18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Fuzzy Meta-Association Rules for Information Fusion

M. Dolores Ruiz*, Juan Gómez-Romero*, Maria J. Martin-Bautista*, *Member, IEEE*,
Daniel Sánchez*†, *Senior Member, IEEE*, Miguel Delgado*, *Senior Member, IEEE*
*CITIC-UGR, Dept. Computer Science and A.I. University of Granada (Spain)
†European Centre for Soft Computing, Mieres (Spain)
Email: {mdruiz, jgomez, mbautis, daniel}@decsai.ugr.es, mdelgado@ugr.es

***Abstract*—Nowadays, data volume, distribution, and volatility makes it difficult to apply traditional Data Mining techniques in the search of global patterns in a domain under observation. This is the case of the methods for discovering associations, which typically require a single uniform dataset. To address the scenarios in which satisfying this requirement is not practical or even feasible, we propose a new method for fusing information extracted from individual and partially heterogeneous databases in the form of association rules. This method produces meta-association rules; i.e., rules in which the antecedent or the consequent may contain rules as well. In this paper, we describe the formulation and the implementation of two alternative frameworks that obtain, respectively, crisp meta-rules and fuzzy meta-rules. The comparison of both frameworks shows that the fuzzy approach offers several advantages: it is more accurate, 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.**

I. Introduction

In the last years, the increasing number of sensor devices, the paradigm shift from lower-level object recognition to higher-level situation assessment, and the need of integrating heterogeneous sources (including *soft information* in textual form) have drawn the attention of Information Fusion researchers to Data Mining and Knowledge Discovery. This research area aims to the development of intelligent methods for automatic extraction of hidden knowledge from available datasets. Accordingly, data mining processes can be used to obtain elaborated knowledge that can be afterwards fused with sensor-based data and other information. Furthermore, data mining can be applied on fused data to achieve better inferences towards situation and threat assessment. Likewise, the exponential growth of available data poses a great challenge to the classical approaches in Data Mining and Knowledge Discovery. Moreover, 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 any application domain in which it is necessary to interpret more information sources more efficiently.

Association rules extraction is a well-established Data Mining technique for discovering information from structured databases. Association rules have the form of implications $X \to Y$, which represent the joint co-occurrence of X and Y. They are easy to interpret for final users and offer a wide variety of variations and extensions, which makes them suitable in several scenarios. As a matter of example, they have been used to extract information from sensor networks [4], [29], to recognize human activities in Ambient Intelligence [8], [23], to analyze market and financial data [5], [14], and to detect crime and fraud [25], [26].

Typically, association rules algorithms work on the assumption that a fully-available uniform dataset is available 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 addition, it is more and more frequent having distributed datasets with similar semantics but different structure, 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 mining over patterns/models derived from one or more large and/or complex datasets" [22].

In a recent work [24], we proposed a new HOM technique for fusing association rules obtained from several different databases. This was achieved by means of what we have called *meta-association rules*. These meta-association rules are built from regular association rules that have been previously extracted from individual databases, in such a way that they can contain rules in the antecedent and in the consequent. The semantics of the meta-association rules are different from the regular rules, since they denote associations between associations. However, the basic process to obtain meta-rules has a limitation, because it only considers if an association has been previously mined from the original dataset or not; i.e. a regular rule mined with a confidence of 0.99 has the same importance than another regular rule with a confidence of 0.3.

To address this problem, in this paper we present a new approach for mining meta-rules that allows defining a degree of fulfillment of the regular rules –and by extension, to any attribute considered in the meta-association rule extraction process. Our proposal is based on the Fuzzy Sets theory, which offers a sound framework for the management of imprecise information. Generally speaking, we use fuzzy association rules to solve the problem of a strict interval-based discretization of continuous values. Fuzzy sets relax this restriction by considering fuzzy sets instead of intervals, 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. Therefore, to build fuzzy meta-association rules, regular rules are assigned a degree corresponding to their satisfiability, and this value is used to calculate the intensity of the derived meta-rules. Analogously, any other attribute, either contextual or present in the original databases, can be expressed in a more human-friendly way by using fuzzy sets. In the paper, we describe how to obtain fuzzy meta-association rules by using a two-stage version of a fuzzy rule mining algorithm. We show that this kind of rules convey new relevant information that cannot be obtained by regular rules, and illustrate their application in a use case in crime data analysis.

The paper is structured as follows. In Section II we review the crisp and fuzzy data mining techniques that are employed in the remainder of the paper. Section III describes our approach for fusing information by means of crisp and fuzzy meta-rules, and Section IV shows some interesting experimental results. Section V compare our proposal to other related works. Finally, Section VI points out some conclusions and prospective directions for future research.

