



HAL
open science

Extraction de motifs spatio-temporels fréquents dans les données Corine Land Cover

Assaad Oussama Zeghina, Aurélie Leborgne, Florence Le Ber, Antoine Vacavant

► **To cite this version:**

Assaad Oussama Zeghina, Aurélie Leborgne, Florence Le Ber, Antoine Vacavant. Extraction de motifs spatio-temporels fréquents dans les données Corine Land Cover. conférence Spatial Analysis and GEomatics (SAGEO), May 2025, Avignon, France. <hal-05553116>

HAL Id: hal-05553116

<https://hal.science/hal-05553116v1>

Submitted on 14 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

Extraction de motifs spatio-temporels fréquents dans les données Corine Land Cover

Une approche exploitant la méthode Multi-SPMiner

Assaad ZEGHINA¹, Aurélie LEBORGNE¹,
Florence LE BER¹, Antoine VACAVANT²

1. Université de Strasbourg, ENGEES, CNRS, ICube UMR 7357
`{aozeghina,aurelie.leborgne,florence.le-ber}@unistra.fr`

2. Université Clermont Auvergne, Clermont Auvergne INP, CNRS
`antoine.vacavant@uca.fr`

ABSTRACT. This study explores the extraction of frequent spatiotemporal patterns from the Corine Land Cover (CLC) dataset using the Multi-SPMiner approach. Multi-SPMiner leverages Multi-Graph Convolutional Networks (MGCN) to encode spatiotemporal relationships and identify meaningful motifs efficiently. The method was applied to CLC data, highlighting key temporal, spatial, and spatiotemporal motifs, and demonstrating robust performance in analyzing land use dynamics across Europe.

MOTS-CLÉS : extraction de motifs, graphes spatio-temporels, réseaux neuronaux, Corine Land Cover

KEYWORDS: pattern extraction, spatiotemporal graphs, neural networks, Corine Land Cover

1. Introduction

L'extraction de motifs fréquents dans les données graphes est une méthode clé pour identifier des sous-graphes récurrents au sein de structures complexes, révélant des informations importantes sur les relations sous-jacentes dans les données. Largement mise en œuvre dans des domaines tels que les réseaux sociaux et la biologie (Jiang *et al.*, 2013), cette tâche devient plus complexe dans les multigraphes en raison de la diversité des types d'arêtes (Zeghina *et al.*, 2023). Les graphes spatio-temporels jouent un rôle essentiel dans la modélisation de phénomènes évoluant dans le temps et l'espace, comme les changements d'utilisation des sols ou les dynamiques écologiques (Leborgne *et al.*, 2019; Del Mondo *et al.*, 2013). Toutefois, leur complexité impose des approches avancées pour une extraction efficace des motifs. Multi-SPMiner exploite des réseaux neuronaux convolutifs pour capturer les interactions dans ces graphes complexes. En combinant précision et scalabilité, ce modèle est particulièrement adapté aux données réelles, telles que la base Corine Land Cover (CLC), qui offre une vue détaillée des dynamiques d'utilisation des sols en Europe. Cet article explore l'application de Multi-SPMiner sur le jeu de données CLC, mettant en avant ses performances dans l'extraction de motifs significatifs pour des analyses géographiques.

2. Méthodologie

La méthode Multi-SPMiner. La méthode Multi-SPMiner traite les multigraphes avec des attributs spatio-temporels. Cette méthode suit un processus en deux étapes pour l'extraction de motifs fréquents. Tout d'abord, elle utilise un réseau de neurones de multi-graphes (Multi Graph Convolutional Network, MGCN) pour encoder les voisinages des nœuds dans un espace ordonné, où les relations entre sous-graphes sont préservées. Cette phase d'encodage garantit une identification systématique des motifs d'intérêt en fonction de leur similarité structurelle. Dans la deuxième étape, Multi-SPMiner effectue une recherche itérative de motifs en parcourant l'espace. À partir d'un nœud initial choisi au hasard, des nœuds voisins sont ajoutés de manière incrémentale au motif. La sélection des nœuds est guidée par la minimisation d'une pénalité, qui évalue l'adéquation d'un motif candidat à la structure globale du graphe. Ce processus se poursuit jusqu'à ce que des motifs de la taille souhaitée soient identifiés, garantissant des valeurs de support élevées et une pertinence accrue.

