

Post-processing Association Rules: A Network Based Label Propagation Approach

Renan de Padua¹(✉), Veronica Oliveira de Carvalho²,
and Solange Oliveira Rezende¹

¹ Instituto de Ciências Matemáticas e de Computação,
USP - Universidade de São Paulo, São Carlos, Brazil
{padua,solange}@icmc.usp.br

² Instituto de Geociências e Ciências Exatas,
UNESP - Univerdidade Estadual Paulista, Rio Claro, Brazil
veronica@rc.unesp.br

Abstract. Association rules are widely used to find relations among items in a given database. However, the amount of generated rules is too large to be manually explored. Traditionally, this task is done by post-processing approaches that explore and direct the user to the interesting rules. Recently, the user’s knowledge has been considered to post-process the rules, directing the exploration to the knowledge he considers interesting. However, sometimes the user wants to explore the rule set without adding his prior knowledge BIAS, exploring the rule set according to its features. Aiming to solve this problem, this paper presents an approach, named *PAR_{LP}* (**P**ost-processing **A**ssociation **R**ules using **L**abel **P**ropagation), that explores the entire rule set, suggesting rules to be classified by the user as “Interesting” or “Non-Interesting”. In this way, the user is directed to analyze the rules that have some importance on the rule set, so the user does not need to explore the entire rule set. Moreover, the user’s classification is propagated to all the rules using label propagation approaches, so the most similar rules will likely be on the same class. The results show that the *PAR_{LP}* succeeds to direct the exploration to a set of rules considered interesting, reducing the amount of association rules to be explored.

1 Introduction

Association is widely used in data mining due its simplicity and comprehensibility. This task aims to extract the correlations among items in a given database [1]. However, a large number of rules can be generated even on small data sets. Therefore, the manual exploration of the rules to find interesting patterns is unfeasible. Generally, the number of interesting rules is very small compared to the total number of patterns and, in most of the times, the user must search many rules to find the rules that are considered interesting to him/her.

Some research has been done to direct the users on the exploration and help them finding the interesting rules. Some authors propose the use of networks to

direct the user's exploration, as [6, 10]. Networks are well known on the literature by its capability to model data, preserving the connection among the items on the data set. The networks are used as a mean to facilitate the rules' exploration, modeling it and pruning the rules that are not interesting to the user. The problem of these approaches is that the user must inform, beforehand, what he considers interesting, forcing him to have the knowledge apriori.

Based on the exposed, this paper presents a subjective post-processing approach, named *PAR_{LP}*, that interacts with the user to extract his knowledge according to the importance of the association rules on the rule set. This interaction is iterative made during the exploration, suggesting rules that are considered most relevant according to some network measures. The approach can be divided into 4 steps: (1) the most important rules, according to a network measure, are discovered; (2) user assigns labels to the extracted rules; (3) label propagation is performed to obtain the interesting rules according to the user classification; (4) the user decides if the approach needs to execute again or finish the processing. The approach is based on the idea of using classification algorithms to post-processes association rules, as proposed by [2, 7], learning with the prior interactions.

On the second iteration, the user interaction is made considering only the rules that are classified as "Interesting" by the approach. Networks are well known to model relationships among data set objects, extract properties about the importance of the objects and patterns of the data set [5]. This peculiarity allows extracting important rules for labeling and also allows propagating the labels to classify other rules. After the user interaction classifying rules, the current labels are propagated to all the other rules using a label propagation algorithm. The classification algorithm, used in this work, takes as input a data set containing a huge amount of unlabeled data and a small amount of labeled data and propagates the labels to all the data, classifying the entire data set. The *PAR_{LP}* uses these characteristics to propagate the user's knowledge to the entire rule set.

In summary, the main contribution of this paper is a new post-processing approach that helps the user to find the most relevant rules according to his knowledge and the rule's importance in the rule set. The user interaction is made in a way that he does not need to have a prior knowledge on the data set, suggesting a few rules to be classified by him based on the rule's relevance in the network. By doing so, the approach excludes the necessity to have all the knowledge beforehand, making the classification process easier based on only a few rules, not considering the entire rule set. Besides, this paper also shows that post-processing association rules using label propagation is effective to find the interesting rules according to user's knowledge. This paper presents an experimental evaluation to demonstrate that *PAR_{LP}* is capable of finding the rules that the user considers as interesting. The experiments also show that the proposed approach can reduce the amount of rules to be explored by the user to find those interesting rules.

