

Modélisation des hiérarchies territoriales multiples

Vers la gestion d'informations spatio-temporelles évolutives

Christine Plumejeaud*, Jérôme Gensel*, Marlène Villanova-Oliver*, Maher Ben Rebah, Guillaume Vergnaud*****

* UMR 5217 Laboratoire d'Informatique de Grenoble,
681 rue de la passerelle,
38402 Saint Martin d'Hères
<mailto:{prenom.nom}@imag.fr>

** UMR 8504 Géographie-cités, 13 rue du four, 75006 Paris
Post Doc. UMS 2414 RIATE, Université Paris Diderot Paris 7
maher.benrebah@ums-riate.fr

*** UMR 5600 Environnement Ville Société, Équipe Bio-Géophile, ENS-LSH, 15
parvis René Descartes, 69007 Lyon
EHEHI, Casa de Velázquez, Calle Paul Guinard, 3, 28040 Madrid (Espagne)
guillaumevergnaud@hotmail.com

RÉSUMÉ. L'information statistique disponible aujourd'hui sur tous les espaces géographiques est un potentiel de richesse encore largement inexploité du fait de l'hétérogénéité et de l'évolutivité des sources d'information. L'objectif de cet article est de proposer un modèle de données permettant de gérer des données issues de sources variées, sur une longue période de temps, et qui soit un support efficace pour l'estimation, l'alignement ou la combinaison d'indicateurs statistiques. Ce modèle se focalise sur la gestion de maillages territoriaux de natures multiples, s'organisant dans des hiérarchies territoriales évolutives. Le cœur du modèle est centré sur l'identité des unités géographiques, qui sont les composantes élémentaires des maillages. Les événements du changement territorial et leurs conséquences sur les attributs d'une unité géographique sont décrits de manière à capturer la connaissance et l'expertise humaine sur ces événements.

MOTS-CLÉS : maillage – hiérarchie - agrégation – changement – hétérogénéité - modèle.

ABSTRACT. The diversity and the quantity of statistical information that is nowadays available on every territory, at every scales, is a great potential of knowledge. Yet it is still badly exploited, because of the heterogeneity of the data sources and their evolutivity through time. This paper proposes a model that can handle data issued from various nomenclatures, on a long term, and that could be an efficient support for estimates and combination of statistical indicators. The model is centered on the notion of identity, life and motion of geographical units, that are organised into multiple hierarchical structures. Moreover, change events and their effects on geographic units are described, in a manner that let experts record their knowledge into the database.

KEY WORDS: mesh – hierarchy - agregation – indicator – change - heterogeneity - model.

1. Introduction

L'actualité économique et sociale montre l'importance d'un aménagement du territoire à long-terme pour faire face aux défis de la mondialisation, de la décroissance démographique, de la raréfaction des ressources énergétiques, de la dégradation de l'environnement, etc. Ces faits sont mesurés à partir de statistiques qui rendent compte de façon discrète (en certains lieux, à certaines dates) de phénomènes en évolution constante. Il serait souhaitable de constituer un réservoir de données statistiques permettant d'étudier ces phénomènes à toutes les échelles géographiques, sur un vaste territoire, comme, par exemple, l'espace européen, afin de pouvoir produire des scénarios prédictifs et d'anticiper au mieux le changement.

Cependant, la difficulté réside dans l'hétérogénéité et l'évolutivité des données. Hétérogénéité, car les données sont collectées par des instituts différents, sur des échelles géographiques variables, et de façon très incomplète. Évolutivité, car leur collecte s'effectue sur des territoires qui changent d'emprise spatiale et d'identité, et pour des indicateurs qui changent de sémantique et de qualité. Notre objectif est de proposer une modélisation adéquate de ces données, permettant ultérieurement leur analyse prospective.

La section 2 situe notre modèle par rapport aux précédents travaux [WOR 98], notamment les implémentations nombreuses de SIG temporels [GRE 02], [NOR 03] [PAQ 04], [ALB 07]. Dans la section 3, nous décrivons le système territorial et les relations qui existent entre les unités territoriales. La section 4 présente le modèle sur lequel repose notre Système d'Information Spatio-Temporelle Evolutive (SISTE), avec ses différents composants. La section 5 expose les perspectives que ce modèle générique permet d'entrevoir pour l'estimation spatio-temporelle, et la réalisation d'un suivi spatio-temporel de différents types d'unités géographiques.

2. Modéliser l'information spatio-temporelle : état de l'art

Notre proposition s'appuie sur les résultats des travaux de recherche portant sur la modélisation spatio-temporelle du changement territorial et, plus généralement, sur la représentation du temps dans les SIG. Ces travaux mettent en exergue l'importance que revêt une gestion appropriée du temps, non pas comme un simple attribut de l'espace géographique, mais comme une dimension à part entière dans un système d'information spatio-temporel (SIST). Ainsi, [YUA 99], [KAV 01], et [PEU 02] soulignent qu'un tel système doit supporter la représentation simultanée de trois domaines (ou dimensions) : la dimension spatiale (les unités géographiques), la dimension temporelle (les événements) et la dimension thématique (les valeurs des variables statistiques). Il doit surtout permettre de **pivoter** de la dimension spatiale à la dimension temporelle ou thématique, en montrant par exemple :

- pour une unité géographique donnée, sa courbe d'évolution sur une variable thématique donnée, et sa généalogie (de quelles recompositions territoriales est-elle issue ?)

- pour une date donnée, le découpage territorial de l'espace et les valeurs des attributs thématiques associés.

En effet, l'information statistique peut se modéliser comme un **cube multi-dimensionnel**, avec au moins trois dimensions (voir figure 1) :

- les indicateurs appartiennent à la dimension thématique et s'organisent dans des hiérarchies thématiques,
- les territoires forment des hiérarchies spatiales, et les niveaux de la hiérarchie sont les échelles géographiques, matérialisées par des maillages
- le temps étant la dernière dimension, qui présente différents niveaux de granularité (de la seconde à l'année).

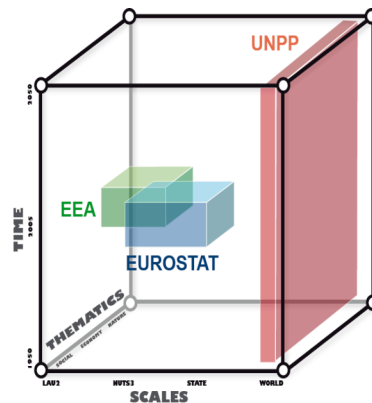


Figure 1. Le cube de l'information statistique.

Le cube multidimensionnel est généralement représenté sous la forme d'une coupe spatiale à un instant précis, avec la cartographie de la variable thématique choisie. Cette vision de couche spatiale date de la création des SIG, et elle a entraîné une vision trop simplifiée des modèles spatio-temporels. Cette première famille de modèles superpose les couches spatiales dans le cube, à plusieurs instants t , [ARM, 88]. Mais elle ne permet à aucun moment de suivre l'évolution d'un territoire en particulier. À la suite de ces travaux, une seconde famille de modèles est apparue, nommée *Space-Time Composite* qui s'attachait à représenter le changement sur un territoire. [LAN 98]. Ces modèles proposent de chercher le plus petit commun dénominateur spatial (*PPCD_spatial*) entre toutes les unités géographiques, et de l'utiliser comme table de passage entre chaque version. Ce type de modèle est souvent implémenté [GRE 02], [NOR 03] [PAQ 04], [ALB 07]. Cependant, ces modèles posent trois problèmes :

- un problème d'implémentation, car, dès qu'une nouvelle version de maillage est à acquérir, il faut recoder le *PPCD_spatial* s'il a été calculé par intersection des maillages, et renuméroter toutes les tables de passage.

