



HAL
open science

Les graphes pour étudier les dynamiques spatiales à partir des séries fiscales médiévales et modernes. Etat des lieux de l'expérience Modelespace

Samuel Leturcq, Romain Raveaux

► To cite this version:

Samuel Leturcq, Romain Raveaux. Les graphes pour étudier les dynamiques spatiales à partir des séries fiscales médiévales et modernes. Etat des lieux de l'expérience Modelespace. Bulletin du Centre d'études médiévales d'Auxerre. Hors-série, 2016, Géolocalisation et sources anciennes? Actes des journées d'études de Dijon, Maison des Sciences de l'Homme, 13-14 novembre 2014, 9, 10.4000/cem.13805 . hal-01311067

HAL Id: hal-01311067

<https://hal.science/hal-01311067>

Submitted on 3 May 2016

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.



Bulletin du centre d'études médiévales d'Auxerre | BUCEMA

Hors-série n° 9 (2016)
Géolocalisation et sources anciennes ?

Samuel Leturcq et Romain Raveaux

Les graphes pour étudier les dynamiques spatiales à partir des séries fiscales médiévales et modernes. Etat des lieux de l'expérience Modelespace

Avertissement

Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.

Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.

Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

revues.org

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

Référence électronique

Samuel Leturcq et Romain Raveaux, « Les graphes pour étudier les dynamiques spatiales à partir des séries fiscales médiévales et modernes. Etat des lieux de l'expérience Modelespace », *Bulletin du centre d'études médiévales d'Auxerre | BUCEMA* [En ligne], Hors-série n° 9 | 2016, mis en ligne le 17 décembre 2015, consulté le 03 mai 2016.
URL : <http://cem.revues.org/13805> ; DOI : 10.4000/cem.13805

Éditeur : Centre d'études médiévales Saint-Germain d'Auxerre
<http://cem.revues.org>
<http://www.revues.org>

Document accessible en ligne sur :
<http://cem.revues.org/13805>

Document généré automatiquement le 03 mai 2016. La pagination ne correspond pas à la pagination de l'édition papier.

Les contenus du Bulletin du centre d'études médiévales d'Auxerre (BUCEMA) sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

Samuel Leturcq et Romain Raveaux

Les graphes pour étudier les dynamiques spatiales à partir des séries fiscales médiévales et modernes. Etat des lieux de l'expérience Modelespace

D'où venons-nous, où allons-nous ?

- 1 Quel historien, travaillant sur des terriers, des cadastres, des compoix..., c'est-à-dire cette documentation comptable foncière privée (pour les terriers) ou publique (cadastres, compoix...) qui foisonne dans les fonds d'archives, n'a pas eu envie de se lancer dans une restitution du paysage ancien au travers des déclarations, parcelle par parcelle, qui y sont consignées ? Dès les XIIIe-XIVe siècle émergent en effet des descriptions très fines, à la parcelle, des possessions foncières. Ces descriptions sont faites sur la base d'enquêtes menées soit au profit d'un seigneur désireux d'assurer la rentrée de ses revenus fonciers (terriers), soit au profit du pouvoir royal dans le cadre de la levée du fouage ou de taille (compoix, cadastres), impôts prélevés dans la partie méridionale du royaume de France sur une base réelle, c'est-à-dire sur la base de la terre possédée. Les contribuables sont contraints de déclarer, devant un agent public, les parcelles qu'ils détiennent, en précisant communément :
 - leur identité (prénom, nom, et parfois surnom, accompagné parfois de formules désignant leur dignité particulière)
 - leur activité
 - leur lieu de résidence
 - la superficie de la parcelle
 - la nature de leur usage (bâti, jardin, verger, vigne, terre labourable, bois, pré...)
 - le lieu dans le territoire (microtoponyme)
 - les confronts (au nombre de 4 le plus souvent), très souvent orientés
 - la somme qui doit être versée, appelée « allivrement » dans les compoix.
- 2 La standardisation de l'enregistrement de ces longues séries (fréquemment plusieurs centaines à plusieurs milliers d'items) rend particulièrement propice le travail statistique, puisqu'il est possible de reporter les attributs de chaque parcelle dans une matrice. Mais surtout, dans la mesure où l'ensemble de l'information est spatialisée au moyen d'indices de localisation que sont les microtoponymes et les confronts, il est théoriquement envisageable de replacer les parcelles les unes par rapport aux autres au sein d'un territoire, en s'aidant du plan-terrier s'il existe (pour les terriers et compoix de la période moderne), sinon du cadastre napoléonien grâce à la microtoponymie, par travail régressif minutieux.
- 3 La dimension spatiale des terriers et compoix est reconnue depuis très longtemps, notamment dans l'historiographie germanique. Marc Bloch soulignait, dès 1929, dans le cadre d'une enquête sur les plans-terriers pour la toute nouvelle revue des *Annales*, la richesse potentielle de cette documentation :

« Car les plans parcellaires, comme tous les documents, ne demeurent monotones et exsangues que jusqu'au jour où le coup de baguette de l'intuition historique leur a rendu une âme. En leurs traits figés, une vie mouvante, pleine de travaux et d'aventures, s'est inscrite et se révèle, toute chaude, à qui a l'art de les saisir : la vie rurale dans ses péripéties et l'infini de ses variétés régionales... ».
- 4 Et Marc Bloch précise plus loin :

« ...le plan parcellaire se place au début et à la fin de l'étude agraire : au début comme instrument d'investigation, un des plus pratiques et des plus sûrs qui soient ; à la fin – une fois bien connue et bien comprise la petite société dont le terroir est la carapace – comme la plus immédiatement sensible des réalités sociales profondes »¹.

