Information Fusion from Multiple Databases using Meta-Association Rules

M.D. Ruiz*, J. Gómez-Romero, M. Molina-Solana, M. Ros, M.J. Martin-Bautista

CITIC-UGR, University of Granada (Spain)

Dept. Computer Science and A.I., University of Granada (Spain)

Abstract

Nowadays, data volume, distribution, and volatility makes it difficult to search global patterns by applying traditional Data Mining techniques. In the case of data in a distributed environment, sometimes a local analysis of each dataset separately is adequate but some other times a global decision is needed by the analysis of the entire data. Association rules discovering methods typically require a single uniform dataset and managing with the entire set of distributed data is not possible due to its size. To address the scenarios in which satisfying this requirement is not practical or even feasible, we propose a new method for fusing information, in the form of rules, extracted from multiple datasets. The proposed model produces meta-association rules, i.e. rules in which the antecedent or the consequent may contain rules as well, for finding joint correlations among trends found individually in each dataset. In this paper, we describe the formulation and the implementation of two alternative frameworks that obtain, respectively, crisp meta-rules and fuzzy meta-rules. We compare our proposal with the information obtained when the datasets are not separated, in order to see the main differences between traditional association rules and meta-association rules. We also compare crisp and fuzzy methods for meta-association rule mining, observing that the fuzzy approach offers several advantages: it is more accurate since it incorporates the strength or validity of the previous information, produces a more manageable set of rules for human inspection, and allows the incorporation of contextual information to the mining process expressed in a more human-friendly format.

Keywords: Association rules, Meta-Association rules, Fuzzy Association rules, Information Fusion, High Order Mining, Multiple Databases, Distributed Databases

1. Introduction

In the last years, the attention of Information Fusion researchers has been shifted to consider and integrate Data Mining and Knowledge Discovery techniques in their

*Corresponding author

Email address: {mdruiz,jgomez,miguems,marosiz,mbautis}@decsai.ugr.es (M.D. Ruiz*, J. Gómez-Romero, M. Molina-Solana, M. Ros, M.J. Martin-Bautista)

Fusion systems in order to discover hidden information from the available data. Accordingly, data mining processes can be used to obtain elaborated knowledge that can be afterwards fused with sensor-based data and other information. Likewise, data can be dynamically generated and streamed, and may be as well syntactically and semantically heterogeneous, which are characteristics that have been typically addressed in Information Fusion. Therefore, the advent of new combined proposals is natural, and can be very helpful in application domains where it is necessary to obtain information, in a higher scale, by incorporating previously extracted knowledge, also in the form of patterns, from several sources.

Association rules extraction is a well-established Data Mining technique for discovering information from structured databases. Association rules have the form of implications $A \rightarrow B$, which represent the joint co-occurrence of A and B. To discover such rules the most used framework is that of Support-Confidence. The support measures the percentage of transactions in a dataset containing A and B (i.e. the joint probability $P(A \cap B)$) and the confidence the conditional probability of having B when A is known (i.e $P(B|A)$). Some papers [5, 10] have emphasized some of the drawbacks of the confidence being the most significant that it gives biased value in the presence of very frequent items. To avoid this, an alternative framework was proposed in [10] where the accuracy is measured by means of the *Support-Certainty Factor* approach. The Certainty Factor (CF) yields a value in the interval [-1, 1] and measures how our belief that B is in a transaction changes when we are told that A is in that transaction. Positive values indicate that our belief increases, negative values mean that our belief decreases, and 0 means no change. The CF has been used in many works as for instance in [10, 11, 13, 34, 12] obtaining better results than when employing the confidence, and reducing the number of obtained rules by filtering those rules corresponding to statistical independence or negative dependence.

Typically, association rules algorithms work on the assumption that a fully-available uniform dataset is accessible at the beginning of the mining process. Nevertheless, this is no longer a valid assumption. In some cases, primary data is only available for a short time, as in the case of stream data, which is usually processed in real time and then deleted after storing the analysis results. In other cases, primary data cannot be disclosed, and only summaries of the most relevant conclusions are publicly available. In some cases local patterns offer specialized knowledge that can be useful to find solutions that could not be discovered by the amalgamation of all available data. In addition, it is more and more frequent having distributed datasets with similar semantics (in most cases geographically distributed), in such a way that they cannot be directly merged and must be processed separately. All these situations require a change in our perspective from raw data analysis to pattern analysis. In that regard, Higher Order Mining (HOM) emerges as "the sub-field of knowledge discovery concerned with data mining models that operate upon derived, i.e. non-primary, forms of information" [32].

To that aim we propose in this paper a second-order mining approach that applies association rules in two steps obtaining what we have called *meta-association rules*. The meta-association rules can contain rules, that have been previously extracted from the individual databases, in the antecedent and/or in the consequent. Then a meta-association rule represents association between previously generated association rules and additional attributes containing information about the original databases. Our pro-

posal can thus be employed as an automatic tool for exploring the results provided from multiple databases independently and see their commonalities by means of analysing the already extracted association rules in each of them, and fusing them in what we have called meta-association rules.

The process of mining meta-association rules can be made by conveniently adapting an existent association rule mining approach (e.g. Apriori, Eclat, FP-growth, etc). In a recent work [33] we proposed a mining approach using crisp association rules to discover the meta-rules. However crisp association rule mining approaches have a limitation since they only take into account if an association was previously mined from the databases or not; i.e. an association mined with a confidence of 0.99 has the same importance than another one with a confidence of 0.3.

To address this limitation, in this paper we propose to use the degree of fulfilment of *regular*[1] rules –and by extension, to any attribute considered in the meta-association rule extraction process for computing the assessment of the meta-association rules. For this, Fuzzy Sets theory offers a sound framework for the management of these quantities, allowing a value to belong to the fuzzy set to a degree, and thus avoiding the problem of over or under estimation of values at the boundaries. Specifically, we will employ a recently proposed model using the Representation by Levels theory (RL for short) for managing fuzziness while satisfying all the natural properties of Boolean algebras [35]. Moreover, RLs offer a suitable representation of fuzzy rules that allows the parallelization of the rule mining process [14]. In this paper, we propose a two-step process for mining meta-association rules using the RLs in order to obtain meaningful fuzzy meta-association rules. We show that this kind of meta-rules convey new relevant information that cannot be obtained by regular rules, fusing the information collected by the whole set of regular rules individually extracted from the databases together with information of any other attribute, either contextual or present in the original databases, that can be also expressed in a more human-friendly way by using fuzzy sets. We compare crisp and fuzzy methods and we illustrate their application in a real scenario of crime data analysis.

The paper is structured as follows. Next Section compares our proposal to other related works. In Section 3 we review the fuzzy data mining technique that will be employed for mining fuzzy meta-association rules. Sections 4 and 5 describe our approach for fusing information by means of crisp and fuzzy meta-rules, and Section 6 shows and discusses some interesting experimental results. Finally, Section 7 points out some conclusions and prospective directions for future research.

2. Related Works

The proposal presented in this paper can be of interest in the application areas of association rules mentioned in Section 1. To the best of our knowledge, this is the first attempt to develop an algorithm to fuse rules mined from horizontal distributed datasets. Data mining algorithms, and particularly association rule mining algorithms, have focused their attention on developing efficient methods when data is distributed or

[1] We will employ the adjective *regular* to the association rules discovered in each database individually.

partitioned among different sites. We can distinguish between horizontal and vertical partitions of data. In the former, the transactions are distributed among different sites, on the opposite, in vertical partitions the itemsets are split among sites. Some works have been made in this direction in order to find efficient algorithms for extracting rules in both horizontal and vertically distributed databases. A good comparative study can be found in [20]. Particularly in [8] it is presented the FDM (fast distributed mining) algorithm based on Apriori but trying to minimize the exchange of communications for computing the support of items. In [1] the CDA (count distributed algorithm) was proposed to parallelize the mining process in p partitions, one for each available processor. More works in this line that extract a single set of rules from distributed data are those of [3, 22, 23, 36]. It is also interesting to stress the differences between parallel and distributed association rule mining (ARM for short). In parallel ARM the algorithms divide the data among different nodes, while in distributed ARM each site performs the task independently but must access the whole data set. So distributed ARM usually aims to minimize the communication costs. This is the case of ODAM algorithm proposed in [3] which deals with geographically distributed data sets and tries to reduce the number of exchanged messages.

