DISCOVERING THE STABLE CLUSTERS BETWEEN INTERESTINGNESS MEASURES

Xuan-Hiep Huynh, Fabrice Guillet, Henri Briand
LINA FRE CNRS 2729 - Polytechnic school of Nantes university
La Chantrerie BP 50609 44306 Nantes Cedex 3, France

Keywords: interestingness measure, stable cluster, post-processing, association rules, knowledge quality.

Abstract: In this paper, dealing with association rules post-processing, we propose to study the correlations between 36 interestingness measures (IM), in order to better understand their behavior on data and finally to help the data miner chooses the best IMs. We used two datasets with opposite characteristics in which we extract two rulesets about 100000 rules, and the two subsets of the 1000 best rules according to IMs. The study of the correlation between IMs with PAM and AHC shows unexpected stabilities between the four ruleset, and more precisely eight stable clusters of IMs are found and described.

1 INTRODUCTION

In the framework of data mining, association rules is a key tool aiming at discovering interesting patterns in data. Unfortunately, it often delivers a prohibitive number of rules in which the data miner (or a user) must find the most interesting ones. In order to help him/her during this post-processing step, many IMs have been proposed and studied in the literature (Agrawal et al., 1996) (Gras et al., 1996) (Hilderman and Hamilton, 2001) (Tan et al., 2004) (Blanchard et al., 2005b).

In this paper, we propose a new approach to evaluate the behavior of 36 objective IMs proposed in the literature. We aim at finding the stable clusters representing the different aspects existing in the datasets via the evaluation of the behavior of IMs. Two new views are proposed to evaluate : (1) the strong relation between IMs and (2) the relative distance between clusters of IMs. The results of this approach is interesting to validate the quality of knowledge discovered in form of association rules and to help the user differentiates natural aspects existing from the datasets.

The paper is organized as follows. In Section 2, we introduce some related works on knowledge quality. Section 3 introduces the computation of interestingness by presenting the two techniques for analyzing the datasets and the calculation of the dissimilarity between IMs. Section 4 presents the data preparations and 36 used IMs to analyze. Then, we discuss the important results obtained from the evaluation of IM behavior on two original datasets and two sets of best rules extracted.

2 RELATED WORKS

2.1 Evaluation of IM Properties

To discover the principles of a good IM, many authors have examined some properties of the interestingness of association patterns. (Piatetsky-Shapiro, 1991) introduced three principles for an association rule $a \rightarrow b$: ”P1” 0 value when a and b are independent, ”P2” monotonically increasing with $a \cap b$, ”P3” monotonically decreasing with a or b. (Major and Magano, 1995) proposed a property ”P4” monotonically increasing with $a \cap b$ when the confidence value $\frac{p(a \cap b)}{pa}$ is fixed. (Freitas, 1999) evaluated a property ”P5” (asymmetry) if $i(a \rightarrow b) \neq i(b \rightarrow a)$. (Klösgen, 1996) gave four axioms : $Q(a,b) = 0$ if a and b are statistically independent, $Q(a,b)$ monotonically increases in $confidence(a \rightarrow b)$ for a fixed $support(a)$, $Q(a,b)$ monotonically decreases in $support(a)$ for a fixed $support(a \cap b)$, $Q(a,b)$ monotonically increases in $support(a)$ for a fixed $confidence(a \rightarrow b) > support(b)$. (Hilderman and Hamilton, 2001) proposed five principles: minimum

Huynh X., Guillet F. and Briand H. (2006).
DISCOVERING THE STABLE CLUSTERS BETWEEN INTERESTINGNESS MEASURES.
In *Proceedings of the Eighth International Conference on Enterprise Information Systems - AIDSS*, pages 196-201
DOI: 10.5220/0002493701960201

value, maximum value, skewness, permutation invariance, transfer. (Tan et al., 2004) defined five interestingness principles: symmetry under variable permutation, row/column scaling invariance, anti-symmetry under row/column permutation, inversion invariance, null invariance.

2.2 Comparison of IMs

Some researches are also interested in making comparisons between IMs.

(Gavrilov et al., 2000) studied the similarity between the IMs for classifying them.

(Hilderman and Hamilton, 2001) proposed five principles for ranking summaries generated from databases, and performed a comparative analysis of sixteen diversity IMs to determine which ones satisfy the proposed principles. The objective of this work is to gain some insight into the behavior that can be expected from each of the IMs in practice.

