

PRÉSERVATION ET MODIFICATION DES RELATIONS ENTRE LES DONNÉES THÉMATIQUES ET LES DONNÉES TOPOGRAPHIQUES LORS DE MIGRATION DES DONNÉES THÉMATIQUES

par Kusay Jaara, Cécile Duchêne¹, Anne Ruas²

¹Institut géographique national (IGN France), 73 avenue de Paris - 94165 Saint-Mandé
prénom.nom@ign.fr

²Institut français des sciences et technologies des transports, de l'aménagement et des réseaux (IFST-TAR), 14-20 Boulevard Newton, Cité Descartes, Champs sur Marne, F-77447 Marne la Vallée Cedex 2
prénom.nom@ifsttar.fr

Ce papier est une version traduite et légèrement revue de la version originale, qui a été publiée dans le livre *Cartography from Pole to Pole* (éditeur : Springer) dans le cadre de la 26^e Conférence cartographique internationale 2013.

Les utilisateurs se servent souvent des données topographiques pour référencer leurs données thématiques. Si les données topographiques sont mises à jour, ou si l'utilisateur veut remplacer les données topographiques par d'autres, les données thématiques doivent être modifiées pour maintenir la cohérence. Nous appelons ce processus la *migration de données thématiques*. Ce papier propose une nouvelle version du processus de migration de données thématiques précédemment proposée dans (Jaara et al., 2012). Dans cette version, nous ajoutons le cas où les relations entre les données thématiques et les données topographiques doivent être modifiées (et non seulement préservées) entre l'état initial et l'état final. Un modèle pour décrire les relations et leurs modifications attendues est proposé. Une méthode multicritère a été utilisée pour relocaliser les données thématiques sur la nouvelle base topographique en fonction des relations modélisées. Le processus est illustré par un exemple type, sur lequel les résultats obtenus sont présentés et discutés.

1 Introduction

Les utilisateurs créent souvent leurs données thématiques en se basant sur des données géographiques fournies par des instituts publics ou issues d'une source de données collaborative. Ils utilisent alors ces données topographiques comme référence pour localiser leurs propres données thématiques. Un utilisateur peut être un opérateur professionnel ou

toute personne qui souhaite saisir des données géolocalisées (ex. données utilisateur issues d'un relevé de terrain). Une fois les données thématiques saisies, il arrive que les données de référence soient mises à jour, ou que l'utilisateur veuille les remplacer par d'autres, par exemple parce qu'elles sont trop détaillées ou parce que l'utilisateur veut fusionner ses données thématiques avec d'autres données thématiques saisies en utilisant une autre base topographique de référence. Les données thématiques doivent alors être modifiées pour assurer la cohérence avec les nouvelles données de référence. Nous appelons ce processus la migration de données thématiques (Jaara et al., 2012). Si la position relative initiale n'était pas prise en compte, le processus serait une intégration de données. La position relative est décrite par les relations. Les relations ont été utilisées dans d'autres travaux pour améliorer la qualité de données géographiques en modifiant localement les objets selon des contraintes définies au niveau des classes (Wallgrün, 2012 ; Brando-Escobar, 2013).

La migration de données thématiques peut être utilisée dans le but de propager les données thématiques entre les différents niveaux de détail d'une base de données multi-représentation (MRDB). Indépendamment du contexte MRDB, si une base de données contient des couches thématiques et des couches topographiques, et que l'utilisateur souhaite en dériver une base de données moins détaillée par généralisation automatique, une des solutions possibles est de généraliser les données topographiques, et d'effectuer ensuite une migration des données thématiques. Dans tous les cas, quand le

niveau de détail des données topographiques finales est trop grossier par rapport au niveau de détail des données thématiques initiales, la migration des données thématiques peut ne pas suffire, et une généralisation des données thématiques peut être nécessaire afin d'obtenir un résultat cohérent. La généralisation de données thématiques n'est pas étudiée dans ce papier, qui se concentre sur la migration des données thématiques.

Dans (Jaara et al., 2012), nous avons proposé un premier processus pour la migration de données thématiques : les relations spatiales initiales entre données thématiques et topographiques sont d'abord extraites et stockées dans un modèle de référencement dédié, puis une migration des données, basée sur ces relations, est effectuée.

Le présent papier propose une nouvelle version de ce processus de migration de données. Dans cette version, nous considérons le cas où les relations entre les données thématiques et les données topographiques doivent être modifiées, et non seulement préservées, entre l'état initial et l'état final.

Par ailleurs, ce papier explique plus en détail le modèle de représentation des relations, la prise en compte de leur modification éventuelle, et la méthode de relocalisation des données thématiques sur les données topographiques en fonction des relations attendues.

Cet article introduit tout d'abord l'idée de la modification des relations et propose une formalisation du problème étudié (partie 2). Une description générale de la nouvelle version du processus de migration de données thématiques est présentée dans la partie 3. La partie 4 décrit la modélisation des relations et de leurs modifications, ensuite, et la partie 5 décrit en détail la relocalisation des données thématiques en fonction des relations attendues. L'approche est illustrée par un cas d'utilisation des accidents routiers : des détails techniques de l'implémentation, ainsi que des résultats, sont présentés dans la partie 6.

1.1 Introduction de l'exemple type utilisé tout au long du papier

Les deux exemples-jouets suivants seront utilisés dans la suite du papier. Tous les deux concernent des accidents routiers (données thématiques) localisés sur les routes (données topographiques). Dans les deux exemples, le but est d'effectuer une migration de ces données thématiques vers une base topographique moins détaillée.

Le premier exemple contient deux accidents (figure 1a, en haut). L'un des deux est proche du rond-point et l'autre est un peu plus loin. Le rond-point est représenté par un carrefour dans la base topographique finale (figure 1b, en haut). Si ces accidents sont relocalisés vers la route la plus proche dans la base topographique finale, la relation de proximité du plus proche avec le rond-point sera perdue (figure c, en haut), malgré l'importance éventuelle de cette relation pour l'analyse des accidents.

Dans le deuxième exemple (figure 1a, en bas), nous avons un accident en face d'un bâtiment et proche d'une intersection route-rivière. Dans la base topographique finale, le bâtiment est plus éloigné de l'intersection route-rivière (figure 1a, en bas). Si l'accident est relocalisé vers le point le plus proche de la route dans la base topographique finale, il ne sera plus en face du bâtiment, alors que cette relation est peut-être importante pour l'utilisateur.