II. Background

A. Association Rules

Formally, let D be a database constituted by a set of transactions where columns represent the different *attributes*, and rows the transactions where the attributes take values. These attributes can be categorical or quantitative. For categorical attributes, we usually have linguistic values associated to the categories, obtaining *items* of the form $\langle attribute, value\rangle$. We can use the same formalization for quantitative attributes, obtaining items of the form $\langle attribute, numerical\ value\rangle$. Continue attributes are usually split into meaningful intervals, thus obtaining items of the form $\langle attribute, interval\rangle$, which trivially indicates that the value of the attribute lies in that interval. From this formulation, a transactional database can be transformed into a boolean database by considering items as pairs $\langle attribute, value/interval\rangle$, and assigning $0, 1$ if the item is satisfied or not.

Let I be the set of all items. An association rule [1] is an "implication" of the form $A \to B$ that relates the presence of itemsets A and B in transactions of D, assuming that $A, B \subseteq I$, $A \cap B = \emptyset$ and $A, B \neq \emptyset$.

The support of a set of items (or itemset) is defined as the probability that a transaction t contains the itemset, i.e. $\mathrm{supp}(A) = |\{t \in D : A \subseteq t\}| \,/\, |D|$. The intensity of an association rule is measured by the ordinary measures of *support* (the joint probability $P(A \cup B)$) and *confidence* (the conditional probability $P(B|A)$):

$$\mathrm{Supp}(A \to B) = \mathrm{supp}(A \cup B),$$

$$\mathrm{Conf}(A \to B) = \frac{\mathrm{supp}(A \cup B)}{\mathrm{supp}(A)}.$$

Given the minimum threshold values $minsupp$ and $minconf$, which should be defined by the user, we say that $A \to B$ is *frequent* if $\mathrm{Supp}(A \to B) \geq minsupp$, and *confident* if $\mathrm{Conf}(A \to B) \geq minconf$.

Definition 1. [3] An association rule $A \to B$ is *strong* if it exceeds the minimum thresholds $minsupp$ and $minconf$ imposed by the user; i.e., if $A \to B$ is frequent and confident. $A \to B$ is *very strong* if both rules $A \to B$ and $\neg B \to \neg A$ are strong.

In this paper, we use the alternative framework for measuring the intensity of association rules proposed in [3], [7], where the accuracy of a rule is measured by means of Shortliffe and Buchanan's certainty factors [28]:

Definition 2. [7] The *certainty factor* of an association rule, denoted as $CF(A \to B)$, is defined as:

$$\begin{cases} \dfrac{\mathrm{Conf}(A \to B) - \mathrm{supp}(B)}{1 - \mathrm{supp}(B)} & \text{if } \mathrm{Conf}(A \to B) > \mathrm{supp}(B) \\ \dfrac{\mathrm{Conf}(A \to B) - \mathrm{supp}(B)}{\mathrm{supp}(B)} & \text{if } \mathrm{Conf}(A \to B) < \mathrm{supp}(B) \\ 0 & \text{otherwise.} \end{cases}$$

CF yields a value in the interval [-1, 1] that measures how our belief that B is in a transaction changes when we know 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. CF has better properties than confidence and other quality measures (see [9] for more details), and helps to reduce the number of rules obtained by filtering rules corresponding to statistical independence or negative dependence [3], [7]. Analogously to the confidence measure, we say that $A \to B$ is *certain* if $\mathrm{CF}(A \to B) \geq minCF$, where $minCF$ is the minimum threshold for the certainty factor given by the user. In [3] it is proven that the CF fulfils $\mathrm{CF}(A \to B) = \mathrm{CF}(\neg B \to A)$. This means that when using the certainty factor, the rule is also very strong. Then, we reformulate the very strong rules definition to denote rules that are frequent and certain.

B. Fuzzy Association Rules

The concepts of transaction and association rule can be generalized to the fuzzy case, as done in [7]. This work defines a fuzzy transaction as a non empty fuzzy subset $\tilde{\tau} \subseteq I$. Thus, for every item $i \in I$ and every fuzzy transaction $\tilde{\tau}$, an item i belongs to $\tilde{\tau}$ with degree[1] $\tilde{\tau}(i) \in [0, 1]$. By extension, let A be an itemset of I, i.e. a subset of items in a fuzzy transaction $\tilde{\tau}$. The membership degree of $A \subseteq I$ to the fuzzy transaction $\tilde{\tau}$ is

[1]Note that $\tilde{\tau}(i)$ is $\mu_{\tilde{\tau}}(i)$ where $\mu_{\tilde{\tau}} : I \longrightarrow [0, 1]$ is the membership function associated to the fuzzy set $\tilde{\tau}$ defined on I.

defined as $\tilde{\tau}(A) = \min_{i \in A} \tilde{\tau}(i)$. In particular, a crisp transaction is a special case of fuzzy transaction where every item in the transaction has membership degree equal to 1 or 0 depending on if it is in the transaction or not.