This paper is organized as follows. Section 2 describes related research and basic concepts. Section 3 presents the *PAR_{LP}* and its motivation. Section 4 describes the experiments that were carried out to analyze the approach. Section 5 discusses the results obtained in the experiments. Finally, conclusion and future works are given in Sect. 6.

2 Background and Related Works

The transductive post-processing approach, proposed in [7] and extended in this paper, comes up with a different way to explore the rules. The approach, named *PAR_{LP}*, models the association rules into networks. This modeling allows the approach to explore the rules according to their importance on the network, selecting the rules to be classified by the user, directing the user's effort. The approach uses label propagation to find the interesting rules based on the users' classification. The label propagation was chosen considering that if a rule is interesting, the rules similar to it are also interesting.

A network can be characterized by a set of elements and the relations among them. Formally, a network can be represented by $N = (V, E, W)$ where V is the set of vertices (elements), E the set of links between the vertices and W the weight of the links [5]. When all the vertices represent the same kind of object, only web sites, for example, the network is called homogeneous network. On the other hand, when the representation of different kinds of objects is considered, like persons and communities in a social network, the network is called heterogeneous network. Besides, there is a variation of the conventional network called bipartite network. The bipartite networks have two different kinds of vertices: groupers (G) and items (H). The groupers can only connect to items vertices and the items can only connect to groupers vertices. The groupers vertices allow the propagation of information among the items vertices. Those different representations allow a large variety of exploration. One example of a bipartite network is the connection among documents (groupers) and terms (items).

One way to classify all vertices of a network considering user's knowledge is through label propagation. Label propagation algorithms classify a data set based on few classified examples (the training set is composed of labeled and unlabeled instances). The classification is made based on a similarity measure. The elements that are considered similar are classified on the same class whereas the elements that are not considered similar are classified on different classes [13].

In this paper, we use two network types to apply label propagation: homogeneous and heterogeneous. The homogeneous networks are building using similarity measures to connect the rules. The bipartite network considers an objective measure to connect the items to the groupers vertices. To propagate the labels two algorithms were used per network type: Learning with Local and Global Consistency (LLGC) and Gaussian Fields and Harmonic Function (GFHF) for the homogeneous networks and Label Propagation using Bipartite Heterogeneous Network (LPBHN) and GNetMine for the bipartite networks. The propagation is made in a way to minimize a regularization function.

The LLGC and GFHF algorithms used to classify the homogeneous networks, were selected because of their great results obtained on the literature. The GFHF algorithm [12] considers the labeled elements as unchanging truth and propagates the class without the need of a parameter. The LLGC algorithm [11] considers that a labeled element can change its labels based on its neighbors. The LLGC algorithm needs the definition of a learning rate parameter α .

The GNetMine and LPBHN used to classify the bipartite heterogeneous networks, were also selected of their results obtained on the literature. The LPBHN algorithm [9] works in the same way the GFHF does. It considers the element's labels as unchanging truth and propagates these labels without the need of a parameter. The GNetMine [4] works in the same the LLGC does. It can change the original label based on its neighbors and need the definition of a learning rate parameter *alpha*. The algorithms' functions can be seen on their respective papers.

[6] proposed an approach that uses networks as a mean to post-process association rules. In their work, the authors force the user to select an item (let's call it objective item) to be the main objective of the exploration. After the selection, a directed hypergraph is constructed considering the objective item as the consequent of the rules and the antecedents are modeled, constructing the level 1. The level 2 considers the antecedents of all the rules modeled on the level 1 as consequent and repeat the process. The hypergraph is constructed until all the items are modeled or the maximum level is reached (informed by the user). Using that hypergraph, all the rules modeled according to an objective item and the user can see how the items interact with the selected item. This approach forces the user to explore only one item per time and, yet, the user must know beforehand which item he/she wants to study.