The terminology of meta-rule has been also employed in association rule mining but with different meanings. In [17] a meta-rule is considered as a template for guiding the association rule extraction process. In higher order mining we have found two different approaches using also the concept of meta-rule. In [9] the authors called temporal meta-rules to those rules found from temporal rules (first-order rules) by a second inference process. However, the meta-rules obtained have the traditional formulation ($A \rightarrow B$) but with a small variability of the estimated confidence among different pre-established time structures. In [4], the authors mine fuzzy rules from association rules, having each association rule assigned a sequence of support and a sequence of confidence values calculated in different time periods. These fuzzy meta-rules are used to capture the changes of the association rules over time, thus obtaining fuzzy meta-rules of the type: *change in support in a period* t_1 = *Fairly decrease* $\rightarrow$ *change in support next period* = *highly decrease*, where appropriate linguistic labels are defined by hand using fuzzy sets. Our concept of meta-rule is different to those proposed in these works [4, 9], since we may have association rules in the antecedent and/or the consequent of the meta-rule, and we do not consider sequences of support/confidence values.

There are other approaches that also apply a two-step data mining process to fuse information. In [27, 29, 30], we can find various proposals aimed at combining association rules mined over a set of clusters. This kind of methods are very interesting for large databases because they scale very well when the number of rows and columns increases.

Along this work, we will see that the proposed methods for mining meta-association rules offer some important benefits compared to previous works: - proposed meta-rules convey new relevant information that cannot be obtained by regular rules, - the proposed algorithm produces a more manageable set of rules for human inspection, and - it allows the incorporation of contextual information to the mining process expressed in a more human-friendly format.

3. Fuzzy Association Rules using RLs

Fuzzy sets have been used in association analysis in different ways [26]. The most representative proposals were (1) those based on the extraction of fuzzy associations in crisp relational databases [25, 24, 19] including those works mining fuzzy itemsets and fuzzy association rules from quantitative transactions [6, 7], and (2) based on the concept of fuzzy transaction [10]. In general, fuzzy sets in (1) were provided in an attribute level by allowing a "soft" definition of interval boundaries, identifying fuzzy sets with linguistic labels and having therefore a more user-friendly representation of rules, and sometimes providing a different granularity of the data. In this paper we focus in proposal (2) where the provided data can be represented by means of a fuzzy transactional database, i.e. the items are satisfied to some extent in the transaction. Formally, given $I = \{i_1, \ldots, i_m\}$ a finite set of items, a *fuzzy transaction* [10] is defined as a non empty fuzzy subset $\tau \subseteq I$, and an item $i \in I$ is satisfied in τ to a certain degree in the unit interval, i.e. $\tau(i) \in [0, 1]$. By extension, if A is a subset of items of I, the membership degree of A to the fuzzy transaction τ is defined as $\tau(A) = \min_{i \in A} \tau(i)$. Then, given a set of fuzzy transactions $\tilde{D}$, a *fuzzy association rule* $A \rightarrow B$ is satisfied in $\tilde{D}$ if and only if $\tau(A) \leq \tau(B)$ for all $\tau \in \tilde{D}$; that is, the membership degree of B is higher than the membership degree of A for all fuzzy transactions τ in $\tilde{D}$. It is easy to see that with these definitions a crisp transaction is a special case of fuzzy transaction, and a crisp association rule is a special case of fuzzy association rule (we refer to [10] for more details).

Following this framework in [14] it is proposed a general formal model for fuzzy rules using the Representation by Levels theory (RL for short). The RL theory was proposed in [35] as a new framework for managing fuzziness while satisfying all the natural properties of Boolean algebras that can not be fulfilled by any standard fuzzy set triple (t-norm, t-conorm, negation) [35, 15, 16].

In [14] the proposed model considers items $i \in I$, and more generally itemsets $A \subset I$, as RL-representations (Λ_A, ρ_A) defined over the fuzzy database $\tilde{D}$, where $\Lambda_A = \{\alpha_1, \ldots, \alpha_m\}$, $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > 0$ is a predefined sets of levels and $\rho_A(\alpha) = \{\tau \in \tilde{D} : \tau(A) \geq \alpha\}$ is a function which applies each level $\alpha \in \Lambda_A$ into a crisp realization in that level.

For any pair of itemsets A, B in $\tilde{D}$, we consider their associated RLs (by abuse of notation we use A and B to name them) $A = (\Lambda_A, \rho_A), B = (\Lambda_B, \rho_B)$. We define a four fold table noted by $\mathcal{M}_{\alpha_i}(A, B, \tilde{D})$ containing the cardinalities in level α_i associated to the partition of $\tilde{D}$ generated by A and B:

$\mathcal{M}_{\alpha_i}$	B	$\neg B$
A	a_i	b_i
$\neg A$	c_i	d_i

where a_i, b_i, c_i and d_i are non negative integers such that $a_i = |\rho_{A \wedge B}(\alpha_i)|$, $b_i = |\rho_{A \wedge \neg B}(\alpha_i)|$ and analogously for c_i and d_i. To assess the validity of fuzzy association rules we can generalize every interest measure from the crisp to the fuzzy case by

means of the basic probability distribution defined as

$$Prob(Y) = \sum_{\alpha_i \in \Lambda \,:\, Y = \rho(\alpha_i)} \alpha_i - \alpha_{i+1} \tag{1}$$

where Y must be the preimage of at least one $\alpha \in \Lambda$. In particular, the obtained measures for support, confidence and certainty factor are in equations (2), (3) and (4) respectively [14]. The crisp version of these measures can be easily obtained by setting $\Lambda = 1$ in the mentioned equations.

$$\mathrm{FSupp}(A \to B) = \sum_{\alpha_i \in \Lambda_A \cup \Lambda_B} (\alpha_i - \alpha_{i+1}) \left(\frac{a_i}{a_i + b_i + c_i + d_i} \right) \tag{2}$$

$$\mathrm{FConf}(A \to B) = \sum_{\alpha_i \in \Lambda_A \cup \Lambda_B} (\alpha_i - \alpha_{i+1}) \left(\frac{a_i}{a_i + b_i} \right) \tag{3}$$

$$FCF(A \to B) = \sum_{\alpha_i \in \Lambda_A \cup \Lambda_B} (\alpha_i - \alpha_{i+1}) \cdot \begin{cases} \dfrac{a_i d_i - b_i c_i}{(a_i + b_i)(b_i + d_i)} & \text{if } a_i d_i > b_i c_i \\ 0 & \text{if } a_i d_i = b_i c_i \\ \dfrac{a_i d_i - b_i c_i}{(a_i + b_i)(a_i + c_i)} & \text{if } a_i d_i < b_i c_i. \end{cases} \tag{4}$$

The calculation of the confidence may have indeterminations of the form "$\frac{0}{0}$" when $|\rho_A(\alpha_i)| = 0$. This happens when no transactions exists satisfying the antecedent. So, to preserve the definition of fuzzy rule we take the value 1 for that indetermination. For the certainty factor, FCF, the RLs of both itemsets must be normalized (i.e. A or B must be satisfied in at least one transaction at level $\alpha = 1$, if not, their degrees of satisfiability should be normalized by the maximum value reached by A or B).

In the following, we present an example to show the process of computing the support, the confidence and the certainty factor of fuzzy rules using the RL formalization.

3.1. Example

Let $I = \{i_1, i_2, i_3, i_4, i_5\}$ and $\tilde{D}$ be the fuzzy database in Table 1. Let A and B be the itemsets $A = \{i_1, i_3\}$ and $B = \{i_4\}$ defined over $\tilde{D}$. Their degrees of satisfiability in each transaction are computed using the minimum as for instance $\tau_2(A) = min\{1, 0.8\} = 0.8$. We then take their RL-representations (Λ_A, ρ_A) and (Λ_B, ρ_B) and we compute the resulting RL-representations from applying the negation (i.e. the complement) to A and B and the partition[2] of $\tilde{D}$ using the conjunctions of A, B and their negations. In Table 2 we show the RLs associated to some of these conjunctions.

[2] By considering the RL-representations associated to $A \wedge B$, $A \wedge \neg B$, $\neg A \wedge B$ and $\neg A \wedge \neg B$ we obtain a partition of $\tilde{D}$.

Table 1: Set of fuzzy transactions $\tilde{D}$

	i_1	i_2	i_3	i_4	i_5
τ_1	1	0.2	1	0.8	0.9
τ_2	1	1	0.8	0	0
τ_3	0.4	0.1	0.7	0.6	0
τ_4	0.6	0	0.4	0.4	0.5
τ_5	0.4	0.1	0	0	0
τ_6	0	1	0	0	0

Table 2: RL-representations of A, B, $\neg A$ and $\neg A \wedge B$ for the set of levels $\Lambda = \{1, 0.8, 0.6, 0.4, 0.2\}$.

