A Novel Rule Weighting Approach in Classification Association Rule Mining

(An Extended Version of 2007 IEEE ICDM Workshop Paper)

Yanbo J. Wang [1], Qin Xin [2], and Frans Coenen [1]
[1] *Department of Computer Science, The University of Liverpool*
Ashton Building, Ashton Street, Liverpool, L69 3BX, United Kingdom
{jwang, frans}@csc.liv.ac.uk
[2] *Department of Informatics, The University of Bergen*
P.B.7800, N-5020 Bergen, Norway
xin@ii.uib.no

September 1st, 2007

Abstract

Classification Association Rule Mining (CARM) is a recent Classification Rule Mining (CRM) approach that builds an Association Rule Mining (ARM) based classifier using Classification Association Rules (CARs). Regardless of which particular CARM algorithm is used, a similar set of CARs is always generated from data, and a classifier is usually presented as an ordered list of CARs, based on a selected rule ordering strategy. Hence to produce an accurate classifier, it is essential to develop a rational rule ordering mechanism. In the past decade, a number of rule ordering strategies have been introduced that can be categorized under three headings: (1) support-confidence, (2) rule weighting, and (3) hybrid. In this paper, we propose an alternative rule weighting scheme, namely CISRW (Class-Item Score based Rule Weighting), and develop a rule weighting based rule ordering mechanism based on CISRW. Subsequently, two hybrid rule ordering strategies are further introduced by combining (1) and CISRW. The experimental results show that the three proposed CISRW based/related rule ordering strategies perform well with respect to the accuracy of classification.

1. Introduction

Data Mining is a promising area of current research and development in Computer Science, which is attracting more and more attention from a wide range of different groups of people. Data Mining aims to extract various models of hidden and interesting knowledge (i.e. patterns, rules, regularities, trends, etc.) from databases, where the volume of a collected database can be measured in GBytes. Classification Rule Mining (CRM) [24] is a well-known Data Mining technique for the extraction of hidden Classification Rules (CRs) — a common model of mined knowledge — from a given database that is coupled with a set of pre-defined class labels, the objective being to build a classifier to classify "unseen" data records. One recent approach to CRM is to employ Association Rule Mining (ARM) [1] methods to identify the desired CRs, i.e. Classification Association Rule Mining (CARM) [3].

CARM aims to mine a set of Classification Association Rules (CARs) from a class-transaction database, where a CAR describes an implicative co-occurring relationship between a set of binary-valued data attributes (items) and a pre-defined class, expressed in the form of an "⟨antecedent⟩ ⇒ ⟨consequent-class⟩" rule. In [27] Thabtah *et al.* indicate that several literatures (i.e. [3], [4], [5], [8], [23]) show that CARM is able to build more accurate classifiers than other CRM techniques, such as decision trees (i.e. [15]), rule induction (i.e. [16], [19]) and probabilistic approaches (i.e. [20]). In [14] Coenen *et al.* also suggest that results presented in a number of studies (i.e. [19], [21]) show that in many cases CARM seems to offer higher accuracy than traditional classification methods, i.e. C4.5 [24].

In the past decade, a number of CARM approaches have been developed that include: TFPC (Total From Partial Classification) [11] [14], CBA (Classification Based Associations) [21], CMAR (Classification based on Multiple Association Rules) [19], CPAR (Classification based on Predictive Association Rules) [29], etc. Although these CARM approaches employ different ARM techniques to extract CARs from a given class-transaction database, a similar set of CARs is always generated, based on a pair of specific values for both support and confidence thresholds. Regardless of which particular CARM method is utilized, a classifier is usually presented as an ordered list of CARs, based on a selected rule ordering strategy. Hence, it can be indicated that the essential to produce a more accurate CARM classifier is to develop a better (more rational) rule ordering approach.

Coenen and Leng [11], identify:

- **Three Common CARM Case Satisfaction Approaches:** (1) Best First Rule, (2) Best *K* Rules, and (3) All Rules.
- **Five Established CARM Rule Ordering Mechanisms:** (1) Confidence Support size-of-rule-Antecedent (CSA), (2) size-of-rule-Antecedent Confidence Support (ACS), (3) Weighted Relative Accuracy (WRA), (4) Laplace Accuracy (LAP), and (5) Chi-square Testing (χ^2).

These are described in further detail in subsections 2.2, 3.1 and 3.2 below. In [28] the authors divide the above rule ordering mechanisms into two groups: (type 1) support-confidence based which includes CSA and ACS; and (type 2) rule weighting based which includes WRA, LAP and χ^2. Wang *et al.* in [28] also propose a hybrid based ordering approach by combining one rule ordering mechanism taken from the (type 1) group and another rule ordering mechanism taken from the (type 2) group.

