Comparative Analysis of Variations of Ant-Miner by Varying Input Parameters

Sonal P. Rami
ME-CSE Student,
Dept. of Computer Engineering,
Kalol Institute of Technology & Research Centre,
Gujarat, India

Mahesh H. Panchal
Head & Associate Prof,
Dept. of Computer Engineering,
Kalol Institute of Technology & Research Centre,
Gujarat, India

ABSTRACT
Ant Colony Optimization (ACO) studies artificial systems that take inspiration from the behavior of real ant colonies and which are used to solve discrete optimization problems. ACO can be applied to the data mining field to extract rule-based classifiers. This paper presents variations of Ant-Miner named cAnt-Miner (Ant-Miner coping with continuous attributes), which incorporates an entropy-based discretization method in order to cope with continuous attributes during the rule construction process and Ant-Tree-Miner (constructing decision trees based on ACO) which generates classifications rules always in graphical form (Decision Tree). Three algorithms (Ant-Miner, Ant-Tree-Miner and cAnt-Miner) are compared against input parameters with respect to predictive accuracy and simplicity of the discovered rules.

General Terms
Data Mining, Classification, Optimization Problem

Keywords
Ant colony optimization, cAnt-Miner, Ant-Tree-Miner, decision tree.

1. INTRODUCTION
Data mining is defined as "The nontrivial extraction of implicit, previously unknown, and potentially useful information from data."Data mining sometimes called data or knowledge discovery which aims whatever data is available that data find some conclusions in the form of rules data mining can be consider as an effective and efficient way to discover or to transform the invisible to visible data "As knowledge extraction, information discovery, information harvesting, exploratory data analysis, data archeology, data pattern processing, and functional dependency analysis "Basically, the main purpose use of data mining is to manipulate huge amount of data[1].

The data analysis task is classification, where a model or classifier is constructed to predict categorical labels [1]. A classification task begins with training data for which the target values are known. The discovered knowledge is often represented in the form of IF (conditions) THEN (class) classification rules, which has the advantage of representing a comprehensible model to the user Ant Colony Optimization (ACO) concepts inspired by the behavior of natural ants. Ants often find the shortest path between a food source and the nest of the colony without using visual information. In order to exchange information about which path should be followed, ants communicate with each other by means of a chemical substance called pheromone. As ants move, a certain amount of pheromone is dropped on the ground, creating a pheromone trail. The more ants follow a given trail. This paper present an overview of Ant-Miner and variations of Ant-Miner, an ACO algorithm for discovering classification rules in data mining ACO algorithms have been successfully applied to different classification problems. Ant-Miner [2] the first implementation of an ACO algorithm for the classification task of data mining, cAnt-Miner the second implementation of an ACO algorithm to deal with continuous data, Ant-Tree-Miner the third implementation of an ACO algorithm for decision tree induction of data mining. An extension to Ant-Miner, named cAnt-Miner coping with continuous attributes [3], which incorporates an entropy-based discretization method in order to cope with continuous attributes during the rule construction process. cAnt-Miner has the ability to create discrete intervals for continuous attributes taking advantage of all continuous attributes information, rather than requiring that a discretization method be used in a pre-processing step. In Ant-Tree-Miner, each non-terminal node represents a test or decision on the considered data item and can also be interpreted as a special form of a rule set, characterized by their hierarchical organization of rules.

The remaining part of the paper is organized as follows: In section 2, Classification concept has been described along with its approaches. In section 3, Overview of Ant-Miner is given along with its algorithmic steps. In section 4, Experimental setup is shown along with its results derived. In section 6 and 7, conclusion is derived along with its future work respectively.

2. CLASSIFICATION AND VARIOUS APPROACHES
Classification is a data mining function that assigns items in a collection to target categories or classes. The goal of classification is to accurately predict the target class for each case in the data. Classification is method for Classify the rule in between predictor attribute and class label attributes. The class label of each training data is known in advance and new data is classified based on the training set is known as supervised learning. Discovered rule tested on testing dataset. Testing dataset is made up of test tuples and their associated

class labels. These tuples are randomly selected from the general dataset and they are independent to training data. Intersection of training and testing data must be null [3].

