

# Visualisation interactive de graphes de grande taille

Saïd KAROUACH  
IRIT/ SIG, Université Paul Sabatier,  
118 route de Narbonne, 31062 Toulouse cedex 04  
[karouach@irit.fr](mailto:karouach@irit.fr)

## 1 Introduction

Les tableaux de contingence ou de cooccurrence peuvent être considérés comme des matrices valuées attachées à des graphes non orientés. Si la matrice est symétrique, les lignes (et les colonnes) représentent les sommets du graphe et les éléments de la matrices les arêtes (simultanément présence et valeur de chaque arête). Si la matrice est asymétrique, les lignes et les colonnes représentent deux type différents de sommets, le graphe est alors dit biparti (les arêtes du graphe ne relient, dans ce cas, que des sommets de types différents).

Actuellement, nous développons des méthodes de visualisation interactive de graphes de grande taille basées sur la notion d'attraction et de répulsion et paramétrables par l'utilisateur. Cette méthode de restitution est très appréciée par les utilisateurs, car elle ne nécessite pas de connaissances mathématiques particulières. Le tout est de proposer un graphe à la fois fidèle à la réalité, suffisamment lisible et autorisant des manipulations interactives axées sur la découverte de relations pertinentes et de stratégies. Nous avons testé plusieurs modes de représentation qui permettent d'apprécier à la fois l'importance de chaque sommet et de chaque connexion (arête) dans la structure visualisée.

## 2 Principe du codage de l'information

### 2.1 Mesures utilisées

Pour les sommets, il est possible de traduire leur importance de différentes manières :

- valeur diagonale (nombre d'occurrences),
- fréquence,
- norme de leur ligne
- norme de leur colonne, ...

Pour les arêtes

- leur module (cooccurrence ou présence absence),
- éventuellement application d'un seuil.

### 2.2 Codage des sommets

Pour traduire les échelles de valeurs des sommets, nous utilisons :

- des points colorés suivant un nuancier,
- des cercles de taille variable,
- des barres verticales.

## Visualisation interactive de graphes de grande taille

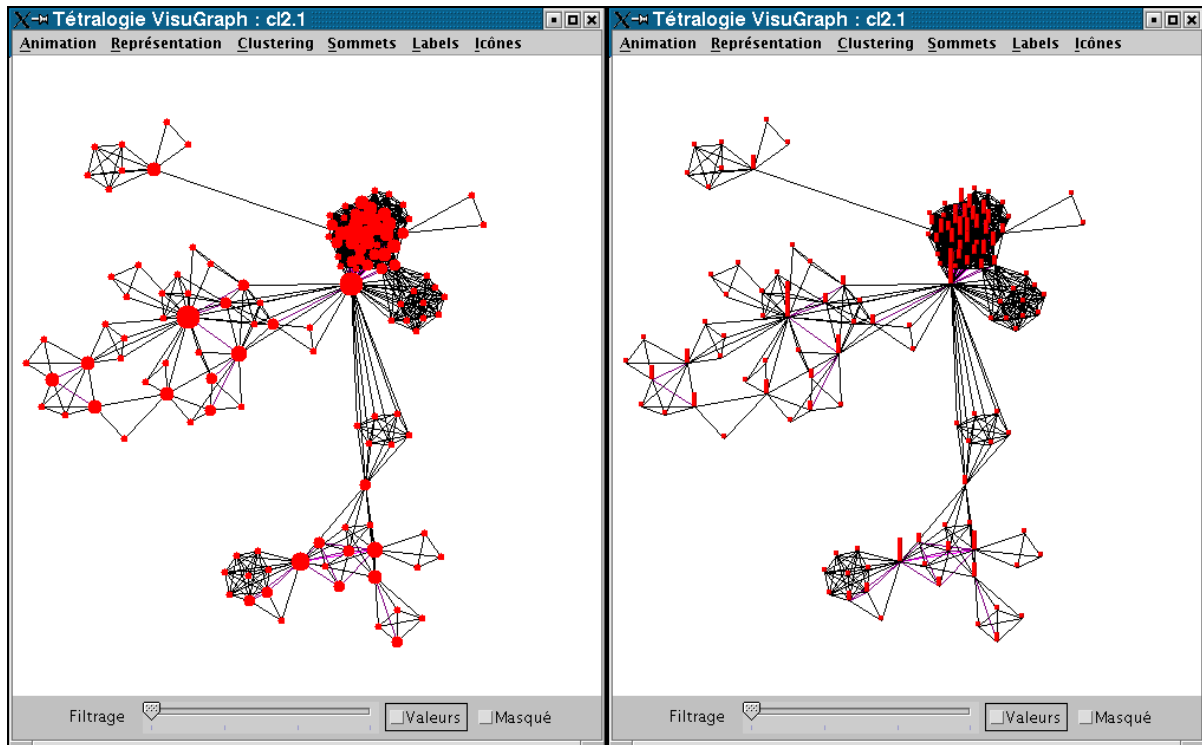


Figure 1. : Tailles des sommets visualisées par des icônes circulaires ou des barres.

### 2.3 Codage des arêtes

Les échelles de valeurs sont ici traduites, soit par des nuances d'une couleur, soit par l'épaisseur. Il est toujours possible d'indiquer la valeur d'une arête en son milieu.