In this paper, we introduce a novel rule weighting scheme, namely CISRW (Class-Item Score based Rule Weighting), which assigns a weighting score to each CAR by assigning a weighting score to each CAR item. Then a rule ordering mechanism is proposed that simply sorts CARs in a descending order, based on their assigned CISRW score. As a consequence, two hybrid rule ordering strategies are further developed: (1) Hybrid CSA/CISRW, and (2) Hybrid ACS/CISRW. The experimental results show good performance regarding the accuracy of classification when using the proposed CISRW based/related rule ordering strategies with the Best First Rule case satisfaction.

The rest of this paper is organized as follows. In section 2 we describe some related work relevant to this study. Section 3 outlines the existing rule ordering strategies in CARM. The proposed rule weighting/ordering approach is described in section 4. In section 5 we present experimental results, and in section 6 our conclusions and open issues for further research.

2. Related Work

2.1. An Overview of CARM Methods

The idea of CARM was first introduced in [3]. Subsequently a number of alternative approaches have been presented. Broadly CARM algorithms can be categorized into two groups according to the way that the CARs are generated:

- **Two Stage Algorithms** where a set of CARs are produced first (stage 1), which are then pruned and placed into a classifier (stage 2). Typical algorithms of this approach include CBA [21] and CMAR [19]. CBA, developed by Liu *et al.* in 1998, is an Apriori [2] based CARM algorithm, which: (1) applies its CBA-GR procedure for CAR generation; and (2) applies its CBA-CB procedure to build a classifier based on the generated CARs. CMAR, introduced by Li *et al.* in 2001, is similar to CBA but generates CARs through a FP-tree [17] based approach.
- **Integrated Algorithms** where the classifier is produced in a single processing step. Typical algorithms of this approach include TFPC [11] [14], and induction systems such as FOIL (First Order Inductive Learner) [25], PRM (Predictive Rule Mining) and CPAR [29]. TFPC, proposed by Coenen and Leng in 2004, is an Apriori-TFP [12] [13] based CARM algorithm, which generates CARs through efficiently constructing both P-tree and T-tree set enumeration tree structures [26]. FOIL is an inductive learning algorithm for generating CARs, introduced by Quinlan and Cameron-Jones in 1993. This algorithm was later developed by Yin and Han to produce the PRM CAR generation algorithm. PRM was then further developed, by Yin and Han in 2003 to produce CPAR.

2.2. Case Satisfaction Mechanisms

In [11] Coenen and Leng summarize three case satisfaction mechanisms that have been employed in a variety of CARM algorithms for utilizing the resulting classifier to classify “unseen” data records. These three case satisfaction approaches are itemized as follows (given a particular case):

- **Best First Rule:** Select the first rule that satisfies the given case according to some ordering imposed on the CAR list. The ordering can be defined according to many different ordering strategies including CSA — combinations of confidence, support and size of rule antecedent, with confidence being the most significant factor (used in CBA, TFPC and the early stages of processing of CMAR); ACS — an alternative mechanism to CSA that considers the size of rule antecedent as the most significant factor; WRA — which reflects a number of rule interestingness measures as proposed in [18]; LAP — as used in PRM and CPAR; χ^2 — as used, in part, in CMAR; etc.
- **Best *K* Rules:** Select the first (top) *K* rules that satisfy the given case and then select a rule according to some averaging process as used for example, in CPAR. The term "best" in this case is defined according to an imposed ordering of the form described in Best First Rule.
- **All Rules:** Collect all rules in the classifier that satisfy the given case and then evaluate this collection to identify a class. One well-known evaluation method in this category is WCS (Weighted χ^2) Testing as used in CMAR.

3. Overview of Rule Ordering Approaches

As noted in the subsection 2.2, five established rule ordering strategies can be identified to support the Best First Rule case satisfaction mechanism. Each can be further separated into two phases: (1) a rule weighting phase where each CAR is labeled with a weighting score that represents the significance of this CAR indicates a pre-defined class; and (2) a rule re-ordering phase, which sorts the original CAR list in a descending manner, based on the weighting score assigned to each CAR in phase (1).

As noted in section 1, Wang *et al.* divide the five of the established rule ordering strategies into two groups [28]: type 1, support-confidence based where the well-established "support-confidence" framework is addressed; and type 2, rule weighting based where an additive weighting score is assigned to each CAR, based on a particular weighting scheme.

In [29] Yin and Han (i) believe that there are only a limited number, say at most *K* in each class, of CARs that are required to distinguish between classes and should be thus used to make up a classifier; (ii) suggest a value of 5 as an appropriate value for *K*; and (iii) employ LAP to estimate the accuracy of CARs.

By incorporating the *K* rules concept of Yin and Han with both type 1 and type 2 groups, a hybrid support-confidence & rule weighting based ordering approach was developed in [28], which fuses both the case satisfaction mechanisms of Best First Rule and Best *K* Rules.

