Mining Literal Correlation Rules from Itemsets

Alain Casali
Laboratoire d'Informatique Fondamentale de Marseille (LIF),
CNRS UMR 6166, Aix Marseille Universités, IUT d'Aix en Provence
Aix en Provence, France
alain.casali@lif.univ-mrs.fr

Christian Ernst
Ecole des Mines de St Etienne
CMP - Site Georges Charpak
Gardanne, France
ernst@emse.fr

***Abstract*—Nowadays, data mining tools are becoming more and more popular to extract knowledge from a huge volume of data. In this paper, our aim is to extract Literal Correlation Rules: Correlation Rules admitting literal patterns given a set of items and a binary relation. If a pattern represents a valid Correlation Rule, then any literal belonging to its Canonical Base represents a valid Literal Correlation Rule. Moreover, in order to highlight only relevant Literal Correlation Rules, we add a pruning step based on a support threshold. To extract such rules, we modify the LHS-CHI2 Algorithm and perform some experiments.**

***Keywords*-Data Mining; Chi^2 Correlation Statistic; Literal Pattern.**

I. Introduction and Motivation

An important field in data mining is the discovery of links between values (items) in a binary relation in reasonable response times. Agrawal *et al.* [1] introduce levelwise algorithms in order to compute association rules. Those latter express directional links ($X \rightarrow Y$ for example), based on the support/confidence platform. From this problem, three sub-problems are particularly interesting.

The first one is an adaptation of the supervised classification [2], [3]. Instead of making unsupervised classification, the authors consider the presence of several target attributes. They only consider associations in which the right hand side of the rule contains at least a value of a target attribute. Moreover, they apply rules (specific to the method) allowing to predict to which class belongs an unknown pattern.

The second one is the introduction of literal patterns by Wu *et al.* [4]. The authors compute positive and/or negative association rules, such as $\neg X \rightarrow Y$. To generate the rules, they still use the support/confidence platform by redefining the support of a literal.

In the third one, Brin *et al.* [5] propose the extraction of Correlation Rules, where the platform is no longer based on the support nor the confidence of the rules, but on the Chi-Squared statistical measure, written χ^2. The use of χ^2 is well-suited for several reasons: (i) It is a more significant measure in a statistical way than an association rule; (ii) The measure takes into account not only the presence but also the absence of the items; (iii) The measure is non-directional, and can thus highlight more complex existing links than a "*simple*" implication.

Unlike Association Rules, a Correlation Rule is not represented by an implication but by a set of items for which the value of the χ^2 function is larger or equal than a given threshold, noted $MinCor$.

Since the crucial problem, when computing correlation rules, is the memory usage required by levelwise algorithms, [5] compute only correlations between two values of a binary relation. In [6], we introduce the LHS-CHI2 algorithm. The objective is to compute Decisional Correlation Rules (Correlation Rules which contain, at least, a value of a target attribute). To achieve such an objective, we change the strategy of the browsing search space. We use a lectic order strategy [7] instead of a levelwise one. Since we could not find a function f linking the correlation rate, related to a pattern X, with any of its supersets, we introduce the concept of Contingency Vector, another representation of a Contingency Table based on partitions. Using this concept, we found a function linking the contingency vector of a pattern X with the ones of any of its supersets. Using the LHS-CHI2 algorithm, we have a gain of execution times between 30% and 70% compared to a levelwise algorithm.

Moreover, when applying Advanced Process Control approaches in semiconductor manufacturing, it is important to highlight correlations between parameters related to production, in order to rectify possible drifts of the associated processes. Within this framework, and in collaboration with STMicroelectronics and ATMEL, our previous work [6] focuses on the detection of the main parameters having an impact on the yield. We extract correlations between the values of some columns and those of a target column (a particular column of the file, the yield).

In this paper, we focus on finding out correlations with literal patterns. Such information are important either to point out fault detection, and to detect what parameters do not have an impact on the yield (while they should have). To solve this problem, we introduce Literal Correlation Rule and Literal Decision Correlation Rule concepts. The former is a Correlation Rule admitting literal patterns, and the latter is a restriction of Literal Correlation Rule containing, at least, one value of a target column. In order to compute

ISBN: 978-1-61208-162-5

those rules:

1) We propose a new formula to compute the χ^2 over literal patterns;
2) We show that the χ^2 value for a pattern is equal to the one of "some" literals;
3) We propose a new constraint in order to highlight relevant Literal (Decision) Correlation Rules;
4) We modify the LHS-CHI2 algorithm to take into account these new constraints.