Definition 3. [7] Let I be a set of items, $\tilde{D}$ a set of fuzzy transactions and $A, B \in I$ two disjoint itemsets. A fuzzy association rule is satisfied in $\tilde{D}$ if and only if $\tilde{\tau}(A) \leq \tilde{\tau}(B)$ for all $\tilde{\tau} \in \tilde{D}$; that is, the membership degree of B is higher than the membership degree of A for all fuzzy transactions $\tilde{\tau}$ in $\tilde{D}$.

This definition maintains the meaning of crisp association rules: $\tilde{\tau}(A) \leq \tilde{\tau}(B)$ expresses that if $A \subseteq \tilde{\tau}$ holds, $B \subseteq \tilde{\tau}$ also holds –to an equal or larger extent. Likewise, since crisp transactions are a special case of fuzzy transactions, a crisp association rule is a special case of fuzzy association rule. It is worth to note that the appearance of a fuzzy rule is the same as a crisp rule, but fuzzy rules have been extracted from a fuzzy transactional database using adapted assessment measures as follows.

We employ the assessment measures proposed in [7]. For that, we compute the intensity of a fuzzy association rule by means of a quantified sentence using a fuzzy quantifier $Q_M(x) = x$, representing the notion of "most":

- The support of an itemset X is the evaluation of the quantified sentence "Q_M of the $\tilde{D}$ are $\tilde{\Gamma}_X$", where $\tilde{\Gamma}_X$ is a fuzzy set defined as $\tilde{\Gamma}_X(\tilde{\tau}) = \tilde{\tau}(X)$.
- The support of a fuzzy rule $A \rightarrow B$ in $\tilde{D}$, noted by $\mathrm{FSupp}(A \rightarrow B)$, is the result of the evaluation of the quantified sentence "Q_M of the $\tilde{D}$ are $(\tilde{\Gamma}_A \cap \tilde{\Gamma}_B)$".
- The confidence of a fuzzy rule $A \rightarrow B$ in $\tilde{D}$, $\mathrm{FConf}(A \rightarrow B)$, is the result of the evaluation of the quantified sentence "Q_M of the $\tilde{\Gamma}_A$ are $\tilde{\Gamma}_B$" .
- The certainty factor of the fuzzy rule $A \rightarrow B$ in $\tilde{D}$, $\mathrm{FCF}(A \rightarrow B)$, is computed as in Definition 2 by using a fuzzy version of the support and confidence measures.

To evaluate the quantified sentence "Q_M of the A are B", we can use the GD method defined in [12]:

$$GD_{Q_M}(B/A) = \sum_{\alpha_i \in \Lambda(B/A)} (\alpha_i - \alpha_{i+1})\, Q_M\left(\frac{|(B \cap A)_{\alpha_i}|}{|A_{\alpha_i}|}\right) \tag{1}$$

where $\Lambda(B/A) = \Lambda(B \cap A) \cup \Lambda(A)$, $\Lambda(A)$ is the set of α-cuts[2] of A, and $\Lambda(B/A) = \{\alpha_1, \ldots, \alpha_p\}$ with $\alpha_i > \alpha_{i+1}\ \forall i \in \{1, \ldots, p-1\}$ and $\alpha_p = 0$. If the fuzzy set A is not normalized, it must be normalized and the same factor of normalization must be applied to $B \cap A$.

Example 1. Let $I = \{i_1, i_2, i_3, i_4, i_5\}$ be a set of items and $\tilde{D}_1$ the fuzzy dataset given by Table I. In particular, we can see that $\tilde{\tau}_6$ is a crisp transaction. Some degrees of membership that we can find in $\tilde{D}_1$ are the following: $\tilde{\tau}_1(\{i_3, i_4\}) = 0.9$, $\tilde{\tau}_1(\{i_2, i_3, i_4\}) = 0.2$ and $\tilde{\tau}_2(\{i_1, i_2\}) = 1$. If we consider the set of α-cuts $\Lambda = \{1, 0.8, 0.6, 0.4, 0.2\}$, some of the fuzzy rules that can be found in $\tilde{D}_1$ are shown in Table II. Let us notice that the fuzzy set $\{i_4\}$ has been normalized for computing the assessment values FConf and FCF, and the same normalization factor has been applied to $\{i_4, i_5\}$.

