Boolean factors as a means of clustering of interestingness measures of association rules ⋆

Radim Belohlavek[1], Dhouha Grissa[2,4,5], Sylvie Guillaume[3,4], Engelbert Mephu Nguifo[2,4], Jan Outrata[1]

[1] Data Analysis and Modeling Lab
Department of Computer Science , Palacky University, Olomouc
17. listopadu 12, CZ-77146 Olomouc, Czech Republic
radim.belohlavek@acm.org,jan.outrata@upol.cz
[2] Clermont Université, Université Blaise Pascal, LIMOS, BP 10448, F-63000 Clermont-Ferrand, France
[3] Clermont Université, Université d'Auvergne, LIMOS, BP 10448, F-63000 Clermont-Ferrand, France
[4] CNRS, UMR 6158, LIMOS, F-63173 Aubiére, France
[5] URPAH, Département d'Informatique, Faculté des Sciences de Tunis, Campus Universitaire, 1060 Tunis, Tunisie
dgrissa@isima.fr,guillaum@isima.fr,mephu@isima.fr

Abstract. Measures of interestingness play a crucial role in association rule mining. An important methodological problem is to provide a reasonable classification of the measures. Several papers appeared on this topic. In this paper, we explore Boolean factor analysis, which uses formal concepts corresponding to classes of measures as factors, for the purpose of classification and compare the results to the previous approaches.

1 Introduction

An important problem in extracting association rules, well known since the early stage of association rule mining [32], is the possibly huge number of rules extracted from data. A general way of dealing with this problem is to define the concept of rule interestingness: only association rules that are considered interesting according to some measure are presented to the user. The most widely used measures of interestingness are based on the concept of support and confidence. However, the suitability of these measures to extract interesting rules was challenged by several studies, see e.g. [34]. Consequently, several other interestingness measures of association rules were proposed, see e.g. [35], [23], [12], [38]. With the many existing measures of interestingness arises the problem of selecting an appropriate one.

⋆ We acknowledge support by the ESF project No. CZ.1.07/2.3.00/20.0059, the project is co-financed by the European Social Fund and the state budget of the Czech Republic (R. Belohlavek); Grant No. 202/10/P360 of the Czech Science Foundation (J. Outrata); and by Grant No. 11G1417 of the French-Tunisian cooperation PHC Utique (D. Grissa).

To understand better the behavior of various measures, several studies of the properties of measures of interestingness appeared, see e.g. [12], [27], [23], [16]. Those studies explore various properties of the measures that are considered important. For example, Vaillant et al. [37] evaluated twenty interestingness measures according to eight properties. To facilitate the choice of the user-adapted interestingness measure, the authors applied the clustering methods on the decision matrix and obtained five clusters. Tan et al. [35] studied twenty-one interestingness measures through eight properties and showed that no measure is adapted to all cases. To select the best interestingness measure, they use both a support-based pruning and standardization methods. By applying a new clustering approach, Huynh et al. [21] classifyed thirty-four interestingness measures with a correlation analysis. Geng and Hamilton [12] made a survey of thirty-eight interestingness measures for rules and summaries with eleven properties and gived strategies to select the appropriate measures. D. R. Feno [10] evaluated fifteen interestingness measures with thirteen properties to describe their behaviour. Delgato et al. [9] provided a new study of the interestingness measures by means of the logical model. In addition, the authors proposed and justified the addition of two new principles to the three proposed by Piatetsky-Shapiro [32]. Finally, Heravi and Zaiane [22] studied fifty-three objective measures for associative classification rules according to sixteen properties and explained that no single measure can be introduced as an obvious winner.

The assessment of measures according to their properties results in a measure-property binary matrix. Two studies of this matrix were conducted. Namely, [17] describes how FCA can highlight interestingness measures with similar behavior in order to help the user during his choice. [16] and [14] attempted to find natural clusters of measures using widely used clustering methods, the agglomerative hierarchical method (AHC) and the K-means method. A common feature of these methods is that they only produce disjoint clusters of measures. On the other hand, one could naturally expect overlapping clusters. The aim of this paper is to explore the possibility of obtaining overlapping clusters of measures using factor analysis of binary data and to compare the results with the results of other studies. In particular, we use the recently developed method from [3] and take the discovered factors for clusters. The method uses formal concepts as factors that makes it possible to interpret the factors easily.

2 Preliminaries

2.1 Binary (Boolean) data

Let X be a set of objects (such as a set of customers, a set of functions or the like) and Y be a set of attributes (such as a set of products that customers may buy, a set of properties of functions). The information about which objects have which attributes may formally be represented by a binary relation I between X and Y, i.e. $I \subseteq X \times Y$, and may be visualized by a table (matrix) that contains 1s and 0s, according to whether the object corresponding to a row has the attribute corresponding to a column (for this we suppose some orders of objects and attributes are fixed). We denote the entries of such matrix by I_{xy}. A data of this type is called *binary data* (or Boolean data). The triplet $\langle X, Y, I \rangle$ is called a *formal context* in FCA but other terms are used in other areas.

Such type of data appears in two roles in our paper. First, association rules, whose interestingness measures we analyze, are certain dependencies over the binary data. Second, the information we have about the interestingness measures of association rules is in the form of binary data: the objects are interestingness measures and the attributes are their properties.

