

Volume 57 Number 9 September 2014

ISSN 1674-733X CN 11-5847/TP

中国科学：信息科学

SCIENCE CHINA Information Sciences

Sponsored by

CHINESE ACADEMY OF SCIENCES

NATIONAL NATURAL SCIENCE FOUNDATION OF CHINA

info.scichina.com

link.springer.com



SCIENCE CHINA PRESS

Springer



Editorial Board

Honorary Editor General

Editor General

Editor-in-Chief

GuangZhao ZHOU (Zhou Guang Zhao)

ZuoYan ZHU *Institute of Hydrobiology, CAS, China*

Wei LI *Beihang University, China*

Advisory Committee

Cor CLAEYS

Interuniversity Microelectronics Centre, Belgium

Hiroshi IWAI

Tokyo Institute of Technology, Japan

YanDa LI

Tsinghua Univ., China

ShengGang LIU

Univ. of Electronic Science & Technology of China, China

T. P. MA

Yale Univ., USA

Paul J. WERBOS

National Science Foundation, USA

Howard M. WISEMAN

Griffith University, Australia

YaQin ZHANG

Microsoft Co., Ltd, USA

Taieb ZNATI

The Univ. of Pittsburgh, USA

Executive Associate Editors-in-Chief

Hong MEI

Peking Univ., China

Dongming WANG

Centre National de la Recherche Scientifique, France

Lei GUO

Academy of Mathematics & Systems Science, CAS, China

Associate Editors-in-Chief

Ru HUANG

Peking Univ., China

XiaoHu YOU

Southeast Univ., China

Members

Cesare ALIPPI

Politecnico di Milano, Italy

Jordan M. BERG

Texas Tech Univ., USA

JianEr CHEN

Texas A&M Univ., USA

JingSheng Jason CONG

Univ. of California, Los Angeles (UCLA), USA

S. Barry COOPER

Univ. of Leeds, U.K.

Simon DELEONIBUS

Laboratorios LETI, France

Richard LiMin DU

Voxeasy Institute of Technology, China

Wen GAO

Peking Univ., China

ShuZhi Sam GE

National Univ. of Singapore, Singapore

JiFeng HE

East China Normal Univ., China

XiaoMing HU

Royal Institute of Technology, Sweden

ZhanYi HU

Institute of Automation, CAS, China

Jie HUANG

The Chinese Univ. of Hong Kong, Hong Kong, China

Amir HUSSAIN

Univ. of Stirling, U.K.

YueFeng JI

Beijing Univ. of Post & Telecommunication, China

Chen JIA

ZhongPing JIANG

Polytechnic Institute of NYU, USA

Hai JIN

Huazhong Univ. of Science & Technology, China

ZhongLiang JING

Shanghai Jiao Tong Univ., China

XueJia LAI

Shanghai Jiao Tong Univ., China

Joshua LeWei LI

Monash Univ., Australia

WeiPing LI

Univ. of Science & Technology of China, China

XueLong LI

Xi'an Institute of Optics & Precision, CAS, China

GuiSheng LIAO

Xidian Univ., China

DongDai LIN

Institute of Information Engineering, CAS, China

ZongLi LIN

Univ. of Virginia, USA

DeRong LIU

Institute of Automation, CAS, China

KePing LONG

Univ. of Science & Technology Beijing, China

Teng LONG

Beijing Institute of Technology, China

Jian LÜ

Nanjing Univ., China

PingXi MA

China Electronics Corporation, China

David Z. PAN

Univ. of Texas at Austin, USA

Marios M. POLYCARPOU

Univ. of Cyprus, Cyprus

Long QUAN

The Hong Kong Univ. of Science & Technology, Hong Kong, China

XianHe SUN

Illinois Institute of Technology, USA

ZhiMin TANG

Institute of Computing Technology, CAS, China

Jie TIAN

Institute of Automation, CAS, China

WeiTek TSAI

Arizona State Univ., USA

ChengXiang WANG

Heriot-Watt Univ., U.K.

JiangZhou WANG

Kent Univ., U.K.

Long WANG

Peking Univ., China

XiaoDong WANG

Columbia Univ., USA

ZiYu WANG

Peking Univ., China

Martin D. F. WONG

Univ. of Illinois, USA

Jie WU

Temple Univ., USA

WeiRen WU

Lunar Exploration and Aerospace Engineering Center, China

XinDong WU

Univ. of Vermont, USA

YiRong WU

Institute of Electronics, CAS, China

Donald C. WUNSCH

Missouri Univ. of Science & Technology, USA

XiangGen XIA

Univ. of Delaware, USA

ChengZhong XU

Wayne State Univ., USA

Jun XU

Tsinghua Univ., China

Ke XU

Beihang Univ., China

ZongBen XU

Xi'an Jiaotong Univ., China

Qiang YANG

The Hong Kong Univ. of Science & Technology, Hong Kong, China

Xin YAO

Univ. of Birmingham, U.K.

MingSheng YING

Tsinghua Univ., China

HuanGuo ZHANG

Wuhan Univ., China

FuChun ZHENG

Univ. of Reading, U.K.

Dian ZHOU

The Univ. of Texas at Dallas, USA

ZhiHua ZHOU

Nanjing Univ., China

Albert Y. ZOMAYA

The Univ. of Sydney, Australia

Editorial Staff

Fei SONG

Jing FENG

JunQing LI

Kai JIANG

Cover Designer

Yu HU

Contents

Vol. 57 No. 9 September 2014

RESEARCH PAPER

Generic regular decompositions for generic zero-dimensional systems.....	092101(14)
TANG XiaoXian, CHEN ZhengHong & XIA BiCan	
Small universal simple spiking neural P systems with weights	092102(11)
ZENG XiangXiang, PAN LinQiang & PÉREZ-JIMÉNEZ Mario J.	
The incremental subgradient methods on distributed estimations in-network	092103(10)
FENG Hui, JIANG ZiDong, HU Bo & ZHANG JianQiu	
Acyclic orientation graph coloring for software-managed memory allocation	092104(18)
WANG Li, XUE JingLing & YANG XueJun	
Lighting virtual objects in a single image via coarse scene understanding	092105(14)
CHEN XiaoWu, JIN Xin & WANG Ke	
Self-adaptive spatial image denoising model based on scale correlation and SURE-LET in the nonsubsampled contourlet transform domain.....	092106(15)
LIANG MeiYu, DU JunPing & LIU HongGang	
Multi-feature hierarchical topic models for human behavior recognition	092107(15)
LI HePing, ZHANG Feng & ZHANG ShuWu	
High-efficiency pipeline design of binary arithmetic encoder.....	092108(8)
SONG Rui, GUI HongFei, LI YunSong & WU ChengKe	
Trust-based service composition in multi-domain environments under time constraint	092109(16)
ZHANG Tao, MA JianFeng, LI Qi, XI Ning & SUN Cong	
DNA-chip-based dynamic broadcast encryption scheme with constant-size ciphertexts and decryption keys.....	092110(10)
FANG XiWen & LAI XueJia	
A new parallel lattice reduction algorithm for BKZ reduced bases.....	092111(10)
LIU XiangHui, FANG Xing, WANG Zheng & XIE XiangHui	
Secure linear system computation in the presence of malicious adversaries.....	092112(10)
ZHANG Bo & ZHANG FangGuo	
Analysis and improvement of a provable secure fuzzy identity-based signature scheme.....	092113(5)
XIONG Hu, CHEN YaNan, ZHU GuoBin & QIN ZhiGuang	
Novel way to research nonlinear feedback shift register	092114(14)
ZHAO DaWei, PENG HaiPeng, LI LiXiang, HUI SiLi & YANG YiXian	
Robust sparse principal component analysis.....	092115(14)
ZHAO Qian, MENG DeYu & XU ZongBen	
Unordered rule discovery using Ant Colony Optimization	092116(15)
KHAN Salabat, BAIG Abdul Rauf, ALI Armughan, HAIDER Bilal, KHAN Farman Ali, DURRANI Mehr Yahya & ISHTIAQ Muhammad	
A trend based investment decision approach using clustering and heuristic algorithm	092117(14)
WU ChungMin, CHOU ShengChun & LIAW HorngTwu	
Characteristic model-based all-coefficient adaptive control for automatic train control systems	092201(12)
GAO ShiGen, DONG HaiRong & NING Bin	
Leader-follower formation control without leader's velocity information	092202(12)
SHEN DongBin, SUN ZhenDong & SUN WeiJie	
Outlier deletion based improvement on the StOMP algorithm for sparse solution of large-scale underdetermined problems	092203(14)
ZHANG WanHong, ZHOU Tong & HUANG BoXue	
Uncovering network traffic anomalies based on their sparse distributions	092204(11)
CHENG GuoZhen, CHEN HongChang, CHENG DongNian, ZHANG Zhen & LAN JuLong	

Ownership by Science China Press; Copyright©2014 by Science China Press, Beijing, China and Springer-Verlag Heidelberg, Germany

Submission of a manuscript implies: that the work described has not been published before (except in the form of an abstract or as part of a published lecture, review, or thesis); that it is not under consideration for publication elsewhere; that its publication has been approved by all co-authors, if any, as well as — tacitly or explicitly — by the responsible authorities at the institution where the work was carried out. The author warrants that his/her contribution is original and that he/she has full power to make this grant. The author signs for and accepts responsibility for releasing this material on behalf of any and all co-authors. Transfer of copyright to Science China Press and Springer becomes effective if and when the article is accepted for publication. After submission of the Copyright Transfer Statement signed by the corresponding author, changes of authorship or in the order of the authors listed will not be accepted by Science China Press and Springer. The copyright covers the exclusive right (for U.S. government employees: to the extent transferable) to reproduce and distribute the article, including reprints, translations, photographic reproductions, microform, electronic form (offline, online) or other reproductions of similar nature.

