

## Research article

# Computation of Support and Confidence for Interval-Valued Fuzzy Association Rules

 Michal Burda<sup>\*</sup> , Viktor Pavliska, Petra Murinová

Institute for Research and Applications of Fuzzy Modeling, CE IT4Innovations, Division University of Ostrava, 30. Dubna 22, 701 03 Ostrava, Czech Republic

## ARTICLE INFO

### Article History

 Received 03 Feb 2020  
 Accepted 13 Jul 2020

### Keywords

 Interval-valued data  
 Fuzzy association rules  
 Support  
 Confidence  
 Algorithm

## ABSTRACT

The aim of this paper is to provide an algorithm for the computation of support and confidence of the association rules on interval-valued fuzzy sets. Each element of the interval-valued fuzzy set has a membership degree defined as an interval. In other words, the membership intervals may be interpreted as partial knowledge when the precise value is not known. The computations of the support and the confidence are discussed with respect to the three most common triangular norms (minimum, product and Łukasiewicz), which act as conjunction in support and confidence definitions.

© 2020 The Authors. Published by Atlantis Press B.V.

 This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

## 1. INTRODUCTION

The main objective of this paper is to continue the development of generalized association rules with missing values, which are a tool for the explanatory analysis of large data sets. In [1], we introduced a generalized algorithm for the association analysis of natural data with undefined values.

Association rules were first introduced by Hájek *et al.* in the late 1960s [2] through their formulation of the General Unary Hypotheses Automaton (GUHA) method [3] for processing binary or categorical data. Agrawal [4] reinvented this approach independently in 1993. An analysis of natural data based on linguistic summaries was developed by Yager in 1982 [5] and later by Kacprzyk [6] and, in many aspects, is similar to association analysis. An extension of association rules for fuzzy data was also developed in 2009 by Ralbovský [7]. An interpretation of data using expressions of natural language, namely using the generalized intermediate quantifiers “Almost all, Most, Many,” was proposed in [8].

In real-world applications, it is very common for data being analyzed to sometimes be missing. There may be many reasons for data unavailability. Some reasons are discussed in [9]. If an analytical method cannot work with missing data, records with missing values are often removed. Another common practice is to *impute* some reasonable values in the empty cells, if the missingness is not dependent on the unobserved or missing value itself [10]. Imputation is performed by substituting the missing values with averages or by applying a more sophisticated approach, such as machine learning techniques [11].

On the other hand, some analytical methods allow for the direct processing of missing values. For instance, an extension of the GUHA method allowing the search for association rules on binary and categorical data with missing values was introduced in [3]. The popular C4.5 algorithm [12] for decision trees can also manage missing values via distribution-based imputation.

The problem of missing values has also been studied from the mathematical logic’s point of view. Fundamental grounds were laid by Kleene, Bochvar, Sobociński and others by establishing three-valued logics, which were also studied by Łukasiewicz [13] in 1920. An overview can be found in [14]. Generalized approaches were introduced in [15], [16] and [17]. A three-valued logic modeling *borderline* or *unknown* values is presented in [18]. A deeper interpretation of the third-truth value (*possible, undefined, half-true, inconsistent*) can be found in [19].

The missingness of a value is a special case of partial knowledge. Other values may be measured imprecisely, which can be described by an interval of possible values. From that perspective, a completely missing value can be characterized with an interval that is identical to the whole domain of the attribute.

Recently, computations with intervals have become popular. For instance, Grzegorzewski [20] introduces the measures of dispersion for interval data, Billard and Diday [21] focuses on regression analysis for interval-valued data and Lima Neto and De Carvalho [22] develops a nonlinear interval-valued regression.

The aim of this paper is to generalize the concept of association rules for interval-valued fuzzy sets as introduced in [23]. The fuzzy set generalizes a classical mathematical set by allowing a partial

<sup>\*</sup>Corresponding author. Email: [michal.burda@osu.cz](mailto:michal.burda@osu.cz)

membership of an element. Each element of the universe is assigned a degree of membership to a given fuzzy set. Interval-valued fuzzy sets further generalize this concept by using intervals of membership degrees. These intervals may be interpreted as partial information where the exact membership degree is not known. Interval-valued fuzzy sets were studied in [24] and [25].

This paper enables the search for association rules on fuzzy data that have missing information. As the fuzziness captures the degree of an object possessing some property, the lack of information in data or imprecision in membership degree assignment may be modeled using the intervals of these membership degrees.

Concretely, this paper focuses on the algorithm for fast computing of the most common quality measures of an association rule: the *support* and *confidence*. The support measures the number of records in the database that satisfy both antecedent  $A$  and consequent  $C$  of the association rule  $A \rightarrow C$ . The confidence is a measure defined as a generalization of conditional probability: it is the fraction of records satisfying consequent  $C$  in a set of records that fulfill antecedent  $A$ . The “and” connective is in a fuzzy association analysis modeled with a triangular norm (t-norm). The presented algorithms are designed for the three most commonly used t-norms: Gödel (minimum), Goguen (product) and Łukasiewicz t-norm.

**Example 1.** Consider a dataset with concentrations of potassium and water pH measured in order to prove their influence on the grow of water plants.  $U$  is the set of experiments conducted. The initial three numeric columns (pH, potassium, grow rate) can be transformed into several fuzzy sets such as “acidic pH,” “neutral pH,” “basic pH,” “low potassium,” “medium potassium,” “high potassium,” etc. From that, the association rules searching algorithm may find the rule “acidic pH”  $\otimes$  “medium potassium”  $\rightarrow$  “high grow rate”.

If the original numeric values were all measured precisely, then also membership degrees to the fuzzy sets would be precise and we obtain a precise confidence of that rule. On the other hand, if some measures were missing or if the measuring tool was imprecise, we might be given intervals of possible values instead of precise numbers, which may lead into intervals of possible membership degrees and also into imprecise confidence.

The rest of the paper is organized as follows: Section 2 recalls basic definitions from interval-valued fuzzy set theory. Section 3 introduces the reader into association rules and Section 4 generalizes association rules for interval-valued fuzzy sets. Section 5 concludes the paper.

## 2. PRELIMINARIES

The main goal of this section is to introduce a mathematical background that will be used for fuzzy association rules.

### 2.1. Interval-Valued Fuzzy Sets

A fuzzy set [26] is defined as a mapping from a universe of discourse  $U$  to a real interval  $[0, 1]$ , i.e.  $F : U \rightarrow [0, 1]$ . Unlike crisp sets, where an object fully belongs or does not belong to a set, fuzzy sets enable an object  $u \in U$  to belong partially to the set  $F$  in a degree  $F(u)$ . A fuzzy set  $X$  is a subset of a fuzzy set  $Y$ ,  $X \subseteq Y$ , if  $X(u) \leq Y(u)$ , for all  $u \in U$ . A size of a fuzzy set  $X$  is  $|X| = \sum_{u \in U} X(u)$ .

Triangular norms (t-norms) are binary operations  $\otimes : [0, 1] \times [0, 1] \rightarrow [0, 1]$ , which fulfill the following axioms for each  $x, y, z \in [0, 1]$ :

- T1 commutativity:  $x \otimes y = y \otimes x$ ;
- T2 associativity:  $x \otimes (y \otimes z) = (x \otimes y) \otimes z$ ;
- T3 monotonicity:  $x \otimes y \leq x \otimes z$  whenever  $y \leq z$ ;
- T4 boundary condition:  $1 \otimes x = x$ .

T-norms have been mainly studied by Klement *et al.* [27] and later elaborated by many others. The most common examples of t-norms are

1. Gödel (minimum) t-norm:  $\min\{x, y\}$ ,
2. Goguen (product) t-norm:  $x \cdot y$ ,
3. Łukasiewicz t-norm:  $\max\{0, x + y - 1\}$ .

**Convention.** In the sequel, we will work with closed intervals  $[x_1, x_2] \subseteq [0, 1]$ . By  $\mathcal{I}$  we will denote the set of all closed intervals on  $[0, 1]$ , i.e.,  $\mathcal{I} = \{[x_1, x_2] \mid [x_1, x_2] \subseteq [0, 1]\}$ . For  $X = [x_1, x_2]$ , the lower and upper bound will be respectively denoted with  $[X]$  and  $\bar{X}$ , i.e.,  $[X] = x_1$  and  $\bar{X} = x_2$ .

An interval-valued fuzzy set  $F$  on the universe  $U$  is a mapping  $F : U \rightarrow \mathcal{I}$  assigning an interval from  $\mathcal{I}$  to any element from the universe  $U$ .

First of all, we start with a definition of the ordering  $\leq_{\mathcal{I}}$  on  $\mathcal{I}$ . Let  $X, Y \in \mathcal{I}$  then  $X \leq_{\mathcal{I}} Y$  iff  $X \vee_{\mathcal{I}} Y = Y$  (equivalently  $X \wedge_{\mathcal{I}} Y = X$ ), where

$$X \vee_{\mathcal{I}} Y = [\max\{[X], [Y]\}, \max\{\bar{X}, \bar{Y}\}],$$

$$X \wedge_{\mathcal{I}} Y = [\min\{[X], [Y]\}, \min\{\bar{X}, \bar{Y}\}].$$

We continue with definitions of basic operations on  $\mathcal{I}$ . We generalize this approach to interval fuzzy logic in definitions below which were introduced in [24,25].

A *t-norm on  $\mathcal{I}$*  is a binary operation  $\otimes_{\mathcal{I}} : \mathcal{I} \times \mathcal{I} \rightarrow \mathcal{I}$ , which satisfies the following axioms for each  $X, Y, Z \in \mathcal{I}$ :

- (T1 $_{\mathcal{I}}$ ) commutativity:  $X \otimes_{\mathcal{I}} Y = Y \otimes_{\mathcal{I}} X$ ;
- (T2 $_{\mathcal{I}}$ ) associativity:  $X \otimes_{\mathcal{I}} (Y \otimes_{\mathcal{I}} Z) = (X \otimes_{\mathcal{I}} Y) \otimes_{\mathcal{I}} Z$ ;
- (T3 $_{\mathcal{I}}$ ) monotonicity:  $X \otimes_{\mathcal{I}} Y \leq_{\mathcal{I}} X \otimes_{\mathcal{I}} Z$  whenever  $Y \leq_{\mathcal{I}} Z$ ;
- (T4 $_{\mathcal{I}}$ ) boundary condition:  $[1, 1] \otimes_{\mathcal{I}} X = X$ .

**Theorem 1. (Representability [28])** Let  $\otimes : [0, 1] \times [0, 1] \rightarrow [0, 1]$  be a t-norm. For  $X, Y \in \mathcal{I}$ , we put

$$X \otimes_{\mathcal{I}} Y = [[X] \otimes [Y], [X] \otimes [\bar{Y}]].$$

Then  $\otimes_{\mathcal{I}}$  is a t-norm on  $\mathcal{I}$ .