3.1. Support-Confidence Based Ordering

- **CSA:** The CSA rule ordering strategy is based on the well-established "support-confidence" framework that was originally introduced for AR interestingness measure. CSA does not assign an additive weighting score to any CAR in its rule weighing phase, but simply gathers both values of confidence and support, and the size of the rule antecedent to "express" a weighting score for each CAR. In its rule re-ordering phase, CSA generally sorts the original CAR list in a descending order based on the value of confidence of each CAR. For these CARs that share a common value of confidence, CSA sorts them in a descending order based on their support value. Furthermore for these CARs that share common values for both confidence and support, CSA sorts them in an ascending order based on the size of the rule antecedent.
- **ACS:** The ACS rule ordering strategy is a variation of CSA. It takes the size of the rule antecedent as its major factor (using a descending order) followed by the rule confidence and support values respectively. In [11] Coenen and Leng state that ACS ensures: "*specific rules have a higher precedence than more general rules*".

3.2. Rule Weighting Based Ordering

- **WRA:** The use of WRA can be found in [18], where this technique is used to determine an expected accuracy for each CAR. In its rule weighting phase, WRA assigns an additive weighting score to each CAR. The calculation of the WRA score of a CAR *R*, confirmed in [11], is:

$$wra\_score(R) = support(R.antecedent) \times (confidence(R) - support(R.consequent\text{-}class))$$

In its rule re-ordering phase, the original CAR list is simply sorted in a descending order, based on the assigned *wra_score* of each CAR.

- **LAP:** The use of the *Laplace Expected Error Estimate* [9] can be found in [29]. The principle of applying this rule ordering mechanism is similar to WRA. The calculation of the LAP score of a CAR R is:

  $$lap\_score(R) = (support(R.antecedent \cup R.consequent\text{-}class) + 1) / (support(R.antecedent) + |C|)$$

  where $|C|$ denotes the number of pre-defined classes.

- **χ^2:** χ^2 Testing is a well-known technique in Statistics [22], which can be used to determine whether two variables are independent of one another. In χ^2 Testing a set of observed values O is compared against a set of expected values E — values that would be estimated if there were no associative relationship between the variables. The value of χ^2 is calculated as: $\sum_{\{i = 1\ldots n\}} (O_i - E_i)^2 / E_i$, where n is the number of entries in the confusion matrix, which is always 4 in CARM. If the χ^2 value between two variables (the rule antecedent and consequent-class of a CAR) above a given threshold value (for CMAR the chosen threshold is 3.8415), thus it can be concluded that there is a relation between the rule antecedent and consequent-class, otherwise there is not a relation. After assigning an additive χ^2 score (value) to each CAR, it can be used to re-order the CAR list in a descending basis.

3.3. Hybrid Rule Ordering Schemes

From the foregoing we can identify six hybrid rule ordering schemes [28]:

- **Hybrid CSA/WRA:** Selects the Best K Rules (for each pre-defined class) in a WRA manner, and re-orders both the best K CAR list and the original CAR list in a CSA fashion. The best K CAR list is linked at front of the original CAR list.
- **Hybrid CSA/LAP:** Selects the Best K Rules (for each pre-defined class) in a LAP manner, and re-orders both the best K CAR list and the original CAR list in a CSA fashion. The best K CAR list is linked at front of the original CAR list.
- **Hybrid CSA/χ^2:** Selects the Best K Rules (for each pre-defined class) in a χ^2 manner, and re-orders both the best K CAR list and the original CAR list in a CSA fashion. The best K CAR list is linked at front of the original CAR list.
- **Hybrid ACS/WRA:** Selects the Best K Rules (for each pre-defined class) in a WRA manner, and re-orders both the best K CAR list and the original CAR list in an ACS fashion. The best K CAR list is linked at front of the original CAR list.
- **Hybrid ACS/LAP:** Selects the Best K Rules (for each pre-defined class) in a LAP manner, and re-orders both the best K CAR list and the original CAR list in an ACS fashion. The best K CAR list is linked at front of the original CAR list.
- **Hybrid ACS/χ^2:** Selects the Best K Rules (for each pre-defined class) in a χ^2 manner, and re-orders both the best K CAR list and the original CAR list in an ACS fashion. The best K CAR list is linked at front of the original CAR list.

4. Proposed Rule Weighting/Ordering

In this section, we describe the proposed CISRW rule weighting scheme, which assigns a weighting score to each CAR, by computing a score for each rule item and averaging the sum of all rule item scores. Then a rule weighting based rule ordering strategy is introduced founded on CISRW. As a consequence, two further hybrid rule ordering strategies that combining either CSA or ACS with CISRW, are proposed.

4.1. Proposed Rule Weighting Scheme

4.1.1. Item Weighting Score

There are n items involved in a given class-transaction database D_{CT} that is coupled with a set of pre-defined classes $C = \{c_1, c_2, \ldots, c_{m-1}, c_m\}$. For a particular pre-defined class A, a score is assigned to each item in D_{CT} that distinguishes the significant items for class A from the insignificant ones.