3 PAR_{LP} : Post-processing Association Rules Using Label Propagation

The main concept behind the PAR_{LP} is that if the user considers a rule R_x as interesting, then the rules similar to it will also be interesting to be explored. The same goes to the case that the user finds a rule R_y not interesting, considering the similar rules also as not interesting. Figure 1 shows how the exploration is made. Consider "1" and "2" as the rules to be found. On (a) user selects "2" as interesting and "5" as not interesting. This classification is made aided by the proposed approach. On (b) the label propagation is done, considering the squares as not interesting and circles as interesting. Note that, after applying the classification algorithm, the interesting rules ("1" and "2", plus "3") are classified as interesting and will be presented to the user. In this case, the reduction of the exploration space is 50 % (the user only needs to explore 50 % of the existing rules).

The PAR_{LP} works as shown in Fig. 2. The input of the approach is a rule set, containing the rules to be explored. The approach starts on step 1, modeling the association rules in a network. To do so, it's necessary to define the

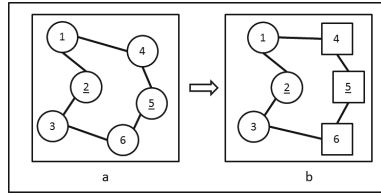


Fig. 1. The PAR_{LP} concept

network type [NT] and the similarity measure [S] to be used. The network type defines the network structure, i.e., homogeneous or heterogeneous networks. In case of homogeneous networks, the user also needs to define a way to link the rules considering their similarity, such as connecting a rule with their K-Nearest Neighbors or connect all rules through an RBF kernel [3]. The similarity measure [S], like hoJaccard (Eq. 2) and hoConfidence (Eq. 3), is used to calculate the similarity between two rules, setting the weight of their connections. In the beginning, the modeled association rules have no defined class, containing only the rules and the connections among them.

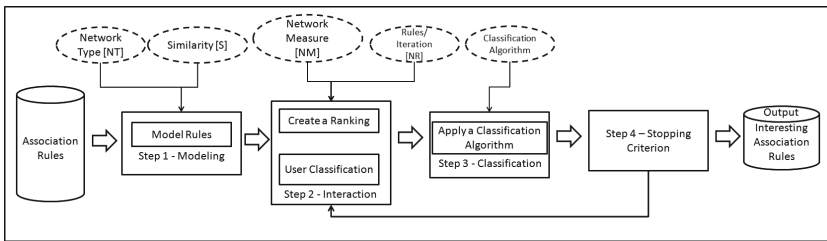


Fig. 2. The PAR_{LP} approach

In step 2, the PAR_{LP} interacts with the user, aiming to classify a few rules that will be used on the label propagation process to discover other interesting and non-interesting rules. This interaction aims to capture the user’s knowledge and direct the exploration to the rules he considers “Interesting”. To do so, a ranking is created to select the best rules according to a centrality measure [NM]. On the bipartite networks, a projection on a homogeneous network was created to apply the [NM]. This project was based on the bibliographic coupling concept [5]. The approach selects the N best rules and the N worst rules to be classified by the user; that way, the selected rules are not in the same dense point in the network. The number of selected rules is defined by the parameter rules/iteration [NR], where $NR = N + N$ (N -best rules and N worst rules). In this experiment, the [NR] was set to 10 to all the configurations. This number was selected so the user can evaluate just a small amount of rules and the classifiers still have a satisfactory amount of classified elements. The selected rules are shown to the

user to be classified in 2 different classes: “Interesting” and “Non-Interesting”. The “Interesting” class contains the rules considered as interesting and will be explored by the user when the approach finishes. On the first iteration, the entire rule set will be explored to make the ranking. From the second iteration, only the rules classified as “Interesting” are considered to make the ranking. Since the ranking is created to suggest rules to be classified by the user, the rules already classified by the user are not considered on the ranking.

In step 3, the unlabeled association rules are classified using a network based label propagation algorithm. It is important to distinguish the classifications made by the user from the ones made by the classifier: the user’s classifications can not be changed, i.e., over the iterations these classifications will be maintained by the approach and used in the training set over all the iterations; on the other hand, the classifier’s classifications can change over the iterations. The change on the rules classified by the label propagation algorithm is made so the approach can refine the results, adapting to the new knowledge informed by the user.