**Proof.** Let  $\otimes$  be a t-norm, hence  $\otimes$  is commutative. Then obviously

$$X \otimes_{\mathcal{I}} Y = [[X] \otimes [Y], [X] \otimes [\bar{Y}]] =$$

$$= [[Y] \otimes [X], [Y] \otimes [\bar{X}]] = Y \otimes_{\mathcal{I}} X,$$

so we see that  $\otimes_I$  is commutative too. We can prove the associativity (T 2<sub>I</sub>) and the boundary condition (T 4<sub>I</sub>) in the similar way.

To prove the monotonicity (T 3<sub>I</sub>), let us consider arbitrary  $Y, Z \in \mathcal{I}$  such that  $Y \leq_I Z$ . By definition,  $Y \vee_I Z = Z$ , which means that  $\max(|Y|, |Z|) = |Z|$  and  $\max(|Y|, |Z|) = |Z|$ , i.e.,  $|Y| \leq |Z|$  and  $|Y| \leq |Z|$ . Since  $\otimes$  is monotonous, we can see that for any  $X \in \mathcal{I}$ ,  $|X \otimes Y| \leq |X \otimes Z|$  and  $|X \otimes Y| \leq |X \otimes Z|$ , which is equivalent to  $X \otimes_I Y \vee_I X \otimes_I Z = X \otimes_I Z$ , which means that  $X \otimes_I Y \leq_I X \otimes_I Z$ , from which we can see the monotonicity of  $\otimes_I$ .

**Convention.** Whenever we will need to perform an operation between an element  $a \in [0, 1]$  and an interval  $X \in \mathcal{I}$  (e.g.,  $a \otimes_I X$ ) then we can promote the element  $a$  to the interval  $[a, a]$  so we can use the operation defined for intervals.

## 2.2. (Noninterval) Fuzzy Association Rules

Let  $\mathcal{O} = \{o_1, o_2, \dots, o_n\}$ ,  $n > 0$ , be a finite set of abstract elements called objects and  $\mathcal{A} = \{a_1, a_2, \dots, a_m\}$ ,  $m > 0$ , be a finite set of attributes. Within the association rules framework, a dataset  $D$  is a mapping that assigns to each object  $o \in \mathcal{O}$  and attribute  $a \in \mathcal{A}$  a truth degree  $D(a, o) \in [0, 1]$ , which represents the intensity of assignment of attribute  $a$  to object  $o$ . We also define

$$D(X, o) = \bigotimes_{a \in X} D(a, o) \quad (1)$$

for each subset  $X \subseteq \mathcal{A}$  and a fixed t-norm  $\otimes$ .

*Association rule* is a formula  $A \rightarrow C$ , where  $A \subseteq \mathcal{A}$  is an *antecedent* and  $C \subseteq \mathcal{A}$  is a *consequent*. It is natural to assume  $A \cap C = \emptyset$  and also  $|C| = 1$ .

As each combination of attributes in antecedent and consequent form a well-formed association rules, an important problem is to identify rules that are relevant to the given dataset  $D$ . So far, there exist a large number of measures of rule relevance. An overview can be found in [29]. Perhaps the most essential association rule quality measures are *support* and *confidence*, which may be defined as follows:

Let  $A \rightarrow C$  be an association rule and  $D$  be a dataset. Then

$$\begin{aligned} \text{supp}(A \rightarrow C) &= \sum_{o \in \mathcal{O}} D(A, o) \otimes D(C, o), \\ \text{conf}(A \rightarrow C) &= \frac{\sum_{o \in \mathcal{O}} D(A, o) \otimes D(C, o)}{\sum_{o \in \mathcal{O}} D(A, o)}. \end{aligned} \quad (2)$$

**Example 2.** Consider a dataset  $D$  in Table 1. Here  $\mathcal{O} = \{o_1, \dots, o_5\}$  are table rows and  $\mathcal{A} = \{a_1, \dots, a_3\}$  are table columns. The numeric values in the table represent membership degrees  $D(a, o)$ , for  $a \in \mathcal{A}$  and  $o \in \mathcal{O}$ . Let  $A = \{a_1, a_2\}$  and let  $\otimes$  be a Gödel (minimum) t-norm. Then  $D(A, o_1) = \min\{0.9, 1.0\} = 0.9$ ,  $D(A, o_2) = 0$ ,  $D(A, o_3) = 0.8$  etc. Let  $C = \{a_3\}$ . The association rule  $A \rightarrow C$ , i.e.,  $\{a_1, a_2\} \rightarrow \{a_3\}$ , has the following support and confidence:

$$\begin{aligned} \text{supp}(A \rightarrow C) &= 0.9 + 0 + 0.8 + 0 + 0 = 1.7, \\ \text{conf}(A \rightarrow C) &= 1.7 / (0.9 + 0 + 0.8 + 0.1 + 0) \doteq 0.94. \end{aligned}$$

The mining process of the association rules starts with the generation of all association rules with antecedents of length one. Support measures are computed for all of the generated rules, and the

**Table 1** | An exemplary noninterval fuzzy dataset  $D$ . Objects from  $\mathcal{O}$  are represented by table rows, attributes from  $\mathcal{A}$  with columns.

	$a_1$	$a_2$	$a_3$
$o_1$	0.9	1.0	0.9
$o_2$	0.7	0.0	0.1
$o_3$	0.9	0.8	1.0
$o_4$	0.1	0.9	0.0
$o_5$	0.0	0.0	0.2

generated rules, which exceed the user-defined minimum support threshold, serve as a basis for the longer rules. The process is based on the fact that the support is nonincreasing with respect to the extension of antecedents by the new attributes. The monotonicity of the support is often the basis for efficient association rules in mining algorithms. Rules that satisfy the minimum support condition are then subjected to a confidence evaluation. A rule with both support and confidence above the user-defined thresholds is considered interesting and it is included in the results. Details of the algorithm may be found in [30].

## 3. ASSOCIATION RULES ON INTERVAL-VALUED FUZZY SETS

In order to extend the association rules framework for data containing *interval-valued* membership degrees, one has to switch the range of membership degrees from  $[0, 1]$  to  $\mathcal{I}$ . In other words, dataset  $D$  becomes a mapping such that  $D(a, o) \in \mathcal{I}$  for each  $a \in \mathcal{A}$  and each  $o \in \mathcal{O}$ . Similarly to (1), we define

$$D(X, o) = D(x_1, o) \otimes_I D(x_2, o) \otimes_I \dots \otimes_I D(x_k, o),$$

for  $X = \{x_1, x_2, \dots, x_k\} \subseteq \mathcal{A}$  and a fixed t-norm  $\otimes_I$  on  $\mathcal{I}$ .

For interval-valued membership degrees, the support and confidence of an association rule cannot be computed directly as in (2). Note that the interval membership degrees represent a partial knowledge—the exact membership degree is only known to be from the given interval. Therefore, the rule's support and confidence must also be represented as an interval. Although the evaluation of support is quite straightforward, the computation of the lower and upper bounds for confidence is an optimization problem that involves finding the exact values of the membership degrees such that the resulting confidence is minimal or maximal. To formalize this problem mathematically, we introduce the notion of a *possible world* as follows.

**Definition 1.** Let  $D$  be a dataset that assigns an interval of truth values to each pair  $\langle a, o \rangle$  of attribute  $a \in \mathcal{A}$  and object  $o \in \mathcal{O}$ , i.e.,  $D(a, o) \in \mathcal{I}$ . A *possible world* of the dataset  $D$  is a mapping  $w_D : \mathcal{A} \times \mathcal{O} \rightarrow [0, 1]$  such that  $w(a, o) \in D(a, o)$  for each  $a \in \mathcal{A}$ ,  $o \in \mathcal{O}$ . A class of all such mappings  $w_D$  will be denoted by  $W_D$ .

**Convention.** For the sake of simplicity, we omit the subscript  $D$  in  $w_D$  and  $W_D$ , if  $D$  is evident from the context. Moreover, we put  $w(A, o) = \bigotimes_{a \in A} w(a, o)$ , for any  $A \subseteq \mathcal{A}$ .

**Example 3.** Table 2 shows a dataset  $D$  with interval-valued fuzzy membership degrees, from which we can see, e.g.,  $D(a_1, o_3) =$

**Table 2** | An exemplary interval fuzzy dataset  $D$ . Objects from  $\mathcal{O}$  are represented by table rows, attributes from  $\mathcal{A}$  with columns.

	$a_1$	$a_2$	$a_3$
$o_1$	[0.9, 1.0]	[1.0, 1.0]	[0.8, 0.9]
$o_2$	[0.0, 0.7]	[0.0, 0.5]	[0.1, 0.1]
$o_3$	[0.5, 0.9]	[0.8, 1.0]	[1.0, 1.0]
$o_4$	[0.0, 0.1]	[0.9, 0.9]	[0.0, 0.0]
$o_5$	[0.0, 0.0]	[0.0, 0.5]	[0.2, 1.0]

[0.5, 0.9]. If we put  $A = \{a_1, a_2\}$  and we let  $\otimes$  to be Goguen (product) t-norm, we obtain  $D(A, o_3) = [0.5 \cdot 0.8, 0.9 \cdot 1.0] = [0.4, 0.9]$ , etc. As one can see, Table 1 may be viewed as a possible world of the interval-valued dataset from Table 2. Another possible world is, e.g., such  $w$  that  $w(a_1, o_1) = 0.95$ ,  $w(a_2, o_1) = 1.0$ ,  $w(a_3, o_1) = 0.88$ ,  $w(a_1, o_2) = 0.0$ , etc.

In the following subsections, we generalize the quality measures of an association rule—the *support* and *confidence*—for interval-valued data. We also propose efficient algorithms for their computation and prove their correctness.

### 3.1. The Support

In this section, a generalization of an association rule’s *support* for interval-valued data is provided. A subsequent theorem gives a hint for a rather straightforward way of computation.

**Definition 2.** Let  $D$  be a dataset, let us have a rule  $A \rightarrow C$  and let  $\otimes$  be a t-norm. Then the support of  $A \rightarrow C$  on a possible world  $w \in W$  is

$$\text{supp}_{\otimes}^{A \rightarrow C}(w) = \sum_{o \in \mathcal{O}} w(A, o) \otimes w(C, o).$$

The interval-valued support of  $A \rightarrow C$  in the dataset  $D$  is

$$\text{supp}_{\otimes}(A \rightarrow C) = [\text{supp}_1, \text{supp}_2],$$

with such maximal  $\text{supp}_1$  and minimal  $\text{supp}_2$  that

$$\forall w \in W : \text{supp}_1 \leq \text{supp}_{\otimes}^{A \rightarrow C}(w) \leq \text{supp}_2.$$

For simplicity, we omit the superscript and write only  $\text{supp}_{\otimes}(w)$  because the rule  $A \rightarrow C$  is fixed throughout the paper. Obviously,  $\text{supp}_{\otimes}(w) \in [0, |\mathcal{O}|]$ .