An author may self-archive an author-created version of his/her article on his/her own website. He/she may also deposit this version on his/her institution's and funder's (funder designated) repository at the funder's request or as a result of a legal obligation, including his/her final version, provided it is not made publicly available until after 12 months of official publication. He/she may not use the publisher's PDF version which is posted on www.springerlink.com for the purpose of self-archiving or deposit. Furthermore, the author may only post his/her version provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The original publication is available at www.springerlink.com".

All articles published in this journal are protected by copyright, which covers the exclusive rights to reproduce and distribute the article (e.g., as offprints), as well as all translation rights. No material published in this journal may be reproduced photographically or stored on microfilm, in electronic data bases, video disks, etc., without first obtaining written permission from the publishers. The use of general descriptive names, trade names, trademarks, etc., in this publication, even if not specifically identified, does not imply that these names are not protected by the relevant laws and regulations.

While the advice and information in this journal is believed to be true and accurate at the date of its going to press, neither the authors, the editors, nor the publishers can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Special regulations for photocopies in the USA: Photocopies may be made for personal or in-house use beyond the limitations stipulated under Section 107 or 108 of U.S. Copyright Law, provided a fee is paid. All fees should be paid to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA, Tel.: +1-978-7508400, Fax: +1-978-6468600, <http://www.copyright.com>, stating the ISSN of the journal, the volume, and the first and last page numbers of each article copied. The copyright owner's consent does not include copying for general distribution, promotion, new works, or resale. In these cases, specific written permission must first be obtained from the publishers.

Unordered rule discovery using Ant Colony Optimization

KHAN Salabat^{1,3*}, BAIG Abdul Rauf^{2,3}, ALI Armughan¹, HAIDER Bilal¹,
KHAN Farman Ali¹, DURRANI Mehr Yahya¹ & ISHTIAQ Muhammad^{1,3}

¹Department of Computer Science, COMSATS Institute of Information Technology, Attock Campus, Pakistan;

²Al Imam Mohammad Ibn Saud Islamic University (IMSIU), College of Computer and Information Sciences,
Riyadh, Saudi Arabia;

³Department of Computer Science, National University of Computer and Emerging Sciences, Islamabad, Pakistan

Received January 12, 2014; accepted April 16, 2014; published online June 27, 2014

Abstract In this article, a novel unordered classification rule list discovery algorithm is presented based on Ant Colony Optimization (ACO). The proposed classifier is compared empirically with two other ACO-based classification techniques on 26 data sets, selected from miscellaneous domains, based on several performance measures. As opposed to its ancestors, our technique has the flexibility of generating a list of IF-THEN rules with unrestricted order. It makes the generated classification model more comprehensible and easily interpretable. The results indicate that the performance of the proposed method is statistically significantly better as compared with previous versions of AntMiner based on predictive accuracy and comprehensibility of the classification model.

Keywords classification, ant colony optimization, data mining, unordered rule set, comprehensibility, pattern recognition

Citation Khan S, Baig A R, Ali A, et al. Unordered rule discovery using Ant Colony Optimization. Sci China Inf Sci, 2014, 57: 092116(15), doi: 10.1007/s11432-014-5133-5

1 Introduction

A classification problem encompasses the assignment of an unseen sample to a predefined class based on the characteristics of the given sample (in the form of feature values). To perform this task, a classifier is first trained. The training is performed using randomly selected samples related to the same concept (class) and termed as training set. Later, the trained model is tested over unseen samples to determine its generalization power. The objective of the trained model is to predict the class labels of unseen patterns as accurately as possible.

Swarm intelligence [1–3], which deals with the collective behavior of small and simple entities, has been used in many application domains. Ant Colony Optimization (ACO) proposed in the early 1990s [4–6] is one of the most famous meta-heuristic under the umbrella of swarm intelligence. Although each individual entity/agent has only limited capabilities, the complete swarm exhibits complex overall behavior [7]. Since its inception, ACO has been used to solve many complex problems, including those related to data

*Corresponding author (email: salabat.khan@ciit-attock.edu.pk)

mining [8–11] as well as other combinatorial optimization problems: for example, Travelling Salesperson Problem (TSP), quadratic assignment, vehicle routing, connection oriented and connectionless network routing, sequential ordering, graph coloring, shortest common super sequence, single machine tardiness, and multiple knapsack.

In this article, we present a new ACO-based algorithm for the discovery of IF-THEN classification rules. The first ant-based classifier system known as AntMiner is reported in the research work of Parpinelli et al. [12]. The authors have focused on two main objectives: first, a classifier system must be accurate and second, it must be comprehensible for different domain experts. The proposed algorithm, called UAntMiner-C, offers both the properties of accuracy and comprehensibility (as its ancestors) and also discovers the rules that can be used in unordered manner (unlike original AntMiner versions). For ordered rule list, say for example containing 100 rules, 99th rule is only meaningful if all the first 98 rules do not get activated/ fired. In other words, 99th rule is only understandable in the context of all the rules preceding it in the ordered rule list. This dependency of the rules in the ordered rule list makes a classifier less comprehensible. On the other hand, UAntMiner-C discovers unordered rule list. In this way, a user has the flexibility of interpreting a rule independently. The proposed algorithm has some unique features (discussed in subsequent sections), which allows it to achieve a higher accuracy rate when compared with some previously proposed ACO-based classification approaches. The main contributions of our work are summarized as follows:

1. A novel heuristic function is proposed based on the correlation of two terms (attribute-value pairs) for the selection of candidate conditions to be added to the antecedent part of IF-THEN rule.
2. A parallel rule learning method across class labels is introduced. For instance, the iterations of FOR-EACH loop (in the UAntMinerC algorithm) can be executed in parallel (for a data set with multiple class labels) to achieve better performance in terms of execution time.
3. The usage of asymmetric matrices to store the pheromone and heuristic values. Due to the usage of asymmetric edges in the search space (designed for rule generation), the search process does not converge prematurely as compared with existing ACO-based classifiers.
4. The rules are discovered without any strict ordering to avoid the dependency between the rules if discovered in order.
5. The usage of three conflict resolution strategies and fruitful discussion on why more conflicts can cause inconsistent behavior in the performance of an unordered rule list classifier?

The remainder of this paper is organized as follows. In Section 2, we review the related research work (i.e., different ACO based classification algorithms known as AntMiner s). In Section 3, the architecture of the proposed solution will be discussed. Subsequently, in Section 4, we present some simulation results to show the promising ability of our technique. Finally, Section 5 will conclude this work.

2 Related work

2.1 AntMiner

The first ant-based classifier system known as AntMiner is reported in the research work of Parpinelli et al. [12]. The idea is to extract an ordered rule list of the form “if rule antecedent then consequent” using training samples selected from a benchmark data set. The antecedent part consists of conjunction of terms/ conditions of the form $\langle \text{Attribute}, \text{operator}, \text{Value} \rangle$ pair e.g. Color = Red and consequent part is a class value predicted by the rule. For example, the rule “IF $\langle \text{Color} = \text{Red} \text{ AND } \text{Shape} = \text{Round} \rangle$ THEN $\langle \text{Apple} \rangle$ ” contains two conditions based on which the class label ‘Apple’ is predicted. In the testing phase, the discovered rule list is applied in sequential order to classify the unseen samples.

The data set is first discretized, as AntMiner [12] works only on categorical attributes. An ant constructs a rule by starting with an empty rule and incrementally adds one term/ condition at a time to the antecedent part of the rule. The selection of a condition is probabilistic and based on two factors: (1) heuristic value and (2) the amount of pheromone associated with the term. The authors use the information gain as the heuristic value for the selection of a condition. After the antecedent part of a rule

is constructed, the consequent of the rule is assigned by a majority vote of the training samples covered by the rule. The constructed rule is then pruned by removing irrelevant terms as an attempt to improve its accuracy. The quality of the rule is determined and used to update the pheromone values on the trail followed by the ant (conditions selected by the ant). After all ants have constructed their rules, the best quality rule is selected and added to the discovered rule list. The training samples correctly classified by the discovered rule are removed from the training set. This process is continued until the number of uncovered samples (in the training set) are less than a user-specified threshold. The final product is an ordered discovered rule list that is used to classify the test data. For more details, readers are kindly requested to refer to [12]. Throughout this paper, we refer to the AntMiner [12] as “original AntMiner”.

