

Découverte interactive de règles d'association via une interface visuelle

Pascale Kuntz, Rémi Lehn
Fabrice Guillet, Bruno Pinaud

Laboratoire d'informatique de Nantes Atlantique (LINA)
Site Ecole Polytechnique
La Chantrerie - rue Christian Pauc
BP 50609, 44306 Nantes Cedex 3
{pascale.kuntz, remi.lehn, fabrice.guillet, bruno.pinaud}@univ-nantes.fr,

Résumé. En nous appuyant sur des hypothèses majoritairement empruntées à des travaux sur les systèmes anthropocentrés d'aide à la décision, nous décrivons dans cet article un environnement interactif de fouille de règles d'association dans lequel l'utilisateur pilote le processus, en jouant le rôle d'une heuristique dans un environnement de recherche complexe. Afin de permettre à la fois une représentation visuelle accessible et une instanciation aisée des outils d'interactivité le modèle choisi est ici un graphe en niveaux - les niveaux étant associés aux cardinaux des sous-ensembles d'attributs des prémisses. Le processus a été déployé dans un logiciel prototype dont l'analyse des résultats ouvre de nouvelles perspectives sur l'analyse comportementale d'un utilisateur en situation de fouille.

1 Introduction

Si les efforts les plus remarquables de la première décennie de l'Extraction de Connaissances dans les Données (ECD) ont principalement porté sur le développement d'algorithmes automatiques performants, le rôle de l'utilisateur est peu à peu devenu un sujet de préoccupation majeur. Ce besoin d'intégration a conduit à l'émergence de nombreux outils de visualisation et d'interaction [Fayyad *et al.*, 2002], [Grinstein, 1996]. Il s'agit d'apporter à l'utilisateur un substrat artificiel qui transcrive un grand nombre d'informations et qui soit un support à ses connaissances et à son intuition pour lui permettre de découvrir de nouvelles relations et d'imaginer de nouvelles questions.

1.1 Visualisation et interaction en ECD

Traditionnellement, le recours à des techniques de visualisation précède et suit les étapes de traitement automatiques. En amont du processus d'ECD, elles assistent les tâches de sélection et de pré-traitement des données en renforçant la convivialité des interfaces. En aval, à l'issue de l'application des algorithmes de découverte de structures, elles visent à présenter des résultats sous des formes intelligibles facilitant leur interprétation. Dans ce contexte, la plupart des outils ont été définis pour des objectifs spécifiques (recherche de règles d'association, classification de données spatiales, etc.).