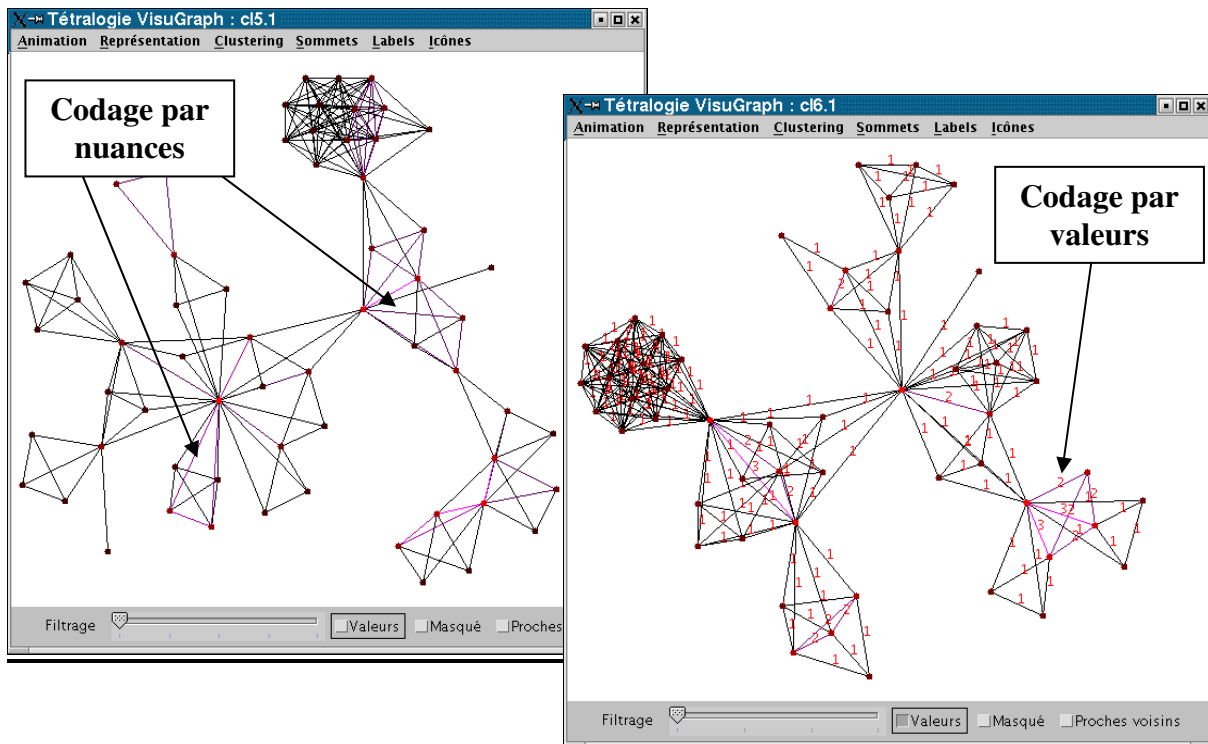


Figure 2. : Codage de la taille des arêtes par des nuances (et valeurs).

### 3 Recherche de classes et de connecteurs

#### 3.1 Cas des matrices symétriques

Ces visualisations permettent de naviguer dans la multitude des liens mis à jour, en proposant des vues partielles (obtenues par seuillage) ou simplifiées (graphes de classes ou d'agrégats) qui permettent de comprendre rapidement et de façon intuitive certaines organisations locales.

Dans l'exemple, ci-dessous, nous proposons d'abord la vue globale puis une de ses simplifications qui permet de détecter les groupes et les individus qui servent d'interfaces. De plus, la coloration des libellés (ici fonction de la valeur diagonale, de la marginale ou de la norme du vecteur) met l'accent sur les chefs de file du domaine et permet de les cibler pour une étude plus approfondie de leur influence.