2.2 AntMiner2 & AntMiner3

The extensions of the original AntMiner algorithm are proposed by Liu et al. in AntMiner2 [13] and AntMiner3 [14]. In AntMiner2, the authors used density estimation as a heuristic function instead of information gain (as was used by AntMiner). They concluded that the new and simpler heuristic function does the same job as was done by the complex one. Hence, AntMiner2 is computationally less expensive as compared with original AntMiner while offering comparable performance. In AntMiner3, the authors used a different pheromone update method. They update and evaporate the pheromone of only those conditions that are present in the constructed rule and do not evaporate the pheromones of unused conditions. In this way, the exploration of the search space is encouraged.

2.3 AntMiner+

Martens et al. [7] proposed a Max-Min Ant system-based algorithm (AntMiner+) that differs from the previously proposed AntMiner’s in several aspects. A different search environment for the ants is designed. Class label of rules is committed prior to the construction of antecedent part. They incorporated the ability to form interval rules (for ordinal attributes). Additionally, the setting of two important user-defined ACO parameters (alpha and beta) is automated. Only the iteration-best ant is allowed to update the pheromone, the range of the pheromone trail is limited within an interval, and a new rule quality measure is used [7].

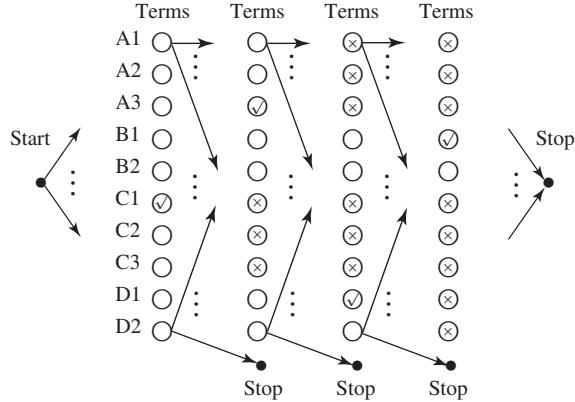
2.4 Unordered AntMiner

The unordered rule list AntMiner is proposed in the work of Smaldon et al. [15]. As per new algorithm, there is no need to apply the rules in order of their discovery. Therefore, the approach can be considered as more modular in terms of interpreting the generated rules independently. The learning flow of original AntMiner algorithm is modified such that an extra FOR-EACH loop is added as the outermost loop which iterates over the values present in the domain of class attribute. Unlike original AntMiner, in unordered AntMiner, all the ants construct the rule for the current class label (selected in the outer most FOR-EACH loop) and then choose only those terms that would result in optimum level of accuracy. So, in unordered AntMiner, class value is selected before starting the rule construction process and does not change afterward (not even during pruning stage). The samples covered by the best rule are removed from the training set and this learning process continues until the length of the positive samples (the samples with the class value equal to the current selected class value) in the training set is greater than a user-defined parameter (`max_uncovered_cases`). In the next iteration of the outermost FOR-EACH loop, the training set is again reinstated.

Since the class value (for which rule construction is being performed) is already known, a new heuristic function based on Laplace-corrected confidence value is used as

$$\eta_{i,j}(t) = \frac{|\text{Value}_{i,j}, k| + 1}{|\text{Value}_{i,j}| + n_C}, \quad (1)$$

where η is the heuristic value at time t , $|\text{Value}_{i,j}|$ is the total number of training samples having $\text{Value}_{i,j}$ (i.e., j th value in the domain of the i th attribute), ‘ k ’ is the current selected class label, and ‘ n_C ’ represents the total number of class labels present in the given data set.

**Figure 1** The UAntMinerC search.

Classification of unseen data in unordered AntMiner is definitely different. Rules are applied to the cases without considering their order in the discovered rule list. In case if antecedent part of more than one rule matches the given test sample (a conflict occurs, i.e., which rule should we fire next?), two conflict resolution strategies are used. First, high-quality rule decision is preferred. Second, CN2 conflict resolution-based decision is made.

3 Proposed method

In this paper, a new ant-based classification technique is proposed, called UAntMiner-C (Unordered AntMiner using Correlation of terms), which is an extension to AntMiner-C [16]. We begin with a general description of the algorithm followed by the discussion about the heuristic function, pheromone update, rule pruning, and other related stages.

3.1 General description

The intention is to build, incrementally set of classification rules, which is represented as follows:

$$\text{IF } \langle T_1 \text{ AND } T_2 \text{ AND } \dots \rangle \text{ THEN } \langle \text{class} \rangle.$$

Each term T_i (in the antecedent part of the rule) is an attribute-value pair related to an operator. In this work, “=” is used as the only relational operator for relating an attribute to one of its domain value. A term, for example, “Weather = Sunny” has three parts: (1) Operator “=”, (2) the attribute name “Weather” and (3) a value in the domain of attribute “Weather”, that is, “Sunny”. Due to the usage of only “=” operator, all continuous-valued attributes in the input data set are discretized before presenting to the proposed classifier.

Search space: search space is designed for producing candidate solution. This space can be defined with the aid of data set provided as input. The search space can be considered as a directed acyclic graph, where the vertices are used to represent the terms (attribute-value pair) extracted from the data set. The exact topology of search space is given in Figure 1. The objective of the ants is to discover IF-THEN unordered rules by visiting the vertices. The selection of vertices (during the tour of an ant) is based on heuristic (discussed later) and pheromone values associated with the edges. The statistic about the total number of terms/ vertices, present in the data set, can be calculated as

$$\text{Total_terms} = \sum_{n=1}^a b_n, \quad (2)$$

where ‘ a ’ represents the size of a given data set in terms of attributes (without considering the target attribute) and ‘ b_n ’ is the number of values that the attribute ‘ A_n ’ can have. The search space is defined in such a way that an ant can pick any of the vertex and there is no fixed ordering in which the vertices are required to be visited. This is to discourage the biasness of vertices ordering introduced in AntMiner+.

For example, a problem's search space is shown in Figure 1. The search space is represented as a DAG and constructed with the help of input data set. The search space contains 10 vertices in a column where each vertex corresponds to a term. The number of columns that contain the vertices (representing the terms) will be equal to the number of attributes present in the given data set. A_1 , A_2 , and A_3 are the labels assigned to the vertices (terms) related to attribute A . The data set consists of four attributes ($A-D$) for which the search space is constructed. Attributes (A , B , C , and D) have (3, 2, 3, and 2) values in their respective domains. For constructing a rule, the ant is placed on vertex labeled "Start." The ant moves from left to right and can visit only a single vertex in a column. As a vertex corresponds to a term (attribute = value pair), the antecedent of the rule is actually being constructed based on the selection of vertices during the tour of an ant. After a term has been selected (i.e., a vertex is visited), the attribute related to that term is flagged as used. During the tour of an ant, once an attribute is flagged used, no other term related to that attribute can be selected further. This restriction is necessary, as we cannot allow conditions of the type "Weather = Cloudy AND Weather = Sunny" (which doesn't make sense both at the same time). As we can see in Figure 1 that ant has selected C_1 (after making a move from the start vertex) in the first column; attribute C is thus flagged as used, and all the terms corresponding to attribute C (C_1 , C_2 , and C_3) are flagged prohibited (shown with 'x' symbols in columns (2–4) of the search space). Next, the ant has chosen the vertices with the labels A_3 , D_1 , and B_1 (in order), in turn the vertices that become prohibited are shown with 'x' symbols. The process of rule construction (for an ant) is terminated if any of vertices with label "Stop" is selected by the ant. This process also terminates prematurely after insertion of a term (into the antecedent of a rule), such that the rule covers only those training samples with homogeneous class labels (it is to be noted that the consequent of a rule or the class label is selected before constructing the rule antecedent part described later).

Suppose ' A ' represents the number of attributes and ' T ' represents the number of total terms present in the data set. The worst case space complexity of the search space (in terms of number of edges to be stored in memory) can be calculated as $O(A, T) = T_1 + (A - 1) * (T * T) + T_2$, where T_1 shows the number of edges originating from 'start' vertex toward the nodes in the first layer (a column of vertices). T_2 shows the number of edges originating from the vertices in the second last layer and which are connected to the 'stop' vertex (in the last layer). It may please be noted that not all the vertices in a layer are connected to all the vertices in the next layer, and thus reduces the actual space complexity of the search space. The complexity of the search space increases substantially if the value of T is large. However, the chance of such a case is low due to the usage of discretized data sets containing only categorical attributes (where each attribute contains only a few values in its domain).

AntMiner [12] uses a mesh topology where all the terms are connected to each other. An ant can freely move within the search space without any particular ordering imposed on attribute selection. The search environment used by AntMiner+ [7] is different in several ways: each of the ants is required to choose one of the classes (majority class is excluded from this set) before constructing its rule, the selection of alpha and beta parameters are made part of the search process, the dimensions (i.e., attributes) are visited in a fixed order but an ant may choose to ignore a dimension, and for ordinal variables a mechanism that enables ants to include interval conditions in their rules is introduced.

