*International Journal of Innovative Computing, Information and Control*
Volume **13**, Number **1**, February **2017**

 ISSN 1349-4198
pp. 289–305

# ENHANCING RELATIONAL ASSOCIATION RULES WITH GRADUALNESS

Istvan-Gergely Czibula, Gabriela Czibula and Diana-Lucia Miholca

Faculty of Mathematics and Computer Science
Babeş-Bolyai University
1, M. Kogălniceanu Street, Cluj-Napoca 400084, Romania
{istvanc; gabis}@cs.ubbcluj.ro; mdic1150@scs.ubbcluj.ro

Received August 2016; revised December 2016

Abstract. *We propose relational association rules enhanced with gradualness as a novel data mining concept. Relational association rules extend the conventional association rules to the ability to express generic relations persisting between the attributes within a data set. We propose a further extension which, by using gradual relations, is also aware of the degree to which a relation is satisfied. Its aim is to create a more expressive and stable approach for detecting relational association rules. We introduce an Apriori-like algorithm, GRANUM, which efficiently discovers all the interesting gradual relational association rules regardless of their length. Two case studies investigate the performance of using gradual relations in the mining process. The ability to capture semantically relevant rules, interpretability and noise-tolerance are evaluated. The results obtained confirm the improved efficiency of using gradual relations within the mining process thus highlighting the potential of our proposal.*
**Keywords:** Data mining, Relational association rules, Gradual relations

1. **Introduction.** Data mining techniques [22] are widely applied in various domains such as medicine, bioinformatics, bioarcheology, software engineering [18] to discover relevant patterns and rules in large volumes of data [26]. Association rule mining [25] consists in identifying attribute-value conditions that frequently appear together in data sets [4, 15, 30]. Although, classical association rules [29] do not consider measurable information that may occur between the attribute values and that can provide significant information regarding the mined data. In order to deal with fuzzy transactions, instead of crisp ones, the classical association rules have been extended by Delgado et al. [10] towards *fuzzy association rules*. The authors have also defined an afferent notion of *gradual rules* as relations of proportionality between degrees of membership. In that context, the membership degree quantifies the degree to which an attribute value can be characterized by a linguistic label.

Ordinal association rules [5] are specific association rules which express ordinal relationships that commonly appear in data. However, different informative relations may exist between the attribute values of the instances in a data set and many of them are different from ordinal. In these cases, ordinal association rules are not effective for expressing various interesting relationships which hold between data attributes. This is why we have previously extended in [27] the concept of *ordinal association rules* towards *relational association rules* (*RAR*s) which are able to depict different types of relations between attribute values. Compared to the classical association rule mining process, *relational* association rules are effective for uncovering much powerful rules that may lead to valuable data mining results.

There are situations in which it is relevant to consider the degree to which a relationship between two attributes is satisfied. For example, when using relational association rules for classification, it would be of interest to know not only that a relation holds, but to consider the magnitude to which the relation is satisfied. Classical *relational association rules* [27] are not powerful enough to describe such situations. In order to express the importance of the relationships which occur between record attributes, the boolean relationships have to be replaced with *gradual* [19] ones. A gradual relation is aware of the extent to which the relation is satisfied. For instance, a gradual ordinal relation, such as *less than*, will also consider how distant the values are. Moreover, when using crisp relational association rules in a data mining process, a drawback may be the sensitivity of the set of interesting relational association rules to noisy data. From this viewpoint, the robustness of relational association rule mining may be increased by using *gradual* relations, specifically *fuzzy* relations, instead of *boolean* ones.

The contribution of the paper is twofold. First, it extends classical *relational association rules* by enhancing them with gradualness [19]. So, we propose in this paper ***gradual*** *relational association rules* (*GRAR*s) as a novel data mining concept which uses *gradual relations*, specifically *fuzzy* [20] relations, instead of conventional boolean (or crisp) relations. *Gradual relations* have not been used so far in the data mining literature. Second, an *Apriori*-like algorithm (called *GRANUM* (*gradual relational association rules miner*)) is proposed, which efficiently discovers within a data set all interesting gradual relational association rules of any length. Two case studies evaluate the efficiency of the proposed approach regarding its ability to capture semantically relevant rules, its stability to noise and the interpretability of the rules discovered. The experimental data is both synthetic (two data sets out of three) and real (one data set out of three). We compare the sets of rules mined using the gradual approach with the ones mined using the *non-gradual* (or *crisp*) approach. Based on this comparison, we outline the benefits of using gradual relations.

To the best of our knowledge, the approach introduced in this paper is novel since there are no existing approaches in the data mining literature related to the discovery of *gradual relational association rules*.

The remainder of the paper is organized as follows. Section 2 gives the motivation behind our approach and its relevance. Section 3 presents the fundamentals of *relational association rules*. We describe our proposal for *gradual relational association rules* and the *GRANUM* algorithm in Section 4. In Section 5 we provide an experimental evaluation of our approach. Section 6 consists of a discussion on the obtained results. We conclude and give several directions for future improvements and research in Section 7.

## 2. Motivation. Potential applications.

*Relational association rules* (*RAR*s) have been applied in domains such as *medicine* (for diagnosis prediction [28]), *bioinformatics* (for predicting whether a DNA sequence contains a promoter region [8]) and *data cleansing* [5]. Mining medical data leads to information which are relevant in diagnosing and treating [7] diseases. In bioinformatics, genomic data analysis [31] is widely used for studying and predicting abnormal medical conditions. Association rules can also be used in data analysis, for outliers detection [5].

An inconvenience of using crisp *RAR*s in data mining tasks is that the set of interesting rules discovered is sensitive to noisy data [8], i.e., the set structure is changing, and the rules substantially vary. This is because the crisp approach fails to overlook the noise.

Many studies have suggested that fuzzy [24] concepts are particularly convenient to model noisy data [21]. However, noisy data is common in real-world problems due to inaccuracies, distortions and contamination occurring during data collection, storage or

transmission [12]. Even small measurements errors can have a big negative impact on the data mining process.

This motivates us to further extend relational association rules. We replace the crisp relations with *gradual relations*, specifically *fuzzy* relations. Consequently, we propose *gradual RAR*s as a novel data mining concept. The experimental section of the present paper evaluates the stability of the *GRAR*s. The evaluation results confirm that the gradual alternative is much more stable than the crisp one. So using gradual relations is preferable when mining possibly noisy data.

We also experiment the *GRAR* approach with imprecise (approximated) data [21]. The results show that *GRAR* mining algorithm succeeds in discovering the relevant rules even if the corresponding crisp (boolean) relations are occasionally violated due to approximations.

The non-gradual approach is unaware of the degree to which a relation is violated. For instance, $1000 < 1$ and $1 < 1$ are both false and therefore equally treated as negative examples. Similarly, it equally treats $1 < 1000$ and $1 < 1.001$. However, what if 1 is the result of an approximation and the real value is 1.0015? We consider that, in a crisp setting, the rigid transition between a positive instance and a negative one causes information loss. Gradual relations take advantage of a membership function which expresses the degree to which a relation is verified. The minor difference between $1 < 1$ and $1 < 1.001$ are reflected by a relative difference in membership values.

*GRAR*s extend the applicability of *RAR*s towards real-world data mining tasks involving impure data.

## 3. **Background.**

This section presents background concepts about *relational association rules*.