2.2 Association rules

An *association rule* [36] over a set Y of attributes is a formula

$$A \Rightarrow B \tag{1}$$

where A and B are sets of attributes from Y, i.e. $A, B \subseteq Y$. Let $\langle X, Y, I \rangle$ be a formal context. A natural measure of interestingness of association rules is based on the notions of confidence and support. The *confidence* and *support* of an association rule $A \Rightarrow B$ in $\langle X, Y, I \rangle$ is defined by

$$\mathrm{conf}(A \Rightarrow B) = \frac{|A^{\downarrow} \cap B^{\downarrow}|}{|A^{\downarrow}|} \quad \text{and} \quad \mathrm{supp}(A \Rightarrow B) = \frac{|A^{\downarrow} \cap B^{\downarrow}|}{|X|},$$

where $C^{\downarrow}$ for $C \subseteq Y$ is defined by $C^{\downarrow} = \{x \in X \mid \text{ for each } y \in C : \langle x, y \rangle \in I\}$. An association rule is considered interesting if its confidence and support exceed some user-specified thresholds. However, the support-confidence approach reveals some weaknesses. Often, this approach as well as algorithms based on it lead to the extraction of an exponential number of rules. Therefore, it is impossible to validate it by an expert. In addition, the disadvantage of the support is that sometimes many rules that are potentially interesting, have a lower support value and therefore can be eliminated by the pruning threshold *minsupp*. To address this problem, many other measures of interestingness have been proposed in the literature [13], mainly because they are effective for mining potentially interesting rules and capture some aspects of user interest. The most important of those measures are subject to our analysis and are surveyed in Section 3.1. Note that association rules are attributed to [1]. However, the concept of association rule itself as well as various measures of interestingness are particular cases of what is investigated in depth in [18], a book that develops logico-statistical foundations of the GUHA method [19].

2.3 Factor analysis of binary (Boolean) data

Let I be an $n \times m$ binary matrix. The aim in Boolean factor analysis is to find a decomposition

$$I = A \circ B \tag{2}$$

of I into an $n \times k$ binary matrix A and a $k \times m$ binary matrix B with $\circ$ denoting the Boolean product of matrices, i.e.

$$(A \circ B)_{ij} = \max_{l=1}^{k} \min(A_{il}, B_{lj}).$$

The inner dimension, k, in the decomposition may be interpreted as the number of factors that may be used to describe the original data. Namely, $A_{il} = 1$ if and only if the lth factor applies to the ith object and $B_{lj} = 1$ if and only if the jth attribute is one of the manifestations of the lth factor. The factor model behind (2) has therefore the following meaning: The object i has the attribute j if and only if there exists a factor l that applies to i and for which j is one of its particular manifestations. We refer to [3] for further information and references to papers that deal with the problem of factor analysis and decompositions of binary matrices.

In [3], the following method for finding decompositions (2) with the number k of factors as small as possible has been presented. The method utilizes formal concepts of the formal context $\langle X, Y, I\rangle$ as factors, where $X = \{1, \ldots, n\}$, $Y = \{1, \ldots, m\}$ (objects and attributes correspond to the rows and columns of I). Let

$$\mathcal{F} = \{\langle C_1, D_1\rangle, \ldots, \langle C_k, D_k\rangle\}$$

be a set of formal concepts of $\langle X, Y, I\rangle$, i.e. $\langle C_l, D_l\rangle$ are elements of the concept lattice $\mathcal{B}(X, Y, I)$ [11]. Consider the $n \times k$ binary matrix $A_\mathcal{F}$ and a $k \times m$ binary matrix $B_\mathcal{F}$ defined by

$$(A_\mathcal{F})_{il} = 1 \text{ iff } i \in C_l \quad \text{and} \quad (B_\mathcal{F})_{lj} = 1 \text{ iff } j \in D_l. \tag{3}$$

Denote by $\rho(I)$ the smallest number k, so-called Schein rank of I, such that a decomposition of I exists with k factors. The following theorem shows that using formal concepts as factors as in (3) enables us to reach the Schein rank, i.e. is optimal [3]:

Theorem 1. *For every binary matrix I, there exists $\mathcal{F} \subseteq \mathcal{B}(X, Y, I)$ such that $I = A_\mathcal{F} \circ B_\mathcal{F}$ and $|\mathcal{F}| = \rho(I)$.*

As has been demonstrated in [3], a useful feature of using formal concepts as factors is the fact that formal concepts may easily be interpreted. Namely, every factor, i.e. a formal concept $\langle C_l, D_l\rangle$, consists of a set C_l of objects (objects are measures of interestingness in our case) and a set D_l of attributes (properties of measures in our case). C_l contains just the objects to which all the attributes from D_l apply and D_l contains all attributes shared by all objects from C_l. From a clustering point of view, the factors $\langle C_l, D_l\rangle$ may thus be seen as clusters C_l with their descriptions by attributes from D_l. The factors thus have a natural, easy to understand meaning. Since the problem of computing the smallest set of factors is NP-hard, a greedy approximation algorithm was proposed in [3, Algorithm 2]. This algorithm is utilized below in our paper.

3 Clustering interestingness measures using Boolean factors

3.1 Measures of interestingness

In the following, we present the interestingness measures reported in the literature and recall nineteen of their most important properties that were proposed in the literature.