- 5 Marc Bloch parle d'« intuition », de « coup de baguette »... et de fait lui-même utilise abondamment les terriers modernes, et notamment leurs plans. Mais il faut bien avouer que c'est de manière très impressionniste, piochant de-ci de-là les informations utiles à ses travaux sur les campagnes françaises, notamment dans le cadre du cycle de conférences qu'il donne à Oslo dans le courant de l'année 1929, prélude aux *Caractères originaux de l'Histoire rurale française*.
- 6 Le coup de baguette de l'intuition, en France, ce sont d'abord les géographes, naturellement sensibles aux thématiques spatiales, qui s'en emparent dans le cadre des travaux régionaux de l'avant et après seconde guerre mondiale. C'est par exemple Roger Dion (le Val de Loire), Etienne Juillard (la plaine d'Alsace), Max Derruau (la plaine de la Limagne)²... Mais les historiens ruralistes ne sont pas en reste, se lançant dans ces approches dès les années 1950-1960, à l'instar de Pierre de Saint-Jacob sur la Bourgogne du Nord³. Mais il faut bien noter que les approches spatiales de ces travaux historiques sont en réalité limitées ; ainsi, la thèse de Pierre de Saint-Jacob ne comporte-t-elle aucune carte, alors que les terriers et plans-terriers ont été largement mis à contribution dans son étude. En réalité, le foisonnement de l'information et l'ampleur du travail nécessaire pour traiter des séries qui comportent couramment plusieurs milliers d'items pour un seul document ont souvent découragé les meilleures volontés, jusqu'à ce que l'outil informatique offre aux chercheurs non seulement les outils performants de cartographie, mais bien plus la possibilité d'automatiser le traitement des données. Dès les années 1970 débutent les premières études systématiques à l'aide de bases de données informatisées⁴. Cependant, jusqu'à ces dix dernières années, les systèmes de base de données utilisés par les historiens ne permettaient pas d'intégrer efficacement la dimension spatiale ; ils utilisaient des outils de traitement statistique des séries qu'offrent les sources fiscales médiévales et modernes, mais ils butaient sur le traitement topologique de ces séries. De fait, durant les années 1970, 1980 et encore 1990, de nombreux historiens s'attaquent courageusement au problème en travaillant de manière « artisanale ». Les historiens modernistes ont su utiliser les plans accompagnant les terriers des XVIIe et XVIIIe siècles, en dressant des atlas thématiques par le report manuel des informations consignées dans les déclarations sur le plan parcellaire, à l'instar de Bernard Bodinier pour analyser la structure de seigneuries dans le département de l'Eure⁵. En absence de plan parcellaire accompagnant les documents (ce qui est le cas presque systématiquement avant le XVIIe siècle), le chercheur est obligé de se lancer dans la reconstitution d'un puzzle en utilisant les précisions topographiques, microtoponymiques, et surtout les mentions de confronts, soit une reconstitution « à la main » du parcellaire, comme le montre par exemple le travail d'Elvis Mallorqui pour analyser les structures d'un village géronais d'un capbreu de 1319⁶. Ce travail extrêmement long et difficile (car il existe des lacunes, des informations manquantes, des ambiguïtés...) est envisageable pour reconstituer le parcellaire d'une zone restreinte, mais en aucun cas l'ensemble d'un terroir. En outre, le chercheur aboutit finalement à des cartes certes intéressantes, mais toujours inertes. En effet, les croisements plus complexes d'informations sont possibles théoriquement, mais impossibles dans la pratique, en absence d'outil qui permette d'automatiser les requêtes et qui en fournisse une interprétation spatiale. L'émergence des Systèmes d'information géographique a offert de nouvelles possibilités dès lors qu'il s'est agi de lier un plan parcellaire vectorisé à une matrice élaborée pour enregistrer les attributs de chaque parcelle⁷ ; l'analyse des données spatiales s'en est trouvée fortement enrichie, permettant des approches anthropologiques novatrices. Pourtant, l'usage des SIG ne lève pas l'intégralité des verrous. D'une part ils ne sont d'aucune utilité pour le traitement des compoix ou terriers qui ne sont pas accompagnés de plans parcellaires ; le chercheur reste contraint d'utiliser le remontage manuel. D'autre part, dans le cas d'une série de compoix et terriers se succédant pour un même territoire, il s'avère très difficile de relier les tables issues de chaque document pour analyser les dynamiques de l'occupation du sol. Il existe là un défi technique et méthodologique que le programme ANR Modelespace a tenté de relever.

Le programme Modelespace : utiliser la théorie des graphes pour exploiter les données spatiales des compoix et terriers

- 7 Au début des années 1980, des historiens ont imaginé un outil de traitement à la fois statistique et topologique des données issues des compoix et terriers⁸. Ainsi débutent des expériences cherchant à automatiser le traitement de ces informations spatiales. Dès 1985, Georges Montpied et Jacques Rouault, travaillant sur deux cadastres de la petite ville d'Embrun dans les Alpes à la fin du Moyen Age (1461), cherchent à mettre en place un système de traitement automatique des informations spatiales des cadastres modernes afin d'aboutir à un système de cartographie automatique à l'aide de la théorie des graphes⁹. Mais ils se heurtent à l'époque à deux obstacles majeurs qui ne permettent pas de faire aboutir le projet : d'une part un obstacle mathématique (l'application topologique de la théorie des grands graphes pour le traitement de l'information spatialisée des sources fiscales médiévales et modernes n'est pas encore connue), d'autre part un obstacle technologique (les SIG, qui permettraient la mise en œuvre pratique de cette application topologique des graphes, ne sont pas encore connus et utilisés).
- 8 Reprenant cette idée d'utiliser le concept de graphe mathématique¹⁰, un programme de recherche a été lancé en 2010, soutenu pendant 3 ans par l'Agence nationale de la Recherche (ANR), portant le nom de MODELESPACE¹¹. Les objectifs de ce programme sont d'aller au-delà des reconstructions manuelles des plans parcellaires en générant un outil qui permette de mener, à terme, d'ambitieux projets d'analyse des dynamiques spatiales de territoires urbains ou ruraux pour lesquels on dispose de plusieurs terriers, compoix et cadastres qui sont autant d'enregistrements successifs de l'organisation du territoire. C'est donc à un changement d'échelle qu'invite cet outil de modélisation des dynamiques spatiales, puisqu'il doit permettre d'intégrer dans une même base de données les états successifs d'un même espace, d'offrir une modélisation graphique de chacun de ces états, et de mesurer les transformations spatiales par comparaison de ces différentes représentations, en faisant sauter l'obstacle de l'hétérogénéité documentaire (notamment l'opposition plan/registre). L'objectif n'est donc pas de créer un outil de reconstitution automatique d'un parcellaire au plus proche de la réalité passée ; ce serait un objectif non seulement peu réaliste, mais encore fondé sur une conception statique des territoires et des paysages. Il s'agit plutôt de créer, à partir de l'information récupérée dans les descriptions de confronts de chaque parcelle, la matrice d'adjacence des parcellaires de chaque époque renseignée par les compoix et terriers successifs d'un lieu, afin de pouvoir réduire l'information spatiale contenue dans chaque document à l'état de graphes comparables les uns avec les autres (fig. 1 et 2).

Fig. 1. Un graphe issu d'un plan terrier

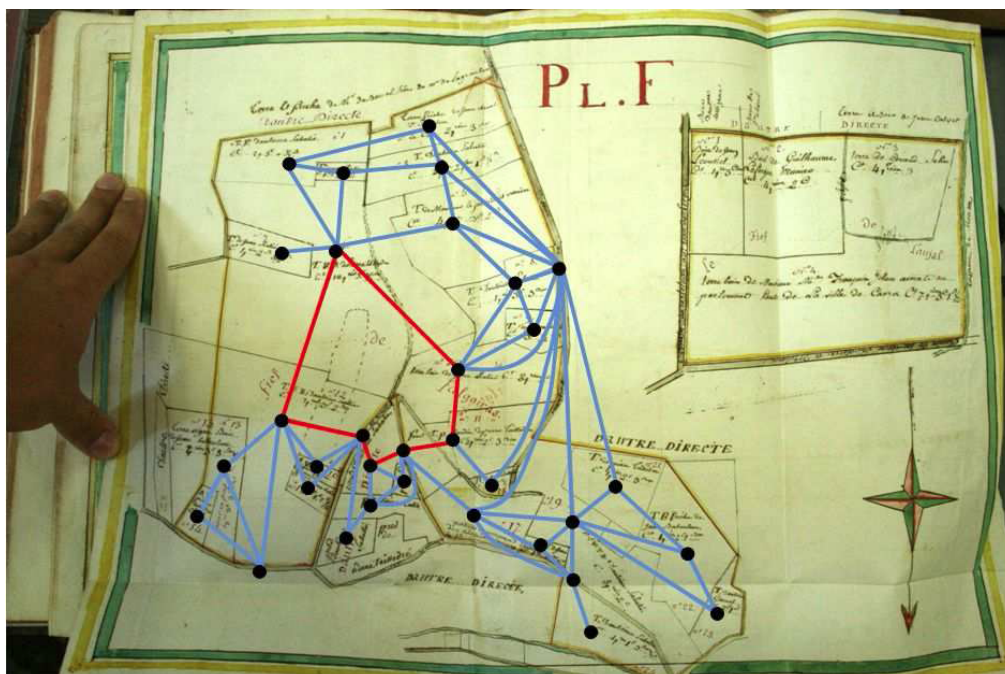
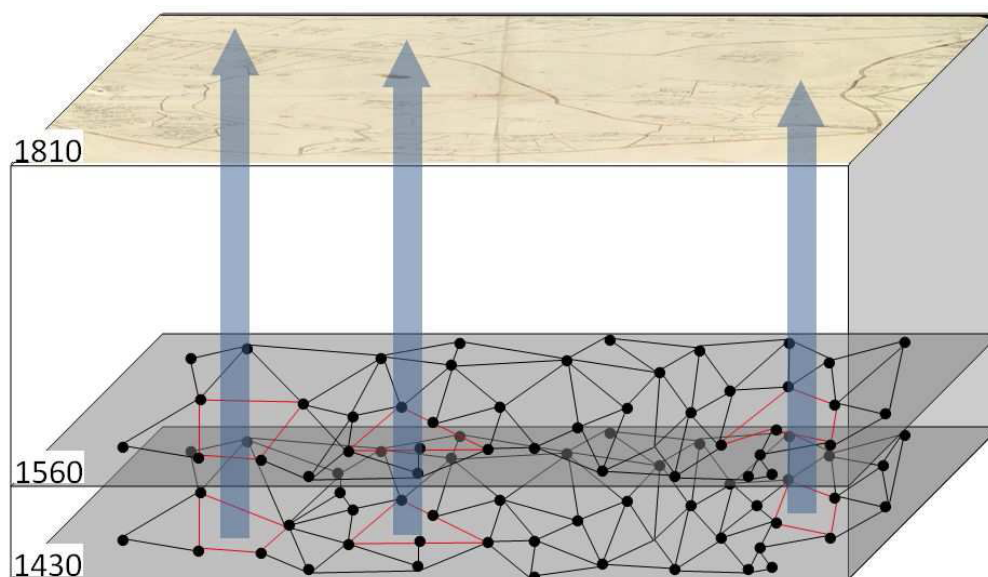