Finally, we carry out experiments on relations provided by the above mentioned manufacturers.

The paper is organized as follows: in Section II, the bases of Literal Patterns and of (Decision) Correlation Rules are recalled. Section III describes the main contribution of the paper. Experiments are detailed in Section IV. As a conclusion, we summarize our contribution and outline some research perspectives.

II. Related Work

In this paper, we use the following notations: let $\mathcal{R}$ be the set of all 1-items and r a binary database relation over $\mathcal{R}$. In our context, $\mathcal{R}$ can be divided into two distinct sets, noted $\mathcal{I}$ and $\mathcal{T}$. $\mathcal{I}$ represents the values of the binary relation used for criteria analysis, and $\mathcal{T}$ is a target attribute. In this section, the concepts of Literal Patterns and Correlation Rules are first recalled.

A. Literal Patterns

Let X, Y be two subsets of $\mathcal{R}$. A literal is a pattern $X\overline{Y}$ in which X is also called the positive part and $\overline{Y}$ the negative part. Literal patterns can be used to extend the well known association rules mining problem: The goal is to obtain new semantics. In a basket market analysis context, the rule $X \rightarrow W$ symbolizes the probability to buy W if one bought X. Using literals, we can extract rules such as $X\overline{Y} \rightarrow W$. This rule materializes the probability to buy W if one bought X but no 1-item (items with cardinality 1) of Y. In [4], to compute rules with literal patterns, Wu *et al.* always use the support-confidence platform by redefining the support of a literal: The number of transactions of the binary relation including X and containing no 1-item of Y.

Example 1: The relation example r given in Table I is used to illustrate the introduced concepts. In this relation, BC and $\overline{B}C$ patterns have a support equal to 4 and 0 respectively. The association rules $B \rightarrow C$ and $\overline{B} \rightarrow C$ have a confidence equal to 1/2 and 0 respectively. This means that half of the transactions including pattern B also contains pattern C and we can not find a transaction which does not contain B and which includes C.

The Canonical Base of a pattern X groups all the possible combinations of literals $Y\overline{Z}$ such that the union between the positive and the negative parts is X, and there is no 1-item in common between those two parts. More precisely, the Canonical Base can be defined as follows:

Table I
Relation example r.

Tid	$\mathcal{I}$	$\mathcal{T}$
1	BCF	G
2	BCF	G
3	DF	G
4	F	G
5	BC	H
6	BC	-
7	BD	-
8	B	-
9	BF	-
10	BF	-

Definition 1 (Canonical Base): Let $X \subseteq \mathcal{R}$ be a pattern, we denote by $\mathbb{P}(X)$ the Canonical Base associated to X. This set is defined as follows: $\mathbb{P}(X) = \{Y\overline{Z}$ such that $X = Y \cup Z$ and $Y \cap Z = \emptyset\} = \{Y\overline{Z}$ such that $Y \subseteq X$ and $Z = X \backslash Y\}$.

By extension, we can define the Canonical Base of a literal $Y\overline{Z}$ as follows: $\mathbb{P}(Y\overline{Z}) = \{Y_1\overline{Z_1}$ such that $Y_1 \subseteq YZ$ and $Z_1 = YZ\backslash Y_1\} = \mathbb{P}(YZ)$.

Example 2: The Canonical Base associated with $X = \{A, B, C\}$ contains the following elements: $\{ABC, AB\overline{C}, AC\overline{B}, BC\overline{A}, A\overline{BC}, B\overline{AC}, C\overline{AB}, \overline{ABC}\}$.

The following property expresses that, if we take two literals belonging to the same Canonical Base, then their associated Canonical Bases are the same.

Property 1: Let X be a pattern and $Y_1\overline{Z_1}, Y_2\overline{Z_2}$ two literal patterns belonging to its Canonical Base. We have: $\mathbb{P}(X) = \mathbb{P}(Y_1\overline{Z_1}) = \mathbb{P}(Y_2\overline{Z_2})$.