As stated in the subsequent theorem, finding the boundaries  $\text{supp}_1$  and  $\text{supp}_2$  is quite straightforward.

**Theorem 2.** Let us have a rule  $A \rightarrow C$  and let  $D$  be a (interval-valued) dataset. Let  $w_1, w_2 \in W$  such that for each  $o \in \mathcal{O}$ ,

$$\begin{aligned} w_1(A, o) &= [D(A, o)], \quad w_1(C, o) = [D(C, o)], \\ w_2(A, o) &= [D(A, o)], \quad w_2(C, o) = [D(C, o)]. \end{aligned}$$

Then for any t-norm  $\otimes$ ,  $\text{supp}_{\otimes}(A \rightarrow C) = [\text{supp}_1, \text{supp}_2]$ , where  $\text{supp}_1 = \text{supp}_{\otimes}(w_1)$  and  $\text{supp}_2 = \text{supp}_{\otimes}(w_2)$ .

**Proof.** The proof evidently follows from the monotonicity of the t-norm  $\otimes$ .

**Example 4.** Let  $\otimes$  be Gödel min t-norm. Accordingly to Theorem 2, the support of rule  $\{a_1, a_2\} \rightarrow \{a_3\}$  in Table 2 can be obtained by evaluating the support of two possible worlds:  $w_1$  of lower bounds and  $w_2$  of upper bounds.

$$\begin{aligned} \text{supp}_1 &= \min\{0.9, 1, 0.8\} + \min\{0, 0, 0.1\} + \dots + \\ &+ \min\{0, 0, 0.2\} = 1.3, \end{aligned}$$

$$\begin{aligned} \text{supp}_2 &= \min\{1, 1, 0.9\} + \min\{0.7, 0.5, 0.1\} + \dots + \\ &+ \min\{0, 0.5, 1\} = 1.9. \end{aligned}$$

**Theorem 3.** Let  $A \rightarrow C$  be a rule,  $A, C \subseteq \mathcal{A}$ , let  $a \in \mathcal{A}$  be an attribute,  $a \notin A \cup C$ , and let  $D$  be a dataset. Then for arbitrary t-norm  $\otimes$ ,

$$\text{supp}_{\otimes}(A \rightarrow C) \geq_I \text{supp}_{\otimes}((A \cup \{a\}) \rightarrow C). \quad (3)$$

**Proof.** Note that accordingly to Theorem 2,

$$\begin{aligned} \lfloor \text{supp}_{\otimes}(A \rightarrow C) \rfloor &= \sum_{o \in \mathcal{O}} \lfloor D(A, o) \rfloor \otimes \lfloor D(C, o) \rfloor, \\ \lfloor \text{supp}_{\otimes}((A \cup \{a\}) \rightarrow C) \rfloor &= \sum_{o \in \mathcal{O}} \lfloor D(A, o) \rfloor \otimes \\ &\otimes \lfloor D(a, o) \rfloor \otimes \lfloor D(C, o) \rfloor. \end{aligned}$$

(Analogous equations hold for upper bounds of the support.) Since  $\otimes$  is monotone, we can immediately see that inequality (3) holds.

Theorem 3 says that with extension of rule’s antecedent by a new attribute, the rule’s support is not increasing. As discussed in Section 2.2, such property of the support is important for association rules searching algorithms. Theorem 3 allows us to use, e.g., the Agrawal’s *Apriori algorithm* [30] to search also for association rules on interval-valued data by constraining the support with user-defined *minSupp* threshold, e.g., as follows:

- $\lfloor \text{supp}_{\otimes}(A \rightarrow C) \rfloor \geq \text{minSupp}$ ;
- $\lceil \text{supp}_{\otimes}(A \rightarrow C) \rceil \geq \text{minSupp}$ ; or
- $\frac{1}{2} \left( \lfloor \text{supp}_{\otimes}(A \rightarrow C) \rfloor + \lceil \text{supp}_{\otimes}(A \rightarrow C) \rceil \right) \geq \text{minSupp}$ .

### 3.2. The Confidence

The objective of this section is to provide an algorithm computing the association rule’s confidence in the interval-valued data. The problem is analyzed from the perspective of *possible worlds* (see Definition 1).

First, the setting of  $w(C, o) \in D(C, o)$  that yields the minimum or maximum confidence is discussed. After that, the setting of  $w(A, o) \in D(A, o)$  is analyzed separately with respect to the Gödel (minimum), Goguen (product) and Łukasiewicz t-norms, and the potential solutions from the interval  $D(A, o)$  are identified. It is shown, that for each  $o \in \mathcal{O}$ , at most two  $w(A, o)$  *suspicious points* exist that need to be examined in the search for the extreme confidence. (The combination of those points is called *the suspicious world*). After that, an efficient algorithm (Algorithm 1) is presented that solves the optimization sub-problem, which is later used to find the extreme confidence (minimum or maximum) based on the suspicious points (Algorithms 2, 3 and 4).

**Algorithm 1:** An algorithm for searching for  $F(I_{min})$ , as introduced in Lemma 11. The input arguments  $G, H$  are assumed to be real numbers,  $H \neq 0$  and input arguments  $\mathbf{g}$  and  $\mathbf{h}$  are assumed to be vectors of length  $n$ ,  $\mathbf{g}=(g_1, \dots, g_n)$ ,  $\mathbf{h}=(h_1, \dots, h_n)$ ,  $h_i \neq 0$ .

```

function (FracMinim) ( $G, H, \mathbf{g}, \mathbf{h}$ )
   $heap \leftarrow$  new empty min-heap
  for all  $i \in \{1, \dots, n\}$  do
    push  $\frac{g_i}{h_i}$  into  $heap$ 
  end for
  while  $heap$  is not empty do
    pop a minimal element  $\frac{a}{b}$  from  $heap$ 
    if  $\frac{a}{b} < \frac{G}{H}$  then
       $G \leftarrow G + a$ 
       $H \leftarrow H + b$ 
    else
      break the while loop
    end if
  end while
  return  $\frac{G}{H}$ 
end function

```

**Algorithm 2:** The computation of  $\text{conf}_{\odot}(A \rightarrow C)$  on the dataset  $D$

```

function (Confidence)  $\odot(D, A, C)$ 
   $G, G', H \leftarrow 0$ 
   $\mathbf{g}, \mathbf{g}', \mathbf{h} \leftarrow$  new empty vector
  for all  $o \in \mathcal{O}$  do
     $al \leftarrow [D(A, o)]$ ,  $ah \leftarrow [D(A, o)]$ 
     $cl \leftarrow [D(C, o)]$ ,  $ch \leftarrow [D(C, o)]$ 
     $G \leftarrow G + al \cdot cl$ 
     $G' \leftarrow G' + al \cdot ch$ 
     $H \leftarrow H + al$ 
    if  $al \neq ah$  then
      append  $(ah \cdot cl - al \cdot cl)$  to  $\mathbf{g}$ 
      append  $(ah \cdot ch - al \cdot ch)$  to  $\mathbf{g}'$ 
      append  $(ah - al)$  to  $\mathbf{h}$ 
    end if
  end for
   $lo \leftarrow \text{FRACMINIM}(G, H, \mathbf{g}, \mathbf{h})$ 
   $hi \leftarrow -\text{FRACMINIM}(-G', H, -\mathbf{g}', \mathbf{h})$ 
  return  $[lo, hi]$ 
end function

```

**Definition 3.** Let  $D$  be a dataset, let us have a rule  $A \rightarrow C$  and let  $w \in W$  such that  $w(A, o) \neq 0$  for some  $o \in \mathcal{O}$ . Let  $\otimes$  be a  $t$ -norm. Then the confidence of  $A \rightarrow C$  in a possible world  $w$  is

$$\text{conf}_{\otimes}^{A \rightarrow C}(w) = \frac{\sum_{o \in \mathcal{O}} w(A, o) \otimes w(C, o)}{\sum_{o \in \mathcal{O}} w(A, o)}. \quad (4)$$

The interval-valued confidence of  $A \rightarrow C$  in the dataset  $D$  is

$$\text{conf}_{\otimes}(A \rightarrow C) = [\text{conf}_1, \text{conf}_2],$$

with such maximal  $\text{conf}_1$  and minimal  $\text{conf}_2$  that

$$\forall w \in W : \text{conf}_1 \leq \text{conf}_{\otimes}^{A \rightarrow C}(w) \leq \text{conf}_2.$$

Obviously,  $\text{conf}_{\otimes}^{A \rightarrow C}(w) \in [0, 1]$ . Again, we omit  $A \rightarrow C$  in the superscript where the rule  $A \rightarrow C$  is obvious from the context.

**Algorithm 3:** Computation of  $\text{conf}_{\min}(A \rightarrow C)$  on the dataset  $D$

```

function (Confidence)  $\min(D, A, C)$ 
   $G, H, M, N \leftarrow 0$ 
   $\mathbf{g}, \mathbf{h} \leftarrow$  new empty vector
  for all  $o \in \mathcal{O}$  do
     $al \leftarrow [D(A, o)]$ ,  $ah \leftarrow [D(A, o)]$ 
     $cl \leftarrow [D(C, o)]$ ,  $ch \leftarrow [D(C, o)]$ 
    if  $ah \leq cl$  then
       $G \leftarrow G + al$ 
       $H \leftarrow H + al$ 
       $\triangleright$  because  $\min(al, cl) = al$ 
    else if  $al \geq cl$  then
       $G \leftarrow G + cl$ 
       $H \leftarrow H + ah$ 
       $\triangleright$  because  $\min(ah, cl) = cl$ 
    else
       $G \leftarrow G + al$ 
       $H \leftarrow H + al$ 
       $\triangleright$  because  $\min(al, cl) = al$ 
      if  $al \neq ah$  then
        append  $(cl - al)$  to  $\mathbf{g}$ 
        append  $(ah - al)$  to  $\mathbf{h}$ 
      end if
    end if
     $a \leftarrow \text{closest}([al, ah], ch)$ 
     $M \leftarrow M + \min(a, ch)$ 
     $N \leftarrow N + a$ 
  end for
   $lo \leftarrow \text{FRACMINIM}(G, H, \mathbf{g}, \mathbf{h})$ 
   $hi \leftarrow \frac{M}{N}$ 
  return  $[lo, hi]$ 
end function

```

**Example 5.** Let  $D$  be a dataset as in Table 2 and  $\otimes$  be Goguen product  $t$ -norm. Let  $A = \{a_1, a_2\}$  and  $C = \{a_3\}$ . Let us define three possible worlds,  $w_1, w_2, w_3$  by taking the lower bounds for  $C$  and lower bound for  $A$  ( $w_1$ ), upper bound for  $A$  ( $w_2$ ) and purposely mixed bounds for  $A$  ( $w_3$ ). Please see Table 3 and compare it with Table 2. However, neither a selection of lower bounds as in  $w_1$  nor upper bounds as in  $w_2$  yield in a minimum confidence of a rule  $A \rightarrow C$  possible to obtain from  $D$ . The minimum confidence is reached for the world  $w_3$ :  $\text{conf}_{\otimes}^{A \rightarrow C}(w_1) \doteq 0.86$ ,  $\text{conf}_{\otimes}^{A \rightarrow C}(w_2) \doteq 0.74$ ,  $\text{conf}_{\otimes}^{A \rightarrow C}(w_3) \doteq 0.67$ .