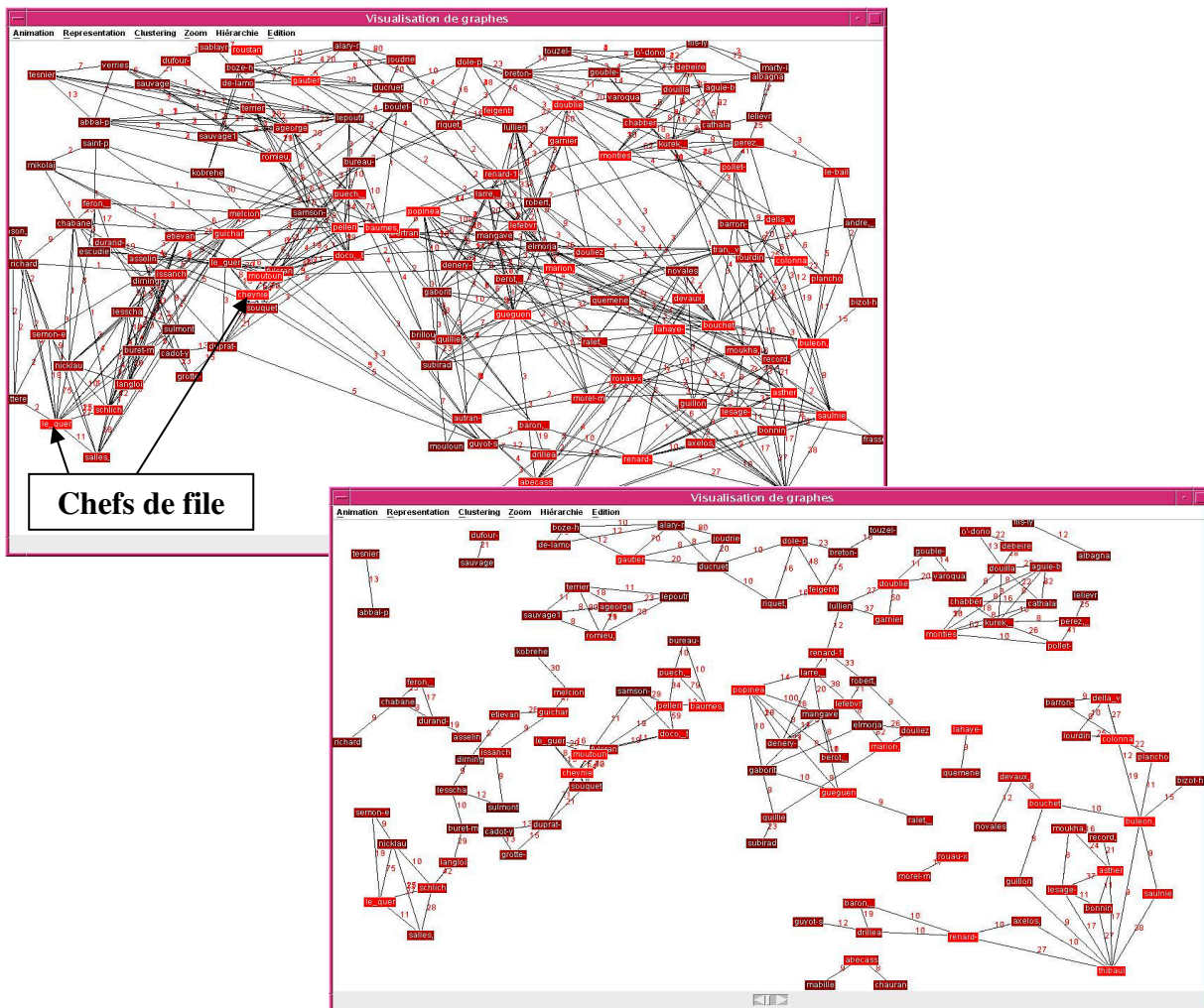


Figure 3. : Graphe de liens avant et après simplification par seuillage.

La classe connexe initiale se décompose, ici, en plusieurs sous-classes pseudo-connexes (déconnexion partielle), les éléments interfaces servent de liens entre des groupes qui, sans leur présence, seraient eux aussi isolés.

#### 3.2 Cas particulier des graphes bipartis

Si le graphe est issu d'une matrice non symétrique croisant deux dictionnaires différents, la recherche de connecteurs s'effectue de façon indirecte. En effet, une ligne connecte des colonnes et une colonne des lignes. Dans l'exemple ci-dessous, où nous avons choisi une matrice Auteurs – Pays : certains auteurs en rouge sont des liens privilégiés entre des pays. Par contre, dans les groupes isolés, le pays en vert devient l'élément fédérateur. Ici encore, la coloration donne une idée de l'importance de

## Visualisation interactive de graphes de grande taille

chacun. Dans le cas de la présence de plusieurs connecteurs (à gauche de la figure suivante) il est donc possible de détecter immédiatement le plus actif d'entre eux (ici celui qui se trouve au centre).