Definition 1. *Let $Ç^A(Item_h)$ denote the contribution of each $item_h \in D_{CT}$ for class A, which represents how significantly $item_h$ determines A, where $0 \leq Ç^A(Item_h) \leq |C|$, and $|C|$ is the size function of the set C.*

The calculation of $Ç^A(Item_h)$ is given as follows:

$$Ç^A(Item_h) = TransFreq(Item_h, A) \times (1 - TransFreq(Item_h, \neg A)) \times (|C| / ClassCount(Item_h, C))$$

where

(1) The $TransFreq(Item_h, A \text{ or } \neg A)$ function computes how frequently that $Item_h$ appears

in class A or the group of classes $\neg A$ (the complement of A). The calculation of this function is: (*number of transactions with* $Item_h$ *in the class or class-group*) / (*total number of transactions in the class or class-group*); and

(2) The *ClassCount*($Item_h$, C) function simply counts the number of classes in C which contain $Item_h$.

The rationale of this item weighing score is demonstrated as follows:

- The weighting score of $Item_h$ for class A tends to be high if $Item_h$ is frequent in A;
- The weighting score of $Item_h$ for class A tends to be high if $Item_h$ is infrequent in $\neg A$;
- The weighting score of $Item_h$ for any class tends to be high if $Item_h$ is involved in a small number of classes in C. (In [7], a similar idea can be found in feature selection for text categorization.)

4.1.2. Rule Weighting Score

Based on the item weighting score, a rule weighing score is assigned to each CAR R in the original CAR list.

Definition 2. *Let $Ç^A(R)$ denote the contribution of each CAR R in the original CAR list for class A that represents how significantly R determines A.*

The calculation of $Ç^A(R)$ is given as follows:

$$Ç^A(R) = \left(\sum_{\{h = 1 \ldots |R.antecedent|\}} Ç^A(Item_h \in R.antecedent)\right) / (|R.antecedent|)$$

where $|R.antecedent|$ is the size function of the antecedent of this CAR.

4.2. Proposed Rule Ordering Strategies

- **CISRW:** In the rule weighing phase, the CISRW weighing score is assigned to each CAR, which represents how significantly the CAR antecedent determines its consequent-class. In the rule re-ordering phase, the original CAR list is simply sorted in a descending order based on the assigned CISRW score of each CAR.
- **Hybrid CSA/CISRW:** Selects the Best K Rules (for each pre-defined class) in a CISRW manner, and re-orders both the best K CAR list and the original CAR list in a CSA fashion. The best K CAR list is linked at front of the original CAR list.
- **Hybrid ACS/CISRW:** Selects the Best K Rules (for each pre-defined class) in a CISRW manner, and re-orders both the best K CAR list and the original CAR list in an ACS fashion. The best K CAR list is linked at front of the original CAR list.

In Figure 1, a common procedure for both proposed Hybrid CSA/CISRW and Hybrid ACS/CISRW strategies is outlined.

Procedure Hybrid CSA(or ACS)/CISRW
Input: (a) A list of CARs $\mathfrak{R}$ (either in CSA or ACS ordering manner);
(b) A desired number (integer value) for the best rules K;
Output: A re-ordered list of CARs $\mathfrak{R}^{NEW}$ (either in Hybrid CSA/CISRW or Hybrid ACS/CISRW ordering manner);
begin
$\mathfrak{R}^{NEW} := \{\varnothing\}$;
$\mathfrak{R}^{CISRW} := \{\varnothing\}$;
for each CAR $\in \mathfrak{R}$ **do**
    calculate the CISRW score (δ) for this CAR;
    $\mathfrak{R}^{CISRW} \leftarrow \mathfrak{R}^{CISRW} \cup (\text{CAR} \oplus \delta)$;
    // the $\oplus$ sign means "with" an additive CAR attribute
end for
sort $\mathfrak{R}^{CISRW}$ in a descending order based on δ;
$\mathfrak{R}^{CISRW} \leftarrow$ **remain** the top K CARs (for each pre-defined class) $\in \mathfrak{R}^{CISRW}$;
sort $\mathfrak{R}^{CISRW}$ either in CSA or ACS ordering manner; *// keep it consistent with $\mathfrak{R}$*
$\mathfrak{R}^{NEW} \leftarrow$ **link** $\mathfrak{R}^{CISRW}$ at front of $\mathfrak{R}$;
return ($\mathfrak{R}^{NEW}$);
end

Figure 1. The Hybrid CSA(or ACS)/CISRW procedure

5. Experimental Results

In this section, we aim to evaluate the proposed CISRW based/related rule ordering strategies with respect to the accuracy of classification. All evaluations were obtained using the TFPC algorithm coupled with the Best First Rule case satisfaction, although any other CARM classifier generator, founded on the Best First Rule mechanism, could equally well be used. Experiments were run on a 1.86 GHz Intel(R) Core(TM)2 CPU with 1.00 GB of RAM running under Windows Command Processor.