B. Correlation Rules and Decision Correlation Rules

In [5], Brin *et al.* propose the extraction of correlation rules. The platform is no longer based on the support nor the confidence of the rules, but on the χ^2 statistical measure. The formula to compute the χ^2 for a pattern X is:

$$\chi^2(X) = \sum_{Y\overline{Z} \in \mathbb{P}(X)} \frac{(Supp(Y\overline{Z}) - E(Y\overline{Z}))^2}{E(Y\overline{Z})} \qquad (1)$$

Such a computation requires (i) the support, and (ii) the expectation value (or average) of all literals belonging to $\mathbb{P}(X)$. The expectation value of a literal $Y\overline{Z}$ measures the theoretical frequency in case of independence of all 1-items included in $Y\overline{Z}$, see Formula (2).

$$E(Y\overline{Z}) = |r| * \prod_{y \in Y} \frac{Supp(y)}{|r|} * \prod_{z \in Z} \frac{Supp(\overline{z})}{|r|} \qquad (2)$$

Each support of each literal belonging to the Canonical Base associated to X is stored in a table called Contingency Table. Thus, for a given pattern X, its contingency table, noted $CT(X)$, contains exactly $2^{|X|}$ cells.

In our context, there is a single degree of freedom between the items. A table giving the centile values with regard to the

χ^2 value for X can be used in order to obtain the correlation rate for X [8].

Example 3: With the relation Example r given in Table I, Table II shows the contingency table of pattern BC.

Table II
CONTINGENCY TABLE OF PATTERN BC.

	B	$\overline{B}$	$\sum_{row}$
C	4	0	4
$\overline{C}$	4	2	6
$\sum_{column}$	8	2	10

Thus, $\chi^2(BC) \simeq 0.28$, which corresponds to a correlation rate of about 45%.

Unlike association rules, a correlation rule is not represented by an implication but by the patterns for which the value of the χ^2 function is larger than or equal to a given threshold.

Definition 2 (Correlation Rule): Let $MinCor$ be a threshold (≥ 0), and $X \subseteq \mathcal{R}$ a pattern. If the value for the χ^2 function for X is larger than or equal to $MinCor$, then this pattern represents a valid Correlation Rule.

In addition to the previous constraint, many authors have proposed some criteria to evaluate whether a Correlation Rule is semantically valid [9]:

1) As the χ^2 computation has no significance for a 1-item, we only examine patterns of cardinality larger than or equal to two;
2) Since the χ^2 function is an increasing function, we impose a maximum cardinality, noted *MaxCard*, on the number of patterns to examine;
3) The Cochran criterion: All literal patterns of a contingency table must have an expectation value different from zero and 80% of them must have a support larger than 5% of the whole population. This criterion has been generalized by Brin et al. [5] as follows: $MinPerc$ of the literal patterns of a contingency table must have a support larger than $MinSupCT$, where $MinPerc$ and $MinSupCT$ are thresholds specified by the end-user.

Example 4: Let $MinCor = 0.25$, then the correlation rule materialized by the BC pattern is valid ($\chi^2(BC) \simeq 0.28$). However, the correlation rule represented by the BH pattern is not valid ($\chi^2(BH) \simeq 0.1$).

The crucial problem, when computing correlation rules, is the memory requirement by levelwise algorithms. For a pattern X, the computation of the χ^2 function is based on a contingency table including $2^{|X|}$ cells. Thus, at level i, C_n^i candidates (where n is the number of values of r) have to be generated and stored, in the worst case scenario, as well as the associated contingency tables. With cells encoded over 2 bytes, corresponding storage space requires 2.5 GB of memory at the 3rd level, and 1.3 TB at the 4th level.

This is why we have changed the browsing search space strategy in [6]. Instead of using a lewelwise algorithm, our algorithm, called LHS-CHI2, browse the search space according to the lectic order [7]. It is based on:

1) The LS algorithm [10]. This algorithm allows the browsing of the powerset lattice using a balanced tree;
2) Contingency vectors, another representation of the contingency tables based on bit vectors;
3) A proposition which links the contingency vector of a pattern X with the ones of its immediate successors, "*i.e.*" contingency vectors of patterns $X \cup y, \forall y \in \mathcal{R} \backslash X$;
4) A pruning step based on the positive border [11], noted BD^+.

Still in order to limit the browsing search space, whatever the browsing strategy used, we only consider Correlation Rules which have a value belonging to the set $\mathcal{T}$.

Definition 3 (Decision Correlation Rule): A Decision Correlation Rule is a Correlation Rule which contains at least one value of the target attribute $\mathcal{T}$.

