Anthropocentric visualization of optimal cover of association rules

Amira Mouakher and Sadok Ben Yahia

Faculty of Sciences of Tunis, Tunis, Tunisia.
sadok.benyahia@fst.rnu.tn

Abstract. Visual mining as an emerging field is gradually witnessing approaches where visual capabilities are of use to mine manageable and useful knowledge. In this respect, the implication of the user at the core of the mining process seems to be one of the key factors in the success of any visual mining system. In this paper, we introduce a new anthropocentric visual mining approach for the extraction and the visualization of a reduced cover of association rules. Given that getting out such a reduced cover is NP-hard problem, we introduce a greedy algorithm that relies on correlation assessment metrics to flag discovered formal concepts as optimal. In addition, we present some snapshots illustrating the key features of the implemented visualization tool that relies on the virtual reality paradigm.
Keywords: Formal concept analysis, Optimal cover of binary relation, Visualization of association rules, Virtual reality

1 Introduction and motivations

Within the traditional framework of association rule mining [1], managing the high number of frequent patterns extracted from real-life datasets becomes an important topic. In addition, providing efficient and easy-to-use graphical tools to users is a promising challenge of data mining, especially in the case of association rules. These tools must be able to generate explicit knowledge and, then, to present it in an elegant way. Visualization techniques have shown to be an efficient solution to achieve such a goal. Even though considered as a key step in the mining process, the visualization step of association rules has received much less attention than that paid to the extraction step. Thus, visual representations have been proposed to tackle some of these problems. At first, visualization was employed mainly to assist end users in exploring the extracted rule set, based on visual representations of the rule space. There are also recent examples of employing visual representations to help end users along the execution of the mining algorithm, e.g., to show intermediate results and to allow user feedback during the process.

In this paper, we introduce a new anthropocentric approach for the visualization of a cover of association rules. The latter cover is obtained after the extraction of an optimal cover of the underlying extraction context. The main thrust of this contribution is twofold:

– Reduction of the user overload during the knowledge exploration. Indeed, whenever faced with a deluge of knowledge, a user may be asked to rely on a tedious

and boring application of its perceptual abilities to explore such set of association rules. To prevent such a situation, we devise the extraction of a reduced cover of association rules that is obtained from optimal cover of the underlying extraction context.
- According to Hackman and Oldham [2], the larger the user's skills involvement during the mining process, the stronger sense of independence and responsibility is. Thus, in the Virtual reality based visualization tool that we have introduced, the user is at the core of the full user-driven visualization process. Indeed, he is acting as a "background" pruning medium for shedding light on relevant association rules. Thanks to virtual reality environment, it is possible to create virtual visual worlds based on the characteristics of the data, so that visual data explorers can be immersed in these worlds, navigate around, and observe the data from within.

The rest of the paper is organized as follows. The next section recalls the key notions used throughout this paper. Section 3 describes the OPTCOVER algorithm devised for the extraction of the optimal cover of a binary relation as well as the results of the carried out results. Section 4 presents the Virtual Reality based visualization of association rules tool. Section 5 concludes the paper and points out our future work.

2 Key notions

In this section, we briefly sketch the key notions used in the remainder of this paper.

Definition 1. (BINARY RELATIONS) *A binary relation $\mathcal{R}$ between two finite sets $\mathcal{D}$ and $\mathcal{T}$ is a subset of the Cartesian product $\mathcal{D} \times \mathcal{T}$. An element in $\mathcal{R}$ is denoted by (x, y), where y denotes the image of x by $\mathcal{R}$. For a binary relation, we associate the following subsets [3]:*

- *The set of images of x is defined by: $x.\mathcal{R} = \{y|(x, y) \in \mathcal{R}\}$, we may also denote it as in FCA notation by x'.*
- *The set of antecedent of y is defined by: $\mathcal{R}.y = \{x|(x, y) \in \mathcal{R}\} = y'$.*

