POST-PROCESSING ASSOCIATION RULES WITH CLUSTERING AND OBJECTIVE MEASURES

Veronica Oliveira de Carvalho
Instituto de Geociências e Ciências Exatas, UNESP - Univ Estadual Paulista, Rio Claro, Brazil

Fabiano Fernandes dos Santos, Solange Oliveira Rezende
Instituto de Ciências Matemáticas e de Computação, USP - Universidade de São Paulo, São Carlos, Brazil

Keywords: Association rules, Post-processing, Clustering and objective measures.

Abstract: The post-processing of association rules is a difficult task, since a large number of patterns can be obtained. Many approaches have been developed to overcome this problem, as objective measures and clustering, which are respectively used to: (i) highlight the potentially interesting knowledge in domain; (ii) structure the domain, organizing the rules in groups that contain, somehow, similar knowledge. However, objective measures don't reduce nor organize the collection of rules, making the understanding of the domain difficult. On the other hand, clustering doesn't reduce the exploration space nor direct the user to find interesting knowledge, making the search for relevant knowledge not so easy. This work proposes the PAR-COM (Post-processing Association Rules with Clustering and Objective Measures) methodology that, combining clustering and objective measures, reduces the association rule exploration space directing the user to what is potentially interesting. Thereby, PAR-COM minimizes the user's effort during the post-processing process.

1 INTRODUCTION

Association rules are widely used in many distinct domain problems (see (Semenova et al., 2001; Fonseca et al., 2003; Aggelis, 2004; Metwally et al., 2005; Domingues et al., 2006; Zhang and Gao, 2008; Rajasekar and Weng, 2009; Changguo et al., 2009)) due to its ability to discover the frequent relationships that occur among sets of items stored in databases. Although this characteristic along with its inherent comprehensibility motivates its use, the main weakness of association technique occurs when it is necessary to analyze the mining result. The huge number of rules that are generated makes the user exploration a difficult task. Many approaches have been developed to overcome this problem, as *Querying* (*Q*), *Evaluation Measures* (*EM*), *Pruning* (*P*), *Summarizing* (*S*) and *Grouping* (*G*) (Baesens et al., 2000; Jorge, 2004; Natarajan and Shekar, 2005; Zhao et al., 2009). These post-processing approaches aid the exploration process by reducing the exploration space (**RES**), as *Q*, *P* and *S*, by directing the user to what is potentially interesting (**DUPI**), as *EM*, or by structuring the domain (**SD**), as *G*.

One of the more popular approaches to estimate the interestingness of a rule is the application of evaluation measures (Natarajan and Shekar, 2005; Zhao et al., 2009). These measures are usually classified as objective or subjective. The objective measures depend exclusively on the structure pattern and the data used in the process of knowledge extraction, while the subjective measures depend fundamentally on the final user's interest and/or needs. Therefore, the objective measures are more general and independent on the domain in which the data mining process is carried out. (Geng and Hamilton, 2006; Ohsaki et al., 2004; Tan et al., 2004) describe many objective measures besides the classics *Support* and *Confidence*. In this approach, the rules are ranked according to a selected measure and an ordered list of potentially interesting knowledge is shown to the user. Although this **DUPI** approach highlights the potentially interesting knowledge, it doesn't reduce nor organize the collection of rules, making the understanding of the domain difficult.

Grouping is a relevant approach related to **SD**, since it organizes the rules in groups that contain, somehow, similar knowledge. These groups improve

Carvalho V., Santos F. and Rezende S..
POST-PROCESSING ASSOCIATION RULES WITH CLUSTERING AND OBJECTIVE MEASURES.
DOI: 10.5220/0003457500540063
In *Proceedings of the 13th International Conference on Enterprise Information Systems* (ICEIS-2011), pages 54-63
ISBN: 978-989-8425-53-9

the presentation of the mined patterns, providing the user a view of the domain to be explored (Reynolds et al., 2006; Sahar, 2002). However, this approach doesn't reduce the exploration space nor direct the user to find interesting knowledge, making the search for finding relevant knowledge not so easy. Grouping can be done: (i) based on a user criteria; (ii) by using a clustering technique. In case (i) the user describes how the groups will be formed; for example, the user can specify that rules that have the same consequent will be grouped together. In case (ii) the user "let the rules speak for themselves" (Natarajan and Shekar, 2005).

Clustering is the process of finding groups in data (Kaufman and Rousseeuw, 1990). A cluster is a collection of objects that are similar to each other within the group and dissimilar towards the objects of the other groups[1]. Many steps have to be done in a clustering process as: (i) the selection of a similarity/dissimilarity measure, used to calculate the proximity among the objects; (ii) the selection/execution of a clustering algorithm, which are basically divided in two families: partitional and hierarchical (Kaufman and Rousseeuw, 1990).

Considering the exposed arguments, this work proposes the PAR-COM (Post-processing Association Rules with Clustering and Objective Measures) methodology that, by combining clustering (**SD**) and objective measures (**DUPI**), reduces the association rule exploration space by directing the user to what is potentially interesting. Thus, PAR-COM improves the post-processing process since it adheres **RES** and **DUPI**. Besides, different from the approaches related to **RES**, PAR-COM doesn't only show the user a reduced space through a small subset of groups but also highlights the potentially interesting knowledge.

The paper is structured as follows: Section 2 presents some concepts and related works; Section 3 the PAR-COM methodology; Section 4 the configurations used in experiments to apply PAR-COM; Section 5 the results and discussion; Section 6 the conclusions and future works.