α_i	ρ_A	ρ_B	$\rho_{\neg A}$	$\rho_{\neg A \wedge B}$
1	$\{\tau_1\}$	$\emptyset$	$\tilde{D} \backslash \tau_1$	$\emptyset$
0.8	$\{\tau_1, \tau_2\}$	$\{\tau_1\}$	$\{\tau_3, \tau_4, \tau_5, \tau_6\}$	$\emptyset$
0.6	$\{\tau_1, \tau_2\}$	$\{\tau_1, \tau_3\}$	$\{\tau_3, \tau_4, \tau_5, \tau_6\}$	$\{\tau_3\}$
0.4	$\{\tau_1, \tau_2, \tau_3, \tau_4\}$	$\{\tau_1, \tau_3, \tau_4\}$	$\{\tau_5, \tau_6\}$	$\emptyset$
0.2	$\{\tau_1, \tau_2, \tau_3, \tau_4\}$	$\{\tau_1, \tau_3, \tau_4\}$	$\{\tau_5, \tau_6\}$	$\emptyset$

Table 3: $\mathcal{M}_{\alpha_i}(A, B, \tilde{D})$ for $A = \{i_1, i_3\}$ and $B = \{i_4\}$ where $\alpha_i \in \Lambda$.

	a_i	b_i	c_i	d_i
$\mathcal{M}_1$	0	1	0	5
$\mathcal{M}_{0.8}$	1	1	0	4
$\mathcal{M}_{0.6}$	1	1	1	3
$\mathcal{M}_{0.4}$	3	1	0	2
$\mathcal{M}_{0.2}$	3	1	0	2

The association rule mining process is the following: for every level in Λ the four fould tables $\mathcal{M}_{\alpha_i}(A, B, \tilde{D})$ are calculated[3] (see Table 3 where each $\mathcal{M}_{\alpha_i}$ is represented in a row) and then the values for the support, the confidence and the certainty factor are computed. The discovered rules are those exceeding the minimum thresholds. In the case of the itemsets A and B previously defined, we compute the FSupp,

[3]Theoretically we can compute the four fold tables for every level in $\Lambda_A \cup \Lambda_B$ but in practice we fix the RL-set into a finite set of levels in the unit interval because the possible different levels that can appear in a database could be very large.

FConf and FCF as follows:

$$\begin{aligned}
\mathrm{FSupp}(A \to B) = \mathrm{FSupp}(\{i_1, i_3\} \to \{i_4\}) &= (1-0.8)\frac{0}{6} + (0.8-0.6)\frac{1}{6} + \\
&+ (0.6-0.4)\frac{1}{6} + (0.4-0.2)\frac{3}{6} + 0.2 \cdot \frac{3}{6} = 0.167 \\
\mathrm{FConf}(A \to B) &= (1-0.8)\frac{0}{1} + (0.8-0.6)\frac{1}{2} + (0.6-0.4)\frac{1}{2} + (0.4-0.2)\frac{3}{4} + \\
&+ 0.2 \cdot \frac{3}{4} = 0.5 \\
\mathrm{FCF}(A \to B) &= (1-0.8)\frac{0}{6} + (0.8-0.6)\frac{4}{10} + (0.6-0.4)\frac{2}{8} + (0.4-0.2)\frac{6}{12} + \\
&+ 0.2 \cdot \frac{6}{12} = 0.33
\end{aligned}$$

4. Fusing Information with Meta-Association Rules

Meta-association rules aim at fusing information obtained from a set of databases whose provenance can be diverse or can proceed from the same database which had been geographically distributed or had been horizontally partitioned for storage, security or other reasons. Specifically, we want to discover possible associations from the set of association rules that have been extracted from each primary dataset. This can be useful in those scenarios in which datasets are too large, complex and heterogeneous. Examples of this type of datasets appear, for instance, when databases collect sensor data with different provenance (e.g. light or movement sensors, video), and when data is partitioned across different collection or storage places. To obtain meaningful meta-rules involving the same kind of items, the primary datasets must have similar structure and semantics, although their structure do not need to be exactly the same.

Let us explain the problem with an example. Let us suppose that a multi-branch organization, e.g. a banking company, has several branches spread across the country. In this case, the data stored by each individual branch will have a similar structure. The organization is then interested in analysing its profits according to customers' data. There are two options to do so: the first one is to analyse data globally, i.e. to analyse the whole available data by compiling a very large dataset by merging all local data, whilst the second one is to analyse and obtain relevant information from local patterns obtained separately in each office. When the second approach is chosen, the number of rules for being inspected can be overwhelming. For this reason, in this paper, we propose the use of meta-association rules to facilitate the analysis in the second kind of scenarios. Fusing extracted knowledge by means of meta-rules has some advantages. First, dealing with the complete dataset is not necessary, which increases the efficiency. Second, it allows working with the results/patterns obtained from the individual databases, reducing the time needed in the rule mining process and obtaining, in most cases, a smaller set of rules which will be easier to examine. Additionally, it is not necessary for all the data sources to have exactly the same structure. Moreover, meta-rules facilitate the analysis of trends at a higher abstraction level, since we can study the differences between the rules rather than the differences between data points. Last but

not least, meta-association rules facilitate the incorporation of contextual knowledge to the fusion process, for example demographic (e.g. population) or economic data (e.g. unemployment rate, income level, etc.) about the places where the divisions are sited. On the other hand, since the meta-rule analysis works on already "summarized" data (in the form of association rules), there is a unavoidable loss of information.

4.1. Meta-Association Rules Mining Process

The general architecture of our proposal is depicted in Figure 1. We mainly distinguish two major steps. First, regular association rules are obtained from each database, represented by D_i. Second, the resulting associations are fused in the form of meta-association rules. For this purpose, we collect the information obtained in the first step in a new database named *meta-database*, $\mathfrak{D}$, that will be fuzzy or boolean (see 2) depending on the use or not of the accuracy of regular rules extracted in the first step. We then can differentiate between fuzzy or crisp meta-association rules. We describe in detail this process right after.

Figure 1: Process flow: From original datasets to final meta-association rules.

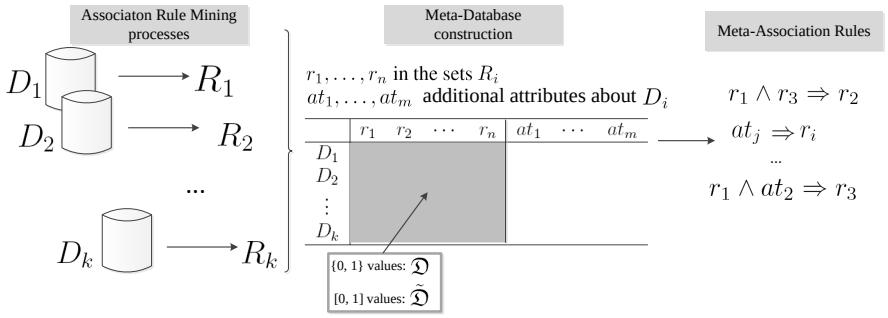


Figure 2: Examples of boolean (left table) and fuzzy (right table) meta-databases.

$\mathfrak{D}$	r_1	$\cdots$	r_n	at_1	$\cdots$	at_m
D_1	1	$\cdots$	0	1	$\cdots$	1
D_2	0	$\cdots$	0	0	$\cdots$	1
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
D_k	1	$\cdots$	1	1	$\cdots$	0

$\tilde{\mathfrak{D}}$	r_1	$\cdots$	r_n	at_1	$\cdots$	at_m
D_1	0.2	$\cdots$	1	0.9	$\cdots$	1
D_2	0	$\cdots$	0.6	0	$\cdots$	0.2
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
D_k	0.9	$\cdots$	0.5	1	$\cdots$	0.1

Step 1. Let $\{D_1, D_2, \ldots, D_k\}$ be a set of databases sharing some of their attributes. After applying a rule extraction procedure to each database[4] a different set of

[4]For the sake of simplicity, we have considered that the regular rules obtained in the first stage are crisp. However, it can be seen that the same process can be also applied to other types of rules. In particular, fuzzy

association rules R_i is obtained for each D_i. Then we have in the sets $R_1, R_2, \ldots, R_k$ information about the discovered association rules and their assessment values. We want to remark that the number of rules within $R_1, R_2, \ldots, R_k$, can be different, as well as the number of items in the antecedent or in the consequent of each rule. For the success of the process, it is worth to notice that there may be some common rules among sets R_i. Without loss of generality, we assume that the same thresholds for the minimum support and certainty factor values have been used when processing each dataset.