Definition 2. (RELATIVE PRODUCT) *Let $\mathcal{R}$ and $\mathcal{R}'$ be two binary relations, we define the relative product of $\mathcal{R}$ and $\mathcal{R}'$ as the relation denoted by: $\mathcal{R}o\mathcal{R}' = \{(x, y)|\exists\, t|(x, t) \in \mathcal{R} \wedge (t, y) \in \mathcal{R}'\}$. "$o$" is denoting the usual associative operator for relational composition, where $\mathcal{I}o\mathcal{R} = \mathcal{R}o\mathcal{I} = \mathcal{R}$, where $\mathcal{I}$ is the identity on the domain $\mathcal{D}$. In addition, we use the following notations:*

- *The inverse of a relation is: $\mathcal{R}^{-1}$=$\{(x, y)|(y, x) \in \mathcal{R}\}$.*
- *The complement of a relation is: $\overline{\mathcal{R}}$=$\{(x, y)|(x, y) \notin \mathcal{R}\}$*
- *The relation $\mathcal{I}$, identity on a set $\mathcal{A}$, is given by the following expression, $\mathcal{I}(\mathcal{A}) = \{(x, x)|x \in \mathcal{A}\}$.*

Definition 3. (FORMAL CONTEXT) *A formal context is a triplet $\mathcal{K}$ = ($\mathcal{O}$, $\mathcal{I}$, $\mathcal{R}$), where $\mathcal{O}$ represents a finite set of objects, $\mathcal{I}$ is a finite set of items and $\mathcal{R}$ is a binary (incidence) relation (i.e., $\mathcal{R} \subseteq \mathcal{O} \times \mathcal{I}$). Each couple $(o,\, i) \in \mathcal{R}$ expresses that the object $o \in \mathcal{O}$ contains the item $i \in \mathcal{I}$.*

	a	b	c	d	e	f	g
1			×		×	×	
2			×		×	×	×
3			×		×		×
4			×			×	
5			×	×		×	
6		×				×	×
7	×	×				×	×

Table 1. A formal context.

Example 1. *In the remainder, we will consider the formal context depicted by Table 1 with $\mathcal{O} = \{1, 2, 3, 4, 5, 6, 7\}$ and $\mathcal{I}$ = {a, b, c, d, e, f, g}.*

Definition 4. (FORMAL CONCEPT) *A pair (A, B) is a formal concept of $(\mathcal{O}, \mathcal{I}, \mathcal{R})$ if and only if $A \subseteq O$, $B \subseteq I$, $A' = B$ and $B' = A$. In other words, (A, B) is a formal concept if the set of all attributes shared by the objects of A is identical with B and on the other hand A is also the set of all objects that have all attributes in B. A is then called the extent and B the entent of the formal concept (A, B). The formal concepts of a given context are naturally ordered by the subconcept-superconcept relation as defined by:*

$$(A_1, B_1) \leqslant (A_2, B_2) \Longleftrightarrow A_1 \subseteq A_2 (\Longleftrightarrow B_2 \subseteq B_1)$$

Definition 5. (PERTINENT FORMAL CONCEPT) *A formal concept $FC_i = (A_i, B_i)$ containing the couple (a, b) of the formal context $\mathcal{K}$ = $(\mathcal{O}, \mathcal{I}, \mathcal{R})$ is said to be* pertinent, *if it maximizes the value of an interestingness measure,* `w.r.t` *to all other formal concepts containing the couple (a, b).*

Definition 6. (PERTINENT COVER) *Given the formal context $\mathcal{K} = (\mathcal{O}, \mathcal{I}, \mathcal{R})$. A pertinent cover associated with the formal context $\mathcal{K}$ is the cover constituted of pertinent formal concepts covering all the couples of $\mathcal{R}$.*

Definition 7. (PSEUDO-CONCEPT) *The pseudo-concept containing the couple (a, b) is the union of all the formal concepts containing (a, b). The pseudo-concept, denoted PFC$_{ab}$, is computed by getting the restriction of $\mathcal{R}$ to the set of examples described by b, i.e., $\phi(b)$ and the set of properties describing the object a, i.e., $\psi(a)$, where (ϕ, ψ) are the Galois connexion operators. Formally,*

$$PFC_{ab} = I(b.\mathcal{R}^{-1}) o\mathcal{R} o I(a.\mathcal{R})$$

Example 2. *Considering the extraction context $\mathcal{K}$ given by Table 1, then the pseudo-concept containing the couple $(1, b)$ is PFC_{1b}=$(67, cef)$.*