2 RELATED WORKS

Since PAR-COM combines clustering and objetive measures, this section presents some works related to the clustering approach. The works regarding objective measures are all associated with the ranking of rules and due to its simplicity are not here described.

[1]The words cluster and group will be used as synonymous in this work.

In order to structure the extracted knowledge, different clustering strategies have been used for post-processing association rules. (Reynolds et al., 2006) propose to group partially classification rules obtained by two algorithms proposed by them. In this case, all the rules have the same consequent, i.e., the clustering is done taking into accounting the antecedent of the rules. Although the kind of rule considered in their work is not association, the idea is the same: the only difference is that all the rules contain the same consequent. Clustering is demonstrated through partitional (K-means, PAM, CLARANS) and hierarchical (AGNES) algorithms using Jaccard as the similarity measure. The Jaccard between two rules r and s, presented in Equation 1, is calculated considering the common transactions (t) the rules match (in our work we refer this similarity measure as Jaccard with Rules by Transactions (J-RT)). A rule matches a transaction t if all the rule's items are contained in t.

$$\text{J-RT(r,s)}=\frac{\{t\ matched\ by\ r\} \cap \{t\ matched\ by\ s\}}{\{t\ matched\ by\ r\} \cup \{t\ matched\ by\ r\}} \quad (1)$$

(Jorge, 2004) demonstrates the use of clustering through hierarchical algorithms (Single Linkage, Complete Linkage, Average Linkage) using Jaccard as the similarity measure. In this case, the Jaccard between two rules r and s, presented in Equation 2, is calculated considering the items the rules share (in our work we refer to this measure as Jaccard with Rules by Items (J-RI)).

$$\text{J-RI(r,s)}=\frac{\{items\ in\ r\} \cap \{items\ in\ s\}}{\{items\ in\ r\} \cup \{items\ in\ r\}} \quad (2)$$

(Toivonen et al., 1995) propose a similarity measure based on transactions and use a density algorithm to do the clustering of the rules. In their work it is considered that all rules contain the same consequent, i.e., as in (Reynolds et al., 2006) the clustering is done taking into account the antecedent of the rules. (Sahar, 2002) also proposes a similarity measure based on transactions considering the (Toivonen et al., 1995) work, although it uses a hierarchical algorithm to do the clustering. However the algorithm is not mentioned and, in this case, it is considered that the rules contain distinct consequents.

It is important to observe that all the described works, related to **SD**, are only concerned with the domain organization. Thus, a methodology as PAR-COM that take it as an advantage to reduce the exploration space, by directing the user to relevant knowledge, is useful.

3 PAR-COM METHODOLOGY

The PAR-COM (Post-processing Association Rules with Clustering and Objective Measures) methodology aims at combining clustering and objetive measures to reduce the association rule exploration space directing the user to what is potentially interesting. For this purpose, PAR-COM considers that there is a subset of groups that contains all the *h*-top interesting rules, so that a small number of groups have to be explored. The *h*-top interesting rules are the *h* rules that have the highest values regarding an objective measure, where *h* is a number to be chosen. Besides, it is also considered that if some rules within a group express interesting knowledge, than the other rules within the same group also tend to express interesting knowledge. This assumption is taken considering the concept of cluster: a collection of objects that are similar to one another. So, if the rules are similar regarding a similarity measure, an interesting rule within a group indicates that its similar rules are also potentially interesting. Based on the exposed arguments, PAR-COM can reduce the exploration space by directing the user to the groups that are ideally interesting. As as consequence, PAR-COM can allow the discovery of additional interesting knowledge inside these groups.

The PAR-COM methodology, presented in Figure 1, is described as follows:

Step A: the value of an objective measure is computed for all rules in the association set.

Step B: the *h*-top rules is selected considering the computed values.

Step C: after selecting a clustering algorithm and a similarity measure the rule set is clustered.

Step D: a search is done to find out the clusters that contain one or more *h*-top rules selected in Step B. These clusters are the ones that contain the potentially interesting knowledge (PIK) of the domain. The more *h*-top rules a cluster has the more interesting it is.

Step E: only the *m* first interesting clusters are shown to the user, who is directed to a reduced exploration space that contains the PIK of the domain, where *m* is a number to be chosen.

As will be noted in the results presented in Section 5, the combination of clustering with objective measures used in PAR-COM aids the post-processing process, minimizing the user's effort.

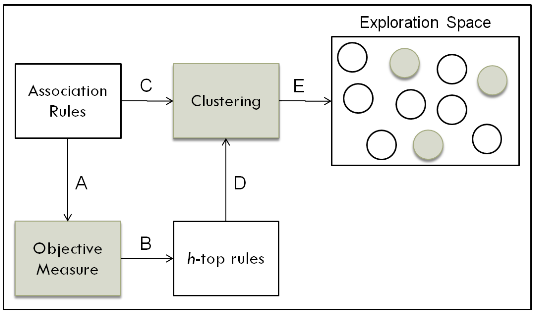


Figure 1: The PAR-COM methodology.

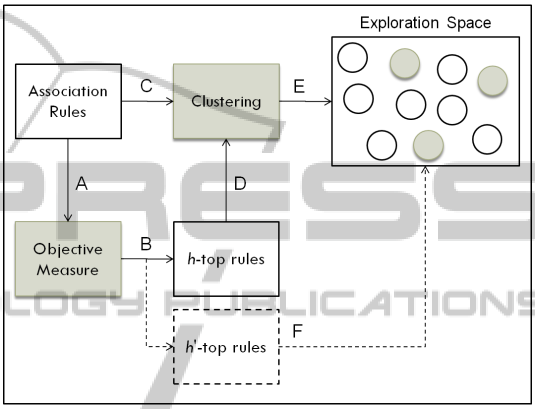


Figure 2: Step F: a validation step in the PAR-COM methodology.