Traditional association rules [13] omit the relations between the attribute values, but in many practical situations such information is relevant, offering a deeper understanding of data. For example, if the attributes are numerical then partial orderings such as $\leq$ or $\geq$ may exist between the attribute values. However, if the attributes are non-numerical (e.g., set-valued attributes or attributes whose values are images) then more general relations (i.e., relations which are able to express different types of associations) are needed to be defined. This is why the traditional association rules have been extended towards *ordinal* and, more general, *relational* association rules. Câmpan et al. have introduced *ordinal association rules* (*OAR*s) in [6]. *OAR*s are able to express ordinal relations. In [27], we have extended the ordinal association rules towards *relational association rules* (*RAR*s). *RAR*s are able to express more general relations including relations between non-numerical attributes. The goal of *relational association rule mining* is to identify frequent generic relations between different combinations of attributes [27].

We briefly review in the following the concept of relational association rules.

Let $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ be a set of *instances* (entities, records), where each instance is characterized by a list of $m$ attributes, $\mathcal{A} = (a_1, \ldots, a_m)$. We denote by $\Phi(e_j, a_i)$ the value of attribute $a_i$ for an instance $e_j$ [23].

Relations which are not necessarily ordinal can be defined between two domains $D_i$ and $D_j$. Examples of ordinal relations are *less or equal* ($\leq$), *equal* ($=$), *greater or equal* ($\geq$). We denote by $\mathcal{R}$ the set of all these possible relations [27].

A *relational association rule* [27] is an expression $(a_{i_1}, a_{i_2}, a_{i_3}, \ldots, a_{i_\ell}) \Rightarrow (a_{i_1}\ R_1\ a_{i_2}\ R_2\ a_{i_3} \ldots R_{\ell-1}\ a_{i_\ell})$, where $\{a_{i_1}, a_{i_2}, a_{i_3}, \ldots, a_{i_\ell}\} \subseteq \mathcal{A} = \{a_1, \ldots, a_m\}$, $a_{i_j} \neq a_{i_k}$, $j, k = 1, \ldots, \ell$, $j \neq k$ and $R_j \in \mathcal{R}$ is a relation over $D_{i_j} \times D_{i_{j+1}}$, $D_{i_j}$ being the domain of the attribute $a_{i_j}$.

The rules which have their *support* and *confidence* greater than or equal to given thresholds are called *interesting*. They are of particular interest in a data mining task.

A complete and efficient algorithm for mining interesting $RAR$s has been introduced in [27]. It is an *Apriori* [1]-like algorithm and adapts the $DOAR$ (*discovery of ordinal association rules*) algorithm which has been proposed in [6].

4. **Gradual Relational Association Rules. Our proposal.** We introduce in the following *gradual relational association rules*. This concept extends the one of (crisp) *relational association rules*, which we have previously introduced in [27], by enhancing it with gradualness through the use of *fuzzy relations*.

4.1. **Theoretical model.** Let $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ be a set of *instances* (entities, records), where each instance is characterized by a list of $m$ attributes (features), $\mathcal{A} = (a_1, \ldots, a_m)$. Each attribute $a_i$ takes values from a non-empty and non-fuzzy domain $D_i$, which also contains a *null* (*empty*) value. We denote by $\Phi(e_j, a_i)$ the value of attribute $a_i$ in the instance $e_j$.

We define a fuzzy binary relation [3] $R$ between two domains $D_i$ and $D_j$ as follows:

$$R = \{<(x, y), \mu_R(x, y)> : x \in D_i, y \in D_j\}$$

$\mu_R : D_i \times D_j \to [0, 1]$ is a membership function which associates to each pair $(x, y)$, $x \in D_i$, $y \in D_j$ the *membership degree* $\mu_R(x, y)$ measuring the degree to which the relation $R$ is satisfied. $\mu_R(x, y) = 1$ means that $x$ and $y$ are completely related according to $R$. $\mu_R(x, y) = 0$ means that $x$ and $y$ are fully unrelated according to $R$ [3]. We denote by $\mathcal{F}$ the set of all possible binary fuzzy relations which can be defined between any two crisp attribute domains.

**Definition 4.1.** *A **gradual relational association rule**, gRule, is a sequence* $(a_{i_1}\ R_1\ a_{i_2}\ R_2\ a_{i_3} \ldots R_{\ell-1}\ a_{i_\ell})$, *where* $\{a_{i_1}, a_{i_2}, a_{i_3}, \ldots, a_{i_\ell}\} \subseteq \mathcal{A} = \{a_1, \ldots, a_m\}$, $a_{i_j} \neq a_{i_k}$, $j, k = 1, \ldots, \ell$, $j \neq k$ *and* $R_j \in \mathcal{F}$ *is a binary fuzzy relation over* $D_{i_j} \times D_{i_{j+1}}$. $D_{i_j}$ *is the domain of the attribute* $a_{i_j}$.

*The membership degree of the gradual relational association rule gRule for an instance* $e \in \mathcal{E}$ *is defined as* $\mu_{gRule}(e) = \min\{\mu_{R_j}(\Phi(e, a_{i_j}), \Phi(e, a_{i_{j+1}})), j = 1, 2, \ldots, \ell - 1\}$ *and expresses the grade or the magnitude to which the relation is satisfied.*

*a) If* $a_{i_1}, a_{i_2}, a_{i_3}, \ldots, a_{i_\ell}$ *are non-missing in* $m$ *instances from the data set then we call* $s = \frac{m}{n}$ *the support of the rule.*

*b) If we denote by* $E' \subseteq \mathcal{E}$ *the set of instances where* $a_{i_1}, a_{i_2}, a_{i_3}, \ldots, a_{i_\ell}$ *are non-missing and* $\mu_{gRule}(e) > 0$ *for each instance* $e$ *from* $E'$, *then we call* $c = \frac{|E'|}{n}$ *the confidence of the rule.*

*c) Using the notation from b), we call* $m = \dfrac{\sum_{e \in E'} \mu_{gRule}(e)}{n}$ *the rule's membership.*

The number of attributes in a *gradual relational association rule* is called its *length*. We note that the maximum length of a *gradual relational association rule* is equal to the number $m$ of the attributes characterizing the data.

We mention that the *membership degree* of a *gradual relational association rule* of length greater than 2 was defined using the min function which is a *t-norm* function [11].

The membership degree of a crisp (non-gradual, Boolean) relational association rule of any length is 1. It particularizes the membership degree of a *gradual* relational association rule to a rigid setting. The *interestingness* of a gradual relational association rule is defined similarly with the one of the non-gradual relational association rule.

**Definition 4.2.** *We call a GRAR interesting if its support s and confidence c are greater than or equal to user-specified thresholds, i.e.,* $s \geq s_{\min}$ *and* $c \geq c_{\min}$.

**Definition 4.3.** *The inverse of the binary fuzzy relation* $R = \{<(x,y), \mu_R(x,y)> : x \in X, y \in Y\}$ *will be denoted in the following by* $R^{-1}$ *and is defined as* $R^{-1} = \{<(x,y), 1 - \mu_R(x,y)> : x \in X, y \in Y\}$.

We mention that the crisp *RAR* is a particular case of *GRAR* whose membership and confidence coincide. The membership of a *GRAR* differs from its confidence, refining it and expressing the overall degree to which the relational rule is satisfied.

The membership is also a measure of *interestingness*. Consequently, it can be considered as a criterion for filtering the interesting *GRAR*s.

4.2. ***GRANUM* algorithm for gradual relational association rule mining.** We have extended the *Apriori*-like algorithm *DOAR* introduced in [6] to *GRANUM* which is able to discover interesting *gradual relational association rules* within data.

