally exhaustive search is a complete search, in which every possible subset is analyzed, but the complexity of the search is exponential to the cardinality of feature space. Sequential search algorithms include pied greedy search algorithm, such as sequential forward selection, sequential backward selection, and bidirectional selection. Sequential search gives up completeness, thus work at risk to lose the optimal solutions. Random search starts with randomly selected subsets and may be proceed in a sequential way as random-restart hill climbing algorithm or generate next subset in a completely random manner. Randomness helps these methods to escape the local optima [2].

Subset Evaluation

Each newly generated subset is evaluated by a classifier dependent or independent evaluation criterion. Classifier independent algorithms are used in filter model and the goodness of the selected subset is exploited by the intrinsic characteristics of the training data. Some popular independent evaluation criteria are distance measures, dependency measures, and consistency measures. A dependant criterion used in the wrapper model requires a predetermined mining algorithm, the performance of the algorithm is considered as the evaluation criteria. Performance of wrapper model is relatively high when compared to filter model, but the complexity of the model is very high [2], [3]. The proposed algorithm uses wrapper model of feature selection and accuracy of the decision tree classifiers is used to evaluate the selected subset.

Stopping Criteria

Stopping criteria of the algorithm determines when the algorithm should stop. The stopping criteria may be formulated based on one of the following:

i. The number of iterations

ii. When subsequent addition or deletion of features does not improve the performance of the classifier

iii. When a minimum required performance is reached.

Result Validation

In real world applications, the selected features are evaluated by monitoring the changes in the mining algorithms performance.

C. Decision Trees

Decision tree is a tree structure, where internal nodes denote a test on an attribute, each branch represents the outcomes of the test and the leaf node represents the class labels. Decision tree induction is the learning of decision trees from class-labeled training tuples. Construction of decision trees is simple and fast, and does not need any domain knowledge and hence appropriate for exploratory knowledge discovery. In general, decision tree classifiers have good accuracy, but successful use of it depends on the data at hand. Decision trees are used for classification and classification rules are easily generated from them. An unknown tuple X can be classified, given its attribute values by testing the attribute values against the decision tree. The three popular methods ID3, C4.5, and CART adopt a greedy, non-backtracking approach in which decision trees are constructed in a top-down recursive divide-and-conquer manner [9],[13]. Analysis were made by constructing decision trees using all the three algorithms and it was determined that C4.5 algorithm construct tree that classifies the recruitment data well.

C4.5 algorithm, successor of ID3 uses gain ratio as splitting criterion to partition the data set. ID3 is an iterative algorithm that uses information gain as splitting criterion to construct the tree. For each attribute A, the method calculates the information gain as the difference between the information required to classify the data set based on just the proportion and the information required to classify after partitioning on A. The expected information needed to classify a tuple in the training set D is given by (1) [9],[12]:

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i) \tag{1}$$

Where p_i is the probability that an arbitrary tuple in D belongs to class C_i, and is estimated as the ratio of number of instances in class C_i in D to the total number of instances in D. The amount of information still required to classify D, after splitting them using A with v possible values is calculated using (2).

$$Info_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} \times Info(D_j) \tag{2}$$

Information gain obtained by branching the training set on the attribute A is given as in (3).

$$Gain(A) = Info(D) - Info_A(D) \tag{3}$$

The C4.5 algorithm applies a kind of normalization to information gain using a “split information” value. Split information for an attribute A with v values is defined as in (4) [9]:

$$Split\inf(A) = -\sum_{i=1}^{v} \frac{|D_i|}{|D|} \times \log_2\left(\frac{|D_i|}{|D|}\right) \tag{4}$$

Where $|D_i|$ is the number of instances in the training set D with i^{th} value for the attribute A and |D| is the total number of instances in the training set. Gain ratio is defined as in (5) and the attribute with maximum gain ratio is selected as the splitting attribute [9].

$$Gainratio(A) = \frac{Gain(A)}{Split\inf(A)} \tag{5}$$

D. Heuristic Information

Heuristic information for each attribute A in the feature space gives the approximate importance of A for the task of classification. The maximum of the conditional probabilities of the attribute A for its possible values with respect to the categories is used as the heuristic information for the problem. The heuristic information of an attribute H_A, when attributes A has V values with respect to K categories is defined as in (6).

$$H_A = \max_{\substack{i=1 \\ j=1}}^{\substack{i=v \\ j=k}} (p(A_i / C_j) \tag{6}$$

E. Ant Colony Optimization

Ant Colony Optimization is a metaheuristic algorithm that guides and modifies other heuristics to produce solutions beyond those that are normally generated in a quest for local optimality [1]. Identifying the optimal feature subset is a NP-hard problem; in worst case exact algorithms need exponential time to find the optimal solution, hence approximate algorithms also called as heuristic methods obtain good, which are near-optimal solutions at relatively low computational cost.