4 EXPERIMENTS

Some experiments were carried out to evaluate the performance of PAR-COM. However, in order to validate the results shown in Section 5 an additional step was added to the methodology, as presented in Figure 2. **Step F** considers all the *h*'-top interesting rules to be also selected in Step B. The *h*'-top rules are the first *h* rules that immediately follow the previously selected *h*-top rules. Thus, the aim of Step F is to demonstrate that the *m* clusters shown to the user really contain PIK. For this purpose, a search is done to find out if these *m* clusters contain one or more *h*'-top rules. It is expected that these *m* clusters cover all the *h*'-top rules, since by definition a cluster is a collection of objects that are similar to one another. So, as mentioned before, if the rules are similar regarding a similarity measure, an interesting rule inside a group indicates that its similar rules are also potentially interesting. It is important to note that PAR-COM doesn't aid the exploration as an ordered list of PIK, which is the case when objective measures are used. For that reason, PAR-COM can allow the discovery of additional interesting knowledge inside the *m* groups.

The two data sets used in experiments are pre-

Table 1: Details of the data sets used in experiments.

Data set	# of transactions	# of distinct items	Brief description
Adult	48842	115	This set is a R pre-processed version for association mining of the "Adult" database available in UCI (Frank and Asuncion, 2010). It was originally used to predict whether income exceeds USD 50K/yr based on census data.
Income	6876	50	This set is also a R pre-processed version for association mining of the "Marketing" database available in (Hastie et al., 2009). It was originally used to predict the anual income of household from demographics attributes.

sented in Table 1. These data sets are available in *R Project for Statistical Computing*[2] through "arules" package[3]. For both data sets the rules were mined using an *Apriori* implementation developed by Christian Borgelt[4] with a maximum number of 5 items per rule and excluding the rules of type $TRUE \Rightarrow X$, where X is an item in the data set. With the Adult data set 6508 rules were generated using a minimum support of 10% and a minimum confidence of 50% and with Income 3714 rules considering a minimum support of 17% and a minimum confidence of 50%. These parameter values, as those presented below, were chosen experimentally.

Since the works described in Section 2 only use one family of clustering algorithms and one similarity measure to cluster the association rules, it was decided to apply PAR-COM with one algorithm of each family and with the two most used similarity measures (J-RI and J-RT (Equations 1 and 2)). The Partitioning Around Medoids (PAM) was chosen within the partitional family and the Average Linkage within the hierarchical family. In the partitional case, a medoid algorithm was chosen because the aim is to cluster the more similar rules in one group; thus, the ideal is that the centroid group be a rule and not, for example, the mean, as in the K-means algorithm. In the hierarchical case, the traditional algorithms were applied (Single, Complete and Average) and the one that had the best performance is here presented. PAM was executed with k ranging between 6 to 15. The dendrograms generated by Average Linkage were cut in the same ranges (6 to 15).

To apply PAR-COM it was also necessary to choose the values of h (Step B), m (Step E) and an objective measure (Step A). h was set to 15, the highest value of k, because we want to evaluate if the 15-top rules were spread among the groups (one in each group) or concentrated in little groups (as expected by the PAR-COM methodology). m was set to 3, half of the minimum value of k, because we want to evaluate the exploration space reduction considering only 50% of the groups. To evaluate the behavior of the objective measures in the PAR-COM methodology, 6 measures were chosen among the ones described in (Tan et al., 2004): *Certainty Factor* (CF), *Collective Strength* (CS), *Gini Index* (GI), *Laplace* (L), *Lift* (also known as *Interest Factor*) and *Novelty* (Nov) (also known as *Piatetsky-Shapiro's*, *Rule Interest* or *Leverage*). These measures were chosen because they are more used than the others in the post-processing works found in literature (see (Zhao et al., 2009)). Besides, it is expected that any measure produces good results. Table 2 summarizes the configurations applied to evaluate PAR-COM.

Table 2: Configurations used to evaluate PAR-COM.

Data sets	Adult; Income
Algorithms	PAM; Average Linkage
Similarity measures	J-RI; J-RT
k	6 to 15
h	15
m	3
Objective measures	CF; CS; GI; L; Lift; Nov

5 RESULTS AND DISCUSSION

Considering the configurations presented in Table 2, PAR-COM was applied and the results are presented in Figures 4, 5, 6 and 7. The results were grouped by algorithm for each data set. Figures 4 and 6 present the results for the Adult data set using, respectively, PAM and Average Linkage and Figures 5 and 7 for Income also using, respectively, the same algorithms. Each figure contains 12 sub-figures: 6 related to the

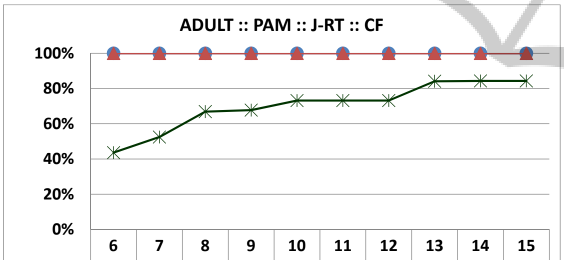
[2]http://www.r-project.org/.

[3]http://cran.r-project.org/web/packages/arules/index.html.

[4]http://www.borgelt.net/apriori.html.