To identify interesting association rules and to enable the user to focus on what is interesting for him, about sixty interestingness measures [20], [35], [10] were proposed in the literature. All of them are defined using the following parameters: $p(XY)$, $p(\bar{X}Y)$, $p(X\bar{Y})$ and $p(\bar{X}\bar{Y})$, where $p(XY) = \frac{n_{XY}}{n}$ represents the number of objects satisfying XY (the intersection of X and Y), and $\bar{X}$ is the negation of X. The following are important examples of interestingness measures:

Lift *[6]:* Given a rule $X \to Y$, *lift* is the ratio of the probability that X and Y occur together to the multiple of the two individual probabilities for X and Y, i.e.,

$$Lift(X \to Y) = \frac{p(XY)}{p(X) \times p(Y)}.$$

If this value is 1, then X and Y are independent. The higher this value, the more likely that the existence of X and Y together in a transaction is not just a random occurrence, but because of some relationship between them.

Correlation coefficient *[31]: Correlation* is a symmetric measure evaluating the strength of the itemsets' connection. It is defined by

$$Correlation = \frac{p(XY) - p(X)p(Y)}{\sqrt{p(X)p(Y)p(\bar{X})p(\bar{Y})}}.$$

A correlation around 0 indicates that X and Y are not correlated. The lower is its value, the more negatively correlated X and Y are. The higher is its value, the more positively correlated they are.

Conviction *[6]: Conviction* is one of the measures that favor counter-examples. It is defined by

$$Conviction = \frac{p(X)p(\bar{Y})}{p(X\bar{Y})}$$

Conviction which is not a symmetric measure, is used to quatify the deviation from independence. If its value is 1, then X and Y are independent.

M_{GK} [15]: M_{GK} is an interesting measure, which allows the extraction of negative rules.

$$M_{GK} = \frac{p(Y/X)-p(Y)}{1-p(Y)}, \quad \text{if } X \text{ favorise } Y$$
$$M_{GK} = \frac{p(Y/X)-p(Y)}{p(Y)}, \quad \text{if } X \text{ defavorise } Y$$

It takes into account several situations of references: in the case where the rule is situated in the attractive zone (i.e. $p(Y/X) > p(Y)$), this measure evaluates the distance between independence and logical implication. Thus, the higher the value of M_{GK} is close to 1, the more the rule is close to the logical implication and the higher the value of M_{GK} is close to 0, the more the rule is close to the independence. In the case where the rule is located in the repulsive zone (i.e. $p(Y/X) < p(Y)$), M_{GK} evaluates this time a distance between the independence and the incompatibility. Thus, the closer the value of M_{GK} is to -1, the more similar to incompatibility the rule is; and the closer the value of M_{GK} is to 0, the closer to the independence the rule is.

As was mentioned above, several studies [35], [23], [25], [13] were reported in the literature on the various properties of interestingess measures to be able to characterize and evaluate the interestingness measures. The main goal of researchers in the domain is then to provide a user assistance in choosing the best interestingness measure meeting his needs. For that, formal properties have been developed [32], [24], [35], [12], [4] in order to evaluate the interestingness measures and to help users understanding their behavior. In the following, we present nineteen properties reported in the literature.

3.2 Properties of the measures

Figure 1 lists 19 properties of interestingness measures. The properties are described in detail in [16]; we omit details due to lack of space.

The authors in [14] proposed an evaluation of 61 interestingness measures according to the 19 properties (P_3 to P_{21}). Properties P_1 and P_2 were not taken into account in this study because of their subjective character. The measures and their properties result in a binary measure-property matrix that is used for clustering the measures according to their properties. The clustering performed in [14] using the agglomerative hierarchical method and the K-means method revealed 7 clusters of measures which will be used in the next section in a comparison with the results obtained by Boolean factor analysis applied on the same measure-property matrix.

3.3 Clustering using Boolean factors

The measure-property matrix describing interestingness measures by their properties is depicted in Figure 2. It consists of 62 measures (61 measures from [14] plus one more that has been studied recently) described by 21 properties because the three-valued property P_{14} is represented by three yes-no properties

No.	Property	Ref.
P_1	Intelligibility or comprehensibility of measure	[25]
P_2	Easiness to fix a threshold to the rule	[23]
P_3	Asymmetric measure.	[35], [23]
P_4	Asymmetric measure in the sense of the conclusion negation.	[23], [35]
P_5	Measure assessing in the same way $X \rightarrow Y$ and $\bar{Y} \rightarrow \bar{X}$ in the logical implication case.	[23]
P_6	Measure increasing function the number of examples or decreasing function the number of counter-examples.	[32], [23]
P_7	Measure increasing function the data size.	[12], [35]
P_8	Measure decreasing function the consequent/antecedent size.	[23], [32]
P_9	Fixed value a in the independence case.	[23], [32]
P_{10}	Fixed value b in the logical implication case.	[23]
P_{11}	Fixed value c in the equilibrium case.	[5]
P_{12}	Identified values in the attraction case between X and Y.	[32]
P_{13}	Identified values in the repulsion case between X and Y.	[32]
P_{14}	Tolerance to the first counter-example.	[23], [38]
P_{15}	Invariance in case of expansion of certain quantities.	[35]
P_{16}	Desired relationship between $X \rightarrow Y$ and $\bar{X} \rightarrow Y$ rules.	[35]
P_{17}	Desired relationship between $X \rightarrow Y$ and $X \rightarrow \bar{Y}$ antinomic rules.	[35]
P_{18}	Desired relationship between $X \rightarrow Y$ and $\bar{X} \rightarrow \bar{Y}$ rules.	[35]
P_{19}	Antecedent size is fixed or random.	[23]
P_{20}	Descriptive or statistical measure.	[23]
P_{21}	Discriminant measure.	[23]

Fig. 1. Interestingness measures properties.