Using all the constraints mentioned above, it results a gain of time between 30% and 80% using our strategy than using a levewise one. The pseudo-code of the LHS-CHI2 Algorithm is given below. The pseudo-code of the procedure CREATE_CV can be found in [6]. The predicate *CtPerc* expresses the satisfiability of the Cochran criterion. The first call to LHS-CHI2 is made with $X = \mathcal{I}$ and $Y = \emptyset$.

If we want to extract all the Correlation Rules, and not only the Decision Correlation Rules, we have to prune the test "$\exists t \in \mathcal{T} : t \in X$" from line 1 of the LHS-CHI2 Algorithm.

Algorithm 1: LHS-CHI2 Algorithm.

```
input : $X$ and $Y$ two patterns
output: $\{Z \subseteq X$ such that $\chi^2(Z) \geq MinCor\}$
1  if $Y = \emptyset$ and $\exists t \in \mathcal{T} : t \in X$ and $|X| \geq 2$ and
   $\chi^2(X) \geq MinCor$ then
2      Output X, $\chi^2(X)$
3  end
4  $A := max(Y)$ ;
5  $Y := Y \backslash \{A\}$ ;
6  LHS-CHI2(X,Y) ;
7  $Z := X \cup \{A\}$ ;
8  if $\forall z \in Z, \exists W \in BD^+ : \{Z \backslash z\} \subseteq W$ then
9      CV(Z) := CREATE_CV(VC(X),Tid(A)) ;
10     if $|Z| \leq MaxCard$ and
       $CtPerc(CV(Z), MinPerc, MinSupCT)$ then
11         $BD^+ := max_{\subseteq}(BD^+ \cup Z)$ ;
12         LHS-CHI2(Z,Y) ;
13     end
14 end
```

Example 5: The results of the LHS-Chi2 algorithm with the relation example r (cf. Table I) using thresholds $MinSupCT = 0.2$, $MinPrec = 0.3$ and $MinCor = 1.8$ are given in Table III.

Table III
Result of LHS-Chi2 Algorithm.

Correlation Rule	χ^2 value
BG	3.75
FG	4.44
BCF	4.24
BCG	9.10
BDF	10.14
CFG	5.74
DFG	4.93
BCFG	20.09

III. Literal Correlation Rules

In this section, we present our contribution. The aim is to build Literal Correlation Rules (Correlation Rules over Literal Patterns) only from (i) the relation r and (ii) the set of items $\mathcal{R}$. Mining such rules with the help of the relation $\overline{r}$ and of the set of items $\overline{\mathcal{R}}$ is not suitable because:

- Since the relation r is often sparse, the relation $\overline{r}$ is dense. As a result, the relational operators (union, intersection, ...) have slow performances over $r \cup \overline{r}$.
- It is possible to find an item $A \in \mathcal{R}$ such that the pattern $A\overline{A}$ satisfies all the constraints over a Correlation Rule. It is the case of the pattern B in our example relation. As a consequence, the set of solutions is polluted by inconsistent patterns.

We first define the concept of Literal Correlation Rule: an extension of Correlation Rules. We show, in a second step, that two literal patterns, which belong to the same Canonical Base, have the same χ^2 value and satisfy the same set of constraints (see Section II-B). As a consequence, the set of Correlation Rules can be used as a base for the Literal Correlation Rules. Then we show that the number of Literal Correlation Rules is exponential with regard to the number of Correlation Rules. We modify the LHS-Chi2 Algorithm in order to compute Literal Correlation Rules, and we add another pruning step in order to limit the number of results. We finally define the χ^2 function for a literal pattern $X\overline{W}$ as follows:

$$\chi^2(X\overline{W}) = \sum_{Y\overline{Z} \in \mathbb{P}(X\overline{W})} \frac{(Supp(Y\overline{Z}) - E(Y\overline{Z}))^2}{E(Y\overline{Z})} \qquad (3)$$

Definition 4 (Literal (Decision) Correlation Rule): Let $X\overline{W}$ be a pattern, $MinCor$, $MinPerc$, $MinSupCT$ and $MaxCard$ thresholds specified by the end-user. According to the criteria introduced in Section II-B, the literal $X\overline{W}$ is a valid Literal Correlation Rule if and only if:

1) $\chi^2(X\overline{W}) \geq MinCor$;
2) $2 \leq |X\overline{W}| \leq MaxCard$;
3) $MinPerc$ cells of its contingency table have a support greater or equal than $MinSupCT$.