(Tan et al., 2004) introduced twenty-one IMs using Pearson's correlation and has found two situations in which the IMs may become consistent with each other, namely, the support-based pruning or table standardization are used. In addition, they also used five proposed interestingness properties to capture the utility of an objective IM in terms of analyzing k-way contingency tables.

(Carvalho et al., 2003) (Carvalho et al., 2005) evaluated eleven objective IMs for ranking them by their functionality of interesting effective of a decider.

(Choi et al., 2005) used an approach of multi-criteria decision aide for finding the best association rules.

(Blanchard et al., 2005a) classified eighteen objective IMs in four groups according to three criteria: independent, equilibrium, characteristic descriptive or statistic.

(Huynh et al., 2005) proposed a classification approach by correlation graph that can identify eleven classes on thirty-four IMs.

3 INTERESTINGNESS COMPUTATION

3.1 Techniques for Analyzing the Datasets

We use two data analysis techniques to illustrate: agglomerative hierarchical clustering (AHC) and partitioning around medoids (PAM) (Kaufman and Rousseeuw, 1990). Each of these techniques is used as a means for achieving the results.

These two techniques are used with a $q \times q$ dissimilarity matrix, where $d(i, j) = d(j, i)$, measuring the difference or dissimilarity between two IMs m_i and m_j. AHC, is used in our work, finds the most similar clusters according to the average linkage method. PAM, is more robust than the k-means method, is to find a subset $m_1, m_2, ..., m_k \subset 1, ..., q$ which minimizes the objective function $\sum_{i=1}^{q} min_{t=1,...,k} d(i, m_t)$.

3.2 Dissimilarity Between IMs

Let $R(D) = \{r_1, r_2, ..., r_p\}$ denote input data as a set of p association rules derived from a dataset D. Each rule $a \to b$ is described by its itemsets (a, b) and its cardinalities $(n, n_a, n_b, n_{a\bar{b}})$.

Let M be the set of q available IMs for our analysis $M = \{m_1, m_2, ..., m_q\}$. Each IM is a numerical function on rule cardinalities: $m(a \to b) = f(n, n_a, n_b, n_{a\bar{b}})$.

For each IM $m_i \in M$, we can construct a vector $m_i(R) = \{m_{i1}, m_{i2}, ..., m_{ip}\}$, $i = 1..q$, where m_{ij} corresponds to the calculated value of the IM m_i for a given rule r_j.

The matrix ($p \times q$) of interestingness values:

$$\mathbf{m} = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1q} \\ m_{21} & m_{22} & \dots & m_{2q} \\ \dots & \dots & \dots & \dots \\ m_{p1} & m_{p2} & \dots & m_{pq} \end{pmatrix}$$

The correlation value between any two IMs m_i, m_j, $\{i, j = 1..q\}$ on the ruleset R will be calculated by using a Pearson's correlation coefficient $\rho(m_i, m_j)$ (Ross, 1987), where $\overline{m_i}, \overline{m_j}$ are the average calculated values of vector $m_i(R)$ and $m_j(R)$ respectively.

Definition 1. The dissimilarity d between two IMs m_i,m_j is defined by:

$$d(m_i, m_j) = 1 - |\rho(m_i, m_j)|$$

where: $\rho(m_i, m_j) =$

$$\frac{\sum_{k=1}^{p}[(m_{ik} - \overline{m_i})(m_{jk} - \overline{m_j})]}{\sqrt{[\sum_{k=1}^{p}(m_{ik} - \overline{m_i})^2][\sum_{k=1}^{p}(m_{jk} - \overline{m_j})^2]}}$$

As correlation is symmetrical, the $q(q-1)/2$ dissimilarity values can be stored in one half of a matrix $q \times q$.

$$\mathbf{d} = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1q} \\ d_{21} & d_{22} & \dots & d_{2q} \\ \dots & \dots & \dots & \dots \\ d_{q1} & d_{q2} & \dots & d_{qq} \end{pmatrix}$$

with: $d_{ii} = 0$, and $\forall (i, j), i \neq j, d_{ij} \geq 0, d_{ij} = d_{ji}$

