

# Fuzzy Meta-Association Rules

M.D. Ruiz, J. Gómez-Romero, M.J. Martín-Bautista, D. Sánchez, M.A. Vila, M. Delgado

CITIC-UGR, Dept. Computer Science and A.I., University of Granada (Spain)  
{mdruiz, jgomez, mbautis, daniel, vila}@decsai.ugr.es, mdelgado@ugr.es

## Abstract

Association rules is a useful tool to extract new information from raw data expressed in a comprehensive way for decision makers. However, in some applications raw data might not be available for several reasons. First, stream data are only temporarily available for their processing or if it is stored, only summaries or representations of the extracted knowledge are kept. Second, under some circumstances primary data cannot be disclosed due to privacy or legal restrictions. In the light of these observations we propose fuzzy meta-association rules for mining association rules over already discovered rules in a set of databases sharing common information. We compare this proposal with a previous one using crisp meta-rules showing that fuzzy meta-association rules discover interesting knowledge obtaining a more manageable set of rules for human inspection and allowing the use of fuzzy items to express additional knowledge about the original databases.

**Keywords:** Fuzzy association rules, Meta-association rules, Higher Order Mining.

## 1. Introduction

Data mining techniques usually assumed that *primary data* have been captured by some application, cleaned and preprocessed according to the necessities of the mining algorithm. However, in some applications, primary data is only available for a short time. That is the case of stream data which are processed in real time and usually deleted after storing the results. Besides, primary data might not be available in several scenarios, or the owners are not authorized to share it. This is the case for instance of law enforcing agencies facing security threats and other institutions that cannot disclose personal data. However, these organizations may be able to publish summarized or extracted knowledge. In order to bring into play the advantages of knowledge discovery to these scenarios, we need to change our perspective to address pattern analysis instead of data analysis. In this line a new field of knowledge discovery emerges with the name of Higher Order Mining (HOM) concerned with applying mining techniques over patterns/models derived from one or more large and/or complex datasets [1]. Accordingly, several combinations of mining techniques (association discovery, clustering, classifica-

tion, trend analysis, etc.) can deal with this kind of scenarios.

Association rules are a well-established technique for mining information from structured databases. They provide support for the identification of novel, potentially-useful, and comprehensive knowledge in the form of implications  $X \rightarrow Y$ , which represent the joint co-occurrence of  $X$  and  $Y$  in the database. Our proposal combines association rule mining techniques in both the primary and the processed data, thus producing what we have named *meta-association rules*. Meta-association rules are rules about rules; i.e. they are built from already obtained association rules that have been extracted from several databases about the same topic or with a similar structure. Therefore, meta-association rules can contain rules in the antecedent and/or in the consequent. Generally speaking, meta-association rule mining is a second order mining technique that can be employed when the available information is in the form of association rules.

In a recent work [2], we presented an early approach for meta-association rule mining based on crisp rule mining. Crisp meta-rules are discovered by taking into account only if an association has been previously mined from the original dataset or not, and not its degree of fulfillment. This means that, to build meta-rules, a rule mined with a confidence of 0.9 has the same importance than another rule with a confidence of 0.5. Consequently, it is convenient to consider the available measures, like the confidence, to quantify the importance of the rule in the HOM process. As it is often done in association rules, these numerical values can be discretized into intervals with crisp boundaries. However, this usually leads to an over estimation or under estimation of the boundary values. Fuzzy set theory overcomes this problem by considering the degrees of satisfiability themselves. For this reason, we propose in this work the use of fuzzy meta-association rules, which allow us to work directly with degrees expressed in the unit interval.

The paper is structured as follows. In Section 2 we review the concepts of crisp and fuzzy association rule and the derived assessment measures. Section 3 describes our proposal for crisp and fuzzy meta-association rule-mining. The algorithm for mining meta-rules is then described in Section 4. Section 5 shows the experiments made joint with a discussion of the results. Section 6 discusses some related works and finally section 7 points out

some conclusions and prospective lines for future research.

## 2. Background

### 2.1. Association Rules

Given a set  $I$  (“set of items”) and a database  $D$  constituted by a set of transactions, each one being a subset of  $I$ , association rules [3] are “implications” of the form  $A \rightarrow B$  that relate the presence of itemsets  $A$  and  $B$  in transactions of  $D$ , assuming  $A, B \subseteq I$ ,  $A \cap B = \emptyset$  and  $A, B \neq \emptyset$ .