A general description of the UAntMiner-C algorithm is shown in Algorithm 1. In the outermost loop, a class label is selected for which the rules will be constructed. The Pos_Cases(TrainingSet) is a function that returns the total number of cases in the training set with the class selected in the outermost FOR-EACH loop. The basic structure of the UAntMiner-C algorithm is that of AntMiner [12] and the flow of learning is based on unordered AntMiner [15]. Iterating over each class value in the class attribute domain, the algorithm is presented with all the training data. Several rules are constructed within an iteration of REPEAT-UNTIL loop, and the one with the best quality is selected to be included in the discovered rule list. The best rule correctly classifies (i.e., rule antecedent and consequent part if matched with the training patterns) a subset of training samples which are removed from the training set.

This process continues until the length of the positive cases (the cases with class value equal to current selected class) in the training set is greater than a user-defined parameter max_uncovered_cases. The training set is then again reinstated for the next class value selected in (outermost FOR-EACH loop) so

Algorithm 1: The UAntMinerC algorithm

```

1: DiscoveredRuleList={}; /* rule list is initialized as empty list */
2: FOR EACH Class value in class attribute domain
3:   (TrainingSet) = {all training samples};
4:   WHILE (Pos_Cases(TrainingSet) > Max_uncovered_samples)
5:     t = 1; /* counter for ants */
6:     rcc = 1; /* counter for rule convergence test */
7:     Construct the search space for reduced training set;
8:     Initialize pheromone values on all edges;
9:     Calculate heuristic values for all edges;
10:    REPEAT
11:      Send an Ant which constructs a classification rule Rt for the selected class;
12:      Assess the quality of the rule and update the pheromone of all trails;
13:      IF (Rt == Rt-1) /* update convergence test */ THEN
14:        rcc = rcc + 1;
15:      ELSE
16:        rcc = 1;
17:      END IF
18:      t = t + 1;
19:    UNTIL (t ≥ No_of_ants) OR (rcc ≥ No_rules_converg)
20:    Choose the best rule Rbest among all rules Rt constructed by all the ants;
21:    Prune the best rule Rbest;
22:    Add the pruned best rule Rbest to DiscoveredRuleList;
23:    Remove the training samples correctly classified by the pruned best rule Rbest;
24:  END WHILE
25:  Add a default rule in the DiscoveredRuleList;
26: END FOR
27: Prune the DiscoveredRuleList (Optional);

```

that a maximal number of negative samples are available to the algorithm. In the original AntMiner, max_uncovered_cases referred to all the cases in the training set, rather than positive cases only, as interpreted in the UAntMiner-C.

The algorithm terminates when the outer FOR-EACH loop iterates over all the possible class values. The algorithm produces an unordered rule list that can be used for predicting the class labels of test samples.

3.2 Rule construction

To construct a rule, first, a class label is committed and then rules antecedent part is constructed. The entire procedure is discussed as follows.

- **Class commitment.** Ant selects a class label (as the consequent of IF-THEN rule) prior to the construction of rule antecedent part. This beforehand selection is necessary to use the heuristic function discussed later. This selection is performed in the outermost FOR-EACH loop of the algorithm. The selected class label remains fixed for a swarm run during WHILE loop.

Method of class assignment is different as compared with AntMiner and AntMiner+. AntMiner first construct the rule and then a class is assigned. The rule is assigned with the majority class of the current training samples that it covers. For AntMiner+, an ant assigns a class label to a rule before the construction of its antecedent part (ignoring the class in majority).

- **Pheromone initialization.** The pheromone values on all the edges in the search space are initialized at the start of WHILE loop. Pheromone value on edge(i, j), that is, the edge connecting term ‘i’ and term ‘j’ is initialized using the following:

$$\tau_{ij}(t=1) = \frac{1}{\sum_{n=1}^a x_n b_n}, \quad (3)$$

where ($t = 1$ means the start of the WHILE loop), a represents the size of data set in terms of attributes, and b_n is the length of unique values that A_n can have in its domain. x_n is 0 if term_i is related to attribute A_n , otherwise, it will be 1. No pheromone is placed over the edges between the terms relating

to same attribute. Based on the fact that the pheromone concentration for all the edges is same (except zero-valued pheromone edges), no historical information (in terms of pheromone values) is available to guide the search process of 1st ant.

- **Term selection.** The antecedent of a rule that an ant is constructing usually consists of several terms. The selection of terms must be done intelligently so that the generated rule set becomes a better representative of the relationship between predictor attributes and target attribute. The terms are added incrementally one by one to the antecedent part of the rule that can be conceived as a partial trail of the ant. For a particular ant trail, no more than one term can be selected related to an attribute A_n . The selection probability of a term depends on its pheromone and heuristic values as compared with the other candidate terms and is calculated using the following:

$$P_{ij}(t) = \frac{\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(s)}{\sum_{j=1}^{\text{Total_terms}} x_j \{\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(s)\}}, \quad (4)$$

The pheromone value presents on the edge between the two terms T_i and T_j is represented with $\tau_{ij}(t)$ for the t th ant (the ant for which selection is being made), and $\eta_{ij}(s)$ specifies the heuristic value associated between terms T_i and T_j during s th iteration of WHILE loop. For all the terms, term_j , which are not flagged prohibited, x_j will be set to 1, otherwise it is set to 0.

- **Heuristic function.** Heuristic value is another important learning dynamic of ACO algorithm. For deciding whether to select a particular term, a term is evaluated based on its heuristic and pheromone values. Heuristic function can guides the swarm toward the right direction. A correlation-based heuristic function is used. Correlation between the term that is just added to the partial tour of an ant with other candidate terms is calculated as follows:

$$\eta_{ij} = \frac{|\text{term}_i, \text{term}_j, \text{class}_k|}{|\text{term}_i, \text{class}_k|} \cdot \frac{|\text{term}_j, \text{class}_k|}{|\text{term}_j|}, \quad (5)$$

where term_i is the term that is recently added by the ant to its partial tour (representing a partial rule) and term_j is one of the candidate terms that is being evaluated to be selected next. The remaining training samples (which are not covered by any of the rules discovered so far) associated with class label k and that contain term_i and term_j are represented with $|\text{term}_i, \text{term}_j, \text{class}_k|$. $|\text{term}_i, \text{class}_k|$ are the training samples that are associated with class label k and contain term_i as one of its predictor attribute value. For the selection of first column of vertices (directly connected to outgoing edges from ‘start’ vertex), the following Laplace-corrected confidence [10] is used to calculate the heuristic value for a C -class problem:

$$\eta_{\text{Start},j} = \frac{|\text{term}_j, \text{class}_k| + 1}{|\text{term}_j| + m}, \quad (6)$$

‘ C ’ represents the number of class values in the domain of target attribute. The heuristic function penalize those terms that would lead to overfitting by producing very specific rules (poor generalization is expected). As an example, let’s assume that there exists a term which appears in only a single training instance with the class label equivalent to the class label chosen by an ant. For such a term, the confidence will turn out to be 1 without applying (6). However, with the usage of (6), the confidence becomes equal to 0.7 for a binary class problem.

3.3 Quality of a rule

ACO requires the evaluation of generated solutions and uses them as a feedback for generation of new solutions. The correctness of this evaluation has direct consequences on the search process. For classification rule discovery problem, we would like to have rules with high confidence and large coverage. The quality Q of a rule is measured by adding its confidence and coverage [15]:

$$Q = \frac{\text{TP}}{\text{Covered}} + \frac{\text{TP}}{N}, \quad (7)$$

where TP represents the set of training examples that are correctly classified by the discovered rule. Covered represents the set of training examples that match the antecedent part of the rule and N is taken equivalent to the length of yet uncovered training examples.

3.4 Pruning of rule

Pruning is performed the same way as done in the original AntMiner. One term is temporarily deleted (at a time) from the antecedent of the rule and its quality is checked. If doing so increases the quality of rule, change to the rule is committed, otherwise there is no need to remove that term. Repeat this for all the remaining terms until either single term is left in the pruned rule or no more improvement is possible. The only difference in UAntMiner-C is: the consequent part of the rule does not change. After removing each term, speculatively, only the rule quality for the current positive class is computed.

3.5 Rules conflict

In UAntMiner-C, there may be more than one rule that might cover (i.e., the rule antecedent part matches with the attribute values of the unseen data instance) an unseen data instance during classification of testing data which will result in a conflict. One of the following scenarios may occur [15]:

1. If none of the discovered rules cover the test case, the case is assigned with the majority class in the training data set.
2. If only one of the discovered rules covers the test case, the case is assigned with the class predicted by that rule.
3. If more than one of the discovered rules cover the test case (e.g., R_1, \dots, R_n), but all those rules predict the same class, the case is assigned with the class predicted by the covering rules.
4. If more than one of the discovered rules cover the test case (e.g., R_1, \dots, R_n), but the rules do not predict the same class, a rule conflict strategy is used to determine which class should the case be assigned with. For the time being, just to keep the discussion simple, we assume that the classification problem being solved contains two classes only (Yes and No). The discussion is valid and extendible easily for multiclass problem.

