
Modélisation, Simulation et Analyse d'Environnements Urbains Peuplés

Approche multi-agent pour l'étude des déplacements multimodaux

W. Chaker * — M.-J. Proulx ** — B. Moulin * — Y. Bédard **

* Département d'informatique et de génie logiciel – Université Laval, 1065, av. de la Médecine, Local 3908, Québec (QC), Canada G1V 0A6

waliid.chaker@ift.ulaval.ca ; bernard.moulin@ift.ulaval.ca

** Centre de Recherche en Géomatique – Université Laval, Pavillon Louis-Jacques-Casaul, Québec (QC), Canada G1V 0A6

Marie-Josée.Proulx@scg.ulaval.ca ; Yvan.Bedard@scg.ulaval.ca

RÉSUMÉ. Cet article présente une nouvelle approche de conception d'un Environnement Urbain Virtuel permettant une modélisation multi-échelle grâce à l'utilisation d'un patron de conception qui met en relation les lieux d'activité et les différents réseaux de transport. Cet environnement est ensuite enrichi avec une population synthétique qui l'occupe pour former l'Environnement Urbain Peuplé. Nous avons développé le méso-simulateur de déplacements TransNetSim et l'avons couplé avec le micro-simulateur Vissim afin qu'il puisse tirer partie de tous les avantages du modèle. L'utilisation d'outils SOLAP a permis d'une façon simple et efficace d'analyser les résultats des simulations réalisés à la ville de Québec.

ABSTRACT. In this paper, we propose a novel multiscale approach to model Virtual Urban Environments thanks to the introduction of a design pattern which links activity places and transportation networks. This environment is also augmented with a synthetic population in order to build the Populated Urban Environment. The meso-simulator TransNetSim has been developed and coupled with the micro-simulator Vissim to benefit from the model capabilities. We used SOLAP tools to carry out the analysis of the simulations' results in a simple and efficient manner for Quebec city.

MOTS-CLÉS : environnement urbain virtuel, population synthétique, modélisation multi-échelle, mobilité urbaine, transport multi-modal, Spatial OLAP, simulation multi-agent du trafic.

KEYWORDS: virtual urban environment, synthetic population, multiscale modelling, urban mobility, multi-modal transport, Spatial OLAP, multi-agent simulation of traffic.

1. Introduction

Depuis les années 1990, motivés par l'essor que prenaient les *Systèmes d'Information Géographiques* (SIG), plusieurs travaux ont défini de nouveaux concepts comme l'*Environnement Urbain Virtuel* (Donikian, 1997)(Jiang, 1999), l'*Environnement Géographique Virtuel* (Gong *et al.*, 1999), ou les *Villes Fractales* (Batty *et al.*, 1994). L'objectif commun est de modéliser l'espace urbain afin de pouvoir simuler certains phénomènes liés aux *dynamiques urbaines* (Albeverio *et al.*, 2008). On distingue principalement deux catégories: D'abord, les *simulations urbaines* à base d'*Automates Cellulaires* (Benenson *et al.*, 2004) (Waddell, 2002) qui couvrent à petites échelles des espaces larges. Ensuite, les *simulations de la mobilité quotidienne* des individus (Moulin *et al.*, 2003) (Timmermans *et al.*, 2009), couvrant des espaces détaillés mais restreints, traitées par des approches *multi-agents* en complément des approches macroscopiques traditionnelles. Aujourd'hui, grâce à la croissance quasi-exponentielle des capacités informatiques et de la disponibilité des données SIG, on assiste à une certaine convergence des approches de simulation des dynamiques urbaines. Les simulations qui regroupent un très grand nombre d'agents capables de réagir avec autonomie dans des environnements spatiaux larges et détaillés sont plus que jamais des sujets d'intérêt, comme en témoignent les *systèmes multi-agents massifs* (Ishida *et al.*, 2004). Cependant, même avec un ordinateur très puissant et des données SIG détaillées et exhaustives, on ne peut garantir la faisabilité et la vraisemblance de telles simulations.

Dans le cadre de ce travail, nous avons abordé ces problèmes pour une application particulière : la simulation des déplacements multimodaux des habitants de la ville de Québec. D'abord, nous avons noté la nécessité de faire varier l'échelle de l'*Environnement Urbain Virtuel* (EUV) pour que celui-ci puisse servir à différents simulateurs ayant des modèles comportementaux plus ou moins détaillés. Par conséquent, nous proposons un modèle multi-échelle de l'EUV permettant de regrouper les concepts de *lieux*, de *réseaux de transport* et leurs liens. Nous avons aussi remarqué que les approches existantes ne rassemblent pas l'offre et la demande en transports dans un même modèle de l'environnement, c'est-à-dire l'environnement spatial et la population qui l'occupe. Nous proposons alors un modèle plus générique que nous appelons l'*Environnement Urbain Peuplé* (EUP) et qui est construit à partir de l'EUV. La section 2 de l'article définit formellement un EUV grâce au concept de *T-graphe* hiérarchique et pondéré que nous avons introduit. L'EUP est ensuite obtenu en produisant une population synthétique associée à l'EUV. Une démarche générique pour produire de telles populations est définie dans la section 3 ainsi que la démarche d'assignation des lieux d'occupation aux personnes et ménages. Nous présentons ensuite la plateforme TransNetSim de simulation multi-agent que nous avons développée et couplée avec un outil SOLAP (Section 4). Le SOLAP a permis d'inspecter et d'analyser les données de la population synthétique ainsi que les résultats des

4 Revue Internationale de Géomatique. Volume X – n° x/2009

simulations des déplacements effectuées sur l'EUP de Québec. La section 5 discute de l'intérêt de l'approche proposée et conclut l'article.

2. Modélisation multi-échelle d'Environnements Urbains Virtuels

Les patrons de conception, qui ont connu un grand succès dans le domaine du génie-logiciel (Gamma *et al.*, 1995), permettent d'augmenter la qualité du résultat, notamment en appliquant des solutions déjà existantes à des problèmes courants de conception. Ces patrons sont décrits sous une forme abstraite, sans s'attacher aux détails du problème à résoudre. Pour la modélisation multi-échelle de l'*Environnement Urbain Virtuel* (EUV) nous avons déjà proposé dans (Chaker *et al.*, 2007) un patron qui intègre à la fois les lieux et le réseau de transport multi-modal. Ce sont les deux composantes de base de l'EUV dans notre approche (Figure 1). Un lieu est un concept générique qui matérialise la localisation spatiale d'une activité bien déterminée. Le lieu, tout comme l'activité, ne prennent de sens que si l'échelle spatio-temporelle est précisée. Par exemple, à une échelle *A*, on parle des activités de travail, magasinage, loisir et donc des activités quotidiennes qui se déroulent à l'intérieur de la ville. À une autre échelle spatio-temporelle *B*, l'activité pourrait signifier le déplacement d'une personne d'une ville à une autre. À l'échelle *A*, un lieu est un bâtiment, une parcelle ou tout ce à quoi une adresse civique pourrait référer, alors qu'à l'échelle *B* la ville elle-même est un lieu. Par conséquent, le réseau de transport doit coïncider avec la sémantique du lieu car, avant tout, le réseau de transport n'est que le moyen de se déplacer d'un lieu à un autre.