Moreover, if a value of the target attribute is present either in the positive part of the literal either in its negative part, the rule is called a Literal Decision Correlation Rule.

Example 6: Let us consider the following thresholds $MinCor = 1.8$, $MinSupCT = 0.2$ and $MinPrec = 0.3$, and the literal $B\overline{G}$. The contingency table associated to this literal pattern is:

	B	$\overline{B}$
G	2	2
$\overline{G}$	2	4

Since the four cells of this contingency table are greater than 2, we satisfy the third condition. We have $\chi^2(B\overline{G}) \simeq 3.75 \geq MinCor$ and the first condition is valid. Literal $B\overline{G}$ has a cardinality equal to 2, thus the second condition is checked. Moreover, Literal $B\overline{G}$ contains a value of the target attribute. As a consequence, the Literal Decision Correlation Rule materialized by $B\overline{G}$ is valid.

Let $X \cup \overline{A}$ and $X \cup A$ be two literal patterns, where X does not contain a negative part. The following lemma shows that the χ^2 values for both literal patterns are equals.

Lemma 1: Let X be a pattern and A a 1-item. We have: $\chi^2(X \cup \overline{A}) = \chi^2(X \cup A)$

The following proposition shows that any literal pattern belonging to the same Canonical Base has the same χ^2 value.

Proposition 1: Let $X \subseteq \mathcal{R}$ be a pattern, then we have:

$$\forall Y\overline{Z} \in \mathbb{P}(X), \chi^2(X) = \chi^2(Y\overline{Z}) \qquad (4)$$

The following lemma indicates how we can build valid Literal Correlation Rules given only valid Correlation Rules.

Lemma 2: If a pattern X is a valid Correlation Rule (its χ^2 value is greater or equal than the threshold $MinCor$ and X satisfies all the constraints given in Section II-B), then any literal pattern belonging to its Canonical Base represents a valid Literal Correlation Rule.

Consequences of Proposition 1 and of Lemma 2 are very attractive. When mining Literal Correlation Rules, we do not need, as input of our algorithm, the set $\mathcal{R} \cup \overline{\mathcal{R}}$ but only the set $\mathcal{R}$. We just have to modify the processing done on the leaves of the LHS-Chi2 execution tree, in order to explore the Canonical Base associated with the current pattern. Moreover, as expected in introduction, we do not need the relation $\overline{r}$. Finally, the following results holds:

Corollary 1: Correlation Rules are a lossless representation for Literal Correlation Rules.

The concept of lossless representation applied to association rules [12] or to literal association rules mining [13], are very helpful to reduce the number of rules. However, we cannot predict an exact value for the expected gain. With the following lemma, we show that the number of Literal

Correlation Rules depends on the number of Correlation Rules having cardinality i ($i \in [2, MaxCard]$).

Lemma 3: Let us denote by *Sol* the set of solutions related to the problem of finding all the Correlation Rules satisfying all the constraints. Let Sol_i be the subset of *Sol* which contains only rules of cardinality i. Let *Sol'* be the set of solutions related to the problem of finding all the Literal Correlation Rules satisfying the same set of constraints than *Sol*. Then we have: $|Sol'| = \sum_{i=2}^{i=MaxCard} |Sol_i| * 2^i$.

A drawback highlighted by decision makers using the *MineCor* software (the software which implements the LHS-Chi2 Algorithm) is that the extracted rules which have a large χ^2 value could appear seldom in the relation. As a consequence, they consider that the obtained information is not of great quality. To answer their expectations, we modify the LHS-Chi2 Algorithm by adding a pruning step based on the support (using a threshold $MinSup$) and by extracting Literal Correlation Rules. The changes only affect the first three lines of the LHS-Chi2 Algorithm. The new algorithm is called LHS-LChi2. Like in the LHS-Chi2 Algorithm, if we want to extract the Literal Correlation Rules and not only the Literal Decision Correlation Rules, we have to prune the test "$\exists t \in \mathcal{T} : t \in X$" for line 1 of the LHS-LChi2 Algorithm.