J-RI similarity measure and 6 to J-RT; each group of these 6 figures corresponds to an objective measure. The *x* axis of each graphic represents the range considered for *k*. The *y* axis represents the percentage of *h*-top and *h*'-top rules contained in the *m* first interesting clusters (lines *h*-top and *h*'-top) and also the percentage of reduction in the exploration space (line R). Each graphic title indicates the configuration used.

In order to facilitate the interpretation of the graphics consider Figure 3 (an enlarged version of Figure 4(g)). It can be observed that: (i) the first 3 interesting clusters (*m*=3) contain, for each *k*, all (100%) the 15-top rules (*h*=15) using J-RT with CF; (ii) the first 3 interesting clusters contain, for each *k*, all (100%) the 15'-top rules (*h*=15); thus, by the validation step (Step F), these 3 clusters are ideal the 3 most interesting subsets; (iii) for *k*=15, for example, the first 3 interesting clusters cover 16% (100%-84%) of the rules, leading to a reduction of 84% in the exploration space; in order words, if the user explores these 3 clusters, he will explore 16% of the rule's space.



	6	7	8	9	10	11	12	13	14	15
h-top	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
h'-top	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
R	44%	52%	67%	68%	73%	73%	73%	84%	84%	84%

Figure 3: PAM result in the Adult data set using J-RT and CF.

Evaluating the results, in relation to PAM algorithm, it can be noticed that:

- in Figure 4 the J-RT similarity measure presented better results compared with J-RI in relation to the *h*-top and *h*'-top rules (compare 4(a) with 4(g), 4(b) with 4(h), 4(c) with 4(i), 4(d) with 4(j), 4(e) with 4(k) and 4(f) with 4(l)). However, J-RI and J-RT had a similar behavior regarding the exploration space reduction. In J-RT all the objective measures presented similar results regarding *h* and *h*' different from J-RI that had the worst performance in *Laplace* and *Lift*. Besides, in both cases, high values of *k* give high reductions and a good performance related to the *h* and *h*'-top rules.
- in Figure 5 the J-RT similarity measure presented, in almost all the cases, better results compared with J-RI in relation to the *h*-top and *h*'-top rules (compare 5(a) with 5(g), 5(b) with 5(h), 5(c) with 5(i), 5(d) with 5(j), 5(e) with 5(k) and 5(f) with 5(l)). However, J-RI and J-RT had a similar behavior regarding the exploration space reduction. *Certainty Factor* and *Laplace* generated better results than the others in J-RI regarding *h* and *h*'; in J-RT, *Gini Index*, *Lift* and *Novelty* generated better results than the others regarding *h* and *h*'. Besides, in both cases, high values of *k* give high reductions and a good performance related to the *h* and *h*'-top rules.

Summarizing the results in Figures 4 and 5, it can be seen that with the PAM algorithm the similarity measure that had the best performance was J-RT regarding *h* and *h*'. However, considering the exploration space reduction, both similarity measures presented similar behavior. On the other hand, in relation to Average algorithm, it can be noticed that:

- in Figure 6 both J-RI and J-RT presented good results in relation to the *h*-top and *h*'-top rules in all the used objective measures. In almost all the cases, J-RI had a little better performance than J-RT considering the exploration space reduction for high values of *k* (compare 6(a) with 6(g), 6(b) with 6(h), 6(c) with 6(i), 6(d) with 6(j), 6(e) with 6(k) and 6(f) with 6(l)). Besides, in both cases, high values of *k* give high reductions and a good performance related to the *h* and *h*'-top rules.
- in Figure 7 both J-RI and J-RT presented good results in relation to the *h*-top and *h*'-top rules in almost all the used objective measures (exceptions were Figures 7(k) and 7(l)), although J-RI had a better performance than J-RT (compare 7(a) with 7(g), 7(b) with 7(h), 7(c) with 7(i), 7(d) with 7(j), 7(e) with 7(k) and 7(f) with 7(l)). In all the cases J-RT had a better performance than J-RI considering the exploration space reduction. Besides, in both cases, high values of *k* give high reductions and a good performance related to the *h* and *h*'-top rules.

Summarizing the results in Figures 6 and 7, it can be seen that with the Average algorithm the similarity measure that had the best performance was J-RI regarding *h* and *h*', although J-RT had presented similar behavior in many cases. However, considering the exploration space reduction, none of the similarity measures won in both data sets. Thus, since PAM had a better performance with J-RT and Average with J-RI, comparing the results of PAM using J-RT (Figures 4(g) to 4(l) and 5(g) to 5(l)) with Average using