Step 2. In the second stage, we collect the information from the first step by creating a structured database, $\mathfrak{D}$, namely *meta-database*. All the different rules r_j pertaining to at least one of the sets R_i, $i = 1, \ldots, k$ will be attributes (i.e. columns) in the produced meta-database. An additional feature of $\mathfrak{D}$ is that it can be enriched with data describing further characteristics of D_i, which are represented in $\mathfrak{D}$ by means of new attributes $at_1, \ldots, at_m$. The meta-database is finally used as the input of the meta-rule extraction process. Depending on the type of values of the meta-database we distinguish two different strategies to extract the meta-rules.

4.2. Crisp Meta-Association Rules

Crisp meta-association rules are discovered by following a crisp association rule mining strategy. This first approach was first presented in [33], where the created meta-database $\mathfrak{D}$ is a boolean database. Therefore, in this case $\mathfrak{D}$ concentrates information about the presence/absence of a rule r_j in the sets $R_1, \ldots, R_k$, representing only if a given regular rule has been generated or not in the first step of the process. Additionally, $\mathfrak{D}$ can also contain some extra information in the form of crisp attributes. We depict this case in the left table of Figure 2, in which we assign to the position (i, j) of $\mathfrak{D}$ the value 1 if the rule r_j is found in D_i, and 0 otherwise. Afterwards, *crisp meta-association rules* are extracted from $\mathfrak{D}$ by a classical association rule mining algorithm. Crisp meta-association rules thus represent the co-occurrence of rules, rules and attributes, or attributes in the meta-database.

Formally, we can obtain three types of meta-association rules[5]

- $ar_i \Rightarrow ar_j$, where ar_i, ar_j are conjunction of rules (or a rule when ar only involves a rule). For example, if $ar_i = r_1 \wedge r_3$ and $ar_j = r_2$ then $r_1 \wedge r_3 \Rightarrow r_2$ is a meta-association rule.
- $ats_i \Rightarrow ats_j$, where ats_i, ats_j are conjunction of attributes (or an attribute when ats only contains one attribute).
- $ar_i \Rightarrow ats_j \wedge ar_k$ or $ats_j \wedge ar_k \Rightarrow ar_i$, where ar_i, ar_k represent conjunction of rules and ats_j a conjunction of attributes, and they can be mixed; e.g. we can find a meta-association rule of the form $r_1 \wedge at_2 \Rightarrow r_3 \wedge at_4$.

rules can be used, because they provide the same kind of information: the rule itself, the (fuzzy) support and the (fuzzy) CF.

[5]To differentiate between *regular* association rule and *meta*-association rule we will employ the symbols $\rightarrow$ and $\Rightarrow$ respectively.

As it was mentioned in the introduction, this procedure has limitations, because it does not collect all the available information (i.e. support and validity of the rules) since it only takes into account if a rule has been previously mined from a database or not. This implies that regular rules found with different intensity (e.i., $\mathrm{CF}(r_i) \gg \mathrm{CF}(r_j)$) have the same importance in the boolean meta-database. To solve this issue, assessment measurements can be incorporated into the meta-database by using intervals, thus obtaining items of the form $\langle r_i, interval \rangle$. For instance, if the CF is used, the intervals are subsets of $[minCF, 1]$, and the items have the form $\langle r_i, (\mathrm{CF}_1, \mathrm{CF}_2] \rangle$, where $\mathrm{CF}_j \in [minCF, 1]$, $j = 1, 2$. However, this approach is problematic due to the crisp boundaries of intervals. For example, given the values $\mathrm{CF}(r_i) = 0.75$ and $\mathrm{CF}(r_j) = 0.76$, and the intervals $(0.5, 0.75]$ and $(0.75, 1]$, the CF values would lie in different intervals even though they are very close. This issue motivates our proposal for using RLs to represent these continuous values lying in the unit interval in a gradual form.

4.3. Fuzzy Meta-Association Rules using RLs

Fuzzy transactions consist on a fuzzy subset of items, where the degree of each item is interpreted as its membership degree to the transaction. In the case of a fuzzy meta-database $\tilde{\mathfrak{D}}$, the transactions are the initial databases D_i and the items are the discovered rules $r_j \in R_i$. It seems thus coherent to use the assessment value –in our framework we will employ the certainty factor– as the degree of satisfiability of the rule in the dataset. The general form of the fuzzy meta-database is depicted in right table of Figure 2, where the value in position (i, j), i.e. in row D_i and column r_j is the certainty factor of the rule r_j in database D_i. In contrast to the crisp case, the values of the fuzzy meta-database are in the unit interval –more specifically, in the interval $[minCF, 1]$.

In this framework, the supplementary information can be supplied by means of crisp and fuzzy attributes describing some characteristics or information about the original databases. Particularly interesting are the suitability of fuzzy sets to represent imprecise properties by using linguistic terms plus the degree of fulfilment of the object of that property. They are also useful to represent geographic, demographic or economical statistics which are usually provided by percentages. Moreover, we can express expert knowledge in a comprehensive way for the experts and the final user of the system. For example, we can say that the poverty index of a region is low with a degree of 0.9, which means that the poverty index is low with a high degree.

Once the fuzzy meta-database has been built, the fuzzy meta-rules are obtained by applying a fuzzy association rules mining algorithm. We use the framework using RLs described in Section 3 considering the FSupp and FCF measures which is detailed in next section. Similarly to the crisp case, the fuzzy meta-rules can be classified in three different types

- fuzzy meta-rules that relate only rules/conjunction of rules,
- fuzzy meta-rules that relate only attributes/conjunction of attributes, and
- fuzzy meta-rules that relate conjunction of rules and attributes.

The external appearance of fuzzy meta-rules is the same as the crisp meta-association rules, but the mode of discover them is different because we utilize different information (the assessment degree of the rules to the databases) and the assessment measures (FSupp, FCF) conveniently adapted to that aim. More important is also to notice that the extracted fuzzy meta-rules represent associations that have a high certainty factor in the original datasets, rather than just presence as in the crisp case.

5. Algorithm

In this section, we describe the algorithm that implements the complete extraction procedure from the original datasets to the output: the set of crisp/fuzzy meta-association rules. In the literature we can find several approaches for discovering crisp or fuzzy association rules. Most of these algorithms work in two steps (like Apriori [2]). In the first one, the so-called *frequent itemsets* or *candidate set of items* are calculated. This step can be seen in lines 6-12 of Algorithm 1. In a second step, those rules with accuracy above a user-defined threshold are extracted (lines 13-15). The first step is the most computationally expensive, and different strategies and heuristics have been proposed to reduce the time expended in the mining process. In our proposal we use a binary representation of items using bit strings to speed up the computations of conjunctions [13]. Likewise, using the binary representation the memory footprint is not high, decreasing the memory requirement of the system.

The complete process is described in Algorithm 1. We start mining the association rules satisfying the support and certainty factor requirements in each of the original databases $D_1, \ldots, D_k$. Next, the meta-database $\mathfrak{D}$ or the fuzzy meta-database $\tilde{\mathfrak{D}}$ is created. In this step, the supplementary information may be added as attributes to the meta-database. For the case of the fuzzy meta-database the attributes of $\tilde{\mathfrak{D}}$ can be modelled as fuzzy sets, as previously explained. Finally, crisp or fuzzy meta-association rules are respectively extracted. Clearly, when the initial datasets are not available, the meta-rule extraction procedure would start at step 18.

Step 22 starts when the fuzzy meta-database is available, and then any association rule mining process can be applied. In particular, we execute once steps from 2 to 16 for $\mathfrak{D}$ as input (instead of D_i), obtaining thus fuzzy meta-association rules as output. In our implementation, we use the algorithm presented in [14] and briefly described in Algorithm 2, which computes the fuzzy assessment measures FSupp and FCF of equations (2) and (4) by parallelizing on the set of levels in Λ, specifically in step 9, the crisp support and the certainty factor are computed independently for each $\alpha \in \Lambda$ and then the weighted sums of equations (2) and (4) are calculated, like it is explained in Example 3.1.

