



Science Arts & Métiers (SAM)

is an open access repository that collects the work of Arts et Métiers Institute of Technology researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <https://sam.ensam.eu>
Handle ID: <http://hdl.handle.net/10985/15018>

To cite this version :

Meihan JIN, Christophe CLARAMUNT - Un modèle spatio-temporel sémantique pour la modélisation de mobilités en milieu urbain - Revue Internationale de Géomatique - Vol. 28, n°3, p.311-338 - 2018

Any correspondence concerning this service should be sent to the repository

Administrator : scienceouverte@ensam.eu



Un modèle spatio-temporel sémantique pour la modélisation de mobilités en milieu urbain

Meihan JIN¹, Christophe Claramunt²

1. School of Architecture and Urban Planning, Harbin Institute of Technology
HIT Campus, The University Town of Shenzhen,
Xili, Nanshan District, Shenzhen, P.R.China
jmeihan@hotmail.com

2. Institut de Recherche de l'Ecole Navale
IRENav, BCRM de Brest, Ecole navale
CC 600 - Lanveoc
christophe.claramunt@gmail.com

RESUME. La croissance rapide et la complexité de nombreuses villes contemporaines offrent de nombreux défis de recherche pour les scientifiques à la recherche d'une meilleure compréhension des modèles de mobilité qui se produisent dans l'espace et dans le temps. A l'heure où de très grandes séries de données de trajectoires en milieu urbain sont disponibles grâce à la profusion de nombreux capteurs de positionnement et de services, de nouvelles opportunités sont offertes à la recherche en géomatique. Cependant, une bonne intégration des données de mobilité nécessite au préalable l'élaboration de cadres méthodologiques et conceptuels tout comme des bases de données spatio-temporelles qui offriront les capacités de représentation et de manipulation des données. La recherche présentée dans cet article de synthèse développe une modélisation conceptuelle et une approche de gestion de base de données pour représenter et analyser les trajectoires humaines dans des espaces urbains. Le modèle considère les dimensions spatiales, temporelles et sémantiques afin de tenir compte de l'ensemble des propriétés issues des informations de mobilité. Plusieurs abstractions de données de mobilité et des outils de manipulation de données sont développés et expérimentés à partir d'une large base de données de trajectoires disponibles dans la ville de Pékin. L'intérêt de l'approche est double: premièrement, il montre que de larges ensembles de données de mobilité peuvent être intégrés au sein de SGBD spatio-temporels extensibles; deuxièmement, des outils de manipulation et d'interrogation spécifiques peuvent être dérivés à partir de fonctions intégrées au sein d'un langage d'interrogation. Le potentiel de l'approche est illustré par une série d'interrogations qui montrent comment à partir d'une large base de données de trajectoires quelques patrons de comportements peuvent être obtenus.

ABSTRACT. The continuous development and complexity of many modern cities offer many research challenges for urban scientists searching for a better understanding of mobility

patterns that happen in space and time. Today, very large trajectory datasets are often publicly generated thanks to the availability of many positioning sensors and location-based services. However, the successful integration of mobility data still requires the development of conceptual and database frameworks that will support appropriate data representation and manipulation capabilities. The research presented in this paper introduces a conceptual modeling and database management approach for representing and analyzing human trajectories in urban spaces. The model considers the spatial, temporal and semantic dimensions in order to take into account the full range of properties that emerge from mobility patterns. Several object data types and data manipulation constructs are developed and experimented on top of an urban dataset testbed currently available in the city of Beijing. The interest of the approach is twofold: first, it clearly appears that very large mobility datasets can be integrated in current extensible GIS; second, significant patterns can be derived at the database manipulation level using some specifically developed query functions.

Mots-clés :Trajectoires urbaines, Modélisation spatio-temporelle, Bases de données spatio-temporelles.

KEYWORDS: Urban trajectories; Spatio-temporal data modeling; Spatio-temporal database

1. Introduction

Les villes ont toujours été des environnements dynamiques et complexes. En effet, les villes sont le lieu d'un grand nombre d'activités humaines et d'interactions dont leur exploration et leur compréhension sont loin d'être aisées. En particulier, les grandes villes modernes génèrent de nombreux déplacements et des flux de trafic qui nécessitent encore le développement de cadres de modélisation de données spatio-temporelles appropriés afin de fournir aux planificateurs urbains des outils permettant une meilleure compréhension des patrons de déplacement.

L'étude des dynamiques de mobilité en milieu urbain est un sujet de recherche essentiellement abordé depuis le début des années 1980 (Hanson & Huff, 1986; Jiang et al., 2012; Pentland & Liu, 1999). L'une des principales raisons de l'émergence de ce courant de recherche est une conséquence du constat que la prédiction des comportements de déplacement des individus pourrait être un élément essentiel d'une meilleure planification des transports urbains et de l'analyse des politiques publiques de la ville (Kitamura, 1988). Un objectif important de la recherche d'une meilleure compréhension des dynamiques de mobilité passe par le développement d'une capacité de représentation et de prédiction des mobilités afin d'étudier comment en particulier les individus réagissent aux changements dans leur environnement. Cette recherche d'une meilleure compréhension des mobilités des individus reste encore une tâche difficile dans la mesure où cela implique non seulement des processus décisionnels complexes, mais également une série de contraintes sémantiques, spatiales et temporelles. A ce stade, le développement de méthodes et de modèles conceptuels d'exploration de données de mobilité pour une meilleure analyse et compréhension des comportements humains est toujours considéré comme un défi scientifique.

Plusieurs chercheurs ont successivement introduit des cadres méthodologiques et conceptuels pour représenter et analyser des comportements de mobilité (Alvares et al., 2007; Hanson & Huff, 1982; Fernandez & Yuan, 2000; Mennis & Peuquet, 2000; Yuan & Goodchild, 2007; Renso et al., 2013). D'après Andrienko et al. (2008), un mouvement est défini comme une généralisation de plusieurs primitives géométriques et sémantiques dérivées à partir de trajectoires individuelles, et qui peut être analysé à différents niveaux d'abstraction pour en dériver de la connaissance. A un niveau conceptuel, une trajectoire peut être représentée comme l'évolution de la position d'un objet en mouvement dans l'espace pour un intervalle de temps donné et avec un objectif décrit par ses caractéristiques sémantiques (Spaccapietra et al., 2008). Au niveau applicatif, plusieurs taxonomies ont été introduites pour caractériser des patrons de trajectoires (Dodge et al. 2008).

Plusieurs algorithmes géométriques ont également été proposés pour décrire des mobilités. En particulier plusieurs approches de « clustering » de trajectoires ont été introduites par différents auteurs (cf. Nanni, 2013). Un patron de trajectoire sera défini par un ensemble de trajectoires qui recèlent des similarités spatiales, temporelles et sémantiques. Cette notion de similarité s'appréhende par exemple à partir d'une mesure de distance sémantique (Parent et al. 2013). Par ailleurs, Laube introduit plusieurs algorithmes de détection de patrons de mouvement à partir de la

construction d'une matrice de mouvements et de contraintes de distance (Laube, 2005). Une approche plus spécifique appliquée à des mobilités urbaines a été introduite par Buchin et al. (2011). Cette approche est basée sur un calcul de la distance de Fréchet dérivée de deux trajectoires pour un même intervalle de temps donné, et ce, afin d'en dériver s'il y a lieu des patrons de similarité. Hung et Peng (2009) proposent une approche basée sur les probabilités et un algorithme de "clustering" également basé sur une fonction de distance. Une K-distance dérivée du K-voisin tel que défini par (Cover & Hart, 1967) peut également constituer la fondation d'un algorithme de « clustering » (Ong et al. 2010). Globalement, non seulement des patrons de mouvements peuvent être identifiés à partir de données de mobilités, mais aussi des trajectoires anormales ou particulières.

Les patrons de mouvement doivent idéalement révéler toute régularité spatiale et temporelle reconnaissable, sous réserve d'un intérêt qui dépend du domaine d'application (Dodge, 2008). Les irrégularités sont également des tendances à explorer. Par exemple, la découverte de déplacements aléatoires et particuliers au sein de larges bases de données de trajectoires est une autre direction qui a été explorée par Gonzalez et al. (2008). Dans de nombreux cas, les humains ont tendance à effectuer des déplacements routiniers, c'est un premier contexte d'exploration, un deuxième serait celui d'analyser des comportements particuliers ou même anormaux. Dans le cas des comportements réguliers, les contraintes institutionnelles, culturelles et obligations qui influencent les mobilités humaines jouent un rôle déterminant (Buliung et al. 2008, Schlich & Axhausen, 2003). Schlich a montré, qu'en général, l'étude de similarités dans les déplacements humains à travers de longues périodes peut révéler des tendances significatives (Schlich & Axhausen, 2003). De telles tendances émergent en général de l'étude de déplacements journalier, voire de différences entre patrons de mobilité en semaine ou en week-ends. Ces aspects ont été étudiés dans de nombreuses villes canadiennes et européennes (Axhausen et al., 2002; Buliung et al., 2008; Susilo & Axhausen, 2007). D'autres études ont montré une relative instabilité de patrons de déplacements durant les fins de semaines (Srivastava & Schönfelder, 2003; Susilo & Kitamura, 2005).

