

# Représentation de graphes par ACP granulaire

Bruno Gaume\*, Louis Ferré\*\*

\*IRIT-UPS, 118 route de Narbonne, F-31062 Toulouse Cedex 4, France  
gaume@irit.fr,  
<http://dilan.irit.fr>

\*\*GRIMM-EA2254, Université Toulouse II, 5 Allées Antonio Machado, 31058  
Toulouse Cedex  
loferre@univ-tlse2.fr  
<http://www.univ-tlse2.fr/grimm/smash/ferre/index.html>

**Résumé.** L'extraction d'information de grands graphes repose le plus souvent sur leur représentation dans des espaces de dimension réduite et on utilise généralement des méthodes factorielles appliquées à des mesures de dissimilarités calculées à partir des matrices associée du graphe ou l'analyse spectrale de leur Laplacien discret. Efficaces pour dégager les structures globales, ces représentations sont parfois peu exploitables dès lors que l'on s'intéresse à une perspective du graphe à partir de certains sommets privilégiés. Or l'information recherchée a souvent un caractère "local". Pour représenter le graphe du point de vue d'un ou de plusieurs sommets sélectionnés, nous proposons une méthode d'Analyse en Composantes Principales "Granulaire" consistant à appliquer une A.C.P. "filtrée" à un tableau de proximités. La visualisation d'un graphe de dictionnaire dont la mesure de proximité est obtenue à partir d'un algorithme original illustre notre propos.

## 1 Introduction

De par l'étude des grands réseaux, la visualisation de grands graphes est un problème crucial qui se rencontre dans des domaines aussi variés que la linguistique, l'électronique, la biologie, la "webologie" (étude et représentation du Web), les neurosciences etc... Selon la problématique, on cherche à révéler tel ou tel aspect privilégié du graphe au travers de sa représentation. Souvent, on vise à le représenter globalement et à identifier les composantes connexes ou les composantes "les plus" connexes.

Pour atteindre cet objectif, il est possible de représenter un graphe par les coordonnées des sommets dans un espace propre de son Laplacien (voir, e.g., [Chung, 1997], [Kuntz *et al.*, 2001], [Mohar, 1991]): il est notamment connu que la multiplicité de la valeur propre nulle est égale au nombre de composantes connexes du graphe. Citons également, l'analyse de contiguïté [Lebart, 2001] qui conduit à une représentation du graphe par Analyse des Correspondances de la matrice associée (voir également [Benzecri *et al.*, 1973]) ce qui revient encore à projeter sur des sous-espaces propres du Laplacien. Une deuxième solution est fournie par la construction d'une mesure de similarité ou de dissimilarité entre sommets du graphe à partir de sa matrice associée (en utilisant un indice de type Dice, Jacquard,...). On peut ainsi appliquer à cette