The first step in the mining process consists in defining the binary fuzzy relations which could occur between the attributes domains. This step is data dependent and the experimental part of the paper exemplifies it (Section 5).

Given a data set, *GRANUM* identifies, using an iterative process, the interesting gradual relational association rules of any length. The algorithm starts by computing the interesting 2-length gradual rules. Then, as in the *Apriori* algorithm, *GRANUM* generates candidate rules of length $k$ ($k \geq 3$) from the rules of length $k-1$. It prunes next the candidates which violates the minimum support and confidence constraint. As a consequence, only the interesting $k$-length rules are obtained. These are used in the next iteration. The iterative process stops when no new interesting rules are found.

The generation of candidate rules (denoted by *GenCandidates* in the algorithm from Figure 1) is a key element of *GRANUM*. The sub-algorithm is similar to the candidates generation process of the *DOAR* described in [6]. It joins two $k-1$ length rules $gradualRule_1$ and $gradualRule_2$ to obtain a $k$-length candidate rule *candRule*. Table 1 gives the three possibilities for joining.

TABLE 1. The process for candidate generation in the *GRANUM* algorithm

| Nr. | $gradualRule_1$ | $gradualRule_2$ | $candRule$ |
|---|---|---|---|
| 1. | $(a^1R^1a_{i_1}R_1a_{i_2}\ldots R_{k-3}a_{i_{k-2}})$ | $(a_{i_1}R_1a_{i_2}\ldots R_{k-3}a_{i_{k-2}}R^2a^2)$ | $(a^1R^1a_{i_1}R_1a_{i_2}\ldots R_{k-3}a_{i_{k-2}}R^2a^2)$ |
| 2. | $(a^1R^1a_{i_1}R_1a_{i_2}\ldots R_{k-3}a_{i_{k-2}})$ | $(a^2R^2a_{i_{k-2}}R_{k-3}^{-1}\ldots a_{i_2}R_1^{-1}a_{i_1})$ | $(a^1R^1a_{i_1}R_1a_{i_2}\ldots R_{k-3}a_{i_{k-2}}(R^2)^{-1}a^2)$ |
| 3. | $(a_{i_1}R_1a_{i_2}\ldots R_{k-3}a_{i_{k-2}}R^1a^1)$ | $(a_{i_{k-2}}R_{k-3}^1\ldots a_{i_2}{R_1}^{-1}a_{i_1}R^2a^2)$ | $(a^2(R^2)^{-1}a_{i_1}R_1a_{i_2}\ldots R_{k-3}a_{i_{k-2}}R^1a^1)$ |

The candidate generation process used by the *GRANUM* algorithm leads to a significant pruning of the exponential search space of all possible rules. It guides the search to those regions where interesting rules may be found, pruning the other regions.

The *GRANUM* algorithm is given in Figure 1. Our current implementation provides two features: (a) it discovers all interesting *GRARs* of any length; (b) it finds all maximal rules of any length, i.e., if a rule of a certain length can be extended with one attribute and it remains interesting then only the extended rule is kept.

Similar to the proof presented in [6], it is provable that the *GRANUM* algorithm preserves the completeness and correctness which characterize the *DOAR* algorithm.

5. **Experimental Results.** This section provides an experimental evaluation of the *gradual* relational association rule mining approach, emphasizing its advantages over the non-gradual approach.

**Algorithm** $GRANUM$ **is:**
**Input:**
- the data set $\mathcal{E}$ of $m$-dimensional entities
- the set $\mathcal{F}$ of binary fuzzy relations used in the mining process
- the minimum support ($s_{\min}$) and confidence ($c_{\min}$) thresholds

**Output:**
- the set $\mathcal{G}Rules$ of all interesting gradual relational association rules that hold over $\mathcal{E}$

**Begin**
$Cand \leftarrow \{ (a_{i_1}\ R\ a_{i_2}) \mid a_{i_1}, a_{i_2} \in A,\ i_1, i_2 = 1, \ldots, m,\ i_1 < i_2,\ R \in \mathcal{F} \}$;
Scan $\mathcal{E}$ and compute the support and confidence of candidates from $Cand$;
Keep in $Rules_2$ the interesting rules from $Cand$;
$k \leftarrow 2$;
While ($Rules_k \neq \emptyset$ and $k < m$) do
  $Cand \leftarrow GenCandidates(Rules_k)$;
  Scan $\mathcal{E}$ and compute the support and confidence of candidates from $Cand$;
  Keep in $Rules_{k+1}$ the interesting rules from $Cand$;
  $k \leftarrow k + 1$;
End;
$\mathcal{G}Rules \leftarrow \bigcup_{i=2}^{k} Rules_i$;
**End** $GRANUM$

FIGURE 1. The $GRANUM$ algorithm

5.1. **First case study.** We give in the following two easy-to-follow examples to illustrate that the gradual approach is able to capture rules which are inaccessible to the non-gradual one.

5.1.1. *Area experiment.* In the following we compare $RAR$s and $GRAR$s through a first numerical experiment. Our objective is to comparatively test the adaptability of the two approaches in the case in which data is altered by small approximation errors. Ideally, the set of interesting rules is stable, i.e., the interesting rules remain the same.

We propose as a first mining task to find equalities among numerical characteristics of triangles. Table 2 describes the data set.

TABLE 2. Attribute information in *Area* experiment

| | Attribute | Meaning |
|---|---|---|
| 1. | $a_1$ | Length of side a |
| 2. | $a_2$ | Length of side b |
| 3. | $a_3$ | Length of side c |
| 4. | $a_4$ | Length of height $h$, considering $c$ as base side |
| 5. | $a_5$ | $Area_h = (c * h)/2$ |
| 6. | $a_6$ | $Area_{Heron's} = \sqrt{p(p-a)(p-b)(p-c)}$ |

The data set is synthetically built by approximating to two decimal places the length of the height (attribute $a_4$). This approximation is natural assuming that the measurements are taken manually using a ruler. Such measurements approximations are very likely to appear in real cases. Moreover, an information system, due to its limited precision, can also cause minor variations in the results of the calculations it performs. Consequently, since many real data sets are obtained using such a system, small errors are very likely to appear.

TABLE 3. *Area* sample

| $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|
| 15 | 12 | 4 | 8.85 | 17.700 | 17.662 |
| 14 | 6 | 12 | 5.96 | 35.760 | 35.777 |
| 15 | 14 | 11 | 13.36 | 73.480 | 73.485 |
| 22 | 14 | 21 | 13.51 | 141.855 | 141.936 |
| 6 | 6 | 6 | 5.20 | 15.600 | 15.588 |
| 2 | 4 | 3 | 1.94 | 2.910 | 2.905 |
| 41 | 44 | 8 | 39.23 | 156.920 | 156.895 |
| 8 | 6 | 7 | 5.80 | 20.300 | 20.333 |
| 5 | 11 | 12 | 4.58 | 27.480 | 27.495 |
| ... | ... | ... | ... | ... | ... |

The actual values for the first nine instances are given in Table 3.

The interesting relation is that, as described in Table 2, $a_5$ and $a_6$ are numerical values representing the area of the corresponding triangle. So it is expected that their equality will be reported by the mining process.

However, the crisp approach *DOAR* fails to find any relational association rule. It fails even if we drastically decrease the minimum confidence threshold. This is explained by the fact that, because of the rounding errors, the area computed by using the height ($a_5$) is only approximately equal with the one given by Heron's formula ($a_6$).