$P_{14.1}$, $P_{14.2}$, and $P_{14.3}$. We computed the decomposition of the matrix using Algorithm 2 from [3] and obtained 28 factors (as in the case below, several of them may be disregarded as not very important; we leave the details for a full version of this paper). In addition, we extended the original 62×21 binary matrix by adding for every property its negation, and obtained a 62×42 binary matrix. The reason for adding negated properties is due to our goal to compare the results with the two clustering methods mentioned above and the particular role of the properties and their negations in these clustering methods. From the 62×42 matrix, we obtained 38 factors, denoted $F_1, \ldots, F_{38}$. The factors are presented in Figures 3 and 4. Figure 3 depicts the object-factor matrix describing the interestingness measures by factors, Figure 4 depicts the factor-property matrix explaining factors by properties of measures. Factors are sorted from the most important to the least important, where the importance is determined by the number of 1s in the input measure-property matrix covered by the factor [3]. The first factors cover a large part of the matrix, while the last ones cover only a small part and may thus be omitted [3], see the graph of cumulative cover of the matrix by the factors in Figure 5.

4 Interpretation and comparison to other approaches

The aim of this section is to provide an interpretation of the results described in the previous section and compare them to the results already reported in the literature, focusing mainly on [14]. As was described in the previous section, 38 factors were obtained. The first 21 of them cover 94 % of the input measure-property matrix (1s in the matrix), the first nine cover 72 %, and the first five

	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14.1	P15	P16	P17	P18	P19	P20	P21	P14.2	P14.3
correlation	0	1	1	1	1	1	1	0	0	1	1	0	0	1	1	1	0	0	1	1	0
Cohen	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	1	1	0
confidence	1	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0
causal confidence	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0
Pavillon	1	1	0	1	1	1	1	0	0	1	1	0	0	0	1	0	0	0	1	1	0
Ganascia	1	1	1	1	0	0	0	1	1	0	0	0	0	0	1	0	0	0	1	1	0
causal confirmation	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
descriptive confirmation	1	1	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	1	0
conviction	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	1	0	1
cosine	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
coverage	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
dependency	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0
causal dependency	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
weighted dependency	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Bayes factor	1	1	0	1	1	1	1	0	0	1	1	0	1	0	0	0	0	0	1	0	1
Loevinger	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	1	1	0
collective strength	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	1	0	1
Fukuda	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
information gain	0	1	0	1	1	1	1	0	0	1	1	1	0	0	0	0	0	0	1	0	0
Goodman	0	1	1	1	1	0	1	1	0	1	1	0	0	1	1	1	0	0	1	1	0
implication index	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0
IPEE	1	1	1	1	0	0	0	0	1	0	0	1	0	0	0	0	1	1	0	0	0
IP3E	1	1	1	1	0	0	0	0	1	0	0	1	0	0	0	0	1	1	1	0	0
PDI	0	1	1	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0
II	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	0	1	1	0	0	0
EII	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	0
REII	1	1	1	1	1	0	1	0	0	1	1	1	0	0	0	0	1	1	1	0	0
likelihood index	0	1	1	1	1	1	1	0	0	1	1	1	0	0	0	0	1	1	0	0	0
interest	0	1	0	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1	0
Jaccard	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
Jmeasure	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	1
Klosgen	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	1
Laplace	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0
Mgk	1	1	1	1	1	0	1	1	0	1	1	0	0	0	1	0	0	0	1	1	0
least contradiction	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0
Pearl	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	1	0
Piatetsky-Shapiro	0	1	1	1	1	1	1	0	0	1	1	0	0	1	1	1	0	1	1	1	0
precision	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0
prevalence	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
YuleQ	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	0	1	0	0
recall	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
Gini	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1
relative risk	1	1	0	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	1	0	1
Sebag	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1
support	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
one way support	1	1	0	0	1	1	1	0	0	1	1	0	0	0	0	0	0	0	1	0	1
two way support	0	1	0	0	1	1	1	0	0	1	1	0	0	0	0	0	0	0	1	0	1
examples and counter-examples rate	1	1	1	1	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0
VT100	0	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	0	1	1	0
variation support	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	1
YuleY	0	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	0	0	1	0	1
Zhang	1	1	1	1	1	0	1	1	0	1	1	1	1	0	1	0	0	0	1	0	0
causal confirmed confidence	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0
Czekanowski	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
negative reliability	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0
mutual information	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1
Kulczynski	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
Leverage	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
novelty	0	1	1	1	1	1	1	0	0	1	1	0	0	1	1	1	0	0	1	1	0
odds ratio	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	1	0	1
specificity	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
causal support	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0