To handle the conflicts, the following strategies are used:

- **High quality rule.** The predicted class of the conflicting rules (from R_1, \dots, R_n) which has the high quality ‘Q’ is assigned to the test case. This is compactly defined below, where $Q(R_i)$ represents the quality of R_i and argmax selects the maximum value from the values passed to it as arguments.

$$\text{HQ_Rule} = \arg \max(Q(R_1), \dots, Q(R_n)), \quad \text{Test_Case_Class} = \text{ClassOf}(\text{HQ_Rule}).$$

- **CN2.** The class frequencies of the conflicted rules (R_1, \dots, R_n) are calculated using the following.

$$C(\text{Yes}) = \sum_{i=1}^n \text{covers}(R_i, \text{Yes}), \quad C(\text{No}) = \sum_{i=1}^n \text{covers}(R_i, \text{No}),$$

where $\text{covers}(R_i, C_j)$ means the number of cases of class C_j that are covered by rule R_i .

$$\text{Test_Case_Class} = \arg \max \left(C_j \in \text{Classes} \sum_{i=1}^n \text{covers}(R_i, C_j) \right).$$

So, the class for which the class frequency value of the conflicted rules is maximum is assigned to the test case.

- **Naive Bayes.** The Naive Bayes theorem is also used to resolve the conflicts during classification of an unseen pattern. The Bayes theorem is given as

$$P(C_i | R_1 \wedge \dots \wedge R_n) = P(C_i) \frac{P(R_1 \wedge \dots \wedge R_n | C_i)}{P(R_1 \wedge \dots \wedge R_n)}.$$

Since the probability $P(R_1 \wedge \dots \wedge R_n)$ does not affect the relative order of different hypotheses [17], we can assume naively that $P(R_1 \wedge \dots \wedge R_n | C_i) = P(R_1 | C_i) \dots P(R_n | C_i)$. The classification is done as

$$\text{Test_Case_Class} = \arg \max \left(C_j \in \text{Classes} \left(P(C_j) \prod_{i=1}^n P(R_i | C_j) \right) \right),$$

where $P(C_j)$ represents the probability of the total number of training cases having the class label C_j and $P(R_i|C_j)$ is the probability of the total number of training cases that have the class label C_j and are covered by the rule R_i . Note that there is a problem if a rule from the set of conflicting rules (R_1, \dots, R_n) does not cover any instances of a particular class, say for example, C_h . In this case, the chances to select class C_h will become 0 even if there exist several other conflicting rules that cover the instance with high probability for class C_h . Laplace-1 correction [18] as defined below is used to overcome this limitation of Naive Bayes conflict resolution strategy:

$$\text{if } P(R_i|Cj) = 0 \text{ then set it to } \left[\frac{N+1}{N+k} \text{ Laplace-1 Correction} \right],$$

where N is the total number of training cases covered by the rule R_i (only antecedent part) and k is the total number of class labels present in the data set.

4 Differences with existing ACO-based classifiers

Following UAntMiner [15], the proposed classifier (UAntMiner-C) is the second attempt to generate unordered rule list using ACO. In follows, main differences between the proposed algorithm and existing ACO-based classifiers are summarized.

Differences with AntMiner, UAntMiner, AntMiner2, and AntMiner3, where AntMiner-C has	Differences with AntMiner+, where AntMiner-C has
<p>1. Prior selection of class label, whereas the above-mentioned competitors (except UAntMiner) select the class label after the construction of rule antecedent part.</p> <p>2. A novel correlation-based heuristic function.</p> <p>3. A term has multiple heuristic values by considering its correlation with other terms. In the above three competitors, each term has only one value of heuristic function.</p> <p>4. Pruning of only best rule in an iteration of FOR-EACH loop, whereas on the other hand, all the rules are pruned in the competitors which increases the computational complexity.</p> <p>5. The termination of rule construction is based on only positive training samples currently present in the training set. For the competitors (except UAntMiner), <code>max_uncovered_cases</code> parameter is used and all the samples in the current training set are considered.</p> <p>6. An asymmetric pheromone matrix to decrease the chance of premature convergence.</p> <p>7. A different pheromone update equation (an exception is AntMiner3 which has the same update equation).</p> <p>8. A different method of pheromone normalization.</p> <p>9. A different equation for assessing the quality of rule which is also used for rule pruning.</p>	<p>1. A different search space.</p> <p>2. The capability of handling only categorical attribute, whereas AntMiner+ can handle ordinal valued attributes as well.</p> <p>3. A novel correlation-based heuristic function.</p> <p>4. For AntMiner+, state transition (i.e., selection of next term) and pheromone update are according to the Ant System [5].</p> <p>5. A different method of class selection which is done only once and subsequently the class label remains fixed for all the ants in an iteration. For AntMiner+ during an iteration, the ants construct the rules for different class labels.</p> <p>6. Does not exclude the majority class from the class selection choices.</p> <p>7. For AntMiner+, early stopping criterion may be used and which is settled based on the user input.</p> <p>8. UAntMiner-C has several iterations of REPEAT-UNTIL loop for the extraction of a best rule. Each iteration consists of only one ant run. The best rule is pruned after exiting from the REPEAT-UNTIL loop. On the other hand, for AntMiner+ the REPEAT-UNTIL loop is executed until the pheromone values on one path converge to t_{\max} (maximum pheromone value) and for all other paths are t_{\min}. There are multiple ants run per iteration (1000 ants are used in the experiments reported in [7]). The best solution from each iteration is pruned. In other words, several rules are pruned before exiting from the REPEAT-UNTIL loop.</p> <p>9. Unordered rule list discovery.</p>

5 Results

In our experiments, a suite of 26 data sets, that are collected from UCI ML repository¹⁾ is used and is presented in Table 1. The selected data sets are of different complexities and are commonly used

1) Hettich S, Bay S D. The UCI KDD archive. University of California, Irvine, 1996. <http://kdd.ics.uci.edu>.

Table 1 Data set characteristics

Date set	Short name	Attribute	Instance	Class
Balance-scale	bs	5	625	3
Breast-cancer-w	bew	10	699	2
Car	car	7	1728	4
Credit-aus	Crd-aus	16	690	2
Credit-german	Crd-ger	21	999	2
Dermatology	Derm	35	366	6
Ecoli	Ecoli	8	336	8
Glass	glass	10	214	7
Haberman	Haberman	4	307	2
Hayes_Roth	Hayes_Roth	6	132	3
Heart	heart	14	270	2
Hepatitis	hepatitis	20	155	2
House_Votes	House_Vot	18	435	2
Iris	iris	5	150	3
Mammo_masses	Mamm_m	6	961	2
Pima_Indians	Pima_Ind	9	769	2
Segmentation	Segm	20	210	7
Spect_Heart	Spect_H	23	267	2
Tae	tae	6	151	3
Tic-tac-toe	TTT	10	958	2
Transfusion	Trans	5	749	2
Vehicles	Veh	19	282	4
WDBC	WDBC	31	569	2
Wine	wine	14	178	3
Zoo	Zoo	18	101	7
IonoSphere	IonoSphere	34	351	2

Table 2 Parameter values used in experiments

Parameter value	[AntMiner, UAntMiner, UAntMiner-C]
Swarm size	[1000, 1000, 1000]
Max. uncovered samples	[10, 5, 5]
Evaporation rate	[0.15, 0.15, 0.15]
No. of rules converged	[10, 10, 10]
Minimum cases per rule	[10, 10, 1]

as benchmarks for evaluating the performance of classification algorithms. All the continuous-valued attributes are discretized (using an unsupervised filter of Weka-3.4 machine learning tool [10]) before submitting to the proposed algorithm.

The results of UAntMiner-C are compared with AntMiner and Unordered AntMiner. Both of these algorithms and UAntMiner-C are implemented in Visual Studio 2008 C#. The experiments are executed on a machine that has 2.8 GHz processors with 1 GB of main memory. Performance indicators include predictive accuracy, number of rules, and number of terms per rule. The experiments are conducted with 10-fold cross-validation.

The setting of user-defined parameters is summarized in Table 2 for all the three competing algorithms with alpha and beta (set to 1). These values have been chosen because they have been used in earlier versions of AntMiner as reported in the literature [7, 12–15, 19] and we make no claim that these values produce the optimal results.