2 Modification des relations spatiales entre les données thématiques et les données topographiques

2.1 Pourquoi certaines relations doivent elles être modifiées ?

La représentation du monde réel varie d'une base topographique à une autre. La différence peut concerner la taille, la localisation ou la forme des objets. Les objets peuvent être agrégés avec d'autres ou supprimés. La dimension de leur géométrie peut changer, par exemple une rivière est représentée tantôt par une surface tantôt par une ligne. Ces changements induisent une modification des relations spatiales entre les données thématiques et les données topographiques dans la combinaison finale. Par exemple, la relation initiale de proximité avec le rond-point dans la figure 1 doit être modifiée, parce que le rond-point est représenté par un carrefour dans la base topographique finale. Dans certains cas, les relations doivent être modifiées même si l'objet de référence est identique dans les deux bases topographiques. Par exemple, si on imagine un bâtiment disjoint mais proche d'une zone d'avalanche dans les données initiales, il peut être pertinent de modifier sa relation en une relation d'adjacence dans les données finales, si la base topographique finale a un niveau de détail moindre et si la distance initiale était très petite par rapport au nouveau niveau de détail.

2.2 Description formelle de la problématique

Nous formalisons le problème étudié comme suit. Les objets topographiques servent de référence pour les objets thématiques. Le référencement d'un objet thématique par rapport aux objets topographiques se traduit par le fait qu'il existe des relations spatiales porteuses de sens entre l'objet thématique et les objets topographiques. Un objet thématique peut avoir plusieurs relations avec un ou plusieurs objets topographiques. Inversement, un objet topographique peut avoir plusieurs relations avec un ou plusieurs objets thématiques. Nous considérons que chaque relation implique seulement deux objets : un objet thématique et un objet topographique.

Nous adoptons les notations suivantes :

- A est l'ensemble des objets thématiques dans les données initiales,
- B est l'ensemble des objets topographiques dans les données finales,
- \mathcal{R} est l'ensemble des relations existant entre les objets et qui traduisent le référencement des objets thématiques par rapport aux objets topographiques, dans les données initiales,
- A' est l'ensemble des objets thématiques dans les données finales,
- B' est l'ensemble des objets topographiques dans les données finales,
- \mathcal{R}' est l'ensemble des relations existant entre les objets et qui traduisent le référencement des objets thématiques par rapport aux objets topographiques, dans les données finales.

Notre problème est le suivant :

- 1) Connaissant les données initiales thématiques (A) et les données topographiques (B), d'extraire et de caractériser (par des attributs) l'ensemble des relations \mathcal{R} .
- 2) Une fois les relations \mathcal{R} extraites, et connaissant les données topographiques (B'), d'en déduire (1) les relations \mathcal{R}' attendues dans l'état final, et (2) la localisation des objets thématiques dans l'état final, c'est-à-dire l'ensemble A' respectant les relations \mathcal{R}' avec B' :

$$A \mathcal{R} B \wedge B' \rightarrow ? \mathcal{R}' ? A' / A' \mathcal{R} B'$$

3 Processus général de migration de données thématiques permettant de modifier les relations

En figure 2, nous présentons une version enrichie du processus de migration de données thématiques. La première version était présentée par Jaara et al. (2012). Dans cette version, nous intégrons la phase de modification des relations.

Comme indiqué sur le schéma, la migration de données thématiques se fait en 4 étapes :

a) Extraction des relations : les types de relations porteuses de sens sont identifiés selon le cas d'application, puis les instances pertinentes de ces relations sont extraites des données initiales (ex. : l'accident $a1$ est sur la route $r1$ et l'accident $a2$ est proche du rond-point $rp1$). L'extraction et la modélisation des relations sont décrites plus en détail en 4.1 et 4.2.

b) Appariement : les données topographiques initiales et finales sont appariées afin d'explicitier les liens de correspondances et les différences entre objets initiaux et finaux (ex. un rond-point homologue d'un carrefour comme dans notre premier exemple-type illustré en figure 1). Le choix de la méthode d'appariement à utiliser doit être guidé par les caractéristiques des données topographiques considérées. Dans nos exemples-type, nous avons besoin d'une méthode d'appariement de réseau comme celle proposée par Mustière et Devogele (2008).

c) Identification des relations finales : les relations attendues à l'état final sont inférées. Plus précisément, les relations correspondant aux relations initiales sont identifiées en prenant en considération les objets topographiques finaux, et des règles de modification définies *a priori* en fonction du cas d'application considéré. Par exemple, dans nos exemples-type, une relation de proximité entre un accident et un rond-point devient une relation d'inclusion (relation *porté par*) avec le carrefour si le rond-point a pour homologue le carrefour, et que la distance initiale de l'accident au rond-point est suffisamment petite. L'étape d'identification des relations finales est décrite plus en détail en 4.3.

d) Relocalisation de données thématiques : les relations attendues dans l'état final sont utilisées pour contrôler le processus de relocalisation. Ce processus est décrit plus en détail en partie 5.

4 La modélisation des relations spatiales et leur modifications

4.1 Quelles sont les relations à extraire ?

Cette section donne une vue d'ensemble de types de relations considérés et des travaux existants sur leur modélisation, puis nous nous focalisons sur les particularités des relations thématique-référence. Plusieurs classifications des relations spatiales existent. Les relations sont classifiées par Ruas (1999) en trois catégories : relations topologiques, relations d'angularité et relations de voisinage. Pour Jones (1997), les relations sont classifiées comme relations topologiques, relations directionnelles ou relations de

proximité. Dans le modèle proposé par Jaara et al. (2012), les catégories sont classifiées dans trois catégories : topologiques, métriques (ex. : proximité, orientation) et autres (ex. : la forme de la route au voisinage d'un accident).

En ce qui concerne les relations topologiques, le modèle le plus connu est le modèle des 9 intersections par Egenhofer et al. (1991) pour les objets surfaciques, le même modèle a été utilisé pour les relations topologiques objet linéaire-objet linéaire et objet linéaire-objet surfacique par Egenhofer et Herring (1990). La formulation mathématique a été étudiée par Wallgrün (2012). Des exemples de modèles pour les relations directionnelles sont proposés par Papadias and Theodoridis (1997) pour deux objets surfaciques, Kurata and Shi (2009) pour un objet linéaire et un objet surfacique. Les relations de proximité sont souvent représentées par une distance. Dans la catégorie autre selon la classification de Jaara et al. (2012), nous avons des relations nommées « relations de mouvement » par Touya et al. (2012), telles que celles identifiées par Mathet (2000). Par exemple, la relation contourne : une rivière contourne une ville.