2.1 Applications of Classification

- Routing in telecommunication networks
- Traveling Salesman
- Graph Coloring
- Scheduling
- Constraint Satisfaction

2.2 Different approaches of classification

Decision trees are graphical model like tree structure that classifies instances by sorting them based on attribute selection methods. Each node in a decision tree represents a feature in an instance to be classified, and each branch represents a value that the node can assume. Instances are classified starting at the root node (topmost node) and sorted based on their feature values. Each terminal node holds class label represent outcome of the test (leaf node) and internal node denotes a test on an attribute (nonleafnode).

Bayesian classification, Bayesian classifiers are simple probabilistic classifiers based on applying bayes'theorem. They can predict class label value. A naive bayes classifier assumes that the presence or absence of a particular feature of a class is unrelated to the presence or absence of any other feature, given the class variable.

Neural network learn the classification rules by layered graph with output of one node feeding into one or many other nodes in the next layer. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation.

CN2 [4],[7] a well-known data mining algorithm for classification. It searches for a rule list in an incremental fashion. It discovers one rule at a time. Both Ant-Miner and CN2 construct a rule by starting with an empty rule and incrementally add one term at a time to the rule. The comparison of CN2 and Ant-Miner was carried out across two criteria, namely the predictive accuracy of the discovered rule lists and their simplicity

Ripper [5] is rule induction algorithm that employs a global optimization step in order to produce a set of rules, which takes into account both the quality and length of the rules.

Ant colony optimization (ACO) can be applied to the data mining field to extract rule-based classifiers based on the behavior of real ant colonies and on data mining concepts.

3. ANT-MINER

Real ant colony is that of a parallel search over several constructive computational threads based on local problem data and on a dynamic memory structure containing information on the quality of previously obtained result. Natural ants find the shortest path between a food source and the nest without using visual information. In order to exchange information about which path should be followed, ants communicate with each other by means of a chemical substance called pheromone. As ants move, a certain amount of pheromone is dropped on the ground creating a pheromone trail. The more ants follow a given trail, the more attractive that trail becomes to be followed by other ants. This process involves a loop of positive feedback in which the probability that an ant chooses a path is proportional to the number of ants that have already passed by that path. Hence, individual ant following very simple rules, interact to produce an intelligent behavior at the higher level of the ant colony. Ant Colony Optimization (ACO) studies artificial systems that take inspiration from the behavior of real ant colonies and which are used to solve discrete optimization problems. ACO can be applied to the data mining field to extract rule-based classifiers. Shortest path is discovered via pheromone trails.

3.1 Applications

- Efficiently Solves NP hard Problems- Routing, TSP (Traveling Salesman Problem) Vehicle Routing, Sequential Ordering
- Assignment - QAP (Quadratic Assignment Problem),Graph Coloring, Generalized Assignment, Frequency Assignment, University Course Time Scheduling
- Scheduling - Job Shop, Open Shop, Flow Shop Total tardiness (weighted/non-weighted), Project Scheduling, Group Shop
- Subset- Multi-Knapsack, Max Independent Set, Redundancy Allocation, Set Covering, Weight Constrained Graph Tree partition
- Machine Learning - Classification Rules, Fuzzy systems, Bayesian networks
- Network Routing - Connection oriented network routing, Connection network routing, Optical network routing

3.2 Advantages of ACO

Inherent parallelism, positive feedback accounts for rapid discovery of good solutions, efficient for Traveling Salesman Problem and similar problems, can be used in dynamic applications (adapts to changes such as new distances, etc) are some of the advantages of ACO.

3.3 Disadvantages of ACO

Theoretical analysis is difficult, sequences of random decisions (not independent) probability distribution changes by iteration, research is experimental rather than theoretical, time to convergence uncertain are some of the disadvantages.

3.4 Basic Ant-Miner Algorithm

Ant-Miner Algorithm [4] provide step by solution for discover rule .The goal of Ant-Miner is to extract classification rules from data in the form of:

IF <term1 AND term2 AND … > THEN < class >.

The algorithm is inspired by both researches on the behavior of real ant colonies and some data mining concepts as well as principles. Algorithm consists of several steps:

3.4.1 Rule construction

First Ant starts with empty rule and Ant adds one term at a time to rule choice depends on two factors: Heuristic function (problem dependent) η Pheromone associated with term τ.

3.4.2 Rule pruning

Some irrelevant terms may be added during previous phase so, Remove irrelevant, unduly included terms in rule Thus, improving simplicity of rule Iteratively remove one-term-at-a-time Test new rule against rule-quality function:

$$Q = \frac{TP}{TP + FN} * \frac{TN}{FP + TN}$$

This Process repeated until further removals no more improve quality of the rule.