4 DATA PREPARATION AND USED MEASURES

4.1 Dataset

To facilitate the evaluation of stable clusters, we use two opposite datasets D_1 and D_2 to discover the interactions between studied IMs. The categorical mushroom dataset (D_1) comes from the Irvine machine-learning database repository (Newman et al., 1998) and a synthetic dataset (D_2). The latter is obtained by stimulating the transactions of customers in retail businesses (Agrawal et al., 1996). We also generate the set of association rules (ruleset) R_1 (resp. R_2) from the dataset D_1 (resp. D_2) using the the algorithm Apriori (Agrawal et al., 1996). R_2 has the typical characteristic of the Agrawal dataset T5.I2.D10k (T5: average size of the transactions is 5, I2: average size of the maximal potentially large itemsets is 2, D10k: number of items is 100). For an evaluation of the IM behavior of the ”best rules” from these two rulesets, we extracted $R_1^{'}$ (resp. $R_2^{'}$) from R_1 (resp. R_2) as the union of the first 1000 best rules ($\approx 1\%$, descending with interestingness values) issued from each IM (see Tab. 1).

Table 1: Description of the datasets.

Dataset	Items (Avg. length)	Transactions	Number of rules	$R(D)$
D_1	118 (22)	8416	123228	R_1
			10431	$R_1^{'}$
D_2	81 (5)	9650	102808	R_2
			7452	$R_2^{'}$

4.2 Used Measures

Many IMs can be found in the literature (Hilderman and Hamilton, 2001) (Tan et al., 2004). We added this list with four IMs: implication intensity (II), (Gras et al., 1996), entropic implication intensity (EII(α)), (Blanchard et al., 2003), information ratio modulated by contrapositive (TIC) (Blanchard et al., 2005b) and probabilistic index of deviation from equilibrium (IPEE) (Blanchard et al., 2005a) (see Tab. 4.2).

5 RESULTS

For discovering the stable clusters of IMs, two specific views are introduced : the strong relation and the relative distance between IMs. These two view are applied to four matrix of dissimilarity calculated from the four rulesets $R_1, R_1^{'}, R_2, R_2^{'}$ respectively. The result obtained is interesting to differentiate the aspects existing in the datasets or *the stable behaviors of IMs*. Two techniques AHC and PAM are used for each of these views respectively.

Table 2: IMs determining by negative examples.

