

Proposition de stages M2 informatique

Analyse de données spatio-temporelles réelles

ICube, équipe SDC

Mars à Septembre 2021

Contexte

Valoriser les grandes **masses de données spatio-temporelles** disponibles dans différents domaines est crucial. Ceci nécessite de concevoir et de développer des approches innovantes aptes à traiter conjointement les aspects spatiaux et les aspects temporels, ce qui n'est que peu le cas avec les méthodes actuelles. Si les graphes, outils puissants à la fois théoriquement et méthodologiquement, sont utilisés pour modéliser des phénomènes temporels ou spatiaux, les méthodes actuelles de **fouille, d'analyse et d'extraction de connaissances** n'exploitent, en général, qu'une seule dimension de l'information, spatiale versus temporelle. Cela implique souvent une perte de précision et possibilité d'interprétation des résultats. C'est pourquoi nous nous proposons d'exploiter un modèle de graphe spécifique intégrant différents types de relations, les **graphes spatio-temporels** [2]. Un exemple est présenté sur la figure 1.

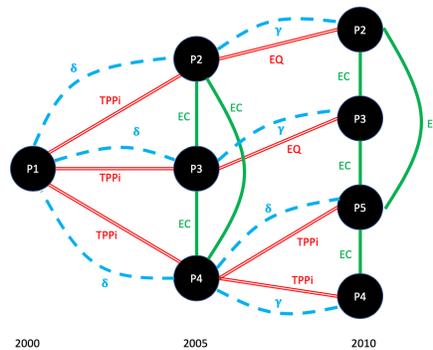


FIGURE 1 – Exemple de graphe spatio-temporel : en vert les relations spatiales, en rouge les relations spatio-temporelles, en bleu pointillé les relations de filiation (conservation ou partage de l'identité)

Sujet

L'objectif général est d'étudier et développer des techniques pour trouver dans de grands graphes spatio-temporels, des motifs, ou répétitions, auxquels les experts du domaine concerné peuvent donner sens. Plus précisément, il s'agit de mettre en œuvre des techniques de **recherche de motifs (sous-graphes fréquents)** [3] sur des jeux de données réelles : images satellitaires, bd géographiques, IRM fonctionnelle (en relation avec des experts de ces domaines). Un exemple est montré sur la figure 2.

La ou le stagiaire s'appuiera pour cela sur les outils disponibles au laboratoire (pour la construction des graphes [4] et pour l'extraction des motifs [3]) mais devra aussi adapter ou améliorer ces outils. Les tâches suivantes seront à réaliser :

- Bibliographie
- Sélection et prétraitement des données
- Transformation : génération de graphes spatio-temporels, incluant la modélisation des relations spatiales et spatio-temporelles adéquates [5, 1]
- Recherche des motifs fréquents; extension de l'approche aux motifs inexacts (approximation des étiquettes, ou du nombre d'arêtes et de sommets)
- Interprétation et comparaison des résultats sur (au moins) deux jeux de données

En fonction de l'avancée, on pourra aussi envisager de comparer les résultats obtenus en recherchant des motifs inexacts dans des données précises ou des motifs exacts dans des données imprécises (par exemple en utilisant des étiquettes floues ou plus générales).

L'objectif final est de construire une chaîne de traitement incluant l'ensemble des opérations sur les données d'un domaine.



FIGURE 2 – Exemple de données réelles à traiter, issues du Registre Parcellaire Graphique de 2015.

Aspects pratiques

Compétences attendues : master informatique, autonome en programmation (python), connaissances sur les graphes, intérêt pour l'aspect expérimental

Conditions : accueil au laboratoire ICube, Illkirch ; rémunération : 3,90 € / heure (à ajuster en 2021), pc linux ou mac.

Encadrantes : Aurélie Leborgne, Florence Le Ber et Stella Marc-Zwecker

Contacts : aurelie.leborgne@unistra.fr, florence.leber@engees.unistra.fr, stella@unistra.fr

À noter : si le travail réalisé le permet, il sera valorisé par l'écriture d'un article dans lequel vous serez co-auteur. Le stage pourrait se prolonger par une thèse.

Références

- [1] Eliseo Clementini and P. Di Felice. Approximate topological relations. *International Journal of Approximate Reasoning*, 16(2) :173–204, 1997.
- [2] Géraldine Del Mondo, J. G. Stell, C. Claramunt, and R. Thibaud. A graph model for spatio-temporal evolution. *Journal of Universal Computer Science*, 16 :1452–1477, 2010.
- [3] Vijay Ingalalli, Dino Ienco, and Pascal Poncelet. Mining frequent subgraphs in multigraphs. *Information Sciences*, 451-452, 04 2018.
- [4] Aurélie Leborgne, Adrien Meyer, Henri Giraud, Florence Le Ber, and Stella Marc-Zwecker. Un graphe spatio-temporel pour modéliser l'évolution de parcelles agricoles. In *Conférence SAGEO*, Clermont-Ferrand, France, 2019.
- [5] David A. Randell, Zhan Cui, and Anthony G. Cohn. A spatial logic based on regions and connection. In *Proceedings 3rd International Conference on Knowledge Representation and Reasoning*, 1992.