As previously mentioned, association rule mining algorithms typically work in two steps. The computational complexity of this kind of algorithms is $\mathcal{O}(|D|\, 2^{|I|})$. In our case, due to the parallel processing of the original D_i, the first stage (lines 1-17) has the same complexity as an standard algorithm, obviously depending in each case on the number of transactions $|D_i|$ and the number of items (that can be in principle different for each D_i). The complexity of the second stage (lines 18-22) will depend on k, the number of initial databases; m, the number of additional attributes; and r, the number of different rules obtained in the first step. Then the complexity for both cases (crisp

Algorithm 1 Meta-association rules mining

Input: $D_1, \ldots, D_k, at_1, \ldots, at_m, minsupp, minCF$
Output: $R_1, \ldots, R_k$ and MR (the set of meta-association rules)

for all D_i such that $1 \leq i \leq k$ **do**
  # D_i *preprocessing*
  Read D_i and store the items I
  Transform D_i into a boolean database(or a set of boolean datasets when D_i is fuzzy)
  # Mine association rules
  if $Supp(X) \geq minsupp$ **then**
    $X \in C$ (C = candidate set)
  end if
  for all $X, Y \in C; X \cap Y = \emptyset$ **do**
    if $Supp(X \rightarrow Y) \geq minsupp$ **then**
      $X \wedge Y \in C$ **and** $X \rightarrow Y$ is frequent
    end if
    if $CF(X \rightarrow Y) \geq minCF$ **then**
      $X \rightarrow Y \in R_i$ **and** $X \rightarrow Y$ is certain
    end if
  end for
end for
# Meta-database $\mathfrak{D}$ *creation*
  Compile all different rules from $R_1, \ldots, R_k$
  Create $\mathfrak{D}$ using compiled rules and additional attributes
 # Mining meta-association rules
Repeat steps 2-16 with $\mathfrak{D}$ as input to mine meta-association rules

Algorithm 2 Fuzzy Association Rule Mining Procedure using RLs

Input: Λ, Fuzzy $\tilde{D}$, $minFSupp$, $minFCF$,
Output: FR (Set of fuzzy association rules)

 # $\tilde{D}$ preprocessing
  Read $\tilde{D}$ and store the items I
  Transform $\tilde{D}$ into p boolean databases (p is the number of levels in Λ)
  Encode databases into p vectors of `BitSets` (one vector for each $\alpha_i \in \Lambda$)
 # Mining fuzzy rules
for all $\alpha_i \in \Lambda$ such that $1 \leq i \leq p$ **do**
  Repeat steps 6-16 of Algorithm 1 to obtain the set of crisp rules in each level.
  Read all the found rules in level α_i
  Compute progressively the measures FSupp and FCF using formulas (2) and (4)
end for
Collect those rules satisfying the FSupp and FCF requirements (i.e. FSupp and FCF values above $minFSupp$ and $minFCF$ respectively)

and fuzzy) is $\mathcal{O}(k2^{m+r})$, since for the fuzzy case (see Algorithm 2) the process is made in parallel for each $\alpha \in \Lambda$.

6. Experiments

In order to emphasize the main differences and compare traditional ARM approach with the MARM, we have performed several experiments with synthetic and real data. A more detailed description of different datasets is in the subsequent section.

Depending on the analysis to be performed, the conducted experimentation can be divided in several groups:

1. To analyse and compare the results obtained using traditional ARM algorithms with those obtained by employing MARM. To that purpose, we extract regular association rules from a database composed of all datasets, and compare the results with the meta-rules obtained with the partitioned data.

2. To compare different aspects of the two presented approaches for discovering the meta-association rules: (a) using the crisp approach, that builds the crisp meta-database ($\mathfrak{D}$), and (b) using the fuzzy meta-database ($\tilde{\mathfrak{D}}$).

3. To make an experimental study about the scalability of the employed MARM algorithms.

All the experimentation has been made in a desktop computer equipped with a 2.5GHz Pentium Dual Core processor and 3GB of RAM running Java 8 on Windows 7. Without lack of generality, in order to obtain readable rules, we have limited the rules obtained in the first step of the process to have one item in the antecedent and one in the consequent. For the meta-association rules, we allow two items at most in the antecedent and the consequent.

6.1. Datasets description

6.1.1. Synthetic dataset

The `synth` dataset contains artificial data divided into 8 databases, each one with 15 transactions and 6 attributes. Every attribute has 6 different possible values, giving a total number of 36 items of the form $\langle attribute, value \rangle$.

The `synth` dataset is a small dataset employed to study and explain the main differences between traditional association rule mining algorithms and proposed meta-association rule mining for fusing information.

6.1.2. Real Data about crimes

The (`chicago`) dataset includes data of crime incidents by district reported by the Chicago police[6] in 2012, plus additional socio-economical data about the neighborhood where the crime happened. Specifically, we have selected six different attributes (quarter of the year in which the incident happened, day period, crime description, location description, arrest and domestic-crime) obtaining around 300 items (pairs $\langle attribute, value \rangle$). The dataset comprises 22 databases[7] each one corresponding to

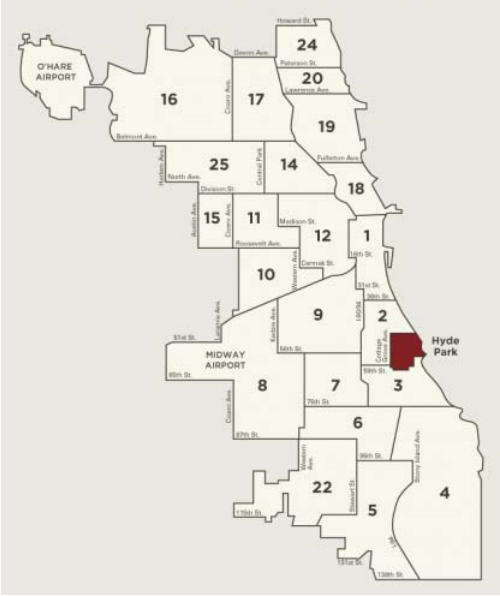
[6] `https://data.cityofchicago.org/`

[7] Note that district ids are not consecutive and there are 22 different districts instead of 25.

Table 4: Description of additional attributes added for each police district.

Additional attributes	Possible values
Number-of-Students	Low, Medium, High, Very High
Number-of-Misconducts	Low, Very low, Medium, High, Very High
Safety-Index	Low, Medium, High

a different police district of the city (see Figure 3). For the meta-database, we have considered additional attributes of schools aggregated by district (number of students, number of misconducts and perceived safety index) which add 12 new items to the meta-database (see Table 4). Note that, in total, we are dealing with more than 330000 transactions and around 6400 items.



District Id	# transactions	# items
01	12 160	268
02	13 481	286
03	17 730	297
04	19 847	302
05	15 296	290
06	19 301	303
07	20 195	298
08	22 493	333
09	16 704	305
10	15 071	301
11	21 873	311
12	15 930	296
14	12 580	296
15	14 434	285
16	10 851	296
17	9 702	282
18	14 247	272
19	15 689	284
20	5 694	252
22	10 797	275
24	9 538	267
25	19 726	301

Figure 3: **Left:** Image from `https://commonsense.uchicago.edu/page/crime-information-and-statistics` about police districts in the city of Chicago. **Right:** Description of datasets composing the `chicago` database.

6.1.3. Synthetic datasets for scalability purposes

Several synthetic databases called `nX-Synth` for $n \in \{1, 2, 3, 4, 5, 6, 8, 10\}$ have been generated using available data randomly taken from the `Chicago` dataset in order to obtain 8 different sets of databases. Every `nX-Synth` is composed by 22 databases

whose size has been increased n times, i.e. the number of transactions of `2X-Synth` is twice the number of transactions of `1X-Synth`, and analogously for the reminder synthetic databases.

6.2. Results with Synthetic Data

Meta-association rule mining aims to produce fused information in the form of association rules that may have in the antecedent and/or the consequent a previously extracted rules. In order to being able to compare the rules obtained when all data is merged using traditional Association Rule mining (ARM) with the meta-rules discovered using the Meta-Association Mule Mining (MARM) when data is composed of a set of databases, we have employed the `synth` dataset.

For the comparison, two experiments have been carried out: (A) we have mined regular rules in a database that aggregates all the transactions appearing in the 8 datasets; and (B) we have mined meta-rules using the corresponding (fuzzy) meta-database $\mathfrak{D}$ or $\tilde{\mathfrak{D}}$. In both cases we imposed the same thresholds: $minSupp = minFSupp = minCF = minFCF = 0.2$.