Fig. 2. Un empilement de graphes, états successifs d'un territoire



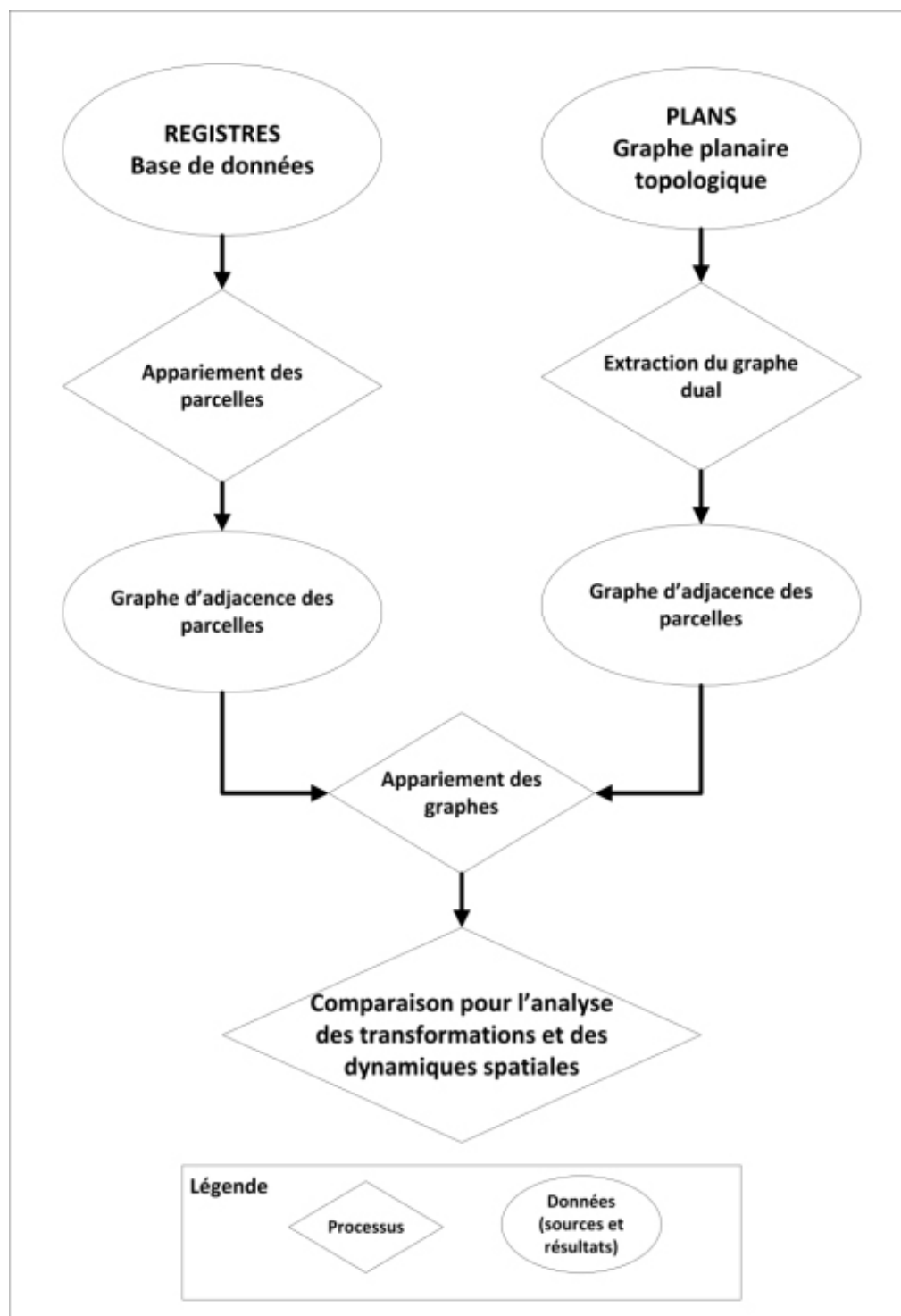
- 9 Pour mener à bien l'expérimentation de la méthode, le choix s'est porté sur la commune d'Odars, installée dans la grande couronne de Toulouse. On dispose en effet, pour cette commune, d'un dossier particulièrement fourni de compoix et de terriers, avec et sans plan, permettant de travailler sur les méthodes de mesure des dynamiques spatiales ; tous ces documents furent dépouillés pour saisir l'intégralité des attributs des parcelles. Trois compoix se succèdent en 1476¹², 1497¹³, 1551¹⁴ et 1598¹⁵, puis nous disposons d'un terrier avec son plan dressé en 1759¹⁶, et enfin du cadastre dit « napoléonien » réalisé en 1811. Les méthodes et les résultats de ce programme ont été publiés dans plusieurs articles¹⁷, de sorte que nous ne nous attarderons pas sur les résultats de l'ANR Modelespace, pour développer une présentation de quelques pistes actuellement suivies pour résoudre la question de la comparaison de graphes. Par ailleurs, Florent Hautefeuille, porteur de l'ANR Modelespace, dans un article intitulé « Géolocalisation des sources fiscales pré-révolutionnaires : au-delà de la quadrature du cercle » et présenté dans le présent dossier du BUCEMA, complète cette réflexion en développant d'autres axes de prolongement de Modelespace.