Lorsque l'un des deux objets en relation est thématique et l'autre topographique, nous constatons deux particularités :

- La première est que les relations ont une importance liée à la nature des données thématiques. En effet, certaines relations sont sémantiquement plus importantes pour certain type de données thématiques. Par exemple, si un accident de bateau a deux relations, une relation (accident de bateau-route de référence) et (accident de bateau-rivière de référence), l'accident de bateau est lié plus avec la rivière qu'avec la route, donc la relation avec la rivière est plus importante.
- La deuxième particularité est la relation particulière *porté par*, que nous définissons comme une relation topologique d'inclusion entre un objet thématique et un objet topographique avec lequel il y a lien sémantique privilégié. Par exemple, les accidents routiers sont *portés par* les routes : ils existent uniquement sur les routes.

4.2 Extraction des relations et leur modélisation

Les relations à extraire sont choisies en fonction du cas d'utilisation et des préférences de l'utilisateur. Les relations sont modélisées et stockées en utilisant un modèle de référencement dédié. Un modèle d'ontologie de relations et de contraintes relationnelles a été

proposé par Touya et al. (2012) dans le but d'unifier les définitions de relations et de contraintes de généralisation automatique. De ce modèle (présenté en *figure 3*), nous utilisons les éléments suivants :

- Les relations spatiales sont soit quantitatives (ex. : distance ou orientation) soit binaires, c'est-à-dire correspondant à un prédicat vérifié ou non (ex. : parallélisme, inclusion ou toute autre relation topologique).
- Les relations binaires sont extraites en testant si certaines conditions sont vérifiées. Les relations quantitatives ont des conditions de pertinence.
- Chaque relation a des propriétés. Par exemple, la relation de proximité entre un objet thématique ponctuel et un bâtiment peut être caractérisée par une propriété de distance entre l'objet et la bordure du bâtiment.
- Une relation a des membres qui sont les objets en relation.
- Les relations sont détectées et caractérisées par des méthodes dédiées.
- Dans le cas des relations binaires (correspondant à des prédicats), il existe des familles de relations telles qu'un couple d'objets ne peut pas vérifier en même temps deux relations de la même famille. Par exemple, dans le modèle de 4 intersections proposé par Egenhofer et al. (1991), nous considérons une famille de 8 relations : contient, inclus dans, couvre, couvert par, intersecte, égal, touche, disjoint.
- Il existe des cas où il n'est pas clair si une relation binaire est vérifiée ou non (par ex. un bâtiment est presque parallèle à une route), ou, dans le cas d'une famille de relations, il est difficile de dire clairement quelle relation parmi deux relations de la même famille est vérifiée. Par exemple, si deux objets surfaciques sont séparés par une distance trop petite en fonction des seuils de perception, il peut y avoir un doute sur le fait qu'ils doivent être considérés comme « disjoints » ou « se touchant ». Aussi, une relation floue peut être définie en lien avec chaque relation binaire (ou en lien avec deux relations binaires dans le cas des familles de relation). Par exemple, « presque parallélisme », « flou entre disjoint et se touchant ».

Dans le modèle présenté par Jaara et al. (2012) il s'agit d'explicitier et d'utiliser des objets caractéristiques, qui sont définis comme des objets additionnels qui aident à obtenir une meilleure description de la localisation des données thématiques. Les objets caractéristiques sont extraits dans les deux bases topographiques (l'initiale et la finale), puis ils sont appariés. Par exemple, les ronds-points et les carre-

fours sont des objets caractéristiques. Un rond-point peut être extrait de la base topographique initiale, ensuite, il peut être apparié avec un carrefour extrait de la base finale. L'explicitation d'objets caractéristiques (de même que l'extraction des relations) constitue une étape d'enrichissement des données, classiquement utilisée dans les processus complexes comme la généralisation. (Brassel et Weibel, 1988) appellent cette étape « reconnaissance de structures »).

Les relations ne lient pas seulement les objets thématiques et les objets topographiques. Comme le montre la figure 4, une relation peut lier un objet thématique avec des objets topographiques, et elle peut lier un objet thématique avec un objet caractéristique. Tout objet topographique ou caractéristique qui est impliqué dans une relation avec un objet thématique est appelé objet support.

Nous proposons une extension du modèle défini par Touya et al. (2012) présenté en figure 3, afin d'y inclure notre modèle de référencement de données thématiques. Avec cette extension, un élément spatial peut être caractéristique, topographique (appelé ObjBDTopo dans la figure 4) ou thématique. Chaque relation a deux membres, *Membre1* qui est un objet thématique et *Membre2* qui est un objet support (topographique ou caractéristique). La figure 5 montre la partie haut niveau du modèle résultant.

4.3 Modification des relations en fonction des données topographiques finales

Les modifications nécessaires sur relations sont identifiées en se basant sur l'appariement entre les bases topographiques initiale et finale. L'appariement peut mettre en évidence la disparition d'un objet topographique, un changement de taille, une agrégation (par exemple un bâtiment agrégé avec d'autres), une différence de représentation (par exemple un objet surfacique devient linéaire), un changement de position (cas du bâtiment dans la figure 1, deuxième exemple). Selon les transformations mises en évidence par l'appariement, certaines relations initiales doivent être légèrement modifiées par ajustement de leurs valeurs d'attributs (par ex. la distance entre le bout d'un bâtiment et un accident si le bâtiment est devenu plus grand dans la base topographique finale). Certaines relations doivent être transformées en une relation d'un autre type (par ex. si une rivière surfacique est devenue linéaire, les relations topologiques dans lesquelles elle était impliquée seront modifiées). Des relations peuvent être complètement ignorées, par exemple, une relation avec un bâtiment qui n'est pas représenté dans la base topographique finale.