3.4.3 Pheromone updating

Increase pheromone in trail followed by current ant according to Quality of found rule.

$$\tau_{ij}(t+1) = \tau_{ij}(t) + \tau_{ij}(t) * Q, \ \ \forall i, j \in R$$

3.4.4 Normalization

Normalize the amount of pheromone value at iteration for each predictor attribute

3.4.5 Stopping Criteria

1. Num. of rules >= Num. of ants
2. Convergence is met
   a. Last *k* ants found exactly the same rule, k = No_rules_converg
3. List of discovered rules is updated
4. Pheromones reset for all trails

3.5 Algorithm

Algorithm 1: Basic Ant-Miner Algorithm

Step1: The training set is classifying the rule based on Predictor attributes and class label.
Step2: Count the value of predictor attributes and Class label
Step3: Find the probability of term based on class Label.
Step4: Find the entropy (Information gain) of value For predictor attributes based on class label
Step5: Find the Heuristic function for each value of Predictor attributes
Step6: Multiply Heuristic function with amount of Pheromone (for single ant, in first iteration Amount of pheromone is same of different Value)
Step7: Choose best attribute value by maximum
Step8: Repeat for each predictor attribute
Step9: Discover the Rule.
Step10: Find Quality of rule [Quality= Sensitivity * Specificity]
Step11: Update pheromone value (Iteration=2)
Step12: Normalize the amount of pheromone value At second iteration for each predictor Attribute.

3.6 Limitation of Ant-Miner

Ant-Miner's limitation – It cannot cope up with continuous attributes.

3.7 Solution

The solution – cAnt-Miner There are numerous discretization methods for handling continuous attributes dynamically create thresholds on continuous attributes' domain values during the rule construction process.

3.8 Variations of Ant-Miner [5] [6]

1. Extend Ant-Miner to cope with continuous attributes(cAnt-Miner)
2. Inducing Decision Trees with an ACO(Ant-Tree-Miner)
3. Extend Quality Measures for classification
4. New Multi-class rule Quality measures
5. Modification for Multi-Label classification
6. Discovering fuzzy classification rules
7. Hierarchical classification
8. Fixing in advance the class predicated by a rule

3.9 Ant-Miner Toolkits

GUI Ant-Miner is a tool for extracting classification rules from data. The data input file used by GUI Ant-Miner complies with the ARFF (Attribute-Relation File Format) of the Weka tool. Data input file is standardized with the well-known Weka system, and runs on virtually any operating system since it is written in Java. It is important to mention that GUI Ant-Miner can only handle nominal attributes.

Myra is a cross-platform Ant Colony Optimization framework written in Java. It provides a specialized data mining layer to support the application of ACO to classification problems. Including the implementation of Ant-Miner, cAnt-Miner, Ant-Tree-Miner

4. EXPERIMENTALS SETUP AND RESULTS

The experiments have been performed in Myra toolkit. Two data sets are used in experiments for accuracy measurement. Ant-Miner and Ant-Tree-Miner use Soybean with 36 Attributes and 683 Instances and cAnt-Miner use Wine with 14 Attributes and 178 Instances.

Data sets are divided in two parts training data and testing data. For soybean data set training data with instances 454 and testing data with instances 229. For wine data set training data with instances 99 and testing data with 79. The data and input parameters described above are shown in table 1 and table 2.

Table 1: Description datasets used in Experiments [11]

Dataset	Size	Attributes	Classes
Soybean.arff	683	36	19
Wine.arff	178	14	3

Table 2: Description of input parameters used in Experiments

Ant-Miner	cAnt-Miner	Ant-Tree-Miner
Number of ants	Number of ants	Number of ants
No. of iterations	No. of iterations	No. of iterations
Minimum no. of covered cases	Minimum no. of covered cases	Minimum no. of instances to be covered by a branch
Maximum no. of uncovered cases	Maximum no. of uncovered cases	-----

Following Experimental results are shown in graphs below.