As can be seen from the example, finding boundaries  $\text{conf}_1$  and  $\text{conf}_2$  of the confidence of the association rule  $A \rightarrow C$  is not as straightforward as for the support. It is an optimization problem, which is equivalent to finding global extremes (minima and maxima) of function (4), i.e., to finding such worlds that yield minimal (resp. maximal) possible confidence.

Worlds  $w \in W$  with optimal values of  $w(C, o)$  can be obtained accordingly to the following lemma.

**Lemma 4.** Let  $A \rightarrow C$  be a rule,  $D$  be a dataset,  $W_1, W_2 \subseteq W$  such that

$$\begin{aligned} W_1 &= \{w \mid w(C, o) = [D(C, o)], \forall o \in \mathcal{O}\}, \\ W_2 &= \{w \mid w(C, o) = [D(C, o)], \forall o \in \mathcal{O}\} \end{aligned}$$

**Table 3** | Three possible worlds for dataset  $D$  in Table 2.  $w_1$  takes lower bounds for  $A$ ,  $w_2$  takes upper bounds for  $A$ , and  $w_3$  is mixed so that it yields the minimum confidence of rule  $A \rightarrow C$  that can be obtained from  $D$ .

	$w_1$		$w_2$		$w_3$	
	$A$	$C$	$A$	$C$	$A$	$C$
$o_1$	0.9	0.8	1.0	0.8	1.0	0.8
$o_2$	0.0	0.1	0.35	0.1	0.35	0.1
$o_3$	0.4	1.0	0.9	1.0	0.4	1.0
$o_4$	0.0	0.0	0.09	0.0	0.09	0.0
$o_5$	0.0	0.2	0.0	0.2	0.0	0.2

and let

$$\text{conf}_{\otimes}(A \rightarrow C) = [\text{conf}_1, \text{conf}_2].$$

Then for any t-norm  $\otimes$ , there exist  $w_1 \in W_1$  and  $w_2 \in W_2$  such that  $\text{conf}_1 = \text{conf}_{\otimes}(w_1)$  and  $\text{conf}_2 = \text{conf}_{\otimes}(w_2)$ .

**Proof.** The proof evidently follows from the monotonicity of the t-norm.

Lemma 4 gives us a straightforward advice of how to select values of  $w(C, o)$  for each  $o \in \mathcal{O}$  in order to minimize (resp. maximize) the confidence  $\text{conf}_{\otimes}(w)$ . Now, let us sketch the way of finding the values of  $w(A, o)$  (for all  $o \in \mathcal{O}$ ) such that the confidence becomes extreme.

Searching for a possible world  $w \in W$  with globally maximal (resp. minimal) confidence  $\text{conf}_{\otimes}(w)$  is equivalent to searching for global extremes of a function  $f(x_1, \dots, x_n)$  with  $n = |\mathcal{O}|$  parameters, where  $x_i$  corresponds to  $w(A, o_i)$ , for  $o_i \in \mathcal{O}$  and for all  $i \in \{1, \dots, n\}$ . Formally, function  $f$  is a mapping

$$f : D(A, o_1) \times \dots \times D(A, o_n) \rightarrow [0, 1],$$

defined as

$$f(x_1, \dots, x_n) = \frac{\sum_{i=1}^n x_i \otimes w(C, o_i)}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n t_i(x_i)}{\sum_{i=1}^n x_i}, \quad (5)$$

where for each  $i \in \{1, \dots, n\}$ ,  $t_i : [0, 1] \rightarrow [0, 1]$  is a function defined as

$$t_i(x) = x \otimes w(C, o_i),$$

since  $w(C, o_i)$  are fixed.

The extremes of function  $f$  may be found:

- at boundaries of intervals  $D(A, o_i)$ , or
- at critical points, where partial derivatives of  $f$  with respect to  $x_i$  are equal to zero or do not exist.

**Lemma 5.** Let function  $f$  be defined for  $\sum_{i=1}^n x_i \neq 0$  as

$$f(x_1, \dots, x_n) = \frac{\sum_{i=1}^n t_i(x_i)}{\sum_{i=1}^n x_i},$$

where  $t_i : [0, 1] \rightarrow [0, 1]$  is for each  $i \in \{1, \dots, n\}$  a differentiable function. Then for  $j \in \{1, \dots, n\}$ ,

$$\frac{\partial f(x_1, \dots, x_n)}{\partial x_j} = 0 \quad \text{if} \quad \frac{dt_j}{dx_j}(x_j) = f(x_1, \dots, x_n). \quad (6)$$

The function  $f(x_1, \dots, x_n)$  is increasing with respect to  $x_j$  if

$$\frac{dt_j}{dx_j}(x_j) > f(x_1, \dots, x_n) \quad (7)$$

and decreasing with respect to  $x_j$  if

$$\frac{dt_j}{dx_j}(x_j) < f(x_1, \dots, x_n). \quad (8)$$

**Proof.** Let  $\frac{\partial f(x_1, \dots, x_n)}{\partial x_j} = 0$ , i.e.,

$$\left( \frac{\partial \sum_{i=1}^n t_i(x_i)}{\partial x_j} \left( \sum_{i=1}^n x_i \right) - \left( \sum_{i=1}^n t_i(x_i) \right) \right) \left( \sum_{i=1}^n x_i \right)^{-2} = 0,$$

which yields in

$$\frac{dt_j}{dx_j}(x_j) = \frac{\sum_{i=1}^n t_i(x_i)}{\sum_{i=1}^n x_i} = f(x_1, \dots, x_n).$$

The inequalities (7) and (8) easily follow from the fact that a function  $f$  is increasing with respect to  $x_j$  if  $\frac{\partial f(x_1, \dots, x_n)}{\partial x_j} > 0$  and decreasing with respect to  $x_j$  if  $\frac{\partial f(x_1, \dots, x_n)}{\partial x_j} < 0$ .

Lemma 5 provides a general framework for finding the points suspicious of extrema for arbitrary t-norm  $\otimes$ . Below we examine suspicious points for the most commonly used t-norms: Goguen product ( $\odot$ ), Gödel minimum (min) and Łukasiewicz ( $\otimes_L$ ).

**Lemma 6. (Suspicious worlds related to product t-norm)** Let  $A \rightarrow C$  be a rule and  $D$  be a dataset. Let

$$P = \{w \in W \mid \forall o \in \mathcal{O} : w(A, o) \in \{[D(A, o)], [D(A, o)]\}\}.$$

Then there exist  $w_1, w_2 \in P$  such that

$$\begin{aligned} [\text{conf}_{\odot}(A \rightarrow C)] &= \text{conf}_{\odot}(w_1) \quad \text{and} \\ [\text{conf}_{\odot}(A \rightarrow C)] &= \text{conf}_{\odot}(w_2). \end{aligned}$$

**Proof.** For  $w \in W$ ,  $\text{conf}_{\odot}(w)$  can be viewed as a function  $f$  in (5). If we put  $t_i(x_i) = x_i \odot w(C, o_i)$ , where  $w(C, o_i)$  is fixed accordingly to Lemma 4, we obtain from equation (6) of Lemma 5 the following equality:

$$\begin{aligned} \frac{dt_j}{dx_j}(x_j) &= w(C, o_j) \\ &= \frac{\sum_{i=1}^n x_i \odot w(C, o_i)}{\sum_{i=1}^n x_i}, \quad \forall j \in \{1, \dots, |\mathcal{O}|\}. \end{aligned}$$

Since  $\odot$  is simply the product, the equation above can be further simplified to

$$\begin{aligned}
 w(C, o_j) \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i \cdot w(C, o_i), & (9) \\
 w(C, o_j)x_j + w(C, o_j) \sum_{\substack{i=1 \\ i \neq j}}^n x_i &= x_j \cdot w(C, o_j) \\
 &+ \sum_{\substack{i=1 \\ i \neq j}}^n x_i \cdot w(C, o_i) \\
 w(C, o_j) &= \frac{\sum_{\substack{i=1 \\ i \neq j}}^n x_i \cdot w(C, o_i)}{\sum_{\substack{i=1 \\ i \neq j}}^n x_i}.
 \end{aligned}$$

So, the partial derivative  $\frac{\partial f}{\partial x_j}$  is equal to zero for all  $x_j \in D(A, o_j)$  if (9) holds. In other words, the local extreme either does not exist (if (9) is not fulfilled) or the extreme is at any point  $x_j \in D(A, o_j)$  (if (9) holds). It means that the function is monotonic with respect to  $x_j$ . In each case, the global extreme of  $\text{conf}(w)$  is thus reached if  $w(A, o_j)$  is equal to either lower or upper bound of the interval  $D(A, o_j)$ , i.e.,  $\lfloor D(A, o_j) \rfloor$  or  $\lceil D(A, o_j) \rceil$ .

**Example 6.** Consider data from Example 3 and Table 2. Naturally, there exists an infinite number of possible worlds that are represented by that interval-valued data. However, if we are searching for worlds with minimum or maximum confidence of the rule  $A \rightarrow C$ ,  $A = \{a_1, a_2\}$ ,  $C = \{a_3\}$ , and product t-norm is used for computations, there exists, accordingly to Lemmas 4 and 6, a finite set  $P$  of possible worlds, which contains both extreme worlds. Any combination of values of suspicious points from Table 4 is a possible world from the set  $P$ . As can be seen,  $|P| = 64$ .

**Definition 4. (Closest element)** A function  $\text{closest} : I \times [0, 1] \rightarrow [0, 1]$  is defined as

$$\text{closest}(X, y) = \begin{cases} \lfloor X \rfloor, & \text{if } \lfloor X \rfloor < y, \\ \lfloor X \rfloor, & \text{if } \lfloor X \rfloor > y, \\ y, & \text{otherwise,} \end{cases}$$

for  $X \in I$  and  $y \in [0, 1]$ .