[2]The α-cut of a fuzzy set A, is defined as $A_\alpha = \{x \in X : \mu_A(x) \geq \alpha\}$. The set of α-cuts of A, $\Lambda(A) = \{\alpha \in [0,1] : \mu_A(x) = \alpha \text{ for some } x \in X\}$.

TABLE I
SET OF FUZZY TRANSACTIONS $\tilde{D}_1$

	i_1	i_2	i_3	i_4	i_5
$\tilde{\tau}_1$	1	0.2	1	0.9	0.9
$\tilde{\tau}_2$	1	1	0.8	0	0
$\tilde{\tau}_3$	0.5	0.1	0.7	0.6	0
$\tilde{\tau}_4$	0.6	0	0	0.5	0.5
$\tilde{\tau}_5$	0.4	0.1	0.6	0	0
$\tilde{\tau}_6$	0	1	0	0	0

TABLE II
SOME FUZZY RULES OBTAINED IN $\tilde{D}_1$

Association rule	FSupp	FConf	FCF
$\{i_1, i_2\} \rightarrow \{i_3\}$	0.167	0.8	0.6
$\{i_4\} \rightarrow \{i_5\}$	0.233	0.767	0.68

III. FUSING INFORMATION WITH META-ASSOCIATION RULES

A. *Problem Statement*

Meta-association rules aim at fusing information obtained from distributed databases. Specifically, we want to combine association rules that have been extracted from each 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 we have databases of 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 they do not need to be equal.

Let us explain the problem with an example. Let us suppose that a banking company has several offices spread across the country that manage similar data. The company is interested in analyzing customer behavior from this data. There are two options to do so: the first one is to compile a very large dataset by merging all offices data, whilst the second one is to study summarized information obtained by each office. Traditionally, the first approach has been prevalent; in this paper, we propose the use of meta-association rules to implement the second one. Fusing extracted knowledge by means of meta-rules has some advantages. First, the complete dataset is not necessary, which increases the efficiency and allows working with very large databases. Additionally, it is not necessary that all the data sources 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 data about the places where the offices are sited. On the other hand, since the meta-rule analysis works on already summarized data, there is a unavoidable loss of information, as explained in Section IV-C.

B. Architecture

The general process flow of our method is depicted in Figure 1. It encompasses two sequential steps. First, we obtain regular association rules from each database. Second, we fuse this information and obtain meta-association rules. Depending on the use of crisp or fuzzy values as the input of this second step, we talk about meta-association and fuzzy meta-association rules, respectively.

Step 1. Let $D_1, D_2, \ldots, D_k$ be k databases that may share some of their attributes. After applying a rule extraction procedure[3], we obtain k different sets of association rules $R_1, R_2, \ldots, R_k$ (each R_i corresponds to a different D_i). The number of rules in $R_1, R_2, \ldots, R_k$, noted $nr_1, nr_2, \ldots, nr_k$, 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 in $R_1, R_2, \ldots, R_k$. 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 create a structured database, namely the meta-database $\mathfrak{D}$, which represents the different rules $r_1, \ldots, r_n$ found in the sets R_i. The meta-database can be enriched with data describing additional features of D_i, which are aggregated by means of new attributes $at_1, \ldots, at_m$. The meta-database is used as the input of the meta-rule extraction process. The two strategies to extract the meta-rules are described next.

C. Crisp Meta-Association Rules

In [24], we proposed a first strategy for discovering meta-rules based on crisp association rules mining. In this approach, the created meta-database $\mathfrak{D}$ is a boolean database representing only if a given regular rule has been generated or not the original databases. Additionally, $\mathfrak{D}$ can also contain some extra information in the form of crisp attributes. We depict this case in Table III, in which we assign to $\mathfrak{D}_{j,i}$ value 1 if the rule r_i is found in D_j, and 0 otherwise. Afterwards, *crisp meta-association rules* are extracted from $\mathfrak{D}$ by a classical association rules mining algorithm, in our case based on the *support-CF* framework. 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:

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

TABLE III
EXAMPLE OF BOOLEAN META-DATABASE REPRESENTING PREVIOUSLY EXTRACTED ASSOCIATION RULES AND ADDITIONAL ATTRIBUTES

$\mathfrak{D}$	r_1	r_2	$\cdots$	r_n	at_1	$\cdots$	at_m
D_1	1	1	$\cdots$	0	1	$\cdots$	1
D_2	0	1	$\cdots$	0	0	$\cdots$	1
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
D_k	1	0	$\cdots$	1	1	$\cdots$	0

- $r_i \to r_j$, where r_i, r_j can be rules or a conjunction of rules; for example: $r_i = r_{i1} \wedge \ldots \wedge r_{is}$.
- $at_i \to at_j$, where at_i, at_j can be attributes or a conjunction of attributes.
- $r_i \to at_j \wedge r_k$ or $at_j \wedge r_k \to r_i$, where r_i, r_k represent conjunction of rules and at_j a conjunction of attributes, and they can be mixed; e.g. we can find a meta-rule of the form $r_1 \wedge at_2 \to r_3 \wedge at_4$.