Bien que de nombreux travaux antérieurs aient notamment montré l'existence de comportements de routine dans les mobilités urbaines, il reste encore nécessaire de développer des modèles à la fois sémantiques et spatio-temporels permettant d'intégrer des données de mobilités et les associer à une description fine de leurs caractéristiques. De telles approches devraient permettre non seulement de donner un cadre précis et manipulable de ces données, mais aussi de faciliter une plus large palette de manipulation et d'analyse à partir du moment où cet ensemble sera intégré au sein d'un SGBD extensible disposant de tous les outils de manipulation et de visualisation, et ce, à différents niveaux d'abstraction et d'échelle dans l'espace et le temps. L'objectif d'une telle approche consiste à pouvoir aborder des questionnements tels que l'étude des comportements habituels des usagers de la ville à différents niveaux d'abstractions dans l'espace et dans le temps, par quartier, par période temporelle, ou l'étude des modalités de déplacements et leur prise en compte par les réseaux de transport existants. Ces quelques exemples illustrent le

type d'analyse qu'un tel modèle de base de données spatio-temporelle de trajectoires doit pouvoir aborder. Pour ce faire, nous proposons de combiner une approche de modélisation conceptuelle avec une démarche d'implémentation au sein d'un SGBD spatio-temporel extensible et une expérimentation appliquée à une large base de données de trajectoires en milieu urbain extraite du projet Geolife développé ces dernières années dans la ville de Beijing en Chine.

La suite de l'article est organisée en plusieurs sections. La section 2 développe l'approche de modélisation conceptuelle de représentation de mobilités urbaines à partir de plusieurs modules complémentaires qui permettent de caractériser les propriétés sémantiques, spatiales et temporelles. La section 3 propose une série de manipulations appliquées à ces trois dimensions et qui permettent d'illustrer le potentiel de l'approche. Une implémentation à partir du SGBD spatio-temporel Postgres/PostGIS et expérimentation aux données Geolife de la ville de Beijing est développée dans la section 4. L'article se termine par une conclusion dans la section 5 qui résume l'apport et quelques perspectives pour de futures recherches et applications.

2. Approche de modélisation

Notre objectif de modélisation consiste à concevoir un modèle générique de base de données sémantique et spatio-temporelle pour la représentation de mobilités en milieu urbain. La composante sémantique du modèle décline la notion de trajectoire telle qu'elle apparaît en domaine urbain, elle se complète de caractéristiques spécifiques permettant de qualifier également les dimensions spatiales et temporelles.

2.1. Modèle de mobilité

Nous proposons un schéma conceptuel de base de données des mobilités à partir de plusieurs schémas étroitement liés les uns aux autres (appelés modules ci-après). Les principes de conception de ce modèle sont inspirés d'une ontologie de trajectoire proposée par Yan et al. (2008), la construction du modèle est développée selon les principes de l'outil de conception MADS (Modeling Application Data with Spatio-temporal features) développé par Spaccapietra et Parent (2006). Ces trois schémas complémentaires sont les suivants:

- Un module Trajectory qui caractérise les déplacements urbains à partir du réseau de transport urbain.
- Un module Transportation Network qui représente les différents réseaux de transport de la ville.
- Un module Travel Pattern qui qualifie la sémantique des mobilités en milieu urbain.

Ces trois modules (Trajectory module, Transportation Network module, and Travel Pattern module) ne sont pas indépendants dans la mesure où les entités identifiées dans chacun des modules peuvent être en relation avec des entités

d'autres modules. L'union de ces trois modules et schémas donne un modèle global de représentation Travel Behavior schema des mobilités dans la ville (Figure 1).

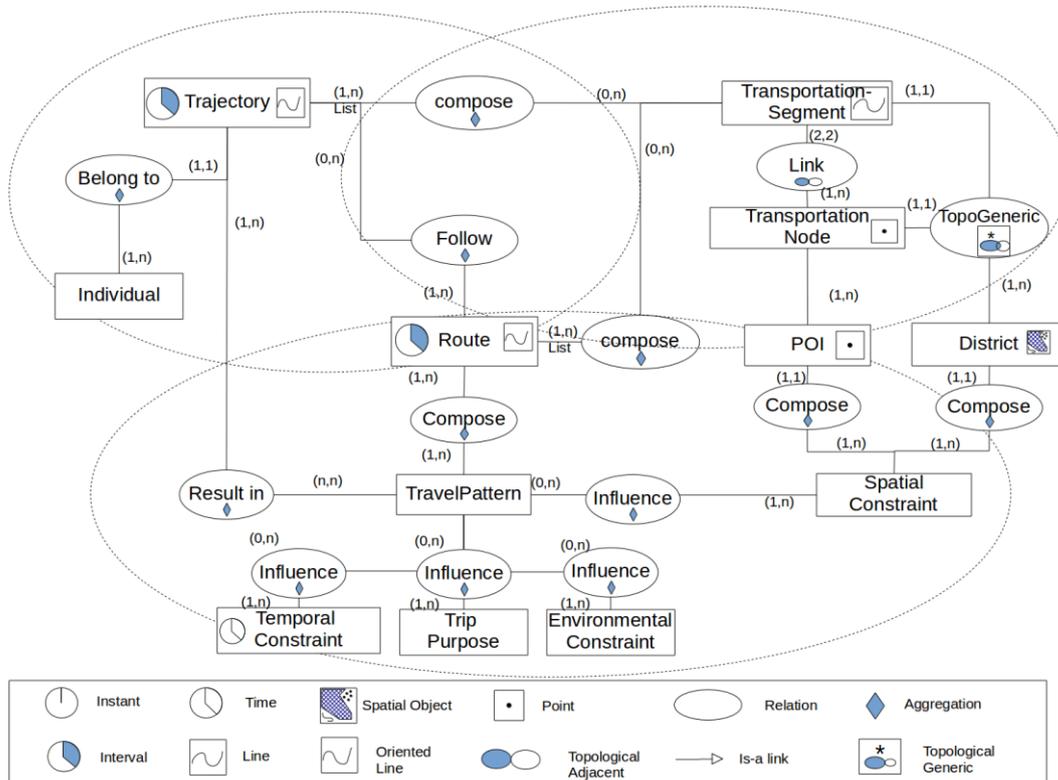


Figure 1 Modèle de représentation des mobilités

2.2. Module Trajectoire

Le module trajectoire est basé sur le modèle initialement proposé par Spaccapietra et al. (2008), dans lequel le concept de trajectoire est basé sur une séquence d'arrêts et de déplacements, comme suggéré par les concepts issues de la Time Geography (Hägerstrand, 1970). Nous considérons et nous nous limitons dans notre approche aux mobilités humaines qui ont lieu le long d'un réseau de transport, une trajectoire est donc ici indivisible d'un réseau de transport. Alors que les premières études en géographie humaine étudiaient les comportements de transport au niveau agrégé, un changement conceptuel est apparu au début des années 70 avec l'émergence de la théorie de la Time Geography (Hägerstrand, 1970). Alors qu'il étudiait les migrations humaines en Suède, Hägerstrand développa une base conceptuelle où les comportements humains sont modélisés au niveau désagrégé.

Selon les principes identifiés par la Time Geography, un individu est associé à un niveau de granularité donné à une notion de trajectoire et d'activité qui représentent une unité de modélisation.

Plus formellement une *Trajectory* représente un objet spatio-temporel. La relation *Compose* entre *TransportationSegment* et *Trajectory* est une aggregation de type *List*. Une *Trajectory* est une sequence de *one-to-many TransportationSegments*. Une *Trajectory* est décrite par un attribut temporel, *Duration*, spécifié par un intervalle et dérivé respectivement par les instants caractérisant le premier noeud et le dernier noeud du premier *TransportationSegment* TR_1 et dernier segment *TransportationSegment* TR_n de cette trajectoire *Trajectory*.

Définition 1: Individual

Un *Individual* représente un humain qui se déplace le long d'un réseau de transport et réalise quelques activités dans la ville. Un *Individual* est modélisé par plusieurs attributs qui qualifient son identité et plusieurs de ses caractéristiques notamment son nom *Name* typé par un *String*, *Age* typé par un *Integer*, *Gender* typé par un *String* et deux valeurs possibles (i.e., *Male* or *Female*), *Occupation* typé par une *Enumeration* (i.e., liste de possibles emplois) et *FamilyRole* typé par une *Enumeration* (e.g., mère, père, enfant).

Définition 2 Trajectory

Une *Trajectory* se réalise à partir de plusieurs composants du réseau urbain et en particulier des types de transport utilisés. Une *Trajectory* est explicitement représentée par une séquence de points qualifiés temporellement (instant, latitude, longitude and altitude), et plus précisément modélisé par une séquence ordonnée de *TransportationSegments* : $Trajectory_i = [TS_1, TS_2, \dots, TS_n]$ où TS_1, TS_2, \dots, TS_n représente une séquence de *TransportationSegments*. Notons qu'un *TransportationSegment* peut être parcouru par plusieurs individus et par plusieurs trajectoires. Une *Trajectory* a implicitement un noeud de départ et un noeud d'arrivée respectivement donnés par le noeud de départ de son premier et dernier *TransportationSegments*.

2.3. Module Transportation Network

Le module *Transportation Network* représente le réseau urbain (réseau des rues, réseau de tramway, ...).

Définition 3: TransportationNode

Un *TransportationNode* modélise une intersection de ligne de transport ou un *Point d'Intérêt (POI)*. Une intersection de ligne de transport matérialise une jonction entre plusieurs *TransportationSegments* (au moins deux). Un *POI* est un point d'intérêt matérialisé dans le réseau par une approximation de sa localisation.

Le type d'objet *TransportationNode* a un attribut spatial *Node* défini comme une géométrie de type *Point* qui représente sa localisation dans le réseau de transport. Un attribut Boolean *IsStop* distingue un *TransportationNode* qui modélise une intersection du réseau de transport (i.e., valué par 1) ou un *POI* (i.e., valué par la

valeur nulle). *IsRoad*, *IsMetro* et *IsRailway* sont les trois attributs qui permettent de distinguer les réseaux de transport sous-jacent auquel un noeud appartient.