The experiments were conducted using a range of datasets taken from the LUCS-KDD Discretised/Normalised ARM and CARM Data Library [10]. The chosen databases are originally taken from the UCI Machine Learning Repository [6]. These datasets have been discretized and normalized using the LUCS-KDD Discretised Normalised (DN) software, so that data are then presented in a binary format suitable for use with CARM applications. It should be noted that the datasets were re-arranged so that occurrences of classes were distributed evenly throughout. This then allowed the datasets to be divided in half with the first half used as the training set and the second half as the test set. Although a "better" accuracy figure might have been obtained using Ten-cross Validation, it is the relative accuracy that is of interest here and not the absolute accuracy.

The first set of evaluations undertaken used a confidence threshold value of 50% and a support threshold value 1% (as used in the published evaluations of CMAR [19], CPAR [29], TFPC [11] [14], and the hybrid rule ordering approach [28]). The results are presented in Table 1 where 120 classification accuracy values are listed based on 20 chosen datasets. The row labels describe the key characteristics of each dataset: for example, the label *adult.D97.N48842.C2* denotes the "adult" dataset, which comprises 48,842 records in 2 pre-defined classes, with attributes that for the experiments described here have been discretized and normalized into 97 binary categories.

Table 1. Classification accuracy — support-confidence & rule weighting strategies vs. the CISRW strategy

DATASETS	CSA	ACS	WRA	LAP	χ^2	CISRW
adult.D97.N48842.C2	80.83	73.99	81.66	76.07	76.07	81.61
breast.D20.N699.C2	89.11	89.11	87.68	65.62	65.62	87.68
chessKRvK.D58.N28056.C18	14.95	14.95	14.95	14.95	14.95	14.95
connect4.D129.N67557.C3	65.83	64.83	67.93	65.83	65.83	66.94
flare.D39.N.1389.C9	84.44	83.86	84.15	84.44	84.44	84.44
glass.D48.N214.C7	58.88	43.93	50.47	52.34	50.47	55.14
heart.D52.N303.C5	58.28	28.48	55.63	54.97	54.97	57.62
horseColic.D85.N368.C2	72.83	40.76	79.89	79.89	63.04	79.89
ionosphere.D157.N351.C2	85.14	61.14	86.86	64.57	64.57	83.43
iris.D19.N.150.C3	97.33	97.33	97.33	97.33	97.33	97.33
led7.D24.N3200.C10	68.38	61.38	63.94	63.88	66.56	60.50
letRecog.D106.N20000.C26	31.13	26.21	26.33	26.33	28.52	26.38
mushroom.D90.N8124.C2	99.21	65.76	98.45	98.45	49.43	98.45
nursery.D32.N12960.C5	80.35	55.88	70.17	70.17	70.17	70.17
pageBlocks.D46.N5473.C5	90.97	90.97	90.20	89.80	89.80	91.56
pima.D38.N768.C2	73.18	71.88	72.92	65.10	65.10	72.92
soybean-large.D118.N683.C19	86.22	79.77	36.36	36.07	77.42	63.93
ticTacToe.D29.N958.C2	71.61	36.12	68.06	65.34	65.34	68.27
waveform.D101.N5000.C3	61.56	47.96	56.24	57.84	57.28	56.08
zoo.D42.N101.C7	80.00	42.00	56.00	42.00	42.00	86.00
Average	**72.51**	**58.82**	**67.26**	**63.55**	**62.45**	**70.16**

From Table 1 it is clear that with a 50% confidence threshold and a 1% support threshold the CSA mechanism worked better than other alternative non-hybrid rule ordering strategies. When applying CSA, the average accuracy of classification throughout the 20 datasets is 72.51%. The performance rank of the five established rule ordering mechanisms is specified as follows: (1) CSA — the average accuracy of classification is 72.51%; (2) WRA — the accuracy is 67.26%; (3) LAP — 63.55%; (4) χ^2 — 62.45%; and (5) ACS — 58.82%. It should be noted that this ranking result corroborates to the results presented in [28] although both investigations involve different datasets in their experimentation section. With respect to the group of rule weighting based rule ordering mechanisms, the proposed CISRW performed better than other rule weighting mechanisms, where its average accuracy of classification throughout the 20 datasets is 70.16%.

The second set of evaluations undertaken used a confidence threshold value of 50%, a support threshold value of 1%, and a value of 5 as an appropriate value for *K* when selecting the Best *K* Rules (*K* = 5 was also used in [29]). The results are demonstrated in Table 2 where 80 classification accuracy values are listed based on 20 chosen datasets.