Ant Colony Optimization algorithm begins and proceeds randomly, at each point of choice it randomly choose ‘n’ points and analyze their probabilistic value. Node with the highest probability value is chosen as next point. The probabilistic value for each point in the search space is determined, based on its heuristic information and pheromone value [1], [5].

Ant colony optimization for feature selection regard attributes of the training data as nodes in the feature space. Each of them possesses an initial pheromone value and heuristic information according to the importance of the attribute for classification. The algorithm randomly starts with a feature, and then chooses ‘b’, (b is the branching factor at each point) features randomly and determines their probabilistic value using (7)

$$P_i^k(t) = \begin{cases} \dfrac{[\tau_i(t)]^\alpha [\eta_i]^\beta}{\sum_{u \in S} [\tau_u(t)]^\alpha [\eta_u]^\beta} & \text{if } i \in S \\ 0 & \text{Otherwise} \end{cases} \tag{7}$$

Where τ_i denotes the pheromone value and η_i denotes heuristic information for the feature i, S is the set of features chosen randomly. α and β determines the importance of the pheromone value and heuristic information respectively. The feature with the highest probabilistic value is included in the constructed feature subset. The probabilistic rule for the problem should balance the exploration and exploitation of the search. Exploration refers to the tendency of the algorithm to analyze the unvisited nodes and exploitation refers to the utilization of the search experience gathered from previous iterations. Using heuristic information is important since it gives the possibility of exploiting, problem specific knowledge [1]. Generally, two types of heuristic information static and dynamic are used. In static method, heuristic information is computed once at initialization and remains unchanged throughout the whole algorithms run. Dynamic method, changes the heuristic information according to the partial solution constructed and computed at each step of the ant's walk [1].

Pheromone contents of all the paths decrease with time, to chance on the evaporation rate and the pheromone contents of the paths taken by best performing ants are increased. The pheromone update rule is used as in (8)

$$\tau_i(t+1) = (1-\rho)\tau_i(t) + \Delta\tau_i^{best}(t) \tag{8}$$

Where $\tau_i(t)$ denotes the pheromone value at the ith iteration, $\rho \in (0,1)$ denotes the evaporation rate. The role evaporation is to avoid stagnation that is the situation in which all the ants constructs the same solution. The quantity of pheromone that each ant has to deposit is determined, by the quality of the chosen features. The quality of the selected features may be determined using the performance of the classifier.

$$\Delta\tau_i^k(t) = \begin{cases} \phi.\left(\gamma(S^k(t)\right) + \dfrac{\varphi(n-|S^k|)}{n} & \text{if i} \in S^k(t) \\ 0 & \text{Otherwise} \end{cases} \tag{9}$$

The equation to determine the deposit of pheromone is as in (9), Where $\gamma(S^k(t)$ is the performance of the classifier. ϕ and φ are the two parameters that control the relative weight of the classifier performance and length of the feature subset. Pheromone update is done for the features selected by the iteration best ant global best ant, when high exploration of search space is required and the convergence time is more for this model. Pheromone update is done only for features selected by the global best ant while fast convergence is required.

III. PROPOSED ACO FOR FEATURE SELECTION

In the feature selection algorithm in Table I, the attributes of the recruitment domain are treated as nodes; each node is associated with heuristic information and an initial pheromone value. Each ant starts with an empty subset and start at a random point. The proposed algorithm involves design of heuristic information, update function for pheromone value of the terms, and the list of features whose pheromone value is updated.

A. Update Pheromone

Pheromone values of the selected terms are updated after estimating the performance of the classifier with the selected features. The features selected by the ants are taken as input attributes and decision trees are constructed using the training set, accuracy of the constructed decision trees is determined with the testing set. The update function deposits pheromone proportional to the accuracy of the classifier. The proposed algorithm deposits pheromone on the features selected by both the global best and iteration best ant.