Définition 4: TransportationSegment

Un *TransportationSegment* modèle un lien entre deux *TransportationNodes* dans le réseau de transport. Un *TransportationSegment* est valué par un attribut spatial (i.e., de type segment de ligne) le long du réseau de transport, et terminé par deux *TransportationNodes*. La longueur de ce segment de ligne est aussi représentée par un attribut. L'attribut *Type* définit les différents types de *TransportationSegment*, c'est à dire route, metro ou train. Nous assumons que quand plusieurs *TransportationSegments* de différents *Types* se superposent ils sont représentés comme des objets différents. La relation entre un *TransportationSegment* et un *TransportationNode* est aussi définie comme une relation topologique *Touch* entre un segment de ligne et un point. Un *TransportationNode* est en relation avec one-to-many *TransportationSegment* alors qu'un *TransportationSegment* doit avoir deux *TransportationNodes* qui représentent respectivement l'origine spatiale et la destination de ce segment.

En sus des différents concepts et abstractions de modélisation présentés par la Figure 1, une contrainte spatio-temporelle est prise en compte afin de garantir la connectivité du réseau (d'autres contraintes dépendantes de l'application pourraient être définies selon les mêmes modalités):

Contrainte 1

Les successifs *TransportationSegments* qui composent une *Route* doivent être connectés, c'est à dire, soit le noeud de départ *TransportationNode* ou/et le noeud d'arrivée *TransportationNode* de ce *TransportationSegment* doivent être soit le noeud de départ ou de fin *TransportationNode* d'un autre *TransportationSegment* de cette *Route*.

2.4. Module Travel Pattern

Le module *TravelPatterns* permet de qualifier les primitives de déplacement extraites des trajectoires.

Définition 7: Route

Une *Route* représente un chemin suivi régulièrement par un ou plusieurs *Individuals*. Nous considérons un *Individual_i* en déplacement le long du réseau de transportation et qui donc génère une *Trajectory*. Une *Route* est généralisée à partir de *Trajectories* régulières. Une *Route* démarre à un *Node* et se termine à un *Node*. Les *TransportationSegments* et *TransportationNodes* qui constituent cette route sont implicitement ordonnés. Plus formellement une *Route* est modélisé comme une séquence de *TransportationSegments*. Afin de modéliser la spatialité d'une *Route*, nous introduisons le concept de *Path* défini comme une séquence de *TransportationSegments*.

$$Path_i = [TransportationSegment_1, TransportationSegment_2, \dots, TransportationSegment_n];$$

Afin de représenter des déplacements réguliers, nous introduisons la notion d'intervalle régulier qui représente des cycles d'intervalles ou en d'autres termes un intervalle temporel qui se répète :

$TemporalCycle = \langle Regularity, Regular\ Interval \rangle$

Un exemple de *TemporalCycle* : $TimeSegment = \langle workdays, 08h-09h \rangle$.

Une *Route* est donc définie par:

$Route_i = \langle Path_i, TemporalCycle_i, TransportationMode \rangle$

Alors qu'un *Individual* réalise une *Trajectory* donnée une seule fois, il peut réaliser un *Path* plusieurs fois, et une même *Route* régulièrement.

Définition 8: TravelPattern

Une *Route* d'un *Node* N_1 vers un noeud *Node* N_m peut être partagée par plusieurs *Individuals* $\{i_1, \dots, i_n\}$. Une *Route* est le résultat implicite de plusieurs contraintes *Constraints* (*BehaviorActivity*, *TemporalConstraint*...). Au final, des *Routes* peuvent générer des patrons de déplacement *TravelPatterns* lorsqu'elles sont réalisées régulièrement ou même rarement (ce qui permettra d'étudier des tendances ou des particularités de déplacement).

Un même ensemble de *Trajectories* peut générer plusieurs *Routes* dont l'analyse va permettre d'étudier plusieurs modalités de patrons de déplacement *TravelPatterns*. Alors que les *Trajectories* et les *Routes* sont explicitement représentés dans l'espace et dans le temps, les patrons de déplacement *TravelPatterns* sont dérivés de *Routes* à partir de contraintes de fréquence.

$TravelPattern = \langle Routine: [\langle Route_1 \rangle, \langle Route_2 \rangle, \dots, \langle Route_n \rangle], TemporalCycle, Frequency, TypeNumber \rangle$

Un patron de déplacement *TravelPattern* devrait de fait être la conséquence de plusieurs facteurs influençant des décisions et facteurs de déplacement d'un ou plusieurs *Individuals*. Par exemple, considérons l'exemple illustratif d'un jour de la semaine et différentes conditions météorologiques. Selon que le temps soit pluvieux ou pas un *Individual* choisira de se déplacer en voiture ou à pieds pour se rendre à son travail distant de quelques centaines de mètres. Ces contraintes et propriétés sont spécifiées par des attributs qui qualifient un patron de déplacement *TravelPattern* soit les attributs qui permettent d'identifier l'*Individual* et ces différentes contraintes. Ces contraintes incluent les contraintes environnementales (conditions météorologiques, paysages, état de la route ...), les contraintes temporelles (matin, après-midi, jours de la semaine ...), les contraintes spatiales (localisation, voisinage, points d'intérêts ...) et le but du déplacement (loisir, travail ...).

Généralement, un patron de déplacement *TravelPattern* sera généré par des contraintes internes et externes. Ce sont ces contraintes et ces influences qui vont constituer la trame des caractéristiques comportementales de notre approche de modélisation.

Les comportements de déplacements des *Individuals* traduisent souvent une large diversité d'actions et d'activités, reflets de différentes caractéristiques socio-culturelles et économiques. Bien que notre objectif consiste à développer une approche de modélisation permettant de prendre en compte les possibles propriétés

influant des patrons de déplacement – un objectif très ambitieux – et afin d’illustrer le potentiel de notre approche, nous choisissons de retenir quelques propriétés élémentaires et attributs d’*Individual* souvent utilisés en sciences humaines comme le sexe, l’âge et l’emploi. Selon Hägerstrand (1970), les comportements humains de déplacement sont souvent contraints par des obligations sociales. Il est donc raisonnable de considérer lors de la modélisation d’un profil personnel de catégoriser socialement les *Individuals* afin de mieux générer des classes provenant de groupes homogènes et potentiellement exprimant des comportements convergents.

Par exemple, considérons les différentes caractéristiques économiques et occupations d’une famille. Les modalités de transport sont susceptibles de varier selon les catégories sociales, les emplois occupés, les activités réalisées et la typologie du foyer. Dans la mesure où un même *Individual* est susceptible de réaliser plusieurs *Trajectories* sur une longue période de temps, ce même *Individual* est susceptible de générer plusieurs patrons de déplacement *TravelPatterns* sur cette même période.

Plusieurs facteurs influencent des décisions de déplacement d’un même *Individual* générant des patrons de déplacement *TravelPattern*. Nous classifions ces facteurs en quatre catégories qui prennent en compte des contraintes temporelles *TemporalConstraint*, des typologies d’activité *BehaviorActivity*, des contraintes environnementales *EnvironmentConstraint*, et des contraintes spatiales *SpatialConstraint* (d’autres catégories de contraintes pourraient être définies selon les mêmes modalités). Comme précisé plus haut, plusieurs caractéristiques sociales, telles par exemple que celles appréhendées par notre modèle, peuvent jouer un rôle dans la façon qu’aura un *Individual* de se comporter et se déplacer dans la cité. Dans l’instanciation de notre modèle et pratiquement un *Individual* est susceptible de « posséder » plusieurs *Trajectories* et ces *Trajectory* sont aussi susceptibles de générer des patrons de déplacements *TravelPatterns*. Ces catégories de contraintes matérialisent différents facteurs qui influencent ces comportements de déplacement. Dans le cas des contraintes spatiales *SpatialConstraint* nous considérons notamment des POInts d’Intérêt *POIs* et des quartiers *District*. Les *POIs* matérialisent des lieux qui présentent un intérêt pour un *Individual* ou des points de repère ou de convergence de ces déplacements (e.g., un supermarché, un bureau de tabac), ces *POIs* jouant le rôle de contraintes spatiales. Un autre exemple de contrainte spatiale provient de cette notion de *District* dont le rôle consiste notamment à caractériser les origines et les destinations des trajectoires de déplacement, cela potentiellement permettant d’étudier des phénomènes de déplacement origine-destination à différents niveaux d’agrégation dans la cité, les tendances générales et les modalités de transport utilisées, pour ne citer que quelques exemples de manipulation possible que ce type de modélisation permettra.

Considérons une deuxième catégorie de contrainte à savoir les contraintes temporelles *TemporalConstraints* comme des constituants de patrons de déplacement *TravelPattern*. Ces contraintes temporelles *TemporalConstraints* peuvent caractériser différentes périodes temporelles comme jour/nuit, heure de pointe, jours de semaine/jours de fin de semaine, journée de travail/journée de vacances etc. Ces types de contraintes temporelles jouent un rôle important dans

l'expression et la réalisation de patrons de déplacement. Ces contraintes temporelles *TemporalConstraints* peuvent être spécifiées selon les contraintes de l'application à partir de différents attributs et granules temporels.

Une activité comportementale *BehavioralActivity* est utilisée pour représenter le but du déplacement d'un *Individual*. Prenons l'exemple d'un déplacement sur son lieu de travail, dans ce cas la contrainte *BehavioralActivity* pourra être représentée par une valeur 'Travail' dans le module *TravelPattern*. Les activités d'un *Individual* peuvent être par exemple classées en plusieurs sous-catégories comme *PersonalActivity*, *FamilyActivity* and *BusinessActivity*. Typiquement ces différentes catégories d'activités représentent des classes de comportement bien particulières. Généralement. Pour une activité *BusinessActivity*, minimiser le temps de déplacement est souvent un critère déterminant à appliquer ce qui ne sera pas forcément ou en tout cas pas avec la même magnitude dans le cas d'une *FamilyActivity*.