Table 2. Classification accuracy — CSA based hybrid strategies vs. the Hybrid CSA/CISRW strategy

DATASETS	Hybrid CSA/WRA	Hybrid CSA/LAP	Hybrid CSA/χ^2	Hybrid CSA/CISRW
adult.D97.N48842.C2	83.33	79.95	79.95	81.46
breast.D20.N699.C2	89.11	88.54	89.11	89.11
chessKRvK.D58.N28056.C18	14.95	14.95	14.95	14.95
connect4.D129.N67557.C3	67.67	65.83	65.83	66.71
flare.D39.N.1389.C9	84.29	54.44	84.44	84.15
glass.D48.N214.C7	66.36	66.36	66.36	65.42
heart.D52.N303.C5	55.63	56.95	58.94	58.28

horseColic. D85.N368.C2	83.15	83.15	79.89	83.15
ionosphere. D157.N351.C2	90.29	89.71	88.00	89.14
iris.D19. N.150.C3	97.33	97.33	97.33	97.33
led7.D24. N3200.C10	68.19	68.19	68.38	68.38
letRecog.D106. N20000.C26	31.49	31.49	31.56	30.87
mushroom.D90. N8124.C2	98.45	98.82	98.45	98.82
nursery.D32. N12960.C5	78.86	78.86	78.86	78.26
pageBlocks.D46. N5473.C5	90.97	90.97	90.97	90.97
pima.D38. N768.C2	73.18	73.18	72.66	73.44
soybean-large. D118.N683.C19	80.94	80.94	82.11	83.58
ticTacToe.D29. N958.C2	74.95	74.74	72.65	73.90
waveform.D101. N5000.C3	57.96	57.96	60.60	58.40
zoo.D42. N101.C7	84.00	90.00	72.00	88.00
Average	**73.56**	**73.62**	**72.65**	**73.72**

From Table 2 it can be seen that with a 50% confidence threshold, a 1% support threshold, and K = 5, the proposed Hybrid CSA/CISRW strategy preformed better than other alternative CSA related hybrid rule ordering mechanisms. When applying Hybrid CSA/CISRW, the average accuracy of classification throughout the 20 datasets is 73.72%. The performances of other CSA related hybrid strategies were ranked as: (1) CSA/LAP — the average accuracy of classification is 73.62%; (2) CSA/WRA — the accuracy is 73.56%; and (3) CSA/χ^2 — 72.65%. This ranking result is consistent to the experimental results shown in [28] even different datasets were used.

With regard to the first two sets of evaluations, the third set of evaluations undertaken used a confidence threshold value of 50% and a support threshold value of 1% as well. Again, a value of 5 was considered as an appropriate value for K. In Table 3, 80 classification accuracy values are listed based on 20 chosen datasets.

Table 3. Classification accuracy — ACS based hybrid strategies vs. the Hybrid ACS/CISRW strategy

DATASETS	Hybrid ACS/WRA	Hybrid ACS/LAP	Hybrid ACS/χ^2	Hybrid ACS/CISRW
adult.D97. N48842.C2	78.56	83.76	80.14	81.95
breast.D20. N699.C2	89.11	88.54	89.11	89.11
chessKRvK.D58. N28056.C18	14.95	14.95	14.95	14.95
connect4.D129. N67557.C3	64.88	64.88	64.88	64.88
flare.D39. N.1389.C9	83.86	83.86	83.86	83.86
glass.D48. N214.C7	65.42	65.42	68.22	63.55
heart.D52. N303.C5	52.32	50.33	50.33	52.32
horseColic. D85.N368.C2	75.00	83.15	71.20	78.80
ionosphere. D157.N351.C2	90.29	89.71	88.00	89.14
iris.D19. N.150.C3	97.33	97.33	97.33	97.33
led7.D24. N3200.C10	62.06	62.06	62.31	65.69
letRecog.D106. N20000.C26	27.39	27.39	28.41	28.58
mushroom.D90. N8124.C2	98.45	98.82	98.45	98.82
nursery.D32. N12960.C5	66.73	66.73	66.73	71.28
pageBlocks.D46. N5473.C5	90.97	90.97	90.97	90.97
pima.D38. N768.C2	73.18	73.18	72.66	71.61
soybean-large. D118.N683.C19	75.66	75.66	78.01	78.59
ticTacToe.D29. N958.C2	60.75	70.35	67.22	67.22
waveform.D101. N5000.C3	59.20	59.20	60.60	58.40
zoo.D42. N101.C7	80.00	80.00	80.00	76.00
Average	**70.31**	**71.31**	**70.67**	**71.15**

From Table 3 it can be seen that with a 50% confidence threshold, a 1% support threshold, and K = 5, the best-performing hybrid ACS related rule ordering strategy is the Hybrid ACS/LAP mechanism. When applying this mechanism, the average accuracy of classification throughout the 20 datasets is 71.31%. The proposed Hybrid ACS/CISRW approach demonstrated a comparable performance to Hybrid ACS/LAP, where its average classification accuracy is 71.15%. The performances of three existing ACS related hybrid strategies were ranked as: (1) ACS/LAP — 71.31%; (2) ACS/χ^2 — the average accuracy is 70.67%; and (3) ACS/WRA — 70.31%. This ranking result corroborates to the experimental results provided in [28] although different datasets were concerned.