L'identification des relations finales attendues en fonction des relations initiales est guidée par des règles. Les règles doivent être identifiées a priori pour chaque type de relation, en fonction des transformations que les objets topographiques sont susceptibles de subir. Il n'est pas nécessaire de connaître la base topographique finale pour définir les règles. Les règles peuvent être liées au message cartographique, en particulier pour les relations qui ont à voir avec les seuils de perception (ex. mentionné en partie 3 de la relation de proximité qui devient une relation « porté par »). Les règles de modification pour les relations topologiques peuvent être définies avec l'aide d'un graphe conceptuel de voisinage (Egenhofer et Mark, 1995). La figure 6 montre un extrait du graphe de voisinage défini par Egenhofer et Mark (1995) pour une ligne et un polygone, qui peut servir à modifier la relation topologique entre une route et un objet surfacique. Une règle basée sur ce graphe pourrait être : « si la relation initiale est (r3) et si la distance entre l'objet surfacique et l'extrémité de la route dans la base topographique finale est inférieure à 100m, alors la relation r sera changée vers (r5) ».

4.4 Besoin d'une évaluation de la satisfaction des relations

Après identification des relations finales attendues, l'objectif est de trouver la meilleure position pour l'objet thématique dans la base finale. Pour cela, nous avons besoin d'une méthodologie pour évaluer une position donnée d'un objet thématique donné dans la base finale. L'évaluation doit montrer le degré du respect des relations finales attendues. Nous proposons d'introduire une mesure de satisfaction d'une relation. Cette mesure quantifie à quel point une relation est vérifiée par une paire d'objets. La méthode de mesure dépend de la nature de la relation et de ses attributs.

Par exemple, pour la relation de proximité entre un accident et un bâtiment, une mesure de la satisfaction S de cette relation peut être définie de la façon suivante :

$$\text{Si } |d'-d| \geq 100m \quad S=0$$

$$\text{Si } |d'-d| \leq 30m \quad S=1$$

$$\text{Sinon} \quad S=1-((|d1-d2|-30)/70)$$

où d est la distance attendue dans l'état final, et d' est la distance correspondant à la localisation de l'accident à évaluer.

Pour les relations binaires, un graphe de voisinage peut aider à détecter le niveau de satisfaction. Si nous reprenons le cas des relations topologiques

ligne-région (fig. 6), on peut par exemple définir la mesure de satisfaction suivante :

Si la même relation existe pour la position donnée
 $S=1$

Sinon si la distance est une transition (ex. de r6 à r4)
 $S=0.5$

Sinon $S=0$

5 Relocalisation de données thématiques basée sur les relations

Après calcul des relations attendues entre les données thématiques et les données topographiques à l'état final, il faut relocaliser les données thématiques pour faire en sorte que les relations attendues soient satisfaites au maximum. Il s'agit d'un problème d'optimisation sous contraintes. La méthode de relocalisation que nous proposons est basée, pour chaque objet thématique à relocaliser, sur une recherche locale autour de sa position initiale en discrétisant l'espace possible. Elle comporte trois étapes :

1) *Détecter un ensemble de positions possibles pour l'objet thématique dans la base finale.* Si l'objet thématique est *porté par* un objet de référence, et si l'objet de référence est surfacique ou linéaire, il est discrétisé afin d'obtenir un ensemble de points. L'ensemble des points venant de la discrétisation est élagué en utilisant un buffer autour de la position initiale de l'objet thématique : cet élagage a pour but d'éliminer les positions trop éloignées donc impossibles. Si l'objet thématique n'est *porté par* aucun objet de référence, l'espace délimité par le buffer autour de la localisation initiale de l'objet thématique sera discrétisé pour obtenir des positions possibles.

2) *Choisir la meilleure position du point de vue des mesures de satisfaction des relations, en utilisant un système multicritère.* Le niveau de satisfaction de chaque relation est calculé pour chaque localisation possible résultant de la discrétisation issue de l'étape 1. Nous considérons la recherche d'une localisation optimale du point de vue des relations attendues comme un problème de décision multicritère, où les satisfactions des relations sont les critères. En effet, il n'est pas toujours possible de satisfaire toutes les relations à la fois. Nous constatons que le fait de s'autoriser à ignorer certaines relations afin de satisfaire les autres au mieux donne un meilleur résultat qu'essayer de trouver absolument un compromis entre toutes les relations.

D'après une étude sur les méthodes multicritère existantes, nous avons décidé d'utiliser la

méthode PROMETHEE II (Brans et Mareschal, 2005), parce qu'elle permet de favoriser certaines solutions selon des préférences données.

PROMETHEE II (Preference Ranking Organization Method for Enrichment Evaluations) est classée parmi les approches par agrégation partielle, ce qui signifie qu'elle compare chaque paire de solutions possibles selon l'ensemble des critères. Après agrégation, le résultat de cette méthode est un classement des positions possibles, ordonnées selon la satisfaction des relations attendues qu'elles permettent d'obtenir.

3) *Raffiner le positionnement au voisinage de la localisation trouvée lors de la 2^e étape.* Un voisinage de la position trouvée à l'étape 2 est discrétisé avec une équidistance plus petite, afin de trouver une solution meilleure que celle qui a été trouvée à la 2^e étape.

6 Application du processus sur le cas d'utilisation des accidents et des routes

Notre méthode a été implémentée dans CartAGen (Renard et al. 2010), module de généralisation de la plate-forme de recherche GeOxygene (Grosso et al. 2012). Cette partie illustre la méthode sur les deux exemples-types de la figure 2 (reproduite ci-dessous en figure 7), et montre les résultats obtenus.

La migration de données thématiques est appliquée comme suit :

a) *Extraction des objets caractéristiques.* Comme il est montré en figure 8, les ronds-points sont identifiés comme des objets caractéristiques. Ils sont extraits du réseau routier par recherche des faces rondes dans le graphe dérivé du réseau routier en utilisant l'indice de Miller, selon la méthode proposée par Sheeren (2005). Dans le deuxième exemple, l'objet caractéristique « intersection routière » est extrait en utilisant une simple intersection entre objets linéaires.

b) *Extraction des relations.* D'abord, nous devons identifier la route qui porte l'accident, ensuite, explorer le voisinage de chaque accident afin de détecter tous les objets support possibles (i.e. tous les objets topographiques ou caractéristiques pouvant avoir avec l'accident une relation traduisant son référencement par rapport aux données initiales). Dans l'exemple 1, le rond-point est trouvé comme un objet caractéristique support parce qu'il est proche de l'accident et adjacent à la route qui

porte l'accident. Dans l'exemple 2, l'intersection route-rivière et le bâtiment sont définis comme des objets supports possibles. L'étape suivante est d'extraire toutes les relations entre les accidents et les objets supports. Toutes les conditions de pertinence des relations sont testées, les relations extraites pour chaque accident sont les suivantes :

- Relation topologique *porté par* entre l'accident et une route, i.e. extraction de la route hôte (porteuse de l'accident).
- Relation topologique et relation de direction entre l'accident et le bâtiment.
- Relations de proximité point-point entre les accidents et les objets caractéristiques ponctuels.
- Relations de proximité point-rond-point entre les accidents et les ronds-points qui sont adjacents à la route hôte.

