

Treillis des concepts SKYLINES : Analyse multidimensionnelle des SKYLINES fondée sur les ensembles en accord

Sébastien Nedjar, Fabien Pesci¹, Lotfi Lakhal & Rosine Cicchetti

Laboratoire d'Informatique Fondamentale de Marseille (LIF), CNRS UMR 6166
Aix-Marseille Université,
IUT d'Aix-en-Provence, Avenue Gaston Berger, 13625 Aix-en-Provence Cedex

Résumé. Le concept de SKYLINE a été introduit pour mettre en évidence les objets « les meilleurs » selon différents critères. Une généralisation multidimensionnelle du SKYLINE a été proposée à travers le SKYCUBE qui réunit tous les SKYLINES possibles selon toutes les combinaisons de critères et permet d'analyser les liens entre objets SKYLINES. Comme le data cube, le SKYCUBE s'avère extrêmement volumineux si bien que des approches de réduction sont incontournables. Dans cet article, nous définissons une approche de matérialisation partielle du SKYCUBE. L'idée sous-jacente est d'éliminer de la représentation les Skycuboïdes facilement re-calculables. Pour atteindre cet objectif de réduction, nous caractérisons un cadre formel : le treillis des concepts ACCORDS. Cette structure combine la notion d'ensemble en accord et le treillis des concepts. À partir de cette structure, nous dérivons le treillis des concepts SKYLINES qui en est une instance contrainte. Le point fort de notre approche est d'être orientée attribut ce qui permet de borner le nombre de nœuds du treillis et d'obtenir une navigation efficace à travers les Skycuboïdes.

1 Introduction

Dans un contexte décisionnel, certaines requêtes ne renvoient aucun résultat. Dans ces requêtes, l'utilisateur recherche les tuples pour lesquels les valeurs de certains critères sont optimales. C'est le caractère « multicritère » de ces interrogations qui les rend généralement infructueuses. En effet, tel tuple peut être optimal pour un critère mais pas pour un autre, il est alors éliminé du résultat alors qu'il aurait pu être pertinent pour l'utilisateur. Par exemple, si l'on considère une base de données immobilières, la recherche du logement « idéal » peut combiner des conditions sur le prix, le plus bas possible, la surface, la plus grande possible, et l'éloignement du lieu de travail, le plus réduit possible. Évidemment il est vraisemblable que ce logement idéal n'existe pas, d'où l'absence de réponse à ce type de requête. Pourtant certains logements pourraient s'avérer pertinents pour l'utilisateur parce que, situés dans une zone proche, mais non voisine, ils réunissent les critères de surface maximale et de prix minimal.

1. Fabien Pesci bénéficie d'une bourse doctorale co-financée par le conseil régional PACA et l'entreprise CA2I.