6. Conclusion

CARM is a recent CRM approach that aims to classify "unseen" data based on building an ARM based classifier. A number of literatures have confirmed the outstanding performance of CARM. It can be clarified that no matter which particular ARM technique is employed, a similar set of CARs is always generated from data, and a classifier is usually presented as an ordered list of CARs, based on an applied rule ordering strategy. In this paper a novel rule weighting approach was proposed, which assigns a weighting score to each generated CAR. Based on the proposed rule weighting approach, a straightforward rule ordering mechanism (CISRW) was introduced. Subsequently, two hybrid rule ordering strategies

(Hybrid CSA/CISRW and Hybrid ACS/CISRW) were further developed.

Table 4. Ranking of classification accuracies for all rule ordering strategies

Rank No.	Rule Ordering Strategy	Average Accuracy	Rank No. presented in [28]
1	Hybrid CSA/CISRW	73.72	—
2	Hybrid CSA/LAP	73.62	**1**
3	Hybrid CSA/WRA	73.56	**2**
4	Hybrid CSA/χ^2	72.65	**3**
5	CSA	72.51	**4**
6	Hybrid ACS/LAP	71.31	**5**
7	Hybrid ACS/CISRW	71.15	—
8	Hybrid ACS/χ^2	70.67	**6**
9	Hybrid ACS/WRA	70.31	**7**
10	CISRW	70.16	—
11	WRA	67.26	**8**
12	LAP	63.55	**9**
13	χ^2	62.45	**10**
14	ACS	58.82	**11**

From the experimental results, it can be seen that the average accuracy of classification, using the 20 chosen datasets, obtained by the proposed Hybrid CSA/CISRW rule ordering strategy can be better than other alternative rule ordering mechanisms, where the accuracy is 73.72%. In Table 4, the rank of classification accuracies for all fourteen rule ordering strategies is presented. The proposed Hybrid CSA/CISRW, Hybrid ACS/CISRW and CISRW rule ordering mechanisms were ranked as No.1, No. 7 and No. 10. Furthermore the performances of eleven existing rule ordering strategies were ranked in accordance with the results presented in [28] although different datasets were used in both investigations.

Further research is suggested to identify the improved rule weighting/ordering approach in CARM to give a better performance.

7. Acknowledgments

The authors would like to thank Prof. Paul Leng and Dr. Robert Sanderson of the Department of Computer Science at the University of Liverpool for their support with respect to the work described here.

8. References

[1] R. Agrawal, T. Imielinski, and A. Swami, "Mining Association Rule between Set of Items in Large Database", In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data (SIGMOD-93)*, ACM Press, Washington, D.C., United States, May 1993, pp. 207-216.

[2] R. Agrawal and R. Srikant, "Fast Algorithm for Mining Association Rules", In *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB-94)*, Morgan Kaufmann Publishers, Santiago de Chile, Chile, September 1994, pp. 487-499.

[3] K. Ali, S. Manganaris, and R. Srikant, "Partial Classification using Association Rules", In *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (KDD-97)*, AAAI Press, Newport Beach, CA, United States, August 1997, pp. 115-118.

[4] M.-L. Antonie and O.R. Zaiane, "An Associative Classifier based on Positive and Negative Rules", In *Proceedings of the 9th ACM SIGMOD International Conference on Management of Data — Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD-04)*, ACM Press, Paris, France, June 2004, pp. 64-69.

[5] E. Baralis and P. Torino, "A Lazy Approach to Pruning Classification Rules", In *Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM-02)*, IEEE Computer Society, Maebashi City, Japan, December 2002, pp. 35-42.

[6] C.L. Blake and C.J. Merz, "UCI Repository of Machine Learning Databases", *Department of Information and Computer Science, University of California, Irvine*, CA, United States, 1998.
http://www.ics.uci.edu/~mlearn/MLRepository.html

[7] C.H. Bong and K. Narayanan, "An Empirical Study of Feature Selection for Text Categorization based on Term Weightage", In *Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence (WI-04)*, IEEE Computer Society, Beijing, China, September 2004, pp. 599-602.

[8] J. Cendrowska, "PRISM: An Algorithm for Inducing Modular Rules", *International Journal of Man-Machine Studies*, 27(4), 1987, pp. 349-370.

[9] P. Clark and R. Boswell, "Rule Induction with CN2: Some Recent Improvement", In *Proceedings of the Fifth European Working Session on Learning (EWSL-91)*, Springer-Verlag, Porto, Portugal, March 1991, pp. 111-116.