N°	Interestingness Measure	$f(n, n_a, n_b, n_{a\bar{b}})$
0	Causal Confidence	$1 - \frac{1}{2}(\frac{1}{n_a} + \frac{1}{n_{\bar{b}}})n_{a\bar{b}}$
1	Causal Confirm	$\frac{n_a+n_{\bar{b}}-4n_{a\bar{b}}}{n}$
2	Causal Confirmed-Confidence	$1 - \frac{1}{2}(\frac{3}{n_a} + \frac{1}{n_{\bar{b}}})n_{a\bar{b}}$
3	Causal Support	$\frac{n_a+n_{\bar{b}}-2n_{a\bar{b}}}{n}$
4	Collective Strength	$\frac{(n_a-n_{a\bar{b}})(n_{\bar{b}}-n_{a\bar{b}})(n_a n_{\bar{b}}+n_b n_{\bar{a}})}{(n_a n_b+n_{\bar{a}} n_{\bar{b}})(n_b-n_a+2n_{a\bar{b}})}$
5	Confidence	$1 - \frac{n_{a\bar{b}}}{n_a}$
6	Conviction	$\frac{n_a n_{\bar{b}}}{n n_{a\bar{b}}}$
7	Cosine	$\frac{n_a-n_{a\bar{b}}}{\sqrt{n_a n_b}}$
8	Dependency	$\mid\frac{n_{\bar{b}}}{n} - \frac{n_{a\bar{b}}}{n_a}\mid$
9	Descriptive Confirm	$\frac{n_a-2n_{a\bar{b}}}{n}$
10	Descriptive Confirmed-Confidence	$1 - 2\frac{n_{a\bar{b}}}{n_a}$
11	EII ($\alpha = 1$)	$\sqrt{\varphi \times I^{\frac{1}{2\alpha}}}$
12	EII ($\alpha = 2$)	$\sqrt{\varphi \times I^{\frac{1}{2\alpha}}}$
13	Example & Contra-Example	$1 - \frac{n_{a\bar{b}}}{n_a-n_{a\bar{b}}}$
14	Gini-index	$\frac{(n_a-n_{a\bar{b}})^2+n_{a\bar{b}}^2}{n n_a} + \frac{(n_b-n_a+n_{a\bar{b}})^2+(n_{\bar{b}}-n_{a\bar{b}})^2}{n n_{\bar{a}}} - \frac{n_b^2}{n^2} - \frac{n_{\bar{b}}^2}{n^2}$
15	II	$1 - \sum_{k=max(0,n_a-n_b)}^{n_{a\bar{b}}} \frac{C_{n_b}^{n_a-k} C_{n_{\bar{b}}}^{k}}{C_n^{n_a}}$
16	IPEE	$1 - \frac{1}{2^{n_a}} \sum_{k=0}^{n_{a\bar{b}}} C_{n_a}^{k}$
17	Jaccard	$\frac{n_a-n_{a\bar{b}}}{n_b+n_{a\bar{b}}}$
18	J-measure	$\frac{n_a-n_{a\bar{b}}}{n} log_2 \frac{n(n_a-n_{a\bar{b}})}{n_a n_b} + \frac{n_{a\bar{b}}}{n} log_2 \frac{n n_{a\bar{b}}}{n_a n_{\bar{b}}}$
19	Kappa	$\frac{2(n_a n_{\bar{b}}-n n_{a\bar{b}})}{n_a n_{\bar{b}}+n_{\bar{a}} n_b}$
20	Klosgen	$\sqrt{\frac{n_a-n_{a\bar{b}}}{n}}(\frac{n_{\bar{b}}}{n} - \frac{n_{a\bar{b}}}{n_a})$
21	Laplace	$\frac{n_a+1-n_{a\bar{b}}}{n_a+2}$
22	Least Contradiction	$\frac{n_a-2n_{a\bar{b}}}{n_b}$
23	Lerman	$\frac{n_a-n_{a\bar{b}}-\frac{n_a n_b}{n}}{\sqrt{\frac{n_a n_b}{n}}}$
24	Lift	$\frac{n(n_a-n_{a\bar{b}})}{n_a n_b}$
25	Loevinger	$1 - \frac{n n_{a\bar{b}}}{n_a n_{\bar{b}}}$
26	Odds Ratio	$\frac{(n_a-n_{a\bar{b}})(n_{\bar{b}}-n_{a\bar{b}})}{n_{a\bar{b}}(n_b-n_a+n_{a\bar{b}})}$
27	Pavillon	$\frac{n_{\bar{b}}}{n} - \frac{n_{a\bar{b}}}{n_a}$
28	Phi-Coefficient	$\frac{n_a n_{\bar{b}}-n n_{a\bar{b}}}{\sqrt{n_a n_b n_{\bar{a}} n_{\bar{b}}}}$
29	Putative Causal Dependency	$\frac{3}{2} + \frac{4n_a-3n_b}{2n} - (\frac{3}{2n_a} + \frac{2}{n_{\bar{b}}})n_{a\bar{b}}$
30	Rule Interest	$\frac{n_a n_{\bar{b}}}{n} - n_{a\bar{b}}$
31	Sebag & Schoenauer	$\frac{n_a}{n_{a\bar{b}}} - 1$
32	Support	$\frac{n_a-n_{a\bar{b}}}{n}$
33	TIC	$\sqrt{TI(a \rightarrow b) \times TI(\bar{b} \rightarrow a)}$
34	Yule's Q	$\frac{n_a n_{\bar{b}}-n n_{a\bar{b}}}{n_a n_{\bar{b}}+(n_b-n_{\bar{b}}-2n_a)n_{a\bar{b}}+2n_{a\bar{b}}^2}$
35	Yule's Y	$\frac{\sqrt{(n_a-n_{a\bar{b}})(n_{\bar{b}}-n_{a\bar{b}})}-\sqrt{n_{a\bar{b}}(n_b-n_a+n_{a\bar{b}})}}{\sqrt{(n_a-n_{a\bar{b}})(n_{\bar{b}}-n_{a\bar{b}})}+\sqrt{n_{a\bar{b}}(n_b-n_a+n_{a\bar{b}})}}$