Definition 8. (ASSOCIATION RULE) *An association rule R is a relation between itemsets and is of the form $R : X \Rightarrow (X \backslash Y)$, such that X and Y are frequent itemsets, and $X \subset Y$. The itemsets X and $(Y \backslash X)$ are, respectively, called the* premise (or antecedent) *and the* conclusion (or consequent) *of the association rule R.*

Definition 9. (SUPPORT OF AN ASSOCIATION RULE) *The support of R, $Supp(R)$, is a measure of statistical significance which represents the number of transactions that include all items in the antecedent and consequent parts of the rule.*

Correlation measures

Many correlation measures have been proposed in the literature [4–6]. Indeed, as shown in [7], the discovery of frequent patterns which satisfy these measures will be of benefit in the reduction of high added value association rule number, since only informative ones will be derived from the highly correlated patterns. Moreover, the extracted correlated patterns have been shown to be very useful in various application domains, text mining, bioinformatics, market basket study, and medical data analysis [8], to cite but a few. In the following, we review some of the most used correlation measures.

2.1 The *confidence* correlation measure

A rule has a measure of its strength called *confidence* defined as the ratio of the number of transactions that include all items in the consequent as well as the antecedent (namely, the support) to the number of transactions that include all items in the antecedent.

$$Conf(R) = \frac{Supp(Y)}{Supp(X)}$$

2.2 The *Lift* correlation measure

The *lift* Ratio of an association rule is defined as follows:

$$Lift(R) = \frac{Conf(R)}{Pr(R)}$$

where $Pr(R)$ is called the *expected confidence* which is defined as the number of transactions having the consequent items divided by the total number of transactions. Comparing *lift* with *confidence*, the *lift* of a rule is a relative measure in the sense that it compares the degree of dependence in a rule versus independence between the consequent items and the antecedent items. The rules that have higher *lift* will have higher dependence in them.

2.3 The *any-confidence* correlation measure

Considering this measure, an association is deemed interesting if **any** rule that can be produced from that association has a confidence greater than or equal to our minimum any-confidence value. The *any-confidence* measure of a non empty pattern $X \subseteq \mathcal{I}$ is defined as follows [7] :

$$any\text{-}conf(X) = \frac{Supp(\wedge X)}{min\{Supp(\wedge i) | i \in X\}}$$

The numerator of the descriptive and symmetric *any-confidence* measure represents the conjunctive support of X. Whereas, the denominator represents the minimum of the conjunctive supports of items $i \in X$.

2.4 The *all-confidence* correlation measure

It is a variation of the previous measure. With this one, an association is deemed interesting if **all** rules that can be produced from that association have a confidence greater than or equal to our minimum all-confidence value. This indicates that there is a dependency between all of the items in the association. The degree of the dependency, of course, is based on the threshold value. The *all-confidence* measure of a non empty pattern $X \subseteq \mathcal{I}$ is defined as follows [7] :

$$\textit{all-conf}(X) = \frac{\textit{Supp}(\wedge X)}{\textit{max}\{\textit{Supp}(\wedge i) | i \in X\}}$$

2.5 The *bond* correlation measure

The *bond* measure [7] (*aka Coherence* [9], *Tanimoto coefficient* [10] and *Jaccard* [11]), computes the ratio between the conjunctive support and the disjunctive one. Thus, the *bond* measure of a non empty pattern $X \subseteq \mathcal{I}$ is defined as follows:

$$\textit{bond}(X) = \frac{\textit{Supp}(\wedge X)}{\textit{Supp}(\vee X)}$$

3 Extracting optimal coverage from a binary relation

Finding optimal cover of binary relation is known to be NP-hard problem [12]. Nevertheless, we witness a large number of works interested in tackling such a problem. Belkhiter et al. [13] introduced an optimal rectangular decomposition of a binary relation as well as an application to documentary databases. The introduced decomposition is based on the election of optimal maximal rectangles (or equivalently formal concepts) that achieve a maximal gain in storage space terms. Formally, the authors introduced the gain function of a maximal rectangle $R = (A, B)$ as follows:

$$gain(R) = (|A| \times |B|) - (|A| + |B|)$$

Later Kcherif et al. [14] introduced a rectangular decomposition approach based on the Riguet's difunctional relation [15]. The computation of this difunctional is reduced to the determination of a set of isolated points allowing the determination of the minimal set of rectangles covering a given binary relation. Recently, Belohlavek and Vychodil [16] tackled the same issue by attempting to solve the *Boolean factor analysis* problem by proposing a new method of decomposition of an $n \times m$ binary matrix I into a Boolean product AoB of an $n \times k$ binary matrix A and a $k \times m$ binary matrix B with k as small as possible.

In this paper, we introduce a new approach, based on a greedy algorithm, for the extraction of an optimal cover of a binary relation. It heavily relies on the formal concept lattice representation. This structure has the advantage of representing the context by formal concepts without information loss. The fact that the lattices's size grows exponentially with the number of attributes and includes a high redundancy of information, seriously hampers its scalability. Thus, the solution is to produce an optimal set of formal concepts representing the lattice. This set which is called *optimal cover of formal* *concepts* has to faithfully represent the context without loss of information. The main criticism that can be of the literature approaches is the fact that a cover is extracted regardless of the quality of knowledge that may be drawn from it. That's why in our approach the gain function is based on the assessment of the correlation (or the strength of the association rules) of the intent part of pertinent formal concepts.

3.1 The OPTCOVER algorithm for building a cover of pertinent formal concepts

In our approach, we have adopted a greedy algorithm for discovering the optimal cover of a binary relation. It consists in building a cover of pertinent formal concepts to reduce the dimensionality of data handled by the user which is heavily involved into this process. Informally, the OPTCOVER algorithm operates as follows :

1. For each couple belonging to the context, compute its pseudo-concept.
2. From each pseudo-concept, extract its enclosed formal concepts.
3. Selection of the concept which maximises the correlation measure's value.
4. Validation or rejection of the chosen formal concept by the user.
5. Elimination of the couples formed by the chosen concept.

At each iteration, the algorithm supplies a pertinent formal concept according to a user set correlation measure. The user is able to navigate into the rule set generated from each formal concept and confirm or reject the choice given by the algorithm and propose another alternative. Finally, we obtain a set of pertinent formal concepts constituting an optimal cover of the formal context. The pseudo-code of OPTCOVER algorithm is sketched by Algorithm 1.

Algorithm 1: OPTCOVER

Data: **- A formal context** $\mathcal{K} = (\mathcal{O}, \mathcal{I}, \mathcal{R})$.
- The correlation measure $\mathcal{H}$.
- The minimum threshold $min_{\mathcal{H}}$.
Results: **- A pertinent coverage of formal concepts** $\mathcal{F}_c$.

Begin
  Set φ to $\{(o, i) \in \mathcal{R}\}$;
  Set $\mathcal{F}_c$ to $\emptyset$;
  Foreach $(o, i) \in \varphi$ **do**
    $PFC_{(o,i)}$=GETPSEUDOCONCEPT($\mathcal{K}, o, i$);
    L_c=GENERATECONCEPTS($\mathcal{K}, PFC_{(o,i)}$);
    (A, B)=COMPUTEPERTINENTCONCEPT($\mathcal{K}, L_c, \mathcal{H}, min_{\mathcal{H}}$);
    $\mathcal{F}_c = \mathcal{F}_c \cup (A, B)$;
    Foreach $(a, b) \in A \times B$ **do**
      Remove (a,b) from φ;
  Return $\mathcal{F}_c$;
End