Dans l'exemple 1, les relations suivantes sont extraites :

- Deux relations *porté par*, une pour chaque accident avec les routes concernées.
- Deux relations de proximité point-rond-point, une pour chaque accident.

Dans l'exemple 2, les relations suivantes sont extraites :

- Une relation *porté par* pour l'accident avec la route concernée.
- Une relation de proximité point-point avec l'intersection route-rivière.
- Une relation de direction entre l'accident et le bâtiment.
- Une relation de proximité avec le même bâtiment.

c) *Extraction des objets caractéristiques finaux*: sont extraits, le carrefour pour le premier exemple, et l'intersection route-rivière pour le deuxième exemple (figure 9).

d) *Appariement*. Un appariement des données *topographiques* initiales et finales est d'abord réalisé. Dans notre implémentation, nous utilisons la méthode d'appariement automatique développée par Mustière et Devogele (2008) pour appairer le réseau routier, et un appariement simple basé sur un identifiant unique pour les bâtiments. Ensuite, les objets *caractéristiques* initiaux et finaux sont appariés. Dans nos exemples, un lien de correspondance est établi entre le rond-point et le carrefour dans le premier exemple, et un lien entre les deux intersections route-rivière dans le deuxième exemple.

e) *Identification des relations finales attendues*. Dans le premier exemple, le rond-point devient un carrefour. Selon les règles qui sont définies a priori sur les relations de proximité point-rond-point, si le rond-point devient un carrefour, la relation devient une relation de proximité point-point. Par ailleurs, parce que nous nous plaçons dans l'hypothèse où la carte résultante est dédiée à l'étude des accidents routiers, si un accident est trop proche d'un carrefour (ici 50 m), le carrefour doit être considéré comme porteur de l'accident. C'est le cas de l'accident *a1* ici, qui est à une distance inférieure à 50m du rond-point. Donc la relation finale attendue avec le carrefour est *porté par*. L'accident *a2* est à une distance supérieure à 50m du rond point, la relation attendue dans l'état final sera donc une relation de proximité avec le carrefour, caractérisée par une distance égale à la distance initiale jusqu'au bout de la route rond-point. Dans le deuxième exemple, les relations finales attendues seront identiques aux relations initiales.

Il est à noter que les modifications de relations effectuées sont dictées par l'usage de la carte finale (ici par hypothèse une étude d'accidentologie). Pour un autre cas d'utilisation tel qu'une carte de navigation en temps réel, il aurait été préférable dans le premier exemple d'augmenter la distance entre l'accident *a1* et le carrefour plutôt que de placer l'accident sur le carrefour, pour bien montrer que l'accident ne bloquait pas la totalité du carrefour.

f) *Relocalisation de données thématiques*. Les routes finales qui portent des accidents sont discrétisées avec un pas de 5 mètres sur leur partie située à moins de 300 mètres de la localisation initiale de l'accident. Les relations attendues sont évaluées pour chaque localisation possible ainsi définie. Les valeurs de l'évaluation sont injectées dans le système multicritère, qui identifie la meilleure solution. Une deuxième discrétisation avec un pas de 1 mètre est faite pour obtenir une localisation plus précise. La figure 10 présente le résultat obtenu, qui correspond au résultat attendu.

7 Conclusion

Le papier présente la suite des recherches sur la migration de données thématiques précédemment introduites par Jaara et al. (2012), où un premier modèle avait été proposé. L'idée de modifier les relations est introduite pour prendre en compte la modification des données de référence et le niveau de détail voulu à l'état final. Un nouveau processus pour la migration de données thématiques est présenté, prenant en considération la modification des rela-

tions. Le processus proposé pour la relocalisation des données thématiques est basé sur une méthode de décision multicritère, où les critères sont les niveaux de satisfaction des relations finales attendues. Le processus est enfin illustré par un cas d'utilisation sur des accidents routiers.

Le processus proposé doit maintenant être testé sur des données réelles, et il est important d'étudier les effets des différents paramètres. Dans un second temps, il serait souhaitable d'adapter le processus pour le cas de données thématiques linéaires.

Les relations formulées peuvent aussi être utilisées pour guider un processus de généralisation des données topographiques. Enfin, les relations peuvent être utilisées afin de contraindre la généralisation des données thématiques, par exemple pour forcer l'agrégation d'un ensemble d'accidents autour d'un rond point, en un seul objet ponctuel de type thématique porté par un carrefour.