In other words,  $\text{closest}(X, y)$  finds an element from interval  $X$  that is the closest to  $y$ .

**Table 4** | Sets of suspicious points for rule  $A \rightarrow C$  from Example 3, if the product t-norm is being used.

	Lower conf.		Upper conf.	
	A	C	A	C
$o_1$	{0.9, 1.0}	{0.8}	{0.9, 1.0}	{0.9}
$o_2$	{0.0, 0.35}	{0.1}	{0.0, 0.35}	{0.1}
$o_3$	{0.4, 0.9}	{1.0}	{0.4, 0.9}	{1.0}
$o_4$	{0.0, 0.09}	{0.0}	{0.0, 0.09}	{0.0}
$o_5$	{0.0}	{0.2}	{0.0}	{1.0}

**Lemma 7. (Suspicious worlds related to minimum t-norm)** Let  $A \rightarrow C$  be a rule and  $D$  be a dataset. Let

$$M = \{w \mid \forall o \in \mathcal{O} : w(A, o) \in \alpha_o\},$$

where

$$\alpha_o = \begin{cases} \{\lfloor D(A, o) \rfloor\}, & \text{if } \lfloor D(A, o) \rfloor \leq \lfloor D(C, o) \rfloor, \\ \{\lceil D(A, o) \rceil\}, & \text{if } \lceil D(A, o) \rceil \geq \lceil D(C, o) \rceil, \\ \{\lfloor D(A, o) \rfloor, \lceil D(A, o) \rceil\}, & \text{otherwise.} \end{cases}$$

Then there exists  $w_1 \in M$  such that  $\lfloor \text{conf}_{\min}(A \rightarrow C) \rfloor = \text{conf}_{\min}(w_1)$ .

Let  $w_2 \in W$  such that

$$w_2(A, o) = \text{closest}(D(A, o), \lceil D(C, o) \rceil). \quad (10)$$

Then  $\lfloor \text{conf}_{\min}(A \rightarrow C) \rfloor = \text{conf}_{\min}(w_2)$ .

**Proof.** For confidence based on minimum t-norm, we can continue similarly as in previous Lemma 6. In this case, we put  $t_i(x_i) = \min(x_i, w(C, o_i))$ , for  $w(C, o_i)$  again fixed accordingly to Lemma 4. From (6) of Lemma 5 we obtain

$$\frac{dt_j}{dx_j}(x_j) = \frac{\sum_{i=1}^n \min(x_i, w(C, o_i))}{\sum_{i=1}^n x_i}, \quad \forall j \in \{1, \dots, |\mathcal{O}|\}.$$

The derivative of  $\min(x_j, w(C, o_j))$  does not exist for  $x_j = w(C, o_j)$ . The derivative equals 1 for  $x_j < w(C, o_j)$ , and 0 for  $x_j > w(C, o_j)$ .

We can deduce from inequalities (7) and (8) that the confidence is increasing with respect to  $x_j$  for  $x_j \leq w(C, o_j)$  and decreasing with respect to  $x_j$  for  $x_j \geq w(C, o_j)$ . From that, we can deduce the following values for  $w(A, o)$ :

- if  $\lfloor D(A, o) \rfloor \leq \lfloor D(C, o) \rfloor$ , the confidence is increasing on the whole  $D(A, o)$  interval. Therefore,  $\lfloor D(A, o) \rfloor$  is the point of global minimum;
- if  $\lceil D(A, o) \rceil \geq \lceil D(C, o) \rceil$ , the confidence is decreasing on the whole  $D(A, o)$  interval. Therefore,  $\lceil D(A, o) \rceil$  is the point of global minimum;
- otherwise, the global minimum is reached either for  $\lfloor D(A, o) \rfloor$  or  $\lceil D(A, o) \rceil$ .

Similarly,

- if  $\lfloor D(A, o) \rfloor \leq \lfloor D(C, o) \rfloor$ , the confidence is increasing on the whole  $D(A, o)$  interval. Therefore,  $\lceil D(A, o) \rceil$  is the point of global maximum;
- if  $\lceil D(A, o) \rceil \geq \lceil D(C, o) \rceil$ , the confidence is decreasing on the whole  $D(A, o)$  interval. Therefore,  $\lfloor D(A, o) \rfloor$  is the point of global maximum;
- otherwise, the global maximum is reached for  $\lfloor D(C, o) \rfloor$ .

**Example 7.** Consider data from Example 3 and Table 2 and minimum t-norm. Accordingly to Lemmas 4 and 7, there are only two possible worlds that are candidates for a lower bound of the confidence of rule  $A \rightarrow C$ . The upper bound of the confidence is given by the world  $w_2$  – see Table 5.

**Lemma 8. (Suspicious worlds related to Łukasiewicz t-norm)** Let  $A \rightarrow C$  and  $D$  be a dataset. Let  $w_1 \in W$  such that

$$w_1(A, o) = \text{closest}(D(A, o), 1 - [D(C, o)]).$$

Then  $[\text{conf}_{\otimes_L}(A \rightarrow C)] = \text{conf}_{\otimes_L}(w_1)$ . Let

$$L = \{w \mid \forall o \in \mathcal{O} : w(A, o) \in \alpha_o\},$$

where

$$\alpha_o = \begin{cases} \{[D(A, o)]\}, & \text{if } [D(A, o)] \leq 1 - [D(C, o)], \\ \{[D(A, o)]\}, & \text{if } [D(A, o)] \geq 1 - [D(C, o)], \\ \{[D(A, o)], [D(A, o)]\}, & \text{otherwise.} \end{cases}$$

Then there exists  $w_2 \in L$  such that  $[\text{conf}_{\otimes_L}(A \rightarrow C)] = \text{conf}_{\otimes_L}(w_2)$ .

**Proof.** By setting  $w(C, o_i)$  fixed accordingly to Lemma 4 and by putting  $t_i(x_j) = x_j \otimes_L w(C, o_i) = \max(0, x_j + w(C, o_i) - 1)$ , we obtain from (6) of Lemma 5, for all  $j \in \{1, \dots, |\mathcal{O}|\}$ :

$$\begin{aligned} \frac{dt_j}{dx_j}(x_j) &= \frac{d \max(0, x_j + w(C, o_j) - 1)}{dx_j} \\ &= \frac{\sum_{i=1}^n \max(0, x_i + w(C, o_i) - 1)}{\sum_{i=1}^n x_i}. \end{aligned}$$

The derivative of  $\max(0, x_j + w(C, o_j) - 1)$  does not exist for  $x_j = 1 - w(C, o_j)$ . The derivative equals 0 (for  $x_j < 1 - w(C, o_j)$ ), and 1 (for  $x_j > 1 - w(C, o_j)$ ).

We can deduce from inequalities (7) and (8) that the confidence is decreasing with respect to  $x_j$  for  $x_j \leq 1 - w(C, o_j)$  and increasing with respect to  $x_j$  for  $x_j \geq 1 - w(C, o_j)$ . From that, we can deduce the following values for  $w(A, o)$ :

- if  $[D(A, o)] \leq 1 - [D(C, o)]$ , the confidence is decreasing on the whole  $D(A, o)$  interval. Therefore,  $[D(A, o)]$  is the point of global minimum;
- if  $[D(A, o)] \geq 1 - [D(C, o)]$ , the confidence is increasing on the whole  $D(A, o)$  interval. Therefore,  $[D(A, o)]$  is the point of global minimum;
- otherwise, the global minimum is reached for  $1 - [D(C, o)]$ .

Similarly,

- if  $[D(A, o)] \leq 1 - [D(C, o)]$ , the confidence is decreasing on the whole  $D(A, o)$  interval. Therefore,  $[D(A, o)]$  is the point of global maximum;

**Table 5** | Sets of suspicious points for rule  $A \rightarrow C$  from Example 3, if the minimum t-norm is being used.

	$M$		$w_2$	
	$A$	$C$	$A$	$C$
$o_1$	{1.0}	{0.8}	{0.9}	{0.9}
$o_2$	{0.0, 0.35}	{0.1}	{0.1}	{0.1}
$o_3$	{0.4}	{1.0}	{0.9}	{1.0}
$o_4$	{0.09}	{0.0}	{0.0}	{0.0}
$o_5$	{0.0}	{0.2}	{0.0}	{1.0}

- if  $[D(A, o)] \geq 1 - [D(C, o)]$ , the confidence is increasing on the whole  $D(A, o)$  interval. Therefore,  $[D(A, o)]$  is the point of global maximum;
- otherwise, the global maximum is reached either for  $[D(A, o)]$  or  $[D(A, o)]$ .

**Example 8.** Consider data from Example 3 and Table 2 and minimum t-norm. Accordingly to Lemmas 4 and 8, there exists a single possible world that is a candidate for a lower, and a different single possible world that is a candidate for the upper bound of the confidence of rule  $A \rightarrow C$  – see Table 6.

Lemmas 6–8 identify the points that are suspected to be global extremes for each object  $o \in \mathcal{O}$ . Recall that for each examined t-norm, there are at most two suspicious alternative values that could be extreme for each  $o_i \in \mathcal{O} = \{o_1, \dots, o_n\}$ . The core of the problem of finding the global extremes lies in selecting the correct value from these alternatives. The subsequent lemma transforms the problem to search for a set of indices  $I = \{1, \dots, n\}$ , for which the objects  $o_i$  pick the greatest suspicious value.

**Lemma 9.** Let  $\mathcal{O} = \{o_1, \dots, o_{n'}, \dots, o_n\}$  be a set of objects,  $n' \leq n$ ,  $\otimes$  be a t-norm,  $A \rightarrow C$  be a rule and  $D$  be a dataset. Let us have a set  $R$  of possible worlds,

$$R = \{w \mid \forall o_i \in \mathcal{O} : w(A, o_i) \in \alpha_i \wedge w(C, o_i) = c_i\},$$

where

$$\alpha_i = \begin{cases} \{a'_i, a''_i\} & \text{for } i \in \{1, \dots, n'\}, \\ \{a'_i\} & \text{for } i \in \{n' + 1, \dots, n\}, \end{cases}$$

$a'_i, a''_i \in D(A, o_i)$ ,  $a'_i < a''_i$ , and  $c_i \in D(C, o_i)$ . Then

$$\begin{aligned} \min\{\text{conf}_{\otimes}(w) \mid w \in R\} &= F(I_{\min}), \\ \max\{\text{conf}_{\otimes}(w) \mid w \in R\} &= F(I_{\max}), \end{aligned}$$

where

$$F(I) = \frac{G + \sum_{i \in I} g_i}{H + \sum_{i \in I} h_i}, \quad I \subseteq \{1, \dots, n'\}.$$

and where

$$\begin{aligned} G &= \sum_{i=1}^n a'_i \otimes c_i, \\ H &= \sum_{i=1}^n a'_i, \\ g_i &= a''_i \otimes c_i - a'_i \otimes c_i, \\ h_i &= a''_i - a'_i, \end{aligned}$$

**Table 6** | Sets of suspicious points for rule  $A \rightarrow C$  from Example 3, if the Łukasiewicz t-norm is being used.