Figure 1: composantes de base du patron de conception proposé

En plus de la représentation sémantique qui est très dépendante de l'échelle comme nous pouvons le constater, notre modèle doit rassembler des descriptions géométriques et topologiques de l'EUV. Cependant, nous ne sommes pas à la recherche d'une description géométrique complète de l'espace. Rappelons que l'objectif de modéliser notre EUV est d'offrir un cadre d'application pour les simulations en transport et que le rendu visuel de l'environnement urbain n'est pas une finalité en soi. Il est important de préciser aussi que nous modélisons l'environnement spatial à partir de données SIG (donc des plans ou des cartes SIG) et que ces données de départ sont déjà une abstraction de l'environnement réel. Chaque carte ou plan est donc définie selon une échelle cartographique et

représente une abstraction à la fois géométrique (forme des éléments) et sémantique (organisation des éléments en couches) à cette échelle. Les approches de généralisation cartographique (Mackaness *et al.*, 2007) ne peuvent alors que renforcer notre démarche en lui permettant de générer systématiquement les données source.

Au niveau de la représentation topologique nécessaire au raisonnement spatial, en particulier la recherche d'itinéraire dans notre cas, nous adoptons le standard du graphe composé de liens (ou arêtes) et de nœuds (ou sommets). Comment alors une représentation aussi simple que le graphe pourrait être adaptée, interprétée ou enrichie pour qu'elle puisse prendre en compte toutes les contraintes de modélisation de l'EUUV ? Les deux sous-sections suivantes traitent cette question. Nous retenons les caractéristiques principales suivantes que le patron de conception doit présenter : 1) Intégration du réseau de transport et des lieux comme étant les composantes de base de l'EUUV ; 2) Prise en compte de la variabilité sémantique du lieu et du réseau dépendamment de l'échelle ; 3) Représentation géométrique adaptée aux besoins de l'EUUV et 4) Représentation topologique du réseau à l'aide de graphes logiques.

2.1. Modélisation conceptuelle

Notre patron de conception est décrit selon la notation UML dans la Figure 2. Nous ne donnons dans cette description que les éléments de base du modèle qui sont indépendants de l'échelle de la représentation. Des liens d'association, d'héritage et d'agrégation relient les différentes classes abstraites du patron de conception. Ni les classes, ni les liens ne sont détaillés au niveau du patron car les détails peuvent varier d'une implémentation à une autre, donc d'une échelle à une autre. Nous pouvons alors remarquer que le lien entre les classes « Nœud » et « Lien » n'a pas de cardinalités et que seulement deux attributs sont donnés pour l'ensemble des classes. Voici une description de chacune des classes abstraites et de leurs relations.

– Classes « *Lien* » et « *Nœud* » : nous considérons que le réseau de transport est dans tous les cas représenté sous forme d'un graphe logique composé de liens et de nœuds. Cependant, il s'agit ici de concepts génériques et un lien peut représenter aussi bien une route qu'une voie faisant partie d'un segment de route (« *Lane-Based Networks* » (Malaikrisanachalee *et al.*, 2005)). Le type d'un lien (attribut *type* de la classe *lien*) permet de différencier le réseau routier du réseau cyclable, piétonnier ou ferroviaire, etc. L'association entre *nœud* et *lien* n'a pas de cardinalité dans ce méta-modèle car elle ne peut se préciser qu'au moment où on précise ces deux concepts dépendamment de l'échelle choisie.

– Classe « *Lieu* » : C'est le concept générique *Lieu* dont la sémantique précise est liée à l'échelle. La Figure 3 complète le modèle en raffinant la notion de lieu. Ainsi, le concept *Lieu* peut être un *Lieu composite* ou un *Lieu élémentaire* (les classes « *Lieu composite* » et « *Lieu élémentaire* » héritent de la classe « *Lieu* »). Cette distinction offre une grande flexibilité au modèle et elle est même

indispensable pour pouvoir desservir à une même échelle, aussi bien des lieux simples qui représentent les origines et destinations de trajets, que des lieux complexes qui sont désagrégés en plusieurs lieux appartenant à une autre échelle plus fine. Le tableau 1 donne un exemple de spécification sémantique des *lieux composites* et *élémentaires* dépendamment des échelles visées.

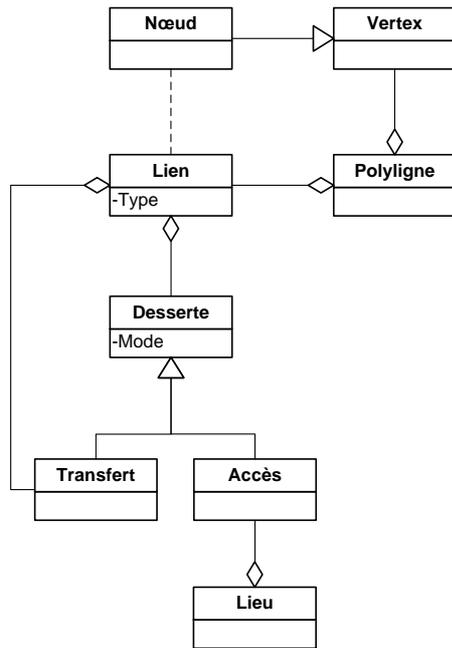


Figure 2: Patron de conception de l'EUUV (notation UML)

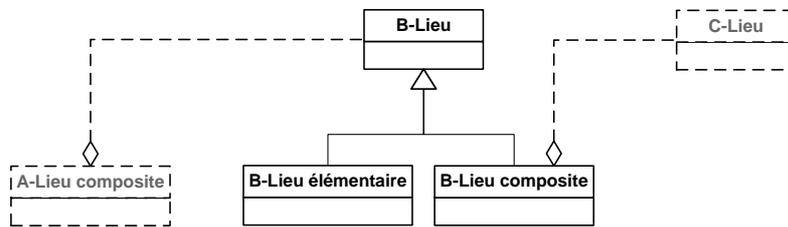


Figure 3: Raffinement de la notion de lieu

Tableau 1: Exemple de spécification sémantique des échelles

Échelle	Sémantique du Lieu Composite	Exemple de Lieu composite	Sémantique du Lieu Élémentaire	Exemple de Lieu Élémentaire
Macro	Agglomération urbaine	Communauté Métropolitaine de Québec	Adresse rurale	route 132, kilomètre 24, direction EST
Méso	Quartier	Université Laval	Adresse urbaine	450, chemin du golf, Québec
Micro	Édifice	Pav. Pouliot de l'université Laval	Lieu extérieur	Terrain de soccer N°3 de l'université Laval

– Classes « *Polyligne* » et « *Vertex* » : La géométrie *Polyligne* (un vecteur de points) est un concept propre au lien et sert à préciser le fait que plusieurs liens de différents types peuvent partager la même géométrie à une certaine échelle. À une échelle *A*, par exemple, le réseau routier peut se confondre avec le réseau piétonnier. Ceci n'est peut être pas valable à une échelle *B* plus fine, dans laquelle un trottoir est représenté à part et avec ses propres liens. Donc un lien possède une et une seule géométrie et inversement, une même géométrie peut être propre à un seul lien, mais peut aussi être partagée par plusieurs liens. Un polyligne est un agrégat ordonné d'au moins deux vertex. Un nœud est aussi un vertex particulier qui permet de rattacher deux ou plusieurs liens, d'où la relation d'héritage entre la classe « *Nœud* » et la classe « *Vertex* ». La définition exacte du vertex dépend du cadre de référence spatial choisi. Le nombre de vertex qui définissent un polyligne dépend non seulement de la géométrie du *Lien* auquel il est associé mais aussi du niveau de l'abstraction géométrique (ou résolution). Nous dissociions ainsi la notion d'échelle de l'EUV de la notion de résolution ou d'abstraction géométrique. En conséquence, la résolution spatiale peut varier tout en maintenant la même échelle de l'EUV.