On the last step, a stopping criterion is checked to decide if the approach will finish or will execute again. If it’s decided to continue, the rules classified by the label propagation algorithm will be explored again from Step 2, considering only the ones classified as “Interesting” to create the ranking. However, all the rules are considered on the label propagation phase. The process continues until the stopping criterion is met. In the end, the rules considered as “Interesting”, either by the classification algorithm and by the user, are outputted to the user.

4 Experiments

The experiments were carried out with 2 different objectives: validate the approach and find the best configurations. The evaluation measure is the exploration space reduction, i.e., the percentage of rules that doesn’t need to be explored in order to find all the interesting rules. High values of exploration space reduction means that the user needs to explore fewer rules to find the ones he is looking for; if the reduction space is 60% for example, the user needs to explore only 40% of the entire rule set. Both the approach validation and the analysis of the best configurations were performed on the exploration space reduction. The analysis is made considering a small set of rules, called “objective set”. This objective set contains the rules to be found on the exploration, simulating the user’s interests.

The approach was analyzed on 8 data sets (balance-scale, breast-cancer, car, ecoli, habermann, iris, tic-tac-toe and zoo), all of them available on UCI¹. These data sets were processed and converted to transactions. The transactions with missing values were removed. The support and confidence values were empirically defined to generate an amount of rules between 1.000 and 2.000. This number of rules was selected to obtain a good trade-off between exploration space (the higher the better) and the number of rules on the objective set (explained below).

¹ <http://archive.ics.uci.edu/ml/>.

Two different kinds of networks were selected to carry out the experiments: homogeneous and bipartite heterogeneous. The homogeneous network allows an exploration based on the similarity among the rules. The bipartite network allows an exploration based on the items shared by the rules and based on the rules shared by the items. Therefore, the homogeneous network have three different network types [NT]: kNN, in which a rule is linked to its K most similar rules; Gaussian, that changes the weight among the connections by applying the weight to a gaussian function; Conventional, that makes no change on the generated network, maintaining all the original connections and weights. No modifiers were used on the bipartite heterogeneous network.

The number of rules per iteration [NR] on both networks were set to 10 (5 best + 5 worst rules per iteration) aiming to reduce the number of rules to be classified by the user without losing performance on the label propagation algorithms. The centralities measures [NM], used to create rankings on both homogeneous and bipartite heterogeneous, and were: degree, which analyses the rule's importance based on its connections to its neighbor; PageRank [8] that measures the rule's importance based on the entire network connections.

To generate the homogeneous networks, the similarities [S] were used: hoJaccard (Eq. 2, that uses as base Eq. 1) that calculates the similarity between two rules based on the items they share, considering the rule's antecedent and consequent separately; hoConfidence (Eq. 3) that calculates the similarity between two rules based on the transactions they share. The confidence was adapted here to work as a similarity measure because of it can calculate how the occurrence of a rule items can contribute to the occurrence of the items on the other rules. In Eqs. 1, 2 and 3, RL_x represents the x -est rule on the rule set, $LHS(RL_x)$ and $RHS(RL_x)$ represent the RL_x antecedent and consequent, respectively, $\#(T(RL_x))$ represents the number of transaction that contains the rule RL_x . To classify the homogeneous networks the GFHF and LLGC classification algorithms were used. These algorithms were selected due to their great results obtained on the literature.

$$Jacc(RL_1, RL_2) = \frac{RL_1 \cap RL_2}{RL_1 \cup RL_2} \tag{1}$$

$$hoJaccard(RL_1, RL_2) = \frac{Jacc(LHS(RL_1), LHS(RL_2)) + Jacc(RHS(RL_1), RHS(RL_2))}{2} \tag{2}$$

$$hoConfidence(RL_1, RL_2) = \frac{\#(T(RL_1) \cup T(RL_2))}{\#T(RL_1)} \tag{3}$$

The bipartite heterogeneous network was modeled so the weight among the connections between items and their respective rules was an objective measure. Two similarity measures [S] were selected: bipartiteJaccard (Eq. 4) that calculates the amount of transactions shared by the rule's LHS and RHS and the traditional confidence measure. The bipartiteJaccard uses the transactions, instead of the items, so it can get the similarity among the items that are on the LHS

and the item on the RHS of the rule. The connection among two different rules was made by the items they share, where the item is a connection point that connects two rules with two different weights. In equation, $LHS(RL_1)$ returns the rule's left-hand side, $RHS(RL_1)$ return the rule's right-hand side and $\#T(RL_X)$ returns the number of transactions that contains the items on RL_X . To classify the bipartite heterogeneous networks the LPBHN and GNetMine algorithms were used. The GNetMine have the same classification bias as LLGC. Also, the LPBHN have the same bias as GFHF. This allows us to make a fair comparison and analysis to which network type provides a better space reduction.