*DOAR* is unaware of the degree to which the equality is violated. So no instance will contribute at the computation of confidence for the rule $a_5 = a_6$ and therefore, it is seen as completely uninteresting.

In contrast to the crisp approach, *GRANUM* is experimentally proved to be able to discover the rule of interest via the *fuzzy equal* ($\approx$) relation. The fuzzy relation is defined using the asymmetric Gaussian membership function, given in Formula (1), for $r = 0.2$ and $m = 2.0$.

$$\approx_{r,m}(x,y) = e^{-\frac{1}{2}\left(\frac{z-c}{s}\right)^m}, \text{ where}$$
$$c = \begin{cases} x, & \text{if } |x| \geq |y|. \\ y, & \text{otherwise.} \end{cases}, \quad z = \begin{cases} x, & \text{if } c = y. \\ y, & \text{otherwise.} \end{cases}, \quad s = r\frac{|x|+|y|}{2} \tag{1}$$

Figure 2 exemplifies graphically the membership function for relation $\approx$ obtained by fixing one of its arguments to 10 and letting the other one to vary between in $[0, 20]$.

Table 4 gives the results of the mining processes.

TABLE 4. Interesting *RAR*s vs. interesting *GRAR*s in *Area* experiment

| Crisp Rule | Confidence | Fuzzy Rule | Confidence | Membership |
|---|---|---|---|---|
| – | – | $a_5 \approx a_6$ | 1 | 0.999 |

As already stated, the non-gradual approach fails in discovering the rule of interest. However, *GRANUM* finds that $a_5 \approx a_6$ is interesting. The value of the membership degree $m = 0.999$ which describes the rule expresses that the two attributes are not exactly ($m < 1$) equal, so the gradual approach avoids loosing such information.

5.1.2. *Drug experiment.* We show by a second experiment that *GRANUM* is able to discover supplementary semantically interesting rules compared to the non-gradual approach.

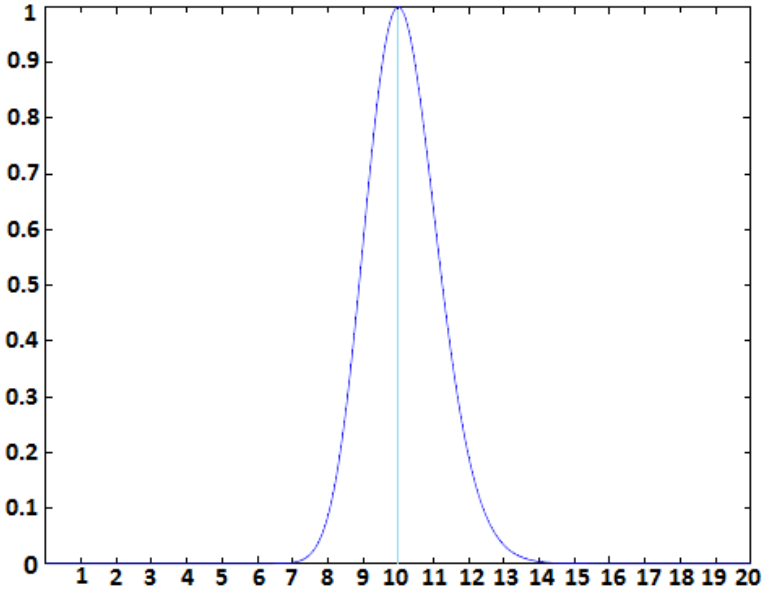


FIGURE 2. $\approx$'s membership function exemplified for $\approx (10, x)$

TABLE 5. Attribute information in *Drug* experiment [17]

| | Attribute | Meaning |
|---|---|---|
| 1. | $a_1$ | Average weight with **standard feed** |
| 2. | $a_2$ | Average weight with **low dose of drug** |
| 3. | $a_3$ | Average weight with **high dose** (twice the low dose) **of drug** |

TABLE 6. *Drug* data set [17]

| $a_1$ | $a_2$ | $a_3$ |
|---|---|---|
| 3.93 | 3.99 | 3.96 |
| 3.78 | 3.96 | 3.94 |
| 3.88 | 3.96 | 4.02 |
| 3.93 | 4.03 | 4.06 |
| 3.84 | 4.10 | 3.94 |
| 3.75 | 4.02 | 4.09 |
| 3.98 | 4.06 | 4.17 |
| 3.84 | 3.92 | 4.12 |

The data set, collected from [17], corresponds to a real experiment which has been developed to compare three different feeding treatments of birds. In an attempt to promote growth, two of the treatments include a drug in addition to standard food.

Table 5 describes the attributes of the data set and Table 6 gives their actual values.

The *GRANUM* algorithm is applied to discovering the binary interesting *GRARs* having the support and confidence greater than or equal to $s_{\min} = 1$ and $c_{\min} = 0.9$, respectively. We also constrain the algorithm to list only those rules whose *membership* is greater than or equal to 0.9 which ensures that the relations obtained are highly satisfied. There are two possibilities of imposing this restriction. We can either incorporate the condition into the algorithm as additional pruning or filter the rules in a further step.

The fuzzy relations considered in this experiment are: $\lesssim$, $\gtrsim$ and $\approx$.

TABLE 7. Interesting *RAR*s vs. interesting *GRAR*s in *Drug* experiment

| Crisp Rule | Confidence | Gradual Rule | Confidence | Membership |
|---|---|---|---|---|
| $a_1 < a_2$ | 1 | $a_1 \lesssim a_2$ | 1 | 0.907 |
| | | $a_1 \approx a_2$ | 1 | 0.923 |
| $a_1 < a_3$ | 1 | $a_1 \lesssim a_3$ | 1 | 0.913 |
| | | $a_2 \approx a_3$ | 1 | 0.967 |

Table 7 comparatively presents the mining results for gradual and non-gradual approaches. The crisp mining process uses relation $<$. The rules highlighted are discovered by the gradual approach, but omitted by the crisp one.

A semantic interpretation of the rules identified by the non-gradual approach is that the drug promotes the growth. However, we are unable to evaluate how much the drug promotes the growth or to measure the influence of the dose used.

By interpreting the results of the mining approach using gradual relations we infer the following additional facts.

- Since $a_2 \approx a_3$ with a high membership $m = 0.967$, the two treatments including the drug are approximately equally effective to promote growth.
- Considering that the membership of rule $a_1 \lesssim a_2$ (0.907) is slightly lower than the one of rule $a_1 \lesssim a_3$ (0.913), we also deduce that doubling the dose of drug has a very slightly positive overall impact in promoting growth. So, although with a higher cost and probably with a negative impact on health, the additional amount of drug improves only a little growth. This is practically useful to justify the disregard of the third treatment.
- Given that $a_1 \approx a_2$ (and $a_2 \approx a_3$), the drug is not very efficient.

All these relevant information are nondeductible by using the crisp approach.

Concluding on the results of this case study, the expressiveness of the gradual approach is dual. On the one hand, it succeeds in discovering additional semantically interesting rules. The crisp approach fails to identify these rules due to minor violations. On the other hand, *membership* is more expressive than *confidence* since it captures the magnitude of rules, i.e., how strong or well satisfied they are.

5.2. **Second case study.** The second case study aims to emphasize that gradual relational association rule mining is more robust to noise than the non-gradual approach.

5.2.1. *Data set.* The values for the attributes of the experimental data set are generated using functions. For an $x \in \mathbb{R}^*$, the value for the attribute $a_i$ of the corresponding data instance is the result of $f_i(x)$, where $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, 4$ are defined in Table 8.