Le contexte environnemental est également une contrainte qui impacte les décisions et donc les patrons de déplacement *TravelPatterns*. Nous classifions en première approximation ces contraintes environnementales *EnvironmentalConstraints* en deux catégories: celles issues de considérations sociales *SocialEnvironment* et celles résultant de l'environnement naturel *NaturalEnvironment*. Les contraintes sociales *SocialEnvironment* modélisent des facteurs sociaux influent sur les décisions ou les conditions de déplacements alors que les contraintes environnementales prennent en compte elles des paramètres physiques de l'espace comme les conditions météorologiques, topographiques etc.. Toutes ces contraintes, à savoir *TemporalConstraints*, *EnvironmentalConstraints* et *BehavioralActivity* sont qualifiées par des attributs et des constructeurs de tuples permettant de les caractériser. A partir du modèle MADS, des liens de généralisation et de spécialisation is-a permettent de spécifier des *Supertypes* et des *Subtypes* qui peuvent respectivement généraliser ou spécialiser des attributs de leurs *Subtypes* et *Supertypes*, respectivement.

La dernière catégorie de contrainte spatiale *SpatialConstraint* modélise les propriétés spatiales qui peuvent avoir un impact sur les patrons de déplacement *TravelPattern* d'un *Individual*. Comme pour les catégories de contrainte précédemment définies, *SpatialConstraint* représente une catégorie de facteurs influençant des décisions de déplacement.

Une *SpatialConstraint* est susceptible d'être définie à partir de *POI* ou de *Districts* que les *Individuals* sont susceptibles de visiter ou de parcourir lors de leurs déplacements. Les deux exemples de contraintes spatiales que nous avons retenus pour illustrer le potentiel de modélisation de notre approche sont les *POIs* et les *Districts*. Un *POI* est une localisation qui présente un potentiel intérêt pour un déplacement donné. Un *POI* est spatialement modélisé comme un point qui matérialise sa position dans l'espace. Ces localisations sont susceptibles de devenir la source ou la destination d'une *Trajectory*. Les attributs d'un *POI* peuvent permettre de rechercher les raisons qui ont conduit une catégorie de personnes à effectuer un tel déplacement d'un *POI A* à un *POI B*. Le rôle de ces districts est important dans le contexte de notre modèle dans la mesure où il va nous permettre de modéliser ces déplacements à différents niveaux d'abstraction, et le niveau des

quartiers est l'un de ces niveaux présentant un intérêt pour l'étude des décisions qui ont générées tel ou tel déplacement. Plus pratiquement un *District* sera considéré comme une primitive spatiale permettant de structurer une ville donnée en plusieurs quartiers représentatifs. En conséquence un *District* et un autre type de contrainte qui joue un rôle important dans les typologies de comportement de déplacement.

Contrainte 2

L'union des Districts du contexte d'étude forme une partition de l'espace.

2.5. Relations entre Différents Modules

Comme dans toute approche de modélisation spatio-temporelle ou même plus générale, les relations Relationships sont des concepts de modélisation importants du schéma Travel Behavior schema. La plupart des connexions sémantiques entre les différents modules sont matérialisés par des relations qui relient des objets identifiés. Par exemple la relation Follow entre un objet Route et un objet Trajectory est une abstraction importante du module TravelPattern, tout comme la relation temporelle entre Route et TransportationSegment. Un patron de déplacement TravelPattern est caractérisé par différents facteurs dans le module TravelPattern. Comme pour la relation entre le module Travel Pattern Module et le module Trajectory, de multiples relations sont modélisées. La relation entre le Trajectory et le module de patron de déplacements Travel Pattern Module matérialise une relation indirecte entre TravelPattern, Individual et représente le fait que chaque Trajectory appartient à un Individual. L'abstraction Individual joue un rôle important dans la modélisation de comportements. Individual est une unité de modélisation essentielle de la modélisation de patrons de déplacement. Un Individual peut posséder une à plusieurs Trajectories qui peuvent refléter plusieurs types d'action et de déplacement. Un District est en relation spatiale avec le réseau de transport. Considérant les possibles relations topologiques comme par exemple celles reliant les TransportationSegments aux Districts (Touch, Within and Intersect), il apparaît clairement que plusieurs configurations topologiques émergent et sont illustrées par le schéma suivant. Ces relations sont représentées dans la base de données initiale et par la relation topologique TopoGeneric (Figure 1). Les POIs jouent aussi un rôle comme les Trajectories. Un POI reflète une source, un intermédiaire ou un nœud de destination d'un déplacement. Un POI est connecté à son réseau et à partir d'une relation topologique avec un nœud du réseau, quand ce POI et les potentiels POIs sont disjoints. Au final, ces POIs jouent un rôle déterminant de qualité de service dans la matérialisation de routes optimales ou préférées.

3. Analyse de Pastrons de Déplacements

3.1. Interrogations de comportements de déplacements (*Travel Displacement Queries*)

Les interrogations de comportements de mobilités (*Travel behavior queries*) devraient idéalement apporter des réponses ou des indications de réponses pour des interrogations orientées vers la recherche d'une meilleure compréhension des mobilités humaines. Il apparaît clairement qu'une catégorisation de ces interrogations devrait permettre de mieux les appréhender et les définir. Selon Dodge et al. (2008), les comportements de mobilité peuvent être catégorisés en trois classes: spatiales, temporelles et spatio-temporelles, auxquelles nous rajoutons la dimension sémantique. A partir de cette classification nous proposons d'étudier les types d'interrogations considérées comme des primitives de base qui pourront être d'une part spécifiées au sein d'un langage de manipulation de bases de données, et ensuite combinées pour exprimer des recherches à haute valeur ajoutée.

3.1.1. Niveau sémantique

A partir de ce niveau sémantique, les interrogations portent généralement sur des abstractions fondamentales exprimées à partir du « Qui », « Quoi », « Comment » ou « Pourquoi » et la manipulation de primitives de base. (Thériault and Claramunt, 1999)

A partir du module *Travel Behavior*, les objets sémantiques et leurs attributs sont pour l'essentiel déclinés dans le module *Travel Pattern* qui répertorie les possibles facteurs qui influencent les décisions de déplacement et qui sont spécifiés à partir de trois catégories complémentaires : *Individual Profile*, *Behavioral Activity* and *Environmental Constraint*.

Bien que le module *Travel Pattern Module* décrive l'essentiel des données sémantiques, ou à minima les plus riches, d'autres opportunités d'extension à partir d'interrogations combinant les informations explicitement représentées au sein des autres mobiles. Par exemple, et à partir de l'objet *Individual*, il est possible de répondre à la question basique "*Qui est cet individu?*". Qualifier and reconnaître qui est un individu en déplacement devrait pouvoir permettre de répondre à de simples interrogations, rechercher des régularités de déplacement et explorer quelles seraient les raisons amenant à cette activité. Plus globalement, dériver une nouvelle part de connaissance des dynamiques et caractéristiques d'une mobilité, et en généralisant une telle approche à un ensemble d'individus se déplaçant dans la cité doit pouvoir permettre d'étudier les phénomènes de déplacement dans la cité. En rattachant par exemple de telles interrogations sémantiques à des détails de parcours et en particulier des *POIs* qui seraient parcourus ou certains types d'activités peut amener à caractériser un ensemble de patrons de déplacements, ou à mieux appréhender le caractère et/ou les particularités d'un ou d'un ensemble d'individus.

3.1.2. Niveau temporel

Les interrogations appliquées au niveau temporel doivent privilégier l'étude de régularités ou d'irrégularités des mobilités d'un ou de plusieurs individus, des origines et/ou des destinations de déplacement, les temps passés à une destination donnée pour réaliser une certaine activité, les temps de parcours pour ne citer que quelques exemples. Ces interrogations sont particulièrement importantes pour des études de transport et de planification en milieu urbain. Du point de vue des abstractions de base utilisées ces interrogations utilisent le plus souvent les qualités « Quand », « Combien de temps », « Combien de fois » et peuvent s'appliquer préférentiellement à partir du module Module.

3.1.3. Niveau spatial

Les interrogations portant sur cette dimension spatiale manipulent des données de localisations et des géométries. En dehors d'expressions basiques telles que « Où se situe cet individu », d'autres catégories d'interrogations peuvent être appliquées à partir de métriques, et d'opérateurs topologiques et de direction, ou leur combinaisons.

3.1.4. Niveau spatio-temporel

Les interrogations combinant les niveaux spatiaux et temporels combinent ces dimensions à partir notamment d'objets représentés dans le module Travel Behavior. Ces interrogations manipulent les abstractions « Où et Quand » et peuvent typiquement porter sur des propriétés de trajectoires (e.g., origine et destination, localisation, parcours) ou sur des contraintes spatiales SpatialConstraints et temporelles TemporalConstraints introduites dans les sections précédentes. Un exemple de patron qui peut être analysé aborde des propriétés fonctionnelles. Par exemple, une comparaison explorerait les caractéristiques spatiales des déplacements journaliers des Trajectory d'un Individual et les Districts croisés lors de ces déplacements. D'autres exemples plus significatifs pourraient par exemple rechercher les différences et les similarités de déplacement entre les trajectoires d'un même individu à plusieurs temps donnés, ou entre plusieurs individus, ou des fréquences de déplacements dans certaines conditions temporelles. Ces mêmes tendances de déplacement pourraient même être étudiées à différents niveaux de granularité temporelle et échelle spatiale.