$$bipartiteJaccard(RL_1) = \frac{\#T(LHS(RL_1)) \cap \#T(RHS(RL_1))}{\#T(LHS(RL_1)) \cup \#T(RHS(RL_1))} \quad (4)$$

In this paper, the user's classification, Step 2 in Fig. 2, was simulated using a set of rules as an objective set. These objective sets are the set of rules to be found on the rule set, simulating the user's interests. Two different groups of objective sets were generated to simulate different types of users. The first group is the random objective set. This set is generated by randomly selecting a total of 1% of the rule set size. The other group consisted on randomly selecting one rule in the rule set and creating a similarity ranking among the selected rule and the entire rule set. The ranking is used to select 1% of the most similar rules to be added into the objective set. The similarity used to calculate the objective set was Jaccard (Eq. 1), that calculates the number of items the rules share divided by the number of distinct items. It is important to emphasize that this measure was not directly used to generate the networks. Because the objective sets were randomly generated, 30 sets were created using each approach. So, for each rule set and each PAR_{LP} configuration, the experiments were executed considering the two groups of user's interests simulation. The random objective set (first group) simulates the users that have a more widely interest on the rule set, with the interesting rules spread across the network. The similarity objective set (second group) simulates the users that have a more specific interest on the rule set.

Based on the objective sets, the user's classification is simulated considering a threshold to be calculated on the first iteration. This threshold is the mean similarity among the rules to be classified by the user and the rules on the objective set, based on the closest similarity of each rule, i.e., for each rule to be classified by the user, the highest similarity is considered to all rules on the objective set. The threshold can be seen in Eq. 5, where $ruleObj$ is the set of rules on the objective set, $ruleCl$ is the set of rules to be classified by the user on the current iteration and nCl is the number of rules to be classified by the user (equals $N + N$ as previously explained). Therefore, in each iteration, the mean similarity among the rules to be classified and the objective set is calculated and compared to the threshold, labeling the rule as "Interesting" if the mean similarity is greater than or equal to the threshold or "Non-Interesting" if the similarity is smaller.

$$Threshold = \frac{1}{nCl} \sum_{i=1}^{nCl} \max(\text{similarity}(\text{ruleCl}_i, \text{ruleObj})) \quad (5)$$

Finally, regarding the stopping criteria, the approach was executed until all the rules on an objective set were classified as “Interesting” either by the user or by the classifier. The experiments were carried out and an analysis was made, aiming to find all the rules on each objective set and looking for the reduction on the exploration space, remembering that the greater the reduction the better the result. In the end, a mean reduction was calculated to each PAR_{LP} configuration (Table 1) and each objective set configuration (random and similarity). Therefore, the results shown in the next section are the mean of 30 executions considering each kind of objective set individually, for each PAR_{LP} configuration.

Table 1. PAR_{LP} configurations

Network	Network type [NT]	Network measure [NM]	Similarity [S]	Classifier
Homogeneous	Knn (K = 10, 20, 30, 40, 50); Gaussian ($\sigma = 0.25, 0.50, 0.75$); conventional	Output degree; PageRank	hoJaccard; hoConfidence	LLGC ($\alpha = 0.1, 0.3, 0.5, 0.7, 0.9$); GFHF
Bipartite heterogeneous	Conventional	Output degree; PageRank	bipartideJaccard; confidence	GNetMine ($\alpha = 0.1, 0.3, 0.5, 0.7, 0.9$); LPBHN

5 Results and Discussion

The experiments were carried out on each data set using all the described PAR_{LP} configurations (Table 1). For each configuration, the PAR_{LP} approach was applied over all the 60 objective sets to test two different strategies of simulating the user’s interests: one containing scattered rules, for users that want to explore the rule set without looking for a specific “theme”, and one containing more similar rules, for users that have a specific area of interest. The results were analyzed considering the mean of the 30 executions on each PAR_{LP} configuration and each group of objective sets, which means that each PAR_{LP} configuration has 2 different results: one for the randomly generated objective sets and one for the objective sets generated based on the similarity.