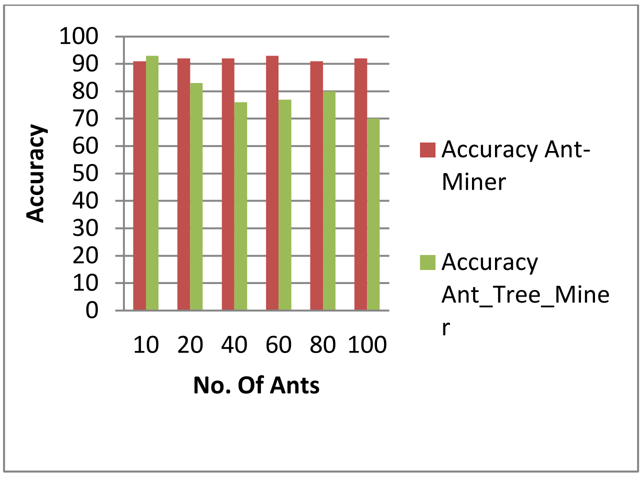

Fig 1: Comparision of Accuracy for Ant- Miner and Ant_Tree_Miner

As shown in the figure 1, it concluds that accuracy of Ant-Tree-Miner decreases with increase in no. of ants, whereas for Ant-Miner the accuracy remains constant.

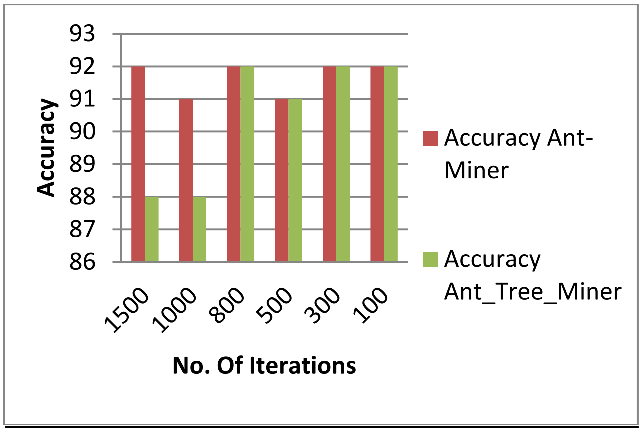

Fig 2: Comparision of Accuracy for Ant- Miner and Ant_Tree_Miner

As shown in figure 2, accuracy for Ant-Miner remains constant, whereas it was decresing for Ant-Tree Miner with increase in no.of iterations

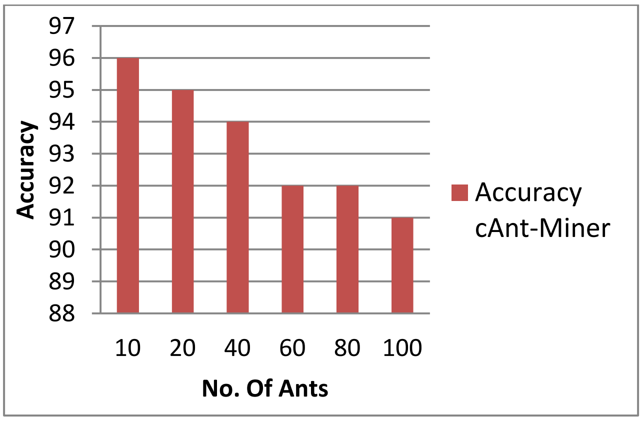

Fig 3: Accuracy vs. No. of Ants for cAnt- Miner

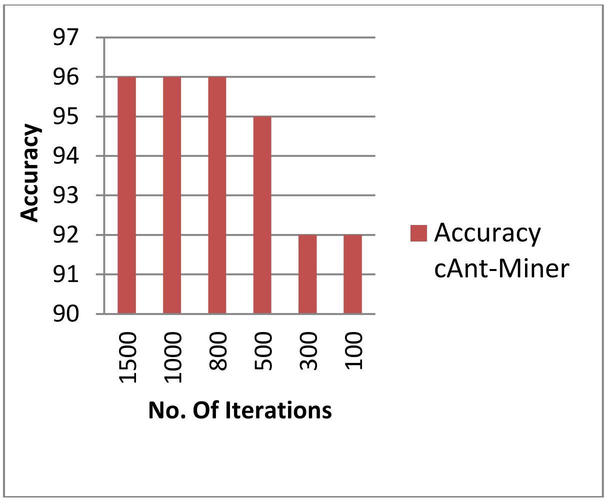

Fig 4: Accuracy vs. No. of Ants for cAnt- Miner

As shown in figure 3, accuracy for cAnt-Miner decreses with increase in no. of Ants and as per figure 4, the accuracy increases with increase in no. of iterations.