Bibliographie

- Brando-Escobar C.**, 2012, *Un modèle d'opérations réconciliables pour l'acquisition distribuée de données géographiques*, Thèse de doctorat en SIG, Université Paris-Est Marne-la-Vallée, soutenue le 5 avril 2013.
- Brans J.P., Mareschal B.**, 2005, « PROMETHEE Methods ». dans Figueira J., Greco S., et Ehrgott M., éditeurs, *Multiple Criteria Decision Analysis: State of the Art Surveys*, p. 163-196, Springer-Verlag.
- Egenhofer M., Herring J.**, 1990, « Categorizing Binary Topological Relationships between Regions, Lines, and Points in Geographic Databases », *Technical Report 90-12*, National Center for Geographic Information and Analysis, University of California, Santa Barbara.
- Egenhofer M.**, 1991, « Reasoning about Binary Topological Relations », Second Symposium on Large Spatial Databases, Zurich, Switzerland, Gunther O, Schek H-J (eds.), *Lecture Notes in Computer Science*, Aug 1991, 525, Springer-Verlag, p. 143-160.
- Egenhofer M., Mark D.**, 1995, « Modeling Conceptual Neighborhoods of Topological Line-Region Relations », *International Journal of Geographical Information Systems* 9(5), p. 555-565.
- Grosso E., Perret J., Brasebin M.**, 2012, GEOXYGENE : « une plate-forme interopérable pour le développement d'applications géographiques », dans Bucher B., Le Ber F., *Développements logiciels en géomatique - Innovations et mutualisation*, Hermès Lavoisier, Chapitre 3.
- Jaara K., Duchêne C., Ruas A.**, 2012, « A model for preserving the consistency between topographic and thematic layers throughout data migration », 15th International Symposium on Spatial Data Handling, Bonn (Germany).
- Jones C.**, 1997, *Geographic Information Systems and Computer Cartography*, Longman.
- Kurata Y., Shi H.**, 2009 « Toward Heterogeneous Cardinal Direction Calculus », Mertsching B, Hund M, and Aziz Z. (ads.) KI, Paderborn, Germany, September 2009, *Lecture Notes in Artificial Intelligence*, 5803, p.450-457.
- Mathet Y.**, 2000, *Étude de l'expression en langue de l'Espace et du Déplacement : analyse linguistique, modélisation cognitive, et leur expérimentation informatique*, Thèse de Doctorat, Université de Caen, France.
- Mustière S., Devogele T.**, 2008, « Matching networks with different levels of detail », *Geoinformatica*, vol. 12, n. 4, p. 435-453.
- Papadias D., Theodoridis Y.**, 1997 « Spatial Relations, Minimum Bounding Rectangles, and Spatial Data Structures », *International Journal of Geographic Information Science*.
- Renard J, Gaffuri J, Duchêne C**, 2010, « Capitalisation problem in research - example of a new platform for generalisation: CartAGen », 12th ICA Workshop on Generalisation and Multiple Representation, Zürich (Switzerland).
- Ruas A.**, 1999, *Modèle de généralisation de données géographiques à base de contraintes et d'autonomie*, Thèse de doctorat, Université de Marne-la-Vallée, France.
- Sheeren D.**, 2005, *Méthodologie d'évaluation de la cohérence interreprésentation pour l'intégration de bases de données spatiales*, Mémoire de thèse.
- Touya G., Balley S., Duchêne C., Jaara K., Regnaud N., Gould N.**, 2012, « Towards an ontology of generalisation constraints and spatial relations », 15th ICA Workshop on Generalisation and Multiple Representation, Istanbul (Turkey).
- Touya G., Bucher B., Falquet G., Jaara K., Steiniger S.**, « Modelling geographic relationships in automated environments », Chapter 3 in D. Burghardt, C. Duchene and W.A. Mackaness (eds): *Abstracting Geographic Information in a Data Rich World*, 2014, LNG&C, Springer, Heidelberg. ISBN 978-3-319-00202-6
- Wallgrün J O.**, 2013, « Topological Adjustment of Polygonal Data », *Advances in Geographic Information Science*, Springer, p. 193-208

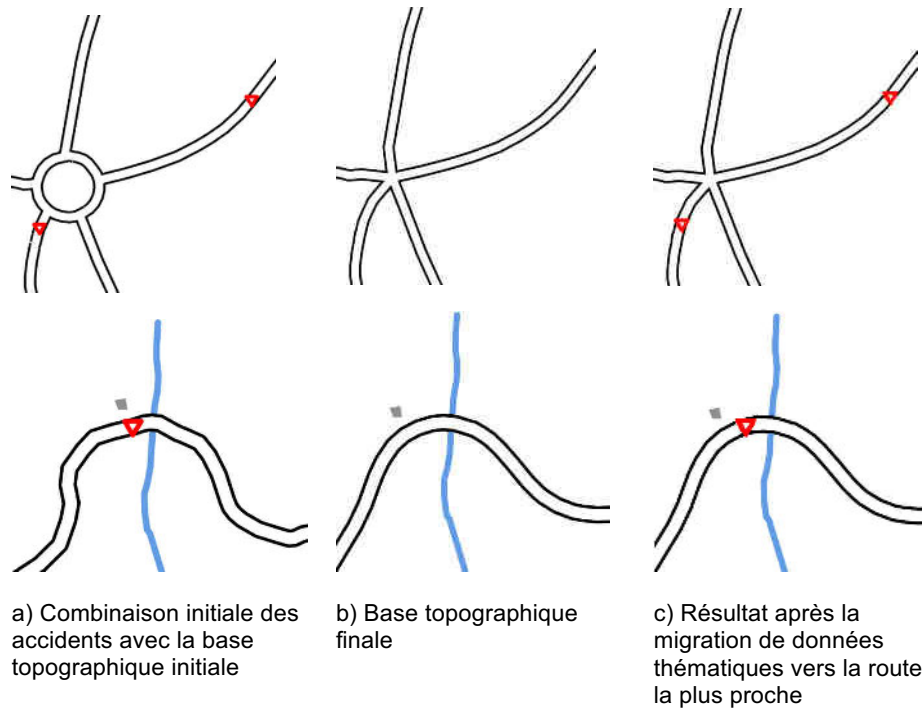


Figure 1 : Notre exemple-jouet : migration d'accidents sur un nouveau réseau routier