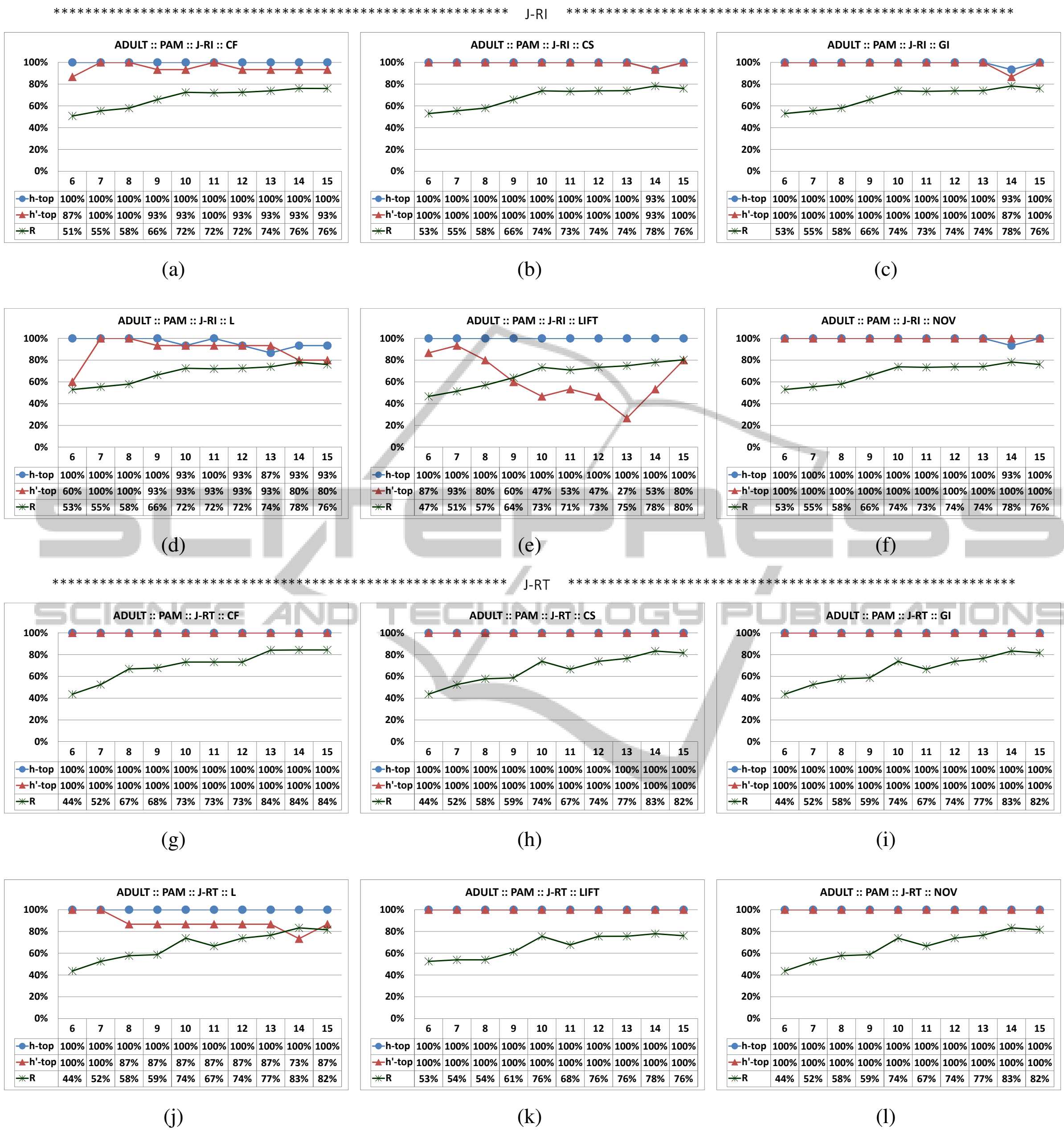


Figure 4: PAM's results in the ADULT data set.

J-RI (Figures 6(a) to 6(f) and 7(a) to 7(f)), it can be noticed that:

- in the Adult data set both algorithms had good results and similar behavior regarding *h* and *h*': in all the cases, 100% of recovery in the 3 interesting clusters (*m*=3) regarding *h*; in almost all the cases, 100% of recovery in the 3 interesting clusters (*m*=3) regarding *h*' (exception to Figure 4(j)). However, PAM had better results considering the reduction exploration space (above 80% for high values of *k*).

- in the Income data set the Average algorithm had good results and a better performance compared with PAM regarding *h* and *h*': in all the cases, 100% of recovery in the 3 interesting clusters (*m*=3) regarding *h*; in almost all the cases, 100% of recovery in the 3 interesting clusters (*m*=3) regarding *h*' (exception to Figures 7(d) and 7(e)). However, PAM had better results considering the reduction exploration space (above 70% for high