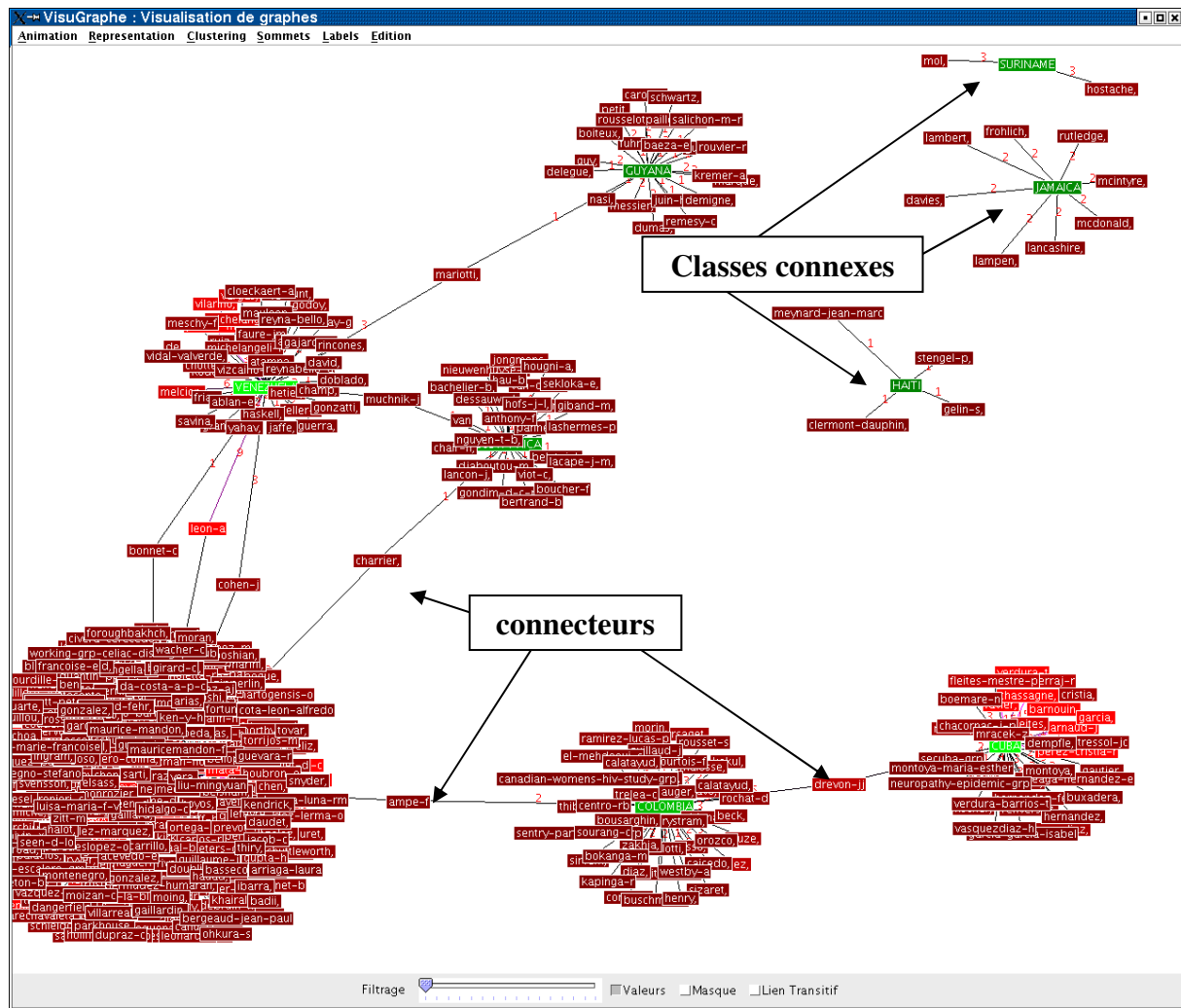


Figure 4. : Graphe biparti : les éléments clés sont soit des auteurs soit des pays.

## 4 Placement dirigé des sommets

Afin de rendre plus lisible la visualisation de graphes complexes, nous avons tenu compte, dans le placement des sommets, de la sérialisation des items obtenue à l'issue d'un tri par blocs diagonaux. Les sommets sont alors distribués de façon circulaire, leurs principales connexions s'effectuent avec leurs voisins et seules les liaisons inter-agrégats attirent maintenant l'attention. Ce principe est applicable aussi bien pour les matrices symétriques (vrai diagonale) qu'asymétriques (pseudo diagonale). Dans ce dernier cas, il convient de synchroniser le placement des lignes (à l'extérieur) avec celui des colonnes (à l'intérieur) afin de générer le moins possible de croisements des arêtes.

Dans l'exemple suivant, où nous avons croisé des Auteurs et des Pays. La matrice est triée par blocs diagonaux en mode relatif. Les Auteurs (lignes de la matrice) sont donc disposés à la périphérie, les pays moins nombreux à l'intérieur. Nous pouvons immédiatement remarquer :

- Les pays totalement isolés,
- Les auteurs isolés dans un pays isolé,
- Les groupes d'auteurs qui collaborent à l'international,
- Les acteurs de ces collaborations à l'intérieur des groupes,
- Les pays liés au travers des auteurs,

## Visualisation interactive de graphes de grande taille

- La force de ces liaisons,
- Les éléments clés (connecteurs uniques, plaques tournantes),
- Le leader de chaque groupe (par la coloration)
- La taille de chaque agrégat qu'il soit ou non isolé,
- L'inclusion de certains pays dans un ou plusieurs autres plus actifs.

En fait, les éléments remarquables (liens importants) sont transversaux, ils représentent, ici, des collaborations internationales inhabituelles (les collaborations principales étant obtenues par voisinage).