Table 3 shows the mean classification accuracy in percentage for all the algorithms with the values after (+/-) denoting the standard deviation based on 10-fold cross-validation. The results of unordered

Table 3 Accuracy rate (HQ, CN2, & Naive)

Conflict resolution strategies = >>>		High quality rule			CN2		Naive	
Date set	AntMinr	UAntMinr	UAntMinerC	UAntMiner	UAntMinerC	UAntMiner	UAntMinerC	
Balance-scale	73.13±1.76	61.92±2.15	89.76±1.62 (+)	80.34±1.76	90.72±1.39 (+)	34.57±2.88	88.00±1.95 (+)	
Breast-cancer-w	92.84±1.37	89.27±1.80	96.00±0.63 (+)	83.98±2.21	96.00±0.63 (+)	85.41±2.31	96.00±0.63 (+)	
Car	84.09±0.83	49.71±1.59	94.16±0.73 (+)	79.63±1.05	94.10±0.75 (+)	31.02±1.17	92.65±0.80 (+)	
Credit-aus	85.22±1.22	85.36±0.82	86.09±1.06	84.06±1.01	84.06±2.09	78.55±1.48	83.33±2.17 (+)	
Credit-german	74.18±1.02	58.15±2.09	75.38±2.71 (+)	72.58±1.59	75.18±2.67	55.76±1.44	74.88±2.61 (+)	
Dermatology	95.35±1.16	83.83±3.03	87.42±1.23	84.65±1.58	87.97±1.10 (+)	76.20±1.57	86.61±1.18 (+)	
Ecoli	69.93±2.87	55.02±2.52	77.88±3.08 (+)	63.90±4.42	77.88±3.08 (+)	33.01±2.04	77.89±2.87 (+)	
Glass	50.84±3.33	35.89±3.67	67.34±3.07 (+)	49.55±5.41	70.15±3.48 (+)	25.56±4.31	64.03±3.26 (+)	
Haberman	70.32±3.20	69.39±3.54	81.45±2.88 (+)	69.32±2.37	80.81±2.85 (+)	55.59±5.07	80.81±2.85 (+)	
Hayes_Roth	65.05±3.73	65.93±2.53	82.53±3.81 (+)	62.97±3.56	85.60±3.30 (+)	35.71±3.72	77.14±4.61 (+)	
Heart	81.48±1.10	77.78±2.92	84.81±2.24 (+)	84.81±2.56	72.22±3.75 (-)	71.85±3.03	76.30±3.32	
Hepatitis	85.88±2.47	88.38±2.54	87.25±2.26	89.63±2.25	81.29±4.16 (-)	89.67±2.61	80.04±3.32 (-)	
House_Votes	97.25±0.82	95.19±1.46	97.02±0.96	92.43±1.41	96.10±1.27 (+)	91.96±1.34	96.79±1.42 (+)	
Iris	88.67±2.99	85.33±2.78	94.67±1.66 (+)	83.33±2.48	90.67±2.27 (+)	75.33±5.07	90.00±2.05 (+)	
Mamm_m	81.78±1.54	79.91±0.93	83.97±1.08 (+)	79.08±2.62	84.18±1.49 (+)	66.80±2.32	83.35±1.44 (+)	
Pima_Indians	68.52±1.71	67.87±1.81	81.92±1.87 (+)	69.56±1.68	82.06±1.87 (+)	61.38±1.42	81.67±1.89 (+)	
Segmentation	61.43±2.50	57.14±2.35	73.33±1.77 (+)	55.24±2.94	73.33±1.77 (+)	43.33±2.97	72.38±1.71 (+)	
Spect_Heart	83.87±3.04	76.35±2.39	82.82±2.29 (+)	86.50±2.25	81.32±2.39 (-)	76.01±1.89	80.57±2.50 (+)	
Tae	50.50±4.46	41.88±4.41	80.17±3.26 (+)	51.83±4.42	80.17±4.07 (+)	25.92±3.68	78.17±3.43 (+)	
Tic-tac-toe	72.13±1.85	61.79±1.14	97.90±0.97 (+)	69.00±1.53	97.80±0.97 (+)	70.67±2.91	97.70±0.86 (+)	
Transfusion	78.64±1.69	36.85±2.92	79.70±1.79 (+)	72.77±1.77	79.70±1.79 (+)	72.90±1.78	79.44±1.76 (+)	
Vehicles	55.63±2.69	56.31±3.93	75.95±1.77 (+)	46.77±2.94	76.66±1.71 (+)	31.19±2.06	75.60±2.15 (+)	
WDBC	87.53±1.21	79.62%±2.32	91.22±1.04 (+)	84.53%±1.22	83.48±5.18	79.97%±1.77	79.11±6.68	
Wine	92.65±2.38	94.38±1.17	99.44±0.56 (+)	92.68±1.68	97.22±2.23 (+)	91.57±2.52	98.27±0.88 (+)	
Zoo	88.18±1.93	88.09±3.60	90.09±2.11	85.00±3.73	90.09±2.98	87.36±3.69	90.09±3.33	
Ionosphere	82.29±2.29	74.86±2.55	87.71±2.83 (+)	81.43±1.96	87.43±2.80 (+)	68.86±2.54	87.43±2.80 (+)	

Table 4 Number of conflicts occurred

Data set	No. of conflicts for UAntMiner	# of conflicts for UAntMiner-C
bs	49.30±1.08	3.10±0.57
car	140.80± 2.12	9.30±0.82
wine	0.40±0.31	0.20±0.13
Zoo	1.10±0.38	0

AntMiner and UAntMiner-C are further compared based on three different conflict resolution strategies. For each of the conflict resolution strategy, under the results of UAntMiner-C, the presence of the symbols (+) or (-) in a cell indicates that the predictive accuracy of the corresponding version of UAntMiner-C is significantly better or worse than the predictive accuracy of the Unordered Rule Set Ant-Miner for the corresponding data set and conflict resolution strategy. Significant difference is declared if the standard deviation intervals of the two accuracies (for UAntMiner vs. UAntMiner-C for a particular data set with a particular conflict resolution strategy) do not overlap.

In all the tables, the best results in a row are highlighted. The (win/loss/tie) situation for all (26) data sets based on average accuracy considering the results in all conflict resolution strategies can be summarized as AntMiner (2/24/0), unordered AntMiner (1/25/0), and UAntMiner-C (23/3/0). It may also be noted that UAntMiner-C performance is consistent; producing almost the same average accuracies (less varying) for different conflict resolution strategies, whereas unordered AntMiner produces highly

Table 5 Average number of rules

Date set	AntMiner	UAntMiner	UAntMinerC
Balance-scale	11.50±0.22	16.50±0.27	59.90±1.00
Breast-cancer-w	9.50±0.31	12.40±0.27	18.50±0.54
Car	15.70±0.37	11.90±0.10	79.40±0.91
Credit-aus	6.50±0.22	8.20±0.20	21.90±1.80
Credit-german	9.20±0.33	13.10±0.28	55.20±4.01
Dermatology	7.70±0.15	7.10±0.10	20.40±0.64
Ecoli	11.70±0.30	17.90±0.23	34.60±0.56
Glass	7.60±0.43	11.10±0.31	25.30±0.45
Haberman	9.30±0.15	18.60±0.48	77.10±3.07
Hayes_Roth	5.80±0.29	9.00±0.26	29.00±0.70
Heart	5.10±0.10	6.60±0.16	15.00±0.80
Hepatitis	5.00±0.00	5.40±0.16	7.40±0.34
House_Votes	5.60±0.16	4.80±0.13	7.10±0.31
Iris	7.40±0.16	11.00±0.37	10.30±0.42
Mamm_m	7.10±0.46	14.40±0.16	34.40±2.03
Pima_Indians	8.40±0.27	13.30±0.33	55.10±4.37
Segmentation	8.50±0.17	12.20±0.25	21.50±0.31
Spect_Heart	6.50±0.17	8.40±0.16	11.70±0.62
Tae	7.20±0.47	11.90±0.18	17.50±0.65
Tic-tac-toe	7.10±0.23	7.10±0.10	29.90±2.57
Transfusion	6.50±0.34	9.10±0.10	16.60±2.57
Vehicles	8.30±0.15	10.90±0.10	36.40±0.43
WDBC	8.20±0.25	8.80±0.25	23.90±0.95
Wine	4.50±0.17	5.00±0.00	7.10±0.18
Zoo	6.00±0.00	7.60±0.27	7.00±0.15
IonoSphere	8.70±0.15	10.20±0.13	19.20±0.88

varying average accuracies under different conflict resolution strategies when tested on some data sets. For example, unordered AntMiner produces highly varying average accuracies using different conflict strategies for balance scale and car data sets. However, it performs consistent for wine, zoo, and some other data sets. UAntMiner-C is consistent because it learns the rules that are more robust and causes very few conflicts on the test data. On the other hand, the rules generated using unordered AntMiner result in high number of conflicts (for some data sets) and the average accuracies are thus more dependent on the conflict resolution strategy rather than the classification algorithm itself. For these data sets, numbers of conflicts that occurred during the testing phase (10-fold cross-validation) are summarized in Table 4.

In Tables 5 and 6, the results are presented for (average number of rules found \pm standard deviation) and (average number of terms per rule \pm standard deviation). These two parameters actually represent the comprehensibility of the classifier built after training. The UAntMiner-C is somewhat expensive in terms of simplicity of the rules for most of the data sets, but this is, of course, expected for some data sets to learn the complex relationships of the feature sets and the corresponding class label. In turn, it generates highly accurate predictions.