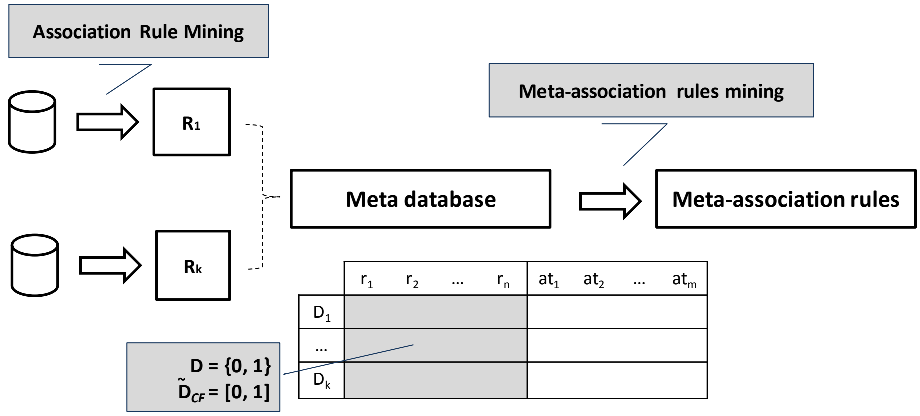
As it was mentioned in the introduction, this procedure has limitations, because it only takes into account if a rule has been previously mined from a dataset or not. This implies that, in the boolean meta-database, regular rules found with different intensity (e.g., $\mathrm{CF}(r_i) \gg \mathrm{CF}(r_j)$) have the same importance. 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 similar. This issue motivates our proposal for a different representation of continuous values.

D. Fuzzy Meta-Association Rules

Fuzzy transactional databases support the extraction of fuzzy association rules from a continuous representation of values. Following the same notation used in the previous section, a fuzzy meta-database $\tilde{\mathfrak{D}}_{CF}$ is created based on the certainty factor of the extracted regular rules. An example of such fuzzy meta-database is depicted in Table IV. In contrast to the crisp case, the values of the fuzzy meta-database are in the unit interval –more specifically, in the interval $[minCF, 1]$. The value in column r_i and row D_j is the certainty factor of the rule r_i in database D_j.

If we want to add additional attributes describing some features of the original databases, we can use crisp or fuzzy sets in the meta-database. Particularly, these attributes can be provided in a comprehensive way for the user by means of fuzzy sets. 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.

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



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 method described in Section II-B considering the FSupp and FCF measures. Similarly to the crisp case, the fuzzy meta-rules can be classified in three different types:

- fuzzy meta-rules that relate only rules,
- fuzzy meta-rules that relate only attributes, and
- fuzzy meta-rules that relate rules and attributes.

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.

TABLE IV
FUZZY META-DATABASE THAT COMPILES THE OBTAINED ASSOCIATION RULES AND THE ADDITIONAL ATTRIBUTES

$\tilde{\mathfrak{D}}$	r_1	r_2	$\cdots$	r_n	at_1	$\cdots$	at_m
D_1	0.2	1	$\cdots$	0	0.9	$\cdots$	1
D_2	0	1	$\cdots$	0.6	0	$\cdots$	0.2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
D_k	0.9	0	$\cdots$	0.5	1	$\cdots$	0.1

IV. Algorithm and Experimental Evaluation

In this section, we describe the algorithm that implements the complete extraction process from the original datasets to the output meta-association rules. Afterwards, we present some experiments carried out with a database on criminal offenses. Our study of the results has two main goals:

- to compare the meta-rules obtained with the crisp and the fuzzy approaches, and
- to compare the regular association rules obtained after merging all datasets with the meta-rules obtained with partitioned data.

A. Algorithm and Implementation

Algorithm 1 describes the complete process depicted in Figure 1. Initially, crisp association rules are extracted from 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, additional features may be added as attributes to the meta-database. As previously explained, in the fuzzy case the attributes of the meta-database can be modeled as fuzzy sets. Finally, crisp or fuzzy meta-association rules are respectively extracted. Obviously, when the initial datasets are not available, the meta-rule extraction procedure would start at step 12.