From this experimentation we observed that some of the rules found in (A) were the same rules obtained separately in each of the 8 datasets, and they thus appear in some of the meta-rules obtained in (B), i.e. imagine that the rule $r_1 : i_1 \rightarrow i_2$ is found in (A), then it could happen that exists some meta-association rules that contain r_1 in the antecedent or the consequent, for instance $r_1 \wedge at_1 \Rightarrow r_2$. This is comprehensive since a rule that is found in the majority of datasets (and thus appear in a meta-rule) is reasonable to be also found in the merged-database.

Nevertheless, there are many rules that are extracted using (A) but they do not appear in any meta-rule of (B). The reason of this is that in the compiled database, the overall support and CF values are high enough to exceed the threshold, whilst in the partitioned datasets the transactions supporting these measures are only concentrated in a few of the partitioned datasets, and as a consequence they do not appear in any meta-rule in (B). In the other way round, there are many rules appearing in meta-rules of (B) that are not extracted when we merge all datasets (A). This is also perfectly reasonable since some rules obtained in different datasets may not have enough support or certainty factor values to exceed the imposed thresholds when the database is compound with the whole set of transactions.

The intuition can also suggests that some groups of items should appear in regular rules found in (A) and meta-rules found in (B), but this is not true in general. Let us extend this idea with an example. For instance, we may expect that items i_1, i_2, i_3 should be related somehow, because the rules $r_1 : i_1 \wedge i_2 \rightarrow i_3$ and $r_2 : i_1 \rightarrow i_2 \wedge i_3$ have been extracted in (A). Although r_1 and r_2 do not appear explicitly in any meta-rule, one could expect that "similar" rules involving previous items will appear in a meta-rule such as for instance $(i_1 \rightarrow i_2) \Rightarrow (i_1 \rightarrow i_3)$ that relates two rules –different to r_1 and r_2– which contain the items i_1, i_2, i_3. This is not true in general. In fact, we have discovered many meta-rules that do not have the same combination of items of any regular rule extracted in (A) and vice versa. The rationale behind this is the same as in the previous case.

6.3. Results with Chicago Real Data

For highlighting the differences between ARM and MARM, we have conducted several experiments by merging all the existent data which is divided by districts. The first issue to address is how to incorporate the additional information given by the attributes described in Table 4. Since this information is already aggregated by districts, it does not make sense to add them as new transactions of the merged-database. However, one possibility could be to add them as new items, but since we only have one value per district we have to repeat the same value for all the transactions (in our case crime incidents) occurred in the same district. We thus compared the rules obtained in the merged-database in both cases, with and without incorporating the additional attributes. In this case we have varied the $minSupp$ and $minCF$ thresholds to different values obtaining the results depicted in Figure 4. We can see that adding the additional attributes leads to the discovery of more association rules; involving, generally, only the additional attributes incorporated. This is reasonable due to the repetition of the values in all the transactions of the same district.

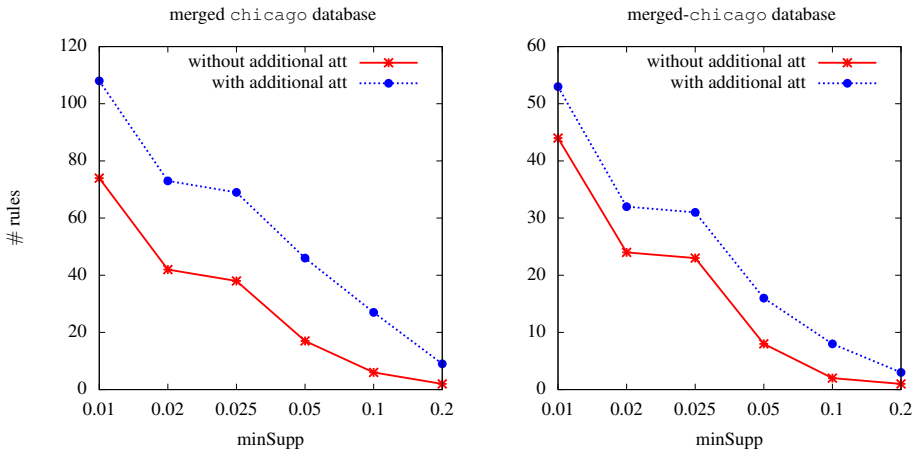


Figure 4: Number of association rules obtained in the merged-chicago database for $minCF = 0.2$ (**left**) and $minCF = 0.5$ (**right**).

In consequence, using the results obtained and the previous analysis made with both, the synth and the chicago datasets, we can conclude that, the meta-rules obtained –and therefore the information that they represent– when the datasets are partitioned are not the same as the traditional association rules discovered in a single database containing all the transactions. The rationale behind this is that meta-rules are obtained from a particular configuration/distribution of data, and will thus depend on that. Beside this, the meaning conveyed in each case is slightly different, because in the former case we have associations between items in the complete set of transactions, and in the second case associations between associations that hold in a significant group of databases (districts for the chicago data). In any case, let us highlight that (1) meta-rules mining supports the incorporation of additional knowledge about the

databases, whilst it is not always possible or does not make sense to add this information for the merged data; and (2) using the complete dataset is not possible in several scenarios (e.g. streaming data which is not stored), and therefore there is no alternative to the use of meta-association rules, or a similar high order mining technique.

A second battery of experiments have been made to analyse the performance of crisp and fuzzy approaches for mining meta-association rules. To this aim we have set the $minSupp$ and $minFSupp$ to 0.1 and 0.05, and we have compared the number of meta-rules discovered and the execution time for different values of $minCF$ and $minFCF$, particularly in $\{0.2, 0.3, \ldots, 0.9\}$. The results are shown in Figures 5 and 6. In the graphs of Figure 5 it is clearly shown that the number of meta-rules in the crisp approach exceeds the number of rules obtained with the fuzzy approach, being significantly larger for small values of $minCF/minFCF$. Besides, the big difference of the number of crisp meta-rules when the $minCF$ threshold is slightly increased, confirms the problems previously mentioned about the graduality when employing the boolean database. In contrast, in the fuzzy case the reduction in the number of rules retrieved is softer, because of the use of the rule's CF to build the fuzzy meta-database. Therefore, we can affirm that fuzzy meta-rules are more appropriate than crisp meta-rules, since the number of rules has less variability and is more manageable for human inspection.

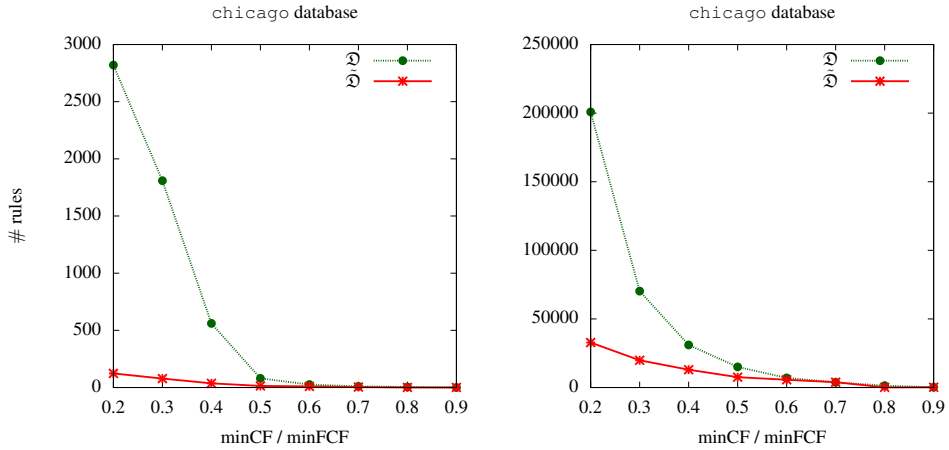


Figure 5: **Left:** Number of crisp/fuzzy meta-association rules (y-axis) for different $minCF/minFCF$ (x-axis) when $minSupp/minFSupp = 0.1$.
Right: Number of crisp/fuzzy meta-association rules (y-axis) for different $minCF/minFCF$ (x-axis) when $minSupp/minFSupp = 0.05$.