TABLE I. PROPOSED FEATURE SELECTION ALGORITHM

Input: D – training set of size N_d;
T – testing set of size $N - N_d$
A – set of Attributes in D
C – set of selection categories in D of size N_c
τ_0 - initial pheromone value
b – branching factor
N_{ants} – number of ants
N_i – number of Iterations
n_c – number of features to be selected
t – threshold for feature inclusion
r – ratio of features to be included
F – set of features for each ant
F_{best} – best subset of features
Output: Selected Features
Procedure ACO_for_Feature_Selection
1: Initialize r and N_i.
2: For all $a_i \in$ A do
3: Initialize a_i.pheromone = τ_0
4: For all $v_i \in$ V, where V is the set of possible values of A
5: For all $c_j \in$ C, where C is the set of categories
6: a_i.heuristic_Value $\leftarrow \max\limits_{i=1, j=1}^{i=v, j=k} (p(A_i / C_j))$
7: End For
8: End For
9: End For
10: $n_c \leftarrow \|A\| \times r$
11: For iter = 1 to N_i

Generate N_{ant} ants
For ant_1 to ant_{Nant}
For 1 to n_c
$L_{cp} \leftarrow$ List of randomly generated b features
For all $L_{cpi} \in L_{cp}$
p $\leftarrow$ Apply probabilistic rule
End For
$ant_j.F \leftarrow ant_j.F \cup$ Feature with highest p value in L_{cp}
End For
Construct decision trees DT with ant_j.F for D
$ACC_j \leftarrow$ accuracy of DT for T
$ACC \leftarrow ACC \cup ACC_j$
End For
Find maximum(ACC) and the iteration best ant
Update Pheromone of Iteration_Best_Ant.F and Global_Best_Ant.F
End For
Analyze the pheromone deposits and determine t
For 1 to N_i
Generate N_{ant} ants
For ant_1 to ant_{Nant}
$L_{cp} \leftarrow$ List of randomly generated b features
For all $L_{cpi} \in L_{cp}$
p $\leftarrow$ Apply probabilistic rule
$L_{thres} \leftarrow L_{thres} \cup L_{cpi}$ if p of $L_{cpi} > t$
End For
$ant_j.F \leftarrow ant_j.F \cup \max(L_{thres})$
Construct decision trees DT with ant_j.F for D
ACC_j accuracy of DT for T
$ACC \leftarrow ACC \cup ACC_j$
End For
Find maximum(ACC) and the iteration best ant
Update Pheromone of nodes in Iteration_Best_Ant.F and Global_Best_Ant.F
End For
$F_{best} \leftarrow$ Global_Best_Ant.F
Return F_{best}

IV. Data Set and Experimental Results

This section describes the dataset used for experimentation and the experiments conducted and results.

A. *Data Set*

Data set required for experimentation was obtained by knowledge acquisition in an industry. The challenge involved in the process is that the knowledge is spread among the domain experts and is not with a single person. Interviews were conducted with the domain experts to understand the problem and the knowledge required to solve the problem. The knowledge acquired is used along with the recruitment database maintained in the industry to form the dataset for experimentation. The data collected from the industry is complex and have noisy, missing and inconsistent data. The data is preprocessed to improve the quality of data and make it fit for the data mining task. The data used are transformed into appropriate formats to support meaningful analysis. Some more attributes are derived using the acquired knowledge to support the mining process. The data set comprised of recruitment details of two years in an IT industry. Data set of one year was used as training set and the other was used for evaluating the constructed model.

The fields in the dataset include the candidates demographical data and academic background such as first name, last name, gender, native state, student ID, year of birth, place of birth, name of high school, percentage of high school marks, name of higher secondary school, percentage of higher secondary marks, stream of diploma (If Applicable), percentage of diploma marks (If applicable), name of the college, stream of under graduation, average of internal marks in each semester, average of internal marks in each semester, consolidate percentage in under graduation, skill set in core industrial competencies rated by themselves in a scale of 1 to 10, marks secured in technical test, marks secured in aptitude test and marks secured in programming test. Totally forty four attributes were outlined and approximately three thousand records were collected. The status of selection of the candidates, which is the target variable includes different values such as recommended, on hold and rejected. The recommended status indicates that the candidate has been selected, on hold status indicate that the candidate will be selected if there is immediate requirement and the status rejected indicate that the candidate is not selected.