To compare UAntMiner-C with some well-established deterministic rule discovery methods, we have selected CBA [20] (Classification based on Predictive Association Rules), CMAR [21] (Classification based on Multiple Association Rules), and CPAR [22] (Classification Based on Associations). These three classifiers are the instances of association rule mining. We have used the software²⁾ for the experiments of these three competitors with (support value = 1% and confidence value = 70%). The data sets are required to be preprocessed to be handled by these association rule-based classifiers. For this purpose, the above mentioned tool is used. The comparison summary based on % average accuracy and average

2) <http://cgi.csc.liv.ac.uk/~frans/KDD/Software/>.

Table 6 Average number of terms per rule in the rule list

Date set	AntMiner	UAntMiner	UAntMinerC
Balance-scale	0.92±0.01	1.00±0.02	3.11±0.01
Breast-cancer-w	1.03±0.02	1.05±0.02	2.19±0.03
Car	1.39±0.04	1.15±0.02	4.68±0.02
Credit-aus	1.55±0.07	2.30±0.05	3.21±0.14
Credit-german	1.74±0.09	2.05±0.07	3.81±0.05
Dermatology	2.70±0.13	4.59±0.12	3.19±0.07
Ecoli	1.35±0.02	1.38±0.01	2.76±0.02
Glass	1.26±0.05	1.90±0.07	3.01±0.04
Haberman	0.89±0.00	0.95±0.00	2.87±0.02
Hayes_Roth	1.06±0.03	0.95±0.03	3.25±0.03
Heart	1.11±0.04	1.58±0.08	2.89±0.09
Hepatitis	1.42±0.05	2.21±0.07	2.73±0.09
House_Votes	1.40±0.08	0.92±0.06	2.82±0.08
Iris	0.86±0.00	0.92±0.01	1.88±0.01
Mamm_m	1.00±0.03	1.74±0.03	2.90±0.01
Pima_Indians	0.91±0.03	1.17±0.03	3.10±0.04
Segmentation	1.46±0.06	2.01±0.04	2.68±0.03
Spect_Heart	2.11±0.10	2.31±0.05	4.04±0.07
Tae	0.97±0.02	1.22±0.03	2.16±0.04
Tic-tac-toe	0.93±0.05	0.86±0.00	4.22±0.03
Transfusion	0.88±0.03	1.03±0.02	2.05±0.19
Vehicles	1.29±0.05	1.44±0.02	3.09±0.02
WDBC	0.89±0.01	1.24±0.07	2.48±0.05
Wine	1.11±0.04	1.46±0.04	2.27±0.03
Zoo	1.13±0.06	1.65±0.05	2.60±0.06
Ionosphere	0.90±0.01	0.99±0.03	2.49±0.03

number of rules for 10-fold experiments is presented in Table 7. For both of these evaluation measures, UAntMiner-C has resulted in best results considering all the data sets.

Based on different parameter settings of support and confidence, we found the competitors to perform inconsistently. For example, CMAR and CBA have resulted in zero percentage accuracy for data sets (glass, hayes-roth, image-segmentation, tae, and vehicle) when parameter setting was (support = 20% and confidence = 80%). We have observed that the results of these three classifiers can be very fluctuating for different settings of support and confidence values. On the other hand, ACO-based classification algorithms are less dependent on the parameter setting as shown in [16].

6 Conclusion

This article has presented an extension to the original AntMiner for unordered IF-THEN rule induction. The proposed method is an unordered rule set generator for classification task which helps in interpreting a rule in the rule list independently, that is, without being associated with the rules preceding it in the rule list and therefore considered more modular and simple approach as compared with its ancestors. As shown in the results, the adaptive reduction of the search space and information about the correlation of attributes in UAntMiner-C help it perform significantly better than the earlier versions of the AntMiner based on different performance dimensions. Among three conflict strategies, UAntMiner-C with HQ rule conflict strategy has significantly outperformed its competitors, keeping in view the average (%) predictive accuracy.

One observation from the results is that the rule conflict strategies play a vital role in the predictive accuracy when the number of conflicts generated by the rule list is very high. As a future direction, new

Table 7 Comparison with CBA, CMAR, and CPAR classifiers

Data set	% Average accuracy (10-fold)				Average rules (10-fold)			
	UAM-C	CBA	CPAR	CMAR	UAM-C	CBA	CPAR	CMAR
Balance-scale	89.76	84.81	72.07	88.65	59.90	67.7	31.0	107.3
Breast-cancer-w	96.00	95.55	96.31	96.13	18.50	62.6	14.3	288.1
Car	94.16	83.51	80.11	80.55	79.40	19.2	54.1	156.7
Credit-aus	86.09	86.23	85.51	86.67	21.90	145.5	21.4	677.6
Credit-german	75.38	47.71	65.67	73.17	55.20	130.2	46.0	1478.3
Dermatology	87.42	81.71	89.65	93.97	20.40	28.8	41.1	156.0
Ecoli	77.88	69.35	70.03	72.34	34.60	78.3	68.3	429.9
Glass	67.34	58.92	59.11	60.74	25.30	70.6	45.7	334.6
Haberman	81.45	71.12	68.90	67.20	77.10	45.4	20.4	15.1
Hayes_Roth	82.53	83.13	80.99	83.85	29.00	18.3	11.6	76.3
Heart	84.81	17.08	74.44	79.96	15.00	3.4	22.1	320.8
Hepatitis	87.25	32.38	84.00	77.21	7.40	2.1	17.9	140.1
House_Votes	97.02	88.13	95.39	95.42	7.10	31.4	6.2	182.5
Iris	94.67	87.33	96.00	95.33	10.30	15.5	9.8	82.9
Mamm_m	83.97	79.27	78.96	79.90	34.40	52.3	17.6	203.0
Pima_Indians	81.92	71.46	73.05	75.12	55.10	177.4	43.2	618.4
Segmentation	73.33	90.69	87.36	84.84	21.50	120.6	111.7	946.3
Spect_Heart	82.82	69.93	68.27	72.55	11.70	86.3	9.3	348.8
Tae	80.17	51.33	49.34	52.00	17.50	40.3	27.6	150.8
Tic-tac-toe	97.90	97.00	65.39	98.95	29.90	9.0	11.3	182.3
Transfusion	79.70	53.81	62.06	75.95	16.60	25.1	16.4	36.9
Vehicles	75.95	68.34	65.49	64.91	36.40	161.3	93.9	1316.6
WDBC	91.22	36.02	90.81	94.38	23.90	3.2	18.2	323.9
Wine	99.44	22.61	93.29	94.84	7.10	2.8	15.7	146.8
Zoo	90.09	40.64	94.09	94.18	7.00	2.0	16.0	35.4
Ionosphere	87.71	18.50	86.04	92.01	19.20	2.0	24.3	263.6
Average >>>	85.62	64.87	78.17	81.95	28.52	53.9	31.35	346.8

rule conflict strategies may be designed that performs robustly for different classification problems. It will also be an interesting research direction to see how a parameterless optimization algorithm called TLBO [23] (Teaching-Learning-Based Optimization) can be used to solve the subject problem.

References

- 1 Engelbrecht A P. Computational Intelligence: an Introduction. 2nd ed. John Wiley & Sons, 2007
- 2 Engelbrecht A P. Fundamentals of Computational Swarm Intelligence. John Wiley & Sons, 2005
- 3 Kennedy J, Eberhart R C, Shi Y. Swarm Intelligence. Morgan Kaufmann/Academic Press, 2001
- 4 Dorigo M, Sttzle T. Ant Colony Optimization. Cambridge: MIT Press, 2004
- 5 Dorigo M, Maniezzo V, Colorni A. Ant system: optimization by a colony of cooperating agents. *IEEE Trans Syst Man Cybern Part B-Cybern*, 26, 1996: 29–41
- 6 Dorigo M, Gambardella L M. Ant colony system: a cooperative learning approach to the travelling salesman problem. *IEEE Trans Evol Comput*, 1, 1997: 53–66
- 7 Martens D, de Backer M, Haesen R, et al. Classification with ant colony optimization. *IEEE Trans Evol Comput*, 11, 2007: 651–665
- 8 Abraham A, Grosan C, Ramos V, eds. Swarm intelligence in Data Mining. Berlin/Heidelberg: Springer-Verlag, 2006
- 9 Han J, Kamber M. Data Mining: Concepts and Techniques. 2nd ed. Amsterdam: Morgan Kaufmann, 2006
- 10 Witten I H, Frank E. Data Mining: Practical Machine Learning Tools and Techniques. 2nd ed. Burlington: Morgan Kaufmann, 2005
- 11 Duda R O, Hart P E, Stork D G. Pattern Classification. Hoboken: Wiley, 2000
- 12 Parpinelli R S, Lopes H S, Freitas A A. Data mining with an ant colony optimization algorithm. *IEEE Trans Evol*