Regarding the execution time (see graphs in Figure 6), it strongly depends on the minimum support used, because this determines the number of regular association rules obtained in the first stage of the algorithm (before constructing the meta-database). On one hand we can observe that in the crisp case the MARM algorithm is not affected by the large number of meta-rules obtained, whilst in the fuzzy case the graphs have the same curve. The rationale of this is that in the fuzzy version the algorithm computes

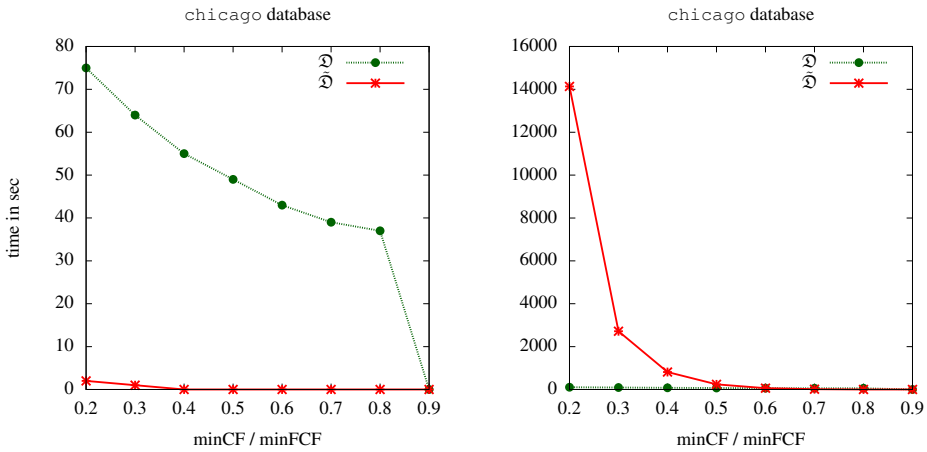


Figure 6: **Left:** Time in sec. for mining crisp/fuzzy meta-association rules (y-axis) for different $minCF/minFCF$ (x-axis) when $minSupp/minFSupp = 0.1$.
Right: Time in sec. for mining crisp/fuzzy meta-association rules (y-axis) for different $minCF/minFCF$ (x-axis) when $minSupp/minFSupp = 0.05$.

the assessment values of the rules by levels, and then combines them to obtain the final FCF values. Low $minFCF$ values result in a notable growth of the rules that need to be considered at each level, and consequently, a larger number of evaluations is performed.

6.4. Scalability study with synthetic data

One of the problems when mining association rules with a considerable size of items is its posterior analysis. Usually, the number of association rules to be inspected contains hundreds or even thousands of rules, incrementing the time-consumed by the expert for finding potentially new and useful knowledge. To tackle this issue is often proposed the tuning of imposed thresholds (minimum support and/or minimum confidence) in order to adjust the number of rules to be discovered, maintaining thus those with a high support and/or confidence. Other procedure, is to change the confidence measure choosing a stronger measure or other type of measures that better fits to our problem. Examples of this are the certainty factor employed here [5], the lift interestingness measure [31], and many others examples that can be found in the literature [18]. There exist other remarkable works dealing with this problem applying other kind of methodologies such as constrained-based frameworks [28], where the analyst can prune the set of association rules by specifying some criteria, or using ranking summaries that order and reduce the resulting association according to some diversity measures [21].

Meta-association rules can be employed to obtain a summarized view of previously extracted rules from multiple databases. In this way, the analyst does not have to analyse the results given in every dataset, since the set of meta-rules fuses the more relevant information since it contains the frequent rules, i.e. those rules that have been

extracted in the majority of the databases. In this sense, our proposal deals with the problem of analysing the big amount of rules that is usually found in association rule mining, particularly when analysing multiple databases.

Concerning the scalability performance of our proposal, we studied the computational complexity of mining meta-association rules (see Section 5), which depends on the number of different rules found in the set of databases to be analysed, i.e. it depends on the cardinality of $R_1 \cap \cdots \cap R_k$. This dependency is also reflected in previous graphs (Figures 4 and 6). Additionally, we have conducted several experiments using different sizes of the `nX-Synth` for $n \in \{1, 2, 3, 4, 5, 6, 8, 10\}$ in order to analyse whether our algorithm is independent on the size of databases. The obtained results are highlighted in Figure 7 where we observe that there is no regularity, since the performance depends on the number of collected rules from the primary datasets.

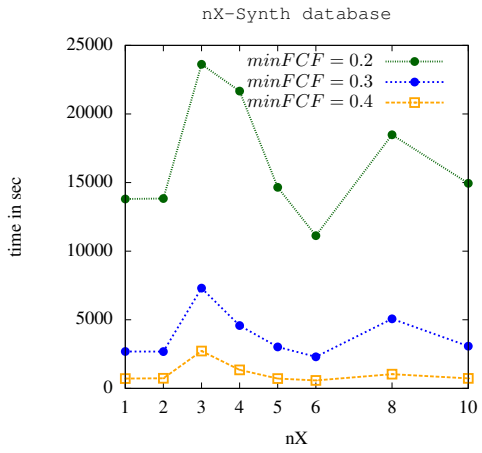


Figure 7: Time in sec. for mining fuzzy meta-association rules (y-axis) for different sizes of `nX-Synth` (x-axis) when $minFSupp = 0.05$.

6.5. Interpretability of Meta-Rules

Table 5 collects some crisp and fuzzy meta-rules discovered in the `chicago` datasets where $Desc.$ stands for Description and f for false. With these meta-rules we exemplify that some meta-rules have been extracted using the crisp approach and not in the fuzzy one and the other way round. Moreover, we can see different types of meta-rules expressing relations between only rules, only attributes, and attributes and rules together. Since meta-rules extracts information from multiple datasets, their meaning in the case of the Chicago database is that we frequently have co-occurrence of some interesting events in many of the districts.

For instance, the first meta-association rule states that there is a very accurate ($FCF = 1$) co-occurrence relation between a low number of students, a very high number of misconducts and a low perception of safety, that is supported by the 9,1% of districts (around 3), which are not a lot. However the second meta-rule has a

higher support (almost a half) and relates three different rules. The last meta-rule relates the very high number of misconducts in a district with the rule (possession of cannabis$\leq$ 30 grams $\rightarrow$ Domestic=false), and the low safety-index. This can be explained as follows: in those districts with a very high number of misconducts, it is frequent and reliable to have (i) a low perception of security; and (ii) a relation between the low possession of cannabis and its occurrence in non-domestic environments. These meta-rules could be useful to the police to determine which actions should be taken to improve, for instance, the safety index perception, by trying to minimize the number of misconducts that are due to the consumption of some kind of drugs.

Antecedent	**Consequent**	$Supp$ / $FSupp$	CF / FCF	$\mathfrak{D}$	$\tilde{\mathfrak{D}}$
Number-of-Students=Low **AND** Number-of-Misconducts=Very high	Safety-Index=Low	0.091	1	X	✓
(Crime-Desc.=$\leq$ 500$ $\rightarrow$ Domestic=f) **AND** (Crime-Desc.=TO VEHICLE $\rightarrow$ Arrest=f)	(Location-Description=STREET $\rightarrow$ Domestic=f)	0.455	0.778	✓	X
(Arrest=f $\rightarrow$ Domestic=f) **AND** Safety-Index=Low	Number-of-Students=Medium	0.064	0.7	✓	X
Number-of-Misconducts=Very high	(Crime-Desc.=POS. CANNABIS $\leq$ 30GMS$\rightarrow$Domestic=f) **AND** Safety-Index=Low	0.136	0.658	X	✓

Table 5: Some meta-association rules discovered in the Chicago dataset for $minSupp = minFSupp = 0.05$ and $minCF = minFCF = 0.5$. (The "✓" is used when the rule has been found and "X" when not)

7. Conclusions and Future Work

In this work, we have elaborated on the concept of meta-association rule to fuse information that has been previously extracted from different sets of data in the form of association rules. We have proposed two different processes to extract meta-association rules, based respectively on the creation of a crisp and a fuzzy meta-database. The fuzzy approach deals appropriately with the continuity inherent to the accuracy measures obtained in the first stage of the framework, and gives the possibility of adding supplementary fuzzy attributes representing additional information expressed in a user-friendly manner. Although the crisp version for mining meta-rules is generally faster, the fuzzy approach obtains a reduced set of meta-rules, and thus it can be faster examined by the end-user or the expert.

The procedure proposed here makes use of a suitable representation of information before being fused in what we called meta-database. Different meta-databases will drive to different "versions" of meta-association rules –or meta-rules with different semantics. In particular, crisp meta-association rules represent relations among the

appearance of rules in some datasets (in several districts for the presented scenario), whilst the fuzzy meta-association rules takes into account not only the presence of rules but also their strength, which is given by the accuracy measure.