- un problème de sémantique, car l'identité des territoires n'est pas capturée : l'espace géographique est considéré comme un ensemble de parcelles géographiques codées, sans lien de généalogie ou d'évolution entre elles.
- la complexité inhérente que présentent les requêtes qui doivent reconstituer l'état du maillage et les relations topologiques à un instant précis.

Enfin, ces modèles ne prennent pas en compte la gestion de hiérarchies multiples, qui sont évolutives.

La modélisation de hiérarchies sur la dimension spatiale a également été abordée dans le cadre des systèmes d'information décisionnels spatiaux dits « SOLAP » [BED 97]. La question de l'évolution des maillages ou des supports géographiques est traitée [TCH 05] à l'aide d'une grille régulière fine de l'espace qui sert de *PPCD_spatial* pour organiser les transferts de données entre les différentes couches temporelles. Des fonctions de mappage relativement simples sont utilisées pour réaffecter les indicateurs statistiques entre les différentes couches temporelles (suivant la proportion à la surface intersectée entre une couche temporelle et le *PPCD_spatial*). Ce type d'approche règle partiellement le problème de l'implémentation puisque c'est ici un unique support matriciel prédéfini à l'avance qui sert de *PPCD_spatial*. Mais le choix de l'échelle de cette matrice peut s'avérer plus délicat que prévu, car la finesse des maillages varie spatialement et temporellement, et on doit procéder à une approximation de la forme irrégulière d'un polygone. Ce traitement du changement spatial ne tient pas non plus compte des changements d'organisation dans la hiérarchie, et enfin, aucun lien sémantique n'est établi entre les unités géographiques : on ne sait pas si ce sont les mêmes ou bien des nouvelles. Or dans l'étude de l'évolution d'un phénomène démographique par exemple, l'identité du territoire est particulièrement importante.

Une approche radicalement différente s'appuie sur l'étude de l'évolution des unités élémentaires du support de l'information que sont les unités géographiques, auxquelles une identité a été attachée [CHE 97], [LAR 99], [THE 99]. À chaque unité géographique, possédant une géométrie (de type quelconque) à un instant t , est attaché un ensemble d'attributs thématiques (les indicateurs statistiques) qui évoluent dans le temps, et le tout (géométrie plus thématique) est identifié par un identifiant informatique unique. Les attributs de l'unité, tels que son nom, son code, son centre peuvent évoluer, mais l'identifiant unique permet d'établir sa trajectoire spatio-temporelle (apparition, évolution et disparition). Bien que l'établissement de l'identité d'une unité géographique dépende de paramètres subjectifs, elle a l'avantage de conférer un sens à l'évolution des unités géographiques. C'est pourquoi, nous nous appuyons sur ce type d'approche pour proposer un modèle où le territoire est représenté sous la forme d'un graphe spatio-temporel modélisant les relations qu'entretiennent les unités géographiques. De plus, dans notre approche, nous enrichissons le contexte spatio-temporel de chaque unité avec les relations d'agrégation qu'elle peut entretenir avec d'autres unités dans le cadre d'organisation territoriales multiples et hiérarchiques.

3. Le système territorial : une organisation complexe et dynamique

Notre objectif est de permettre de collecter des données, issues de sources hétérogènes, qui évoluent dans le temps. Les données sont les valeurs d'indicateurs (ou variables) statistiques qui sont associées à des géométries qui forment, par composition, le support géographique.

3.1 Des supports hétérogènes

L'hétérogénéité des données réside tout d'abord dans leur support géographique. Nous limitons notre modélisation au support de type polygonal irrégulier, le plus fréquent dans le cas des données socio-économiques. Par exemple, le recueil de données statistiques au niveau européen est harmonisé dans la Nomenclature des Unités Territoriales Statistiques (NUTS). Cette nomenclature reprend les découpages territoriaux historiques, qui sont de formes polygonales irrégulières et structure l'information en cinq niveaux, du plus fin au plus large : LAU2 (communes), LAU1 (cantons), NUTS3 (départements), NUTS2 (régions), NUTS1 (grandes régions), et NUTS0 (pays).

Mais, nous visons aussi à l'intégration de supports plus particuliers :

- les formes morphologiques des villes (*Urban Morphological Zones*, UMZ) ;
- les groupements de communes en structures de coopération intercommunale (EPCI à fiscalité propre et/ou territoire de Pays en France) ;
- l'ensemble des bassins versants d'un espace d'étude ;
- la nomenclature des États reconnus par l'ONU (les WUTS).

Dans cette première étude, nous nous limitons à des objets vectoriels de forme polygonale, et nous ne prenons pas en compte les grille régulières (formats *raster*).. Par exemple, nous écartons la grille Corine Land Cover dont les pixels couvrent des zones d'un hectare, et sur lesquels on identifie le type majoritaire d'occupation du sol.

Ces supports de type polygones irréguliers sont définis comme des partitions spatiales de l'espace géographique qui forment alors un *maillage* du territoire. Un maillage est un découpage ou une partition de l'espace. Il peut servir à l'appropriation, la gestion, l'aménagement ou la connaissance de l'espace. Il existe en effet différents types de maillages, selon les acteurs et les fonctions : maillage politico-administratif, maillage statistique, zonage d'aménagement du territoire, etc. Un maillage peut être couvrant (partition totale de l'espace) par rapport à un espace d'étude déterminé, qu'on appelle l'aire d'étude, ou bien non couvrant. Dans le premier cas, on dira que le maillage est continu, dans le second discontinu. Les cas de maillages non couvrants sont fréquents, et sur ces maillages, des données socio-économiques sont aussi collectées, ou bien analysées. On peut citer par exemple, les bassins d'emplois en Europe, les formes morphologiques des villes (nomenclature UMZ), et les structures intercommunales. On peut noter que les données sont encore très rarement collectées directement au niveau des structures intercommunales, mais plutôt au niveau communal, et l'on doit agréger en fonction de la composition communale des structures intercommunales.

Très fréquemment, les nomenclatures qu'utilisent les producteurs de données définissent des niveaux hiérarchiques constitués de maillages qui s'imbriquent les uns dans les autres à partir d'un maillage élémentaire, le maillage le plus fin de l'espace dans la nomenclature considérée. Ces différents niveaux constituent une *organisation hiérarchique* du territoire dans laquelle chaque unité appartient à une ou plusieurs unités de niveau supérieur. Le niveau constitue aussi une échelle d'observation et d'analyse du territoire [GRA 07].

Lorsque chaque unité possède au moins une et une seule entité supérieure, et que toute unité de niveau non élémentaire englobe au moins une unité, la hiérarchie est dite *stricte* et *onto*. Ce cas d'organisation arborescente des unités territoriales est intéressant et étudié [RIG 95], car il permet l'agrégation ascendante depuis les feuilles vers la racine de l'arbre des données, et offre des possibilités de vérification avec les invariants suivants :

- (i) la géométrie d'une unité non élémentaire est l'union des géométries des unités qui la composent ;
- (ii) la valeur d'une variable (quantitative absolue) d'une unité non-élémentaire est la somme des variables associées aux unités qui la composent.