The support of an itemset is defined as the probability that a transaction contains the itemset, i.e.  $\text{supp}(A) = |\{t \in D \mid A \subseteq t\}| / |D|$ . The ordinary measures to assess association rules are the *support* (the joint probability  $P(A \cup B)$ ) and the *confidence* (the conditional probability  $P(B|A)$ ):

$$\begin{aligned} \text{Supp}(A \rightarrow B) &= \text{supp}(A \cup B), \\ \text{Conf}(A \rightarrow B) &= \frac{\text{supp}(A \cup B)}{\text{supp}(A)}. \end{aligned}$$

Given the minimum thresholds *minsupp* and *minconf*, that should be imposed by the user, we will say that  $A \rightarrow B$  is *frequent* if  $\text{Supp}(A \rightarrow B) \geq \text{minsupp}$ , and *confident* if  $\text{Conf}(A \rightarrow B) \geq \text{minconf}$ .

**Definition 1.** [4] An association rule  $A \rightarrow B$  is *strong* if it exceeds the minimum thresholds *minsupp* and *minconf* imposed by the user, i.e. if  $A \rightarrow B$  is frequent and confident.

An alternative framework was proposed in [5, 4] where the accuracy is measured by means of Shortliffe and Buchanan’s certainty factors [6], as follows:

**Definition 2.** [5] Let  $\text{supp}(B)$  be the support of the itemset  $B$ , and let  $\text{Conf}(A \rightarrow B)$  be the confidence of the rule. The *certainty factor* of the rule, denoted as  $CF(A \rightarrow B)$ , is defined as

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

The 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. CF has better properties than confidence and other quality measures (see [7] for more details), and helps to solve some of the confidence drawbacks [4, 5]. In particular, it helps to reduce the number of rules obtained by filtering those rules corresponding to statistical independence or negative dependence. Analogously, we will say that  $A \rightarrow B$  is *certain* if  $CF(A \rightarrow B) \geq \text{minCF}$ , where *minCF* is the minimum threshold for the certainty factor given by the

user. The definition for strong rules can be reformulated when using CF as a rule which must be frequent and certain.

### 2.2. Fuzzy Association Rules

In [5] the concepts of transaction and association rule are generalized to the fuzzy case. The proposal is based on the definition of fuzzy transaction as a non empty fuzzy subset  $\tau \subseteq I$ . For every item  $i \in I$  and every fuzzy transaction  $\tau$ , an item  $i$  will belong to  $\tau$  with degree<sup>1</sup>  $\tau(i)$  where  $\tau(i)$  is a real number in the interval  $[0, 1]$ . Let  $A$  be an itemset of  $I$ , that is, a subset of items, in a fuzzy transaction  $\tau$ . The membership degree of  $A \subseteq I$  to the fuzzy transaction  $\tau$  is defined as  $\tau(A) = \min_{i \in A} \tau(i)$ . In particular, a crisp transaction is a special case of fuzzy transaction where every item in the transaction will have membership degree equal to 1 or 0 depending on if it is in the transaction or not.

**Definition 3.** [5] Let  $I$  be a set of items,  $\tilde{D}$  a set of fuzzy transactions and  $A, B \in I$  two disjoint itemsets, i.e.  $A \cap B = \emptyset$ . A fuzzy association rule 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  $\tau$  in  $\tilde{D}$ .

This definition maintains the meaning of crisp association rules because if we need that  $A \subseteq \tau$  is satisfied, we also need that  $B \subseteq \tau$  be satisfied, in our case this can be translated to  $\tau(A) \leq \tau(B)$ . In this way, as a crisp transaction is a special case of fuzzy transaction, a crisp association rule will be a special case of fuzzy association rule.