This work opens several interesting issues to be addressed in the future. For instance, in many occasions the attributes description do not exactly fit in all the datasets although their meaning is similar. Then, one promising work consists on studying how to fuse this information during the MARM process. A possible solution, could utilise a knowledge repository that would assist in the process of mining meta-rules by matching similar items. Another open issue is how to conveniently adapt the MARM process for streaming data considering different time granularities. It also arouses our interest the suitability of the proposed algorithm for handling with Multiple Big Databases using the up-to-date services available for extremely big datasets.

Acknowledgement

The research reported in this paper was partially supported by the Andalusian Government (Junta de Andalucía) under projects P11-TIC-7460 and P10-TIC-6109; by the Spanish Ministry for Science and Innovation by the project grants TIN2012-30939; and by the ePOOLICE project FP7-SEC-2012-312651, funded from the European Union in the Seventh Framework Programme under grant agreement No 312651.

References

[1] R. Agrawal and J.C. Shafer. Parallel mining of association rules: Design, implementation and experience. In *IBM Research Report*, 1996.

[2] R. Agrawal and R. Srikant. Fast algorithms for mining association rules in large databases. In *Proceedings of the Twentieth International Conference on Very Large Databases*, pages 487–499, Santiago, Chile, 1994.

[3] M. Z. Ashrafi, D. Taniar, and K. Smith. ODAM: an optimized distributed association rule mining algorithm. *IEEE Distributed Systems Online*, 5(3), 2004.

[4] W.H. Au and K.C.C. Chan. Mining changes in association rules: a fuzzy approach. *Fuzzy Sets and Systems*, 149:87–104, 2004.

[5] F. Berzal, M. Delgado, D. Sánchez, and M.A. Vila. Measuring accuracy and interest of association rules: A new framework. *Intelligent Data Analysis*, 6(3):221–235, 2002.

[6] C-H. Chen, A-F. Li, and Y-C. Li. A fuzzy coherent rule mining algorithm. *Applied Soft Computing*, 13(2), 2013.

[7] C-H. Chen, A-F. Li, and Y-C. Li. Actionable high-coherent-utility fuzzy itemset mining. *Soft Computing*, 18:2413–2424, 2014.

[8] D.W. Cheung, V.T. Ng, A.W. Fu, and Y. Fu. Efficient mining of association rules in distributed databases. *IEEE Transactions on Knowledge and Data Engineering*, 8(6):911–922, 1996.

[9] P. Cotofrei and K. Stoffel. From temporal rules to temporal meta-rules. In Kambayashi et al., editor, *Data Warehousing and Knowledge Discovery*, volume 3181 of *Lecture Notes in Computer Science*, pages 169–178. Springer Berlin Heidelberg, 2004.

[10] M. Delgado, N. Marín, D. Sánchez, and M.A. Vila. Fuzzy association rules: General model and applications. *IEEE Transactions on Fuzzy Systems*, 11(2):214–225, 2003.

[11] M. Delgado, M.D. Ruiz, and D. Sánchez. A restriction level approach for the representation and evaluation of fuzzy association rules. In *proceedings of the IFSA-EUSFLAT*, pages 1583–1588, Lisbon, Portugal, 2009.

[12] M. Delgado, M.D. Ruiz, and D. Sánchez. Studying interest measures for association rules through a logical model. *Int.l J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 18(1):87–106, 2010.

[13] M. Delgado, M.D. Ruiz, and D. Sánchez. New approaches for discovering exception and anomalous rules. *Int. J. of Uncert., Fuzziness and Knowledge-Based Systems*, 19(2):361–399, 2011.

[14] M. Delgado, M.D. Ruiz, D. Sánchez, and J.M. Serrano. A formal model for mining fuzzy rules using the RL representation theory. *Information Sciences*, 181(23):5194–5213, 2011.

[15] D. Dubois and H. Prade. New results about properties and semantics of fuzzy set-tehoretic operators. *Fuzzy Sets: Theory and Applications to Policy Analysis and Information Systems*, pages 59–75, 1980.

[16] D. Dubois and H. Prade. An introduction to possibilitic and fuzzy logics. *Non-Standard Logics for Automated Reasoning*, pages 287–315, 1988.

[17] Y. Fu and J. Hah. Meta-rule-guided mining of association rules in relational databases. In *In Proc. 1st Int'l Workshop on Integration of Knowledge Discovery with Deductive and Object-Oriented Databases (KDOOD'95*, pages 39–46, 1995.

[18] L. Geng and H.J. Hamilton. Interestingness measures for data mining: A survey. *ACM Comput. Surv.*, 38(3), September 2006.

[19] A. Gyenesei. A fuzzy approach for mining quantitative association rules. *Acta Cybern.*, 15(2), 2001.

[20] Cornelia Győrödi. A comparative study of distributed algorithms in mining association rules. In *SINTES 11, International Symposium on System Theory –XI Edition, Craiova, Romania*, volume 1, pages 339–345, 2003.

[21] R.J. Hilderman. Assessing the interestingness of discovered knowledge using a principled objective approach. In *UBDM06*, 2006.

[22] W. Jian and L.Xing ming. An efficient association rule mining algorithm in distributed databases. In *First Int. Workshop on Knowledge Discovery and Data Mining, WKDD 2008*, pages 108–113, 2008.

[23] R. Jin, G. Yang, and G. Agrawal. Shared memory parallelization of data mining algorithms: Techniques, programming interface and performance. *IEEE Transactions on Knowledge and Data Engeneering*, 17(1):71–89, 2005.

[24] C.M. Kuok, A.W. Fu, and M.H. Wong. Mining fuzzy association rules in databases. *SIGMOD Record*, 27(1):41–46, 1998.

[25] J.H. Lee and H.L. Kwang. An extension of association rules using fuzzy sets. In *IFSA'97*, Prague, Czech Republic, 1997.

[26] N. Marín, M.D. Ruiz, and D. Sánchez. Fuzzy frameworks for mining data associations: fuzzy association rules and beyond. *WIREs Data Mining and Knowledge Discovery*, 6:3422–3428, 2016.

[27] A. Nanopoulos, N. Apostolos, and Y. Manolopoulos. Mining association rules in very large clustered domains. *Information Systems*, 32:649–669, 2007.

[28] R.T. Ng, L.V.S. Lakshmanan, J. Han, and A. Pang. Exploratory mining and pruning optimizations of constrained associations rules. *SIGMOD Rec.*, 27(2):13–24, 1998.

[29] M. Plasse, N. Nianga, G. Saporta, A. Villeminot, and L. Leblond. Combined use of association rules mining and clustering methods to find relevant links between binary rare attributes in a large data set. *Computational Statistics & Data Analysis*, 52(1):596–613, 2007.

[30] T.T. Quan, L.N. Ngo, and S.C. Hui. An effective clustering-based approach for conceptual association rules mining. In *International Conference on Computing and Communication Technologies, 2009. RIVF '09*, pages 1–7, 2009.

[31] S. Ramaswamy, S. Mahajan, and A. Silberschatz. On the discovery of interesting patterns in association rules. In *Proceedings of the 24th International Conference on Very Large Data Bases*, pages 368–379, San Francisco: Kaufmann, 1998.

[32] J.F. Roddick, M. Spiliopoulou, D. Lister, and A. Ceglar. Higher order mining. *SIGKDD Explorations*, 10(1):5–17, 2008.

[33] M.D Ruiz, J. Gómez-Romero, M.J. Martin-Bautista, and D. Sánchez. Meta-association rules for fusing regular association rules from different databases. In *Proceedings of the 17th Int. Conf. on Information Fusion (FUSION'2014)*, 2014.

[34] M.D. Ruiz, M.J. Martin-Bautista, D. Sánchez, M.A. Vila, and M. Delgado. Anomaly detection using fuzzy association rules. *Int. J. Electronic Security and Digital Forensics*, 26(1):25–37, 2014.

[35] D. Sánchez, M. Delgado, M.A. Vila, and J. Chamorro-Martínez. On a non-nested level-based representation of fuzziness. *Fuzzy Sets and Systems*, 192:159–175, 2012.

[36] N. Saxena, R. Arora, R. Sikarwar, and A. Gupta. An efficient approach of association rule mining on distributed database algorithm. *Int. Journal of Information and Computation Technology*, 3(4):225–234, 2013.