A HYBRID OF ANT COLONY OPTIMIZATION ALGORITHM AND SIMULATED ANNEALING FOR CLASSIFICATION RULES

RIZAUDDIN SAIAN

DOCTOR OF PHILOSOPHY
UNIVERSITI UTARA MALAYSIA
2013
Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok
Abstrak


Kata Kunci: Pengoptimuman koloni semut, Simulasi penyepuhlindapan, Ant-miner, Penginduksian petua
Abstract

Ant colony optimization (ACO) is a metaheuristic approach inspired from the behaviour of natural ants and can be used to solve a variety of combinatorial optimization problems. Classification rule induction is one of the problems solved by the Ant-miner algorithm, a variant of ACO, which was initiated by Parpinelli in 2001. Previous studies have shown that ACO is a promising machine learning technique to generate classification rules. However, the Ant-miner is less class focused since the rule’s class is assigned after the rule was constructed. There is also the case where the Ant-miner cannot find any optimal solution for some data sets.

Thus, this thesis proposed two variants of hybrid ACO with simulated annealing (SA) algorithm for solving problem of classification rule induction. In the first proposed algorithm, SA is used to optimize the rule's discovery activity by an ant. Benchmark data sets from various fields were used to test the proposed algorithms. Experimental results obtained from this proposed algorithm are comparable to the results of the Ant-miner and other well-known rule induction algorithms in terms of rule accuracy, but are better in terms of rule simplicity. The second proposed algorithm uses SA to optimize the terms selection while constructing a rule. The algorithm fixes the class before rule's construction. Since the algorithm fixed the class before each rule's construction, a much simpler heuristic and fitness function is proposed. Experimental results obtained from the proposed algorithm are much higher than other compared algorithms, in terms of predictive accuracy. The successful work on hybridization of ACO and SA algorithms has led to the improved learning ability of ACO for classification. Thus, a higher predictive power classification model for various fields could be generated.

Keywords: Ant colony optimization, Simulated annealing, Ant-miner, Rule induction
Acknowledgement

First and foremost, the author would like to express his gratitude to Allah S.W.T., who has permitted him to complete this thesis.

The author gratefully acknowledges his supervisor, Prof. Dr. Ku Ruhana Ku Mahamud, who has patiently supervised his work; for her continuous encouragement, patience, guidance and promptness in expecting this academic work to anchor the voyage.

The author also wishes to express his gratitude to Ministry of Higher Education and Universiti Teknologi MARA for the study leave granted.

To his beloved wife, Dr. Zeti Zuryani Mohd Zakuan and his three princesses; Rini Barizah, Rini Bazilah and Rini Basyirah; the author appreciates their understanding and for being there with him while he is sailing through the arduous journey.

To his family and friends; the author values their words of encouragement.
Table of Contents

Permission to Use .............................................................................................................. i
Abstrak .......................................................................................................................... ii
Abstract ........................................................................................................................ iii
Acknowledgement ......................................................................................................... iv
Table of Contents .......................................................................................................... v
List of Tables ................................................................................................................... viii
List of Figures ............................................................................................................... x
List of Appendices ....................................................................................................... xii
List of Abbreviations .................................................................................................... xiii

CHAPTER ONE INTRODUCTION ................................................................. 1
1.1 Problem Statement ................................................................................................. 4
1.2 Research Objectives ............................................................................................... 6
1.3 Significance of the Research .................................................................................. 6
1.4 Scope, Assumptions and Limitations of the Research ............................................. 7
1.5 Structure of the Thesis ........................................................................................... 8