In order to assess a fuzzy association rule we employ a proposal based on quantified sentence evaluation using the fuzzy quantifier  $Q_M(x) = x$  representing the quantifier “most” in the following way:

- The support of an itemset  $A$  is the evaluation of the quantified sentence “ $Q_M$  of the  $\tilde{D}$  are  $\Gamma_A$ ” where  $\Gamma_A$  is a fuzzy set defined as  $\Gamma_A(\tau) = \tau(A)$ .
- The support of a fuzzy rule  $A \rightarrow B$  in  $\tilde{D}$ , noted by  $\text{FSupp}(A \rightarrow B)$ , is the evaluation of the quantified sentence “ $Q_M$  of the  $\tilde{D}$  are  $(\Gamma_A \cap \Gamma_B)$ ”.
- The confidence of a fuzzy rule  $A \rightarrow B$  in  $\tilde{D}$ ,  $\text{FConf}(A \rightarrow B)$ , is the evaluation of the quantified sentence “ $Q_M$  of the  $\Gamma_A$  are  $\Gamma_B$ ”.
- The certainty factor of the fuzzy rule  $A \rightarrow B$  in  $\tilde{D}$ ,  $\text{FCF}(A \rightarrow B)$ , is computed using the fuzzy versions of support and confidence using Definition 2.

To evaluate the quantified sentence “ $Q$  of the  $A$  are

<sup>1</sup>Note that  $\tau(i)$  is  $\mu_\tau(i)$  where  $\mu_\tau : I \rightarrow [0, 1]$  is the membership function associated to the fuzzy set  $\tau$  defined on  $I$ .

$B$ ” we employ the GD method defined in [8] as

$$GD_Q(B/A) = \sum_{\alpha_i \in \Lambda(B/A)} (\alpha_i - \alpha_{i+1}) Q \left( \frac{|(B \cap A)_{\alpha_i}|}{|A_{\alpha_i}|} \right) \quad (1)$$

where  $\Lambda(B/A) = \Lambda(B \cap A) \cup \Lambda(A)$ , being  $\Lambda(A)$  the set of  $\alpha$ -cuts of  $A$ , and  $\Lambda(B/A) = \{\alpha_1, \dots, \alpha_p\}$  for  $\alpha_i > \alpha_{i+1}$  for every  $i \in \{1, \dots, p-1\}$  where  $\alpha_{p+1} = 0$ . The fuzzy set  $A$  must be normalized, if not,  $A$  will be normalized and the same factor of normalization will be applied to  $B \cap A$ . Some important properties of this proposal are given by the formulas (2) and (3).

$$\text{FConf}(A \rightarrow B) = 1 \Leftrightarrow \tau(A) \leq \tau(B), \forall \tau \in \tilde{D}. \quad (2)$$

$$\text{FCF}(A \rightarrow B) = 1 \Leftrightarrow \text{FConf}(A \rightarrow B) = 1. \quad (3)$$

**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 1. In particular, we can see that  $\tau_6$  is a crisp transaction. Some degrees of membership that we can find in  $\tilde{D}_1$  are the following:  $\tau_1(\{i_3, i_4\}) = 0.9$ ,  $\tau_1(\{i_2, i_3, i_4\}) = 0.2$  and  $\tau_2(\{i_1, i_2\}) = 1$ . If we consider the set of  $\alpha$ -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 in Table 2. Let us notice that to compute the assessment values for the FConf and the FCF the fuzzy set  $\{i_4\}$  has been normalized and the same normalization factor has been applied to  $\{i_4, i_5\}$ .

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$
$\tau_1$	1	0.2	1	0.9	0.9
$\tau_2$	1	1	0.8	0	0
$\tau_3$	0.5	0.1	0.7	0.6	0
$\tau_4$	0.6	0	0	0.5	0.5
$\tau_5$	0.4	0.1	0.6	0	0
$\tau_6$	0	1	0	0	0

Table 1: Set of fuzzy transactions  $\tilde{D}_1$

Rule	FSupp	FConf	FCF
$\{i_1, i_2\} \rightarrow \{i_3\}$	0.167	0.8	0.6
$\{i_4\} \rightarrow \{i_5\}$	0.2	0.767	0.68

Table 2: Some fuzzy rules obtained in  $\tilde{D}_1$

### 3. Meta-Association Rules

Meta-association rules are discovered from association rules previously mined from a set of databases. These databases must have similar structure and semantics, in order to obtain rules involving the same kind of items. The objective of meta-association rules is to extract globally-valid additional knowledge from previously extracted patterns in the form of association rules representing the information obtained from the original databases (primary data).