5. CONCLUSION

The goal of Ant-Miner is to discover classification rules in data sets.and Ant-Tree-Miner, for the induction of decision trees in the context of the classification task in data mining. An extension to Ant-Miner, named cAnt-Miner, which copes with continuous attributes during the rule construction process. In this paper, experiment analysis is done of the basic Ant-Miner, cAnt-Miner, Ant-Tree-Miner. The results show that, concerning predictive accuracy and simplicity of discovered rules, Ant-Miner and Ant-Tree-Miner gives somewhat better results in one categorical data set (Soybean.arff), whereas cAnt-Miner gives a considerably better result in one numerical data set (wine.arff). In general, all three algorithms has consistently found much simpler (smaller) rule lists This paper seems particularly advantageous when it is important to minimize the number of discovered rules and rule terms (conditions) in order to improve comprehensibility of the discovered knowledge. It can be argued that this point is important in many (probably most) data-mining applications, where discovered knowledge will be shown to a human user as a support for intelligent decision making.

6. FUTURE WORK

Two important directions for future research are as follows: Firstly, it would be interesting to extend Ant-Miner with Quality Measures for classification. Quality of the rule can be measured by sensitivity * confidence because confidence have more impact than specificity. Secondly, class label may contains combination of numeric and categorical (fuzzy) values, so considering membership of samples in all possible fuzzy sets, Ant miner should discover the fuzzy rules.

7. REFERENCES

[1] Han J., Kamber M.: *Data Mining – Concepts and Techniques*

[2] Dorigo, M. & Stutzle, T. (2004). *Ant Colony Optimization. Cambridge, MA*: MIT Press.

[3] Singler J., Atkinson B.: *Data Mining using Ant Colony Optimization*

[4] Witten, H., Frank, E.: *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd Edn. Morgan Kaufmann (2005)

[5] Freitas, A., Parpinelli, R., Lopes, H.: *Ant colony Algorithms for data mining.* To apper in Encyclopedia of Info. Sci. & Tech. 2nd Ed (2008)

[6] Parpinelli, R., Lopes, H., Freitas, A.: *Data mining With an ant colony optimization algorithm.* IEEE Transactions on Evolutionary Computation 6(4) (2002) 321–332

[7] Quinlan, J.: C4.5: *Programs for Machine Learning.* Morgan Kaufmann (1993)

[8] Clark, P., Niblett, *T.: The CN2 rule induction Algorithm. Machine Learning* 3(4) (1989) 261–283

[9] Swaminathan, S.: *Rule induction using ant colony Optimization for mixed variable attributes.*Master's thesis, Texas Tech University (2006)

[10] Liu, H., Hussain, F., Tan, C., Dash, M.: Discretization: *An enabling technique Data Mining and Knowledge Discovery* 6 (2002) 393–423

[11] Asuncion, A., Newman, D.: *UCI machine Learning repository.*

[12] A. Abdelhalim, I. Traore, and B. Sayed. Rbdt-1:*A New Rule-Based Decision Tree Generation Technique.* In Proceedings of the 2009 International Symposium on Rule Interchange and Applications, pages 108–121, 2009. R.C. Barros, M.P. Basgalupp, A.C.P.L.F. de

[13] Carvalho and A.A. Freitas. *A Survey of Evolutionary Algorithms for Decision-Tree Induction.* To appear in IEEE Transactions Systems, Man, and Cybernetics, Part C: Applications and Reviews,2011

[14] L. Breiman, J. Friedman, C.J. Stone, and R.A. Olshen *Classification and Regression Trees.*Chapman and Hall, 1984.

[15] Martens, D.; De Backer, M.; Haesen, R., Baesens, B. & Holvoet, T. (2006). Ants constructing rule-based classifiers. In: Agraham, A.; Grosan, C. and Ramos, V. (eds.) *Swarm Intelligence in Data Mining*, 21-43. Berlin: Springer.

[16] Wang, Z. & Feng, B. (2004). Classification rule mining with an improved ant colony algorithm.*AI 2004: Advances in Artificial Intelligence,LNAI 3339,* 357-367. Berlin, Springer.

[17] Parpinelli, R.S., Lopes, H.S. & Freitas, A.A. (2002b). An ant colony algorithm for classification rule discovery. In: Abbass, H., Sarker, R., & Newton, C. (eds.). *Data Mining: a Heuristic Approach*, London: Idea Group Publishing, 191-208.