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)
1: **for all** D_i such that $1 \le i \le k$ **do**
2:   # D_i *preprocessing*
3:   Read D_i and store the items I
4:   Transform D_i into a boolean database
5:   *# Mine very strong rules*
6:   Compute the candidate set C of frequent itemsets $Supp(X) \ge minsupp$
7:   Compose the rule with $X, Y \in C$
8:   **if** $Supp(X \Rightarrow Y) \ge minsupp$ **and** $CF(X \Rightarrow Y) \ge minCF$ **then**
9:     The rule is a very strong rule
10:   **end if**
11: **end for**
12: # $\mathfrak{D}$ *creation*
13: Compile all different rules from $R_1, \ldots, R_k$
14: Create $\mathfrak{D}$ using compiled rules and additional attributes
15: *# Mining meta-association rules*
16: Repeat steps 2-10 to mine meta-association rules from $\mathfrak{D}$

When employing the fuzzy meta-database, the algorithm in step 16 extracts fuzzy association rules. In our implementation, we use the algorithm described in [11], which computes the fuzzy assessment measures by means of a parallel process based on the α-cuts. The proposed algorithm uses an itemset representation based on bit strings [10], which allow us to speed up logical operations with boolean data.

The computational complexity of the algorithm depends on the number of transactions and items. The first step (lines 1-11) is $\mathcal{O}(n2^{|I|})$ for each original D_i, being n the number of transactions of D_i and $|I|$ the number of different items. The second step (lines 12-16) is $\mathcal{O}(k2^{m+r})$, being k the number of databases, m the number of additional attributes, and r the number of rules obtained in the first step. The complexity of the analogous fuzzy version using the Algorithm provided in [11] is also $\mathcal{O}(k2^{m+r})$, since it processes the α-cuts in parallel. Regarding memory requirements, the footprint is not high thanks to the use of the binary representation.

B. Dataset description

We have used two datasets for the experimental evaluation of our proposal. The first dataset (`chicago`) includes data of crime incidents reported by the Chicago police[4] 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. We have split the dataset into 22 databases corresponding to the districts of the city. For the meta-database, we have considered additional attributes of schools aggregated by district (number of students, number of misconducts and perceived safety index).

A second small-sized synthetic dataset (`synth`) has been built to compare the rules obtained when all data is merged to the meta-rules mined by using distributed data. This dataset contains artificial data divided into eight 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 $\langle attribute, value \rangle$.

C. Experimental Evaluation

We have conducted two different types of experiments. The first one uses the `chicago` database to: (1) compare the number and the type of the crisp and fuzzy meta-rules obtained; and (2) compare the execution time of both approaches. In the second one, we have employed the merged `synth` dataset to analyze the differences of mining meta-rules in partitioned datasets versus mining regular rules in a single dataset that includes all the transactions.

The experiments have been executed 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.

For the `chicago` database, we have set the $minSupp$ and $minFSupp$ to 0.05, and we have compared the number of meta-rules discovered and the execution time for values $minCF$ and $minFCF$ in $\{0.2, 0.3, \ldots, 0.8\}$. The results are shown in Figure 2. We can see that the number of meta-rules in the crisp approach is generally larger than the number of rules in the fuzzy approach (Figure 2, left). We can also observe that the number of crisp meta-rules is drastically reduced when the $minCF$ threshold is slightly increased. In contrast, in the fuzzy case the reduction of the number of meta-rules is not so high. This is due to the use of the rule's CF in the meta-database $\tilde{\mathfrak{D}}_{CF}$. 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.

Regarding execution time (Figure 2, right), in the crisp case it is low and uniform, while in the fuzzy case it is higher and strongly dependent on the $minFCF$ value. This happens because the algorithm for mining fuzzy rules computes 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.

We want to remark that the meta-rules obtained by each approach are not always the same. There are some examples of crisp meta-rules that are not obtained in the fuzzy case and vice versa. To exemplify this, we have selected two meta-rules that have been extracted in only one of the approaches (see Table V, where $Desc.$ stands for Description and f for false). In this table, we can see two meta-rules that express relations between rules and between rules and attributes. For instance, the first meta-association rule states that there is a co-occurrence relation with a medium-high certainty ($FCF =$ 0.658) between a very high `number of misconducts`, the rule (`possession of cannabis`$\leq$ `30 grams` $\rightarrow$ `Domestic=false`), and the low `safety-index` in a district. That means that in districts with a very high number of misconducts, it is frequent and reliable to have: (a) a low perception of security; and (b) a relation between the low possession of cannabis and its occurrence in a non-domestic environments.