The meta-association rule mining process is depicted in Figure 1. Our starting point is a

set of association rules that have been previously mined from a set of databases. Formally, let  $D_1, D_2, \dots, D_k$  be  $k$  databases that may share some of their attributes. After applying a rule extraction procedure, we obtain  $k$  sets of association rules  $R_1, R_2, \dots, R_k$ , each  $R_i$  corresponding to a different  $D_i$ . We assume that these rules are crisp, but the procedure is analogous for other types of rules. The number of rules in  $R_1, R_2, \dots, R_k$ , noted  $nr_1, nr_2, \dots, nr_k$ , can be different, as well as the number of items in the antecedent or in the consequent of each rule. Interestingly, there may be some common rules in  $R_1, R_2, \dots, R_k$ . Without loss of generality, we also assume that the same threshold for the support and the confidence/certainty factor have been used for each dataset.

In order to discover the meta-association rules we create a structured database, called meta-database  $\mathcal{D}$ , which represents the different rules  $r_1, \dots, r_n$  in the sets  $R_i$ . The meta-database can be also enriched with more data describing features of  $D_i$ , which are aggregated by means of new attributes  $at_1, \dots, at_m$  that participate in the meta-rule mining process. Following the crime data example described in Section 5, each database includes the crime incidents happened in one district of the city of Chicago and the additional information are attributes about district characteristics provided by a statistics office, such as the number of residents or the security perception index.

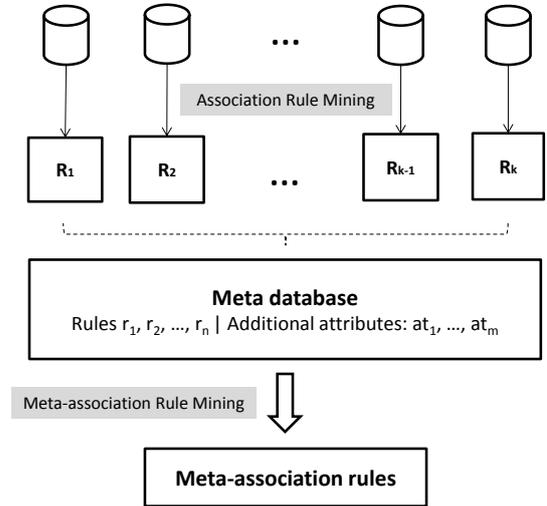


Figure 1: Process flow: From original datasets to final meta-association rules.

We continue with the two different developed strategies based on crisp rules and fuzzy rules respectively. A comparison between them is further included in Section 5.

#### 3.1. Crisp Meta-Association Rules

A first proposal for mining meta-association rules was developed in [2] by using crisp rules. In this

work, a binary meta-database  $\mathcal{D}$  is created by collecting the rules plus the extra information about the original datasets as shown in Table 3: if the rule  $r_i$  is found in  $D_j$ , then the value is 1; otherwise, it is 0. Afterwards, *crisp meta-association rules* are extracted from  $\mathcal{D}$  by employing a classical association rule mining algorithm, in our case using the *support-CF* framework because rules obtained using the certainty factor are more reliable than those extracted by using the confidence measure, as explained in Section 2.1. Using this approach, the obtained crisp meta-association rules 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:

- $r_i \rightarrow r_j$  where  $r_i, r_j$  can be rules or a conjunction of rules; for example:  $r_i = r_{i1} \wedge \dots \wedge r_{is}$ .
- $at_i \rightarrow at_j$  where  $at_i, at_j$  can be attributes or a conjunction of attributes.
- $r_i \rightarrow at_j$  or  $at_j \rightarrow r_i$  where  $r_i, at_j$  are a conjunction of rules and attributes, and can be mixed; i.e. a rule of the form  $r_1 \wedge at_2 \rightarrow r_3$  can be found.

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

Table 3: Example of boolean meta-database compiling previously extracted association rules and the additional attributes

It can be easily seen that this procedure for extracting crisp meta-rules 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  $CF = 0.5$  and with  $CF = 1$  have the same importance. These measurements can be incorporated into the meta-database by using continuous items with values in the interval  $[minCF, 1]$ . If we want to extract crisp meta-rules, the most direct alternative is to discretize the intervals by using items of the type  $\langle r_i, (CF_1, CF_2) \rangle$  where  $CF_j \in [minCF, 1]$ . However, this approach has some problems due to the crisp boundaries of intervals. For example, given the values  $CF(r_i) = 0.75$  and  $CF(r_j) = 0.76$ , 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 a different representation for continuous values. Fuzzy sets are an adequate option, since fuzzy meta-rules can be mined from a meta-database where the items are satisfied to some extent.