Fig. 2. Input binary matrix describing interestingness measures by their properties.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30	F31	F32	F33	F34	F35	F36	F37	F38
correlation	1	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0
Cohen	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
confidence	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
causal confidence	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
Pavillon	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Ganascia	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0
causal confirmation	0	1	0	0	0	1	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
descriptive confirmation	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
conviction	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
cosine	0	1	0	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
coverage	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
dependency	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
causal dependency	0	1	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
weighted dependency	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1
Bayes factor	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Loevinger	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
collective strength	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1
Fukuda	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0
information gain	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Goodman	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	1	0
implication index	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
IPEE	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	0	0	1	0	1	0	0
IP3E	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0
PDI	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	1	0
II	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0
EII	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0
REII	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
likelihood index	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	1	1	0
interest	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
Jaccard	0	0	0	0	1	0	1	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Jmeasure	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1
Klosgen	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
Laplace	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Mgk	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
least contradiction	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Pearl	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0
Piatetsky-Shapiro	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0
precision	0	1	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0
prevalence	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
YuleQ	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1
recall	0	1	0	0	1	0	0	0	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Gini	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1
relative risk	1	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Sebag	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
support	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
one way support	1	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
two way support	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1
examples and counter-examples rate	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1
VT100	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	1	1	0
variation support	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1
YuleY	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1
Zhang	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1
causal confirmed confidence	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
Czekanowski	0	1	0	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
negative reliability	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
mutual information	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Kulczynski	0	0	0	0	1	0	1	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Leverage	0	1	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
novelty	1	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0
odds ratio	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1
specificity	0	1	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
causal support	0	1	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0

Fig. 3. Interestingness measures described by factors obtained by decomposition of the input matrix from Figure 2 extended by negated properties.

	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14.1	P15	P16	P17	P18	P19	P20	P21	P14.2	P14.3	$\overline{P3}$	$\overline{P4}$	$\overline{P5}$	$\overline{P6}$	$\overline{P7}$	$\overline{P8}$	$\overline{P9}$	$\overline{P10}$	$\overline{P11}$	$\overline{P12}$	$\overline{P13}$	$\overline{P14.1}$	$\overline{P15}$	$\overline{P16}$	$\overline{P17}$	$\overline{P18}$	$\overline{P19}$	$\overline{P20}$	$\overline{P21}$	$\overline{P14.2}$	$\overline{P14.3}$
F1	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0
F2	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	1	1	1	1	1	0	0	1	1	0	0	1
F3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	0	1	1	1	1	1	0	1	0	0	0	0
F4	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	1	1	0	1	0	0	0	0	0
F5	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	1	1	1	1	1	0	0	0
F6	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	1
F7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	0	1	1	0	1	0
F8	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0	1	1
F9	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
F10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	1	1	0	0	0	1
F11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	1	1	0	0	0	0	0	0
F12	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1
F13	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
F14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1	1	1	1	1	0	0	1	1	0	0	0
F15	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	1	0
F16	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
F17	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	0	0	0
F18	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
F19	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
F20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	1	0	1	1	0	0	0	0
F21	0	1	0	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
F22	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	1	1
F23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0
F24	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	0	0	0
F25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	1
F26	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	1
F27	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	0	0	0	0	0	0	0	1
F28	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	1	0	0	1	1	1	1	0	1	0	0	0	0
F29	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	0	1	1	1	0	0	0	0	1
F30	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	1	1	0	0	0	0	0	0	0
F31	0	1	1	1	1	0	1	1	0	1	1	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	1	0	1	1
F32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	1
F33	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	1	0	1	0
F34	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
F35	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
F36	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
F37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
F38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0

Fig. 4. Factors obtained by decomposition of the input matrix from Figure 2 extended by negated properties. The factors are described in terms of the original and negated properties.

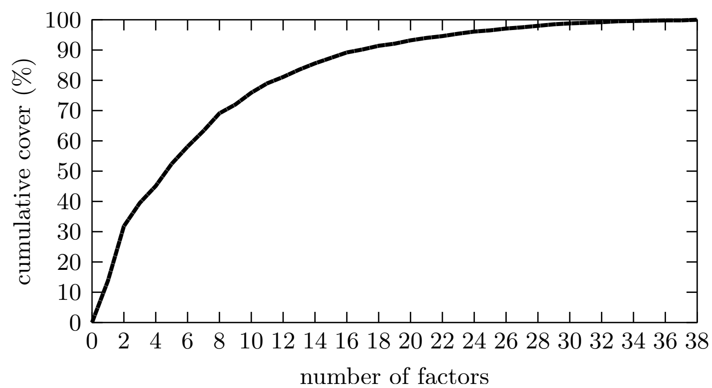

Fig. 5. Cumulative cover of input matrix from Figure 2 extended by negated properties by factors obtained by decomposition of the matrix.