CHAPTER TWO LITERATURE REVIEW ........................................ 10
2.1 Data Mining and Classification .............................................................................. 10
2.2 Classification Using Rule Induction ........................................................................ 11
2.3 Ant Colony Optimization Metaheuristic ............................................................... 15
2.4 Applications of Ant Colony Optimization .............................................................. 17
2.5 Ant Colony Optimization for Rule Induction ......................................................... 22
    2.5.1 Train by Fixing Classes .................................................................................. 23
    2.5.2 New Heuristic Functions .............................................................................. 24
    2.5.3 New Pheromone Updating Procedure ........................................................ 24
    2.5.4 Pseudorandom Proportional Transition Rule ............................................. 25
    2.5.5 Remove Pruning Procedure ........................................................................ 25
2.6 Simulated Annealing Algorithm ............................................................................ 26
    2.6.1 Applications of Simulated Annealing .......................................................... 28
    2.6.2 Hybrid ACO and SA Algorithm Variants .................................................... 30
2.7 Hybrid ACO for Rule Induction ................................................................. 32
2.8 Summary ................................................................................................. 35

CHAPTER THREE RESEARCH METHODOLOGY ............................................ 36
3.1 Data Set Development ........................................................................... 37
3.2 Algorithm Formulation .......................................................................... 43
3.3 Rule Validation ....................................................................................... 43
3.4 Summary .................................................................................................. 46

CHAPTER FOUR ATTRIBUTE SELECTION METHODS FOR
DIMENSIONALITY REDUCTION ........................................................................ 48
4.1 Attribute selection method ...................................................................... 49
4.2 Best Attribute Selection Method ............................................................ 51
4.3 Performance of Ant-miner on Reduced Attributes Data Sets .................... 57
4.4 Summary .................................................................................................. 60

CHAPTER FIVE SIMULATED ANNEALING AS LOCAL SEARCH IN ANT
COLONY OPTIMIZATION FOR RULE INDUCTION ........................................ 62
5.1 Simulated Annealing as Local Search ..................................................... 62
5.2 Experiment and Results ......................................................................... 73
  5.2.1 Classification of 17 Data Sets from UCI Repository ............................ 73
  5.2.2 Classification of Web Data Set ............................................................. 90
5.3 Summary .................................................................................................. 91

CHAPTER SIX SIMULATED ANNEALING FOR BEST TERMS
SELECTION ......................................................................................................... 93
6.1 Simulated Annealing for Term Selection ............................................... 94
6.2 Experiment and Results ......................................................................... 108
  6.2.1 Classification of 17 Data Sets from UCI Repository ......................... 109
  6.2.2 Classification of Web Data Set ............................................................. 125
6.3 Summary .................................................................................................. 127

CHAPTER SEVEN CONCLUSION AND FUTURE WORK ......................... 129
7.1 Research Contribution ............................................................................ 129
7.2 Future Work ............................................................................................ 130
# List of Tables

