

Sélection Incrémentale d'un Schéma de Fragmentation Horizontale d'un Entrepôt de Données Relationnel

Rima Bouchakri^{*,**}, Ladjel Bellatreche^{*}

^{*}LIAS/ISAE-ENSMA – Université de Poitiers
Futuroscope, 86960 - France
(rima.bouchakri, bellatreche)@ensma.fr

^{**}Ecole nationale Supérieure d'Informatique (ESI)
Alger, Algérie
r_bouchakri@esi.dz

Résumé. La fragmentation horizontale permet de réduire la complexité des requêtes décisionnelles exécutées sur un entrepôt de données relationnel. Elle se base sur le principe de réorganisation des données qui ne nécessite pas un espace de stockage supplémentaire. Cependant, sélectionner un schéma de fragmentation horizontale d'un entrepôt n'est guère une tâche facile, vu l'espace de recherche très complexe à exploiter. Les algorithmes existants sélectionnent un schéma de fragmentation lors de la phase de conception physique d'un entrepôt, afin d'optimiser une charge de requêtes préalablement connue. Ces algorithmes ne prennent pas en considération les changements au niveau des requêtes. Dans cet article, nous proposons d'effectuer *une sélection d'un schéma de fragmentation dite incrémentale* basée sur les algorithmes génétiques. Notre approche permet l'optimisation de l'exécution de la charge de requêtes décisionnelles et l'adaptation du schéma de fragmentation aux changements de la charge. Nous réalisons une étude expérimentale qui montre l'intérêt de la sélection incrémentale d'un schéma de fragmentation.

1 Introduction

Les entrepôts de données (\mathcal{ED}) permettent le stockage et la consolidation, en une seule localité, d'une quantité gigantesque d'information issues de plusieurs sources hétérogènes. Les entrepôts de données sont souvent modélisés par des schémas relationnels (ex. schéma en étoile). Un entrepôt de données est souvent interrogé par l'intermédiaire de requêtes décisionnelles complexes. Elles renferment des opérations de sélections, d'agrégations et des opérations de jointures dites en étoiles entre la table de faits et une ou plusieurs tables de dimensions. Afin d'optimiser ses opérations, il devient primordial d'employer des techniques d'optimisations, comme la fragmentation horizontale (\mathcal{FH}) qui a déjà fait ses preuves dans le contexte des bases de données traditionnelles et d'entrepôt de données. Brièvement, la fragmentation horizontale permet de répartir les instances d'une table, d'un index ou d'une vue matérialisée en un ensemble de partitions disjointes appelés *fragments horizontaux* [10].