

# Techniques de fouille de données pour la réécriture en présence de contraintes de valeurs

Hélène Jaudoin

LIMOS - UMR 6158 CNRS - Université Clermont-Ferrand II

Dans le cadre de mes travaux de thèse, le problème de la réécriture de requêtes en termes de vues en présence de contraintes de valeurs a été étudié. Les contraintes de valeurs correspondent à la notion de type énuméré dans les bases de données et permettent de spécifier les valeurs autorisées pour un attribut d'un concept donné. Ce travail a pour origine une application dans le domaine du développement durable. Il s'agit plus précisément d'un système d'intégration de sources de données agricoles visant à garantir la traçabilité des pratiques agricoles.

L'étude du problème de la réécriture en présence des contraintes de valeurs a mis en évidence de nouveaux cas de réécritures. Le problème du calcul de ces nouveaux cas de réécritures peut être rattaché à un cadre de découverte de connaissances dans les bases de données. L'exploitation de ce cadre a un double intérêt. Premièrement le cadre de découverte des connaissances nous permet de bénéficier de solutions algorithmiques existantes pour résoudre effectivement le problème de réécriture. Deuxièmement il nous permet de calculer efficacement les réécritures et d'envisager le passage à l'échelle en adaptant des algorithmes efficaces de fouille de données.

Nous montrons ici comment le problème de la réécriture de requêtes en présence de contraintes de valeurs se rattache au cadre de découverte de connaissances introduit par Mannila et Toivonen. Puis nous décrivons une implémentation de cette approche et donnons quelques résultats expérimentaux.