– Classes « *Desserte* », « *Accès* » et « *Transfert* » : Nous considérons que la desserte est un concept générique qui formalise aussi bien la desserte d'un lieu à partir d'un réseau de transport (sous-classe « *Accès* ») que la desserte d'un réseau à partir d'un autre réseau (sous-classe « *Transfert* »). Dans les deux cas, la desserte est accompagnée d'un mode de transport (attribut *Mode* dans la classe *desserte*) qui caractérise le mode de transport avec lequel la desserte s'effectue. Il est important de distinguer ici le mode de transport (attribut de la classe *desserte*) et le type du réseau de transport (attribut de la classe *Lien*). En effet, pour un type particulier comme « réseau routier » il peut y avoir plusieurs dessertes utilisant différents modes comme « autobus » et « automobile ». À une échelle *A*, le premier accès sera une station de bus alors que le deuxième sera un espace de stationnement. À une échelle *B* plus grande (donc plus précise), il sera plus commode de considérer la station de bus comme un transfert de type « autobus » permettant de relier un lien de type « réseau routier » à un lien de type « réseau piétonnier ».

Nous avons déjà donné dans (Chaker *et al.*, 2009) des exemples montrant l'expressivité et la généralité du patron de conception que nous proposons. La notion de desserte, qui permet à la fois de modéliser le transfert modal et l'accès aux lieux, joue un rôle central dans notre modèle. Ce rôle est similaire au rôle du *Verbe* en langage naturel qui est celui d'organiser une phrase en indiquant l'état, le changement, l'action qui se produit. Dans notre cas, la *Desserte* organise une portion d'une scène spatiale en indiquant comment les éléments sont connectés, comment ils forment un espace navigable et comment on utilise ou on change de mode de transport au sein de cet espace. La généralité du patron proposé se matérialise aussi dans son indépendance vis-à-vis de l'échelle spatiale. Nous avons alors donné dans (Chaker *et al.*, 2009) un exemple complet de modélisation aux échelles macro, méso et micro et l'exemple montrait bien que le patron subissait peu de changements lors de son instanciation.

2.2. Formalisation en théorie des graphes

La théorie des graphes offre un moyen simple de définir formellement notre EUV. Nous passons ainsi de la représentation conceptuelle à la représentation logique de l'EUV. À l'aide d'un graphe orienté nous définissons le **T-graphe** dans lequel chaque sous-graphe ou composante est relative à un type de réseau bien déterminé. Le T-graphe regroupe aussi les concepts de lieu, d'accès et de transfert. Nous introduisons ensuite deux variantes du T-graphe, à savoir le T-graphe pondéré et le T-graphe hiérarchique qui utilise la définition de forêt, ce qui nous permet de définir un EUV.

Définition 1 (rappel) : Graphe orienté

Un graphe orienté G est un couple (V, A) , où : V est un ensemble dont les éléments sont des sommets ; A est un ensemble de couples ordonnés de sommets appelés arcs.

Le formalisme de graphe orienté est souvent utilisé pour définir de nombreux problèmes en recherche opérationnelle, notamment ceux qui s'appliquent aux transports (Mathis, 2003) comme le problème du voyageur de commerce ou celui du plus court chemin. Dans ces problèmes, les sommets du graphe sont considérés comme des lieux de visite, des villes le plus souvent. Cette définition est insuffisante pour définir notre EUV multi-échelle car elle confond la notion de lieu avec la notion d'intersection qu'on ne peut représenter qu'avec des sommets. Par ailleurs, la notion de graphe orienté ne permet pas la multiplicité des arcs reliant les mêmes sommets, ce qui rend difficile la définition de plusieurs modes de transport dans un même graphe. Cependant, une variante du graphe orienté, appelée *multi-graphe*, permet de définir plusieurs arcs entre les mêmes sommets. Comme les intersections ne sont pas souvent partagées par différents réseaux, le *multi-graphe* est aussi peu adéquat pour la représentation logique de notre EUV. Il faut alors différencier non seulement les arcs, mais aussi les sommets. Dans la littérature, certains auteurs utilisent le formalisme de *hypergraphe* ou *h-graphe* pour définir des *hypergraphes multimodaux* (Lozano *et al.*, 2002). Les

hypergraphes généralisent la notion de graphe dans le sens que les arêtes ne relient plus un ou deux sommets, mais un nombre quelconque de sommets. Plus récemment, Ayed et ses collègues ont défini le formalisme de *Transfer Graph* dans lequel chaque composante est un réseau représenté par un graphe orienté auquel on rajoute des transferts entre des sommets appartenant à des composantes différentes (Ayed *et al.*, 2008). Ces transferts sont des arcs particuliers qui représentent des transferts modaux. Cependant, ni la nature multi-échelle, ni la notion d'accès aux lieux n'est prise en compte dans les *Transfer graphs*. La définition 3.3 suivante introduit donc notre notion de *T-graphe* (T venant de 'Transport') qui servira de support pour la représentation logique d'un EUV multi-échelle.

Définition 2 : T-graphe

Un T-graphe G est un quadruplet (GS, T, L, P), où : 1) *GS est un ensemble dont les éléments sont des graphes orientés appelés composantes ; 2) T est un multiensemble dont le support est un ensemble de couples ordonnés d'arcs, appelés transferts et tel que si $((i, j), (k, l)) \in T$ alors (i, j) et (k, l) n'appartiennent pas au même graphe de GS ; 3) L est un ensemble dont les éléments sont appelés lieux élémentaires; 4) P est un multiensemble dont le support est un ensemble de couples (arc, lieu élémentaire) appelés accès ou portes.*

Dans le T-graphe on définit des composantes dont chacune d'elle n'est autre qu'un graphe orienté qu'on pourra associer à un type de réseau (routier, piétonnier, ferroviaire, etc.). Ces composantes ne sont reliées entre elles que par des liens spéciaux, les *transferts*. Ces transferts se définissent entre deux arcs appartenant à deux composantes différentes, c'est-à-dire deux réseaux différents. En théorie des graphes, le graphe est souvent pondéré par une fonction de coût qui associe à chaque arc une valeur numérique qui représente la distance ou le temps de parcours nécessaire pour atteindre le 'sommets puits' à partir du 'sommets source'. De la même façon, nous introduisons le T-graphe pondéré dans la définition 3 suivante.