Les acquis du programme MODELESPACE

La construction des graphes

10 La démarche consiste à transformer l'information spatiale, issue d'une part des registres, d'autre part des plans parcellaires anciens, en graphes (appelés « graphes duals ») dans lesquels chaque parcelle correspond à un sommet et chaque relation de voisinage (adjacence visible sur un plan ou décrite dans les documents par la mention des confronts) correspond à une arête. Suivant le type de source, cartographiée ou non, le travail est mené selon deux processus différents (fig. 3).

Fig. 3. Processus de travail



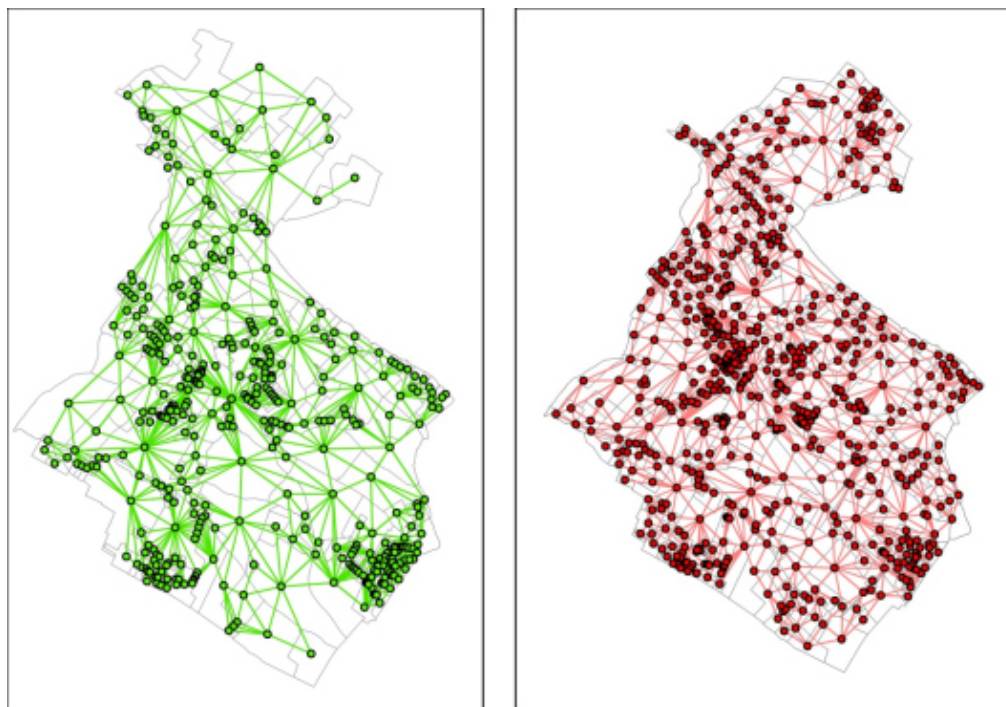
- 11 Le premier processus a consisté à extraire le graphe dual¹⁸ des plans parcellaires anciens dans un Système d'Information Géographique (SIG). À partir des plans originaux, il s'agit d'abord de vectoriser le plan selon une structure topologique de manière à obtenir le graphe planaire topologique du parcellaire¹⁹. Les sommets de ce graphe sont les points d'intersection des frontières des parcelles et les arcs sont les frontières des parcelles. De ce graphe, on extrait le graphe dual dans lequel chaque parcelle est représentée par un sommet et deux parcelles voisines sont reliées par un arc. Dans ce graphe dual, on affecte les attributs des parcelles aux sommets qui les représentent afin d'obtenir ce que nous appelons le graphe d'adjacence des parcelles (fig. 1).
- 12 Le second processus a consisté à concevoir un outil de saisie et d'analyse des données spatiales des registres fiscaux (sans plan parcellaire associé) permettant la construction de graphes. Un outil informatique a été développé par Florent Hautefeuille au sein de l'UMR 5136 FRAMESPA (Toulouse) dans l'optique d'automatiser ce processus pour des registres fiscaux dépourvus de plans²⁰. L'outil, dénommé TERCOMP, est basé sur le logiciel de base de donnée Filemaker. Il est organisé en deux volets distincts : la saisie et le formatage des données d'une part, leur traitement de l'autre.
- 13 La partie saisie propose une interface qui permet d'acquérir l'essentiel des informations des compoix et terriers. Elle découle des expériences acquises depuis les premières tentatives du colloque de 1985²¹ sur le traitement informatique des cadastres anciens et d'une analyse fine de la structure même de ces documents. La base est découpée en six tables principales :
- une table de données générales dans laquelle on enregistre les identifiants de la source et des caractéristiques communes à l'ensemble du registre, comme les unités de surfaces, ou le mode de désignation des confronts.
 - une table « propriétaire » qu'il est nécessaire de remplir en amont des autres pour limiter au maximum les risques de doublonnage et de confusion.
 - une table « parcelle fiscale » dans laquelle sont enregistrées les données attachées à l'unité géographique de base décrite dans le registre et associée aux confronts.
 - une table « parcelle réelle » qui peut parfois recouvrir exactement une parcelle fiscale, mais qui peut également constituer une subdivision de la parcelle fiscale dépourvue d'éléments de localisation propre. C'est le cas d'un mas décrit comme une tenure-bloc, mais qui est constitué d'une maison, de terres et de différentes autres parcelles. Contrairement au cadastre napoléonien où la parcelle est définie par l'unité de propriété et de nature, les compoix et terriers retiennent avant tout l'unité de propriété.
 - une table « confront ». Celle-ci distingue clairement les confronts de personnes qui renvoient à d'autres parcelles décrites par ailleurs dans le registre, et les autres confronts, appelés « invariants » qui désignent les chemins, rivières, bâtiments publics ou seigneuriaux, cimetières... Les invariants sont généralement des repères stables dans le paysage et jouent un rôle essentiel dans la phase de modélisation des données.
 - une table « entité toponymique » qui contient l'ensemble des informations d'ordre toponymique permettant de localiser un bien. Il peut s'agir d'un lieu-dit, du nom d'une parcelle ou de l'appartenance à une paroisse dans le cas de juridictions à plusieurs paroisses.
- 14 Outre cette architecture, le logiciel propose de multiples outils pour faciliter la saisie et limiter les erreurs (liste dans des menus déroulants, messages d'erreur associés à des incohérences). TERCOMP permet de saisir dans la même base de données l'ensemble des registres fiscaux d'une même commune (terriers et compoix), mais aussi l'état des sections du cadastre napoléonien correspondant.
- 15 La seconde partie du logiciel a pour finalité la modélisation automatique de l'espace géographique décrit. Elle repose sur un principe d'appariement des parcelles deux à deux. Lorsqu'une parcelle d'un propriétaire P1 a comme confront un propriétaire P2, le logiciel, par le biais d'une série d'algorithmes, identifie l'éventuelle parcelle du propriétaire P2 qui correspond au confront recherché. Elle génère alors une fiche dans une table d'appariement qui relie les deux parcelles. Cette série d'algorithmes permet d'identifier une part importante des