Dans le cadre de la constitution d'un réservoir de données hétérogènes, la connaissance de ces relations d'agrégation est nécessaire car les opérations d'agrégation vont permettre de rapporter dans un maillage commun des indicateurs mesurés sur des niveaux de maillages différents. On peut également s'appuyer sur des contraintes de désagrégation utiles : par exemple, un effectif compté sur un département se répartit sur les effectifs des communes composant ce département. On souhaite donc se ramener à la gestion de hiérarchies de maillage, pour lesquelles la disponibilité des variables et des géométries sur le maillage élémentaire permet de constituer, par agrégation, un ensemble d'informations multi-échelles.

Cependant, il apparaît que l'organisation hiérarchique stricte des unités territoriales n'est pas systématique. Les récentes recompositions territoriales locales en France en sont un bon exemple. Pour rationaliser la gestion territoriale, il s'agit de relancer l'intercommunalité et de construire des territoires plus vastes que la commune, d'une taille suffisante pour atteindre une masse critique en matière de population. De nouveaux Établissements Publics de Coopération Intercommunale (EPCI) sont créés, mais également ce qu'on appelle des « territoires de projet » (politique des Pays). Même si ce n'est pas explicite dans les lois, une forme de structure hiérarchique avec emboîtement spatial se dessine. Par exemple, en milieu rural, un premier niveau est constitué par les communautés de communes (un des nouveaux EPCI) ; un second par le Pays, fédérant plusieurs communautés de communes. Le Pays est le niveau de réflexion et de conceptualisation du projet de développement local ; les communautés de communes sont chargées des réalisations concrètes. Mais la création de ces maillages intercommunaux est progressive, puisque les différentes structures ne sont pas obligatoires et reposent pour partie sur l'initiative des acteurs locaux. En conséquence, il existe un certain nombre de cas ne respectant pas un emboîtement hiérarchique strict. La figure 2 présente certains de ces cas, et illustre comment ce que nous avons défini comme « nomenclature des

intercommunalités », constituée de trois niveaux (les communes, les EPCI et les Pays), forme une hiérarchie non *stricte* et non *onto* :

1. Une unité d'un certain niveau peut avoir deux unités de niveau directement supérieur. Sur l'exemple, l'EPCI *e4* est partagé entre le Pays *p3* et le Pays *p4*. Dans ce cas, *e4* n'adhère pas à la fois à *p3* et *p4* ; ce sont une partie des communes constituant *e4* qui adhèrent à *p3*, et l'autre partie des communes qui adhèrent à *p4*. Ces cas devraient être transitoires, les EPCI devant normalement respecter, à terme, les limites des Pays, et ceux-ci étant définis, en général, par leurs EPCI adhérents, et non par les communes adhérentes.
2. On observe des appartenances multiples à plusieurs niveaux de maillages. Sur l'exemple, la commune *c7* adhère à un EPCI *e4* mais aussi à un pays *p3* (cas possible si *e4* n'adhère pas dans son ensemble à *p3*).
3. Une unité d'un niveau peut n'avoir aucune unité supérieure dans la nomenclature considérée. Sur l'exemple, la commune *c4* n'a pas d'unité supérieure.
4. Les sauts de niveau sont fréquents. Sur l'exemple, l'unité *c6* n'a pas d'unité supérieure dans les EPCI, mais appartient directement à un pays *p2*. Une autre variante est la hiérarchie non complète avec manque du niveau supérieur : cas de *c5*, qui appartient à *e3*, mais il n'y a pas de Pays au niveau supérieur.

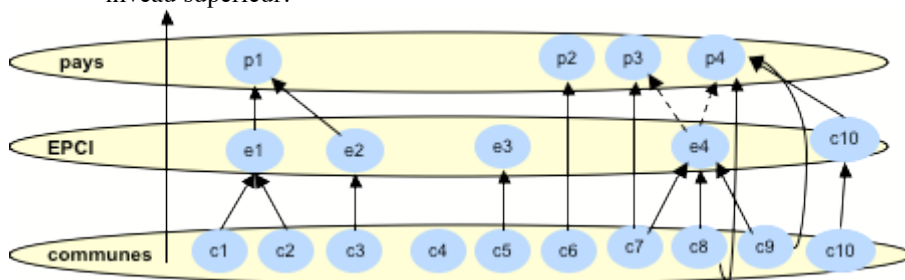


Figure 2. Nomenclature des intercommunalités sur une région française.

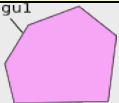

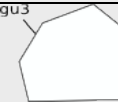

Le modèle que nous proposons vise à produire de l'information agrégée sur des niveaux non élémentaires. Cela reste possible dans ces cas, à condition de définir plus précisément les relations d'agrégation dans la nomenclature. La relation d'agrégation doit lier les niveaux concernés, chaque maillage appartenant à une certaine nomenclature, et cette relation d'agrégation pourra être typée. Le typage le plus simple est l'agrégation totale : les unités considérées appartiennent entièrement à leur unité supérieure. Dans ce cas, les prédicats (i) et (ii) donnés plus haut s'appliquent directement pour l'évaluation des unités supérieures. Sinon, on est dans le cas d'une situation d'agrégation partielle. C'est le cas de l'EPCI *e4* qui appartient aux pays *p3* et *p4*. Mais dans ce cas, la géométrie des unités de niveau supérieure ne peut être constituée par agrégation des unités lui appartenant partiellement. Il faut donc chercher dans un maillage de niveau inférieur une relation d'agrégation totale pour constituer l'unité. Sur l'exemple, les communes *c8* et *c9* et *c10* forment par agrégation totale le pays *p4*.

3.2. Vers une formalisation du changement des mailles territoriales : approche par l'identité

Quel que soit le type de nomenclature, les maillages connaissent des évolutions, soit sur le plan de la forme des unités territoriales au niveau le plus élémentaire, soit sur le plan de l'organisation (ce qui affecte les niveaux non-élémentaires). Les effets des évolutions sont visibles au niveau de chaque unité géographique. De précédents travaux [CHE 97] soulignent que l'attribution d'une *identité* à une unité géographique peut permettre de retracer son évolution spatio-temporelle, puisque tous ses attributs sont susceptibles de varier de façon indépendante : son empreinte spatiale, son nom, son centre, son code, son statut, son appartenance, etc. Certains auteurs proposent une ontologie du changement [HOR 98], basée sur la reconnaissance d'une identité aux unités géographiques, mais celle-ci n'a pas d'implémentation immédiate parce que l'identité d'une unité géographique est subjective. En effet, concrètement, c'est la version de nomenclature qui tranche de façon arbitraire sur la continuité de l'identité d'une unité.

3.2.1. Les événements du changement

Grâce à l'identité d'une unité géographique (*gu*), il est possible de suivre la variation des attributs qui la composent, c'est-à-dire l'évolution de l'unité. Mais pour donner une explication, un sens à cette évolution, il faut la relier aux événements territoriaux. Dans le cadre de l'étude que nous avons menée sur les recompositions territoriales des maillages administratifs en Tunisie [BEN 08], nous soutenons l'hypothèse selon laquelle les changements correspondent à des logiques et des stratégies politiques, et qu'ils sont la conséquence des opérations menées par un pouvoir politique. L'importance de retranscrire les événements territoriaux (les causes du changement) en plus des effets du changement a été largement argumentée dans la littérature [CLA 95]. Il faut globalement en retenir qu'un événement est décrit par une date, un état territorial initial et un état territorial final. En nous appuyant sur le vocabulaire défini dans [SPE 01], nous présentons la liste d'événements à gérer dans le tableau 1.