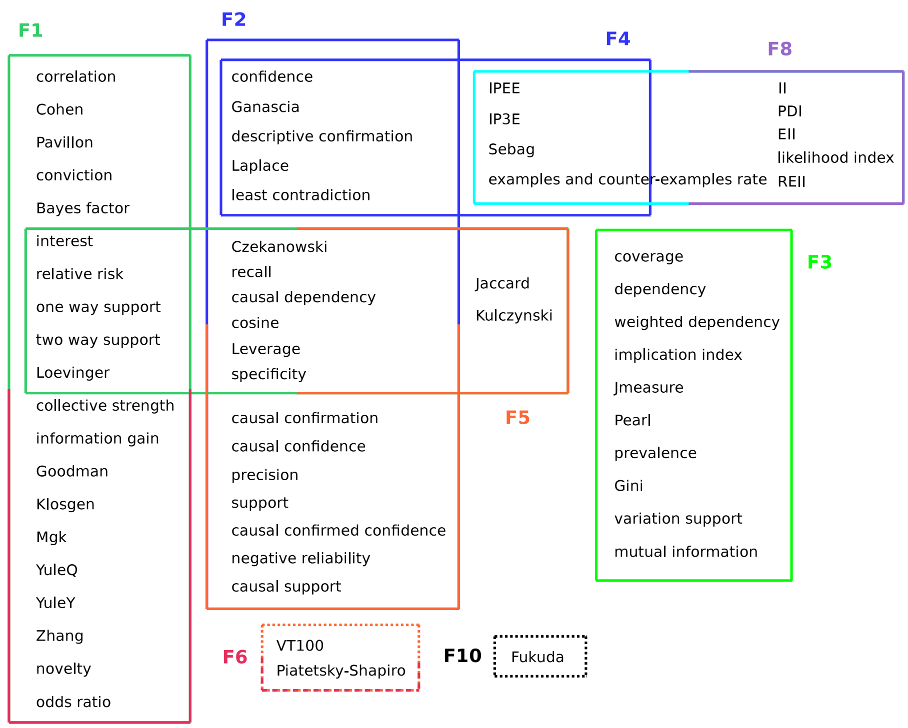

Fig. 6. Venn diagram of the first five factors (plus the eighth and part of the sixth and tenth to cover the whole set of measures) obtained by decomposition of the input matrix from Figure 2 extended by negated properties.

cover 52.4 %. Another remark is that the first ten factors cover the whole set of measures.

Note first that the Boolean factors represent overlapping clusters, contrary to the clustering using the agglomerative hierarchical method and the K-means method performed in [14]. Namely, the clusterings are depicted in Figure 6 describing the Venn diagram of the first five Boolean factors (plus the eighth and part of the sixth and tenth to cover the whole set of measures) and Figure 7, which is borrowed from [14], describing the consensus on the classification obtained by the hierarchical and K-means clusterings. This consensus refunds the classes C_1 to C_7 of the extracted measures, which are common to both techniques.

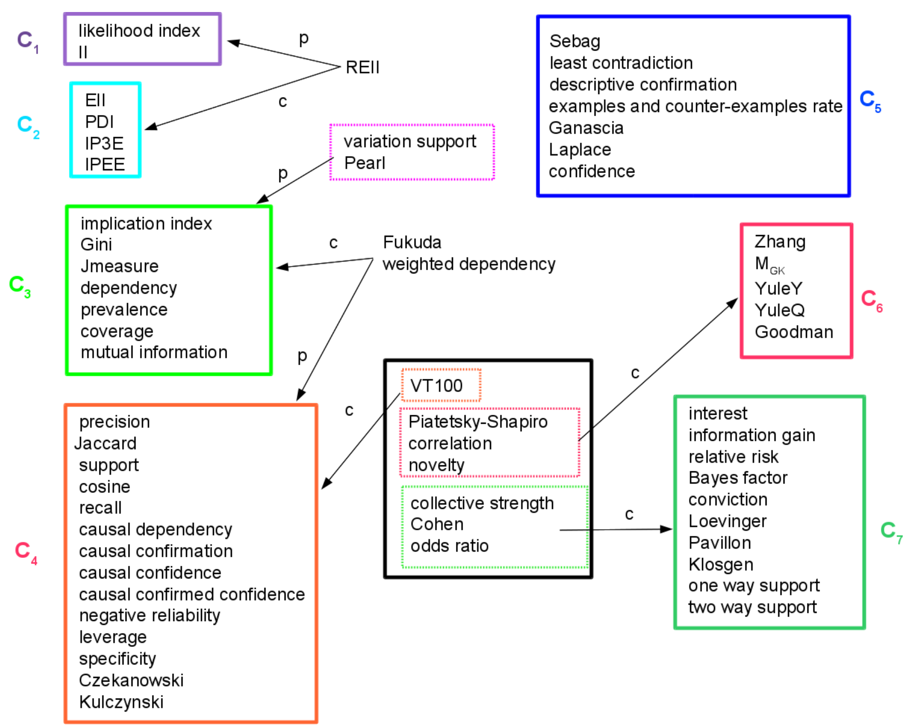

Fig. 7. Classes of measures obtained by the hierarchical and K-means clusterings.

Due to lack of space, we focus on the first four factors since they cover nearly half of the matrix (45.1 %), and also because most of the measures appear at least once in the four factors.

Factor 1. The first factor F_1 applies to 20 measures, see Figure 3, namely: correlation, Cohen, Pavillon, conviction, Bayes factor, Loevinger, collective strength, information gain, Goodman, interest, Klosgen, Mgk, YuleQ, relative risk, one way support, two way support, YuleY, Zhang, novelty, and odds ratio. These measures share the following 9 properties: P4, P7, P9, not P11, P12, P13, not P19, not P20, P21, see Figure 4.

Interpretation. The factor applies to measures whose evolutionary curve increases w.r.t the number of examples and have a fixed point in the case of independence (this allows to identify the attractive and repulsive area of a rule). The factor also applies only to descriptive and discriminant measures that are not based on a probabilistic model.

Comparison. When looking at the classification results reported in [14], F_1 covers two classes from [14]: C_6 and C_7, which together contain 15 measures. Those classes are closely related within the dendrogram obtained with the agglomerative hierarchical clustering method used in [14]. The 5 missing measures form a class obtained with K-means method in [14] with Euclidian distance.