Provided that $x \neq 0$, the following relations occur between the attributes: $a_1 < a_2$, $a_1 < a_3$, $a_1 < a_4$, $a_2 < a_3$, $a_2 < a_4$ and $a_3 < a_4$.

TABLE 8. Attribute description in the second case study

| | Attribute | Meaning |
|---|---|---|
| 1. | $a_1$ | $f_1(x) = -x^2$ |
| 2. | $a_2$ | $f_2(x) = x^2$ |
| 3. | $a_3$ | $f_3(x) = x^2 + \lvert x \rvert$ |
| 4. | $a_4$ | $f_4(x) = x^2 + 2 * \lvert x \rvert$ |

TABLE 9. Synthetic data set

| x | $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|---|
| −1.285951416224598 | −1.653671045 | 1.653671045 | 2.939622461 | 4.225573877 |
| 1.926702979318315 | −3.712184371 | 3.712184371 | 5.63888735 | 7.565590329 |
| 2.136271084620817 | −4.563654147 | 4.563654147 | 6.699925232 | 8.836196316 |
| 0.9539822448311869 | −0.910082123 | 0.910082123 | 1.864064368 | 2.818046613 |
| 4.915958577765329 | −24.166648738 | 24.166648738 | 29.082607316 | 33.998565894 |
| −3.965867399697877 | −15.728104232 | 15.728104232 | 19.693971632 | 23.659839031 |
| 7.119361480481611 | −50.68530789 | 50.68530789 | 57.80466937 | 64.924030851 |
| 15.26207236237816 | −232.930852794 | 232.930852794 | 248.192925157 | 263.454997519 |
| −3.286034146377554 | −10.798020411 | 10.798020411 | 14.084054558 | 17.370088704 |
| 2.365309560849711 | −5.594689319 | 5.594689319 | 7.959998879 | 10.32530844 |
| ... | ... | ... | ... | ... |

TABLE 10. Interesting *RAR*s vs. *GRAR*s in the initial data set

| Crisp Rule | Confidence | Gradual Rule | Confidence | Membership |
|---|---|---|---|---|
| $a_1 < a_2$ | 1 | $a_1 \lesssim a_2$ | 1 | 1 |
| $a_1 < a_3$ | 1 | $a_1 \lesssim a_3$ | 1 | 1 |
| $a_1 < a_4$ | 1 | $a_1 \lesssim a_4$ | 1 | 1 |
| $a_2 < a_3$ | 1 | $a_2 \lesssim a_3$ | 1 | 0.9696 |
| $a_2 < a_4$ | 1 | $a_2 \lesssim a_4$ | 1 | 0.9899 |
| $a_3 < a_4$ | 1 | $a_3 \lesssim a_4$ | 1 | 0.9653 |

The functions from Table 8 are evaluated for 1000 distinct points. The points are randomly generated from the normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 10$. Table 9 gives the first 10 instances from the resulting data set.

Table 10 depicts the *RARs* discovered for a confidence threshold $c_{\min} = 0.95$ and using $<$ and $\lesssim$ relations for the crisp and gradual mining processes, respectively. Unlike the crisp approach, the fuzzy adaptation is able to discover all the ordinal rules expected.

5.2.2. *Addition of noise.* To test the stability of the gradual and non-gradual mining approaches, we incrementally introduce Gaussian noise of mean 0 and increasing standard deviation. We insert noise by perturbing the functions generating the values of the attributes. The value of the attribute $a_i$ in an instance $x$ is updated by adding a value $\triangle a_i(x) = n * x$, $n \sim \mathcal{N}(0, stdev)$, $i = 1, \ldots, 4$.

We successively consider the following standard deviations: 0.5, 0.75, 1, 1.25, 1.5 and 1.75. Figure 3 shows the impact of the Gaussian noise with a standard deviation $stdev = 0.5$. The first 10 instances from the noisy data set obtained for $stdev = 0.5$ are given in Table 11.

5.2.3. *Evaluation measure.* We define a measure which expresses the *stability* of a set of rules and we use it to evaluate how robust non-gradual *RAR*s and *GRAR*s are.

Let us consider a data set $\mathcal{D}$ containing a set of (crisp or gradual) interesting relational association rules $\mathcal{R}ules = \{R_1, \ldots, R_k\}$ (the crisp rules are discovered using the *RAR* method and the gradual rules are discovered using the *GRANUM* method). Each rule $R_i$ from the set $\mathcal{R}ules$ has a confidence $c_i$ within $\mathcal{D}$ and, additionally, a membership degree $m_i$ if it is a *gradual* rule. After introducing noise with a certain standard deviation *stdev* in the data set $\mathcal{D}$, we mine the rules $\mathcal{R}ules_{stdev}^{noise}$ in the noisy data.

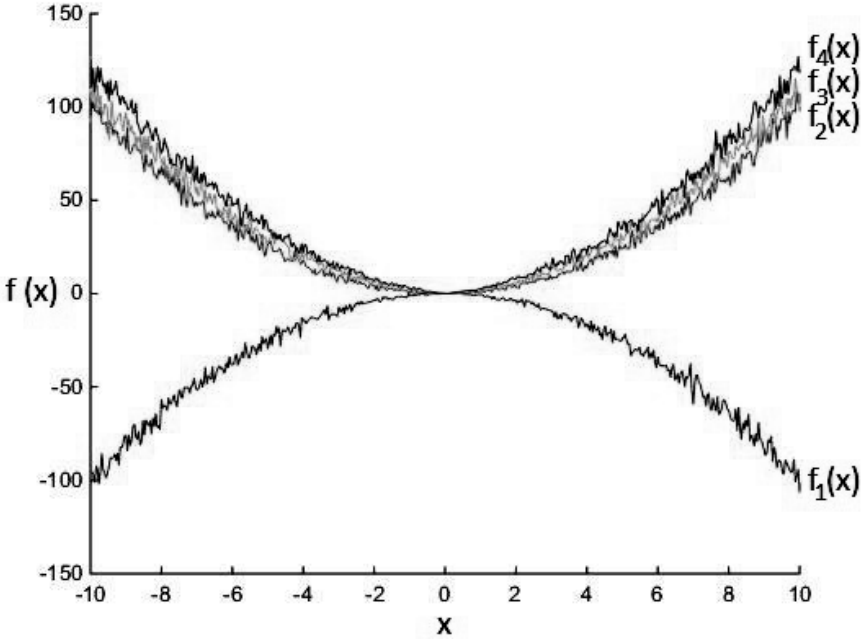


FIGURE 3. Graphical representation of the functions given in Table 8 perturbed by Gaussian noise with $stdev = 0.5$

TABLE 11. Noisy data set generated for Gaussian noise with $stdev = 0.5$

| $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|
| −1.56446388 | 2.26186933 | 3.270616422 | 4.39979761 |
| −3.775955537 | 2.874227317 | 5.410598751 | 8.263267824 |
| −3.539165637 | 4.255660892 | 6.919224393 | 8.176143604 |
| −0.553751097 | 0.689368714 | 1.364894073 | 2.184253559 |
| −23.241562208 | 23.118459399 | 29.173979267 | 35.569777021 |
| −19.247477901 | 15.896638853 | 18.425601202 | 22.042695618 |
| −56.633075185 | 48.707363757 | 57.246696425 | 65.191846898 |
| −219.513832455 | 237.818665209 | 254.283245247 | 259.507882212 |
| −11.138121732 | 10.880560075 | 13.868560545 | 15.682657759 |
| −5.368169183 | 6.43938355 | 7.546232536 | 11.133250622 |
| ... | ... | ... | ... |

Definition 5.1 expresses the degree to which the set of interesting relational association rules identified in a data set $\mathcal{D}$ remains stable (or unchanged) after introducing noise with standard deviation *stdev*.