Iteration 1	**Iteration 2**
$PFC_{(1,c)} = (12345, cef)$ L_c={{**(123,ce)**}; $\{(12345, c)\}; \{(1245, cf)\}\}$	$PFC_{(1,f)} = (124567, cef)$ L_c = {{**(12,cef)**}; $\{(124567, f)\}; \{(1245, cf)\}\}$
Iteration 3	**Iteration 4**
$PFC_{(2,g)} = (2367, cefg)$ L_c ={{**(23,ceg)**}; $\{(2367, g)\}; \{(267, fg)\}\}$	$PFC_{(4,c)} = (12345, cf)$ L_c ={{**(12345,c)**}; $\{(1245, cf)\}\}$
Iteration 5	**Iteration 6**
$PFC_{(4,f)} = (124567, cf)$ L_c = {{**(124567,f)**}; $\{(1245, cf)\}\}$	$PFC_{(5,d)} = (5, cdf)$ L_c ={{**(5,cdf)**}}
Iteration 7	**Iteration 8**
$PFC_{(6,f)} = (67, bfg)$ L_c = {{**(67,bfg)**}}	$PFC_{(7,a)} = (7, abfg)$ L_c = {{**(7,abfg)**}}

Table 2. Running track of the OPTCOVER algorithm

Example 3. *Let us consider the formal context depicted by Table 1. Applying the* OPTCOVER *algorithm for Bond as a correlation measure and minbond = 0.75, the running track is sketched by Table 2. At first, all the couples of the context are decreasingly sorted with respect to the associated pseudo-concept support value. Then, for each couple (o, i), we compute its pseudo-concept $PFC_{(o,i)}$, from which we extract the set of formal concepts denoted L_c. Depending of the Bond value of each formal concept, the* OPTCOVER *algorithm elects the pertinent formal concept. The latter is chosen as the one that maximizes the Bond value from those belonging to L_c. Pertinent formal concepts are marked in bold in Table 2. After keeping track of the pertinent formal concept in the $\mathcal{F}_c$ list, we remove all the couples covered by the pertinent formal concept of the current iteration. The* OPTCOVER *algorithm comes to an end after performing eight iterations and checking that all the couples of the formal context are covered. Thus, the* OPTCOVER *algorithm yields the following output:*

$$\mathcal{F}_c = \{(123, ce), (12, cef), (23, ceg), (12345, c), (124567, f), (5, cdf), (67, bfg), (7, abfg)\}$$

3.2 Experimental results

We compare, through various experiments, the size of our optimal cover *vs* that of the set of formal concepts drawn from the considered datasets. Characteristics of tested benchmark datasets are summarized in Table 3 (1).

As stated in the well known “No Free Lunch Theorem“[17], no interestingness measure can be considered as the outstanding one in compacity terms. Nevertheless, the Bond based cover gives in most cases the most reduced cover as sketched by Table 3. In addition, the compacity rates provided by all considered correlation measure are worth of mention `w.r.t.` to the number of all formal concepts.

At a glance, Table 4 shows that Bond’s based cover is the most reduced one in compacity terms for all considered datasets. Similar results were obtained by Belohlavek’s

[1] These datasets are available at the UC Irvine Machine Learning Repository: *http://archive.ics.uci.edu/ml/*.

Dataset	# **items**	# **objects**	#**formal concepts**	$\#(Lc)$	$\#(Cc)$	$\#(Bc)$
SHUTTLE-LANDING-CONTROL	15	24	52	**18**	23	19
ADULT-STRETCH	20	10	89	14	16	**9**
LENSES	24	12	128	23	32	**12**
ZOO	101	28	377	52	40	**25**
HAYES-ROTH	132	18	380	53	45	**17**
SERVO	167	19	432	59	72	**19**
POST-OPERATIVE	90	25	1521	60	55	**22**

Table 3. Dataset characteristics and the obtained covers.(Lc): Lift based Cover, (Cc): Confidence based Cover, (Bc): Bond based Cover.