Factor 2. F_2 applies to 18 measures, namely: confidence, causal confidence, Ganascia, causal confirmation, descriptive confirmation, cosine, causal dependency, Laplace, least contradiction, precision, recall, support, causal confirmed confidence, Czekanowski, negative reliability, Leverage, specificity, and causal support. These measures share the following 11 properties: P4, P6, not P9, not P12, not P13, P14.2, not P15, not P16, not P19, not P20, P21.

Interpretation. The factor applies to measures whose evolutionary curve increases w.r.t. the number of examples and has a variable point in the case of independence, which implies that the attractive and repulsive areas of a rule are not identifiable. The factor also applies only to measures that are not discriminant, are indifferent to the first counter-examples, and are not based on a probabilistic model.

Comparison. F_2 corresponds to two classes, C_4 and C_5 reported in [14]. $C_4 \cup C_5$ contains 22 measures. The missing measures are: Jaccard, Kulczynski, examples and counter-examples rate and Sebag. Those measures are not covered by F_2 since they are not indifferent to the first counter-examples.

Factor 3. F_3 applies to 10 measures, namely: coverage, dependency, weighted dependency, implication index, Jmeasure, Pearl, prevalence, Gini, variation support, and mutual information. These measures share the following 10 properties: not P6, not P8, not P10, not P11, not P13, not P14.1, not P15, not P16, not P17, not P19.

Interpretation. The factor applies to measures whose evolutionary curve does not increase w.r.t. the number of examples.

Comparison. F_3 corresponds to class C_3 reported in [14], which contains 8 measures. The two missing measures, variation support and Pearl, belong to the same classes obtained by both K-means and the hierarchical method. Moreover, these two missing measures are similar to those from C_3 obtained by the hierarchical method since they merge with the measures in C_3 at the next level of the generated dendrogram. Here, there is a strong correspondence between results obtained using Boolean factors and the ones reported in [14].

Factor 4. F_4 applies to 9 measures, namely: confidence, Ganascia, descriptive confirmation, IPEE, IP3E, Laplace, least contradiction, Sebag, and examples and counter-examples rate. These measures share the following 12 properties: P3, P4, P6, P11, not P7, not P8, not P9, not P12, not P13, not P15, not P16, not P18.

Interpretation. The factor applies to measures whose evolutionary curve increases w.r.t. the number of examples and has a fixed value in the equilibrium case. As there is no fixed value in the independence case, we can not get an identifiable area in the case of attraction or repulsion.

Comparison. F_4 mainly applies to measures of class C_5 obtained in [14]. The two missing measures, IPEE et IP3E, belong to a different class.

5 Conclusions and further issues

We demonstrated that Boolean factors provide us with clearly interpretable meaningful clusters of measures among which the first ones are highly similar to other clusters of measures reported in the literature. Contrary to other clustering methods, Boolean factors represent overlapping clusters. We consider this an advantage because overlapping clusters are a natural phenomenon in human classification. We presented preliminary results on clustering the measures using Boolean factors. Due to limited scope, we presented only parts of the results obtained and leave other results for a full version of this paper.

An interesting feature of the presented method, to be explored in the future, is that the method need not start from scratch. Rather, one or more clusters, that are considered important classes of measures, may be supplied at the start and the method may be asked to complete the clustering. Another issue left for future research is the benefit of the clustering of measures for a user who is interested in selecting a type of measure, rather than a particular measure of interestingness of association rules. In the intended scenario, a user may use various interestingness measures that belong to different classes of measures.

References

1. Agrawal R., Imielinski T., Swami A.: Mining association rules between sets of items in large databases. *Proc. ACM SIGMOD* 1993, 207–216.
2. Agrawal R., Srikant R.: Fast algorithms for mining association rules. *Proc. VLDB Conf.* 1994, 478–499.
3. Belohlavek R., Vychodil V.: Discovery of optimal factors in binary data via a novel method of matrix decomposition. *J. of Computer and System Sciences* **76**(1)(2010), 3–20.
4. Blanchard J., Guillet F., Briand H., Gras R.: Assessing rule with a probabilistic measure of deviation from equilbrium. In *Proc. Of 11th International Symposium on Applied Stochastic Models and Data Analysis ASMDA* 2005, Brest, France, 191–200.
5. Blanchard J., Guillet F., Briand H., Gras R.: IPEE: Indice Probabiliste d'Écart à l'Équilibre pour l'évaluation de la qualité des règles. *Dans l'Atelier Qualité des Données et des Connaissances* 2005, 26–34.
6. Brin S., Motwani R., Silverstein C.: Beyond Market Baskets: Generalizing Association Rules to Correlations. In Proc. of the ACM SIGMOD Conference, Tucson, Arizona, 1997, 265–276.

