

# Schéma relationnel distribué pour les arbres

Cyril Gavoille\* et Arnaud Labourel\*

---

## Représentation distribuée

Représenter un graphe en mémoire est un problème basique mais néanmoins fondamental dans le domaine des structures de données. Il existe deux manières élémentaires d'accomplir cette tâche, utiliser une matrice d'adjacence ou bien des listes d'adjacence. Cette deuxième représentation est quasi-optimale en terme d'espace mémoire utilisé mais les requêtes d'adjacence requièrent une recherche dans une liste. En revanche, dans la première représentation, les requêtes d'adjacence se font en temps constant mais par contre l'espace de stockage est sur-linéaire en la taille du graphe.

Une autre technique, connue sous le nom de *représentation distribuée*, consiste à distribuer localement la représentation globale du graphe parmi ses sommets. Le but est de stocker une information utile sur chaque sommet afin de permettre l'implantation de requêtes à partir de données locales vues sous forme d'étiquettes binaires. Concrètement, pour le cas de la requête d'adjacence, il doit être possible de tester la présence d'une arête entre deux sommets en examinant seulement leurs informations locales. Le problème est de minimiser la longueur des étiquettes tout en gardant un calcul des requêtes rapide.

## Notre contribution

Dans un arbre enraciné, deux sommets  $u$  et  $v$  ayant comme plus petit ancêtre commun  $w$  sont dit  $(k_1, k_2)$ -reliés si la distance de  $u$  à  $w$  est  $k_1$  et la distance de  $v$  à  $w$  est  $k_2$ . Pour un entier  $k$ , on définit un schéma  $k$ -relationnel comme étant une représentation distribuée d'arbres enracinés supportant des requêtes afin de déterminer si  $u$  et  $v$  sont  $(k_1, k_2)$ -reliés pour  $k_1, k_2 \leq k$ .

Par exemple, un schéma 1-relationnel permet de tester si deux sommets sont  $(0, 0)$ ,  $(1, 0)$ ,  $(0, 1)$  ou  $(1, 1)$ -reliés, et donc supporte les requêtes d'adjacence, de parenté et de fraternité. Plus généralement, un schéma  $k$ -relationnel permet le calcul des distances entre deux sommets à distance au plus  $k$ .

On améliore dans ce papier le meilleur schéma  $k$ -relationnel qui utilise  $\log n + O(k^2 \log(k \log n))$  bits par sommet [ABR05] en un schéma utilisant  $\log n + O(k \log(k \log(n/k)))$  bits par sommets.

**Théorème 1** *La famille des arbres enracinés de  $n$  sommets admet un schéma  $k$ -relationnel avec des étiquettes de  $\log n + O(k \log(k \log(n/k)))$  bits. De plus, les requêtes s'effectuent en temps constant et l'étiquetage prends un temps  $O(n \log n)$  pour  $k$  fixé.*

## RÉFÉRENCES

- [ABR05] Stephen Alstrup, Philip Bille, and Theis Rauhe. Labeling schemes for small distances in trees. *SIAM Journal on Discrete Mathematics*, 19(2) :448–462, 2005.

---

\*Laboratoire Bordelais de Recherche en Informatique, Université de Bordeaux, Bordeaux, France. {gavoille,labourel}@labri.fr. Les trois auteurs sont financés par le projet "GeoComp" de l'ACI Masse de Données.