Dataset	#**formal concepts**	# **"Bond" cover**	# **Belkhiter's cover**	# **Belohlavek's cover**
SHUTTLE-LANDING-CONTROL	52	**19**	20	22
ADULT-STRETCH	89	**9**	14	10
LENSES	128	**12**	21	12
ZOO	377	**25**	43	26
HAYES-ROTH	380	**17**	49	17
SERVO	432	**19**	60	19
POST-OPERATIVE	1521	**22**	52	22

Table 4. Comparaison with literature approaches in terms of cover compacity.

cover [16] which has been shown to provide very interesting rates. Nevertheless, this latter cover do not provide the most interesting association rules from a knowledge extraction point of view. In fact, during the generation of the Bond based cover, the stress is put on the compacity of the cover, which may hamper the extraction of significant or valuable knowledge for end users.

4 Virtual reality based Visualization of association rules

Intensive studies have been made about efficient association rule mining algorithms from large data sets in the last decade. Data visualization currently plays an important role in knowledge discovery process as it helps miners to create and validate hypotheses about the data and also to track and understand the behavior of mining algorithms [18].

Visual data mining is a young and emerging discipline that combines knowledge and techniques from a number of areas. The ultimate goal of visual data mining is to devise visualizations of large amounts of data that facilitate the interpretation of the data. The goal of the visual data mining is to help a user to get a feeling for the data, to detect interesting knowledge, and to gain a deep visual understanding of the data set [19]. Data visualization currently plays an important role in knowledge discovery. Indeed the KDD process is by nature highly iterative and interactive and requires user involvement. Visualization techniques are a very effective means of introducing the necessary human subjectivity into each step of the process while taking advantage of the human perceptual and cognitive capabilities [20]. Interestingly enough, for data mining to be effective, it is important to include the human in the data exploration process and combine the flexibility, creativity, and general knowledge of the human with the enormous storage capacity and the computational power of todays computers [21]. In this respect, the virtual reality based technology was proposed in order to improve visualization and users interaction. This approach uses 3D representation based on a metaphor such as a landscape or benches [22]. A metaphor consists in substituting a domain by another domain in order to simplify it [23].

The main added value of a virtual reality based solution stands on the fact that it sets itself the target to edit the result of a data mining process in that way, that the user gets a widespread and clear view of the relevant data interactively visualized in 3D. This allows the user to navigate trough and manipulate a scene in which the extracted knowledge is represented.

The visualization approach that we introduce falls within the fully user-driven rule extraction process classification. In fact, an active participation is allowed to users to input previous knowledge about the domain and/or knowledge acquired during the mining task into the process itself. Thus, the user has the possibility to interfere, e.g., to choose the “pertinent“ formal concept, as well as a full control of the mining process, i.e., the user is allowed to backtrack and resume the process with different parameter settings.

Description of some snapshots :

During each iteration, the acting scene is dynamically split by line to separate the different concepts which are represented by cones. The latter are scrambled on the virtual ans their displaying is in snugness dependency with their pertinence, i.e., the higher the pertinence value, the higher the cone’s visibility is. Thus, the user is encouraged to get insights through the most pertinent concepts. During the exploration process, the user can also visualize association rules which plot as 3D histograms: the premise part is shown along the abscisses axis while the conclusion part is drawn on the y-axis. The height of the histogram indicates the value of the pertinence value associated to the considered association rule. Figure 1 (Up) represents the main interface that the user would obtain when he runs the visualization tools. The interface is split into three subspaces:

1. **Settings panel** : The user is asked to set some key settings before proceeding with the extraction process, e.g., choice of the input file, interestingness measure, thresholds and the gadget of visualization manipulation (keyboard, mouse, gant, wiimote).
2. **Panel of mining process control** : this panel permits to control the mining process by offering to launch the process or to pause it. In addition, the user has the ability