7. Carpineto C., Romano G.: *Concept Data Analysis. Theory and Applications.* J. Wiley, 2004.
8. Davey B. A., Priestley H.: *Introduction to Lattices and Order.* Cambridge University Press, Oxford, 1990.
9. Delgado M., Ruiz D.-L., Sanchez D.: Studying Interest measures for association rules through a logical model. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **18**(1)(2010), World Scientific, 87–106.
10. Feno D.R.: Mesures de qualité des règles d'association: normalisation et caractérisation des bases. PhD thesis, Université de La Réunion, 2007.
11. Ganter B., Wille R.: *Formal Concept Analysis. Mathematical Foundations.* Springer, Berlin, 1999.
12. Geng L., Hamilton H.J.: Choosing the Right Lens: Finding What is Interesting in Data Mining. *Quality Measures in Data Mining* 2007, ISBN 978-3-540-44911-9, 3–24.
13. Geng L., Hamilton H. J.: Interestingness measures for data mining: A Survey. *ACM Comput. Surveys* **38**(3)(2006), 1–31.
14. Guillaume S., Grissa D., Mephu Nguifo E.: Catégorisation des mesures d'intérêt pour l'extraction des connaissances. *Revue des Nouvelles Technologies de l'Information*, 2011, to appear (previously available as Technical Report RR-10-14, LIMOS, ISIMA, 2010).
15. Guillaume S.: Traitement des données volumineuses. Mesures et algorithmes d'extraction des règles d'association et règles ordinales. PhD thesis. Université de Nantes, France, 2000.
16. Guillaume S., Grissa D., Mephu Nguifo E.: Propriétés des mesures d'intérêt pour l'extraction des règles. *Dans l'Atelier Qualité des Données et des Connaissances, EGC'2010*, 2010, Hammamet-Tunisie, *http://qdc2010.lri.fr/fr/actes.php*, 15–28.
17. Grissa D., Guillaume S., Mephu Nguifo E.: Combining Clustering techniques and Formal Concept Analysis to characterize Interestingness Measures. *CoRR* **abs/1008.3629**, 2010.
18. Hájek P., Havránek T.: *Mechanizing Hypotheses Formation.* Springer, 1978.
19. Hájek P., Holeňa, Rauch J.: The GUHA method and its meaning for data mining. *J. Computer and System Sciences* **76**(2010), 34–48.
20. Hilderman R. J., Hamilton H. J.: Knowledge Discovery and Measures of Interest, *Volume 638 of The International Series in Engineering and Computer Science* **81**(2)(2001), Kluwer.
21. Huynh X.-H., Guillet F., Briand H.: Clustering Interestingness Measures with Positive Correaltion. **ICEIS** (2) (2005), 248–253.
22. Heravi M. J., Zaïane O. R.: A study on interestingness measures for associative classifiers. **SAC** (2010), 1039–1046.
23. Lallich S., Teytaud, O.: Évaluation et validation de mesures d'intérêt des règles d'association. *RNTI-E-1, numéro spécial* 2004, 193–217.
24. Lenca P, Meyer P., Picouet P., Vaillant B., Lallich S.: Critères d'évaluation des mesures de qualité en ecd. *Revue des Nouvelles Technologies de l'Information (Entreposage et Fouille de données)* (1)(2003), 123–134.
25. Lenca P., Meyer P., Vaillant B., Lallich, S.: A multicriteria decision aid for interestingness measure selection. Technical Report LUSSI-TR-2004-01-EN, Dpt. LUSSI, ENST Bretagne 2004 (chapter 1).
26. Liu J., Mi J.-S.: A novel approach to attribute reduction in formal concept lattices. RSKT 2006, *Lecture Notes in Artificial Intelligence* **4062** (2006), 522–529.

27. Maddouri M., Gammoudi J.: On Semantic Properties of Interestingness Measures for Extracting Rules from Data. *Lecture Notes in Computer Science* **4431** (2007), 148–158.
28. Maier D.: *The Theory of Relational Databases.* Computer Science Press, Rockville, 1983.
29. Pawlak Z.: Rough sets. *Int. J. Information and Computer Sciences* **11**(5)(1982), 341–356.
30. Pawlak Z.: *Rough Sets: Theoretical Aspcets of Reasoning About Data.* Kluwer, Dordrecht, 1991.
31. Pearson K.: Mathematical contributions to the theory of evolution, regression, heredity and panmixia. Philosophical Trans. of the Royal Society A (1896).
32. Piatetsky-Shapiro G.: Discovery, Analysis and Presentation of Strong Rules. In G. Piatetsky-Shapiro & W.J. Frawley, editors: *Knowledge Discovery in Databases.* AAAI Press, 1991, 229–248.
33. Polkowski L.: *Rough Sets: Mathematical Foundations.* Springer, 2002.
34. Sese J., Morishita S.: Answering the most correlated n association rules efficiently. In *Proceedings of the 6th European Conf on Principles of Data Mining and Knowledge Discovery* 2002, Springer-Verlag, 410–422.
35. Tan P.-N., Kumar V., Srivastava J.: Selecting the right objective measure for association analysis. *Information Systems* **29**(4)(2004), 293–313.
36. Tan P.-N., Steinbach M., Kumar V.: *Introduction to Data Mining.* Addison-Wesley, 2005.
37. Vaillant B., Lenca P., Lallich S.: A Clustering of Interestingness Measures. *DS'04, the 7th International Conference on Discovery Science* **LNAI 3245** (2004), 290–297.
38. Vaillant B.: Mesurer la qualité des règles d'association: études formelles et expérimentales. PhD thesis, ENST Bretagne, 2006.
39. Wang X., Ma J.: A novel approach to attribute reduction in concept lattices. RSKT 2006, *Lecture Notes in Artificial Intelligence* **4062** (2006), 522–529.
40. Wille R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival I.: *Ordered Sets.* Reidel, Dordrecht, Boston, 1982, 445–470.
41. Zhang W.-X., Wie L., Qi J.-J.: Attribute reduction in concept lattices based on discernibility matrix. RSFDGrC 2005, *Lecture Notes in Artificial Intelligence* **3642** (2005), 157–165.