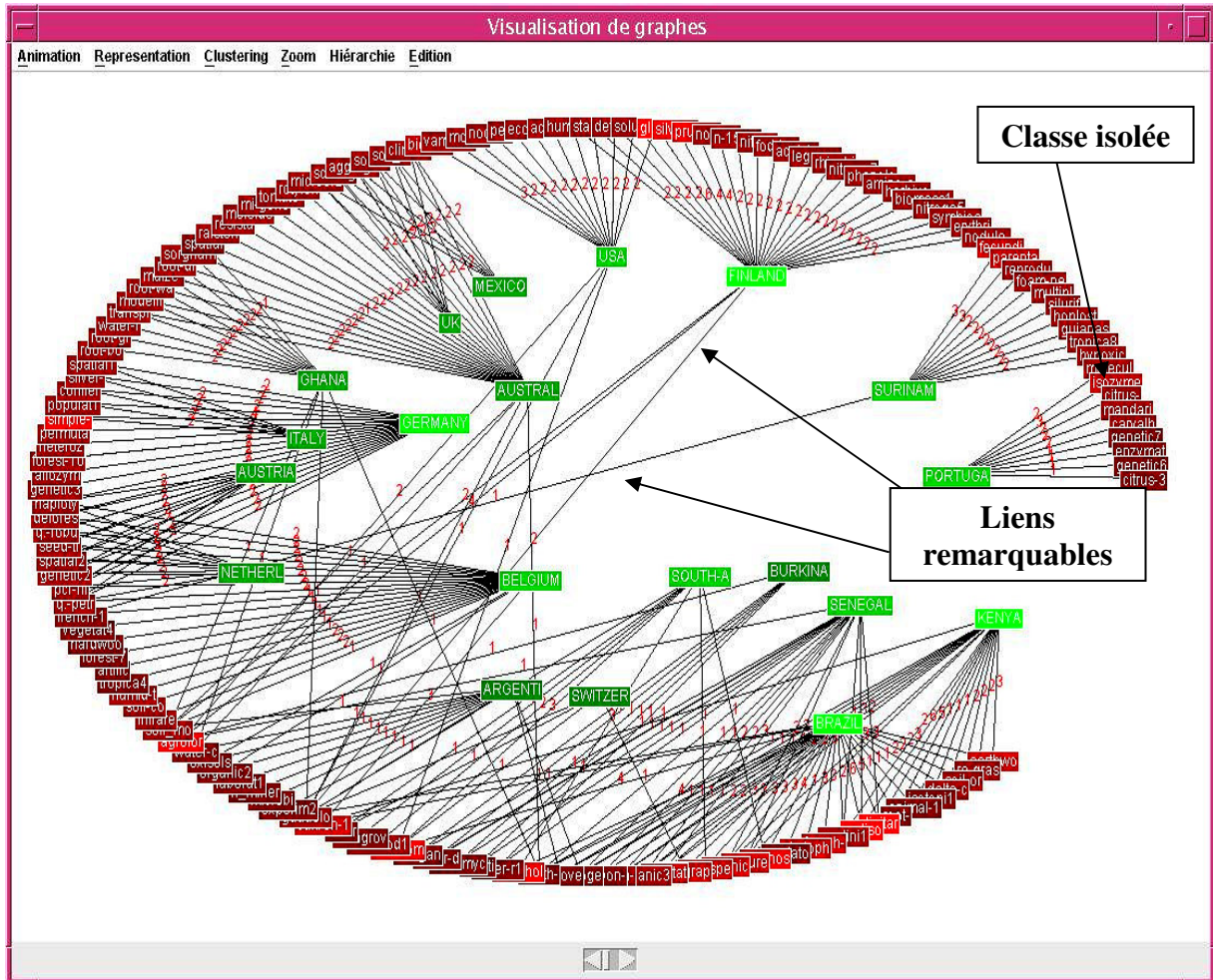


Figure 5. : Placement des items en fonction d'un tri par blocs diagonaux.

## 5 Cas particulier des arbres partiels

Une méthode radicale de simplification d'un graphe connexe est de ne conserver qu'un de ses arbres partiels extremum (ici maximum). Pour cela, il suffit, pour chaque sommet, de garder sa connexion la plus forte (première arête de chaque co-cycle). Le graphe obtenu permet de savoir quels sont les éléments clés du domaine (connecteurs privilégiés). De plus, le placement des sommets obtenu à partir d'un de ses arbres planaires peut servir de base pour la reconstruction du graphe complet ou de l'une de ses simplifications moins poussées.

En voici une illustration qui permet de juger de la lisibilité de cette méthode dans le cas d'une matrice asymétrique (graphe biparti).