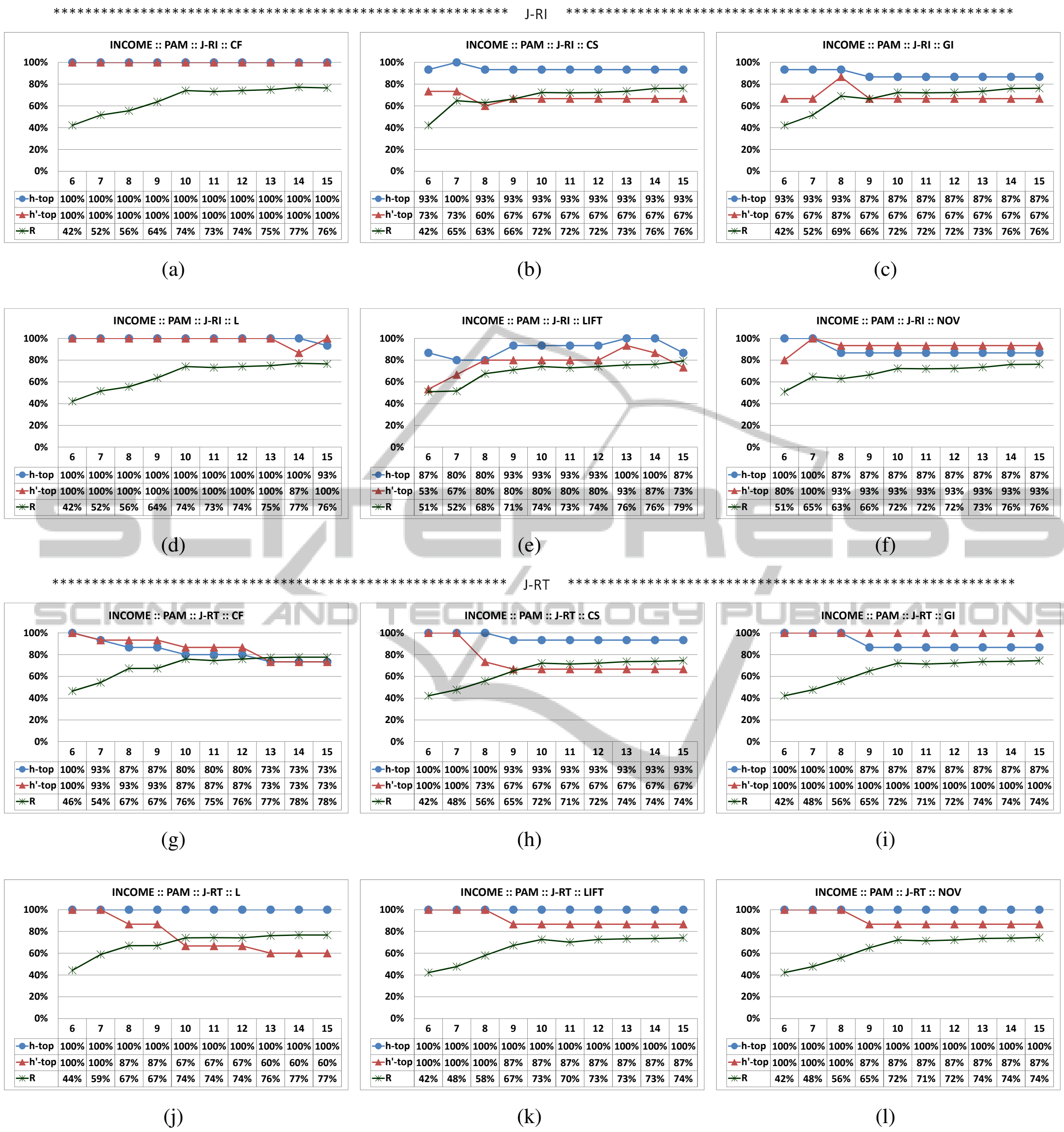


Figure 5: PAM's results in the INCOME data set.

values of k).

Based on the exposed discussion, it can be seen that the user can apply PAR-COM considering the combination PAM:J-RT or Average:J-RI. However, it is important to note that these similarity measures have a semantic that needs to be explored. That way, an evaluation with final users has to be done to find out which of them better recover the more adequate subset of groups related to the PIK.

Still discussing the results, it can be observed that the used objective measures had, broadly, a good performance regarding the h and h'-top rules. The exceptions, considering percentages below 70, were Figures 4(d), 4(e), 5(b), 5(c), 5(e), 5(h), 5(j) and 7(k), which represents approximately only 17% of the cases. Besides, high values of k give high reductions and a good performance related to the h and h'-top rules. Thus, we can reduce the exploration space when using high values of k also maintaining an interesting subset of rules.