Table 2 shows the best and the worst results obtained on each data set. The first column presents the data set used to generate the rules. The second column shows the best result obtained using the random objective sets. The third column shows the worst result obtained using the random objective sets. The fourth and fifth columns show, respectively, the best and worst results obtained using the similarity objective sets. Remember that the experiments were analyzed

based on the exploration space reduction, which means that the values on these columns represent the percentage of rules that the user doesn't need to explore. For example, a 60 % exploration space reduction means that the user needs to explore only 40 % of the rules to find all the interesting ones.

The results show that the objective sets generated by the similarity strategy obtained better results in comparison to the random objective sets. These results can show that an exploration guided by some "theme" or by some related topics will result in a higher reduction than an exploration where the user explores by selecting dissimilar rules as "Interesting". These better results occurred because of the label propagation algorithms, used to classify the rules, since these algorithms put similar knowledge on the same class and dissimilar knowledge on different classes. By simulating the user's classification and selecting dissimilar rules as "Interesting" the label propagation algorithms classify different parts of the network as "Interesting", increasing the number of rules that are considered "Interesting" in comparison to the objective sets that classify only one point on the network. Also, the best reduction was 66.49%, which means that the user will only have to explore about $\frac{1}{3}$ of the entire rule set. The table also shows a variation among the data sets, which means that the domain can contribute to the exploration. The best data set (iris) reduced almost 50 % more in comparison to the worst data set (breast-cancer).

As seen, Table 2 doesn't link the results to the PAR_{LP} configurations, since the analysis performed were based on the overall quality of the results. Therefore, a second analysis was performed aiming to find the best PAR_{LP} configurations on the rule sets. Initially, all the results (considering all the 8 data sets, PAR_{LP} configurations and both objectives sets strategies) were divided into 20 groups considering network, network type [NT] and classifier. From these 20 groups, the 2 best results of each group were selected through a statistical test, coming up with 40 PAR_{LP} configurations. After that, another statistical test was done, based on these 40 PAR_{LP} configurations, through Friedman $N \times N$ with Nemenyi as post-test, selecting the 20 best configurations. Figure 3 shows the results of the final statistical analysis. The horizontal lines show that the PAR_{LP} configurations have no statistical difference. The configurations are described using the following pattern: network type [NT] - network parameter

Table 2. Best and worst exploration space reductions obtained on each data set

Data set	# Gen. Rules	Best Rd	Worst Rd	Best Sim	Worst Sim
balance-scale	1746	40.66%	4.81%	63.69%	29.50%
breast-cancer	1602	19.98%	5.37%	42.45%	4.56%
car	1326	15.91%	4.68%	52.64%	22.17%
ecoli	1685	28.66%	4.87%	51.57%	21.01%
habermann	1006	46.12%	9.15%	58.45%	29.72%
iris	967	51.71%	10.13%	66.49%	39.50%
tic-tac-toe	1317	37.05%	4.02%	61.88%	16.02%
zoo	1658	30.88%	4.40%	46.38%	17.13%

(when applied) - classifier - classifier parameter (when applied) - similarity measure [S] - rank measure [NM]. The first configuration, for example, is kNN, with $k = 50$, using the GFHF classifier, hoJaccard as a similarity measure and PageRank as the ranking measure. It is possible to see that the kNN network, together with the GFHF classifier, obtained the overall best results, being on 9 out of 10 best results.

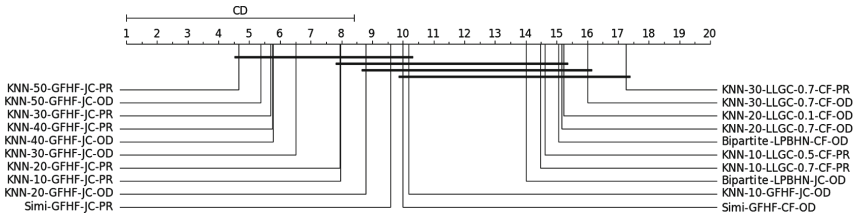


Fig. 3. The statistical analysis results

6 Conclusion

This paper proposed the PAR_{LP} , an iterative and interactive post-processing association rules approach that uses networks and label propagation algorithms to extract interesting and non-interesting rules according to user’s knowledge. The paper presented the PAR_{LP} structure and carried out some experiments using a set of configurations and a user classification simulation. The obtained results were shown and discussed according to the total reduction on the exploration space, i.e., the number of rules not to be exploited by the user.