3.1.5 Niveaux sémantiques, spatiaux et temporels

Afin de mieux comprendre et appréhender les comportements de mobilité dans la cité il est naturel de vouloir combiner les dimensions sémantiques, spatiales et temporelles. Les opportunités d'exploration sont alors démultipliées. Elles peuvent aborder et étudier les conditions de déplacement d'un ou plusieurs *Individual*. Les interrogations portant sur le niveau d'un *Individual* sont susceptibles de rechercher des régularités ou des irrégularités. Plus généralement les interrogations portant sur plusieurs *Individuals* pourraient par exemple recherche « qui sont les individus qui

sont passés par telle route à un temps donné » ou « quels sont les lieux qu'un individu donné a visité à tel moment »

4. Implementation et Validation

La mise en œuvre de ce modèle de comportement de voyage est réalisée à partir d'un système de base de données spatiales PostgreSQL/PostGIS. PostgreSQL est une base de données Open source qui a l'avantage de permettre d'utiliser des extensions spécifiées par l'utilisateur. PostGIS est l'extension spatiale de la base de données relationnelle objet PostgreSQL qui fournit des types de données spatiaux, des index spatiaux et des fonctions de manipulation de données spatiales. Nous avons donc développé une base de données de représentation de mobilités à partir de cet environnement PostgreSQL/PostGIS et instanciées à partir du modèle de données mentionné dans les sections précédentes. Sont notamment spécifiées les objets trajectoires, les individus et les relations les associant. Comme dans de nombreuses approches de modélisation de base de données, toutes les entités identifiées par le modèle de mobilité sont « mappées » vers des tables relationnelles. Des index spatiaux GIST-index et temporels b-Tree offerts par PostgreSQL/PostGIS ont été appliqués pour accélérer les requêtes spatiales et temporelles. Des opérations spécifiques de manipulation de données ainsi que les opérations temporelles spatiales ont été implémentées en tant que fonctions au niveau de l'interface.

4.1. Conception de Fonctions

Les fonctions constituent des éléments de base des requêtes relativement évoluées des langages d'interrogation de données. Une base de données spatiales doit habituellement offrir une série d'opérateurs et de fonctions qui manipulent des données quantitatives (e.g., métrique) ou qualitatives (e.g., topologique). De nombreux opérateurs, fonctions spatiales et temporelles de base sont le plus souvent disponibles dans les systèmes de base de données spatiaux actuels, tel qu'Oracle Spatial, PostGIS etc. Ces fonctions comprennent notamment les métriques et les constructeurs de géométrie (e.g., distance, longueur, buffer), les opérateurs topologiques (e.g., *cross*, *within*, *Touch*). Cependant, ces opérateurs et ces fonctions ne sont pas toujours suffisants pour des applications spécifiques à sémantique de haut niveau, comme pour l'analyse à effectuer par notre étude orientée vers l'analyse des mobilités. Il en résulte que plusieurs catégories de fonctions spécifiques doivent être dérivées et agrégées et définies à partir de sémantiques explicites pour les utilisateurs et qui s'adapteront aux fonctions requises par notre approche. De telles fonctions devraient directement et explicitement faciliter la recherche de modalités de mobilités de certaines personnes, et devraient même être construites pour à la fois traiter et manipuler des données spatiales, temporelles et sémantiques.

Avec l'objectif de s'appliquer préférentiellement à partir de notre modèle conceptuel et logique de représentation des mobilités, les fonctions à construire ont pour objectif d'aider à découvrir des comportements de mobilité dans la ville. La conception de ces fonctions est basée sur les trois catégories décrites dans la section

précédente. Afin de tenir compte de la complexité des analyses possibles de comportements de mobilité, la plupart des fonctions identifiées combinent les dimensions spatiales, temporelles et sémantiques. Des opérateurs géométriques et topologiques spécifiques ont été dérivés d'une catégorisation suggérée dans une précédente recherche (Wu et al., 2014). Le tableau 1 donne une liste des fonctions conçues à ce stade, y compris leur signature et une brève description.

Table 1. Liste de Fonctions

N o.	Name	Signature	Type	Description
1	Ind_At_Instant	Individual x Timestamp-> Geometry(POINT)	Manipulation spatio-temporelle	Retourne la position d'un individu à un temps donné
2	Ind_Traj_During_Period	Individual x Time_Interval-> geometry(LINESTRING)	Manipulation spatio-temporelle	Retourne les segments de transport qu'un individu réalise pour un intervalle temporel donné
3	Ind_Enter_At	Individual x Geometry-> Timestamp	Interrogation topologique	Retourne le temps quand un individu rentre dans une Geometry(LINESTRING or POLYGON)
4	Ind_Exit_At	Individual x Geometry-> Timestamp	Interrogation topologique	Retourne le temps quand un individu est localisé dans une Geometry(LINESTRING or POLYGON)
5	Ind_Confine_In	Geometry(Polygon) x Time_Interval -> Set(Trajectory,Individual)	Interrogation topologique	Retourne les trajectoires d'un individu confinées dans un MBB spatio-temporel
6	Traj_MBR	set(Trajectory) -> Region	Manipulation spatiale	Retourne le minimum rectangle englobant (MBR) d'un ensemble de trajectoires
7	Traj_Pass_District	Trajectory x District -> Boolean	Interrogation topologique	Détermine si une trajectoire <i>CrossIn</i> et puis <i>CrossOut</i> un district
8	Traj_Pass_Transport	Trajectory x (TransportationSegment) -> Boolean	Interrogation topologique	Détermine si une trajectoire <i>CrossIn</i> et puis <i>CrossOut</i> un segment de transport
9	Traj_Nearest_POIs	StartTraj EndTraj(Trajectory) x	Buffer	Trouve les points d'intérêt POIs à une

		Distance -> set(POI)		certaine distance du début ou la fin d'une trajectoire	
0	1	Traj_Direction	Trajectory -> Bearing	Interrogation spatiale	Retourne la direction d'une trajectoire
1	1	Traj_Co-direction	Trajectory x Trajectory -> Boolean	Interrogation spatiale	Determine si deux trajectoires sont orientées dans la même direction
2	1	Ind_Find_Route	Individual x Time_Interval-> Set(Route)	Aggrégation	Retourne les routes extraites des trajectoires d'un individu pour un intervalle de temps donné
3	1	Traj_Fit_Route	Trajectory x Route -> Boolean	Analyse spatio-temporelle	Détermine si une trajectoire suit une route donnée
4	1	Ind_Share_Path	set(individual) -> set(path)	Aggrégation	Retourne les chemins partagés par plusieurs trajectoires
5	1	Trans_Avg_Speed	set(TransportationSegments) x Timestamp -> Speed	Analyse spatio-temporelle	Retourn la vitesse moyenne des trajectoires passant par un segment de transport et pour un intervalle de temps donné

4.2. Données expérimentales de trajectoires

Les requêtes implémentées sont exécutées à partir d'une large base de données de trajectoires extraites de la ville de Pékin. Ce jeu de données de trajectoires dérivées de signaux GPS a été collecté par le projet Geolife (Microsoft Research Asia) (Zheng et al., 2008; Zheng et al., 2009; Zheng et al., 2010) par 178 utilisateurs sur une période de quatre ans d'avril 2007 à octobre 2011. Une trajectoire GPS dans ce référentiel est représentée par une séquence de points estampillés dans le temps et dont les attributs spatiaux sont la latitude, la longitude et l'altitude. Ce jeu de données contient 17 621 trajectoires couvrant une distance totale de 1 251 654 kilomètres et une durée totale de 48 203 heures. Ces trajectoires ont été enregistrées par différents utilisateurs et capteurs GPS ou de téléphones équipés de GPS, et couvrent une grande variété de taux d'échantillonnage : 91% des trajectoires sont enregistrées dans une représentation dense, par exemple toutes les 1 ~ 5 secondes ou tous les 5 ~ 10 mètres par point.

Ce jeu de données Geolife a enregistré un large éventail de mouvements dans la ville de Beijing, en qualifiant notamment les modes de transport utilisés. Alors que certains utilisateurs ont porté un enregistreur GPS pendant quelques années, d'autres

n'ont utilisé cette configuration similaire que pour quelques semaines. Les noms des individus dans la base de données et les requêtes sont simulés dans la mesure où les informations personnelles des utilisateurs ne sont pas disponibles. Les données de transport et d'autres données spatiales de la ville de Pékin sont issues de données spatiales du domaine public, y compris les réseaux routiers et de métro, les districts, et une série de POI extraits du service de cartographie Baidu (Lbsyun Baidu, 2016).

4.3. Extraction des trajectoires

Afin d'associer les trajectoires extraites du projet Géolife au réseau de transport, un processus algorithmique de mise correspondance de ces trajectoires avec le référentiel cartographique a été mis en œuvre. L'implémentation de cet algorithme est principalement basée sur une méthode basée sur des contraintes proposée par Brakatsoulas et al. (2005). Le principe qui sous-tend cette approche consiste à faire correspondre les segments de trajectoires vers les segments du réseau en tenant compte de la connectivité des graphes et des contraintes spatio-temporelles (i.e., la distance couverte par une trajectoire entre deux segments de transport doit être plausible). Ce processus a été exécuté par des fonctions pl/pgsql spécifiquement développées à partir de PostgreSQL. Cette approche permet de générer les segments de transport *TransportationSegments* et les nœuds de transport *TransportationNodes* modélisant le réseau de transport.

4.4. Résultats

Nous introduisons une série de requêtes illustrant le potentiel de notre approche de modélisation. Ces requêtes combinent les fonctionnalités du langage PostgreSQL/PostGIS avec les fonctions définies dans les sections précédentes. Ces requêtes manipulent les dimensions sémantiques, spatiales et temporelles.