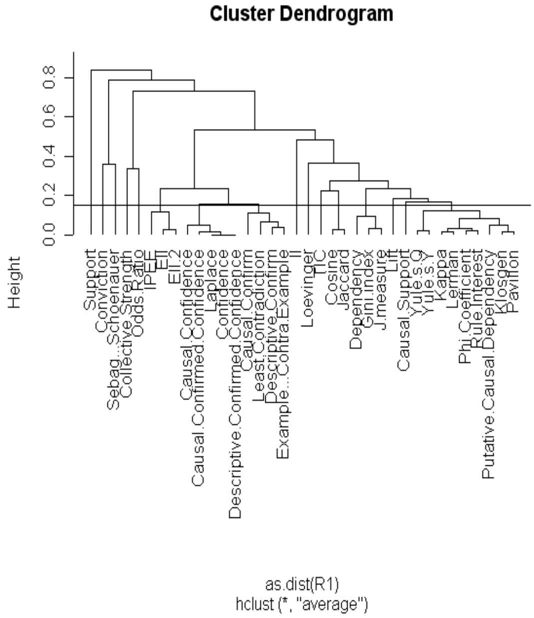

Figure 1: View on the strong relation between IMs.

5.1 View of Strong Relation

The strong relations between IMs are obtained by cutting the dendrogram (AHC) from the bottom with a small value of dissimilarity $d = 0.15$[1] (see Fig. 1 for R_1). The clusters are formed by the hierarchy of IMs in the zone under the horizontal line. The same result can be seen in Tab. 3 in which each column representing for a ruleset.

This view is helpful because the user can choose the clusters of IMs representing the strong agreement between IMs. In each cluster, a represented IM can be selected as a representative IM for all the IMs in the cluster. For example, one can select Laplace as the representative IM for the confidence cluster (Causal Confidence, Causal Confirmed-Confidence, Laplace, Confidence, Descriptive Confirmed-Confidence). This cluster is useful to discover the rules having the strong effect of confidence value.

Tab. 4 illustrates the five comparisons between the four rulesets (vertically). The first four columns show the comparison for each pair of rulesets. The last column illustrates the comparison results obtained from the four rulesets.

[1]The value $\rho = 0.85$ is used because of its widely acceptable in the literature.

Table 3: Clusters of IMs with the strong relations (IMs represented by their orders).

R_1	$R_1^{'}$	R_2	$R_2^{'}$
0,2,5,10,21	0,2,5,10,21	0,2,5,8,10,11,12, 16,21,25,27,29	0,2,5,8,10, 21,25,27,29
1,9,13,22	1,9,13,22	1,9	1,9
3	3,19,20,23, 24,27,28,30	3	3
4	4	4	4,32
6	6	6,31	6,31
7,17	7,17	7,17,19,23,28	7,17,19,23, 28,30,34,35
8,14,18	8,14,18		
11,12,16	11,12		11,12,16
		13	13
		14,18	14,18,20
15	15	15,34,35	15
	16		
19,20,23,27, 28,29,30,34,35			
		20	
		22	22
24		24	24,26
25	25,29		
26	26	26	
		30	
31	31		
32	32	32	
33	33	33	33
	34,35		

5.2 View of Relative Distance

Consider the number of clusters calculated from the precedent view, we can apply the PAM method to see the relative distance[2] between IMs. Fig. 2 illustrates the clusters of IMs in ellipse shape. The number represents the cluster order and each cluster has its specific symbol.

The complementary information obtained from the clusters in this view is useful to the user. By observing the diameter (smallest, biggest, ...) or the other parameters (separation, maximum distance, minimum distance, ...) one can choose the clusters to examine as different aspects from the dataset, each cluster is then represented by a representative IM. This representative IM is calculated as a medoid in the cluster. For example in Fig. 2, Tab. 5, cluster 2 (Least Contradiction, Example & Contra-Example, Causal Confirm, Descriptive Confirm) one can have Example & Contra-Example as the representative IM having the strong effect of negative examples.

Tab. 6, the same way to compare the results between rulesets as Tab. 4, gives all common clusters obtained from the four rulesets evaluated. At the fifth column, we can see an interesting cluster with only one IM : TIC (33), is the original IM for capturing an aspect of informational ratio modulated by the contrapositive.

[2]By using the PCA (Principal Component Analysis) technique.

Table 4: Cluster comparison from the strong relation view (IMs represented by their orders).