Dataset1 consists of 812 records and dataset2 consists of 2192 records. The final status of the candidates after the recruitment process is tabulated in Table II.

TABLE II. Final Status of Candidates in the dataset

Final Status	Percentage of Records in Dataset1	Percentage of Records in Dataset2
1	4.3	2.91
2	1.01	0.96
3	94.68	96.13

B. Experimental Results and Discussions

In the experimental setup of the ACO algorithm, the parameters α , β , ρ , ϕ and φ must be assigned appropriate values. In our experiments, equal importance was given for the heuristic information as well as the pheromone concentration on nodes hence both α and β is set to 0.5. Evaporation rate ρ is set to 0.8 and ϕ =0.8 and φ =0.2. The algorithm was tested with 500 ants by running upto 300 iterations. Decision trees were constructed by embedding the data mining tool weka, unpruned trees were constructed since pruning conceals most of the vital knowledge. Dataset 2 was used for training and dataset1 was used for testing. From Table 1, it may be observed, that the dataset consists of more than 95% of records to be in the rejected category; hence the constructed trees were very excellent in recognizing the rejected data however they were not able to identify selected records to a large extent. Therefore the dataset was premeditated and decision trees were constructed with almost equal number of records in both the categories. When the records were chosen for the learning process, the distribution of the status in the original data was maintained. The constructed decision trees were used to classify all the records in the other dataset. Table III shows the performance of the proposed ACO algorithm.

TABLE III. ASPECTS OF PERFORMANCE OF THE PROPOSED ACO ALGORITHM

Number of Iterations	Number of attributes selected	Details of the attributes selected	Accuracy of the classifier
50	27	Marks obtained in technical, aptitude and programming tests, Stream of study in under graduation, College name, native city, gender, percentage of marks secured in high school and higher secondary school, internal and external marks in all semester and total percentage in under graduation	67.12%
100	20	Marks obtained in technical, aptitude and programming tests, percentage of marks secured in high school and higher secondary school, internal and external marks in all semester and total percentage in under graduation	69.14%
150	12	Marks obtained in technical, aptitude and programming tests, percentage of marks secured in high school and higher secondary school, stream of diploma and under graduation, college name, gender, native city, native state and total percentage in under graduation	72.27%
200	8	Marks obtained in technical, aptitude and programming tests, percentage of marks secured in high school and higher secondary school, stream of under graduation, college name and total percentage in under graduation	75.31%
250	9	Marks obtained in technical, aptitude and programming tests, percentage of marks secured in high school and higher secondary school, stream of diploma and under graduation, native city and total percentage in under graduation	76.23%
300	6	Marks obtained in technical, aptitude and programming tests, percentage of marks secured in high school and higher secondary school, and total percentage in under graduation	79.77%

Accuracy of the constructed decision trees for the test dataset is as shown in Fig 1. The algorithm was stopped after 300 iterations since there was no significant increase or decrease in performance of the classifier.

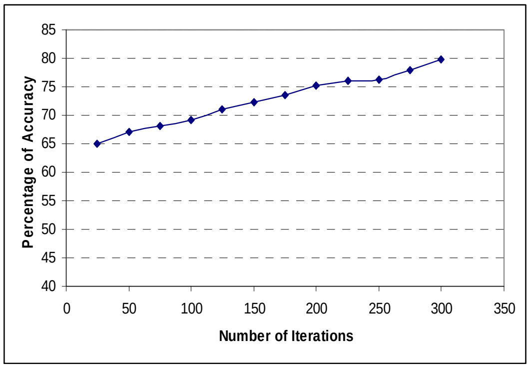

Fig 1 Accuracy of the Classifier at Different Iterations of the proposed algorithm

The performance of the decision tree classifier with all the attributes was around 61%, significant improvement in the performance of the classifier was observed with feature selection as tabulated in Table 3. Analyses were made with the decision trees with good accuracy and 15 rules were deduced. The deduced rules were checked for viability with the domain experts and used in the recruitment process. table IV lists few such rules.