Définition 3 : T-graphe pondéré

Soient : G = (GS, T, L, P) un T-graphe ; GS = {G₁, G₂, ... , G_q} les composantes de G ; V_g = {v^g₁, v^g₂, ... , v^g_n} les sommets de V_g ; (i,j)^g l'arc allant de v^g_i à v^g_j et N l'ensemble des entiers naturels. G est un T-graphe pondéré si et seulement si :

1. chaque composante G_g = (V_g, A_g) est associée aux fonctions suivantes : une fonction de coût de parcours $\varphi_g : A_g \rightarrow N$; Fonction de capacité $\psi_g : A_g \rightarrow N$; Fonction de pénalité $\chi_g : A_g \times A_g \rightarrow N$ telle que $\chi_g((i,j)^g, (k,l)^g) = 0$ si $j \neq k$.

2. G est associé aux fonctions suivantes : une fonction de coût d'accès $\alpha : P \rightarrow N$; Fonction de coût de transfert $\beta : T \rightarrow N \times N$ telle que si $\beta(t) = (x_1, x_2)$ avec $t = ((i,j)^g, (k,l)^{g'})$ alors $x_1 \leq \varphi((i, j)^g)$ et $x_2 \leq \varphi((k, l)^{g'})$

Les coûts de parcours sont les coûts classiques qui représentent des temps ou des distances, alors que la capacité d'un arc $(i,j)^g$ appartenant au graphe $G_g = (V_g, A_g)$

et notée $\psi_g((i,j))$ est le nombre maximal d'entités (véhicules, piétons, etc.) que l'arc $(i,j)^g$ peut contenir. Ces informations sont utiles lors de la planification des itinéraires et aussi lors des simulations des déplacements. La fonction de pénalité est aussi une fonction importante qui associe des pénalités aux intersections, donc aux sommets. Dans la définition 3 nous introduisons la fonction de pénalité sur des couples d'arcs consécutifs. Autrement dit, les pénalités ne sont définies que pour un couple d'arcs consécutifs, c'est-à-dire deux arcs tels que le sommet destination de l'un est le sommet origine de l'autre. Nous aurions pu définir la fonction de pénalité sur l'ensemble des sommets en associant à chaque sommet une matrice dont la taille dépend du degré de chaque sommet, ce qui rendrait cette définition mathématique plus compliquée dans ce cas. Les fonctions de coût d'accès et de coût de transfert, quant à elles, sont des fonctions sur les accès (l'ensemble P) et les transferts (l'ensemble T). Ces coûts sont des coûts de parcours mais réservés aux accès et aux transferts. Dans la définition 3, le coût de transfert est un couple de deux valeurs, la première valeur indique le temps ou la distance au niveau de l'arc origine du transfert et la deuxième valeur est celle relative à l'arc destination. On précise aussi que ces valeurs doivent être inférieures ou égales aux coûts de parcours des arcs auxquels elles sont reliées (l'arc origine pour la première valeur et l'arc destination pour la deuxième valeur).

Définition 4 (rappel) : Forêt

Une forêt est un graphe (non orienté) acyclique.

La définition de Forêt sert de support à la définition 5 dans laquelle on formalise le T-graphe hiérarchique. On rappelle qu'un graphe acyclique est un graphe ne contenant aucun cycle et que dans un graphe non orienté, on appelle cycle une suite d'arêtes consécutives (chaîne) dont les deux sommets extrémités sont identiques. Chacune des composantes connexes d'une forêt est un arbre. Dans un graphe non orienté, le degré d'un sommet est le nombre d'arêtes auxquelles ce sommet appartient. On distingue deux types de sommets dans un arbre : les feuilles dont le degré est 1 ; les nœuds internes dont le degré est supérieur à 1.

Définition 5 : T-graphe hiérarchique

Un T-graphe $G = (GS, T, L, P)$ est hiérarchique si G est associé à une forêt F dont l'ensemble des feuilles est l'ensemble L des lieux élémentaires de G . Les nœuds de F sont appelés lieux composites.

La figure 4 suivante schématise un T-graphe hiérarchique ayant un seul lieu composite. Remarquons que le lieu composite ne peut pas avoir d'accès. Le lieu composite est un concept abstrait qui permet d'agréger des lieux élémentaires. La définition 5 le montre bien d'ailleurs car les nœuds de la forêt associée au T-graphe ne sont utilisés nulle part dans le modèle. Par ailleurs, dans l'optique d'un modèle évolutif, un lieu élémentaire peut devenir un lieu composite. Dans le cas d'un tel raffinement, les « anciens » accès du lieu deviennent des transferts. Ces derniers sont des transferts d'échelle alors que les autres sont des transferts de mode. Les deux types de transferts se caractérisent logiquement de la même manière, raison

pour laquelle on n'en fait pas de distinction dans les définitions précédentes. Nous pouvons aussi remarquer que la notion d'échelle introduite lors du modèle conceptuel se traduit dans le modèle logique par le niveau de profondeur (ou génération) de la forêt F associée au T-graphe. L'échelle est donc une propriété intrinsèque de l'EUV tel que nous le définissons ici.

Définition 6 : *Environnement Urban Virtuel*

Un Environnement Urban Virtuel (EUV) est un T-graphe pondéré et hiérarchique. On note $EUV = (G, F, \Omega)$ l'environnement urbain virtuel où $G = (GS, T, L, P)$ est son T-graphe, F sa forêt et Ω l'ensemble de ses fonctions de pondération.

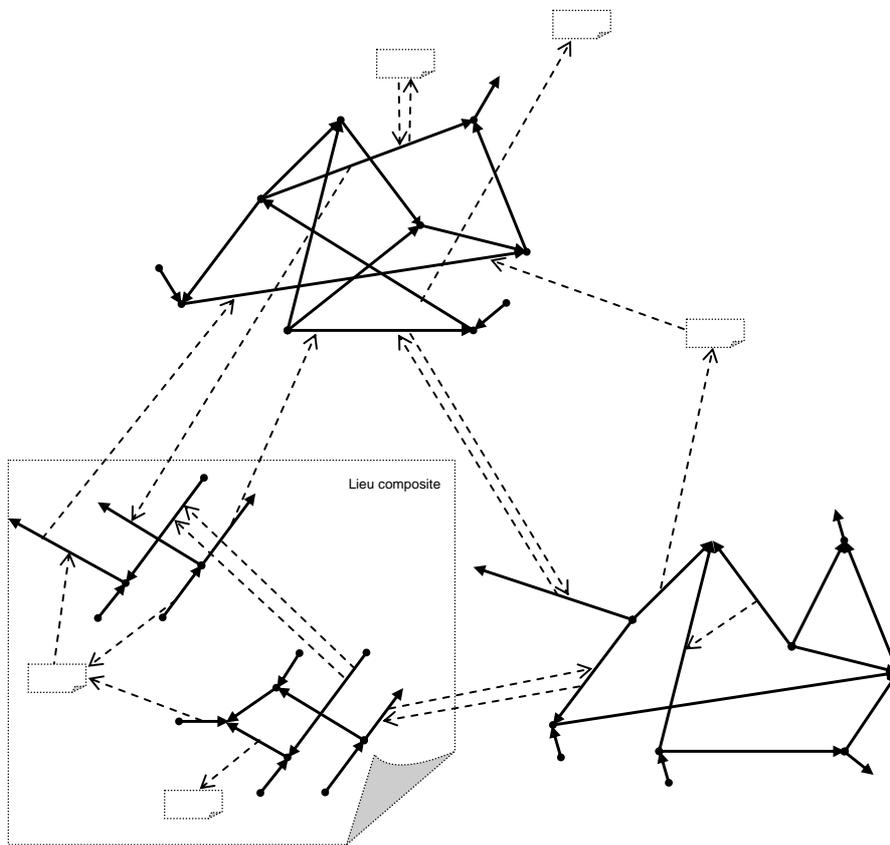


Figure 4: Exemple d'un T-graphe hiérarchique

En conclusion, nous avons pu définir formellement un EUV (définition 6) grâce à la définition du T-graphe (définition 2). Cette formalisation en théorie des

graphes permet une implantation rigoureuse de ces concepts, tout en tenant compte des principales fonctions classiquement utilisées pour la simulation des parcours dans les réseaux de transports.