Évènement		Etat initial	Etat final
ÉVÈNEMENTS ÉVOLUTIFS	apparition : l'unité (<i>gu1</i>) est créée avec une empreinte spatiale et un code la rattachant à au moins une nomenclature.		
	disparition : l'unité (<i>gu1</i>) termine son existence – tous ses attributs (empreinte spatiale, nom, etc.) achèvent leur période de validité avec elle.		
	transformation : un des attributs identitaires (nom, centre, empreinte spatiale, code, etc.) de l'unité est modifié sans changement de son identité.		

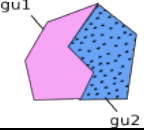

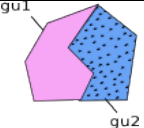
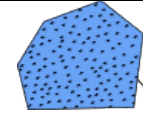
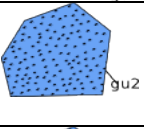
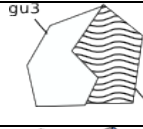
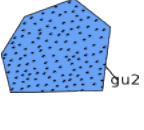
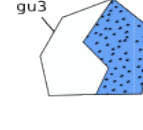
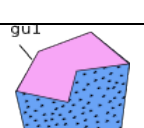
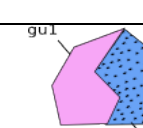
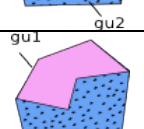
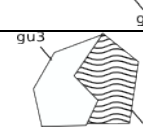
ÉVÈNEMENTS GÉNÉALOGIQUES	fusion : toutes les unités impliquées (<i>gu1</i> , et <i>gu2</i>) disparaissent au profit d'une nouvelle (<i>gu3</i>) agrégeant les empreintes spatiales des disparues.		
	intégration : toutes les unités impliquées disparaissent, sauf une (<i>gu2</i>), qui s'agrandit en agrégeant l'empreinte spatiale des disparues (<i>gu1</i>).		
	scission : une unité (<i>gu2</i>) disparaît au profit de plusieurs nouvelles entités (<i>gu3</i> et <i>gu4</i>), qui récupèrent l'espace couvert par l'empreinte spatiale de la disparue.		
	extraction : une ou plusieurs unités apparaissent (<i>gu3</i>) au détriment, en termes de surface, d'une unité qui existait (<i>gu2</i>) avant l'événement et qui continue d'exister.		
	rectification : aucune unité n'apparaît ou ne disparaît, mais les unités impliquées (<i>gu1</i> et <i>gu2</i>) échangent du territoire, tout en conservant leur identité.		
	réallocation : une portion du territoire est réallouée et une partie ou toutes les unités disparaissent (<i>gu1</i> et <i>gu2</i>) au profit de nouvelles unités (<i>gu3</i> et <i>gu4</i>) qui apparaissent et se partagent l'espace occupé par les disparues.		

Tableau 1. Les évènements du changement territorial, adapté de Spery et al.

Les évènements cités dans le tableau 1, et classés en deux catégories, « évolutifs », et « généalogiques », peuvent avoir les conséquences suivantes : le changement de nom ou de géométrie d'une unité, le changement de statut (qui induit la modification de composition de deux maillages), le changement d'appartenance d'une unité (qui entraîne la modification d'une relation hiérarchique), et le changement de centre. Le modèle doit alors décrire ces conséquences, en les reliant à la réalisation d'un certain évènement.

4. Modèle d'information spatio-temporelle hiérarchique évolutive

Le système d'information spatio-temporel que nous proposons repose sur un modèle orienté objet du système territorial, qui propose de *gérer simultanément plusieurs nomenclatures*, et qui intègre *l'unité géographique* comme objet central d'étude à l'intérieur de chaque nomenclature. En effet, chaque nomenclature définit un niveau de découpage élémentaire du territoire, par la donnée d'un ensemble

d'unités géographiques élémentaires. Chaque unité géographique apparaît, évolue et disparaît, suivant un critère d'identité qui dépend de la nomenclature étudiée. L'historique d'une unité est enregistré, et sa durée de vie est estampillée avec un intervalle temporel (*validityPeriod*). Nous lions l'unité géographique aux différents événements de recomposition territoriale qui peuvent engendrer son apparition, son évolution, ou sa disparition. Le modèle comprend deux parties qui peuvent être considérées indépendamment l'une de l'autre :

- la *partie identitaire* décrit un ensemble générique d'attributs qui peuvent participer à la définition de l'identité d'une unité géographique, pour chaque type de nomenclature.
- la *partie événementielle* décrit l'ensemble des processus de transformations survenus dans une nomenclature et leur impact sur la généalogie et les attributs identitaires des unités.

Le modèle est décrit par un diagramme de classes dans le formalisme UML [BOO 99], dans lequel nous n'avons pas employé les notations d'agrégation ou de composition forte, afin d'éviter toute ambiguïté avec les relations de composition ou d'agrégation explicitées comme des classes d'association entre les unités géographiques. De même, des contraintes spatio-temporelles entre les entités pourraient être formulées, en utilisant l'OCL comme le propose [GRU 01]. Afin de ne pas surcharger le modèle, les annotations OCL ne sont pas incluses, et toutes les cardinalités non indiquées sont des cardinalités *0:n*. Les cardinalités indiquées sont celles définies pour un instant d'observation quelconque.

4.1. Identification d'une unité géographique

L'identification d'une unité géographique se fait par la description complète de l'unité, à un instant *t*, à l'intérieur d'une nomenclature (voir figure 3). Tous les attributs ou les entités que nous proposons dans cette partie sont susceptibles d'évoluer indépendamment les uns des autres.

L'unité géographique (*GeographicUnit* ou *gu*) est identifiée par un identifiant spatio-temporel unique (*id*), qui correspond à sa période d'existence. Chaque unité géographique s'inscrit dans au moins une organisation territoriale définie par une Nomenclature (*Nomenclature*). La nomenclature porte un nom (*nom*), un code (*code*), et est définie par un institut (*institut*), pour une certaine période de validité (*validityInterval*). Lorsque cet institut veut rendre compte de changements territoriaux qui se sont produits, il émet une nouvelle version de Nomenclature : chaque version a donc une période de validité (*validityInterval*). Chaque fournisseur de données statistiques (*Provider*) s'appuie sur une version de Nomenclature lorsqu'il publie des données via un support physique (la *Source*). Pour chaque nomenclature, l'unité est associée à un code (*Code*), qui peut changer suivant les versions de nomenclature. Il est nécessaire de conserver ces codes, car les données statistiques sont fournies en lien avec une version de nomenclature dans laquelle l'unité géographique est identifiée par son code.

