

PCAR : Nouvelle Approche de Génération de Règles d'Association Cycliques

Eya BEN AHMED*, Mohamed Salah GOUIDER**

Laboratoire BESTMOD, Institut Supérieur de Gestion, Tunis

* eya.benahmed@gmail.com,

** ms.gouider@isg.rnu.tn

Résumé. *Les règles d'association cycliques* vise la découverte de nouvelles relations entre des produits qui varient d'une façon régulièrement cyclique dans le temps. Dans ce cadre, nous introduisons, un nouvel algorithme nommé PCAR caractérisé par sa performance et son aspect incrémental. L'étude empirique que nous avons menée montre la robustesse et l'efficacité de notre algorithme proposé vs. ceux de la littérature.

1 Introduction

Dans le cadre de ce papier, nous étudions le problème de la découverte des règles d'association caractérisées par une variation régulière cyclique dans le temps. En effet, si nous générons les règles d'association à partir des données de ventes mensuelles, nous constatons la variation saisonnière telle que certaines règles sont vraies approximativement le même mois chaque année.

2 Nouvelle approche PCAR

Le nouvel modèle de génération de règles d'association, que nous introduisons, nommé PCAR pour désigner PARTITION CYCLIC ASSOCIATION RULES, se base principalement sur le mécanisme de segmentation de la base de données en un nombre de partitions. Dans une optique incrémentale, le parcours de la base se fera ainsi pas à pas. Ce qui en découle, une prise en compte de la caractéristique capitale des bases de données à savoir leurs dynamisations.

Le processus de notre algorithme peut se décortiquer de la manière suivante : Ayant introduit la base de données, le seuil minimum de support, le seuil minimum de la confiance, le nombre de partitions et la longueur du cycle, l'algorithme PCAR se charge de fragmenter la base selon le nombre de partitions passé en entrée. Ensuite, il parcourt l'ensemble des partitions une à une. Pour chaque partition, il extrait l'ensemble des itemsets temporels fréquents cycliques relatifs à la partition en cours se comparant au nouveau seuil de support relatif à la cardinalité de la partition. Aboutissant à la dernière partition, seuls les itemsets cycliques fréquents seront pris en considération lors de la génération des règles d'association cycliques qui représente la dernière étape de notre algorithme.

3 Évaluation expérimentale

Dans le but d'évaluer notre approche, nous avons mené une série d'expérimentations sur diverses bases benchmark. De ce fait, à travers la variation des paramètres clés à savoir le seuil minimal de support, le nombre de partitions et la longueur du cycle, il en découle, en premier lieu que plus on augmente le seuil minimum de support ou le nombre de partitions, plus la performance de PCAR s'amplifie. En second lieu, nos cas de test mettent en exergue que l'augmentation de la longueur du cycle améliore la performance de PCAR vu que plus la longueur du cycle est faible, plus on sera emmené à balayer les partitions une à une, ce qui engendrera la dégradation de la performance de PCAR. Dans la vue d'évaluer radicalement

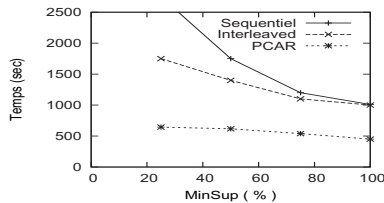


FIG. 1 – Comparatif des temps d'exécution des algorithmes PCAR, INTERLEAVED et SÉQUENTIEL.

notre apport, on se compare avec les algorithmes SÉQUENTIEL et INTERLEAVED étant les approches innovatrices consacrées à la génération des règles d'association cycliques (figure 1).

Par conséquent, PCAR s'avère deux fois plus performant que l'INTERLEAVED et 2,22 fois plus rapide que le SÉQUENTIEL.

4 Conclusion

Évidemment, en termes d'efficacité, PCAR dépasse amplement les deux autres algorithmes et se montre meilleur pour les divers cas de tests.

Ces résultats encourageants obtenus dans ce travail rendent notre approche innovatrice pouvant être étendue dans plusieurs autres travaux ultérieurs. En conséquence, nous envisageons comme perspectives à creuser : (i) Utiliser des cubes de données pour générer des règles cycliques multidimensionnelles, (ii) Aborder la problématique de la maintenance incrémentale de règles d'association cycliques.

Summary

The *cyclic association rules* aims to discover new relationships between items that display regular cyclic variation over time. In this paper, we propose to optimize their extraction efficiently by introducing the PCAR ALGORITHM. The carried out experimental results shed the light on the outperformance of our proposal vs. the pioneering approaches in this trend.