- appariements de parcelles. Par ailleurs TERCOMP isole également les confronts qui n'ont pas de correspondance dans une autre parcelle, ou ceux qui au contraire, en ont plus d'une. Pour améliorer le taux d'identification, des algorithmes de contournement ont été développés. Ils s'appuient sur des informations connexes du type entité toponymique, ou sur des proximités de voisins. Dans le cas d'un de nos registres test, le pergeat de la commune d'Odars de 1598²², le taux de résolution final dépasse légèrement les 75 %. Le registre décrit 632 parcelles fiscales pour 140 propriétaires. Les confronts nominatifs permettent d'envisager théoriquement 1233 appariements. 928 ont été résolus automatiquement. Les vérifications manuelles montrent que le taux d'erreur est très faible. Grâce à ce système, il est possible de relier en un graphe unique 620 des 632 parcelles (soit 98,1 %). Ce système est adapté non seulement aux documents dotés de confronts orientés, mais aussi aux registres (généralement plus anciens) où les confronts ne sont pas orientés. Le taux de résolution automatique est dans ce dernier cas moindre (fig. 5).
- 16 Une fois ce tableau d'appariements établi, l'outil génère un fichier spécifique lisible par un autre logiciel de visualisation de graphe : YED. Dans ce graphe, apparaissent sous la forme de nœuds toutes les parcelles fiscales décrites dans un registre. Tous les confronts « invariants » de type chemin, ruisseau ou juridiction voisine sont également modélisés sous la forme d'un nœud. Lorsque les coordonnées de ces invariants sont connues, elles sont prises en compte dans le graphe et constituent des ancrages du graphe dans l'espace géographique (ces ancrages ont reçu le nom de « points d'amer »). Les arêtes du graphe sont définies par les voisinages de parcelles désambiguïsées. L'objet mathématique ainsi obtenu est désormais comparable à ceux qui sont issus de la modélisation d'autres registres, mais aussi des plans cadastraux.

L'ajustement spatial des graphes

- 17 Chaque état de la documentation trouve sa traduction en un graphe. La comparaison des graphes, à partir de laquelle on peut envisager de mesurer les dynamiques spatiales, nécessite toutefois de parvenir à ajuster spatialement ces graphes. Les graphes issus du plan cadastral napoléonien (1811) et du plan-terrier de 1759 (fig. 4) ne posent aucun problème d'ajustement, dans la mesure où le support cartographique de l'information autorise une géolocalisation très précise de chaque sommet des deux graphes.

Fig. 4. Graphes issus des registres fonciers d'Odars, 1759 (gauche) et 1811 (droite)



- 18 Il n'en va pas de même pour les graphes construits à partir de l'information récoltée dans les registres qui ne sont accompagnés d'aucun plan, à l'instar du compoix d'Odars de 1598

(fig. 5) ; ces graphes doivent impérativement trouver un ancrage spatial, sans lequel ils sont difficilement exploitables.

Fig. 5. Graphe aspatial (sans calage géographique) issu du compoix de 1598



19 L'ancrage passe par la définition de « points d'amer », à savoir des éléments de géolocalisation invariants pour l'ensemble ou partie de la période concernée par la documentation mise en œuvre. Ces points d'ancrage peuvent être des éléments :

- ponctuels : église, chapelle, cimetière, château, carrefour ou maison remarquable. Cette première catégorie de points d'amer est très hétérogène et ne peut venir qu'en complément des deux suivantes
- linéaires : cours d'eau ou route remarquable dont on est certain de la stabilité. Cette catégorie, essentiellement constituée de linéaires, offre l'avantage de couvrir de manière relativement homogène l'ensemble de l'espace étudié. Il présente cependant deux

compte de l'orientation des confronts dans la visualisation du graphe, elles ne permettront pas d'aboutir à une réelle spatialisation des données. L'objectif n'est pas là. Il est dans un premier temps d'arriver à générer des graphes d'aux des parcelles à partir de sources aussi différentes que des registres et des plans. Cet objectif est aujourd'hui atteint ; il reste à développer les outils de comparaison des graphes ainsi obtenus.

La comparaison des graphes : des pistes en cours d'expérimentation

- 21 Les trois années n'ont malheureusement pas suffi pour faire aboutir le projet initial ; la question de la comparaison des graphes n'a pu être qu'effleurée, au travers d'une analyse des mécanismes de transformation spatiale du parcellaire. La réflexion continue toutefois, avec des expériences qui sont lancées dans le cadre de projets de fin d'étude (PFE) d'étudiants de l'école polytechnique de l'université de Tours encadrés par Romain Raveaux, chercheur au Laboratoire d'informatique-LI (EA 6300). Deux pistes particulières sont explorées : d'une part la simplification des grands graphes par la recherche de communautés, d'autre part l'analyse de distorsions entre l'édition de deux graphes par le calcul des coûts pour passer de l'un à l'autre.

Simplifier les graphes : une segmentation des grands graphes en petites communautés.

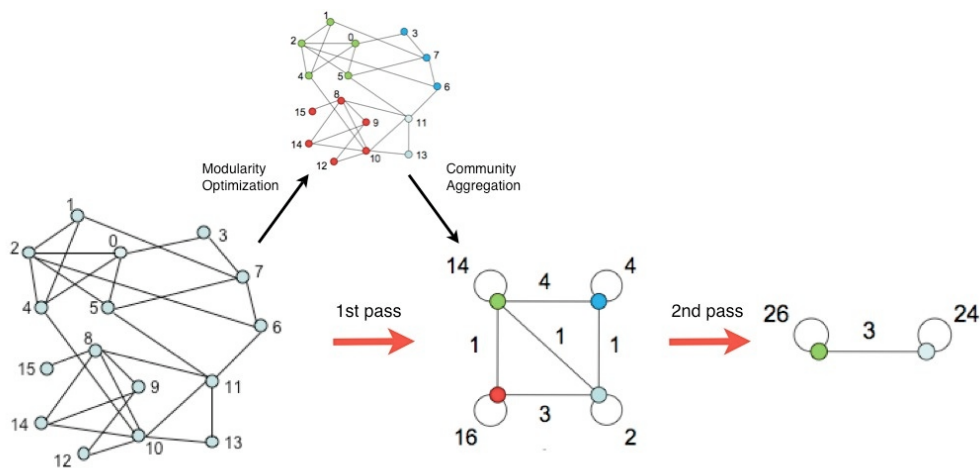
- 22 L'un des verrous techniques majeurs tient dans la complexité des calculs à réaliser pour analyser les transformations d'un état de graphe à un autre. De fait, une expérience menée sur deux graphes d'Odars a montré qu'il fallait 6 heures pour calculer les transformations entre deux graphes de 769 et 497 nœuds, soit des graphes de relativement petite taille, comparativement à des documents qui concernent couramment des milliers de parcelles. Il faut impérativement travailler à réduire la durée des temps de calcul²⁴. La piste suivie est celle de la simplification des graphes, par une segmentation des grands graphes en sous-graphes ; réduire la taille des graphes entraîne une diminution des temps de calculs. Une étude préliminaire a été réalisée dans le courant du printemps 2014 par Léa Lehnebach, étudiante à l'école polytechnique de Tours, dans le cadre d'un mémoire de projet de fin d'étude pour l'obtention d'un diplôme d'ingénieur informaticienne, sous la direction de Romain Raveaux²⁵. Léa Lehnebach a travaillé sur le partitionnement des grands graphes en sous-graphes, en utilisant la méthode dite « de Louvain », à savoir un algorithme de détection de communautés reposant sur l'amélioration de la modularité du graphe en regroupant les nœuds en communautés²⁶. La méthode consiste à isoler les blocs parcellaires qui entretiennent la meilleure connectivité, c'est-à-dire regrouper et isoler l'ensemble des nœuds qui sont fortement connectés. Le partitionnement est optimisé selon le critère de la modularité, c'est-à-dire une mesure, issue d'une formule de modularité (fig. 7), comprise entre -1 et 1 qui caractérise la « qualité » du partitionnement, à savoir la proportion entre arêtes reliant des nœuds de même communauté et arêtes reliant des nœuds appartenant à des communautés différentes. Ainsi, plus la modularité est élevée, plus le graphe présente une structure en communautés significative (le seuil de 0,3 est donné pour caractériser un graphe à structure en communautés significative).