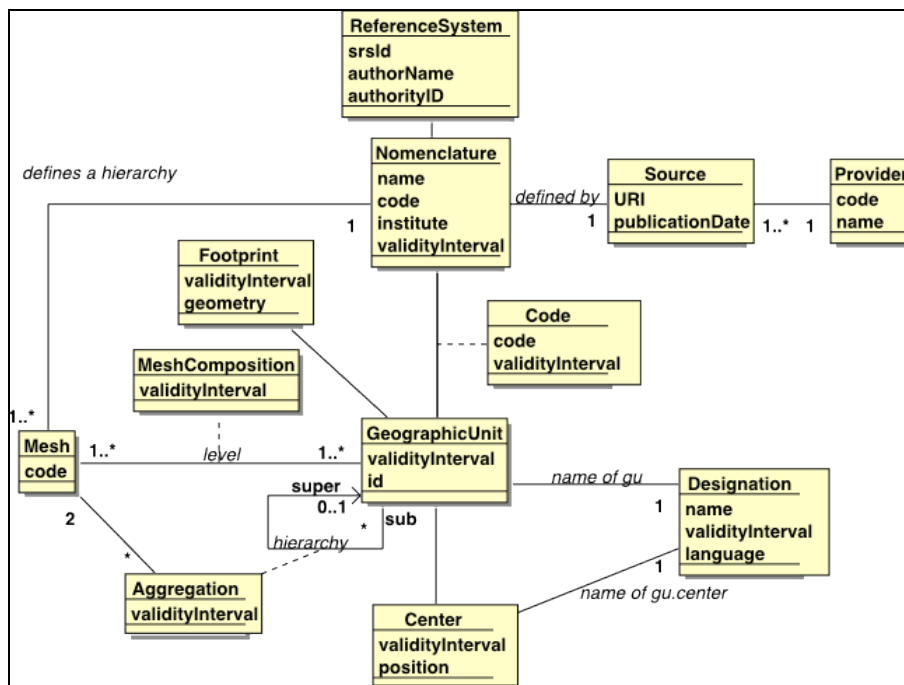


Figure 3. L'identité des unités dans une nomenclature.

À l'intérieur de chaque Nomenclature, l'unité géographique peut posséder plusieurs désignations (*Designation*) en fonction des périodes de temps et des langues.

L'unité géographique peut aussi posséder un centre (*Center*), défini par une position dans l'espace et un nom (*Designation*). La géométrie du centre dépend du type de nomenclature considéré. Ce sera, par exemple, le point le plus bas d'un bassin versant, un point représentant une capitale dans les WUTS, une adresse géographique pointant sur le siège d'un EPCI, ou bien une unité géographique de niveau inférieur dans le cas des départements.

Chaque version de nomenclature est associée à un certain fonds de carte géographique, utilisant un unique système de projection (*Reference System*). Ce fond est constitué par le maillage (*Mesh*) de niveau élémentaire, et permet d'associer chaque unité à son empreinte spatiale (*Footprint*). Cette dernière peut évoluer au gré des événements du changement, et est donc estampillée avec une période de validité (*validityInterval*). Nous ne conservons qu'une seule empreinte spatiale par version de nomenclature, car notre modèle a pour objectif l'estimation, et non pas la représentation cartographique qui nécessite très souvent de stocker divers niveaux de généralisation de l'information géographique. Les empreintes spatiales des unités géographiques de niveau élémentaire feront l'objet d'un processus d'agrégation géométrique lorsque la nomenclature fournit plusieurs niveaux de hiérarchies. Par exemple, l'union des géométries des communes de l'Isère produit le contour géométrique du département de l'Isère.

Chaque nomenclature (*Nomenclature*) définit au moins un maillage (*Mesh*), couvrant ou non, pour lequel on précise la composition : c'est l'objectif de la classe d'association *MeshComposition* entre une unité géographique (*GeographicUnit*) et un maillage (*Mesh*). L'appartenance à un maillage peut changer dans le temps si une unité change de statut : *MeshComposition* précise par une période de validité la durée effective de la relation.

La nomenclature peut définir aussi plusieurs niveaux de maillage, strictement emboîtés ou non. Comme nous l'avons argumenté, la description des relations d'agrégation hiérarchique doit mettre en jeu les unités, mais aussi les maillages auxquels elles appartiennent, ainsi que les nomenclatures impliquées. Cette relation verticale est notée *Aggregation* dans le modèle. Il s'agit d'une classe d'association entre deux unités, l'une (*super*) étant l'unité supérieure de l'autre (*sub*). Elle est datée par la période de validité (*validityInterval*) du rattachement, et elle est liée à deux niveaux de maillages différents (*Mesh*) : l'un pour le niveau inférieur de l'unité, et l'autre pour le niveau supérieur de l'unité, chaque maillage étant défini dans une version de Nomenclature.

Pour illustrer l'utilisation de cette relation dans un schéma relationnel, on propose la relation suivante : *Agregation* (*gu_inf_id*, *mesh_inf_id*, *gu_sup_id*, *mesh_sup_id*, *nomenclature_id*, *validityPeriod*). Son instantiation est illustrée avec le cas de la commune *c7* représentée dans la figure 1:

- *agregation* (*c7*, *communes*, *e4*, *epci*, *interco*, [*1998*, *now*]) indique que la commune *c7* fait partie de l'EPCI *e4* dans la nomenclature des intercommunalités depuis 1998.

- *agregation* (*c7*, *communes*, *p3*, *pays*, *interco*, [*2000*, *now*]) indique que la commune *c7* adhère aussi au pays *p3* dans la nomenclature des intercommunalités, depuis 2000.

- *agregation* (*c7*, *communes*, *dep38*, *département*, *NUTS*, [*1989*, *now*]) indique que la commune *c7* appartient au département *dep38* dans la nomenclature des NUTS depuis 1989.

On remarque ainsi que certaines nomenclatures peuvent partager des niveaux géographiques communs : les intercommunalités et les NUTS partagent le même niveau communal. Dans ce cas, la commune n'est définie qu'une fois dans le système, par son empreinte spatiale, son nom, son centre, etc., mais elle porte un code (*Code*) spécifique à chacune des nomenclatures.

4.2. La généalogie et la transformation des unités

Le modèle explicite aussi les transformations d'identité que connaissent les unités qui sont causées par les événements du changement (voir le diagramme de la figure 4, qui est une extension du diagramme de la figure 3, à partir de l'objet *GeographicUnit*). Chaque événement possède un identifiant unique, et il est daté. Il existe deux types d'évènements (voir le tableau 1) :

- les événements de type « généalogique » (*GenealogyEvent*) associent les unités géographiques qui précèdent un évènement (*pred*), et les unités qui succèdent à l'évènement (*suc*).

- les évènements de type « évolutif » (*LifeEvent*) ne sont associés qu'à une seule unité à la fois, et sont utilisés pour décrire l'évolution d'une unité. Cette évolution peut être liée à un événement généalogique, mais pas obligatoirement.