Table 3.1: Data Sets Used in the Experiments.................................................................39  
Table 4.1: Search Methods for Attribute Selection........................................................50  
Table 4.2: Attribute Evaluation Methods for Attribute Selection......................................50  
Table 4.3: The Numbers of Attributes Generated by Various Attribute Selection Methods .54  
Table 4.4: Comparison Between C4.5 and Ant-miner for Average Predictive Accuracy ......55  
Table 4.5: Comparison Between C4.5 and Ant-miner for Average Number of Rules ........56  
Table 4.6: The Number of Attributes Before and After Reduction ....................................58  
Table 4.7: Comparison of The Average Predictive Accuracy for Models Constructed by Ant-
miner on Original and Reduced UCI Data Sets ..................................................................58  
Table 4.8: Comparison of The Average Number of Rules for Models Constructed by Ant-
miner on Original and Reduced UCI Data Sets ...............................................................59  
Table 4.9: Comparison of The Average Number of Terms for Models Constructed by Ant-
miner on Original and Reduced UCI Data Sets ...............................................................60  
Table 5.1: Average Predictive Accuracy (%) of Ant-miner and Proposed Algorithm 1 ......75  
Table 5.2: Average Number of Rules of Ant-miner and Proposed Algorithm 1 ...............76  
Table 5.3: Average Number of Terms of Ant-miner and Proposed Algorithm 1 ..............77  
Table 5.4: Average Predictive Accuracy (%) of Ant-miner and Proposed Algorithm 1 on Reduced Attributes Data Sets .................................................................................82  
Table 5.5: Average Number of Rules of Ant-miner and Proposed Algorithm 1 on Reduced Attributes Data Sets .........................................................................................83  
Table 5.6: Average Number of Terms of Ant-miner and Proposed Algorithm 1 on Reduced Attributes Data Sets .........................................................................................84  
Table 5.7: Average Predictive Accuracy (%) of Conjuctive Rule, Decision Table, DTNB,  
JRip, PART, ACO/PSO2 and Proposed Algorithm 1 .........................................................89  
Table 5.8: Average Number of Rules of JRip, PART, PSO/ACO2 and Proposed Algorithm 1  
.................................................................................................................................90  
Table 5.9: Performance Comparison for Reduced Web Data..............................................91  
Table 6.1: Average Predictive Accuracy of Ant-miner and Proposed Algorithm 2 ..........110  
Table 6.2: Average Number of Rules of Ant-miner and Proposed Algorithm 2 ..............111  
Table 6.3: Average Number of Terms of Ant-miner and Proposed Algorithm 2 ............112  
Table 6.4: Average Predictive Accuracy of Ant-miner and Proposed Algorithm 2 on Reduced Attributes Data Sets .........................................................................................117
Table 6.5: Average Number of Rules of Ant-miner and Proposed Algorithm 2 on Reduced Attributes Data Sets ........................................................................................................................................118
Table 6.6: Average Number of Terms of Ant-miner and Proposed Algorithm 2 on Reduced Attributes Data Sets ........................................................................................................................................119
Table 6.7: Average Predictive Accuracy of Conjuctive Rule, Decision Table, DTNB, JRip, PART and Proposed Algorithm 2 ........................................................................................................................................124
Table 6.8: Average Number of Rules of JRip, PART, PSO/ACO2 and Proposed Algorithm 2 ........................................................................................................................................125
Table 6.9: Performance Comparison for Reduced Web Data........................................................................................................................................126
List of Figures

Figure 1.1: Classification Task General Framework .......................................................... 1
Figure 2.1: An Example of a Classification Rules ............................................................. 12
Figure 2.2: Experimental Setup for the Double Bridge Experiment ................................ 17
Figure 3.1: Research Phases .............................................................................................. 36
Figure 3.2: k-fold Cross Validation Procedure .................................................................. 44
Figure 4.1: The Process of Generating Rules ..................................................................... 52
Figure 5.1: Sequential Covering Algorithm ........................................................................ 63
Figure 5.2: SA as Local Search in ACO Flow Chart .......................................................... 66
Figure 5.3: SA Flow Chart to Construct Best Rule for an Ant ......................................... 68
Figure 5.4: Terms Selection Procedure Flow Chart ............................................................ 72
Figure 5.5: Comparison of Average Predictive Accuracy Between Ant-miner and Proposed Algorithm 1 ................................................................................................................. 78
Figure 5.6: Comparison of Average Number of Rules Between Ant-miner and Proposed Algorithm 1 ......................................................................................................................... 79
Figure 5.7: Comparison of Average Number of Terms Between Ant-miner and Proposed Algorithm 1 ......................................................................................................................... 80
Figure 5.8: Comparison of Average Predictive Accuracy Between Ant-miner and Proposed Algorithm 1 on Reduced Attributes Data Sets ................................................................. 85
Figure 5.9: Comparison of Average Number of Rules Between Ant-miner and Proposed Algorithm 1 on Reduced Attributes Data Sets ........................................................................... 86
Figure 5.10: Comparison of Average Number of Terms Between Ant-miner and Proposed Algorithm 1 on Reduced Attributes Data Sets ........................................................................... 87
Figure 5.11: Performance Comparison for Reduced Web Data .......................................... 91
Figure 6.1: Sequential Covering Algorithm with Pre-Defined Class ................................ 95
Figure 6.2: Sequential Covering with Pre-Defined Class Flow Chart ................................. 96
Figure 6.3: ACO Algorithm to Extract One Rule ............................................................... 98
Figure 6.4: ACO Algorithm to Extract One Rule Flow Chart .......................................... 99
Figure 6.5: Terms Selection Procedure ............................................................................... 103
Figure 6.6: Terms Selection Procedure Flow Chart .......................................................... 104
Figure 6.7: SA Algorithm to Select One Term Flow Chart .............................................. 105
Figure 6.8: Comparison of Average Predictive Accuracy between Ant-miner and Proposed Algorithm 2 ....................................................................................................................... 113
Figure 6.9: Comparison of Average Number of Rules between Ant-miner and Proposed Algorithm 2 ................................................................. 114
Figure 6.10: Comparison of Average Number of Terms between Ant-miner and Proposed Algorithm 2 ........................................................................................................ 115
Figure 6.11: Comparison of Average Predictive Accuracy between Ant-miner and Proposed Algorithm 2 on Reduced Attributes Data Sets ...................................................... 120
Figure 6.12: Comparison of Average Number of Rules between Ant-miner and Proposed Algorithm 2 on Reduced Attributes Data Sets ................................................................. 121
Figure 6.13: Comparison of Average Number of Terms between Ant-miner and Proposed Algorithm 2 on Reduced Attributes Data Sets ...................................................... 122
Figure 6.14: Performance Comparison for Reduced Web Data ............................................. 127
List of Appendices