	$w_1$		$L$	
	$A$	$C$	$A$	$C$
$o_1$	{0.9}	{0.8}	{1.0}	{0.9}
$o_2$	{0.35}	{0.1}	{0.0}	{0.1}
$o_3$	{0.4}	{1.0}	{0.9}	{1.0}
$o_4$	{0.09}	{0.0}	{0.0}	{0.0}
$o_5$	{0.0}	{0.2}	{0.0}	{1.0}

and  $I_{\min}, I_{\max} \subseteq \{1, \dots, n'\}$  such that for all  $I \subseteq \{1, \dots, n'\}$ ,

$$F(I_{\min}) \leq F(I) \leq F(I_{\max}).$$

**Proof.** Let  $w \in \mathbb{R}$  and let  $I = \{i \mid w(A, o_i) = a_i''\}$ . Then

$$\begin{aligned} \text{conf}_{\otimes}(w) &= \frac{\sum_{o_i \in \mathcal{O}} w(A, o_i) \otimes w(C, o_i)}{\sum_{o_i \in \mathcal{O}} w(A, o_i)} \\ &= \frac{\sum_{i \notin I} a_i' \otimes c_i + \sum_{i \in I} a_i'' \otimes c_i}{\sum_{i \notin I} a_i' + \sum_{i \in I} a_i''} \\ &= \frac{\sum_{i=1}^n a_i' \otimes c_i + \sum_{i \in I} (a_i'' \otimes c_i - a_i' \otimes c_i)}{\sum_{i=1}^n a_i' + \sum_{i \in I} (a_i'' - a_i')} \\ &= \frac{G + \sum_{i \in I} g_i}{H + \sum_{i \in I} h_i} = F(I). \end{aligned}$$

**Lemma 10.** Let  $G, H, g, h \in \mathbb{R}, H, h > 0$ . If  $\frac{g}{h} < \frac{G}{H}$  then

$$\frac{g}{h} < \frac{G+g}{H+h} < \frac{G}{H}.$$

**Proof.** From  $\frac{g}{h} < \frac{G}{H}$  and  $h > 0$  we get  $G > \frac{gH}{h}$ . By substituting that  $G$  into  $\frac{G+g}{H+h}$  we obtain

$$\begin{aligned} \frac{\frac{gH}{h} + g}{H+h} &< \frac{G+g}{H+h}, \\ \frac{\frac{g(H+h)}{h}}{H+h} &< \frac{G+g}{H+h}, \end{aligned}$$

which yields in  $\frac{g}{h} < \frac{G+g}{H+h}$ . The second inequality can be proven analogously.

**Lemma 11.** Let  $G, H, \in \mathbb{R}, H > 0$ , let  $g_i, h_i \in \mathbb{R}, h_i > 0$ , for all  $i \in \{1, \dots, n\}$ , and let

$$F(I) = \frac{G + \sum_{i \in I} g_i}{H + \sum_{i \in I} h_i}, \quad I \subseteq \{1, \dots, n\}. \quad (11)$$

Moreover, let  $I_{\min}, I_{\max} \subseteq \{1, \dots, n\}$  such that for all  $I \subseteq \{1, \dots, n\}$ ,

$$F(I_{\min}) \leq F(I) \leq F(I_{\max}). \quad (12)$$

Then  $\frac{g_j}{h_j} < F(I_{\min})$  implies  $j \in I_{\min}$  and  $\frac{g_j}{h_j} > F(I_{\min})$  implies  $j \notin I_{\min}$ . Similarly,  $\frac{g_j}{h_j} > F(I_{\max})$  implies  $j \in I_{\max}$  and  $\frac{g_j}{h_j} < F(I_{\max})$  implies  $j \notin I_{\max}$ .

**Proof.** Let us prove only the first implication, i.e.,  $\frac{g_j}{h_j} < F(I_{\min})$  implies  $j \in I_{\min}$ . (The other implications can be proven analogously.) Assume to the contrary that  $\frac{g_j}{h_j} < F(I_{\min})$  and  $j \notin I_{\min}$ . Let  $G' = G + \sum_{i \in I_{\min}} g_i$  and  $H' = H + \sum_{i \in I_{\min}} h_i$ . Then  $F(I_{\min}) = \frac{G'}{H'} > \frac{g_j}{h_j}$ . According to Lemma 10, it is evident that

$$F(I_{\min}) = \frac{G'}{H'} > \frac{G' + g_j}{H' + h_j} = F(I_{\min} \cup \{j\}),$$

which yields that  $F(I_{\min})$  is not minimal, which is a contradiction.

Next, we present the FRACMINIM algorithm that searches for  $F(I_{\min})$  (see Algorithm 1). The algorithm uses the so-called minimum heap (min-heap), which is a data structure that allows us to effectively obtain the smallest element of the set of stored values. The algorithm first puts all  $\frac{g_i}{h_i}$  into the heap and then iteratively pops the elements in ascending order from the smallest. Such approach is functionally equivalent to iterating through the collection of values sorted in ascending order by  $\frac{g_i}{h_i}$ . The benefit of using the heap instead of a sorted collection is that a heap has a better time complexity as the elements that are never popped from the heap never need to be fully ordered.

**Lemma 12.** Let  $G, H \in \mathbb{R}, H > 0$ , let  $\mathbf{g}, \mathbf{h}$  be vectors of real numbers, i.e.,  $\mathbf{g} = (g_1, \dots, g_n), \mathbf{h} = (h_1, \dots, h_n)$ , where  $g_i, h_i \in \mathbb{R}, h_i > 0$ , for all  $i \in \{1, \dots, n\}$ . Then the call FRACMINIM( $G, H, \mathbf{g}, \mathbf{h}$ ), as defined in Algorithm 1, returns  $F(I_{\min})$  as in (12).

**Proof.** We are going to prove this lemma inductively by proving a simultaneous validity of the following statements:

- (a) if  $\frac{g_i}{h_i} \geq \frac{G}{H}$  for all  $i \in \{1, \dots, n\}$  then FRACMINIM( $G, H, \mathbf{g}, \mathbf{h}$ ) =  $F(I_{\min})$  where  $\mathbf{g} = (g_1, \dots, g_n)$  and  $\mathbf{h} = (h_1, \dots, h_n)$ ;
- (b) let  $j \in \{1, \dots, n\}$  such that  $\frac{g_j}{h_j} \leq \frac{g_i}{h_i}$  for all  $i \in \{1, \dots, n\}$ ; if  $\frac{g_j}{h_j} < \frac{G}{H}$  then FRACMINIM( $G, H, \mathbf{g}, \mathbf{h}$ ) = FRACMINIM( $G', H', \mathbf{g}', \mathbf{h}'$ ) where  $G' = G + g_j, H' = H + h_j, \mathbf{g}' = (g_1, \dots, g_{j-1}, g_{j+1}, \dots, g_n), \mathbf{h}' = (h_1, \dots, h_{j-1}, h_{j+1}, \dots, h_n)$ .

Ad (a): Let  $\frac{g_i}{h_i} \geq \frac{G}{H}$  for all  $i \in \{1, \dots, n\}$ , possibly also  $n$  may be equal to 0, then obviously the Algorithm 1 returns  $\frac{G}{H}$ . A proof is needed that  $\frac{G}{H}$  is the correct value for  $F(I_{\min})$ . Assume to the contrary that  $F(I_{\min}) < \frac{G}{H}$ . Then from the definition of  $F(I)$  in (11) obviously follows that  $I_{\min} \neq \emptyset$ . On the other hand, it follows from assumptions  $\frac{g_i}{h_i} \geq \frac{G}{H} > F(I_{\min})$  and from Lemma 11 that  $i \notin I_{\min}$ , for all  $i \in \{1, \dots, n\}$ , which is in contradiction with  $I_{\min} \neq \emptyset$ .

Ad (b): Without loss of generality, let us assume the vectors  $\mathbf{g}$  and  $\mathbf{h}$  to be indexed according to the ascending order of values  $\frac{g_i}{h_i}$ , i.e.,  $\frac{g_1}{h_1} \leq \frac{g_2}{h_2} \leq \dots \leq \frac{g_m}{h_m}$ . Now, let  $j \in \{1, \dots, n\}$  such that  $\frac{g_j}{h_j} \leq \frac{g_i}{h_i}$  for all  $i \in \{1, \dots, n\}$  and let  $k = \left\{ \left\{ i \mid \frac{g_i}{h_i} = \frac{g_j}{h_j} \right\} \right\}$ , i.e.,  $k$  is the number of equally smallest fractions. If  $k = 1$  then obviously also  $j = 1$  and calls FRACMINIM( $G, H, \mathbf{g}, \mathbf{h}$ ) and FRACMINIM( $G', H', \mathbf{g}', \mathbf{h}'$ ) return evidently the same result. If  $k > 1$  then  $j \in \{1, \dots, k\}$  and it is clear that the call FRACMINIM( $G, H, \mathbf{g}, \mathbf{h}$ ) will process at least all first  $k$  elements from the vectors  $\mathbf{g}$  and  $\mathbf{h}$ , i.e., after  $k$  steps of the “while” loop, the variables  $G$  and  $H$  become incremented by  $\sum_{i=1}^k g_i$  and  $\sum_{i=1}^k h_i$ , respectively. This is due the fact that  $\frac{g_j}{h_j} < \frac{G}{H}$  implies accordingly to Lemma 10 that  $\frac{g_j}{h_j} < \frac{G+g_j}{H+h_j}$ . Hence, the “if” condition inside of the “while” loop remains true for all first  $k$  fractions in the heap. Then obviously, the call FRACMINIM( $G', H', \mathbf{g}', \mathbf{h}'$ ) will surely process all first  $k - 1$  fractions and thus increment  $G'$  and  $H'$  by  $\sum_{i=1}^{k-1} g'_i = \sum_{i=1, i \neq j}^k g_i$  and  $\sum_{i=1}^{k-1} h'_i = \sum_{i=1, i \neq j}^k h_i$ ,

respectively. This proves  $\text{FRACMINIM}(G, H, \mathbf{g}, \mathbf{h}) = \text{FRACMINIM}(G', H', \mathbf{g}', \mathbf{h}')$ .