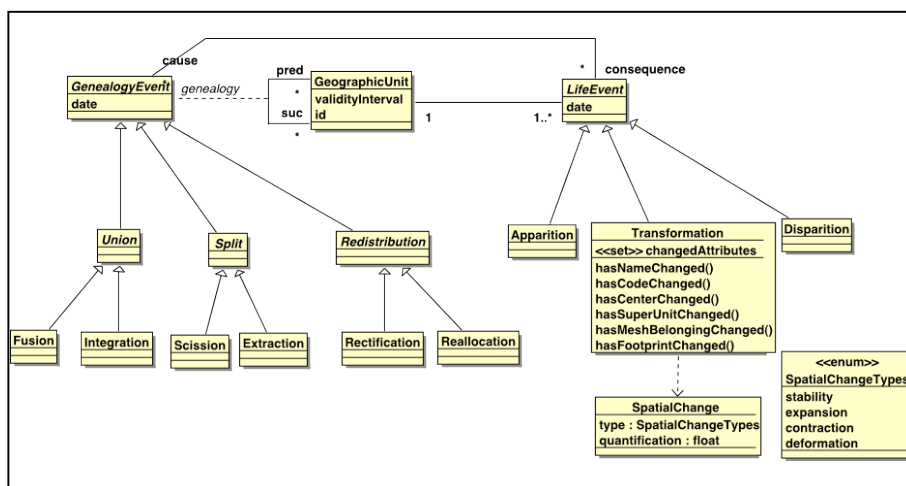


Figure 4. Association des unités aux évènements historiques et leurs conséquences.

Un événement de type généalogique permet de faire le lien entre l'état initial du système et l'état final du système. On note *pred(event_id)*, la fonction renvoyant l'ensemble des gu qui précèdent l'événement « *event_id* », et *suc(event_id)*, celles qui lui succèdent. Les événements de type généalogiques peuvent avoir des impacts sur l'évolution individuelle des unités impliquées dans l'événement : elles peuvent apparaître (*Apparition*), ou bien se transformer sans perte de leur identité (*Transformation*) ou bien disparaître (*Disparition*). La classe *Transformation* indique quels sont les attributs de l'unité géographique qui ont changé. En particulier, si l'information est calculée, elle peut utiliser la classe *SpatialChange* pour préciser le type de changement spatial qui a eu lieu, et dans quelle mesure (*quantification* donne le ratio de la surface finale sur la surface initiale).

Voici quelques exemples, de complexité croissante, montrant comment cette modélisation permet de rendre compte du changement.

Une unité peut apparaître, disparaître ou se transformer en dehors d'un événement de généalogie. Elle peut, par exemple, changer de nom, comme la commune de Malleval qui devient Malleval-en-Vercors en 2005. Cette *Transformation*, caractérisée par un changement de nom, est associée seulement à la commune (voir figure 5).

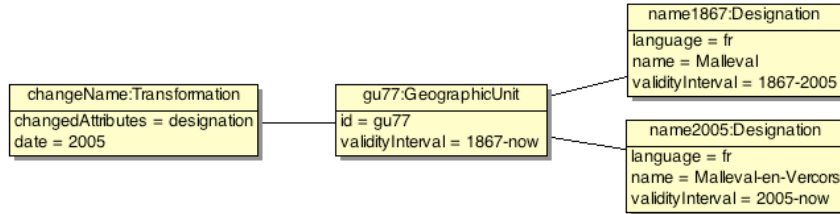


Figure 5. Diagramme d'instances du changement de nom de Malleval (2005).

Le modèle permet également d'enregistrer la modification des relations hiérarchiques à l'occasion d'une réallocation. Prenons, par exemple, le changement d'appartenance de St Priest, commune de l'Isère, qui passe en 1967 dans le département voisin du Rhône. C'est un événement de *réallocation* liant les trois unités, qui continuent leur existence, mais sont transformées (voir figure 6). Une *Transformation* est définie pour St Priest, sur son attribut d'appartenance hiérarchique. À chacun des départements de l'Isère et du Rhône, une *Transformation* de type spatiale est associée pour indiquer d'une part la *contraction* de l'Isère, et, d'autre part, l'*extension* du Rhône (voir figure 7).

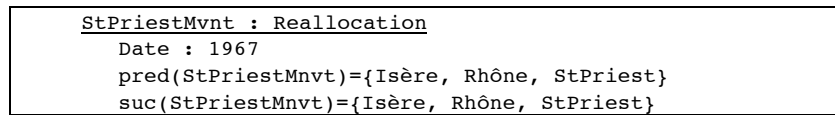


Figure 6. L'évènement Réallocation instancié pour le changement d'affectation de Saint Priest (1967).

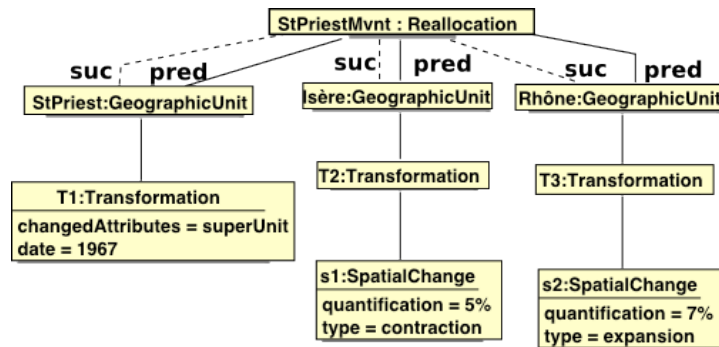


Figure 7. Diagramme d'instances du changement d'appartenance de St Priest (1967).

À une autre échelle, la réunification de l'Allemagne en 1990 est un évènement de type *fusion* qui fait disparaître la RFA et la RDA, au profit d'une nouvelle unité, Allemagne, ces trois unités sont liées par l'évènement « Einigungsvertrag » (voir figure 8). Mais, la RFA et la RDA sont aussi liées chacune à un évènement « *Disparition* », tandis que l'Allemagne est liée à un évènement « *Apparition* ».

Toutes les unités appartenant directement à la RDA et à la RFA changent d'unité supérieure : elles sont aussi impliquées dans cet événement, mais comme changeant d'unité supérieure (c'est une *Transformation* portant sur la relation hiérarchique).

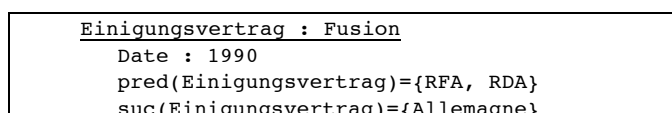


Figure 8. L'évènement *Fusion* instancié pour la réunification allemande (1990).

L'instanciation de la relation *consequence* indique les effets pour chaque unité des événements de généalogie sur chacune des unités impliquées dans l'évènement. Dans cette illustration, nous avons omis l'effet de transformation sur l'unité supérieure pour les Länder allemands, afin d'alléger le texte (voir figure 9).

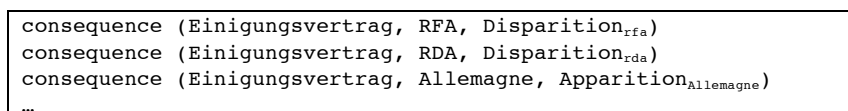


Figure 9. Conséquences de la réunification allemande (1990).

Enfin, l'opération de *réallocation* peut-être illustrée par le remembrement de trois communes de l'Isère conduisant à la création de Chamrousse. L'opération, identifiée « ISE89 » et illustrée par la figure 10, se produit en 1989 entre les trois communes iséroises de Vaulnaveys-le-Haut (gu_1), Séchilienne (gu_2), et Saint-Martin-d'Uriage (gu_3), et fait apparaître une nouvelle commune, nommée « Chamrousse » (gu_4). La figure 11 montre l'instanciation de l'objet ISE89 correspondant.

