

GENERATION AUTOMATIQUE DE RESEAUX EN BIBLIOMETRIE

E. BOUTIN, P DUMAS, L. QUONIAM, H. ROSTAING, H. DOU

Laboratoire Le Pont Université de Toulon et du Var. BP 132. 83957 LA GARDE Cedex:
CRRM. Université Aix-Marseille III, 13397 Marseille Cedex 20

OBJECTIF

L'objectif est de présenter un outil qui permet la création automatique de réseaux à partir de données bibliométriques. On entend par réseau une représentation figurative des différentes composantes d'un ensemble et des relations qui les unissent. Une telle représentation peut s'appliquer aux champs codes, auteurs, mots clés de notices bibliométriques extraites de bases de données brevets, scientifiques, techniques, technico-économiques. Les réseaux, une fois constitués mettent en évidence la notion de cluster de mots clé ou de collaborateurs d'auteurs.

L'intérêt porté aux réseaux en bibliométrie est actuellement limité par une technique de construction souvent manuelle qui rend l'opération fastidieuse. Dans ces conditions, un algorithme débouchant sur la représentation automatique d'un réseau satisfaisant certaines contraintes esthétiques présente un certain intérêt. **METHODOLOGIE.** L'algorithme suit une démarche itérative qui peut être représentée par la figure 1. A partir d'un positionnement initial, une évaluation est faite du caractère esthétique du graphe. De là un moteur va assurer la génération d'un nouveau positionnement des points qui seront à leur tour évalués. La boucle est répétée jusqu'à l'obtention d'une solution acceptable. Le choix d'un positionnement initial des points peut être réalisé de différentes façons: détermination aléatoire, points confondus au départ, choix d'un positionnement particulier. Or ce choix ne doit pas être laissé au hasard car il conditionne la vitesse de convergence de l'algorithme. Pour déterminer cette situation de départ, il est intéressant d'utiliser certains résultats de la théorie des graphes et du calcul matriciel. L'algorithme de sériation, développé au CRRM, conduit à découper un corpus en plusieurs sous groupes. D'autre part, les algorithmes de détermination des isthmes permettent d'identifier certains acteurs qui jouent le rôle de pivots entre les groupes. On considérera que, dans la situation initiale, les points d'un même groupe seront confondus:

Le choix d'une fonction objectif. Cette fonction va permettre de mesurer le caractère esthétique d'un positionnement. Cette fonction intègre un certain nombre de critères permettant d'apprécier la lisibilité d'un réseau. Ces critères sont représentés figure 2. On considère qu'une représentation de réseau est lisible si ses intersections arêtes/arêtes, arêtes/sommets et sommets/sommets sont limitées.

Le moteur est la partie centrale car il permet de passer d'un positionnement des sommets à un autre en garantissant la convergence vers une solution optimale. Pour ce faire, deux manipulations sont effectuées successivement. La première consiste à reproduire d'une génération à l'autre les points les mieux placés et à les combiner entre eux pour générer de nouveaux poids. Cette démarche est complétée par un algorithme «tectonique». Il consiste, à partir de la solution obtenue précédemment, à étudier l'effet de modifications marginales du positionnement des points sur l'esthétique du réseau.