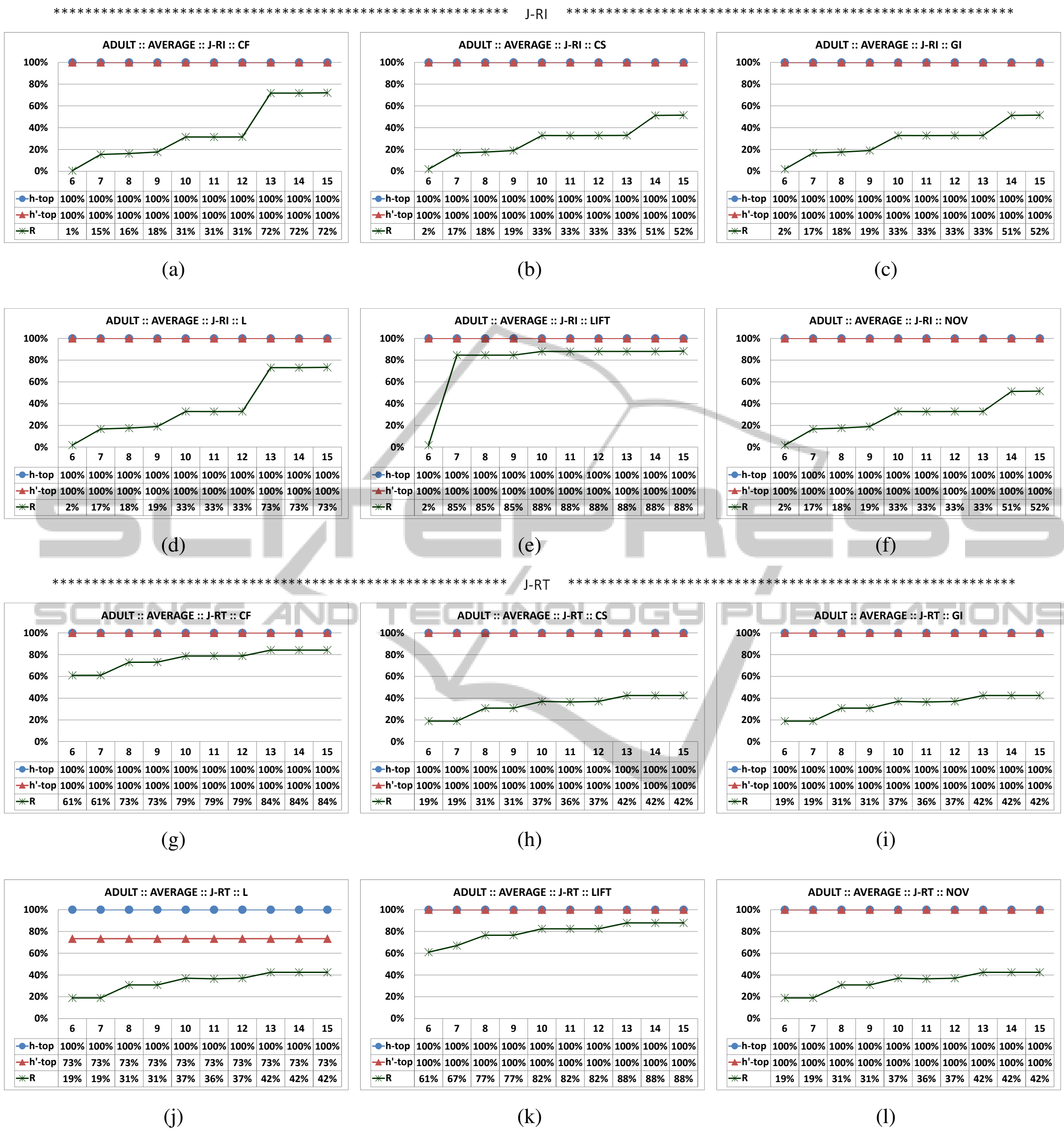


Figure 6: AVERAGE's results in the ADULT data set.

6 CONCLUSIONS

This work presented the PAR-COM methodology that by combining clustering (**SD**) and objective measures (**DUPI**) provides a powerful tool to aid the post-processing process, minimizing the user's effort during the exploration process. PAR-COM can present to the user only a small subset of the rules, providing a view to what is really interesting. Thereby, PAR-COM adheres **RES** and **DUPI**. PAR-COM has a good performance, as observed in Section 5, in: (i) highlighting the potentially interesting knowledge (PIK), demonstrated through the *h*'-top rules; (ii) reducing the exploration space. Thus, PAR-COM can reduce the exploration space without losing PIK, being a good methodology for post-processing association rules.

As a future work, some labeling methodologies will be studied and implemented that, along with PAR-COM, will direct the user to the potentially inte-

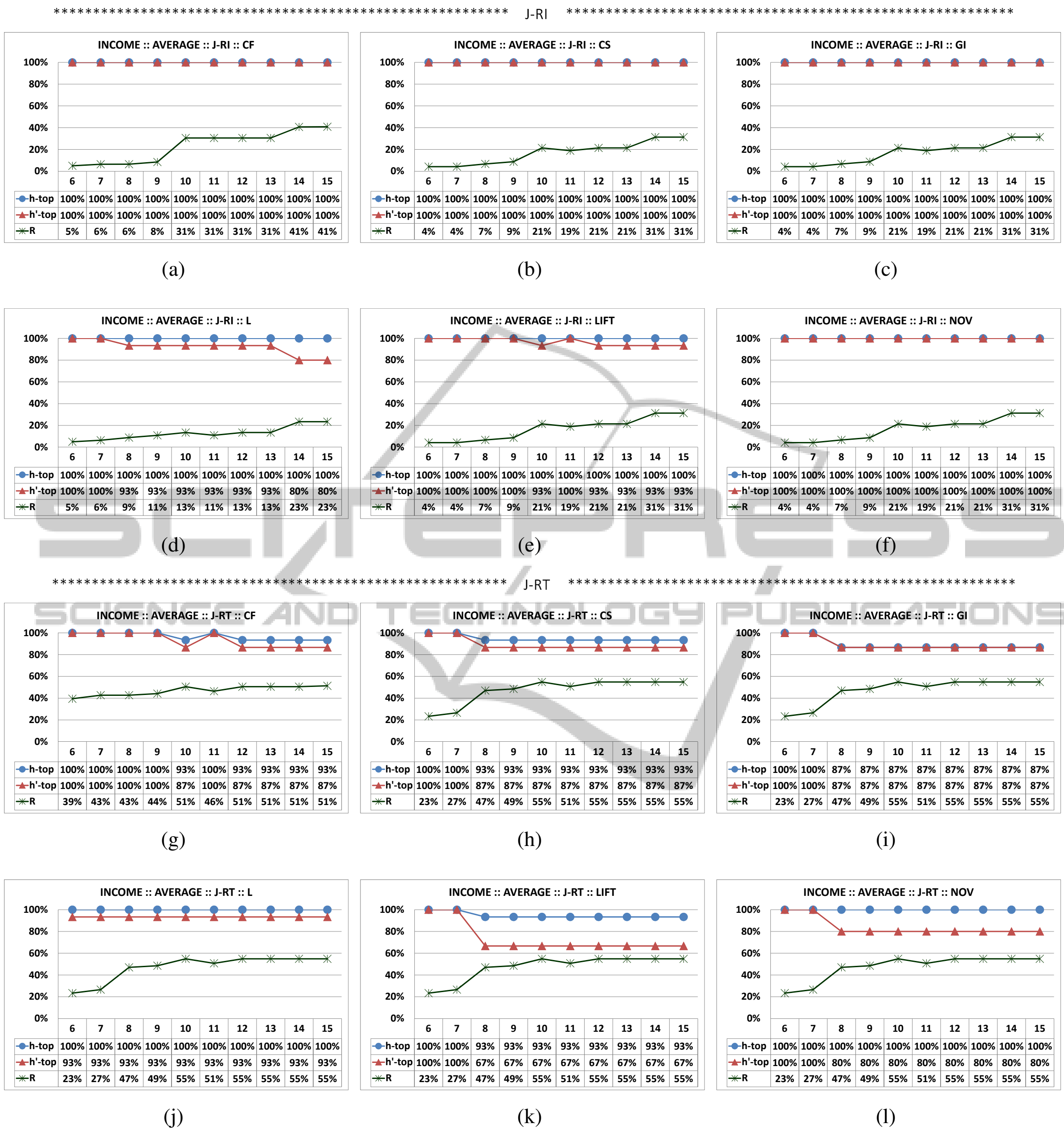


Figure 7: AVERAGE's results in the INCOME data set.

resting "topics" (PIT) in the domain.

