

Fusion de connaissances imparfaites pour l'appariement de données géographiques – Proposition d'une approche s'appuyant sur la théorie des fonctions de croyance

Ana-Maria Olteanu-Raimond

De nos jours, il existe de nombreuses bases de données géographiques (BDG) couvrant le même territoire. Les données géographiques sont modélisées différemment (par exemple une rivière peut être modélisée par une ligne ou bien par une surface), elles sont destinées à répondre à plusieurs applications (visualisation, analyse) et elles sont créées suivant des modes d'acquisition divers (sources, processus). Tous ces facteurs créent une indépendance entre les BDG, qui pose certains problèmes à la fois aux producteurs et aux utilisateurs.

Ainsi, une solution est d'explicitier les relations entre les divers objets des bases de données, c'est-à-dire de mettre en correspondance des objets homologues représentant la même réalité. Ce processus est connu sous le nom d'appariement de données géographiques.

La complexité du processus d'appariement fait que les approches existantes varient en fonction des besoins auxquels l'appariement répond, et dépendent des types de données à apparier (points, lignes ou surfaces) et du niveau de détail. Nous avons remarqué que la plupart des approches sont basées sur la géométrie et les relations topologiques des objets géographiques et très peu sont celles qui prennent en compte l'information descriptive des objets géographiques. De plus, pour la plupart des approches, les critères sont enchaînés et les connaissances sont à l'intérieur du processus.

Suite à cette analyse, nous proposons une approche d'appariement de données qui est guidée par des connaissances et qui prend en compte tous les critères simultanément en exploitant à la fois la géométrie, l'information descriptive et les relations entre eux. Afin de formaliser les connaissances et de modéliser leurs imperfections (imprécision, incertitude et incomplétude), nous avons utilisé la théorie des fonctions de croyance [Shafer, 1976].

Notre approche d'appariement de données est composée de cinq étapes : après une sélection des candidats, nous initialisons les masses de croyance en analysant chaque candidat indépendamment des autres au moyen des différentes connaissances exprimées par divers critères d'appariement. Ensuite, nous fusionnons les critères d'appariement et les candidats. Pour finir, une décision est prise.

Nous avons testé notre approche sur des données réelles ayant des niveaux différents de détail représentant le relief (données ponctuelles) et les réseaux routiers (données linéaires).

Mots-clés : appariement, données géographiques, fusion, connaissances, imperfection