Appendix A List of Stop Words ................................................................. 144
Appendix B Bash Script for Creating Train/Test Sets .................................. 146
Appendix C Web Classification Sample Data Set ........................................... 147
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACO</td>
<td>Ant colony optimization</td>
</tr>
<tr>
<td>AD</td>
<td>Air defense</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial neural network</td>
</tr>
<tr>
<td>ASA</td>
<td>Adaptive simulated annealing</td>
</tr>
<tr>
<td>C2</td>
<td>Command and control</td>
</tr>
<tr>
<td>DFR</td>
<td>Distribution feeder reconfiguration</td>
</tr>
<tr>
<td>DGs</td>
<td>Distributed generators</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic algorithm</td>
</tr>
<tr>
<td>IIR</td>
<td>Infinite-impulse-response</td>
</tr>
<tr>
<td>IR</td>
<td>Information retrieval</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum likelihood</td>
</tr>
<tr>
<td>MMAS</td>
<td>Max-Min ant system</td>
</tr>
<tr>
<td>MSER DFE</td>
<td>Minimum symbol-error-rate decision feedback equalizer</td>
</tr>
<tr>
<td>ODP</td>
<td>DMOZ Open Directory Project</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle swarm optimization</td>
</tr>
<tr>
<td>SA</td>
<td>Simulated annealing</td>
</tr>
<tr>
<td>SAM</td>
<td>Surface to air missile</td>
</tr>
<tr>
<td>STWTS/SDS</td>
<td>Single machine total weighted tardiness with sequence-dependent setups</td>
</tr>
<tr>
<td>TAP</td>
<td>Target assignment problem</td>
</tr>
<tr>
<td>TS</td>
<td>Tabu search</td>
</tr>
<tr>
<td>TSP</td>
<td>Travelling salesman problem</td>
</tr>
<tr>
<td>Web-&gt;KB</td>
<td>CMU World Wide Knowledge Base</td>
</tr>
</tbody>
</table>
The contents of the thesis is for internal user only
CHAPTER ONE
INTRODUCTION

The tremendous growth in computing power and storage capacity, the availability of increased access to data from Web navigation and intranets, the explosive growth in data collection, the storing of the data in data warehouses, and the competitive pressure to increase market share in globalized economy stimulated the development of data mining. Data mining acts as a tool to extract or yield important information from raw data.

Classification is a data mining task of finding the common properties among different objects and classifying the objects into classes. Figure 1.1 depicts the general framework of classification task. The classification model contains a set of classification rules. The classification model categorizes new unseen example data, by predicting a class label for the example. One way of presenting the classification model is by representing the information as a set of IF-THEN rules (classification rules).

![Figure 1.1: Classification Task General Framework](image)

*Figure 1.1: Classification Task General Framework*
REFERENCES


International Conference on Recent Advances and Future Trends in Information Technology (iRAFIT 2012), iRAFIT(6), 7–13.