### 3.2. Fuzzy Meta-Association Rules

Fuzzy databases support the extraction of fuzzy association rules from a continuous representation of values. Following the same notation used in the previous section, we create a fuzzy meta-database,  $\widehat{\mathcal{D}}_{CF}$ , based on the certainty factor. The structure of a generic fuzzy meta-database is depicted in Table 4 where the value in column  $r_i$  and row  $D_j$  is the certainty factor of the rule  $r_i$  in database  $D_j$ , which is a value in the interval  $[minCF, 1]$ . Furthermore, we can use fuzzy values in the additional attributes; e.g., a “high security perception” with a degree of 0.8, representing that the population perceives a high security with a high degree.

$\widehat{\mathcal{D}}$	$r_1$	$r_2$	$\dots$	$r_n$	$at_1$	$\dots$	$at_m$
$D_1$	0.3	1	$\dots$	0	0.9	$\dots$	1
$D_2$	0	1	$\dots$	0.7	0	$\dots$	0.2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$D_k$	0.8	0	$\dots$	0.9	1	$\dots$	0

Table 4: Fuzzy meta-database compiling the obtained association rules and the additional attributes

Once the fuzzy meta-database has been constructed, a fuzzy association rule mining algorithm is applied to obtain the fuzzy meta-rules. We can also distinguish three different types of fuzzy meta-rules; namely, meta-rules that relate only rules, meta-rules that relate only attributes, and meta-rules that relate rules and attributes. Then, the extracted fuzzy meta-rules represent the relation between rules and/or attributes that have a high certainty factor in the original datasets.

### 4. Algorithm and Implementation Issues

In order to evaluate and analyse the proposal we design a process to include the entire flow from the original datasets to the final meta-rules. Our main goal in the experimental part is to analyse and compare the obtained meta-rules employing crisp and fuzzy approaches.

The complete algorithm for mining crisp meta-rules is described in Algorithm 1. When employing fuzzy meta-rules, the algorithm must be modified in step 18 for mining fuzzy association rules. For the experiments we have employed the algorithm described in [9] that computes the fuzzy assessment measures by means of a parallelization of the process by the  $\alpha$ -cuts. The proposed algorithm uses an itemset representation based on bit strings [10], which allows us to speed up logical operations with boolean data. Algorithm 1 implements the complete process depicted in Figure 1. From the original databases  $D_1, \dots, D_k$ , crisp association rules<sup>2</sup>

<sup>2</sup>We have considered that the regular rules are always crisp, but the same process can be also applied if fuzzy rules

are extracted. Next, the meta-database  $\mathfrak{D}$  or the fuzzy meta-database  $\mathfrak{D}$  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 (primary data) are not available, the meta-rule extraction procedure would start at step 14 (from the initial sets of rules).

---

**Algorithm 1** Meta-association rules mining

---

**Input:**  $D_1, \dots, D_k, at_1, \dots, at_m, minsupp, minCF$

**Output:**  $R_1, \dots, R_k$  and  $MR$  (set of meta-association rules)

```

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

```

---

The computational complexity of the algorithm depends on the number of transactions and items. The first step (lines 1-13) is  $\mathcal{O}(n2^i)$  for each original  $D$  (being  $n$  the number of transactions of  $D$  and  $i$  the number of different items), whereas the second step (lines 14-18) 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). Likewise, the complexity of the analogous fuzzy version using the Algorithm provided in [9] is also  $\mathcal{O}(k2^{m+r})$  after a parallelization for processing each  $\alpha$ -cut independently.

## 5. Experimental Evaluation

### 5.1. Parameters and datasets description

For the experimental evaluation of our proposal, we have used an open dataset of crime incidents reported by the police in the city of Chicago<sup>3</sup> in 2012

are provided since the same kind of information is available: rule, fuzzy support and fuzzy CF.

<sup>3</sup><https://data.cityofchicago.org/>

to study the possible relation between crimes in a neighborhood and the educational system. 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 for each database (see Table 5). We have split the dataset in 22 databases according to the districts of the city which is described in Table 5. For the meta-database we have also consider some crisp attributes of schools aggregated by district (number of students, number of misconducts and perceived safety index).

Table 5: Description of the databases

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

Let us notice that although the data used for the experimentation is real, the obtained results do not correspond and should not be extrapolated in any case to the real world.

### 5.2. Results

We have conducted diverse experiments in order to (1) compare the number of crisp and fuzzy meta-rules, and (2) compare the execution times for both approaches. For the experiments, we have used 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.