Algorithm 2: LHS-LChi2 Algorithm.

```
1 if $Y = \emptyset$ and $\exists t \in \mathcal{T} : t \in X$ and $|X| \geq 2$ and $\chi^2(X) \geq MinCor$ then
2     foreach $Y\overline{Z} \in \mathbb{P}(X)$ do
3         if $Supp(Y\overline{Z}) \geq MinSup$ then
4             Output $Y\overline{Z}$, $\chi^2(X)$
5         end
6     end
7 end
8 ...
```

Let us emphasize that the addition of the constraint "$Supp(Y\overline{Z}) \geq MinSup$" has the negative effect of making false Corollary 1 and Lemma 3 unless $MinSup$ equals 0.

Example 7: Continuing our example with parameters $MinSupCT = 0.2$, $MinPrec = 0.3$, $MinCor = 1.8$ and $MinSup = 0.4$, the results of the LHS-LChi2 Algorithm are given in Table IV.

Table IV
Result of LHS-Chi2 Algorithm.

Correlation Rule	χ^2 value	Support
$B\overline{G}$	3.75	6
FG	4.44	4
$B\overline{C}\overline{G}$	9.10	4
$BF\overline{D}$	10.14	4

IV. Experimental Evaluations

Some representative results of the LHS-LChi2 Algorithm are presented below. As emphasized in Section I, the experiments were done on different CSV files of real value measures supplied by STMicroelectronics (STM) and ATMEL (ATM). These files have one or more target columns, resulting from the concatenation of several measurement files. The characteristics of the relations used can be found in Table IV. All experiments were conducted on an HP Workstation (1.8 GHz processor with a 4 Gb RAM). To carry out pre-processing and transformation of these files into a binary relation, we implemented methods described in [14].

Table V
Dataset Examples

Name	Number of Columns	Number of Rows
STM File	1 281	297
ATM File	749	213

Figure 1. Number of Literal Decision Correlation Rules. Results with 4 intervals, $CtPerc = 0.34$, $MinCorr = 1.6$, $MinSupCT = 0.24$ (STM File - target1)

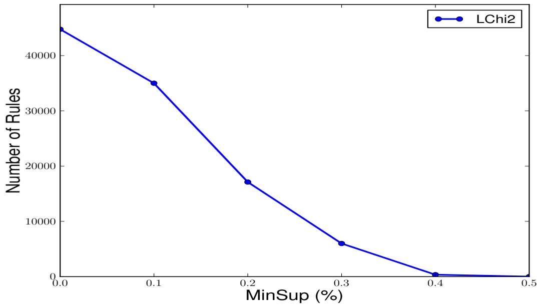
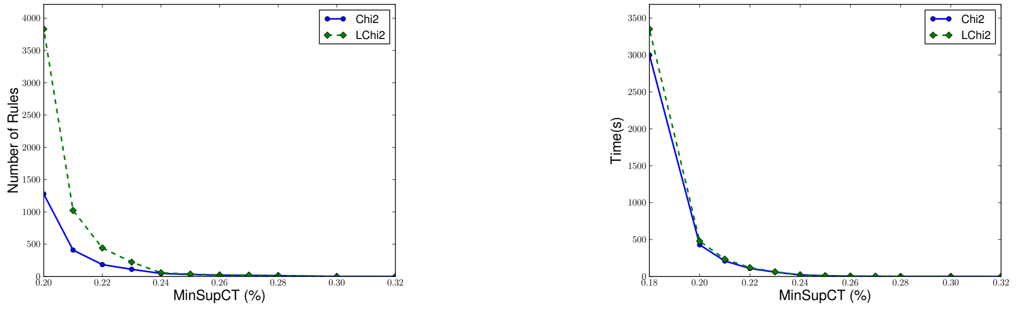


Figure 1 shows the impact of the $MinSup$ threshold over Literal Decision Correlation Rules. In the same way, when extracting frequent pattern, if the threshold $MinSup$ is large, no rule is produced. The lower $MinSup$ is, the more we can approach the bound given in lemma 3.