## Visualisation interactive de graphes de grande taille

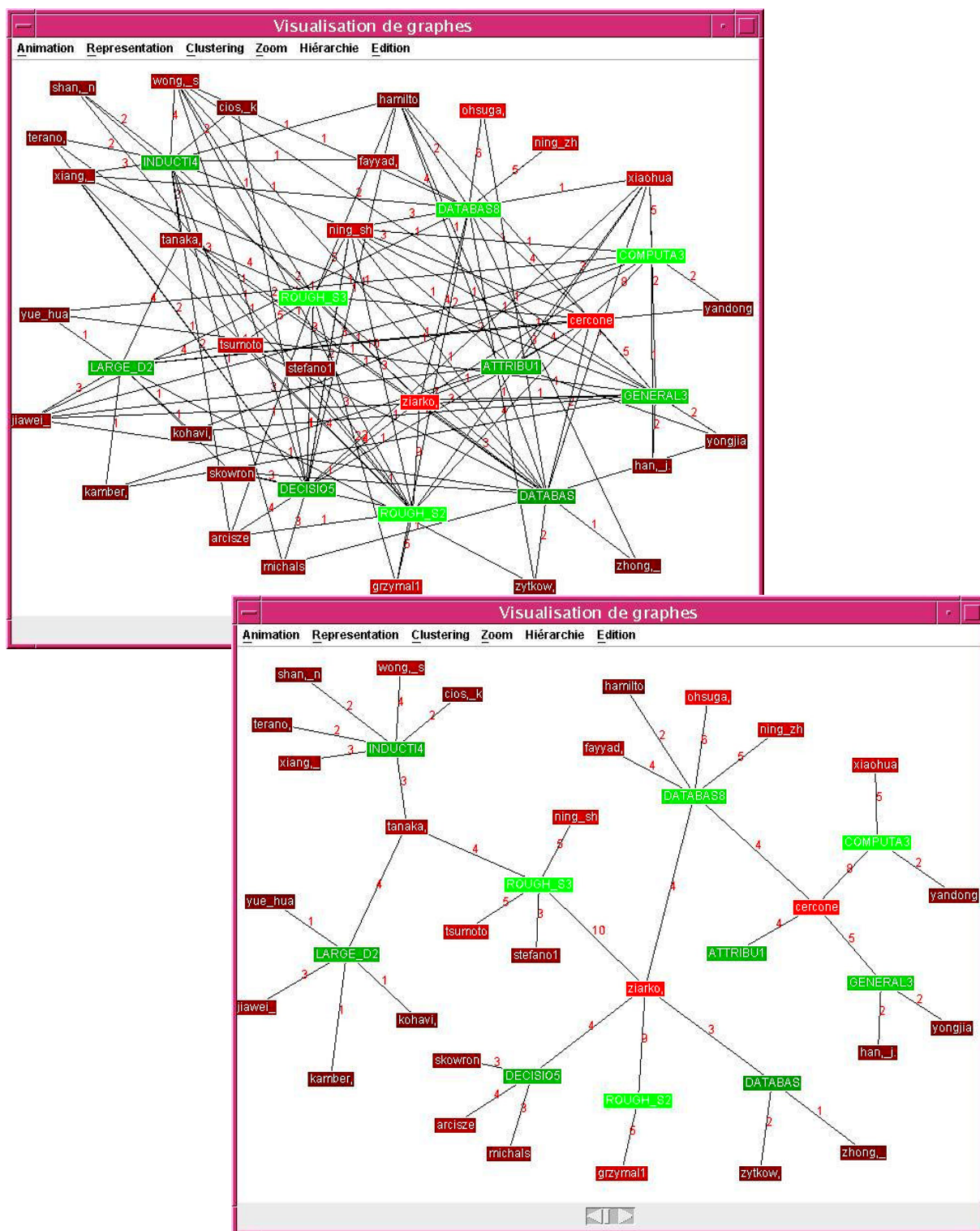


Figure 6. : Graphe simplement connexe et son arbre partiel maximum..

## 6 Extension du graphe depuis un sommet

Afin de réaliser une navigation locale dans le graphe initial souvent trop complexe, il est possible de travailler sur un sous-graphe. Pour cela, nous partons d'un sommet particulier choisi dans une liste alphabétique et nous étendons progressivement le graphe, depuis ce sommet, par transitivité. Cette technique permet, par un changement de focus, de nous concentrer sur un extrait pertinent issu d'une information ciblée (acteur, mot-clé, concept).

## Visualisation interactive de graphes de grande taille

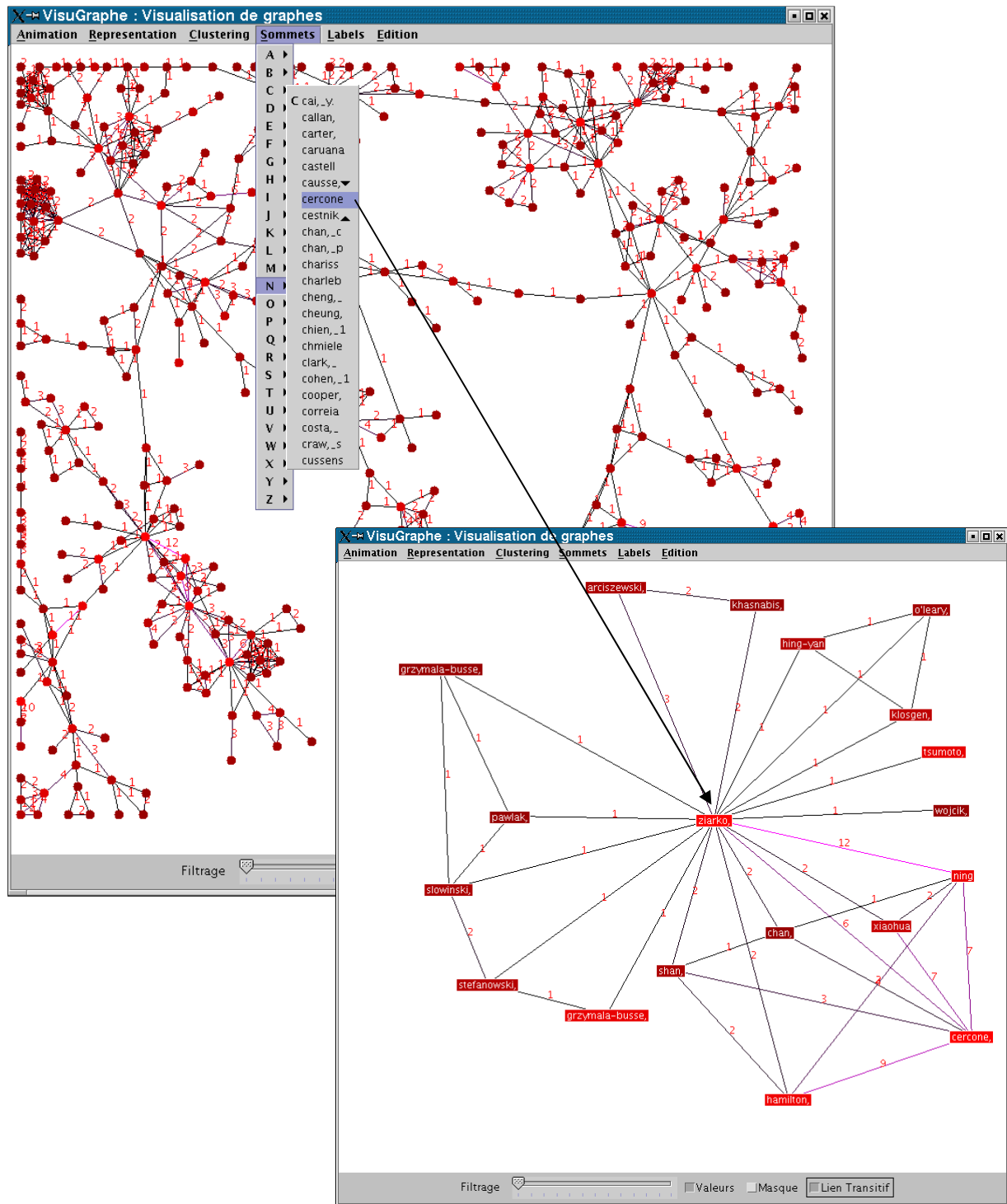


Figure 7. : Extension d'un sous graphe par transitivité depuis un sommet sélectionné.