Fig. 7. Formule de modularité

$$Q = \sum_C \frac{1}{2m} \sum_{i,j \in C} \left(A_{ij} - \frac{k_i \cdot k_j}{2m} \right)$$

- 23 A_{ij} : Valeur de la matrice d'adjacence entre les sommets i et j
 k_i : Somme des poids des arêtes d'adjacentes à i
 m : Nombre d'arêtes du graphe
- 24 L'algorithme consiste à parcourir successivement chaque nœud du graphe et à essayer de rapprocher ce nœud de la communauté, parmi celles de ses voisins, pour laquelle le gain de modularité est strictement positif et maximal (fig. 8). A chaque itération, chaque nœud tente de changer de communauté en fonction du calcul du gain de modularité. Le gain de modularité est la variation de modularité constatée lorsqu'on « enlève » le nœud considéré de sa communauté et qu'on le place dans la communauté de son voisin. Si le gain de modularité est négatif, aucun changement n'intervient. S'il est positif, le graphe est modifié. D'itération en itération, on recherche un équilibre correspondant au meilleur partitionnement, indiqué par la maximisation de la modularité du graphe. Lorsque plus aucun nœud ne peut être « déplacé » d'une communauté à une autre, si la modularité du graphe a augmenté, alors on est parvenu à réduire la taille de ce graphe. En effet le nouveau graphe a pour nœuds les communautés détectées, donc un nombre réduit de nœuds.

Fig. 8. Processus d'optimisation de la modularité du graphe



- 25 C'est selon cette méthode que les graphes des parcelles d'Odars issus du cadastre napoléonien (1811) et du terrier de 1759 ont été traités (fig. 9), permettant de simplifier de manière importante les graphes, puisque pour 1759, le graphe passe de 497 à 85 nœuds, tandis que celui du parcellaire de 1811 passe de 769 à 128 nœuds.

détectées ; cette variation peut être due à la caractéristique non déterministe de l'algorithme. En effet, le parcours des nœuds dans un sens ou dans un autre peut impliquer que les nœuds considérés à un certain stade du parcours basculent d'une communauté à une autre. Cette ultime remarque débouche évidemment sur un questionnement en matière d'appariement des graphes : comment comparer des graphes ?

Apparier les graphes : le calcul de la distance d'édition entre deux graphes.

27 Dans le cadre d'un projet de fin d'étude mené à l'école polytechnique de l'université François-Rabelais de Tours²⁷, Antoine Crèche, encadré par Romain Raveaux (LI), a travaillé sur la méthode de comparaison de deux graphes en utilisant la distance d'édition, c'est-à-dire en calculant les degrés de similitude ou de dissimilitude entre deux graphes. Fondamentalement, il s'agit de comprendre les transformations spatiales dont témoigne l'écart existant entre les graphes G_1 et G_2 de deux états t_1 et t_2 d'un même territoire, en analysant les distorsions entre l'édition de ces deux graphes, sur la base du calcul des coûts pour passer de G_1 à G_2 par le chemin minimal (fig. 11) :

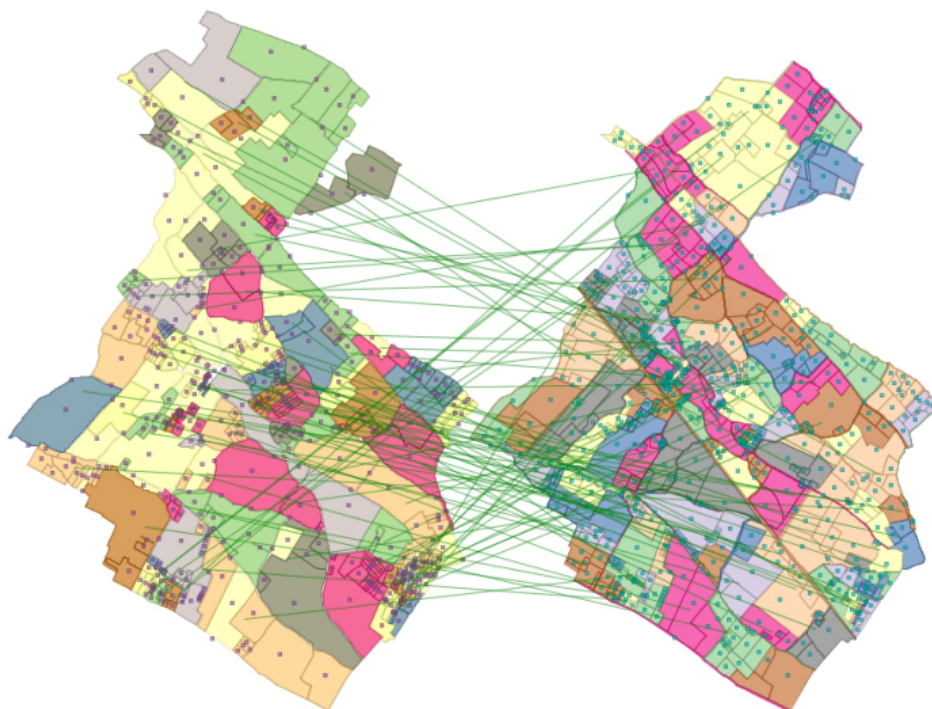
- en supprimant et/ou ajoutant des nœuds et des arcs
- en substituant des nœuds
- en calculant des différences entre les attributs attachés aux nœuds ou aux arcs

Fig. 11. Décomposition du passage de G_1 à G_2 par le chemin minimal



28 La distance d'édition entre G_1 et G_2 (chemins et coûts) est calculée à partir d'une table des distorsions entre les deux graphes. La méthode a été appliquée par Léa Lehnebach sur les graphes issus des parcellaires de 1759 et 1811 d'Odars, après qu'ils aient été partitionnés selon la méthode de Louvain (fig. 12).

Fig. 12. Appariement du graphe-source de 1759 (à gauche) vers le graphe-cible de 1811 (à droite) après leur partitionnement selon la méthode de Louvain



29 Le résultat de ce test demeure confus, ne permettant pas une exploitation scientifique des données. Plusieurs causes peuvent être relevées, qui sont autant de pistes pour améliorer la méthode :

- Les paramètres choisis ne sont pas forcément les plus adaptés : on ne prend en effet pas en compte les arêtes (dans le calcul de la distance d'édition, le coût est de 0 dans tous les cas) afin d'enlever certaines entraves liées à la structure des graphes.
- De tels résultats peuvent en partie s'expliquer par la structure des graphes. Une piste pourrait être de considérer d'autres éléments, comme le poids des arêtes par exemple, ou d'agir directement dès la méthode de Louvain, en apportant des modifications à l'algorithme afin de l'adapter davantage à notre problème.
- L'appariement s'appuie également dans notre situation sur la spatialité des communautés : le barycentre de la communauté n'est peut-être pas le critère déterminant (cas d'une communauté convexe avec un barycentre « hors » de la communauté qui fausserait les résultats).

Conclusion

30 Le programme Modelespace est parvenu à faire sauter des verrous techniques majeurs. Désormais, on dispose d'une méthode de construction de séries de graphes à partir de plans et de matrices issues de cadastres, compoix, terriers... pour des zones possédant des séries documentaires longues. La recherche bute toutefois sur la question de la comparaison de ces graphes, dans la mesure où leur référencement géographique est souvent très imparfait.