**Definition 5.1.** ***Stability to noise – STAB.*** *Let us denote by $\mathcal{R}(\mathcal{D})$ the set of interesting relational association rules discovered in a data set $\mathcal{D}$ and by $\mathcal{R}ules_\sigma(\mathcal{D})$ the set of interesting relational association rules discovered in $\mathcal{D}$ after adding noise of standard deviation $\sigma$. We denote by $STAB(\mathcal{D}, \sigma)$ the stability of the set of interesting relational association rules discovered in $\mathcal{D}$ with respect to noise of standard deviation $\sigma$ and we define it as*

$$STAB(\mathcal{D}, \sigma) = \frac{|\mathcal{R}(\mathcal{D}) \cap \mathcal{R}ules_\sigma(\mathcal{D})|}{|\mathcal{R}(\mathcal{D}) \cup \mathcal{R}ules_\sigma(\mathcal{D})|} \tag{2}$$

$STAB$ has the following two properties.

- $0 \leq STAB(\mathcal{D}, \sigma) \leq 1$.

- $STAB(\mathcal{D}, \sigma)$ is equal to 0 iff $\mathcal{R}(\mathcal{D}) \cap \mathcal{R}ules_{\sigma}(\mathcal{D}) = \emptyset$, i.e., the set of relational association rules discovered after adding the noise is completely different from the set of rules discovered in the initial data set.
- $STAB(\mathcal{D}, \sigma)$ is equal to 1 iff $\mathcal{R}(\mathcal{D}) = \mathcal{R}ules_{\sigma}(\mathcal{D})$, i.e., the set of relational association rules discovered remains unchanged after adding the noise.

**Remark 5.1.** *Larger values for STAB indicate a better stability to noise. So, STAB has to be maximized in order to obtain a more robust mining model.*

5.2.4. *Results.* We apply both crisp and gradual mining approaches to six noisy data sets obtained as described in Section 5.2.2. Tables 12-17 give the rules mined and highlight the changes with respect to the interesting rules mined in the initial data set (Table 10).

TABLE 12. Interesting *RAR*s vs. *GRAR*s for noise with $stdev = 0.5$

| Crisp Rule | Confidence | Gradual Rule | Confidence | Membership |
|---|---|---|---|---|
| $a_1 < a_2$ | 0.969 | $a_1 \lesssim a_2$ | 0.981 | 0.9737 |
| $a_1 < a_3$ | 0.998 | $a_1 \lesssim a_3$ | 1 | 0.9981 |
| $a_1 < a_4$ | 1 | $a_1 \lesssim a_4$ | 1 | 1 |
| | | $a_2 \lesssim a_3$ | 0.998 | 0.957 |
| $a_2 < a_4$ | 0.996 | $a_2 \lesssim a_4$ | 1 | 0.9954 |
| | | $a_3 \lesssim a_4$ | 1 | 0.9636 |

TABLE 13. Interesting *RAR*s vs. *GRAR*s for noise with $stdev = 0.75$

| Crisp Rule | Confidence | Gradual Rule | Confidence | Membership |
|---|---|---|---|---|
| $a_1 < a_2$ | 0.974 | $a_1 \lesssim a_2$ | 0.983 | 0.9779 |
| $a_1 < a_3$ | 0.994 | $a_1 \lesssim a_3$ | 0.996 | 0.9948 |
| $a_1 < a_4$ | 0.999 | $a_1 \lesssim a_4$ | 1 | 0.9991 |
| | | $a_2 \lesssim a_3$ | 0.987 | 0.9166 |
| $a_2 < a_4$ | 0.968 | $a_2 \lesssim a_4$ | 1 | 0.9849 |
| | | $a_3 \lesssim a_4$ | 0.988 | 0.9286 |

TABLE 14. Interesting *RAR*s vs. *GRAR*s for noise with $stdev = 1$

| Crisp Rule | Confidence | Gradual Rule | Confidence | Membership |
|---|---|---|---|---|
| | | $a_1 \lesssim a_2$ | 0.972 | 0.9554 |
| $a_1 < a_3$ | 0.984 | $a_1 \lesssim a_3$ | 0.991 | 0.9872 |
| $a_1 < a_4$ | 0.995 | $a_1 \lesssim a_4$ | 0.999 | 0.9969 |
| | | $a_2 \lesssim a_3$ | 0.965 | 0.8611 |
| | | $a_2 \lesssim a_4$ | 0.992 | 0.9505 |
| | | $a_3 \lesssim a_4$ | 0.983 | 0.8952 |

Table 18 summarizes the numerical evaluation results. We notice the gradual association rules mining approach is much more robust to noise than the *non-gradual* one. Figure 4 illustrates a comparison between the gradual and non-gradual approaches regarding their stability to noise.

We can summarize the results of the second case study as follows. The crisp approach fails to discover all the initial interesting *RAR*s even if the magnitude of the noise is

TABLE 15. Interesting *RAR*s vs. *GRAR*s for noise with $stdev = 1.25$

| Crisp Rule | Confidence | Gradual Rule | Confidence | Membership |
|---|---|---|---|---|
| | | $a_1 \lesssim a_2$ | 0.96 | 0.9527 |
| $a_1 < a_3$ | 0.982 | $a_1 \lesssim a_3$ | 0.992 | 0.9855 |
| $a_1 < a_4$ | 0.998 | $a_1 \lesssim a_4$ | 0.999 | 0.9984 |
| | | $a_2 \lesssim a_3$ | 0.962 | 0.8442 |
| | | $a_2 \lesssim a_4$ | 0.997 | 0.944 |
| | | $a_3 \lesssim a_4$ | 0.977 | 0.8531 |

TABLE 16. Interesting *RAR*s vs. *GRAR*s for noise with $stdev = 1.5$

| Crisp Rule | Confidence | Gradual Rule | Confidence | Membership |
|---|---|---|---|---|
| | | $a_1 \lesssim a_2$ | 0.963 | 0.9465 |
| $a_1 < a_3$ | 0.957 | $a_1 \lesssim a_3$ | 0.969 | 0.9621 |
| $a_1 < a_4$ | 0.99 | $a_1 \lesssim a_4$ | 0.996 | 0.9929 |
| | | $a_2 \lesssim a_4$ | 0.984 | 0.9068 |
| | | $a_3 \lesssim a_4$ | 0.965 | 0.8314 |

TABLE 17. Interesting *RAR*s vs. *GRAR*s for noise with $stdev = 1.75$

| Crisp Rule | Confidence | Gradual Rule | Confidence | Membership |
|---|---|---|---|---|
| | | $a_1 \lesssim a_2$ | 0.956 | 0.9432 |
| | | $a_1 \lesssim a_3$ | 0.972 | 0.9561 |
| $a_1 < a_4$ | 0.981 | $a_1 \lesssim a_4$ | 0.988 | 0.9845 |
| | | $a_2 \lesssim a_4$ | 0.968 | 0.8887 |
| | | $a_3 \lesssim a_4$ | 0.952 | 0.8073 |