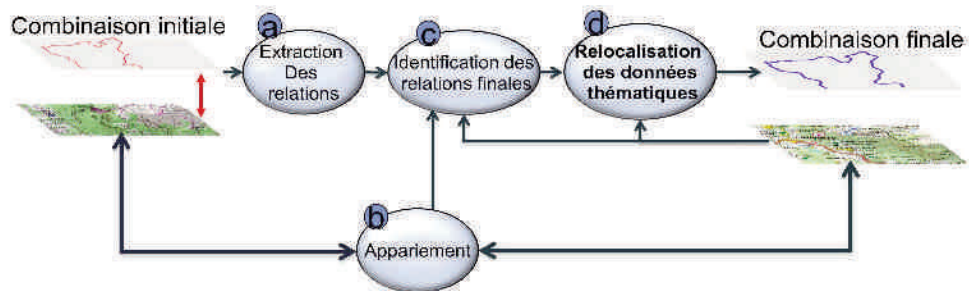


Figure 2 : Processus proposé pour la migration de données thématiques

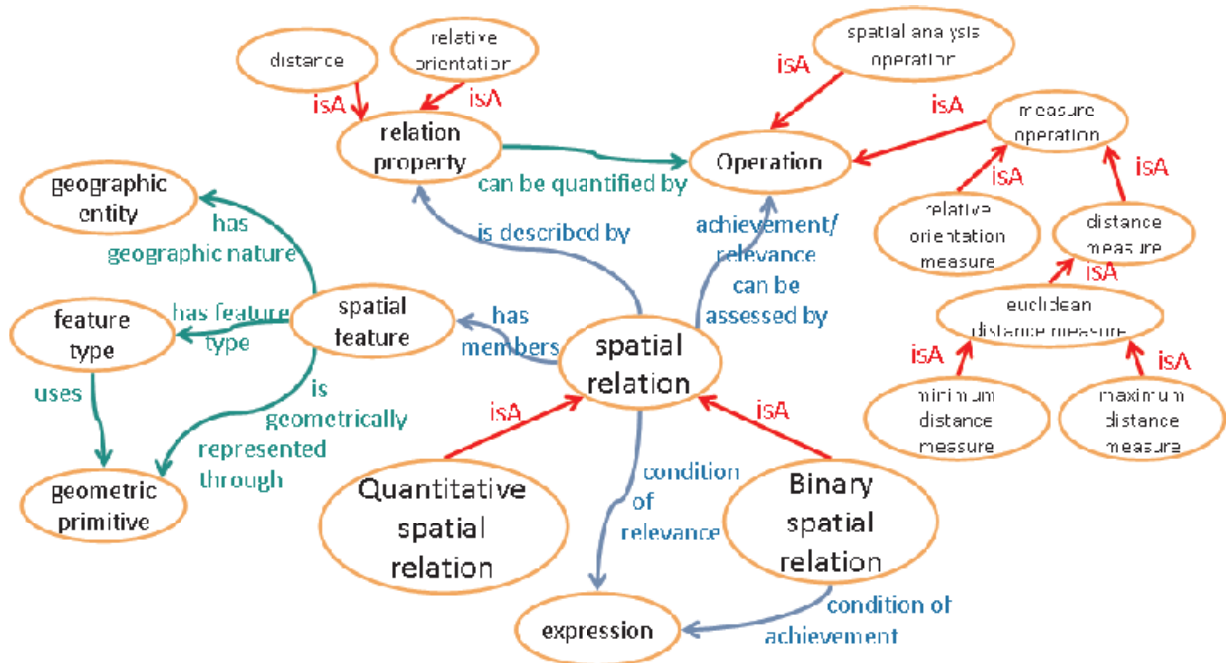


Figure 3 : Modèle pour les relations spatiales, proposé par Touya et al. (2012)

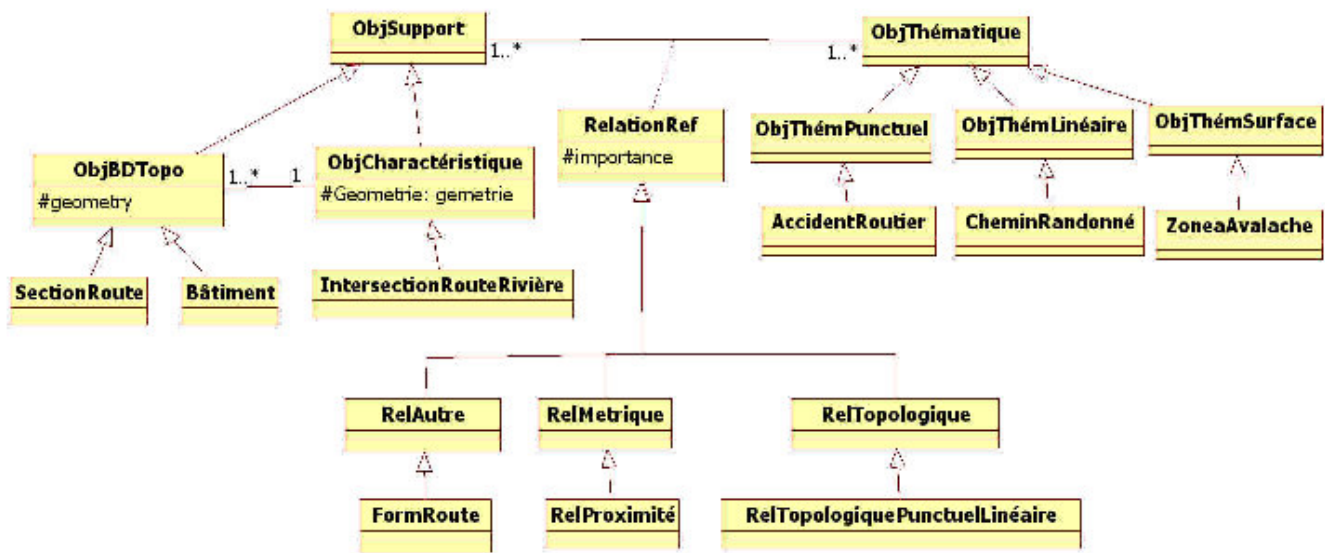


Figure 4 : Diagramme de classes du modèle de référencement proposé par Jaara et al. (2012)

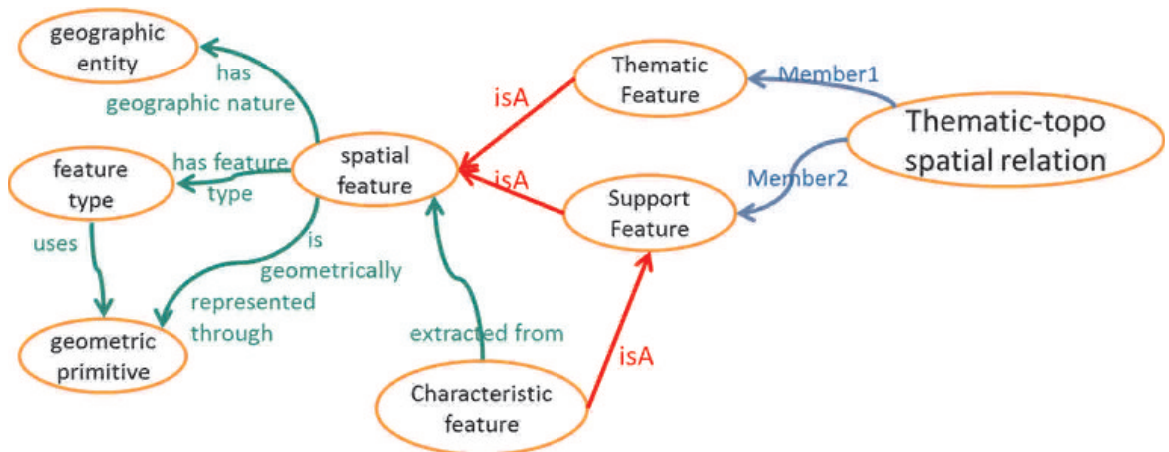


Figure 5 : Extension proposée de l'ontologie des relations définie par Touya et al. (2012)