$R_1 \cap R_1'$	$R_2 \cap R_2'$	$R_1' \cap R_2'$	$R_1 \cap R_2$	$R_1 \cap R_1' \cap R_2 \cap R_2'$
0,2,5,10,21	0,2,5,8,10,21,25,27,29	0,2,5,10,21	0,2,5,10,21	0,2,5,10,21
1,9,13,22	1,9	1,9	1,9	1,9
	3		3	
4			4	
6	6,31			
7,17	7,17,19,23,28	7,17	7,17	7,17
8,14,18				
11,12	11,12,16	11,12	11,12,16	11,12
	13			
	14,18	14,18	14,18	14,18
15		15		
19,20,23,27,28,30		19,23,28,30	19,23,28	19,23,28
	22			
			24	
		25,29		
26			26	
			27,29	
31				
32			32	
33	33	33	33	33
34,35	34,35	34,35	34,35	34,35

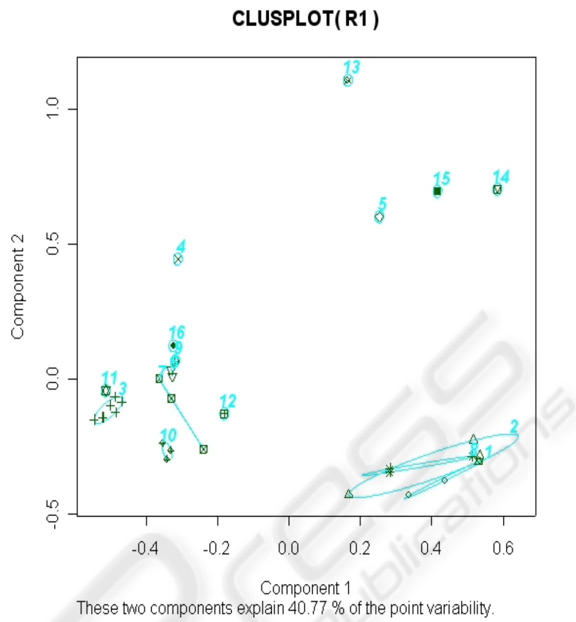

Figure 2: Views on the relative distance between clusters of IMs.

Table 5: Clusters of IMs with the relative distance (IMs represented by their orders).

R_1	R_1'	R_2	R_2'
0,2,5,10,21	0,1,2,5,10,21	0,2,5,8,10,21,25,27,29	0,2,5,8,10,21,25,27,29
1,9,13,22		1,9	1,9
3,19,23,28,30,34,35	3,19,20,23,24,27,28,30	3	3
4	4	4	4,14,18,20
6	6	6,31	6,31
7,17	7,17	7,17,19,23,28	7,17,19,23,28,30
8,14,18	8,14,18		
	9,13,22		
11,12,16	11,12	11,12,16	11,12,16
		13	13
		14,18,30	
15	15	15,34,35	15,34,35
	16		
20,27,29		20	
		22	22
24		24	24,26
25	25,29		
26	26	26	
31	31		
32	32	32	32
33	33	33	33
	34,35		

5.3 Stable Clusters

From the two different evaluations based on the two views of strong relation and relative distance, the more surprising result appears. By analyzing the fifth column from Tab. 4 and Tab. 6, eight stable clusters are found indicating an invariance with the nature of the dataset!.

- The first cluster (Causal Confirmed-Confidence, Laplace, Confidence, Descriptive Confirmed-Confidence, Causal Confidence) has most of the measures issued from the Confidence measure.

- The second cluster (Cosine, Jaccard) has a strong relation with the fifth property proposed by Tan et al. (Tan et al., 2004).

- The third cluster (EII 2, EII) are two measures obtained with different parameters of the same original formula and very useful in evaluating the entropy of implication intensity.

- The fourth cluster (Gini-index, J-measure) is an entropy cluster.

- The fifth cluster (Kappa, Lerman, Phi-Coefficient) is a set of similarity IMs.

- The sixth cluster (Support) indicates the influence of the support values of the rule.

- The seventh cluster (TIC) has only one measure provides the strong evaluation on the information ratio modulated by contrapositive.

- The last cluster (Yule's Y, Yule's Q) gives trivial observation because the measures are all derived from Odds Ratio measure, that is similar to the second property proposed by Tan et al. (Tan et al., 2004).

6 CONCLUSION

Discovering the behaviors of IMs is an interesting research and with the obtained results we can strongly help the user understand different hidden aspects existing on specific datasets. The evaluation of various IMs on the datasets having opposite characteristics is

Table 6: Cluster comparison from the relative distance view (IMs represented by their orders).