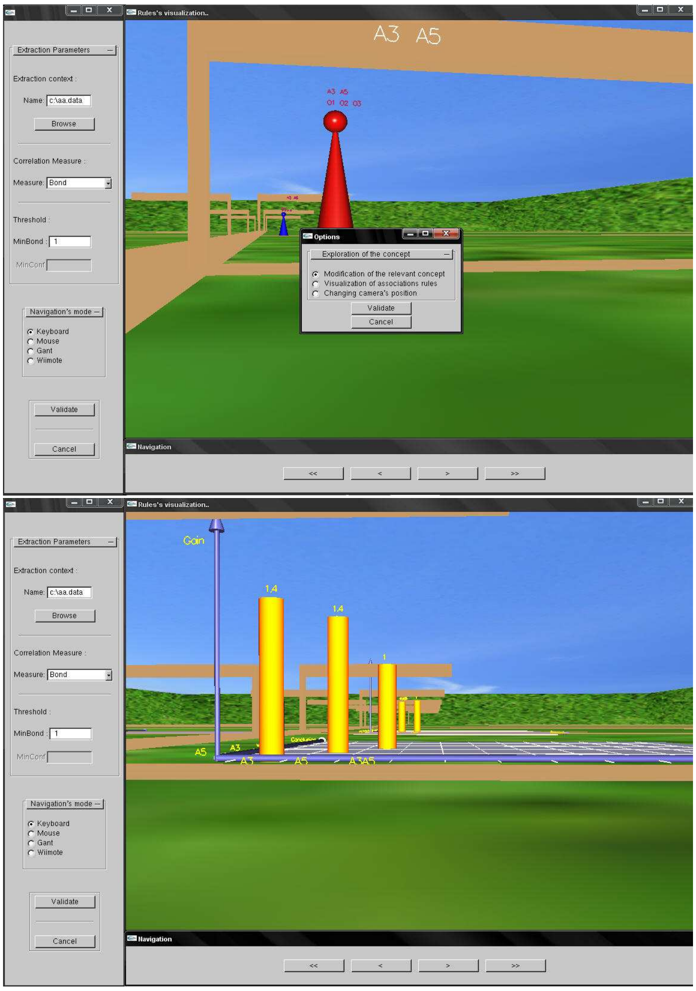

Fig. 1. (**Up**) Formal concept sketched as a cone (**Down**) Getting inside association rules

to backtrack or to move forward without human intervention until the computation of the final output.

3. **Visualization zone**.

Once the setting step is performed, the user can control the mining process. The formal concepts extracted within each iteration are visualized in the dedicated visualization zone. Formal concepts are displayed with the cone metaphor and its height sketches the importance of its gain value. Thus, in each iteration, the pertinent formal concept is the highest one and is automatically recommended to the user. The formal concepts'label is composed by a concatenation of its intent part elements. Formal concepts are put on the scene in a way favoring, in visibility terms, the selection of the pertinent formal concept. In addition, the user can freely navigate through the 3D displaying to scrutinize the remaining formal concepts by getting a closer look as far as the point of view on the scene is modified.

Figure 1 (Up) shows that by performing a double click on the selected formal concept, the user has the possibility to visualize the association rules that may be drawn from such a formal concept. Such set of rules is plot as 3D histograms and clustered `w.r.t` the formal concept used to generate them (cf.,Figure 1 (Down)). By doing so, the user feels free to choose the formal concept after the analysis of the pertinence from the knowledge that may be extracted from such a formal concept. In addition, thanks to dedicated buttons, the user can backtrack in the mining process or decide to skip the process details by asking the prototype to proceed without human intervention.

5 Conclusion and future work

In this paper, we presented a new anthropocentric visualization of a cover of association rules. This cover is obtained after the extraction of an optimal cover of the underlying formal context. The rationale behind the introduced gain function stands on the assessment of the correlation value of the association rule that may be drawn from it. The virtual reality based visualization tool permitted a fully user-driven extraction process.

Other avenues for future work mainly address the following points: (*i*) Extensive evaluation of the intuitiveness and ease of use of the visualization prototype; (*ii*) Study of the derivation of generic basis of association rules from the induced covers as well as the definition of the generation process by means of sound and complete axiomatic system.