- Comput, 2002, 6: 321–332
- 13 Liu B, Abbass H A, McKay B. Density-based heuristic for rule discovery with ant-miner. In: Proceedings of 6th Australasia-Japan Joint Workshop on Intelligent and Evolutionary System, Canberra, 2002. 180–184
 - 14 Liu B, Abbass H A, McKay B. Classification rule discovery with ant colony optimization. In: Proceedings of IEEE/WIC International Conference on Intelligent Agent Technology, Halifax, 2003. 83–88
 - 15 Smaldon J, Freitas A A. A new version of the Ant-Miner algorithm discovering unordered rule sets. In: Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation. New York: ACM, 2006. 43–50
 - 16 Baig A R, Shahzad W. A correlation-based ant miner for classification rule discovery. *Neural Comput Appl*, 2012, 21: 219–235
 - 17 Lindgren T. Methods for rule conflict resolution. In: Proceedings of 15th European Conference on Machine Learning, Pisa, 2004. 262–273
 - 18 Kohavi R, Becker B, Sommerfield D. Improving simple bayes. In: Proceedings of 9th European Conference on Machine Learning, Prague, 1997
 - 19 Khan S, Baig A R, Shahzad W. A novel ant colony optimization based single path hierarchical classification algorithm for predicting gene ontology. *Appl Soft Comput*, 2014, 16: 34–49
 - 20 Yin X, Han J. CPAR: classification based on predictive association rules. In: Proceedings of SIAM International Conference on Data Mining, San Fransisco, 2003. 331–335
 - 21 Li W, Han J, Pei J. CMAR: accurate and efficient classification based on multiple class-association rules. In: Proceedings of IEEE International Conference on Data Mining, San Jose, 2001. 369–376
 - 22 Liu B, Hsu W, Ma Y M. Integrating classification and association rule mining. In: Proceedings of 4th International Conference on Knowledge Discovery and Data Mining, New York, 1998. 80–86
 - 23 Rao R V, Savsani V J, Vakharia D P. Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Comput Aid Des*, 2011, 43: 303–315

Information for authors

SCIENCE CHINA Information Sciences (Sci China Inf Sci), cosponsored by the Chinese Academy of Sciences and the National Natural Science Foundation of China, and published by Science China Press, is committed to publishing high-quality, original results of both basic and applied research in all areas of information sciences, including computer science and technology; systems science, control science and engineering (published in Issues with odd numbers); information and communication engineering; electronic science and technology (published in Issues with even numbers). *Sci China Inf Sci* is published monthly in both print and electronic forms. It is indexed by Academic OneFile, Astrophysics Data System (ADS), CSA, Cabells, Current Contents/Engineering, Computing and Technology, DBLP, Digital Mathematics Registry, Earthquake Engineering Abstracts, Engineering Index, Engineered Materials Abstracts, Gale, Google, INSPEC, Journal Citation Reports/Science Edition, Mathematical Reviews, OCLC, ProQuest, SCOPUS, Science Citation Index Expanded, Summon by Serial Solutions, VINITI, Zentralblatt MATH.

Papers published in *Sci China Inf Sci* include:

REVIEW (20 printed pages on average) surveys representative results and important advances on well-identified topics, with analyses and insightful views on the states of the art and highlights on future research directions.

RESEARCH PAPER (no more than 15 printed pages) presents new and original results and significant developments in all areas of information sciences for broad readership.

BRIEF REPORT (no more than 4 printed pages) describes key ideas, methodologies, and results of latest developments in a timely manner.

Authors are recommended to use *Science China's* online submission services. To submit a manuscript, please go to www.scichina.com, create an account to log in <http://mco3.manuscriptcentral.com/scis>, and follow the instructions there to upload text and image/table files.

Authors are encouraged to submit such accompanying materials as short statements on the research background and area/subareas and significance of the work, and brief introductions to the first and corresponding authors including their mailing addresses with post codes, telephone numbers, fax numbers, and e-mail addresses. Authors may also suggest several qualified experts (with full names, affiliations, phone numbers, fax numbers, and e-mail addresses) as referees, and/or request the exclusion of some specific individuals from potential referees.

All submissions will be reviewed by referees selected by the editorial board. The decision of acceptance or rejection of a manuscript is made by the editorial board based on the referees' reports. The entire review process may take 90 to 120 days, and the editorial office will inform the author of the decision as soon as the process is completed. If the editorial board fails to make a decision within 120 days, please contact the editorial office.

Authors should guarantee that their submitted manuscript has not been published before and has not been submitted elsewhere for print or electronic publication consideration. Submission of a manuscript is taken to imply that all the named authors are aware that they are listed as coauthors, and they have agreed on the submitted version of the paper. No change in the order of listed authors can be made without an agreement signed by all the authors.

Ethical responsibilities of authors: Authors should refrain from misrepresenting research results which could damage the trust in the journal and ultimately the entire scientific endeavour, and follow the COPE guidelines on how to deal with potential acts of misconduct.

Disclosure of potential conflict of interests: Authors must disclose all relationships or interests that could influence or

bias the work. The corresponding author will include a summary statement in the text of the manuscript in a separate section before the reference list.

Once a manuscript is accepted, the authors should send a copyright transfer form signed by all authors to Science China Press. Authors of one published paper will be presented one sample copy. If more sample copies or offprints are required, please contact the managing editor and pay the extra fee. The full text opens free to domestic readers at www.scichina.com, and is available to overseas readers at www.springerlink.com.

Subscription information

ISSN print edition: 1674-733X

ISSN electronic edition: 1869-1919

Volume 57 (12 issues) will appear in 2014

Subscription rates For information on subscription rates please contact: Customer Service

China: sales@scichina.org

North and South America: journals-ny@springer.com

Outside North and South America:

subscriptions@springer.com

Orders and inquiries:

China

Science China Press

16 Donghuangchenggen North Street, Beijing 100717, China

Tel: +86 10 64015683, Fax: +86 10 64016350

Email: informatics@scichina.org

North and South America

Springer New York, Inc.

Journal Fulfillment, P.O. Box 2485

Secaucus, NJ 07096 USA

Tel: 1-800-SPRINGER or 1-201-348-4033

Fax: 1-201-348-4505

Email: journals-ny@springer.com

Outside North and South America:

Springer Distribution Center

Customer Service Journals

Haberstr. 7, 69126 Heidelberg, Germany

Tel: +49-6221-345-0, Fax: +49-6221-345-4229

Email: subscriptions@springer.com

Cancellations must be received by September 30 to take effect at the end of the same year.

Changes of address: Allow for six weeks for all changes to become effective. All communications should include both old and new addresses (with postal codes) and should be accompanied by a mailing label from a recent issue. According to § 4 Sect. 3 of the German Postal Services Data Protection Regulations, if a subscriber's address changes, the German Federal Post Office can inform the publisher of the new address even if the subscriber has not submitted a formal application for mail to be forwarded. Subscribers not in agreement with this procedure may send a written complaint to Customer Service Journals, Karin Tiks, within 14 days of publication of this issue.

Microform editions are available from: ProQuest. Further information available at <http://www.il.proquest.com/uni>

Electronic edition

An electronic version is available at springerlink.com.

Production

Science China Press

16 Donghuangchenggen North Street, Beijing 100717, China

Tel: +86 10 64015683, Fax: +86 10 64016350

Printed in the People's Republic of China

Jointly published by

Science China Press and Springer



SCIENCE CHINA PRESS

SCIENCE CHINA Series | Chinese Science Bulletin

SCIENCE CHINA Series and the *Chinese Science Bulletin* are academic journals supervised by the Chinese Academy of Sciences, co-sponsored by the Chinese Academy of Sciences and the National Natural Science Foundation of China, jointly published by Science China Press and Springer. *SCIENCE CHINA Series* and the *Chinese Science Bulletin* have presented the finest examples of China's development in both natural sciences and technological research to the international scientific community. In order to fully express China's achievements in fundamental scientific and engineering research, *SCIENCE CHINA Series* is published in seven journals, i.e., Mathematics, Chemistry, Life Sciences, Earth Sciences, Technological Sciences, Information Sciences, and Physics (including Mechanics and Astronomy), with the *Chinese Science Bulletin* serving as a multidisciplinary journal.

- Peer-reviewed
- Online submission
- Indexed by SCI, CA, EI, etc.
- Easy access to the electronic version

Honorary Editor General: ZHOU Guangzhao (Zhou Guang Zhao) | Editor General: ZHU Zuoyan

Mathematics	Chemistry	Life Sciences	Earth Sciences
The cover features a fractal spiral pattern in yellow and orange tones.	The cover features a molecular structure and a volcano erupting.	The cover features a microscopic view of green and yellow cells.	The cover features an underwater scene with coral reefs.
Monthly Editor-in-Chief YUAN YaXiang	Monthly Editor-in-Chief WAN LiJun	Monthly Editor-in-Chief WANG DaCheng	Monthly Editor-in-Chief ZHENG YongFei
Technological Sciences	Information Sciences	Physics, Mechanics & Astronomy	Chinese Science Bulletin
The cover features a satellite in space.	The cover features binary code and a digital grid.	The cover features a red and orange abstract design.	The cover features a blue and white abstract design.
Monthly Editor-in-Chief YE HengQiang	Monthly Editor-in-Chief LI Wei	Monthly Editor-in-Chief WANG DingSheng ZHANG jie	Published three times every month

www.scichina.com | www.springer.com/scp

Sponsored by
Chinese Academy of Sciences (CAS)
National Natural Science Foundation of China (NSFC)

Published by
Science China Press & Springer