$R_1 \cap R_1'$	$R_2 \cap R_2'$	$R_1' \cap R_2'$	$R_1 \cap R_2$	$R_1 \cap R_1' \cap R_2 \cap R_2'$
0,2,5,10,21	0,2,5,8,10, 21,25,27,29	0,2,5,10,21	0,2,5,10,21	0,2,5,10,21
	1,9		1,9	
3,19,23, 28,30	3			
4			4	
6	6,31			
7,17	7,17,19,23,28	7,17	7,17	7,17
8,14,18				
9,13,22				
11,12	11,12,16	11,12	11,12,16	11,12
	13			
	14,18	14,18	14,18	14,18
15	15,34,35			
		19,23,28,30	19,23,28	19,23,28
20,27				
	22			
			24	
		25,29		
26			26	
			27,29	
31				
32	32	32	32	32
33	33	33	33	33
34,35		34,35	34,35	34,35

an important method. By calculating the dissimilarity between 36 IMs, we have determined eight stable clusters of IMs as eight different aspects found from the two opposite datasets.

The eight stable clusters denote an interesting relations between IMs because they remark the stable behaviors.

REFERENCES

Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., and Verkano, A. (1996). Fast discovery of association rules. In *Advances in Knowledge Discovery in Databases*. AAAI/MIT Press.

Blanchard, J., Guillet, F., Gras, R., and Briand, H. (2005a). Assessing rule interestingness with a probabilistic measure of deviation from equilibrium. In *ASMDA'05, Proceedings of the 11th International Symposium on Applied Stochastic Models and Data Analysis*.

Blanchard, J., Guillet, F., Gras, R., and Briand, H. (2005b). Using information-theoretic measures to assess association rule interestingness. In *ICDM'05, Proceedings of the 5th IEEE International Conference on Data Mining*.

Blanchard, J., Kuntz, P., Guillet, F., and Gras, R. (2003). Implication intensity: from the basic statistical definition to the entropic version (Chap. 28). In *Statistical Data Mining and Knowledge Discovery*.

Carvalho, D. R., Freitas, A. A., and Ebecken, N. F. F. (2003). A critical review of rule surprisingness measures. In *Proceedings of Data Mining IV - International Confeference on Data Mining*.

Carvalho, D. R., Freitas, A. A., and Ebecken, N. F. F. (2005). Evaluating the correlation between objective rule interestingness measures and real human interest. In *PKDD'05, the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases*.

Choi, D. H., Ahn, B. S., and Kim, S. H. (2005). Prioritization of association rules in data mining: Multiple criteria decision approach. In *ESA'05, Expert Sytems with Applications*.

Freitas, A. (1999). On rule interestingness measures. In *Knowledge-Based Systems, 12(5-6)*. Elsevier.

Gavrilov, M., Anguelov, D., Indyk, P., and Motwani, R. (2000). Mining the stock market: which measure is best? In *KDD'00, Proceedings of the 6th International Conference on Knowledge Discovery and Data Mining*.

Gras, R., Briand, H., Peter, P., and Philippé, J. (1996). Implicative statistical analysis. In *IFCS'96, Proceedings of the Fifth Conference of the International Federation of Classification Societies*. Springer-Verlag.

Hilderman, R. and Hamilton, H. (2001). *Knowledge Discovery and Measures of Interestingness*. Kluwer Academic Publishers.

Huynh, X.-H., Guillet, F., and Briand, H. (2005). Clustering interestingness measures with positive correlation. In *ICEIS'05, Proceedings of the 7th International Conference on Enterprise Information Systems*.

Kaufman, L. and Rousseeuw, P. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley.

Klösgen, W. (1996). Explora: a multipattern and multistrategy discovery assistant. In *Advances in Knowledge Discovery and Data Mining*. AAAI/MIT Press.

Major, J. and Magano, J. (1995). Selecting among rules induced from a hurricane database. In *Journal of Intelligent Information Systems 4(1)*.

Newman, D., Hettich, S., Blake, C., and Merz, C. (1998). *[UCI] Repository of machine learning databases, http://www.ics.uci.edu/∼mlearn/MLRepository.html*. University of California, Irvine, Dept. of Information and Computer Sciences.

Piatetsky-Shapiro, G. (1991). Discovery, analysis and presentation of strong rules. In *Knowledge Discovery in Databases*. MIT Press.

Ross, S. (1987). *Introduction to probability and statistics for engineers and scientists*. Wiley.

Tan, P.-N., Kumar, V., and Srivastava, J. (2004). Selecting the right objective measure for association analysis. In *Information Systems 29(4)*. Elsevier.