3. De l'Environnement Urbain Virtuel à l'Environnement Urbain Peuplé

L'EUV que nous avons défini dans la section précédente, représente la composante spatiale d'un environnement plus global, un environnement qui inclut à la fois l'offre et la demande en transports. Nous appelons cet environnement l'Environnement Urbain Peuplé (EUP). Un EUP, comme le montre la figure 5, est un EUV auquel on rajoute la population qui l'occupe. On parle ici de l'occupation spatiale qui est le plus souvent un lieu. Dans un EUP, nous modélisons la population entière qui occupe l'EUV à l'échelle la plus large. Dépendamment de la précision recherchée et du niveau de raffinement de l'EUV, l'occupation d'une personne peut être un lieu élémentaire, comme elle peut être un lieu composite. On peut alors spécifier qu'une personne habite une ville, ou bien un quartier de la ville, une maison du quartier ou un appartement comme étant un local dans un édifice.

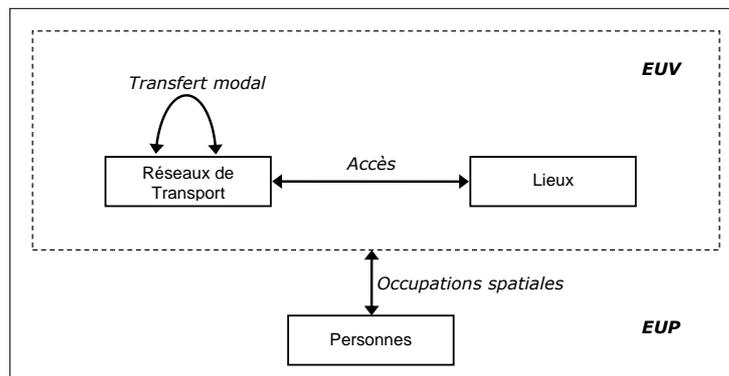


Figure 5: Composantes de l'EUP

La raison pour laquelle on s'intéresse à modéliser au sein de l'EUP l'ensemble de la population occupant un EUV est la suivante. La demande en transports tire son origine de la décision des individus, qui selon leurs besoins et en fonction des moyens dont ils disposent pour se déplacer, choisissent leurs destinations, leurs moments de départ et leurs modes de déplacement. En modélisant une *population synthétique* qui aurait des caractéristiques similaires en terme de composition de ménages et des propriétés des individus, on remonte à la source du problème de modélisation de la demande en transport. Dans les approches classiques de simulations spatiales multi-agents, l'environnement se limite à l'environnement spatial dans lequel les agents évoluent et effectuent leurs mouvements de déplacement. Dans ce cas, la demande en transport fait partie de l'input du

simulateur, ce qu'on appelle le *scénario* dans MAGS (Moulin *et al.*, 2003) par exemple ou les *matrices OD* dans les simulateurs de trafic. D'autres approches en transports font appel aux modèles d'*assignation dynamique du trafic* (DTA de *Dynamic Traffic Assignment*). Ces modèles ne prévoient pas à l'avance le comportement du voyageur en lui laissant la possibilité de modifier ses choix de déplacement (son trajet, son mode, etc.) durant la simulation en fonction de l'environnement (l'état du trafic, le plus souvent). Ces dernières approches favorisent la modélisation de populations synthétiques et leurs associations à leurs environnements spatiaux. En consultant la littérature, nous avons ainsi remarqué un intérêt croissant pour les méthodes de génération de populations synthétiques (Arentze *et al.*, 2007) (Guo *et al.*, 2007), surtout depuis le célèbre projet TRANSIMS. (Barett *et al.*, 2004).

3.1. Démarche de génération de la population synthétique

Nous désignons par le terme *population synthétique*, la population artificielle ou virtuelle associée à l'EUP. Comme l'EUP est un concept qui représente un espace urbain réel (une ville par exemple), la population synthétique doit représenter la population réelle occupant cet espace. La méthode de génération de telles populations doit alors certifier leurs plausibilités, leurs vraisemblances. Il existe de nombreux travaux qui proposent de telles méthodes. On retrouve généralement deux façons de produire une population synthétique : la *reconstruction synthétique* et l'*optimisation combinatoire*. La première est la plus connue et repose sur une démarche d'échantillonnage probabiliste en utilisant des probabilités conditionnelles. L'approche par optimisation combinatoire, quant à elle, repose sur la pondération de données existantes d'une population occupant un espace plus large que celui qu'on veut générer (Williamson *et al.*, 1998). Dans les travaux existants de simulation des déplacements à l'échelle de toute une ville, on fait souvent référence au projet TRANSIMS en ce qui concerne la démarche de production de la population synthétique. Les détails de cette démarche se retrouvent dans (Beckman *et al.*, 1996). Cette méthode qui utilise seulement des données de recensement est composée de deux étapes. D'abord, un tableau croisé des répartitions démographiques est estimé en utilisant la méthode IPF (*Iterative Proportional Fitting*) qui a l'avantage de représenter la corrélation de plus que deux attributs. Ainsi, si on génère une population d'individus avec seulement les attributs a , b et c , on peut facilement ajouter un attribut d en utilisant la probabilité conditionnelle $P(d/a,b,c)$ qui dérive de $P(a,b,c,d)$. La deuxième étape consiste à générer la population synthétique à partir de la distribution obtenue à la manière de l'approche par reconstruction synthétique. La génération de populations synthétiques continue à faire l'objet de recherches très récentes. Dans (Guo *et al.*, 2007) une extension de la méthode de Beckman est proposée alors que dans (Arentze *et al.*, 2007) la méthode IPF est reprise.

La méthode de génération de la population synthétique que nous proposons s'appuie sur la technique de reconstruction synthétique mais intègre d'autres

techniques d'apprentissage automatique dans le but d'obtenir une méthode plus générique s'appliquant à un grand nombre de cas. En effet, nous avons remarqué que chacune des méthodes existantes (Beckman *et al.*, 1996) (Guo *et al.*, 2007) (Arentze *et al.*, 2007) est spécifique à un projet particulier et ne propose que des solutions dépendantes des données utilisées. Contrairement aux approches citées précédemment, notre méthode sépare le problème de génération de la population synthétique en deux sous problèmes : 1) la génération des ménages et individus; 2) l'assignation des ménages et individus à l'environnement spatial, c'est-à-dire aux lieux d'occupation tels que les lieux de résidence, de travail ou d'étude. La première phase est celle que nous détaillons ici et c'est tout simplement la phase de génération de la population synthétique. La deuxième phase de la méthode ne sera détaillée que dans la section 4.4 et c'est la phase qui permet d'obtenir l'EUP en affectant les personnes et ménages aux lieux d'occupations.