Graphes spatio-temporels. Un graphe spatio-temporel est un graphe où les nœuds représentent des entités spatiales à différents instants et les arêtes représentent les relations spatiales, spatio-temporelles et de filiation entre les entités (Del Mondo *et al.*, 2013). Il peut être vu comme l'union de trois sous-graphes (voir figure 1) :

- Le sous-graphe des relations spatiales représente les interactions spatiales entre les entités à un instant donné. Dans ce travail, nous les modélisons par

les relations topologiques de la théorie RCC8 (Randell *et al.*, 1992).

– Le sous-graphe des relations spatio-temporelles représente les interactions spatio-temporelles entre les entités à deux instants temporels successifs. Les relations spatio-temporelles sont également issues de la théorie RCC8.

– Le sous-graphe des relations de filiation représente la transmission de l'identité entre les entités à deux instants temporels successifs.

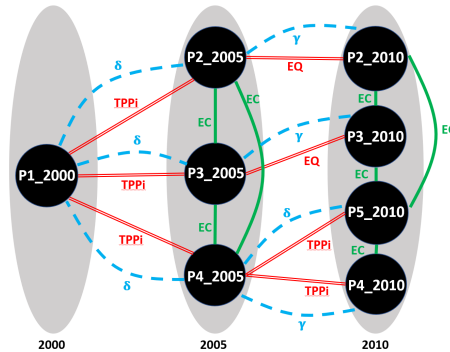


FIGURE 1. Représentation d'un graphe spatio-temporel : les lignes rouges représentent les relations spatio-temporelles, les lignes vertes représentent les relations spatiales, et les lignes bleues représentent les relations de filiation.

Données initiales et construction des graphes. La ressource *Corine Land Cover* (CLC) fournit des informations sur l'occupation des sols dans les pays européens depuis 1990. Elle est constituée de cartes détaillées selon trois niveaux, allant par exemple de 'terre artificialisée' (niveau 1) à 'décharges' (niveau 3). Cinq dates sont disponibles : 1990, 2000, 2006, 2012 et 2018, avec une résolution spatiale de 25 hectares. CLC constitue une ressource essentielle pour étudier les phénomènes spatio-temporels comme les changements d'utilisation des sols et les évolutions environnementales. La construction des graphes spatio-temporels à partir des données CLC s'est déroulée en plusieurs étapes. Des zones urbaines ont d'abord été sélectionnées aléatoirement à travers l'Eu-

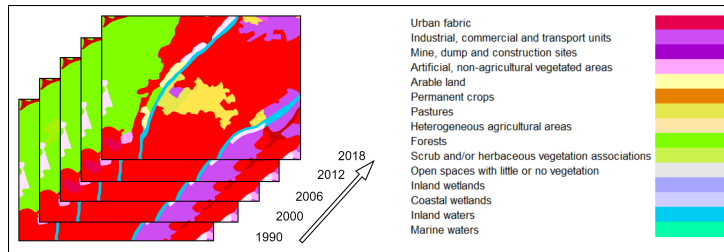


FIGURE 2. Extrait de données CLC illustrant l'évolution de la région parisienne, où les différentes occupations du sol sont codées par des couleurs.

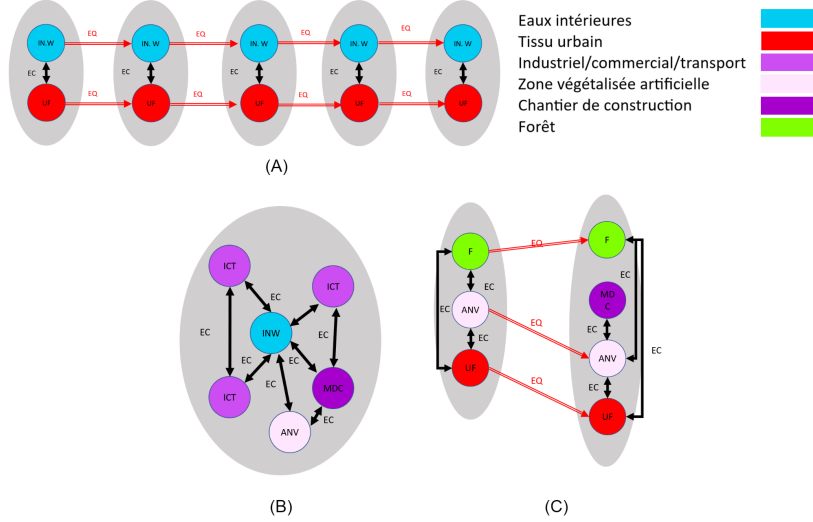


FIGURE 3. Exemples de motifs obtenus : (A) motif spatiotemporel retrouvé dans plusieurs villes Européenne (B) motif spatial partagé entre Paris et Venise (C) motif spatiotemporel provenant des banlieues de Berlin.