**Corollary 13.** Let  $G, H \in \mathbb{R}, H > 0$ , let  $\mathbf{g} = (g_1, \dots, g_n), \mathbf{h} = (h_1, \dots, h_n)$ , where  $g_i, h_i \in \mathbb{R}, h_i > 0$ , for all  $i \in \{1, \dots, n\}$ . Then the call of the function  $\text{FRACMINIM}(-G, H, -\mathbf{g}, \mathbf{h})$ , as defined in Algorithm 1, returns  $-F(I_{\max})$  for  $F(I)$  as in (11) and  $I_{\max}$  as in (12).

**Proof.** The call of  $\text{FRACMINIM}(-G, H, -\mathbf{g}, \mathbf{h})$  searches for  $F'(I'_{\min})$ , where

$$F'(I) = \frac{-G + \sum_{i \in I} (-g_i)}{H + \sum_{i \in I} h_i} = -\frac{G + \sum_{i \in I} g_i}{H + \sum_{i \in I} h_i} = -F(I),$$

for  $F(I)$  as in (11) and where for all  $I \subseteq \{1, \dots, n\}$

$$F'(I'_{\min}) \leq F'(I) = -F(I) \leq F'(I'_{\max}).$$

Hence,  $-F'(I'_{\min})$  is the highest bound of  $F(I)$ , and therefore,  $-F'(I'_{\min}) = F(I_{\max})$ .

**Theorem 14.** Let  $D$  be a dataset and  $A \rightarrow C$  be an association rule. Then the call of function  $\text{CONFIDENCE}_{\ominus}(D, A, C)$ , as defined in Algorithm 2, returns  $\text{conf}_{\ominus}(A \rightarrow C)$ .

**Proof.** Algorithm 2 prepares arguments for calls of  $\text{FRACMINIM}$ . Namely, variables  $G, G', H$  and vectors  $\mathbf{g}, \mathbf{g}', \mathbf{h}$  are initialized as follows (where  $\mathcal{O} = \{o_1, \dots, o_n\}$ ):

$$G = \sum_{i=1}^n [D(A, o_i)] \cdot [D(C, o_i)], \quad (13)$$

$$G' = \sum_{i=1}^n [D(A, o_i)] \cdot [D(C, o_i)], \quad (14)$$

$$H = \sum_{i=1}^n [D(A, o_i)], \quad (15)$$

$$\mathbf{g} = (g_i)_{i=1, g_i \neq 0}^n \text{ where } g_i = [D(A, o_i)] \cdot [D(C, o_i)] - [D(A, o_i)] \cdot [D(C, o_i)], \quad (16)$$

$$\mathbf{g}' = (g'_i)_{i=1, g'_i \neq 0}^n \text{ where } g'_i = [D(A, o_i)] \cdot [D(C, o_i)] - [D(A, o_i)] \cdot [D(C, o_i)], \quad (17)$$

$$\mathbf{h} = (h_i)_{i=1, h_i \neq 0}^n \text{ where } h_i = [D(A, o_i)] - [D(A, o_i)], \quad (18)$$

Let  $n' = |\mathbf{g}| = |\mathbf{g}'| = |\mathbf{h}|$ . Note that  $n' \leq n$  because the vectors do not contain zero elements, which correspond to objects  $o_i$  such that  $D(A, o_i)$  is a single-element interval.

First of all, let us prove that the algorithm computes the correct lower bound  $lo$  of the confidence. According to Lemma 4, the lower bound is obtained for  $w(C, o) = [D(C, o)]$  (for each  $o \in \mathcal{O}$ ). Regarding  $w(A, o)$ , the lower bound of the confidence is obtained for  $w(A, o) \in \{[D(A, o)], [D(A, o)]\}$  (for each  $o \in \mathcal{O}$ ), which follows from Lemma 6. The whole problem of finding the lower bound thus consists in finding the proper combination of  $w(A, o)$  for all  $o$ .

In other words, we have a set of possible worlds

$$R = \{w \mid \forall o_i \in \mathcal{O} : w(A, o_i) \in \{[D(A, o)], [D(A, o)]\} \wedge w(C, o_i) = [D(C, o_i)]\},$$

from which we have to select the one that yields in the minimum confidence. According to Lemma 9, finding a minimum confidence among worlds  $w \in R$  is equivalent to finding  $F(I_{\min})$ , where

$$F(I) = \frac{G + \sum_{i \in I} g_i}{H + \sum_{i \in I} h_i}, \quad I \subseteq \{1, \dots, n\},$$

with  $G, H$  from (13), (15) and  $\mathbf{g} = (g_i), \mathbf{h} = (h_i)$  from (16), (18). The minimum confidence is obtained by calling  $\text{FRACMINIM}(G, H, \mathbf{g}, \mathbf{h})$ , as proven in Lemma 12.

Similarly, finding the upper bound of the confidence is equivalent to finding  $F'(I_{\max})$ , where

$$F'(I) = \frac{G' + \sum_{i \in I} g'_i}{H + \sum_{i \in I} h_i}, \quad I \subseteq \{1, \dots, n\},$$

with  $G', H$  from (14), (15) and  $\mathbf{g} = (g_i), \mathbf{h} = (h_i)$  from (17), (18). According to Corollary 13, the maximum confidence is obtained by calling the function  $\text{FRACMINIM}(-G', H, -\mathbf{g}', \mathbf{h})$ .

**Theorem 15.** Let  $D$  be a dataset and  $A \rightarrow C$  be an association rule. Then the call of function  $\text{CONFIDENCE}_{\min}(D, A, C)$ , as defined in Algorithm 3, returns  $\text{conf}_{\min}(A \rightarrow C)$ .

**Proof.** The proof is analogous to the proof of Theorem 14. It is again based on the search for a possible world  $w \in W$  that yields a globally minimal or maximal confidence. Again, the lower or upper bound of the confidence is respectively reached for  $w(C, o)$  equal to  $[D(C, o)]$  or  $[D(C, o)]$ , see Lemma 4.

Let us prove the correctness of computation of the lower bound  $lo$ . Let

$$S = \{i \mid [D(A, o_i)] \leq [D(C, o_i)]\}$$

$$T = \{i \mid [D(A, o_i)] \geq [D(C, o_i)]\}$$

$$U = \{i \mid [D(A, o_i)] < [D(C, o_i)] < [D(A, o_i)]\}$$

Note that  $S \cup T \cup U = \{1, \dots, n\}$  and  $S, T, U$  are pairwise disjoint. The algorithm initializes the variables  $G, H, \mathbf{g}, \mathbf{h}$  as follows:

$$G = \sum_{i \in S \cup U} [D(A, o_i)] + \sum_{i \in T} [D(C, o_i)],$$

$$H = \sum_{i \in S \cup U} [D(A, o_i)] + \sum_{i \in T} [D(A, o_i)],$$

$$\mathbf{g} = (g_i)_{i \in U, g_i \neq 0} \text{ where } g_i = [D(C, o_i)] - [D(A, o_i)],$$

$$\mathbf{h} = (h_i)_{i \in U, h_i \neq 0} \text{ where } h_i = [D(A, o_i)] - [D(A, o_i)],$$

Note that  $G, \mathbf{g}$  are equivalent to

$$G = \sum_{i=1}^n \min([D(A, o_i)], [D(C, o_i)]),$$

$$\mathbf{g} = (g_i)_{i \in U, g_i \neq 0},$$

$$g_i = \min([D(A, o_i)], [D(C, o_i)]) - \min([D(A, o_i)], [D(C, o_i)]).$$

It is now easy to see that  $G, H, \mathbf{g}, \mathbf{h}$  are equivalent to  $G, H, \mathbf{g}, \mathbf{h}$  from Lemma 9 with  $a'_i, a''_i$  set as in Lemma 7. This all together with Lemma 11 and Lemma 12 leads to the fact that the call of  $\text{FRACMINIM}(G, H, \mathbf{g}, \mathbf{h})$  indeed returns the lower bound of the confidence.

The correctness of upper bound  $hi$  easily follows from (10) in Lemma 7.

**Theorem 16.** *Let  $D$  be a dataset and  $A \rightarrow C$  be an association rule. Then the call of function  $\text{CONFIDENCE}_{\otimes_L}(D, A, C)$ , as defined in Algorithm 4, returns  $\text{conf}_{\otimes_L}(A \rightarrow C)$ .*

**Proof.** The proof is very analogous to proofs of Theorems 14 and 15. It follows from Lemmas 4, 8, 9 and 12 and Corollary 13.

**Theorem 17.** *The time complexity of  $\text{FRACMINIM}$ ,  $\text{CONFIDENCE}_{\odot}$ ,  $\text{CONFIDENCE}_{\min}$  and  $\text{CONFIDENCE}_{\otimes_L}$  is in  $O(n \log n)$ , for  $n = |\mathcal{O}|$ .*

**Proof.** First of all, let us focus on the time complexity of  $\text{FRACMINIM}$ . It is known that pushing an element into the heap as well as popping the minimal element from the heap is in  $O(\log n)$ .  $\text{FRACMINIM}$  does this step  $n$ -times, which results in  $O(n \log n)$ . After that, the elements are removed from the heap sequentially, which is again in  $O(n \log n)$ . Hence, the whole  $\text{FRACMINIM}$  is in  $O(n \log n)$ .

Regarding Algorithms 2–4, the initialization steps can be executed with time complexity in  $O(n)$ . The calls of  $\text{FRACMINIM}$  afterwards are in  $O(n \log n)$ , hence, whole algorithms are in  $O(n \log n)$  too.

## 4. EXPERIMENT

In this section, a practical experiment is conducted to show the efficiency of the algorithm with real settings. As many other papers have addressed the usefulness of the association analysis, we limit ourselves only to the computation of the confidence, which is the main topic of this paper.

To evaluate the performance, we randomly generated (from the uniform distribution) four vectors of equal size, which represent the lower and upper bounds of antecedent  $A$  and consequent  $C$ . Three algorithms were then executed to compute  $A \rightarrow C$ 's confidence based on the Gödel, Goguen and Łukasiewicz t-norm. Since we are not aware of any other algorithm for the computation of the confidence on interval data, we have compared it to the algorithm that computes the confidence on noninterval data which simply involves the calculation of the equation (2). Naturally, the noninterval algorithm is much faster because it has a linear time complexity,  $O(n)$ . The results can be found in Table 7.

As can be seen, the algorithm for Gödel t-norm is about 32 times slower than noninterval computation, the Łukasiewicz variant 43 times, and the confidence based on Goguen t-norm is computed in average 319 times slower than the noninterval variant. However, we are very satisfied with the results as the computation of interval-valued confidence, which is much more complicated than the noninterval variant, still costs less than a second for both Gödel and Łukasiewicz t-norm even for data having more than 13 millions of records with each record being interval-valued, which will often not be the case in real applications.

The single-threaded experiment was performed at Intel(R) Core(TM) i5-3570 CPU at 3.40GHz.