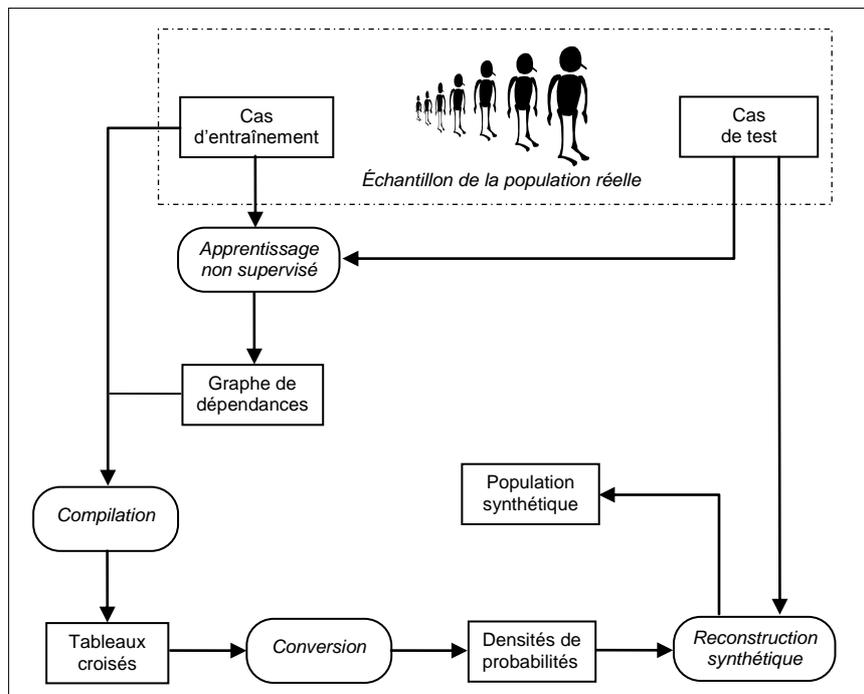


Figure 6: Démarche proposée pour la génération de populations synthétiques

La figure 6 présente une vue schématique de la méthode proposée. Cette méthode prend comme hypothèse le fait qu'on dispose des données relatives à un échantillon réel de la population. Ces données sont généralement le résultat d'enquêtes menées par des organismes gouvernementaux. Comme dans les

approches d'apprentissage automatique, nous séparons les cas d'entrée (c'est à dire les personnes et les ménages de l'échantillon de la population réelle) en deux sous-ensembles : les cas d'entraînements et les cas de test. Le premier ensemble sert à déterminer le modèle de prédiction et le deuxième sert à valider ce modèle. La figure 6 montre qu'un *algorithme d'apprentissage non-supervisé* est utilisé dans une première étape pour générer un *graphe de dépendances*. L'apprentissage non-supervisé est, en effet, une méthode d'apprentissage automatique qui se distingue de l'apprentissage supervisé par le fait qu'elle opère sur des données brutes sans aucun attribut de prédiction. Dans notre cas, l'objectif est de déterminer quels sont les attributs les plus corrélés, les plus dépendants les uns des autres. Le *graphe de dépendances* est donc un graphe qui montre ces corrélations. Un tel graphe peut être obtenu avec un algorithme utilisant la mesure de l'entropie de Shannon qui caractérise l'homogénéité des instances et qui permet de quantifier les interactions.

Ayant obtenu un graphe de dépendances, il est possible de regrouper les attributs fortement corrélés et de générer un tableau croisé pour chaque cluster. Par exemple, si le graphe de dépendances révèle une forte corrélation entre l'âge et l'occupation de la personne (travailleur, étudiant, etc.), on génère systématiquement un tableau qui croise les valeurs de ces deux attributs pour les cas d'entraînement. Cette étape est appelée *l'étape de compilation* dans notre démarche. Les tableaux croisés sont, ensuite, convertis en densités de probabilités conditionnelles. Ces distributions de probabilités forment les entrées de *l'étape de reconstruction synthétique*, par laquelle l'algorithme probabiliste tire des valeurs aléatoires qui respectent ces distributions. La méthode que nous venons de décrire ne concerne que les attributs non spatiaux des personnes et ménages. L'assignation des lieux d'occupation aux personnes et ménages fait l'objet d'une approche séparée, celle que nous décrivons dans la section suivante.

3.2. Assignation des lieux d'occupation

L'assignation des lieux d'occupation aux personnes et ménages de la population synthétique est un problème qui, selon nous, se traite différemment de l'affectation des attributs non spatiaux, donc de la génération de la population synthétique elle-même. Notre hypothèse à ce niveau est que nous disposons d'une population synthétique et d'un EUV. Les trois étapes qui constituent notre démarche de construction d'un EUP sont : ajustement de l'offre à la demande, pondération des relations et application d'un algorithme d'assignation.

3.2.1. Ajustement de l'offre à la demande

Pour pouvoir assigner les lieux aux personnes (ou aux ménages), il faut ajuster le nombre de lieux au nombre de personnes. Pour prendre un exemple concret, pour assigner les lieux de travail aux enseignants, il faut que la somme des postes de travail liés aux lieux d'enseignement coïncide avec le nombre de travailleurs enseignants de la population synthétique. On ne s'intéresse pas ici à la modélisation du processus d'emploi en soi, un exercice fortement intéressant mais

qui sort du cadre de notre étude. Comme les processus de génération de la population synthétique et de construction de l'EUV sont des processus d'approximation, il est très difficile d'obtenir du premier coup une coïncidence parfaite entre l'offre et la demande. C'est la raison pour laquelle nous proposons une étape d'équilibrage durant laquelle le concepteur revient sur les deux processus et essaye d'ajuster certaines variables pour que leurs résultats convergent. La modalité des ajustements dépend de la vraisemblance de l'EUV par rapport à celle de la population synthétique. Cela permet de savoir laquelle des variables ajuster. La vraisemblance de l'EUV relève plutôt de l'exactitude des données SIG alors que la vraisemblance de la population synthétique peut se mesurer par la considération de plusieurs indices, chacun relatif à une étape particulière de la génération et de l'affectation de la population.