## 7 Graphes réduit

### 7.1 Génération du graphe réduit

Une autre méthode pour contourner la complexité de certaines structures est de travailler sur le graphe des classes. Il s'agit alors de décomposer le graphe initial en classes homogènes, afin de regrouper les sommets les plus fortement connectés. Un graphe réduit est alors obtenu : ses sommets sont définis par les classes elles mêmes et ses arêtes par les connexions entre classes. Une fois ce nouveau graphe

## Visualisation interactive de graphes de grande taille

général et étudié, il est possible de revenir au graphe initial en gardant visible l'appartenance de chaque sommet à sa classe grâce à une coloration identique. Cette technique permet donc une approche macroscopique de la structure avec la possibilité de zoomer si nécessaire sur certains détails comme ci-dessous :

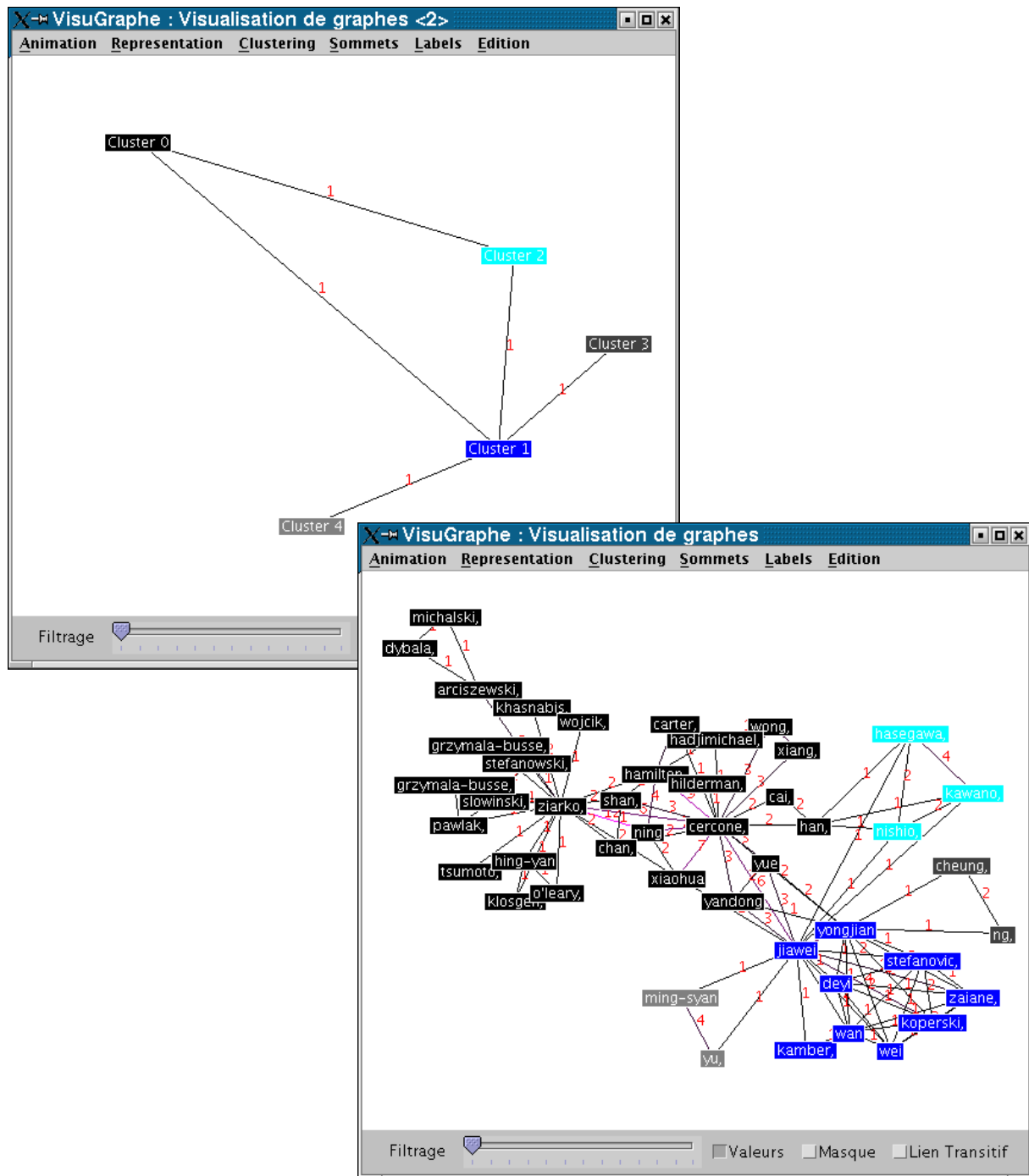


Figure 8. : Graphe réduit et graphe initial associé.

