

Frédéric Flouvat

LIMOS
Clermont Ferrand

Résumé :

Certains travaux ont mis en évidence que pour le problème de la découverte des motifs fréquents maximaux, l'algorithme Apriori restait performant pour des jeux de données dont la taille des motifs fréquents maximaux était relativement petite. Le but de notre approche est d'utiliser efficacement Apriori en changeant de stratégie lorsque celui-ci est mis en difficulté. Un algorithme adaptatif a donc été mis en place, son nom est ABS pour Adaptive Border Search. Cet algorithme s'appuie sur le concept de dualisation pour explorer l'espace de recherche à partir des connaissances déjà découvertes. L'un des intérêts de cette approche, en plus d'être adaptative, est de fournir une caractérisation exacte des différentes bordures permettant ainsi d'effectuer des "sauts" précis dans l'espace de recherche.

Les tests réalisés montrent que cette approche permet selon les cas d'améliorer Apriori par un facteur 10. Même si cette approche n'est pas la plus performante parmi toutes les implémentations disponibles, elle est plus performante que IBE (Irredundant Border Enumerator), la seule s'appuyant sur le concept de dualisation.