3.2.2. Pondération des relations

À ce niveau, nous disposons de deux ensembles L et P de mêmes cardinalités, soient $L = \{L_1, L_2, \dots, L_n\}$ et $P = \{P_1, P_2, \dots, P_n\}$. L représente les lieux d'occupations et P représente les personnes devant occuper ces lieux. L'objet de cette étape est de quantifier les valeurs des relations entre les éléments de L et les éléments de P . Ces relations traduisent des degrés de vraisemblance d'assignation. Si un lieu L_i est fortement relié à une personne P_j , cela veut dire que dans la réalité une personne ayant les mêmes caractéristiques que P_j aura tendance à occuper le lieu L_i . On se propose donc de définir une fonction de poids $\mu : L \times P \rightarrow \mathfrak{R}$. L'approche de calcul des $\mu(L_i, P_j)$ dépend du type d'occupation en question. Par exemple, pour pondérer les relations entre les lieux de travail et les travailleurs, nous proposons une approche basée sur la prédiction ou l'estimation des temps de parcours. Ainsi, on se propose de déterminer pour chaque travailleur de la population synthétique, son *temps prédit* pour aller au travail, c'est-à-dire la durée que cette personne « devrait normalement » mettre pour aller de son domicile à son lieu de travail. La prédiction des temps de parcours est un problème bien connu en transports et plusieurs méthodes existent dans la littérature, notamment des méthodes statistiques comme les méthodes de régression (Wu *et al.*, 2004) et aussi des méthodes d'intelligence artificielle comme les méthodes d'apprentissage automatique (Vanajakshi *et al.*, 2007). Contrairement au cas de la génération de la population synthétique (section 3.1), ce sont les méthodes d'apprentissage supervisé qui s'appliquent ici. En effet, nous supposons que dans l'échantillon réel de la population, le temps de parcours (la variable à prédire) est fourni. Une fois que les *temps prédits* sont déterminés, le calcul des pondérations est directement donné par la formule suivante :

$$\mu(L_i, P_j) = \frac{a}{|TP(P_j) - TM(L_i, LR(P_j))|} + b \quad [1]$$

Avec $\mu(L_i, P_j)$ le poids entre le lieu L_i et la personne P_j ; a et b deux constantes de normalisation; $TP(P_j)$ le Temps Prédit de la personne P_j ; $LR(P_j)$ le Lieu de

Résidence de la personne P_j et $TM(L_i, L_j)$ le Temps Mesuré pour aller du lieu L_i au lieu L_j .

La formule 1 traduit tout simplement le fait que le poids est inversement proportionnel à l'écart absolu entre : 1) le temps de parcours mesuré du lieu de résidence de la personne P_j vers le lieu de travail potentiel L_i ; et 2) le temps de parcours prédit pour que la personne P_j se rende au travail. Donc, plus cet écart entre le temps prédit et le temps mesuré est grand, plus le poids de la relation entre la personne et le lieu potentiel (ou candidat) d'occupation est faible. Dans le cas d'une activité autre que le travail, la formule s'interprète de la même façon pour l'activité en question.

3.2.3. Choix de l'algorithme d'assignation

Ayant les deux ensembles $L = \{L_1, L_2, \dots, L_n\}$ et $P = \{P_1, P_2, \dots, P_n\}$, ainsi que la fonction de poids μ définie sur leur produit cartésien, on s'intéresse maintenant à déterminer n couples (L_i, P_j) qui maximisent la fonction μ . Le problème devient alors typiquement un problème d'optimisation et il existe dans la littérature plusieurs algorithmes candidats à sa résolution. Plus précisément, on est dans le cas du problème du couplage parfait et à moindre coût dans un graphe biparti. Ce problème théorique peut se résoudre au moyen d'un algorithme de recherche de flot de coût minimum. L'une des implémentations les plus efficaces a été proposée par Goldberg (Goldberg, 1997), et c'est celle que nous avons appliquée sur la population synthétique de Québec (section 4 suivante).

4. Simulation des déplacements quotidiens dans la ville de Québec

Grâce à un ensemble de données récoltées sur la ville de Québec, nous avons pu mettre en œuvre notre démarche de construction de l'EUP, ce qui nous a permis par la suite de simuler la mobilité de l'EUP de Québec. TransNetSim est la plateforme de simulation multi-échelle de la mobilité que nous avons développée et que nous présentons à la section 4.2 juste après l'introduction de l'EUP de Québec à la section 4.1. L'analyse de l'EUP de Québec ainsi que l'analyse des résultats des simulations à l'aide de l'outil SOLAP font l'objet de la section 4.3.

4.1. EUP de la ville de Québec

Pour la construction de l'EUP de Québec, l'enquête OD (Origine, Destination) a été notre source principale de données que nous avons jumelée avec des données d'autres sources (recensement, rôles d'évaluation, données d'emploi Québec, etc.) afin de garantir la représentativité de l'EUP généré. L'enquête OD a été administrée par le MTQ (Ministère de Transports de Québec) et le RTC (Réseau de Transports de la Capitale) et donne des informations sur les personnes et les ménages représentant 8% de la population totale de la zone métropolitaine de Québec. Chaque répondant décrivait les activités de déplacement des membres de son ménage pour une journée prise parmi les jours ouvrables de la semaine. Les

répondants donnent alors le motif de chaque déplacement ainsi que les modes de transport utilisés et précisent les lieux d'origine et de destination¹. Les personnes interrogées représentent donc l'échantillon de la population réelle dans notre démarche de génération de la population synthétique (figure 6). Le même échantillon a servi aussi pour la prédiction des temps de parcours domicile-travail afin de pouvoir assigner un lieu de travail à chaque travailleur de la population synthétique selon l'approche de la section 3.2. L'EUP de Québec ainsi construit possède les caractéristiques suivantes :

- un méso-EUV (c'est-à-dire un EUV défini à une échelle qualifiée Méso) couvrant la zone métropolitaine de Québec et constitué de 32071 nœuds et 80855 liens. Des coûts de parcours ont été associés aux liens ainsi que des pénalités aux croisements ont été associés aux nœuds (0.4 minute pour le virage à gauche, 0.2 pour le virage à droite et 0.1 pour aller tout droit). La classification de chaque virage a été faite grâce à une fonctionnalité que nous avons programmée sur les données SIG sources.

- un micro-EUV (c'est-à-dire un EUV défini à une échelle qualifiée Micro) couvrant le campus de l'université Laval qui est un lieu composite du méso-EUV constitué de 13 accès représentant les aires de stationnement du campus, 11 liens de transferts dont 5 entrées et 6 sorties.

- une population synthétique associée au méso-EUV constituée de 325382 ménages et 734860 personnes. Chaque ménage est associé à un lieu de résidence et chaque personne caractérisée par l'âge, le genre, l'occupation, l'emploi, etc. est associée à un lieu d'occupation (un lieu de travail si la personne travaille, un lieu d'étude si la personne étudie et rien sinon).

Au niveau de la génération de population synthétique, les corrélations entre les attributs ont été manuellement identifiées en s'appuyant sur l'expertise d'un spécialiste. Il est certain que le recours à un graphe de dépendances comme le prévoyait notre démarche serait plus sécuritaire sauf que celui-ci nécessite un accès direct aux données confidentielles de l'enquête OD. Pour ce qui est de l'étape de l'assignation des lieux d'occupation, nous avons classé les emplois selon 5 secteurs (éducation, santé, services, commerce en gros, commerce de détail, gouvernement, manufacture et loisir) et 5 types (gestion, professionnel, employé spécialisé ou ouvrier qualifié, employé non spécialisé et ouvrier non qualifié)². L'assignation totale a réussi à maintenir un écart moyen entre les temps prédits et les temps alloués de 4.32 minutes. Les temps prédits ont été estimés avec une approche statistique en utilisant une méthode de régression.