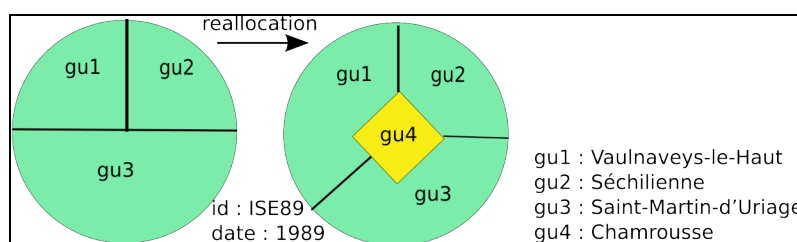


Figure 10. Illustration d'une réallocation.

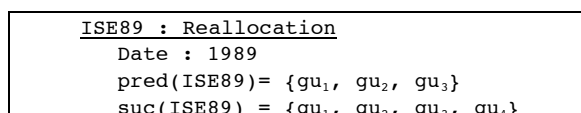


Figure 11. L'évènement *Réallocation* instancié pour la création de Chamrousse.

La commune de Chamrousse est liée à un événement de type « Apparition » pour cette date-là, tandis que les trois autres sont chacune liées à un événement de type « Transformation », où l’empreinte spatiale (*Footprint*) est indiquée comme étant l’attribut changé. Le changement est décrit par la classe *SpatialChange* qu’utilise l’événement de transformation pour décrire les conséquences de l’événement. Par exemple, la relation *consequence(genealogyEvent, geographicUnit, transformation)* peut être instanciée comme suit (voir figure 12), si on considère que *gu1*, *gu2*, et *gu3* sont les trois communes qui ont échangé du terrain dans les transformations d’instances respectives *T1*, *T2*, *T3*.

<i>consequence</i> (<i>ISE89</i> , <i>gu1</i> , <i>T1</i>)
<i>consequence</i> (<i>ISE89</i> , <i>gu2</i> , <i>T2</i>)
<i>consequence</i> (<i>ISE89</i> , <i>gu3</i> , <i>T3</i>)
<i>consequence</i> (<i>ISE89</i> , <i>gu4</i> , <i>A1</i>)

Figure 12. Conséquences du remembrement créant Chamrousse.

Le diagramme d’instance de la figure 13 montre comment les événements d’évolution *T1*, *T2*, *T3*, qui sont trois instances de la classe *Transformation*, décrivent le changement induit par la réallocation ISE89. En particulier, elles indiquent que *Footprint* a changé, et utilisent une instance de *SpatialChange* caractérisant leur changement spatial.

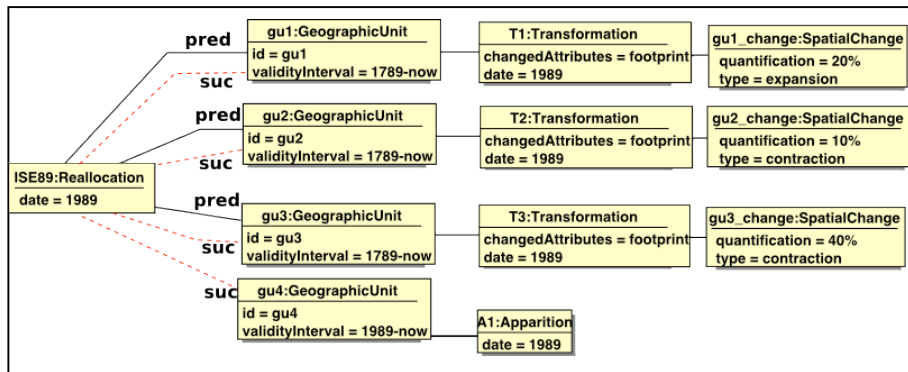


Figure 13. Diagramme d'instances du changement pour l'apparition de Chamrousse (1989).

Ces événements sont modélisés afin de rendre compte sur un axe historique temporel de toutes les évolutions subies par une unité géographique (les événements de type *LifeEvent*). Sur le plan spatial, cette modélisation garde trace des interactions entre unités survenues à un moment donné grâce à l’association à un événement commun de type *GenealogyEvent*.

Dans tous les cas, l’intervalle de validité, noté [*ti*, *tf*], des entités disparues prend fin à la date de l’évènement, et celui des nouvelles unités débute à cette même date. Il existe d’autres invariants que nous détaillons : une unité géographique n’est liée

qu'à un seul événement d'apparition ou de disparition. Les événements de transformation sont toujours datés entre la date d'apparition et celle de disparition. Ensuite, les événements de type *Union* conduisent toujours à l'expansion spatiale d'une des unités : soit la nouvelle, soit celle qui perdure. Les événements de type *Split* morcellent le terrain et réduisent inévitablement les unités spatiales impliquées : il y a contraction spatiale. Pour la *Redistribution* territoriale, aucune des unités impliquées ne reste stable au niveau de l'empreinte spatiale (*footprint*), sauf éventuellement les unités qui les composent, et qui participent à l'opération. Ces invariants peuvent être utilisés afin d'inférer certains changements.

L'instanciation de cette partie du modèle est prévue à partir d'une interface graphique, grâce à laquelle on peut spécifier un événement et ses conséquences. L'ensemble du modèle est, en effet, prévu pour permettre l'acquisition progressive des données, *via* la saisie indépendante de chaque version de nomenclature. En premier lieu, l'acquisition de la partie attributaire des unités est automatique, à partir d'un format de données déjà entièrement spécifié. En second lieu, chaque acquisition peut donner lieu à la saisie des événements de généalogie et d'évolution par l'expert. Cette saisie est facilitée par le système qui infère tout ou partie des événements par comparaison des attributs des unités, pour proposer des événements possibles à l'utilisateur du système. Ce dernier peut corriger et enrichir la liste et la nature des événements dans l'interface graphique, pourvu qu'ils respectent les invariants qui ont été listés, et les lier aux unités concernées.

5. Conclusion et perspectives

Notre objectif est de produire un modèle capable d'enregistrer plusieurs hiérarchies territoriales évoluant simultanément. Nous avons d'abord présenté les principales difficultés et une formalisation du problème. La difficulté est, en premier lieu, liée à l'hétérogénéité des découpages territoriaux, qui ne s'organisent pas tous dans des hiérarchies strictes. En second lieu vient l'évolutivité des supports qui, bien que très étudiée, a rarement été modélisée dans un cadre générique, pour tout type de support irrégulier. Nous avons introduit une modélisation des unités géographiques constituant chaque type de support, la plus complète possible, et surtout intégrant les relations de hiérarchie et de composition multiple de maillage. De plus, l'article expose une modélisation détaillée du changement permettant de capturer dans le modèle l'histoire des unités géographiques, avec les événements du changement et leurs conséquences.

Ce modèle fait l'objet d'une implémentation, mais seulement pour sa partie identitaire, sur une base de données PostgreSQL avec sa cartouche spatiale PostGIS dans le cadre d'un projet Européen, ESPON 2013 database [ESP 06]. Ce projet a pour objectifs :

- le stockage de données statistiques issues de sources variées dans un système facilitant la combinaison d'indicateurs issus de sources hétérogènes ;

- la capture de l'expertise disponible tant sur les supports géographiques (connaissance des évolutions, alignement des nomenclatures), que sur les données statistiques (lignage des valeurs) ;
- la production d'estimation statistiques documentées et reproductibles de valeurs manquantes pour compléter les jeux de données.