[10] F. Coenen, "The LUCS-KDD Discretised/Normalised ARM and CARM Data Library", *Department of Computer Science, The University of Liverpool*, Liverpool, United Kingdom, 2003.
http://www.csc.liv.ac.uk/~frans/KDD/Software/LUCS-KDD-DN

[11] F. Coenen and P. Leng, "An Evaluation of Approaches to Classification Rule Selection", In *Proceedings of the 4th IEEE International Conference on Data Mining (ICDM-04)*,

IEEE Computer Society, Brighton, United Kingdom, November 2004, pp. 359-362.

[12] F. Coenen, P. Leng, and S. Ahmed, "Data Structure for Association Rule Mining: T-trees and P-trees", *IEEE Transactions on Knowledge and Data Engineering*, 16(6), 2004, pp. 774-778.

[13] F. Coenen, P. Leng, and G. Goulbourne, "Tree Structures for Mining Association Rules", *Journal of Data Mining and Knowledge Discovery*, 8(1), 2004, pp. 25-51.

[14] F. Coenen, P. Leng, and L. Zhang, "Threshold Tuning for Improved Classification Association Rule Mining", In *Proceedings of the 9th Pacific–Asia Conference on Knowledge Discovery and Data Mining (PAKDD-05)*, Springer-Verlag, Hanoi, Vietnam, May 2005, pp. 216-225.

[15] W.W. Cohen, "Fast Effective Rule Induction", In *Proceedings of the Twelfth International Conference on Machine Learning (ICML-95)*, Morgan Kaufmann Publishers, Tahoe City, CA, United States, July 1995, pp. 115-123.

[16] Duda, R. and Hart, P., *Pattern Classification and Scene Analysis*, John Wiley & Sons Inc., New York, United States, June 1973.

[17] J. Han, J. Pei, and Y. Yin, "Mining Frequent Patterns without Candidate Generation", In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (SIGMOD-00)*, ACM Press, Dallas, TX, United States, May 2000, pp. 1-12.

[18] N. Lavrac, P. Flach, and B. Zupan, "Rule Evaluation Measures: A Unifying View", In *Proceedings of the 9th International Workshop on Inductive Logic Programming (ILP-99)*, Springer-Verlag, Bled, Slovenia, June 1999, pp. 174-185.

[19] W. Li, J. Han, and J. Pei, "CMAR: Accurate and Efficient Classification based on Multiple Class-association Rules", In *Proceedings of the 2001 IEEE International Conference on Data Mining (ICDM-01)*, IEEE Computer Society, San Jose, CA, United States, 2001, pp. 369-376.

[20] T. Lim, W. Loh, and Y. Shih, "A Comparison of Prediction Accuracy, Complexity and Training Time of Thirty-tree Old and New Classification Algorithms", *Machine Learning*, 40(3), 2000, pp. 203-228.

[21] B. Liu, W. Hsu, and Y. Ma, "Integrating Classification and Association Rule Mining", In *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining (KDD-98)*, AAAI Press, New York City, NY, United States, 1998, pp. 80-86.

[22] Moore, D.S. and McCabe, G.P., *Introduction to the Practice of Statistics (Third Edition)*, W. H. Freeman and Company, United States, 1998.

[23] J.S. Park, M.-S. Chen, and P.S. Yu, "An Effective Hash-based Algorithm for Mining Association Rules", In *Proceedings of the 1995 ACM SIGMOD International Conference on Management of Data (SIGMOD-95)*, ACM Press, San Jose, CA, United States, May 1995, SIGMOD Record 24(2), pp. 175-186.

[24] Quinlan, J.R., *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, CA, United States, 1993.

[25] J.R. Quinlan and R.M. Cameron-Jones, "FOIL: A Midterm Report", In *Proceedings of the 1993 European Conference on Machine Learning (ECML-93)*, Springer-Verlag, Vienna, Austria, April 1993, pp. 3-20.

[26] R. Rymon, "Search through Systematic Set Enumeration", In *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR-92)*, Morgan Kaufmann Publishers, Cambridge, MA, United States, October 1992, pp. 539-550.

[27] F. Thabtah, P. Cowling, and Y. Peng, "The Impact of Rule Ranking on the Quality of Associative Classifiers", In *Proceedings of the Twenty-fifth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence (AI-2005)*, Springer-Verlag, Cambridge, United Kingdom, December 2005, pp. 277-287.

[28] Y.J. Wang, Q. Xin, and F. Coenen, "A Novel Rule Ordering Approach in Classification Association Rule Mining", In *Proceedings of the 5th International Conference on Machine Learning and Data Mining (MLDM-07)*, Springer-Verlag, Leipzig, Germany, July 2007, pp. 339-348.

[29] X. Yin and J. Han, "CPAR: Classification based on Predictive Association Rules", In *Proceedings of the Third SIAM International Conference on Data Mining (SDM-03)*, SIAM, San Francisco, CA, United States, 2003, pp. 331-335.