Requête 1: Quels sont les segments de route que Li Lei a effectués de 7h à 10h ? Et quels sont les segments de route que Li Lei a régulièrement répété de 18h à 21 heures ?

```
Select Ind_find_route(id, '7:00:00', '10:00:00') from IndividualTable where name = 'Li Lei' ;
```

```
Select Ind_find_route(id, '18:00:00', '21:00:00') from IndividualTable where name = 'Li Lei';
```

La fonction *Ind_Find_Route (IndividualId, EndTime)* est encapsulée dans ces deux requêtes dont les résultats sont illustrés par la Figure 2. Cette fonction retourne les itinéraires suivis par Li Lei pendant les intervalles de temps donnés. Ces deux requêtes font implicitement une différence entre les itinéraires effectués par une personne tôt le matin et tard dans l'après-midi. Il apparaît à partir des résultats de la requête illustrés par la figure 2 que les itinéraires effectués en fin d'après-midi sont plus flexibles que ceux effectués tôt le matin. Cela pourrait refléter un ensemble relativement important d'activités effectuées par cette personne après son travail, et

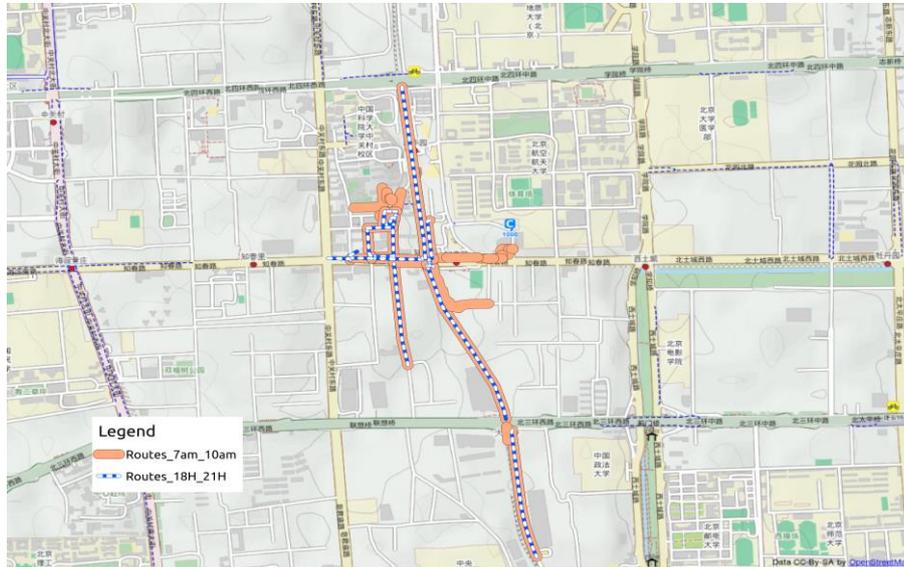


Figure 2 Segments de transportation régulièrement répétés par Li Lei le matin et le soir

d'autre part beaucoup moins de flexibilités pour celles effectués le matin et de plus réguliers patrons de trajectoires. Ce type de requête illustre comment l'exploration de patrons de trajectoire peut aider à l'exploration de comportements type..

Requête 2: Quels sont les segments que Li Lei répète régulièrement pendant les jours de la semaine et les week-ends ?

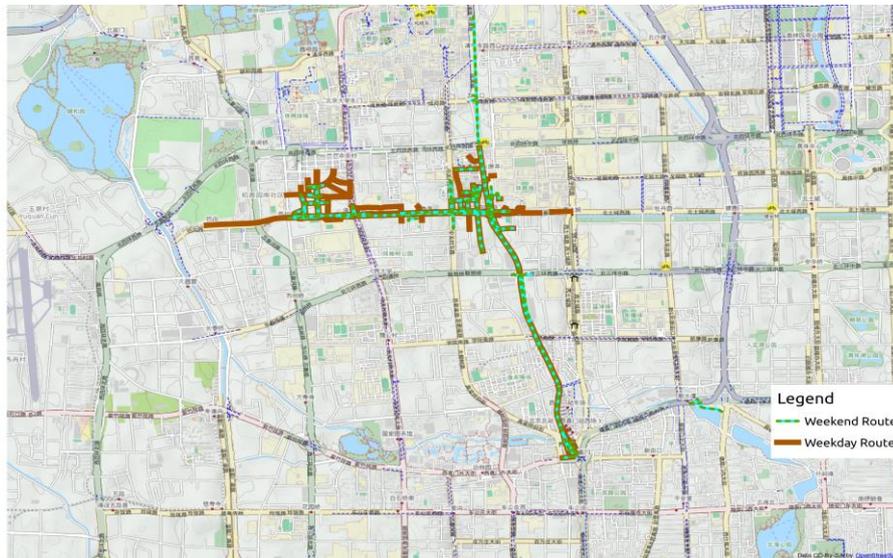


Figure 1 Routes les week-ends et en semaine pour un individu donné

Select **Ind_find_route**(id, 'weekday') from IndividualTable where name = 'Li Lei' ;

Select **Ind_find_route**(id, 'weekend') from IndividualTable where name = 'Li Lei';

La fonction *Ind_find_route* est appliquée dans ces deux requêtes. Elles recherchent les itinéraires effectués par cette personne en semaine et le week-end. Elles permettent de faire une différence entre les comportements de déplacement en semaine et les week-ends et d'en analyser les différences. La figure 3 illustre la palette des itinéraires suivis. Il apparaît clairement que la plage géographique des itinéraires suivis pendant les jours de semaine est plus petite que ceux suivis pendant les week-ends. Cela traduit des patrons de mobilité et des régularités d'activité en semaine et à contrario une plus étroite gamme d'activité pendant les week-ends mais avec une plus grande étendue géographique.

Requête 3: Quelle est l'étendue spatiale des mobilités effectuées par Wang Hua et Li Lei à partir de 9h le 21 septembre 2008?

Select **Traj_MBR**(union(trajecory.id))from TrajectoryTable traj, IndividualTable indiv where traj.individualid = indiv.id and (indv.name = 'Wang Hua' or indiv.name = 'Li Lei') and extract(hour from starttime)>9;

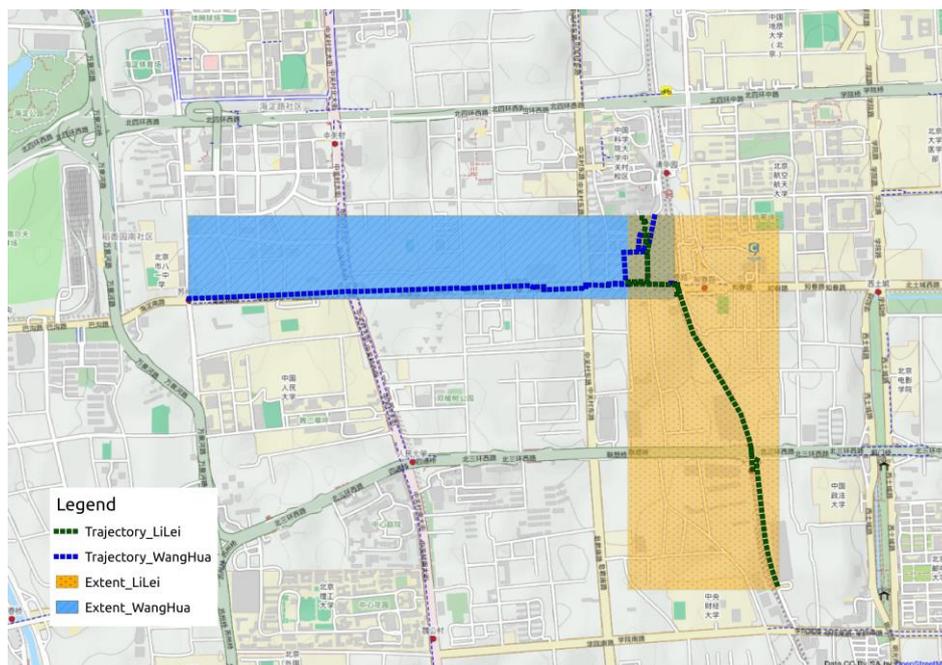


Figure 3 Couverture spatiale des trajectoires effectuées par des individus donnés pour un temps donné

La fonction *Traj_MBR* (*TrajectoryId*) est encapsulée dans cette requête. Cette fonction renvoie le *Minimum Bounding Rectangle* (MBR) des trajectoires données. L'objectif de cette requête est de révéler l'étendue géographique respective des

trajectoires effectuées par certaines personnes pour un intervalle de temps donné. Plutôt que de comparer la géométrie exacte des trajectoires effectuées par une personne, l'idée sous-jacente à cette fonction consiste à comparer visuellement les régions couvertes, ces régions étant dérivées par un MBR. Lorsque certains individus partagent certaines régions géographiques, cela pourrait révéler des modèles de comportement remarquables et des intérêts communs lorsqu'ils se déplacent dans la ville et donc des modalités de co-voiturage ou de transport collectif à étudier. La figure 4 montre le résultat de cette requête appliquée à deux individus donnés. Dans ce cas précis, les figures qui apparaissent montrent que ces deux individus partagent une localisation proche à la fin de leurs trajectoires le matin. Il serait alors possible que les deux individus soient des collègues de travail potentiels.