References

1. Ceglar, A., Roddick, J.F.: Association mining. In ACM Computing Surveys, **volume 38(2),** (2006)
2. Hackman, J., Oldham, G.: Development of the job diagnostic survey. Journal of Applied Psychology **60** (1975) 159–170
3. Schmidt, G., Stroehlein, T.: Relations and Graphs: Discrete Mathematics for Computer Scientits. Springer-Verlag (1993)

4. Ke, Y., Cheng, J., Yu, J.X.: Efficient discovery of frequent correlated subgraph pairs. In: Proceedings of the 9th IEEE International Conference on Data Mining, Miami, Florida, USA (2009) 239–248
5. Kim, W.Y., Lee, Y.K., Han, J.: CCMINE: efficient mining of confidence-closed correlated patterns. In: Proceedings of the 8th International Pacific-Asia Conference on KDD, Sydney, Australia (2004) 569–579
6. Le Bras, Y., Lenca, P., Lallich, S.: Mining interesting rules without support requirement: a general universal existential upward closure property. Annals of Information Systems **8** (2010) 75–98
7. Omiecinski, E.R.: Alternative interest measures for mining associations in databases. IEEE Transactions on Knowledge and Data Engineering **15** (2003) 57–69
8. Xiong, H., Shekhar, S., Tan, P.N., Kumar, V.: Taper: A two-step approach for all-strong-pairs correlation query in large databases. IEEE Transactions on Knowledge and Data Engineering **18** (2006) 493–508
9. Lee, Y.K., Kim, W.Y., Cai, Y.D., Han, J.: COMINE: efficient mining of correlated patterns. In: Proceedings of the 3rd IEEE International Conference on Data Mining, Melbourne, Florida, USA (2003) 581–584
10. Tanimoto, T.T.: An elementary mathematical theory of classification and prediction. Technical Report, I.B.M. Corporation Report (1958)
11. Jaccard, P.: Nouvelles recherches sur la distribution florale. Bulletin de la Société Vaudoise des Sciences Naturelles **44** (1908) 223–270
12. Garey, M., Jonhson, D.: Computers and Intractability: A guide to the Theory of NP-Completeness Relations and Graphs: Discrete Mathematics for Computer Scientits. W.H. Freeman (1979)
13. Belkhiter, N., Bourhfir, C., Gammoudi, M., Jaoua, A., Thanh, N.L., Reguig, M.: Décomposition Rectangulaire Optimale d'une Relation Binaire: Application aux bases de Données Documentaires. INFOR **1** (1994) 33–54
14. Khchérif, R., Gammoudi, M.M., Jaoua, A.: Using difunctional relations in information organization. Information Sciences **125** (2000) 153–166
15. Riguet, J.: Relations binaires, fermetures et correspondances de Galois. Bulletin de la société mathématique de France **76** (1944) 114–155
16. Belohlavek, R., Vychodil, V.: Discovery of optimal factors in binary data via a novel method of matrix decomposition. Journal of Computer and System Sciences **76** (2009) 3–20
17. Wolpert, D.H., Macready, W.: No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation **1** (1997) 67–82
18. de Oliveira, M., Levkowitz, H.: From visual data exploration to visual data mining: A survey. IEEE Transactions On Visualization And Computer Graphics **9** (2003) 378–394
19. Böhlen, M.H., Bukauskas, L., Mazeika, A., Mylov, P.: The 3dvdm approach: A case study with clickstream data. In: Visual Data Mining. (2008) 13–29
20. Blanchard, J., Guillet, F., Briand, H.: Exploratory visualization for association rule rummaging. In: Proceedings of the Third IEEE International Conference on Data Mining, Washington, DC, IEEE Computer Society (2003) 493–496
21. Keim, D.A.: Information visualization and visual data mining. IEEE Transactions On Visualization And Computer Graphics **7** (2002) 100–107
22. Blanchard, J., Guillet, F., Briand, H.: A user-driven and quality-oriented visualization for mining association rules. In: Proceedings of the Third IEEE International Conference on Data Mining, Melbourne, Florida. (2003) 493–496
23. Simoff, S.J.: Towards the development of envorinments for designing visualization support for visual data mining. In: Proceedings of the International Workshop on Visual Data Mining of the 5th International Conference on Principles and Knowledge Discovery in Databases (PKDD'01), Freiburg, Germany (2001)