TABLE 18. Results for the second case study for $s_{\min} = 1$ and $c_{\min} = 0.95$

| **Noise's *stdev* $\sigma$** | **Approach** | **STAB** |
|---|---|---|
| 0.0 | Non-gradual | 1 |
| | Gradual | 1 |
| 0.5 | Non-gradual | 0.6667 |
| | Gradual | **1** |
| 0.75 | Non-gradual | 0.6667 |
| | Gradual | **1** |
| 1 | Non-gradual | 0.3333 |
| | Gradual | **1** |
| 1.25 | Non-gradual | 0.3333 |
| | Gradual | **1** |
| 1.5 | Non-gradual | 0.3333 |
| | Gradual | **0.8333** |
| 1.75 | Non-gradual | 0.1667 |
| | Gradual | **0.8333** |

relatively small ($stdev = 0.5$, Table 12), while the gradual adaption succeeds in keeping the set of interesting *GRAR*s unchanged until the noise's standard deviation triples (Table 16). Furthermore, the number of rules omitted by the crisp approach grows rapidly such

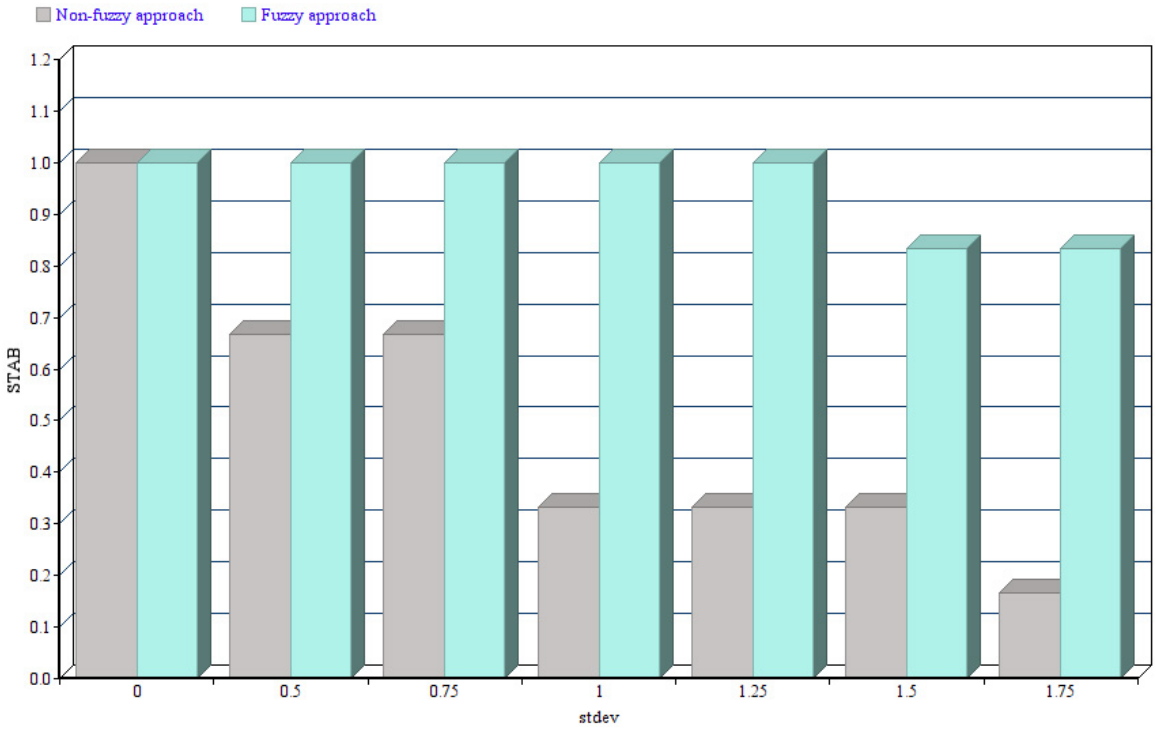


FIGURE 4. Stability to noise. *RAR*s vs. *GRAR*s.

that, when the gradual process starts to lose one rule, the crisp process already omits five rules of six. However, the gradual approach reflects the small changes produced by noise by small changes in memberships. So, it is aware of the small variations in data but it expresses this variation softly, through membership fluctuation, rather than rigidly, through removing elements from the set of interesting rules.

Even for noise of relatively small magnitude (with standard deviation $stdev = 0.5$), the non-gradual approach fails to obtain the same set of interesting *RAR*s as in the initial data set, while, by applying the gradual approach, the *GRAR*s discovered in the noisy data sets remain the same (Table 12).

6. **Discussion.** In this section we examine the results of the experimental evaluation. Based on the experimental results (Section 5), we conclude that the *gradual* approach to *RAR*s mining has two main advantages over the *non-gradual* approach.

Firstly, the gradual adaptation is able to discover relevant rules which the non-gradual approach fails to discover. The additional rules (e.g., several attributes have approximately equal values for most of the instances in the data set) discovered by the gradual approach are often important in the data mining process. For instance, if we think to use the rules for classification [9], these rules, if they are discriminating (i.e., specific to a class and unspecific to the others), increase the accuracy of the classifier.

Secondly, according to the experiment conducted (Section 5.2), gradual relational association rules approach is more robust to noise than non-gradual approach and thus more suitable to model data which is possibly noisy. Table 18 and Figure 4 show that the stability to noise ($STAB$) decreases, for both *RAR*s and *GRAR*s, as the degree of noise increases. This is expected but we notice that the crisp rules are much more affected by noise than the gradual ones. The set of non-gradual rules fluctuates even if the magnitude of the noise altering the data is small. The crisp approach disregards the rules which noise convert from interesting to uninteresting, but the gradual approach expresses the noise by decreasing the membership proportionally. As a consequence, the fluctuation of

the interesting $GRAR$s is delayed which means that the set of interesting gradual rules remains stable with respect to noise of small magnitude, starting to change only when the noise becomes critical. So, the gradual adaption is also affected by noise, but later and less.

An additional advantage of the proposed approach is given by the *membership*. Firstly, the fuzzy relations are highly adjustable. We can adapt the membership functions according to the particularities of the data set considered. Secondly, the membership of $GRAR$s is semantically relevant. It refines the confidence, expressing the overall degree to which the $GRAR$ is satisfied. Accordingly, it can serve as a criterion for comparing and filtering the rules.

6.1. **Comparison to existing methods.** As stated before, the concept of *gradual relational association rules* is a novel one. However, we provide in the following a comparison between $GRAR$s and similar concepts from the data mining literature, highlighting the benefits of our proposal.

Compared to the classical *relational association rules* which are able to describe relationships between the attributes values in a data set [27], *gradual relational association rules* are more expressive. They are able to depict the strength of the relationships thus expressing the degree to which the relationships are satisfied. The experiments from Section 5 emphasize two major advantages of $GRAR$s compared to the classical $RAR$s. First, we conclude here about the ability of the gradual approach to deal better with noisy data than the non-gradual one. The set of interesting $GRAR$s is more stable than the set of $RAR$s when the data instances are affected by noise. The second improvement is the capability of $GRAR$s to uncover additional semantically interesting rules compared to the non-gradual approach.

The concept of *association rules* has been generalized so as to afford *fuzzy* transactions. Therefore, the concept of *fuzzy association rules* [10] has been defined as (non-relational) association rules in a set of fuzzy transactions through the use of membership degrees indicating the degrees of inclusion of the itemsets in the non-crisp transactions. Delgado et al. [10] have also defined the term of *gradual rule* as an expression with the following form: "The more attribute $X$ is in $L_i^X$, the more attribute $Y$ is in $L_j^Y$", with $L_i^X$ and $L_j^Y$ linguistic labels for the attributes $X$ and $Y$, respectively.