TABLE IV. RULES INFERRED FROM DECISION TREES WITH GOOD ACCURACY

If Marks_In_Programming > 25 AND Percentage_In_Higher_Sec_School > 90 AND Marks_In_Technical > 15 then Selected
If College = X OR College = Y AND Mark_In_Aptitude > 15 AND Percentage_In_BE > 70 AND Percentage_In_BE < 80 AND Percentage_In_High_School > 85 AND Percentage_In_Higher_Sec_School > 90 then Selected
If College = not(X) OR College = not(Y) AND Percentage_In_BE > 70 AND Percentage_In_BE < 80 AND Mark_In_Aptitude < 10 then Rejected
If College = X OR College = Y AND Percentage_In_High_School > 90 AND Percentage_In_Higher_Sec_School > 80 AND Percentage_In_Higher_Sec_School < 90 AND Average_Internal_Marks_In_BE>10 AND Average_External_Marks_In_BE> 60 then Selected
If College = not(X) AND College = not(Y) AND Percentage_In_High_School > 90 AND Percentage_In_Higher_Sec_School > 80 AND Average_Internal_Marks_In_BE>15 AND Average_External_Marks_In_BE>55 then Rejected

V. CONCLUSION

The recruitment mining problem was identified and the domain has been well studied by interacting with the domain experts. In the study, it was tacit that the problem is NP-hard since there is no direct optimal solution which would find the criteria for selection. Also the set of input attributes are not same for all the industries and usually varies every year. Decision trees were constructed with historical data collected from an industry and it was observed that the decision trees constructed with subset of attributes was superior to the decision tree constructed with the entire input attribute set. Hence an optimization algorithm based on ACO has been designed and implemented to identify the features that impact the recruitment process. There was considerable increase in the accuracy of the classifier after few hundreds of iterations and rules were deduced from the decision trees with good accuracy. The viability of the deduced rules were verified with the domain experts and applied in the recruitment process.

VI. REFERENCES

[1] Marico Dorgio and Thomas Stutzle, “Ant Colony Optimization”, MIT Press, 2005.

[2] Huan Liu and Lei Yu, “Towards Integrating Feature Selection Algorithms for Classification and Clustering”, IEEE Transactions on Knowledge and Data Engineering, Vol., 17, No. 4, pp. 491-502, April 2005.

[3] M. Dash and H. Liu, “Feature Selection for Classification”, Intelligent Data Analysis, pp. 131-156, 1997.

[4] George Formann, “Feature Selection we’ve Barely scratched the surface”, Hewlett Packard Laboratories, Palo Alto, 2007.

[5] Ahmed Al-Ani, “Ant Colony Optimization for Feature Subset Selection”, Proceedings of World Academy of Science, Engineering and Technology, Vol. 4, pp. 35-38, February 2005.

[6] Thomas Stutzle and Holgar Hoos, “Max-Min Ant System And Local Search For The Traveling Salesman Problem”, proceedings of IEEE International Conference on Evolutionary Computation, pp. 309-314, 1997.

[7] Chen-Fu Chien and Li-Fei Chen, “Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry, Expert Systems with Application”, Elsevier, Vol. 34: pp. 280-290, 2008.

[8] M. Saidi Mehrabad and M. Fathian Brojeny, “The development of an expert system for effective selection and appointment of the jobs applicants in human resource management”, Computers and Industrial Engineering, Elsevier, Vol. 53: 306-312, 2007.

[9] Jiawei Han, Micheline Kamber, “Data Mining Concepts and Techniques”, Second Edition Morgan Kaufmann Publishers, San Francisco, 2006.

[10] Flippo, E. B., “Personnel management”, McGraw-Hill Inc, U.S, 1984.

[11] Scarpello, V. G., and Ledvinca, J. “Personnel human resources management”, South Western publishing company, U.S., 1988.

[12] Lewis, C.D. “Employee selection”, Nelson Thrones Ltd., London, UK, 1987.

[13] Usama M. Fayyad and Keki B. Irani “Multi-interval discretization of continuous valued attributes for classification learning”, Thirteenth

International Joint Conference on Artificial Intelligence, France, pp. 1022-1027, 1993.

[14] Leo Brieman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone, "Classification and Regression Trees", Chapman and Hall/CRC, New York, 1984.