Requête 4: Recherche les districts de provenance des personnes qui prennent la ligne de métro 10 entre 6 heures et 9 heures

```
Select district.id, district.geom from district, TrajectoryTable trj,
TransportationSegmentTable seg, TransportationTable trans where
Traj_PassTransport (trj.id, seg.id) and extract(hour from trj.starttime) between 6am
and 9am and seg.roadid = trans.id and trans.name = 'subway line 10' and
st_within(trj.startpt,district.geom);
```

La fonction *Traj_PassTransport* (*Trajectory.Id*, *TransportationSegment.ID*) est appliquée à cette requête. Cette fonction booléenne retourne 'vrai' lorsqu'une

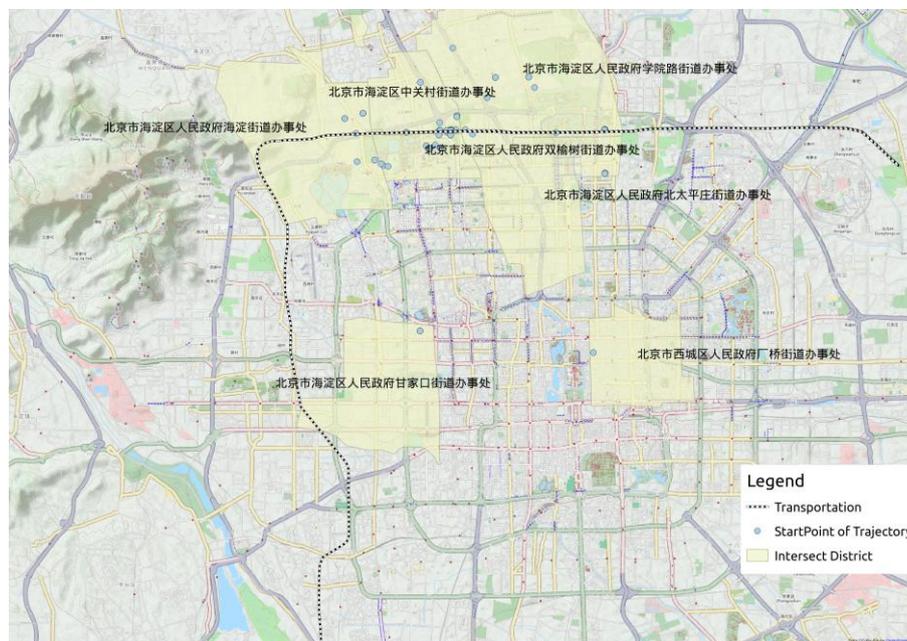


Figure 4 Districts matérialisant des débuts de trajectoires de la ligne de métro 10

trajectoire donnée passe par un segment de transport *TransportationSegment* donné. L'objectif de cette requête est d'explorer et de rechercher les lieux d'où les personnes prennent la ligne de métro 10 le matin. Dans la ville de Beijing, une telle requête pourrait permettre de rechercher des configurations de patrons de transport, une information potentiellement utile pour la planification en transport. Le résultat de cette requête est illustré à la figure 5. Il apparaît que les individus qui prennent la ligne de métro 10 le matin proviennent principalement du nord-ouest de la ville.

Requête 5: Recherche les itinéraires les plus réguliers suivis par les personnes dans l'après-midi de 5 heures à 7 heures. Et qui sont globalement les chemins les plus populaires suivis par la plupart des individus dans l'après-midi de 5 heures à 7 heures ?

```
Select rou.id, num, trans.name, trans.type, seg.geom from (select
sum(array_length(trajs,1)) num, rou.segmentid from RouteTable rou group by
segmentid order by num desc limit 15), TransportationTable trans,
TransportationSegmentTable seg where rou.segmentid = seg.id and trans.id =
seg.roadid;
```

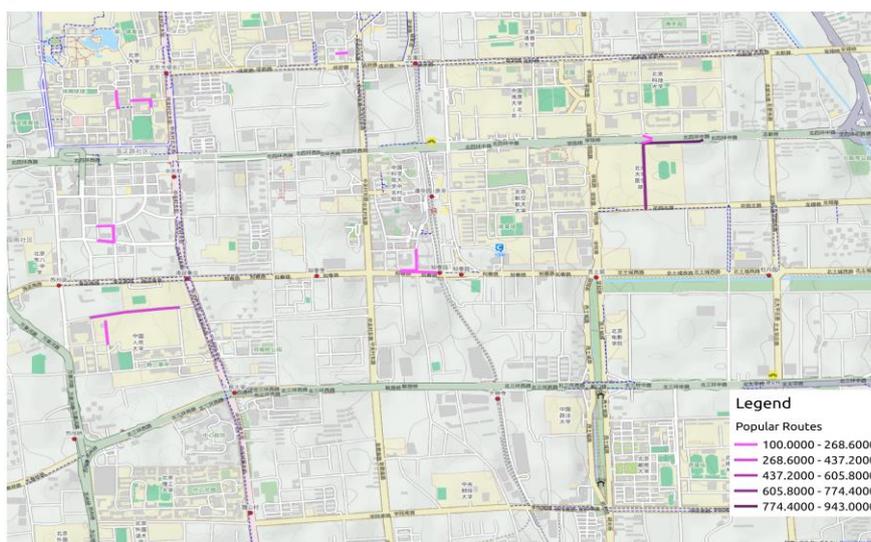


Figure 5 Routes densément fréquentées

```
Select rou.id, num_indiv, trans.name, trans.type, seg.geom from (select
count(rou.individualid) indiv, rou.segmentid from RouteTable rou group by
segmentid order by indiv desc limit 15), TransportationTable trans,
TransportationSegmentTable seg where rou.segmentid = seg.id and trans.id =
seg.roadid;
```

Ces deux requêtes appliquent des agrégations spatio-temporelles. Les résultats sont illustrés par les Figures 6 et 7. Le but de ces requêtes est de trouver les itinéraires les plus populaires dans la ville pour un intervalle de temps donné ainsi

que les plus densément réalisés. Ces requêtes devraient révéler à la fois des parcours fréquents et ceux qui sont très fréquentés à certains temps. Les résultats de ces requêtes révèlent les itinéraires les plus fréquemment exécutés par des individus (première requête, i.e., les 15 premiers itinéraires) et les plus occupés (deuxième requête). La Figure 6 montre les itinéraires les plus densément parcourus, tandis que la Figure 7 décrit les parcours les plus fréquents et révèle des régularités d'itinéraires. Il semble que les itinéraires les plus populaires soient relativement dispersés dans la ville, alors que les itinéraires fréquentés par la plupart des

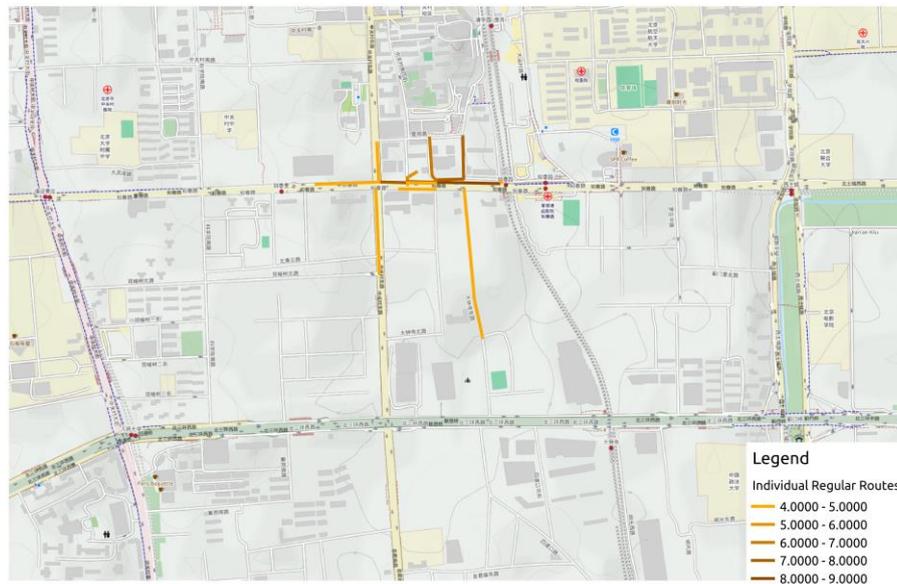


Figure 6 Routes populaires

individus sont plus concentrés dans une même région. Le nombre d'individus agrégés dans la Figure 7 est relativement faible dans la mesure où peu de personnes suivent des itinéraires réguliers. Cela reflète la diversité des comportements de mobilité dans cas.

Requête 6: Où la personne nommée «Li Ming» commence habituellement et termine ses déplacements avant 9:00 le matin? Quelles sont les destinations possibles ?

```
Select Traj_Nearest_POIs((select endpt from TrajectoryTable traj,
IndividualTable indiv where indiv.id = traj.individualid and indiv.name = 'Li Ming'
and extract(hour from traj.starttime) < 9),50);
```

La fonction *Traj_Nearest_POIs* (*Trajectory.StartPT/EndPT*, *Radium*) est appliquée à cette requête. Cette fonction renvoie le point d'intérêt le plus proche du début ou de la fin de la trajectoire pour une distance donnée. Les résultats de la requête sont illustrés par la Figure 8. Ils montrent que les destinations matinales sont relativement limitées à une petite région. On pourrait conclure que ce modèle régulier dans le voisinage de son domicile est susceptible d'être une trajectoire vers son lieu de travail. Les *POIs* proches de cette zone de destination pourraient représenter les lieux de travail possibles de cet individu. On pourrait même explorer les lieux de travail possibles selon les catégories de *POIs* présentes dans cette région. D'autre part, les trajectoires pratiquées le matin pour atteindre le lieu de travail potentiel reflètent une relative diversité, une raison possible étant des arrêts probables pour par exemple prendre son petit déjeuner.