rope afin d'assurer une variabilité géographique suffisante. Pour chaque région ainsi définie et pour chacune des cinq périodes temporelles disponibles, un graphe a été construit : chaque sommet y représente une parcelle homogène, étiquetée selon son type d'occupation du sol, tandis que les arêtes modélisent les relations entre parcelles voisines, à la fois dans l'espace et dans le temps. Les sommets et les arêtes ont été annotés respectivement avec les types d'utilisation des sols et les relations d'adjacence spatiale ou temporelle. Enfin, afin de garantir une complexité suffisante pour l'analyse, les régions sélectionnées ont été progressivement étendues, permettant de générer des graphes comptant environ 1000 sommets par période.

3. Résultats et discussion

Le modèle Multi-SPMiner, utilisant un réseau de neurones convolutifs pour multigraphes (MGCN), s'est montré efficace sur le jeu de données CLC, atteignant jusqu'à 75% de motifs trouvés, selon les tailles de motifs et les graphes explorés. La vérité terrain a été obtenue à l'aide d'une approche algorithmique exhaustive, permettant de comparer précisément les performances du modèle. Dans cette étude, des motifs de tailles 6, 8 et 10 (en nombre de nœuds) ont été considérés pour évaluer la capacité du modèle à identifier à la fois des structures petites et simples ou plus grandes et complexes. Les expériences ont été menées sur une machine équipée d'un processeur Ryzen 7 9800X3D et d'une

carte graphique RTX 4080 16Go. Multi-SPMiner obtient ces résultats avec un temps d'exécution moyen inférieur à 30 secondes pour des graphes de taille 1000 avec des motifs de taille 10, en faisant l'une des approches les plus efficaces pour l'analyse des graphes spatio-temporels. L'étude a identifié trois types principaux de motifs : les motifs temporels, les motifs spatiaux et les motifs spatio-temporels combinés, comme illustré respectivement dans la figure 3. Les motifs temporels mettent en évidence les changements au fil du temps, tels que l'expansion urbaine progressive dans les zones environnantes, reflétant comment les nœuds urbains empiètent sur les espaces verts ou agricoles. Par exemple, le motif A (taille 10) illustre une région où les relations entre les zones d'eau et les surfaces urbaines restent constantes, représentant ainsi un motif partagé entre plusieurs grandes villes européennes. Des motifs similaires ont également été obtenus dans des zones rurales en Finlande (avec des variations). Cela reflète une certaine stabilité, bien que ces résultats puissent être influencés par le faible niveau d'activité humaine dans ces régions ou par des zones urbaines où les infrastructures sont déjà établies. Ces zones vertes ou surfaces d'eau, souvent stables, apparaissent dans les motifs sans apporter de réelle valeur en termes d'identification ou de tendances. Le motif B est un motif spatial partagé entre Paris et Venise. Il met en évidence des relations autour des lacs et rivières, avec une présence significative (> 25) de ce motif dans les zones concernées. Ces zones entourant les surfaces d'eau incluent également des sites commerciaux, de construction ou d'exploitation minière, ce qui souligne leur importance dans les dynamiques spatiales des deux villes. Enfin, le motif C (avec une fréquence > 25) révèle une dynamique spatiotemporelle liée à l'urbanisation dans les banlieues de Berlin. On y observe une transformation d'une zone artificielle, non agricole et végétalisée en chantier de construction. De plus, des sous-motifs plus petits, issus de la même composante, mettent en lumière les interactions complexes entre ces zones en mutation rapide.

Bibliographie

- Del Mondo G., Rodríguez M., Claramunt C., Bravo L., Thibaud R. (2013). Modeling consistency of spatio-temporal graphs. *Data & Knowledge Engineering*, vol. 84, p. 59–80.
- Jiang C., Coenen F., Zito M. (2013). A survey of frequent subgraph mining algorithms. *The Knowledge Engineering Review*, vol. 28, n° 1, p. 75–105.
- Leborgne A., Meyer A., Giraud H., Le Ber F., Marc-Zwecker S. (2019). Un graphe spatio-temporel pour modéliser l'évolution de parcelles agricoles. In *Sageo 2019*.
- Randell D. A., Cui Z., Cohn A. G. (1992). A spatial logic based on regions and connection. In *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning*.
- Zeghina A., Leborgne A., Le Ber F., Vacavant A. (2023). Multi-spmminer: A deep learning framework for multi-graph frequent pattern mining with application to spatiotemporal graphs. *Procedia Computer Science*, vol. 225, p. 1094–1103.