31 C'est sur ce point précis de la comparaison de graphes qu'il faut désormais concentrer tous les efforts. Tous les espoirs sont permis, dans la mesure où les informaticiens travaillent couramment sur cette thématique, par exemple pour de la reconnaissance de formes. Il faut donc envisager un Modelespace 2 pour faire aboutir une méthodologie capable d'offrir aux historiens et archéologues la possibilité d'exploiter la documentation fiscale médiévale et moderne dans toute sa richesse informative.

Bibliographie

- C. BERGE, *Théorie des graphes et ses applications*, Paris, Dunod, 1958.
 C. BERGE, *Graphes et Hypergraphes*, Paris, Dunod, 1970.
 J.-L. BIGET et alii (éd.), *Les cadastres anciens des villes et leur traitement par l'informatique*, Rome, Ecole française de Rome, 1989.
 M. BLOCH, « Les plans parcellaires », in *Annales d'histoire économique et sociale*, 1(1929), p. 60-70.
 V. D Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, « Fast unfolding of communities in large networks », in *Journal of Statistical Mechanics: Theory and Experiment*, oct. 2008, P10008 (12 p.).
 B. BODINIER, « Seigneurs et propriétaires dans la France d'Ancien Régime. L'exemple des campagnes de l'Eure », in G. BRUNEL, O. GUYOTJEANNIN, J. M. MORICEAU (dir.), *Terriers et plans-terriers du XIII^e au XVIII^e siècle. Actes du colloque de Paris, 23-25 septembre 1998*, Caen, AHSR, 2002, p. 293-308.
 G. BRUNEL, O. GUYOTJEANNIN, J. M. MORICEAU (dir.), *Terriers et plans-terriers du XIII^e au XVIII^e siècle. Actes du colloque de Paris, 23-25 septembre 1998*, Caen, AHSR, 2002.
 P. BRUNET, *Structure agraire et économie rurale des plateaux tertiaires entre la Seine et l'Oise*, Caen, 1960.
 A. CRÈCHE, *Comparaison de graphes pour l'analyse de cadastres*, Rapport d'analyse finale de projet de fin d'étude, Université François-Rabelais de Tours, Ecole polytechnique, 2013.
 M. DERRUAU, *La grande Limagne auvergnate et bourbonnaise*, Grenoble, 1949.
 R. DION, *Le Val de Loire : étude de géographie régionale*, Tours, 1934.
 F. HAUTEFEUILLE, X. RODIER et alii, « De l'espace aux graphes. Mesurer les dynamiques spatiales des terroirs villageois », in *Mesure et histoire médiévale*. Paris, Publications de la Sorbonne, 2013, p. 99-118.
 A. GUERREAU, « Une méthode de représentation graphique d'un ensemble de parcelles repérées seulement par leurs confronts », in *Le médiéviste et l'ordinateur*, 8 (1982), p. 8-10.
 E. JUILLARD, *La vie rurale dans la plaine de Basse Alsace. Essai de géographie sociale*, Strasbourg, 1992.
 M. LE COUÉDIC et alii, « Du cadastre ancien au graphe. Les dynamiques spatiales dans les sources fiscales médiévales et modernes », in *ArchéoSciences, revue d'archéométrie*, 2012, p. 71-84.
 L. LEHNEBACH, *Analyse de plans cadastraux à l'aide de graphes*, Rapport d'analyse finale de projet de fin d'étude, Université François-Rabelais de Tours, Ecole polytechnique, 2014.

- S. LETURCQ, « A la découverte de la dimension spatiale des terriers... Le SIG, outil d'analyse des terroirs d'exploitation », in *Le médiéviste et l'ordinateur*, 44 (2006)- <http://lemo.irht.cnrs.fr/44/terriers.htm>.
- E. MALLORQUI GARCIA, « Les campagnes de Gérone à travers des capbreus du XIV^e siècle », in G. BRUNEL, O. GUYOTJEANNIN, J. M. MORICEAU (dir.), *Terriers et plans-terriers du XIII^e au XVIII^e siècle. Actes du colloque de Paris, 23-25 septembre 1998*, Caen, AHSR, 2002, p. 343-360.
- G. MONTPIED, J. ROUAULT, « Reconstitution et cartographie d'un parcellaire à partir de cadastres textuels du XV^e siècle », in *Le médiéviste et l'ordinateur*, 8 (1982), p. 11-13.
- G. MONTPIED, J. ROUAULT, « Du texte au graphe : établissement d'une carte du parcellaire à partir des données de deux cadastres de la fin du Moyen Age », in J.-L. BIGET et alii (éd.), *Les cadastres anciens des villes et leur traitement par l'informatique*, Rome, Ecole française de Rome, 1989, p. 359-380.
- X. Rodier et alii, « From space to graphs to understand spatial changes using medieval and modern fiscal sources », in E. Graeme et alii (ed.), *40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*, Mar 2012, Southampton, United Kingdom. Amsterdam University Press, 2013, p. 424-431.
- P. DE SAINT-JACOB, *Les paysans de la Bourgogne du Nord au dernier siècle de l'Ancien Régime*, Caen, AHSR, 1995 (1^{ère} éd. 1960).
- B. VUE, *Microtoponymie et archéologie des paysages à Neuilly-L'évêque (52) du XIII^e au XX^e siècle : comportement des microtoponymes au fil des siècles : le nom, l'espace et l'homme qui le nomme*, Université de Nancy 2, Thèse de doctorat d'Histoire, 1997.
- M. ZERNER, *Le cadastre, le pouvoir et la terre. Le Comtat Venaissin au début du XV^e siècle*, Paris, Ecole française de Rome, 1993.