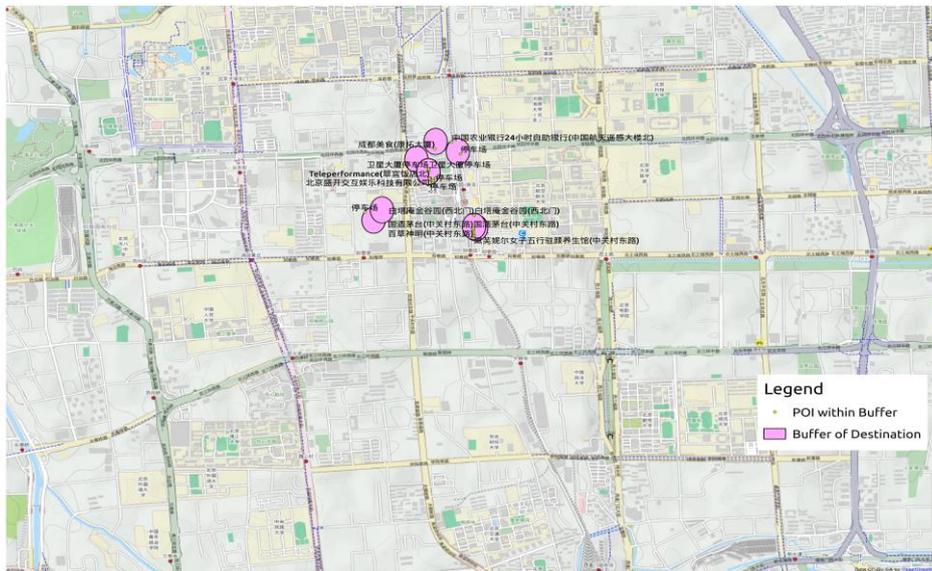


Figure 7 Points d'intérêts POIs proches de destinations en matinée

Au vu des résultats de l'ensemble de ces requêtes précédentes qui ont un caractère à la fois illustratif et exploratoire, il apparaît que dans de nombreux cas, l'analyse des modèles de trajectoire peut révéler des habitudes de mobilités voire certains comportements sociaux, et ce, même si la sémantique des données de base telles que fournies par la base de données Geolife est relativement limitée.

4.5. Evaluation des performances

Les performances de calcul des requêtes présentées ci-dessus montrent que cette implémentation basée sur notre modèle de comportement de mobilités est capable de manipuler efficacement de grandes données de déplacement extraites des données GPS brutes. Toutes les requêtes ont été exécutées sur un processeur Intel Core i7

avec 8 Go de RAM. Les délais d'exécution varient de quelques millisecondes à quelques centaines de millisecondes, ce qui est largement acceptable au niveau d'une interface et d'un utilisateur. Par exemple, les temps d'exécution des deux requêtes 5 sont les plus rapides (i.e., temps d'exécution de 325 ms et 3.149 ms). Ces temps d'exécution très rapides sont le résultat de la manipulation des seuls attributs sémantiques. La requête 3 (i.e., temps d'exécution de 426,244 ms) et la requête 6 (i.e., temps d'exécution de 608.800 ms) sont légèrement plus lentes, bien qu'elles restent encore rapides. Cela s'explique par l'application d'une jointure de table encapsulée dans les fonctions prédéfinies *Traj_MBR* et *Traj_Nearest_POIs*. Inversement, la requête 1 (i.e., temps d'exécution de 1331,423 ms et de 1094,187 ms), la requête 2 (i.e., temps d'exécution de 12707,713 ms et de 6258,714 ms) et la requête 4 (i.e., temps d'exécution 11704.87 ms) sont les requêtes les plus lentes car elles impliquent des fonctions qui encapsulent des opérateurs spatiaux sur plusieurs tables. Ces temps d'exécution restent globalement acceptables bien que des opérateurs spatiaux, sémantiques et temporels soient appliqués, et dont les résultats sont affichés visuellement au niveau de l'interface. Au final, il apparaît clairement que l'extensibilité de PostgreSQL/PostGIS s'est avérée favorable à la mise en œuvre de ce modèle de comportement de mobilités ainsi que les fonctions spécialisées identifiées.

5. Conclusion

La recherche développée dans cet article introduit une approche conceptuelle de base de données spatio-temporelles dont l'objectif est de modéliser les trajectoires humaines dans un environnement urbain et de fournir des capacités de manipulation et d'analyse de ces données. Le modèle de données de mobilités propose plusieurs concepts de modélisation à des niveaux sémantiques, spatiaux et temporels. Le modèle identifie plusieurs notions primitives telles que les abstractions d'individus, de trajectoires et de routes, l'ensemble étant relié par une série de relations et une association aux réseaux urbains sous-jacents. Les principales caractéristiques des réseaux urbains et de transport sont prises en compte, tout comme une série de contraintes sémantiques, temporelles et spatiales. A partir d'une implémentation réalisée sur les systèmes de base de données PostGIS et PostgreSQL, plusieurs fonctions spécifiques ont été identifiées et mises en œuvre. Ces fonctions permettent de manipuler les dimensions sémantiques, temporelles et spatiales à partir de modalités proches des besoins utilisateurs, et ont pu être expérimentées et mises en œuvre. Les principes de cette approche peuvent être étendus à d'autres opérateurs de manipulation de données génériques et spécifiques, susceptibles d'être distribués et partagés avec d'autres utilisateurs.

Bibliographie

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B., & Vaisman, A. (2007). A model for enriching trajectories with semantic geographical information. *Proceedings of the 15th Annual ACM International*

- Symposium on Advances in Geographic Information Systems - GIS '07*, (i), 1. <https://doi.org/10.1145/1341012.1341041>
- Andrienko, N., Andrienko, G., Pelekis, N., & Spaccapietra, S. (2008). Basic concepts of movement data. In *Mobility, Data Mining and Privacy: Geographic Knowledge Discovery* (pp. 15–38). https://doi.org/10.1007/978-3-540-75177-9_2
- Chen, B. Y., Yuan, H., Li, Q., Shaw, S.-L., Lam, W. H. K., & Chen, X. (2016). Spatiotemporal data model for network time geographic analysis in the era of big data. *International Journal of Geographical Information Science*, 30(6), 1041–1071. <https://doi.org/10.1080/13658816.2015.1104317>
- Chen, W., Ji, M. H., & Wang, J. M. (2014). T-DBSCAN: A spatiotemporal density clustering for GPS trajectory segmentation. *International Journal of Online Engineering*, 10(6), 19–24. <https://doi.org/10.3991/ijoe.v10i6.3881>
- Claramunt, C., & Thériault, L. (1996). Towards semantics for modeling spatiotemporal processes within gis. In *Seventh International Symposium on Spatial Data Handling, SHD'96* (pp. 1–12).
- Dodge, S., Weibel, R., & Lautenschütz, A.-K. (2008). Towards a taxonomy of movement patterns. *Information Visualization*, 7(3–4), 240–252. <https://doi.org/10.1057/PALGRAVE.IVS.9500182>
- Goodchild, M. F., Yuan, M., & Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21(3), 239–260. <https://doi.org/10.1080/13658810600965271>
- Hägerstrand, T. (1970). WHAT ABOUT PEOPLE IN REGIONAL SCIENCE? *Papers in Regional Science*, 24(1), 7–24. <https://doi.org/10.1111/j.1435-5597.1970.tb01464.x>
- Huff, J. O., & Hanson, S. (1986). Repetition and Variability in Urban Travel. *Geographical Analysis*, 18(2), 97–114. <https://doi.org/10.1111/j.1538-4632.1986.tb00085.x>
- Inoue, R., & Tsukahara, M. (2016). Travel Pattern Analysis from Trajectories Based on Hierarchical Classification of Stays. *International Conference on GIScience, Short Paper Proceedings, UC Berkeley, 1*. <https://doi.org/10.21433/B3111WV9S715>
- Jiang, S., Ferreira, J., & González, M. C. (2012). Discovering urban spatial-temporal structure from human activity patterns. *The ACM SIGKDD International Workshop*, 95. <https://doi.org/10.1145/2346496.2346512>
- Kitamura, R. (1988). An evaluation of activity-based travel analysis. *Transportation*. <https://doi.org/10.1007/BF00167973>
- Parent, C., Spaccapietra, S., & Zimányi, E. (2006). *Conceptual modeling for traditional and spatio-temporal applications: The MADS approach. Conceptual Modeling for Traditional and Spatio-Temporal Applications: The MADS Approach*. <https://doi.org/10.1007/3-540-30326-X>
- Pentland, A., & Liu, A. (1999). Modeling and Prediction of Human Behavior. *Neural Computation*, 11(1), 229–242. <https://doi.org/10.1162/089976699300016890>

- Renso, C., Baglioni, M., de Macedo, J. A. F., Trasarti, R., & Wachowicz, M. (2013). How you move reveals who you are: Understanding human behavior by analyzing trajectory data. *Knowledge and Information Systems*, 37(2), 331–362. <https://doi.org/10.1007/s10115-012-0511-z>
- Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Porto, F., & Vangenot, C. (2008). A conceptual view on trajectories. *Data and Knowledge Engineering*, 65(1), 126–146. <https://doi.org/10.1016/j.datak.2007.10.008>
- Tang, L., Kan, Z., Zhang, X., Yang, X., Huang, F., & Li, Q. (2016). Travel time estimation at intersections based on low-frequency spatial-temporal GPS trajectory big data. *Cartography and Geographic Information Science*, 43(5), 417–426. <https://doi.org/10.1080/15230406.2015.1130649>
- Zheng, Y., Li, Q., Chen, Y., Xie, X., & Ma, W.-Y. (2008). Understanding mobility based on GPS data. In *Proceedings of the 10th international conference on Ubiquitous computing - UbiComp '08* (p. 312). <https://doi.org/10.1145/1409635.1409677>
- Zheng, Y., Xie, X., & Ma, W. (2010). GeoLife: A Collaborative Social Networking Service among User, Location and Trajectory. *IEEE Data Eng. Bulletin*, 33(2), 32–40. <https://doi.org/10.1.1.165.4216>
- Zheng, Y., Zhang, L., Xie, X., & Ma, W.-Y. (2009). Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of the 18th international conference on World wide web - WWW '09* (p. 791). <https://doi.org/10.1145/1526709.1526816>