Experiments using different values for pairs of  $minSupp - minCF$ , and  $minFSupp - minFCF$

thresholds have been made to compare the number of obtained meta-rules and the execution times. In the left part of Figure 2, we show the number of meta-rules (in logarithmic scale) obtained for  $minSupp = minFSupp = 0.05$ , and  $minCF, minFCF \in \{0.2, 0.3, \dots, 0.8\}$ . We can see that the number of crisp meta-rules is generally larger than the number of fuzzy meta-rules. We can also observe that the number of crisp meta-rules is drastically reduced when the threshold of the certainty factor is slightly increased. In the fuzzy case the reduction of meta-rules is not so high because of the use of the rule importance measures in the meta-database  $\tilde{\mathcal{D}}_{CF}$ . In this regard, we can conclude 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 the execution time, in the right part of Figure 2 we show the time in seconds for the same set of parameters. We can see that the time consumed in the fuzzy case strongly depends on the number of rules obtained in each level. This happens because we have to calculate the crisp values in each  $\alpha$ -cut to compute  $FSupp$  and  $FCF$ . This explains the high execution time for  $\tilde{\mathcal{D}}_{CF}$  with  $minFSupp = 0.05$  and  $minFCF = 0.2$  (right chart in Figure 2), where the number of extracted rules exceed ten thousand rules.

We have selected in Table 6 different examples of meta-rules<sup>4</sup> that have been extracted in the city of Chicago dataset. This shows several types of meta-association rules representing relations among rules, attributes, and rules and attributes together. For instance, first meta-association rule states that there is a co-occurrence relation with certainty ( $FCF = 0.658$ ) between the very high number of misconducts and the low safety-index in a district joint with the rule (possession of cannabis  $\leq 30$  grams  $\rightarrow$  Domestic=false). This means that when there is a very high number of misconducts then it is frequent and reliable to have a low perception of security and a relation between the low possession of cannabis crime and its occurrence in a non-domestic environment.

### 5.3. Discussion

Meta-association rules have several advantages compared to traditional association rules. The first one, and the most evident, is that they can be built over secondary data when primary data is not available. In addition, our approach allows adding more information about the original datasets when available. Secondly, meta-association rules offer a new kind of knowledge, since regular rules may appear in the antecedent and/or the consequent of the meta-rules. For the city of Chicago we were interested in analysing the relationship between

the associations involving crime description data and the educational systems by districts. When a whole database collecting the information is considered we do not obtain information about the appearance of the associations in the majority of districts, instead the rules would represent single association between the items in the city without considering the division into districts. For instance, the third meta-rule of Table 6 reveals that in the 9.1% of districts there is an association relating the rule (if arrested then it is not a domestic crime) and the low safety-index with the medium number of students. This meta-rule differs from a corresponding similar association rule such as **Arrest=t AND Domestic=f AND Safety-Index=Low  $\rightarrow$  Number-of-Students=Medium** that would represent the association among these items in a high percentage of the crimes happened in Chicago. In addition, the attributes corresponding to the districts will be, in general, frequent items in the database collecting all the crimes since they have the same value in all transactions corresponding to the same district and this data is only computed by districts. This is one of the principal reasons that searching for meta-rules also allows to discover associations in data which are originally partitioned.

Moreover, when the meta-rule only involves rules, the meaning is completely different to that of regular rules since the relation represented by the meta-rule also involves previously extracted relations, not only items. This is the case, for instance, of the second meta-rule of Table 6 that relates the *presence* (since  $\mathcal{D}$  is used) of three regular rules in the 45.5% of the districts with  $CF = 0.778$ .

We want also to highlight the subtle difference between crisp and fuzzy meta-rules. Crisp meta-rules represent the co-occurrence of the *presence* of the regular rules and the additional attributes, meanwhile in fuzzy meta-rules the co-occurrence is *weighted* by the CF of the initial rules. Let us illustrate this with an example. Let us suppose that rules  $r_1$  and  $r_2$  are found in every original dataset with a  $CF$  around 0.1. If we mine crisp meta-rules, we probably obtain a meta-rule of the type  $r_1 \rightarrow r_2$  since it only takes into account the presence of both rules in all the datasets; on the opposite, if we mine fuzzy meta-rules using  $\tilde{\mathcal{D}}_{CF}$  it will not be extracted due to its low  $CF$  in all the databases. This is the case of the second rule in Table 6 which has been extracted in  $\tilde{\mathcal{D}}$  but not in  $\tilde{\mathcal{D}}_{CF}$ .