The results show that the PAR_{LP} is capable of finding a set of rules according to the interactions made during the process, meaning that the PAR_{LP} can be used successfully to direct the exploration of the rules according to the interactions made with the algorithm, reducing the amount of rules to be explored and directing the exploration to the rules chosen by the user interaction. This reduction can be increased when the rules to be found by the user are more similar, or share the same “theme”. Even on the random objective sets, that contain more widespread rules, the obtained reductions were satisfactory. Also, the analysis on the 20 best configurations shows that the kNN network, together with GFHF classifier, are the most promising configuration.

As future works a further analysis on the way the rules sets are modeled will be done aiming to find the characteristics that result on greater exploration space reduction. This analysis will be carried out aiming to discover beforehand which similarity measure to be used according to the rule set features. Also, the algorithms complexity will be reduced. The actual configuration is expensive e can not be used on larger rule sets. To solve the “theme” guided exploration, the proposed approach will be extended to consider different user’s interests.

Acknowledgments. We would like to thank CAPES (PROEX-8434242/D) and FAPESP: Grant 2014/08996-0, São Paulo Research Foundation (FAPESP) for the financial aid.

References

1. Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. In: ACM SIGMOD (1993)
2. de Carvalho, V.O., de Padua, R., Rezende, S.O.: Semi-supervised learning to support the exploration of association rules. In: Bellatreche, L., Mohania, M.K. (eds.) DaWaK 2014. LNCS, vol. 8646, pp. 452–464. Springer, Heidelberg (2014)
3. de Sousa, C.A.R., Rezende, S.O., Batista, G.E.A.P.A.: Influence of graph construction on semi-supervised learning. In: Blockeel, H., Kersting, K., Nijssen, S., Železný, F. (eds.) ECML PKDD 2013, Part III. LNCS, vol. 8190, pp. 160–175. Springer, Heidelberg (2013)
4. Ji, M., Sun, Y., Danilevsky, M., Han, J., Gao, J.: Graph regularized transductive classification on heterogeneous information networks. In: Balcázar, J.L., Bonchi, F., Gionis, A., Sebag, M. (eds.) ECML PKDD 2010, Part I. LNCS, vol. 6321, pp. 570–586. Springer, Heidelberg (2010)
5. Newman, M.E.J.: *Networks: An Introduction*. Oxford University Press, Oxford (2010)
6. Pandey, G., Chawla, S., Poon, S., Arunasalam, B., Davis, J.G.: Association rules network: definition and applications. *Stat. Anal. Data Min.* **1**, 260–279 (2009)
7. Padua, R., Carvalho, V.O., Rezende, S.O.: Post-processing association rules using networks and transductive learning. In: 13th ICMLA, pp. 318–323 (2014)
8. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank citation ranking: bringing order to the web. Technical report, Stanford InfoLab (1999)
9. Rossi, R.G., Lopes, A.A., Rezende, S.O.: A parameter-free label propagation algorithm using bipartite heterogeneous networks for text classification. In: ACM SAC, pp. 79–84 (2009)
10. Yang, G., Shimada, K., Maby, S., Hirasawa, K., Hu, J.: A genetic network programming based method to mine generalized association rules with ontology. *JACIII* **12**, 63–76 (2007)
11. Zhou, D., Bousquet, O., Navin Lal, T., Weston, J., Schölkopf, B.: Learning with local and global consistency. *NIPS* **16**, 321–328 (2004)
12. Zhu, X., Ghahramani, Z., Lafferty, J.: Semi-supervised learning using Gaussian fields and harmonic functions. In: ICML, pp. 912–919 (2003)
13. Zhu, X., Goldberg, A.B.: *Introduction to Semi-Supervised Learning*, vol. 6. Morgan & Claypool Publishers, San Rafael (2009)