Le projet aborde à la fois les échelles locales (comme la nomenclature des intercommunalités), les échelles intermédiaires (la classique nomenclature des NUTS), et, enfin, les échelles mondiales avec les bases de l'ONU, qui sont construites sur la nomenclature des WUTS.

La suite de ce travail consiste à intégrer des nomenclatures environnementales et urbaines, comme les bassins versants, et celles des UMZ. Puis, nous imaginons un système basé sur des ontologies spatio-temporelles qui permettra l'alignement des unités issues de différentes nomenclatures. Le nom, le centre, l'empreinte spatiale, le code, les maillages, l'unité supérieure et les unités inférieures, sont un ensemble d'attributs d'une unité géographique qui nous permettront d'établir un dictionnaire de passage inter-nomenclature. Par exemple, la France des WUTS correspond à la France des NUTS, mais sans les DOM-TOM. L'interface graphique permettant d'instancier le modèle du changement est également en cours de spécification et doit être implémentée sur la base de données. Elle doit permettre aux experts de saisir les tables de passage entre versions de nomenclature.

Nous intégrerons ensuite les valeurs des indicateurs, qui sont renseignées par un modèle de métadonnées conçu spécifiquement pour les besoins de données collectées sur des nomenclatures territoriales. Ce qui pourra enfin nous conduire à l'estimation de données manquantes exploitant le graphe spatio-temporel entre unités capturé par le modèle présenté ici.

7. Bibliographie

- [ALB 07] ALBRECHT J. Dynamic GIS. In Wilson J and S Fotheringham (Eds) *The Handbook of Geographic Information Science*, pp. 436-446. London: Blackwell. (2007)
- [ARM 88] ARMSTRONG, M.P. « Temporality in spatial databases ». *GIS/LIS'88, Proceedings Accessing the World Vol 2*. pp. 880-889, (1988)
- [BED 97] BÉDARD, Y., LAM, S., PROULX M.J., CARON, P.Y. and LETOURNEAU, F., « Data Warehousing for Spatial Data: Research Issues », In : *Proceedings of the International Symposium Geomatics in the Era of Rradarsat (GER'97)*, Ottawa pp. 25-30 (1997)
- [BEN 08] BEN REBAH M., « La cartographie dynamique comme outil d'investigation territoriale : le cas du découpage territorial en Tunisie », Thèse de doctorat, Université de Paris VII, (2008)
- [Boo 99] BOOCH G., RUMBAUGH J.E., JACOBSON I. : *The Unified Modeling Language User Guide*. *J. Database Management*. 10(4): 51-52 (1999)

- [CHE 97] CHEYLAN J.-P., LARDON S., MATHIAN H., Sanders L., Les problématiques liées au temps dans les SIG, *Revue Internationale de Géomatique*, n°4, pp. 287-305, (1997).
- [CLA 95] CLARAMUNT, C., THERIAULT M., Managing time in GIS: an event oriented approach, *Recent Advances on Temporal Databases*, Eds.: Clifford, J. and A., Tuzhilin, Zurich: Springer-Verlag, Switzerland, pp. 23-42, (1995)
- [ESP 06] ESPON - ESPON 3.2 DATA NAVIGATOR 2 Final Report Part 1 – Handbook for Data Collection – (2006)
www.espon.eu/mmp/online/website/content/projects/260/716/file_2647/fr-3.2-DN2_Final_Jan2007.pdf
- [GRA 07] GRASLAND, C., « A la recherche d'un cadre théorique et méthodologique pour l'étude des maillages territoriaux » Entretiens Jacques Cartier, « Les découpages du territoire », Lyon, décembre (2007)
- [GRE 02] GREGORY, I. Time-variant GIS Databases of Changing Historical Administrative Boundaries: A European Comparison. *Transactions in GIS*, vol. 6, n°2, pp 161-178 (2002)
- [GRU 01] GRUMBACH S., RIGAUX, P. and SEGOUFIN, L., Spatio-Temporal Data Handling with Constraints, *GeoInformatica* vol. 5, n°1, pp. 95-115. (2001)
- [HOR 98] HORNSBY, K., EGENHOFER M. Identity-Based Change Operations for Composite Objects. Proc. Eighth International Symposium on Spatial Data Handling, Vancouver, Canada, pp. 202-213, POIKER, T. and CHRISMAN, N. (eds). (1998)
- [KAV 01] KAVOURAS M., "Understanding and Modelling Spatial Change", in Frank A. Raper J. and Cheylan J.P. (Eds.): *Life and Motion of Socio-Economic Units*, Chapter 4. London: Taylor & Francis, GISDATA Series 8, (2001)
- [LAN 98] LANGRAN G. and CHRISMAN, N.R., « A Framework for Temporal Geographic Information », *Cartographic*, vol 25. n°3, pp. 11-14, (1998)
- [LAR 99] LARDON, S., LIBOUREL, T., CHEYLAN, J.-P., Concevoir la dynamique des entités spatio-temporelles, *Revue internationale de géomatique*, vol. 9, n°1, pp. 67-99. (1999)
- [NOR 03] NORMAN, P., REES, P., and BOYLE P., Achieving data compatibility over space and time: creating consistent geographical zones. *International Journal of Population Geography* vol. 9 pp. 365-386 (2003)
- [PAQ 04] PAQUE, D., « Gestion de l'historicité et méthodes de mise à jour dans les SIG », *Cybergeo*, Cartographie, Imagerie, SIG, article 278, mis en ligne le 23 juin 2004, modifié le 29 juin 2007. URL : <http://www.cybergeo.eu/index2500.html>. Consulté le 26 juin 2009.
- [PAR 06] PARENT, C., SPACCAPIETRA S., and ZIMÁNYI E. *Conceptual Modeling for Traditional and Spatio-Temporal Applications - The MADS Approach*. Springer, (2006).
- [PEU 02] PEUQUET, D., *Representations of Space and Time*, New York, Guilford Press, (2002).

- [RIG 95] RIGAUX, P. et M. SCHOLL Multi-scale partitions : Application to spatial and statistical databases. In SSD, pp. 170–183.(1995)
- [SPE 01] SPÉRY L., CLARAMUNT C., LIBOUREL T., « A Spatio-Temporel Model for the Manipulation of Lineage Metadata », *GeoInformatica*, Vol. 5, n°1, pp. 51-70, (2001)
- [THE 99] THERIAULT, M., CLARAMUNT, C., La représentation du temps et des processus dans les SIG : une nécessité pour la recherche interdisciplinaire, *Revue internationale de géomatique*, vol 9., n°1, pp. 67-99. (1999)
- [TCH 05] TCHOUNIKINE A., MIQUEL M., LAURINI R., AHMED T., BIMONTE S. et BAILLOT V.. Panorama de travaux autour de l'intégration de données spatio-temporelles dans les hypercubes. *Revue des Nouvelles Technologies de l'Information - Entrepôts de données et l'Analyse en ligne*. France : Cépaduès-Éditions, pp. 21-33 (2005)
- [WOR 98] WORBOYS, M.F., A generic model for spatio-bitemporal geographic information. In M.J Egenhofer and R.G. Golledge (eds.) *Spatial and Temporal Reasoning in Geographic Information Systems*. New York : Oxford University Press, pp. 25-39, (1998)
- [YUA 99] YUAN, M., « Use of a three-domain representation to enhance GIS support for complex spatio-temporal queries », *Transactions in GIS*, Vol. 3, pp. 137-159 (1999)