Besides, the user can also provide part of his/her knowledge about the original datasets by means of new fuzzy items of the form  $\langle attribute, linguistic label \rangle$  and a degree of satisfiability of that item. E.g. in a certain district the item  $\langle poverty-index, low \rangle$  is fulfilled with a degree of 0.9, meaning that we are 90% sure that the poverty index is low in the district, which is more natural and easier to interpret than an item of the type

<sup>4</sup>Desc. stands for Description,  $f$  for false and  $t$  for true.

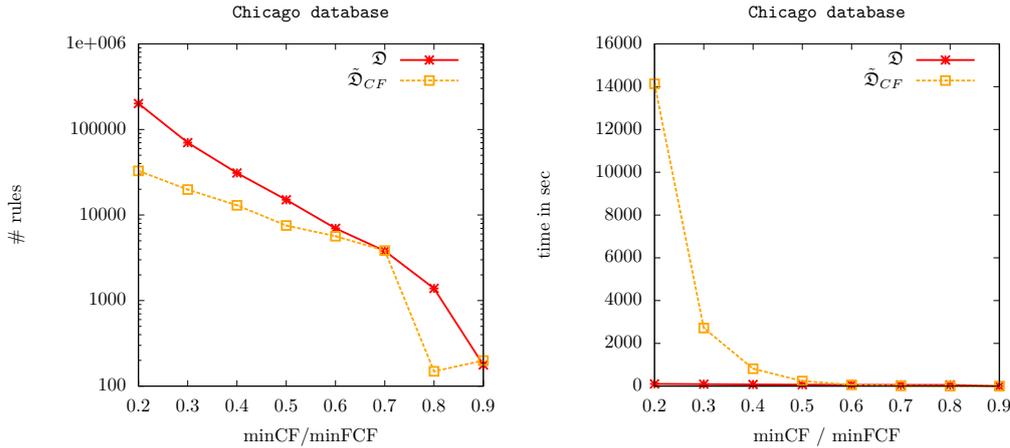


Figure 2: **Left:** Number of crisp/fuzzy meta-association rules (y-axis in logarithmic scale) vs  $minCF/minFCF$  (x-axis) when  $minSupp = 0.05$ . **Right:** Time in sec. for mining crisp/fuzzy meta-association rules (y-axis) vs  $minCF/minFCF$  (x-axis) when  $minSupp = 0.05$ .

Antecedent	Consequent	Supp/FSupp	CF/FCF	$\mathcal{D}$	$\tilde{\mathcal{D}}_{CF}$
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
(Arrest=t $\rightarrow$ Domestic=f) AND Safety-Index=Low	Number-of-Students=Medium	0.064	0.7	✓	x
Number-of-Students=Low AND Number-of-Misconducts=Very high	Safety-Index=Low	0.091	1	x	✓

Table 6: Examples of meta-association rules found in the City of Chicago dataset for  $minSupp = minFSupp = 0.05$  and  $minCF = minFCF = 0.5$ .

$\langle attribute, value \rangle$  such as `poverty-index = 1,02`.

To sum up, we can stress that: (i) According to the results of the crisp and fuzzy algorithms, the number of fuzzy meta-rules is in general more manageable than the number of crisp meta-rules. (ii) The process to mine fuzzy meta-rules is more adequate, since we take into account the degree of importance of the initial rules to build the meta-database. (iii) We can use fuzzy attributes for expressing additional knowledge about the original databases.

## 6. Related Work

Mining meta-association rules is a higher order mining technique since it discovers associations between associations. To the best of our knowledge, in the literature there is only one work that also develops this idea of meta-rules. In [11] fuzzy meta-rules are mined to discover temporal changes in rules that have been previously mined in different time periods. For that, the authors mine fuzzy rules over a set of association rules with a sequence of supports and a sequence of confidences in different time periods. Then, the fuzzy meta-rules capture the regularities governing how each association rule changes over time by using appropriate linguistic labels and

obtaining for each rule a fuzzy meta-rule of the type *Change in support in a period  $t_1 = Fairly decrease \rightarrow Change in support next period = Highly decrease$* . Compared to this approach, our concept of meta-rule is different because we may have association rules in the antecedent and/or in the consequent of the meta-rule. Furthermore, we do not consider sequences of supports/confidences, and conversely, we use the certainty factor.