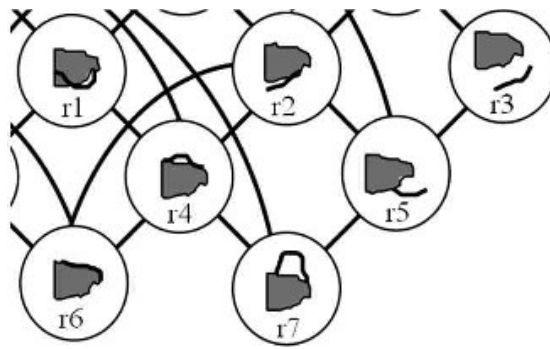


Figure 6 : Une partie du graphe conceptuel de voisinage des relations ligne-polygone (Egenhofer et Mark 1995)

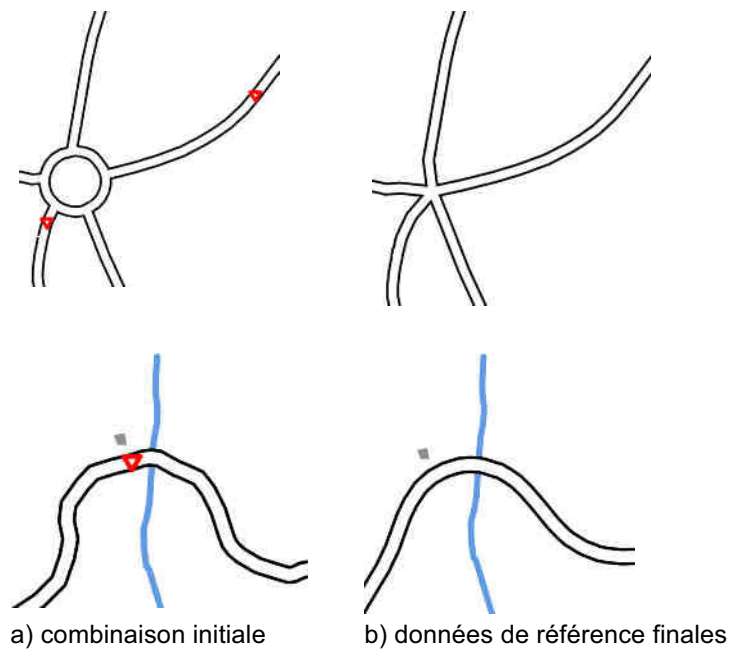


Figure 7 : Données de référence initiales et finales avec la localisation initiale des accidents

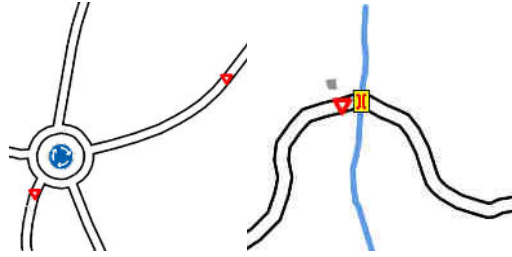


Figure 8 : Extraction des objets caractéristiques initiaux

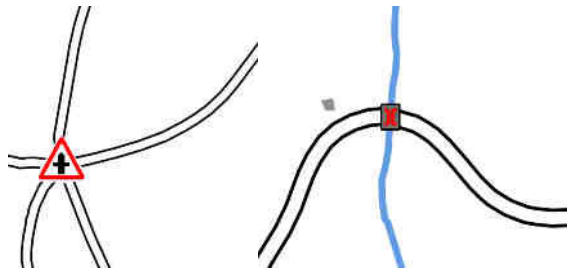
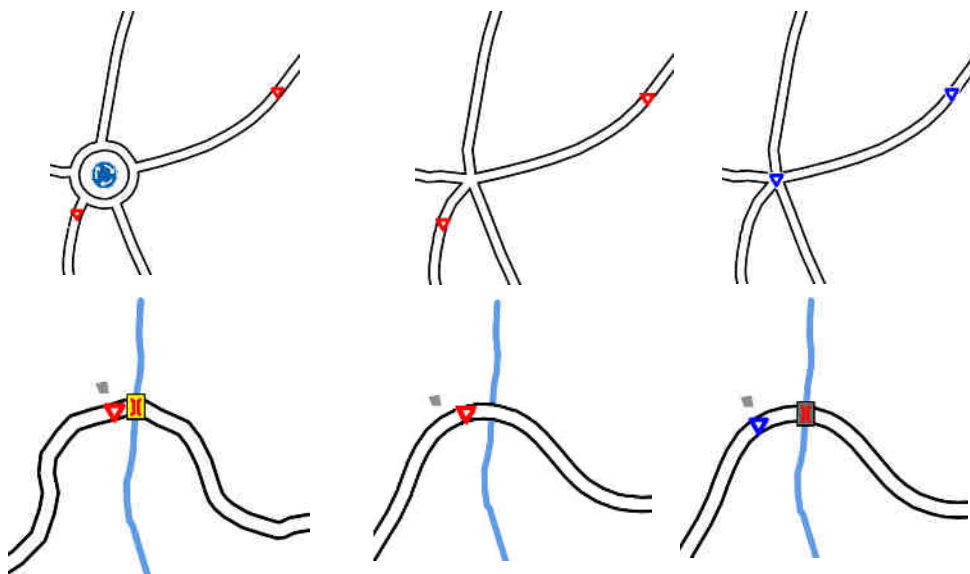


Figure 9 : Extraction des objets caractéristiques finaux



a) Combinaison des accidents avec la base topographique initiale

b) Combinaison finale après une migration vers le point le plus proche de la route

c) Combinaison finale en utilisant notre méthode de migration de données thématiques

Figure 10 : Résultat de migration de données thématiques