The second experimental series were conducted using the artificial dataset `synth`. We have carried out two experiments: (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 meta-database $\tilde{D}_{CF}$. In both cases we employed the same thresholds: $minSupp = minFSupp = minCF = minFCF = 0.2$. After comparing the obtained rules, we have observed the following facts:

- Some of the rules discovered in the first case are part of the antecedent and the consequent of a meta-rule discovered in the second case.
- Other rules found in (A) do not appear in any meta-rule. The reason is that in the compiled database, the overall support and CF values can be high enough to exceed the threshold. In contrast, when the datasets are partitioned, it may happen that a candidate rule do not exceed the minimum support and CF thresholds in a dataset, and therefore it will not count to build a meta-rule.
- Although one can expect that some groups of items should appear in regular rules found in (A) and meta-

[4]https://data.cityofchicago.org/

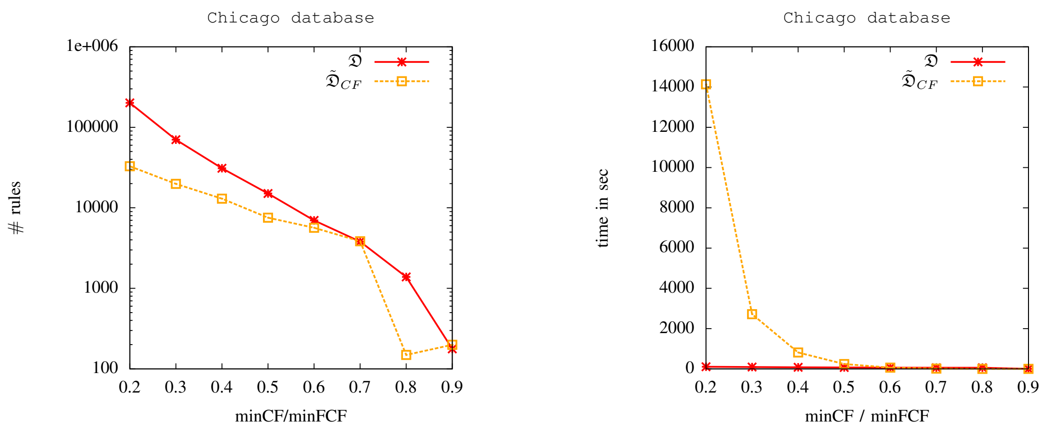


Fig. 2. **Left:** Number of crisp/fuzzy meta-association rules (y-axis in logarithmic scale) vs $minCF/minFCF$ (x-axis) when $minSupp = 0.05$
Right: Time in sec. for mining crisp/fuzzy meta-association rules (y-axis) vs $minCF/minFCF$ (x-axis) when $minSupp = 0.05$.

Antecedent	**Consequent**	$Supp/FSupp$	CF/FCF	$\mathfrak{D}$	$\tilde{\mathfrak{D}}_{CF}$
`Number-of-Misconducts=Very high`	`(Crime-Desc.=POSS: CANNABIS` $\leq$ `30GMS`$\rightarrow$`Domestic=f)` **AND** `Safety-Index=Low`	0.136	0.658	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

TABLE V
SOME META-ASSOCIATION RULES FOUND IN THE CHICAGO DATASET FOR $minSupp = minFSupp = 0.05$ AND $minCF = minFCF = 0.5$.
THE SYMBOL "✓" IS USED WHEN THE RULE HAS BEEN FOUND AND "X" WHEN NOT.

rules found in (B), 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 in the compiled database and vice versa. The rationale behind this is the same as in the previous case.

In consequence, we can conclude that, in general, the meta-rules –and therefore the information that they represent– obtained when the datasets are partitioned are not the same as the rules obtained in a single database containing all the transactions. The reason is that the meta-rules are obtained from already summarized information, and therefore some information loss should be expected. In addition, 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 districts. In any case, let us highlight that using the complete dataset is not possible in several scenarios, and therefore there is no alternative to the use of meta-association rules, or a similar HOM technique.

V. Related Works

The proposal presented in this paper can be of interest in the application areas of association rules mentioned in Section I. To the best of our knowledge, this is the first attempt to develop an algorithm to fuse rules mined from different databases. Most related works are focused on mining association rules in distributed databases; i.e., they extract a single set of rules from the distributed data [6], [15], [16], [27].