There are also other approaches of HOM that use association mining over a set of clusters [12, 13, 14]. This kind of approaches are very interesting when the database is very large, because they scale very well when the number of rows and columns increase. The other way round, there are also approaches that incorporate association mining for processing the primary data, for then employing other mining techniques such as clustering and classification. Some of these works [15, 16], which are extensively described in [1], address the problem of clustering and classification of association rules. More recently, we can also highlight the works [17, 18, 19, 20], in which the creation of groups of rules mitigates one problem that often appears in association rule mining: the very large number of rules generated in the process. A slightly different approach based on a multi-tier granule mining was proposed in [20].

## 7. Conclusions

Meta-association rules are proposed in this paper as rules about rules that can be employed when the available information about several datasets is in the form of association rules. We have proposed different types of meta-association rules: crisp and fuzzy meta-rules. Fuzzy meta-rules take advantage of the assessment measures provided when mining rules from the original datasets. We have compared the different approaches obtaining that, in general, mining fuzzy meta-rules give a more manageable set of rules for its posterior analysis and they allow the use of fuzzy items to express additional knowledge about the original databases.

We have tested the approach in a set of coherent databases sharing the same set of attributes. An interesting issue to be addressed in the future is that of considering rules containing similar attributes that are not related beforehand. In this regard, the information should be semantically integrated allowing to link attributes corresponding to similar semantics. We plan to face this problem in the near future by using a knowledge repository assisting the algorithm in matching similar items.

## Acknowledgement

The research reported in this paper was partially supported by the Andalusian Government 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 7th Framework Programme [FP7/2007-2013] under grant agreement No 312651.

## References

- [1] J.F. Roddick, M. Spiliopoulou, D. Lister, and A. Ceglar. Higher order mining. *SIGKDD Explorations*, 10(1):5–17, 2008.
- [2] 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 *Proc. of the 17th Int. Conf. on Information Fusion*, 2014.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] E. Shortliffe and B. Buchanan. A model of inexact reasoning in medicine. *Mathematical Biosciences*, 23:351–379, 1975.
- [7] M. Delgado, M.D. Ruiz, and D. Sánchez. Studying interest measures for association rules through a logical model. *Int. J. of Uncert., Fuzziness and Knowledge-Based Systems*, 18(1):87–106, 2010.
- [8] M. Delgado, D. Sánchez, and M.A. Vila. Fuzzy cardinality based evaluation of quantified sentences. *Int. J. of Approximate Reasoning*, 23:23–66, 2000.
- [9] 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.
- [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] W.H. Au and K.C.C. Chan. Mining changes in association rules: a fuzzy approach. *Fuzzy Sets and Systems*, 149:87–104, 2004.
- [12] A. Nanopoulos, N. Apostolos, and Y. Manolopoulos. Mining association rules in very large clustered domains. *Information Systems*, 32:649–669, 2007.
- [13] 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.
- [14] T.T. Quan, L.N. Ngo, and S.C. Hui. An effective clustering-based approach for conceptual association rules mining. In *Int. Conf. on Computing and Communication Technologies, 2009. RIVF '09*, pages 1–7, 2009.
- [15] B. Lent, A. Swami, and J. Widom. Clustering association rules. In *13rd Int. Conf. on Data Engineering*, pages 220–231, 1997.
- [16] G.K. Gupta, A. Strehl, and J. Ghosh. Distance based clustering of association rules. In *Intelligent Engineering Systems Through Artificial Neural Networks*, pages 759–764, 1999.
- [17] A. Hashizume, B. Yongguang, X. Du, and N. Ishii. Generating representative from clusters of association rules on numeric attributes. In *IDEAL 2003, LNCS 2690*, pages 605–613, 2003.
- [18] A. An, S. Khan, and X. Huang. Objective and subjective algorithms for grouping association rules. In *3rd Int. Conf. on Data Mining, 2003. ICDM 2003*, pages 477–480, 2003.
- [19] A. An, S. Khan, and X. Huang. Hierarchical grouping of association rules and its application to a real-world domain. *Int. Journal of Systems Science*, 37(13):867–878, 2006.
- [20] Y. Li and J. Wu. Summarization of association rules in multi-tier granule mining. *IEEE Intelligent Informatics Bulletin*, 13(1):21–29, 2012.