Through the current proposal, we redefined the concept of *gradual rule* in the context of *relational association rules*. In contrast to the above mentioned approaches, the data mined remain crisp, while the relations themselves are fuzzy, thus characterized by a membership function expressing gradualness.

An alternative to extend the classical *relational association rules*, besides the current proposal which uses *fuzzy relations* instead of crisp ones, would be to use *fuzzy* sets for representing the attributes characterizing the data. The later perspective would explore a *gradual* viewpoint with respect to the data, not to the relations. In the present approach of exploring the gradualness of relations instead of a fuzzy view of the data, we started from the intuition that the use of *gradual relations* we propose would be more expressive than crisp relations between fuzzified attributes. The extent to which the relations are satisfied may be more significant for expressing relevant patterns, particularly when the gradual relations are learned from data and thus not predefined.

7. **Conclusions and Further Work.** We proposed the concept of *gradual relational association rules* as an extension of relational association rules [27] towards a gradual approach. The approach is novel since there are no other approaches using fuzzy [2] relations in association rules mining.

Based on the results of the experiments performed, we conclude that gradual relations are more comprehensive, more expressive and less sensitive to noise.

We intend to extend the experimental evaluation to investigate the effectiveness of $GRAR$s for classification, to capitalize the use of non-ordinal relations and to hybridize gradual relational association rules and *artificial neural networks* ($ANN$s) [14] in a supervised learning [16] setting in which the $ANN$s will learn the membership function.

**Acknowledgment.** This work was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS-UEFISCDI, project number PN-II-RU-TE-2014-4-0082. The authors also gratefully acknowledge the helpful comments and suggestions of the editor and the reviewers, which have improved the paper and the presentation.

## REFERENCES

[1] R. Agrawal and R. Srikant, Fast algorithms for mining association rules in large databases, *Proc. of the VLDB'20*, pp.487-499, 1994.

[2] M. A. Khanesar and M. Teshnehlab, Direct fuzzy model reference controller for siso nonlinear plants using observer, *International Journal of Innovative Computing, Information and Control*, vol.6, no.1, pp.297-306, 2010.

[3] H. Borzecka, Multi-criteria decision making using fuzzy preference relations, *Operations Research and Decisions*, vol.3, pp.5-21, 2012.

[4] T. Calders, N. Dexters, J. J. M. Gillis and B. Goethals, Mining frequent itemsets in a stream, *Inf. Syst.*, vol.39, pp.233-255, 2014.

[5] A. Câmpan, G. Şerban and A. Marcus, Relational association rules and error detection, *Studia Universitatis Babes-Bolyai Informatica*, no.1, pp.31-36, 2006.

[6] A. Câmpan, G. Şerban, T. M. Truta and A. Marcus, An algorithm for the discovery of arbitrary length ordinal association rules, *DMIN*, pp.107-113, 2006.

[7] K. J. Cios, *Medical Data Mining and Knowledge Discovery*, Springer, 2001.

[8] G. Czibula, M.-I. Bocicor and I. G. Czibula, Promoter sequences prediction using relational association rule mining, *Evolutionary Bioinformatics*, vol.8, pp.181-196, 2012.

[9] G. Czibula, Z. Marian and I. G. Czibula, Software defect prediction using relational association rule mining, *Inf. Sci.*, vol.264, pp.260-278, 2014.

[10] M. Delgado, N. Marin, D. Sanchez and M. A. Vila, Fuzzy association rules: General model and applications, *Trans. Fuz. Sys.*, vol.11, no.2, pp.214-225, 2003.

[11] M. Gagolewski and J. Lasek, The use of fuzzy relations in the assessment of information resources producers' performance, *Proc. of the 7th IEEE Int. Conference Intelligent Systems*, Warsaw, Poland, pp.289-300, 2014.

[12] L. P. F. Garcia, A. C. P. L. F. de Carvalho and A. C. Lorena, Effect of label noise in the complexity of classification problems, *Neurocomputing*, vol.160, pp.108-119, 2015.

[13] A. Gosain and M. Bhugra, A comprehensive survey of association rules on quantitative data in data mining, *IEEE Conference on ICT*, pp.1003-1008, 2013.

[14] D. Graupe, *Principles of Artificial Neural Networks*, World Scientific Publishing Company, 2013.

[15] J. Han, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.

[16] S. B. Kotsiantis, I. Zaharakis and P. Pintelas, *Supervised Machine Learning: A Review of Classification Techniques*, 2007.

[17] D. J. Hand, F. Daly, K. McConway, D. Lunn and E. Ostrowski, *A Handbook of Small Data Sets*, CRC Press, 1993.

[18] H. He, T. Yin, J. Dong, P. Zhang and J. Ren, Efficient mining of high utility software behavior patterns from software executing traces, *International Journal of Innovative Computing, Information and Control*, vol.11, no.5, pp.1779-1793, 2015.

[19] E. Hüllermeier, Association rules for expressing gradual dependencies, *Proc. of PKDD'06*, pp.200-211, 2002.

[20] M. Ksantini, M. A. Hammami and F. Delmotte, On the global exponential stabilization of Takagi-Sugeno fuzzy uncertain systems, *International Journal of Innovative Computing, Information and Control*, vol.11, no.1, pp.281-294, 2015.

[21] F. J. Lopez, A. Blanco, F. Garcia, C. Cano and A. Marin, Fuzzy association rules for biological data analysis: A case study on yeast, *BMC Bioinformatics*, vol.9, no.1, pp.1-18, 2008.

[22] X. Ma, J. Wang, L. Zhao and R. Li, Mutual exclusion role constraint mining based on weight in role-based access control system, *International Journal of Innovative Computing, Information and Control*, vol.12, no.1, pp.91-101, 2016.

[23] A. Marcus, J. I. Maletic and K.-I. Lin, Ordinal association rules for error identification in data sets, *Proc. of CIKM*, New York, NY, USA, pp.589-591, 2001.

[24] A. E. Ougli and B. Tidhaf, Optimal type-2 fuzzy adaptive control for a class of uncertain nonlinear systems using an LMI approach, *International Journal of Innovative Computing, Information and Control*, vol.11, no.3, pp.851-863, 2015.

[25] J. Ren, W. Li, Y. Wang and L. Zhou, Graph-mine: A key behavior path mining algorithm in complex software executing network, *International Journal of Innovative Computing, Information and Control*, vol.11, no.2, pp.541-553, 2015.

[26] J. Ren, H. Wang, Y. Ma, Y. Li and J. Dong, Efficient software fault localization by hierarchical instrumentation and maximal frequent subgraph mining, *International Journal of Innovative Computing, Information and Control*, vol.11, no.6, pp.1897-1911, 2015.

[27] G. Şerban, A. Câmpan and I. G. Czibula, A programming interface for finding relational association rules, *IJCCC*, pp.439-444, 2006.

[28] G. Şerban, I. G. Czibula and A. Câmpan, Medical diagnosis prediction using relational association rules, *Proc. of ICTAMI*, pp.339-352, 2008.

[29] A. Song, X. Ding, M. Li, W. Cao and K. Pu, A novel binary bat algorithm for association rules mining, *ICIC Express Letters*, vol.9, no.9, pp.2387-2394, 2015.

[30] P.-N. Tan, M. Steinbach and V. Kumar, *Introduction to Data Mining*, 1st Edition, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2005.

[31] B. Ycart, F. Ponten and J.-J. Fournie, *Statistical Data Mining for Symbol Associations in Genomic Databases*, 2013.