Notes

- 1 M. BLOCH, « Les plans parcellaires », in *Annales d'histoire économique et sociale*, 1(1929), p. 61-62.
- 2 R. DION, *Le Val de Loire : étude de géographie régionale*, Tours, 1934 ; P. BRUNET, *Structure agraire et économie rurale des plateaux tertiaires entre la Seine et l'Oise*, Caen, 1960 ; E. JUILLARD, *La vie rurale dans la plaine de Basse Alsace. Essai de géographie sociale*, Strasbourg, 1992 ; M. DERRUAU, *La grande Limagne auvergnate et bourbonnaise*, Grenoble, 1949...
- 3 P. DE SAINT-JACOB, *Les paysans de la Bourgogne du Nord au dernier siècle de l'Ancien Régime*, Caen, AHSR, 1995 (1^{ère} éd. 1960).
- 4 Cf. exposé méthodologique de Monique ZERNE, dans sa thèse intitulée *Le cadastre, le pouvoir et la terre. Le Comtat Venaissin au début du XV^e siècle*, Paris, Ecole française de Rome, 1993 (BEFR, 174).
- 5 B. BODINIER, « Seigneurs et propriétaires dans la France d'Ancien Régime. L'exemple des campagnes de l'Eure », in G. BRUNEL, O. GUYOTJEANNIN, J. M. MORICEAU (dir.), *Terriers et plans-terriers du XIII^e au XVIII^e siècle. Actes du colloque de Paris, 23-25 septembre 1998*, Caen, AHSR, 2002, p. 293-308.
- 6 E. MALLORQUI GARCIA, « Les campagnes de Gérone à travers des capbreus du XIV^e siècle », in *Terriers et plans-terriers...*, p. 349.
- 7 S. LETURCQ, « A la découverte de la dimension spatiale des terriers... Le SIG, outil d'analyse des terroirs d'exploitation », in *Le médiéviste et l'ordinateur*, 44 (2006)- <http://lemo.irht.cnrs.fr/44/terriers.htm>.
- 8 A. GUERREAU, « Une méthode de représentation graphique d'un ensemble de parcelles repérées seulement par leurs confronts », in *Le médiéviste et l'ordinateur*, 8 (1982), p. 8-10. G. MONTPIED, J. ROUAULT, « Reconstitution et cartographie d'un parcellaire à partir de cadastres textuels du XV^e siècle », in *Le médiéviste et l'ordinateur*, 8 (1982), p. 11-13.
- 9 G. MONTPIED, J. ROUAULT, « Du texte au graphe : établissement d'une carte du parcellaire à partir des données de deux cadastres de la fin du Moyen Age », in J.-L. BIGET et alii (éd.), *Les cadastres anciens des villes et leur traitement par l'informatique*, Rome, Ecole française de Rome, 1989, p. 359-380.
- 10 C. BERGE, *Théorie des graphes et ses applications*, Paris, Dunod, 1958.
- 11 ANR-09-BLAN-0322-02 « Modélisation des espaces préindustriels ». Ce projet 2010-2012 a été piloté par Florent Hautefeuille (FRAMESPA), Bertrand Jouve (IMT) et Samuel Leturcq (CITERES-LAT). Le site internet est consultable à cette adresse : <http://modelespace.univ-tlse2.fr/>
- 12 AD 31, AMT II 483-1, Estime et cadastre d'Odars (1476).
- 13 AD 31, AMT II 483-2, Estime et cadastre d'Odars écrit en languedocien (1497).
- 14 AD 31, 2E 4356, Compoix d'Odars (1551).
- 15 AD 31, AMT II 487, Compoix d'Odars (1598).
- 16 AD 31, AMT II 493, Terrier d'Odars, accompagné d'un plan parcellaire (1759).
- 17 M. LE COUÉDIC et alii, « Du cadastre ancien au graphe. Les dynamiques spatiales dans les sources fiscales médiévales et modernes », in *ArchéoSciences, revue d'archéométrie*, 2012, p. 71-84. X. RODIER et alii, « From space to graphs to understand spatial changes using medieval and modern fiscal sources »,

in E. GRAEME *et alii* (ed.), *40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*, Mar 2012, Southampton, United Kingdom. Amsterdam University Press, 2013, p. 424-431. F. HAUTEFEUILLE, X. RODIER *et alii*, « De l'espace aux graphes. Mesurer les dynamiques spatiales des terroirs villageois », in *Mesure et histoire médiévale*. Paris, Publications de la Sorbonne, 2013, p. 99-118.

18 C. BERGE, *Théorie des graphes...*, *op. cit.*, p. 206-217.

19 Un plan parcellaire est un graphe planaire topologique. En effet « Un graphe G est planaire s'il est possible de le représenter sur un plan de sorte que les sommets soient des points distincts et les arcs des courbes simples ne se rencontrant pas en dehors de leurs extrémités. Un graphe planaire topologique est une représentation d'un graphe planaire G sur un plan » (C. BERGE, *Graphes et Hypergraphes*, Paris, Dunod, 1970).

20 Possibilité de consulter la base Tercomp en ligne : <http://modelespace.univ-tlse2.fr/>

21 J.-L. BIGET *et alii*, *Les cadastres anciens...*, *op. cit.*

22 AD 31, AMT II 487, Compoix d'Odars (1598).

23 L'expression « entité toponymique » correspond à une aire diffuse au sein de laquelle se concentre un microtoponyme. Il est possible, pour chaque phase, de dresser une liste de ces aires toponymiques, mettant en évidence des apparitions, des disparitions, des élargissements, des rétrécissements et des déplacements (cf. B. VUE, *Microtoponymie et archéologie des paysages à Neuilly-L'évêque (52) du XIII^e au XX^e siècle : comportement des microtoponymes au fil des siècles : le nom, l'espace et l'homme qui le nomme*, Université de Nancy 2, Thèse de doctorat d'Histoire, 1997). Ces entités toponymiques jouent un rôle essentiel dans la construction des graphes issus des registres sans plan, dans la phase de résolution des ambiguïtés.

24 La distance d'édition de graphes est un problème d'optimisation appartenant à la classe de complexité NP. Un problème NP est un problème dont la recherche d'une solution demande *a priori* beaucoup de temps (complexité exponentielle en fonction du nombre de noeuds des graphes), tandis que la vérification que la solution est correcte ne demande pas beaucoup de temps.

25 L. LEHNEBACH, *Analyse de plans cadastraux à l'aide de graphes*, Rapport d'analyse finale de projet de fin d'étude, Université François-Rabelais de Tours, Ecole polytechnique, 2014.

26 V. D BLONDEL, J.-L. GUILLAUME, R. LAMBIOTTE, E. LEFEBVRE, « Fast unfolding of communities in large networks », in *Journal of Statistical Mechanics : Theory and Experiment*, oct. 2008, P10008 (12 p.).

27 A. CRÈCHE, *Comparaison de graphes pour l'analyse de cadastres*, Rapport d'analyse finale de projet de fin d'étude, Université François-Rabelais de Tours, Ecole polytechnique, 2013.

Pour citer cet article

Référence électronique

Samuel Leturcq et Romain Raveaux, « Les graphes pour étudier les dynamiques spatiales à partir des séries fiscales médiévales et modernes. Etat des lieux de l'expérience Modelespace », *Bulletin du centre d'études médiévales d'Auxerre | BUCEMA* [En ligne], Hors-série n° 9 | 2016, mis en ligne le 17 décembre 2015, consulté le 03 mai 2016. URL : <http://cem.revues.org/13805> ; DOI : 10.4000/cem.13805

À propos des auteurs

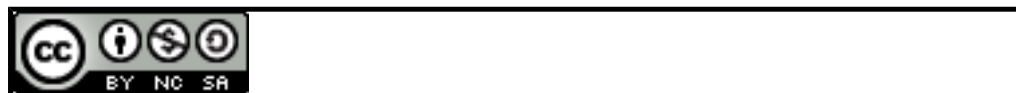
Samuel Leturcq

université François-Rabelais (Tours), UMR 7324 Citeres-LAT

Romain Raveaux

université François-Rabelais (Tours), EA 6300 LI

Droits d'auteur



Les contenus du *Bulletin du centre d'études médiévales d'Auxerre (BUCEMA)* sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

Résumés

De 2010 à 2012, le programme ANR Modelespace a initié une recherche pour mettre au point une méthodologie d'analyse des dynamiques spatiales en partant de la comparaison des jeux de données spatialisées issues des sources fiscales des périodes médiévales et modernes (terriers, compoix, cadastres...) par le biais de la théorie des graphes. Il s'agit de présenter les avancées significatives de la méthodologie, et de définir les pistes envisagées pour établir la méthode de comparaison des graphes.

From 2010 till 2012, the Modelespace program introduced a research to work out a methodology to analyse the spatial dynamics by leaving of the comparison between different sets of spatial data of fiscal documents of medieval and modern periods (fieldbooks and land registry, called *cadastres*, *catasti*, *compoix*, *estimes*... in southern Europe) with the graph theory. We explain the significant advances of the methodology, and we propose different ways to establish the comparison method of the graphs.

Entrées d'index

Index de mots-clés : Terrier, cadastre, Moyen Age, Période moderne, paysage, méthodologie, graphe, dynamique spatiale

Index by keyword : Fieldbooks, Land registry, Medieval Ages, Modern period, Landscape, Methodology, Graphs, Spatial dynamics