In [2], a different idea of meta-rules is developed. 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*. In that regard, they use appropriate linguistic labels defined by means of fuzzy sets. Our concept of meta-rule is different, since we may have association rules in the antecedent and/or the consequent of the meta-rule, and we do not consider sequences of supports/confidences.

There are other approaches that also apply a two-step data mining process to fuse information. In [19]–[21], 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 increase.

There are also some research works that try to compare sets of rules. For instance, in [13] the author proposes general measures for comparing two sets of rules, namely rule overlapping, average support difference, and average confidence difference. Similarly, there are some interesting works aimed at comparing expert knowledge and automatically-extracted knowledge in the form of rules. To name some of them, we can highlight the proposals in [17], [18], where users' knowledge and their impressions are captured in the form of rules (or similar structures), and then, these rules are compared to the rules extracted from data.

VI. Conclusions and Future Work

In this paper, we have elaborated on the concept of meta-association rules to fuse information previously extracted from different and probably heterogeneous databases. We have proposed two different processes to extract meta-association rules, based respectively in the creation of a crisp and a fuzzy meta-database. The fuzzy approach solves some problems of the crisp approach, because it does not discard the actual value of the assessment measures obtained in the initial data mining stage. We have compared these two approaches in practice, and we have concluded that, in general, the fuzzy process leads to a more manageable set of rules and allows the incorporation of additional information to the process in a more natural way. We have also compared the rules obtained by combining the information in a unique database versus the meta-rules obtained by fusing the regular rules using a meta-database. As a result, we have seen that the information obtained is different in both cases with regard to the number and the meaning of obtained rules. An interesting issue to be addressed in the future is to study in more detail how to fuse regular rules including attributes with similar semantics. We plan to use a knowledge repository that would assist the process of mining meta-rules by matching similar items.

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, H. Manilla, R. Sukent, A. Toivonen, and A. Verkamo. *Advances in Knowledge Discovery and Data Mining*, chapter Fast discovery of Association rules, pages 307–328. AAA Press, 1996.

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

[3] 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.

[4] A. Boukerche and S. Samarah. An efficient data extraction mechanism for mining association rules from wireless sensor networks. In *Int. Conference on Communications, ICC'07*, pages 3936–3941, 2007.

[5] M. Chen, A. Chiu, and H. Chang. Mining changes in customer behavior in retail marketing. *Expert Systems with Applications*, 2005.

[6] 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.

[7] 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.

[8] M. Delgado, M. Ros, and M.A. Vila. Correct behavior identification system in a tagged world. *Expert Systems with Applications*, 36(6):9899–9906, 2009.

[9] 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.

[10] 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.

[11] 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.

[12] M. Delgado, D. Sánchez, and M.A. Vila. Fuzzy cardinality based evaluation of quantified sentences. *Int. Journal of Approximate Reasoning*, 23:23–66, 2000.

[13] Damian Dudek. Measures for comparing association rule sets. In *Int. Conference on Artificial Intelligence and Soft Computing, ICAISC 2010, Part I, LNAI 6113*, pages 315–322, 2010.

[14] M. M. A. Ellatif. Association rules technique to diagnosis financial performance for ksa stock market companies. *Working Paper Series, Faculty of Computers and Information. Mansoura University of Egypt*, 2005.

[15] 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.

[16] 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.

[17] B. Liu, W. Hsu, and S. Chen. Using general impressions to analyze discovered classification rules. In *KDD*, pages 31–36, 1997.

[18] B. Liu, W. Hsu, L. Mun, and H. Lee. Finding interesting patterns using user expectations. *IEEE Transactions on Knowledge and Data Engeneering*, 11(6):817–832, 1999.

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

[20] 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.

[21] 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.

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

[23] M. Ros, M.P. Cuéllar, M. Delgado, and M.A. Vila. Online recognition of human activities and adaptation to habit changes by means of learning automata and fuzzy temporal windows. *Information Sciences*, 220:86–101, 2013.

[24] 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.l Conf. on Information Fusion (FUSION'2014)*, 2014.

[25] 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.

[26] D. Sánchez, M.A. Vila, L. Cerda, and J.M. Serrano. Association rules applied to credit card fraud detection. *Expert Systems with Applications*, 36:3630–3640, 2009.

[27] 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.

[28] E. Shortliffe and B. Buchanan. A model of inexact reasoning in medicine. *Mathematical Biosciences*, 23:351–379, 1975.

[29] S.K. Tanbeer, C.F. Ahmed, B. Jeong, and Y. Lee. Efficient mining of association rules from wireless sensor networks. In *11th Int. Conf. on Advanced Communication Technology*, pages 719–724, 2009.