### 7.2 Extraction des classes du graphe réduit

A partir du graphe réduit, nous avons la possibilité de visualiser séparément chacune de ses classes afin de réaliser une analyse détaillée de sa structure et d'en comprendre ses mécanismes. Chaque classe est alors considérée comme un nouveau graphe et bénéficie de l'ensemble des fonctions



## Visualisation interactive de graphes de grande taille

interactives vues plus haut. Comme le nombre de sommets est maintenant réduit, il est beaucoup plus facile de détecter des stratégies et d'identifier des éléments importants tels que les connecteurs, les leaders, les sous groupes, ...

Dans l'exemple ci-dessous, nous avons décomposé un graphe connexe en 3 classes et nous avons, à partir du graphe réduit, demandé l'affichage séparé de chaque classe. Les sommets gardent la couleur qui leur a été attribuée et en leur appliquant une nouvelle migration il est possible de simplifier encore le dessin de ce graphe. Ici, le graphe réduit permet de garder en mémoire la structure initiale : les clusters 1 et 2 ne sont pas connectés, le cluster 0 est donc le cluster central.

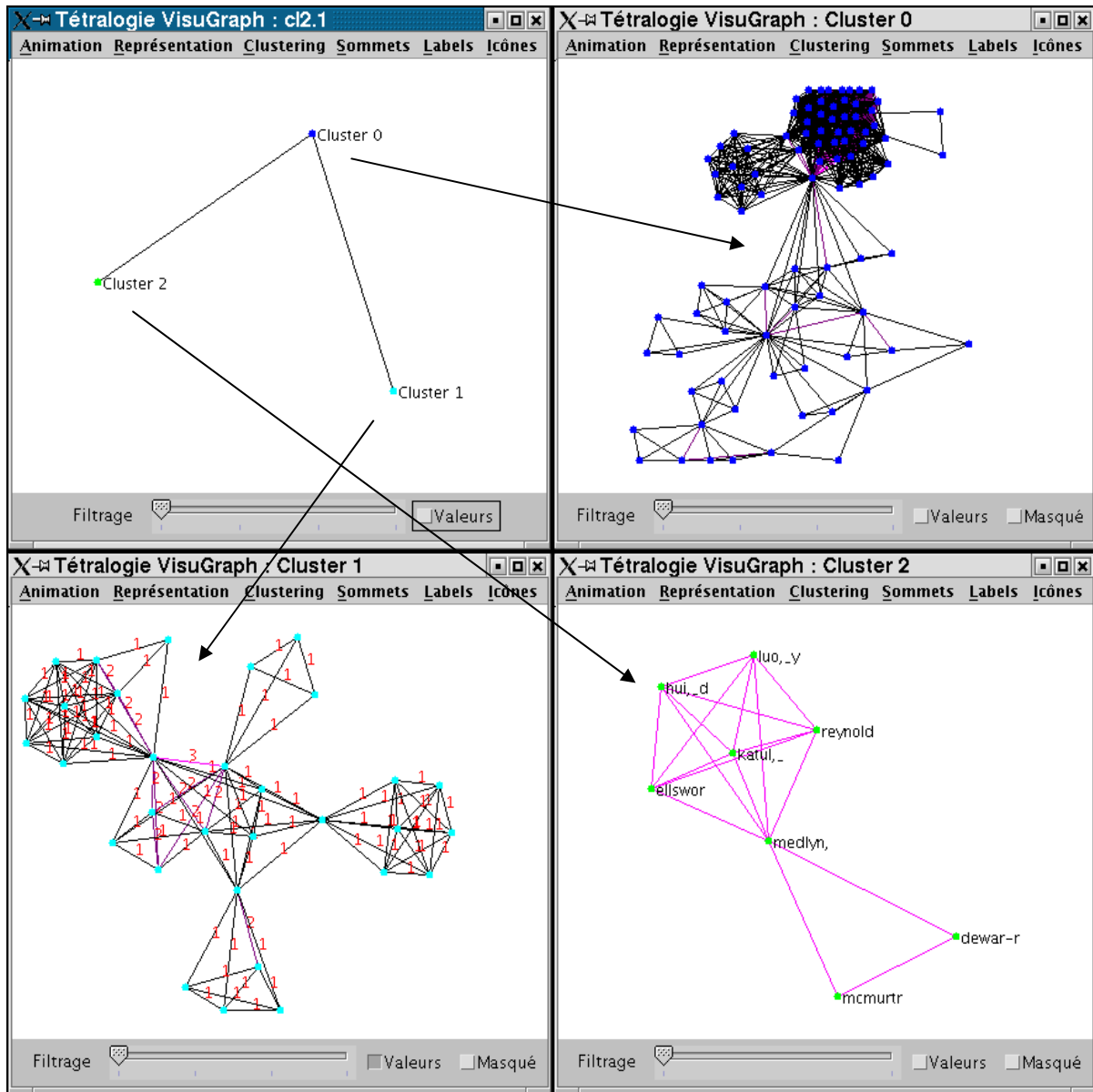


Figure 9. : Graphe réduit et graphes de ses classes.

## 8 Conclusion

Cette technique de visualisation du relationnel est très prisée par les utilisateurs car elle fait intervenir, de façon intuitive, un certain nombre de réflexes déjà acquis. Nous essayons d'en augmenter la convivialité et les possibilités de navigation. Parallèlement, nous avons pour projet de l'étendre aux données évolutives : morphing de graphes, trajectoires, animations, ...