The goal of Figure 2(a) is to compare the number of rules produced by LHS-Chi2 and LHS-LChi2 Algorithms. Since LHS-Chi2 Algorithm does not have a pruning step using the $MinSup$ threshold, we decided to fix it to the value of $MinSupCT$. The number of rules produced by the LHS-LChi2 Algorithm is greater with a factor between 1 and 2.5. In Figure 2(b), we compare the two algorithms over the same hypothesis. As we can see, execution times are very close (less than 12% in the worst case). This can be explained by our specific implementation of the LHS-LChi2 Algorithm: the computation of the χ^2 function and the pruning step using $MinSup$ both require the browsing of a contingency table. During the χ^2 computation, we

Figure 2. Results with 6 intervals, $CtPerc = 0.3$, $MinCorr = 2.8$, $MinSup = MinSupCT$ (ATM file - target3).



(a) Number of Decision Correlation Rules *vs.* Number of Literal Decision Correlation Rules

(b) Execution Time

put into a vector all the literal patterns having a support greater than $MinSup$. As a consequence, line 3 can be resumed as a browsing vector (which contains, in the worst case, $2^{MaxCard}$ elements). Thus, the difference between the execution times can be explained by the number of input/output operations which are more important in the LHS-LCHI2 Algorithm since we extract more rules.

V. Conclusion and Future Work

When mining Correlation Rules, one drawback is that the extracted rules which have a large χ^2 value appear seldom in the relation. As a consequence, we do not know which literal patterns have an important impact on the rules. To solve this problem, we have introduced the concept of Literal Correlation Rules: Correlation Rules admitting literal patterns. We show that the set of Correlation Rules satisfying a set of constraints is a base for the Literal Correlation Rules satisfying the same set of constraints. Thus we provide an upper border for the number of Literal Correlation Rules. In order to highlight only relevant Literal Correlation Rules, we add a pruning step based on the support of a literal, and therefore modified the related algorithm.

To continue our work, we intend to use multi-core strategies. In a first step, one thread could process the leaves of the execution tree while another could explore the branches of the tree. In a second step, our aim is to to parallelize each branch of the LHS-LCHI2 algorithm.

References

[1] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo, "Fast discovery of association rules," in *Advances in Knowledge Discovery and Data Mining*. AAAI/MIT Press, 1996, pp. 307–328.

[2] G. Chen, H. Liu, L. Yu, Q. Wei, and X. Zhang, "A new approach to classification based on association rule mining," *Decision Support Systems*, vol. 42, no. 2, pp. 674–689, 2006.

[3] W. Li, J. Han, and J. Pei, "Cmar: Accurate and efficient classification based on multiple class-association rules," in *ICDM*, IEEE Computer Society, 2001, pp. 369–376.

[4] X. Wu, C. Zhang, and S. Zhang, "Efficient mining of both positive and negative association rules," *ACM Trans. Inf. Syst.*, vol. 22, no. 3, pp. 381–405, 2004.

[5] S. Brin, R. Motwani, and C. Silverstein, "Beyond market baskets: Generalizing association rules to correlations," in *SIGMOD Conference*, 1997, pp. 265–276.

[6] A. Casali and C. Ernst, "Extracting decision correlation rules," in *DEXA*, ser. Lecture Notes in Computer Science, vol. 5690. Springer, 2009, pp. 689–703.

[7] B. Ganter and R. Wille, *Formal Concept Analysis: Mathematical Foundations*. Springer, 1999.

[8] M. Spiegel and L. Stephens, *Outline of Statistics*. McGraw-Hill, 1998.

[9] D. Moore, "Measures of lack of fit from tests of chi-squared type," in *Journal of statistical planning and inference*, vol. 10, no. 2, 1984, pp. 151–166.

[10] M. Laporte, N. Novelli, R. Cicchetti, and L. Lakhal, "Computing full and iceberg datacubes using partitions," in *ISMIS*, 2002, pp. 244–254.

[11] H. Mannila and H. Toivonen, "Levelwise search and borders of theories in knowledge discovery," *Data Min. Knowl. Discov.*, vol. 1, no. 3, pp. 241–258, 1997.

[12] N. Pasquier, R. Taouil, Y. Bastide, G. Stumme, and L. Lakhal, "Generating a condensed representation for association rules," *J. Intell. Inf. Syst.*, vol. 24, no. 1, pp. 29–60, 2005.

[13] G. Gasmi, S. B. Yahia, E. M. Nguifo, and S. Bouker, "Extraction of association rules based on literalsets," in *DaWaK*, ser. Lecture Notes in Computer Science, vol. 4654. Springer, 2007, pp. 293–302.

[14] D. Pyle, *Data Preparation for Data Mining*. Morgan Kaufmann, 1999.