### Algorithm 4: Computation of $\text{conf}_{\otimes_L}(A \rightarrow C)$ on the dataset $D$

```

function (Confidence)  $\otimes_L(D, A, C)$ 
   $G, H, M, N \leftarrow 0$ 
   $\mathbf{g}, \mathbf{h} \leftarrow$  new empty vector
  for all  $o \in \mathcal{O}$  do
     $al \leftarrow \lfloor D(A, o) \rfloor, ah \leftarrow \lfloor D(A, o) \rfloor$ 
     $cl \leftarrow \lfloor D(C, o) \rfloor, ch \leftarrow \lfloor D(C, o) \rfloor$ 
     $a \leftarrow \text{closest}(\lfloor al, ah \rfloor, 1 - cl)$ 
     $M \leftarrow M + \max(0, a + cl - 1)$ 
     $N \leftarrow N + a$ 
    if  $ah \leq 1 - ch$  then
       $H \leftarrow H + al$ 
    else if  $al \geq 1 - ch$  then
       $G \leftarrow G + ah + ch - 1$ 
       $H \leftarrow H + ah$ 
    else
       $H \leftarrow H + al$ 
      if  $al \neq ah$  then
        append  $(ah + ch - 1)$  to  $\mathbf{g}$ 
        append  $(ah - al)$  to  $\mathbf{h}$ 
      end if
    end if
  end for
   $lo \leftarrow \frac{M}{N}$ 
   $hi \leftarrow -\text{FRACMINIM}(-G, H, -\mathbf{g}, \mathbf{h})$ 
  return  $\lfloor lo, hi \rfloor$ 
end function

```

**Table 7** | Experimental running times (in milliseconds) of computations of confidence based on Gödel t-norm (G), Goguen product t-norm (P) and Łukasiewicz t-norm (L). A noninterval confidence computation (non-int) is also provided for comparison.

Data Size	Non-int	G	P	L
400	0.00	0.07	0.19	0.03
800	0.00	0.05	0.35	0.07
1600	0.00	0.09	0.70	0.12
3200	0.01	0.18	1.49	0.24
6400	0.01	0.35	2.93	0.48
12800	0.01	0.37	3.21	0.53
25600	0.02	0.71	6.93	1.02
51200	0.12	1.52	14.83	2.24
102400	0.09	2.84	29.60	4.31
204800	0.21	5.74	65.66	9.13
409600	0.63	12.12	144.46	18.59
819200	1.39	25.11	322.94	38.27
1638400	2.17	52.72	850.75	81.70
3276800	4.44	108.16	1769.73	167.95
6553600	8.60	226.69	4073.44	365.52
13107200	17.43	482.65	9182.41	795.08

## 5. CONCLUSION

In this paper, efficient algorithms were presented that allow computation of the support and confidence of an association rule from interval-valued fuzzy data. Interval values may be used to model a partial knowledge, i.e., values that are unknown or measured with a limited precision. The proposed algorithms compute the minimal and maximal possible support and confidence based on the assumption that the nonprecise data are independent. It is shown that the computation of the confidence on interval-valued data is more complex than the computation for noninterval data. The time complexity of the interval-valued confidence is in  $O(n \log n)$ . For contrast, the time complexity of noninterval data confidence is in

$O(n)$ . However, the most time-consuming part of the association analysis is the search for rules with sufficiently high support, whose complexity remains unchanged (as discussed in Theorem 3).

The computation of minimal and maximal confidence is an optimization problem that depends on the applied t-norm. The algorithms presented here allow computation with the Gödel (minimum) t-norm, Goguen (product) t-norm and Łukasiewicz t-norm. However, several results are applicable to an arbitrary t-norm: see Lemmas 4 and 5. Moreover, the developed algorithms are directly applicable to a t-norm that is proven to be connected to at most two suspicious values (see Lemma 5) per data row. Investigation of other t-norms or families of t-norms is left for future work.

## CONFLICT OF INTEREST

Authors declare no conflict of interest.

## AUTHORS' CONTRIBUTIONS

All authors contributed to the work presented in this paper. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

Authors acknowledge support by project “LQ1602 IT4Innovations excellence in science” and by GAČR 16-19170S.

## REFERENCES

- [1] P. Murinová, M. Burda, V. Pavliska, Fuzzy association rules on data with undefined values, in: J. Medina, M. Ojeda-Aciego, J. Verdegay, I. Perfilieva, B. Bouchon-Meunier, R. Yager (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Applications, IPMU 2018*, vol. 855, Springer, Cham, Switzerland, 2018, pp. 165–174.
- [2] P. Hájek, The problem of a general conception of the GUHA method, *Kybernetika*. 4 (1968), 505–515. <https://dml.cz/handle/10338.dmlcz/125459>
- [3] P. Hájek, T. Havránek, *Mechanizing Hypothesis Formation (Mathematical Foundations for a General Theory)*, Springer-Verlag, Berlin, Heidelberg, Germany, 1978.
- [4] R. Agrawal, R. Srikant, Fast algorithms for mining association rules, in *Proceeding of 20th International Conference on Very Large Databases*, AAAI Press, Santiago de Chile, Chile, Chile, 1994, pp. 487–499. <https://dl.acm.org/doi/proceedings/10.5555/645920>
- [5] R.R. Yager, A new approach to the summarization of data, *Inf. Sci.* 28 (1982), 69–86.
- [6] J. Kacprzyk, R.R. Yager, S. Zadrożny, A fuzzy logic based approach to linguistic summaries of databases, *Int. J. Appl. Math. Comput. Sci.* 10 (2000), 813–834. <https://www.amcs.uz.zgora.pl/?action=paper&paper=930>
- [7] M. Ralbovský, *Fuzzy GUHA*, PhD Thesis, University of Economics, Prague, Czech Republic, 2009. <https://insis.vse.cz/zp/14956/podrobnosti>
- [8] P. Murinová, M. Burda, V. Pavliska, An algorithm for intermediate quantifiers and the graded square of opposition towards linguistic description of data, in: J. Kacprzyk, E. Szmidi, S. Zadrożny, K. Atanassov, M. Krawczak (Eds.), *Advances in Fuzzy Logic and Technology*, Springer, Cham, Switzerland, 2017, pp. 592–603.
- [9] M. Burda, P. Murinová, V. Pavliska, On missing membership degrees: modelling non-existence, ignorance and inconsistency, in: S. Destercke, T. Denoeux, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), *Uncertainty Modelling in Data Science, SMPS 2018*, Springer, Compiègne, France, 2018,
- [10] A. Gelman, J. Hill, *Data Analysis Using Regression and Multi-level/Hierarchical Models*, Volume Analytical Methods for Social Research, Cambridge University Press, New York, NY, USA, 2007.
- [11] Y. Liu, V. Gopalakrishnan, An overview and evaluation of recent machine learning imputation methods using cardiac imaging data, *Data*. 2 (2017), 8.
- [12] J. Ross Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993. <https://dl.acm.org/doi/book/10.5555/152181>
- [13] J. Łukasiewicz, O logice trojwartosciowej, *Ruch Filozoficzny*. 5 (1920), 170–171. <https://philpapers.org/rec/UKAOLT-2>
- [14] G. Malinowski, *The Many Valued and Nonmonotonic Turn in Logic*, North-Holland, Amsterdam, Netherlands, 2007.
- [15] L. Běhounek, V. Novák, Towards fuzzy partial logic, in *Proceedings of the IEEE 45th International Symposium on Multiple-Valued Logics (ISMVL 2015)*, Waterloo, Canada, 2015, pp. 139–144.
- [16] L. Běhounek, M. Daňková, Towards fuzzy partial set theory, in: J. Carvalho, M.J. Lesot, U. Kaymak, S. Vieira, B. Bouchon-Meunier, R. Yager (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2016)*, Springer, Cham, Switzerland, 2016, pp. 482–494.
- [17] P. Murinová, M. Burda, V. Pavliska, Undefined values in fuzzy logic, in: J. Kacprzyk, E. Szmidi, S. Zadrożny, K. Atanassov, M. Krawczak (Eds.), *Advances in Fuzzy Logic and Technology*, Springer, Cham, Switzerland, 2017, pp. 604–610.
- [18] D. Ciucci, D. Dubois, J. Lawry, Borderline vs. unknown: comparing three-valued representations of imperfect information, *Int. J. Approx. Reason.* 55 (2014), 1866–1889.
- [19] D. Ciucci, D. Dubois, Three-valued logics, uncertainty management and rough set, *Int. J. Approx. Reason.* 55 (2014), 1866–1889.
- [20] P. Grzegorzewski, Measures of dispersion for interval data, in: S. Destercke, T. Denoeux, M.Á. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), *Uncertainty Modelling in Data Science*, Springer International Publishing, Cham, Switzerland, 2019, pp. 91–98.
- [21] L. Billard, E. Diday, Regression analysis for interval-valued data, H.A.L. Kiers, J.-P. Rasson, P.J.F. Groenen, M. Schader (Eds.), *Data Analysis, Classification, and Related Methods*, Springer, Berlin, Heidelberg, Germany, 2000, pp. 369–374.
- [22] E.D. Lima Neto, F.D. De Carvalho, Nonlinear regression applied to interval-valued data, *Pattern Anal. Appl.* 20 (2017), 809–824.
- [23] P. Murinová, M. Burda, V. Pavliska, Association analysis on interval-valued fuzzy sets: uncertainty, vagueness and non-existence, in: V. Torra, Y. Narukawa, I. Aguiló, M. González-Hidalgo (Eds.), *Modeling Decisions for Artificial Intelligence*, Springer, Cham, Switzerland, 2018, pp. 89–100.
- [24] Q. Chen, S. Kawase, An approach towards consistency degrees of fuzzy theories, *Fuzzy Sets Syst.* 113 (2000), 237–251.

- [25] C. Deschrijver Cornelis, E.G. Kerre, Implication in intuitionistic fuzzy and interval-valued fuzzy set theory: construction, classification, application, *Int. J. Approx. Reason.* 35 (2004), 55–95.
- [26] L.A. Zadeh, Fuzzy sets, *Inf. Control.* 8 (1965), 338–353.
- [27] E.P. Klement, R. Mesiar, E. Pap, *Triangular Norms*, Kluwer, Dordrecht, Netherlands, 2000.
- [28] S. Kawase, Q. Chen, N. Yanagihar, On interval valued fuzzy reasoning, *Trans. Japan Soc. Ind. Appl. Math.* 6 (1996), 285–296.
- [29] L. Geng, H.J. Hamilton, Interestingness measures for data mining: a survey, *ACM Comput. Surv. (CSUR)*. 38 (2006), 9.
- [30] R. Agrawal, T. Imielinski, A. Swami, Mining association rules between sets of items in large databases, in *Proceeding of the ACM SIGMOD International Conference on Management of Data*, Washington, DC, USA, 1993, pp. 207–216.