ACKNOWLEDGEMENTS

We wish to thank Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for the financial support (process number 2010/07879-0).

REFERENCES

Aggelis, V. (2004). Association rules model of e-banking services. *Data Mining V – Information and Communication Technologies*, 5:46–55.

Baesens, B., Viaene, S., and Vanthienen, J. (2000). Post-processing of association rules. In *KDD'00: Proceedings of the Special Workshop on Post-processing, The 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2–8.

Changguo, Y., Nianzhong, W., Tailei, W., Qin, Z., and Xiaorong, Z. (2009). The research on the application of association rules mining algorithm in network intrusion detection. In Hu, Z. and Liu, Q., editors, *ETCS'09: Proceedings of the 1st International Workshop on Education Technology and Computer Science*, volume 2, pages 849–852.

Domingues, M. A., Jorge, A. M., and Soares, C. (2006). Using association rules for monitoring meta-data quality in web portals. In *WAAMD'06: Proceedings of the II Workshop em Algoritmos e Aplicações de Mineração de Dados – SBBD/SBES*, pages 105–108.

Fonseca, B. M., Golgher, P. B., Moura, E. S., and Ziviani, N. (2003). Using association rules to discover search engines related queries. In *LA-WEB'03: Proceedings of the 1st Conference on Latin American Web Congress*, pages 66–71. IEEE Computer Society.

Frank, A. and Asuncion, A. (2010). UCI machine learning repository. University of California, Irvine, School of Information and Computer Sciences. http://archive.ics.uci.edu/ml.

Geng, L. and Hamilton, H. J. (2006). Interestingness measures for data mining: A survey. In *ACM Computing Surveys*, volume 38. ACM Press.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics. Springer, second edition. http://www-stat.stanford.edu/ tibs/ElemStatLearn/.

Jorge, A. (2004). Hierarchical clustering for thematic browsing and summarization of large sets of association rules. In Berry, M. W., Dayal, U., Kamath, C., and Skillicorn, D., editors, *SIAM'04: Proceedings of the 4th SIAM International Conference on Data Mining*. 10p.

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

Metwally, A., Agrawal, D., and Abbadi, A. E. (2005). Using association rules for fraud detection in web advertising networks. In *VLDB'05: Proceedings of the 31st International Conference on Very Large Data Bases*, pages 169–180.

Natarajan, R. and Shekar, B. (2005). Interestingness of association rules in data mining: Issues relevant to e-commerce. *SĀDHANĀ – Academy Proceedings in Engineering Sciences (The Indian Academy of Sciences)*, 30(Parts 2&3):291–310.

Ohsaki, M., Kitaguchi, S., Okamoto, K., Yokoi, H., and Yamaguchi, T. (2004). Evaluation of rule interestingness measures with a clinical dataset on hepatitis. In Boulicaut, J.-F., Esposito, F., Giannotti, F., and Pedreschi, D., editors, *PKDD'04: Proceedings of the 8th European Conference on Principles and Practice of Knowledge Discovery in Databases*, volume 3202, pages 362–373. Springer-Verlag New York, Inc.

Rajasekar, U. and Weng, Q. (2009). Application of association rule mining for exploring the relationship between urban land surface temperature and biophysical/social parameters. *Photogrammetric Engineering & Remote Sensing*, 75(3):385–396.

Reynolds, A. P., Richards, G., de la Iglesia, B., and Rayward-Smith, V. J. (2006). Clustering rules: A comparison of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modelling and Algorithms*, 5(4):475–504.

Sahar, S. (2002). Exploring interestingness through clustering: A framework. In *ICDM'02: Proceedings of the IEEE International Conference on Data Mining*, pages 677–680.

Semenova, T., Hegland, M., Graco, W., and Williams, G. (2001). Effectiveness of mining association rules for identifying trends in large health databases. In Kurfess, F. J. and Hilario, M., editors, *ICDM'01: Workshop on Integrating Data Mining and Knowledge Management, The IEEE International Conference on Data Mining*. 12p.

Tan, P.-N., Kumar, V., and Srivastava, J. (2004). Selecting the right objective measure for association analysis. *Information Systems*, 29(4):293–313.

Toivonen, H., Klemettinen, M., Ronkainen, P., Hätönen, K., and Mannila, H. (1995). Pruning and grouping discovered association rules. Workshop Notes of the ECML'95 Workshop on Statistics, Machine Learning, and Knowledge Discovery in Databases, 47–52, MLnet.

Zhang, J. and Gao, W. (2008). Application of association rules mining in the system of university teaching appraisal. In *ETTANDGRS'08: Proceedings of the International Workshop on Education Technology and Training & International Workshop on Geoscience and Remote Sensing*, volume 2, pages 26–28. IEEE Computer Society.

Zhao, Y., Zhang, C., and Cao, L. (2009). *Post-Mining of Association Rules: Techniques for Effective Knowledge Extraction*. Information Science Reference. 372p.