4.2. *TransNetSim : une plateforme de simulation multi-agent du trafic*

¹ Plus de détails sur l'enquête OD sont disponibles sur <http://www.mtq.gouv.qc.ca>

² 40 graphes bipartis ont été générés, le plus gros avait 31447 nœuds et son couplage a été résolu par l'algorithme de Goldberg en 30 minutes sur un monoprocesseur de 2.66 GHz

TransNetSim est une plateforme de simulation multi-agent du trafic que nous avons conçue conformément au modèle multi-échelle de l'EUP. Comme le montre la figure 7, dans TransNetSim l'EUP représente la base de l'environnement sur lequel différents simulateurs peuvent se brancher. Chaque simulateur (ou plutôt moteur de simulation) possède son propre modèle comportemental défini par rapport à une échelle particulière. Dans notre cas, comme l'EUP de Québec est construit à partir d'un méso-EUV et d'un micro-EUV, on retrouve un méso-simulateur et un micro-simulateur. Nous avons implémenté nous-mêmes le méso-simulateur car des défis de rapidité d'exécution ont été relevés, vu la taille du réseau considéré et le nombre des agents à simuler. Plus de détails concernant les performances du méso-simulateur ont été donnés dans (Chaker *et al.*, 2008). Pour le micro-simulateur, nous avons pu intégrer dans notre plateforme le logiciel commercial Vissim grâce à son interface programmable et surtout grâce à la comptabilité de son modèle de l'environnement avec le nôtre.

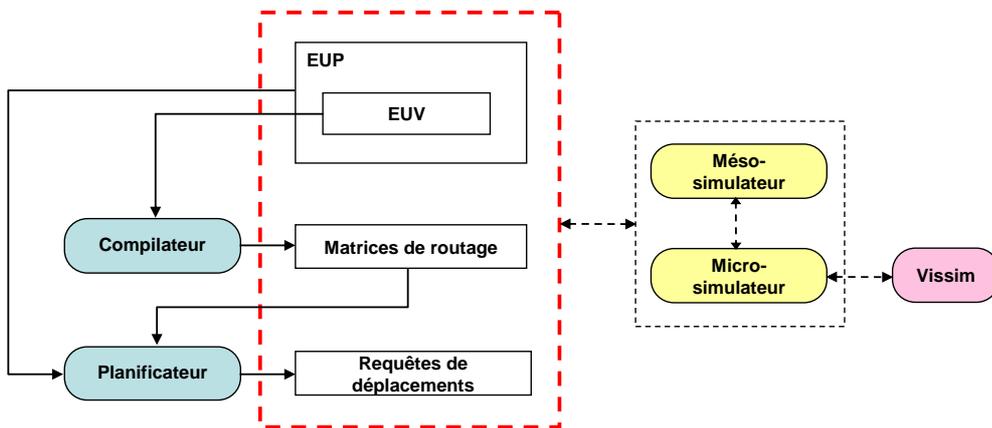


Figure 7: Architecture de la plateforme TransNetSim

À part les simulateurs, la figure 7 montre d'autres processus qui opèrent sur l'environnement de simulation. Ce sont les processus de prétraitement. On retrouve le compilateur qui a pour objet de compiler tous les chemins optimaux possibles dans une structure persistante, la matrice de routage. Ainsi, aucun algorithme de recherche d'itinéraire n'a besoin d'être exécuté durant la simulation. La matrice de routage permet de retourner en un temps linéaire le chemin optimal entre un nœud origine et un nœud destination, ce qui a rendu possible la simulation de 1 million de parcours dans le méso-EUV (graphe de plus de 80 000 liens) en moins d'une minute. Le deuxième processus du prétraitement est le planificateur. Partant de l'EUP, plus précisément de la population synthétique, le planificateur génère les plans de déplacements de chaque personne. Pour l'instant un modèle simple a été pris en compte qui ne permet de planifier que l'activité d'aller au lieu d'occupation.

Le planificateur implémente un modèle décisionnel qui permet de savoir selon le jour de la semaine: 1) si la personne ira ou non à son lieu de travail ou d'étude ; 2) dans quelle tranche horaire elle partira et 3) quel moyen de transport elle utilisera. Le planificateur génère ainsi une liste de déplacements à simuler, chacun défini par :

- une origine et une destination qui sont des *dessertes* dans le modèle de l'EUV. Rappelons qu'une *desserte* peut être aussi bien un *transfert* d'un *lien* à un autre qu'un *accès* d'un *lieu* à partir d'un *lien*.

- soit un temps de départ soit la paire (durée de l'activité précédente, référence du dernier déplacement).

- un mode de transport. On considère un seul mode pour chaque déplacement. Cela ne veut pas dire que la multimodalité n'est pas permise car différents déplacements peuvent s'enchaîner pour former un *voyage*. Il est justement de la responsabilité du planificateur d'enchaîner les déplacements multimodaux.

La simulation des déplacements pour aller au lieu de travail ou d'étude ainsi que les déplacements du retour au domicile ont été effectuées sur l'EUP de Québec pour chaque jour ouvrable de la semaine. La micro-simulation qu'on peut exécuter en parallèle sur le campus de l'université Laval (figure 8) permet d'avoir une visualisation réaliste de l'état du trafic, ce qui nous a permis de nous rendre compte d'une légère surestimation de la demande et de calibrer en conséquence certaines variables. D'autres calibrations ont été apportées suite aux analyses SOLAP faites sur les données de l'EUP et les données résultantes de comptage du trafic. Ces analyses sont présentées dans la section suivante.

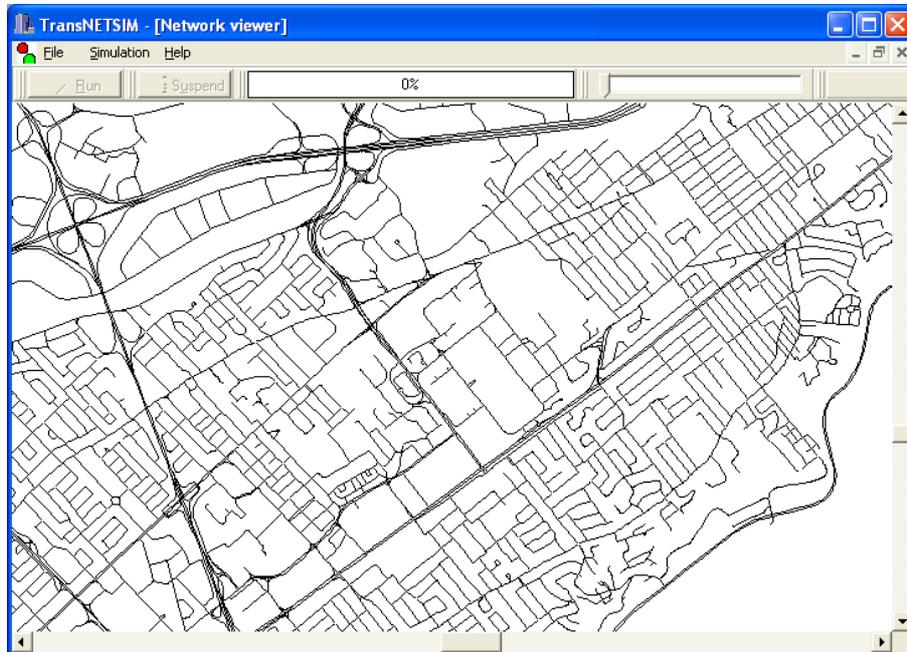


Figure 8.a: Le meso-simulateur de notre plateforme TransNetSim avec le réseau routier de toute la ville de Québec.

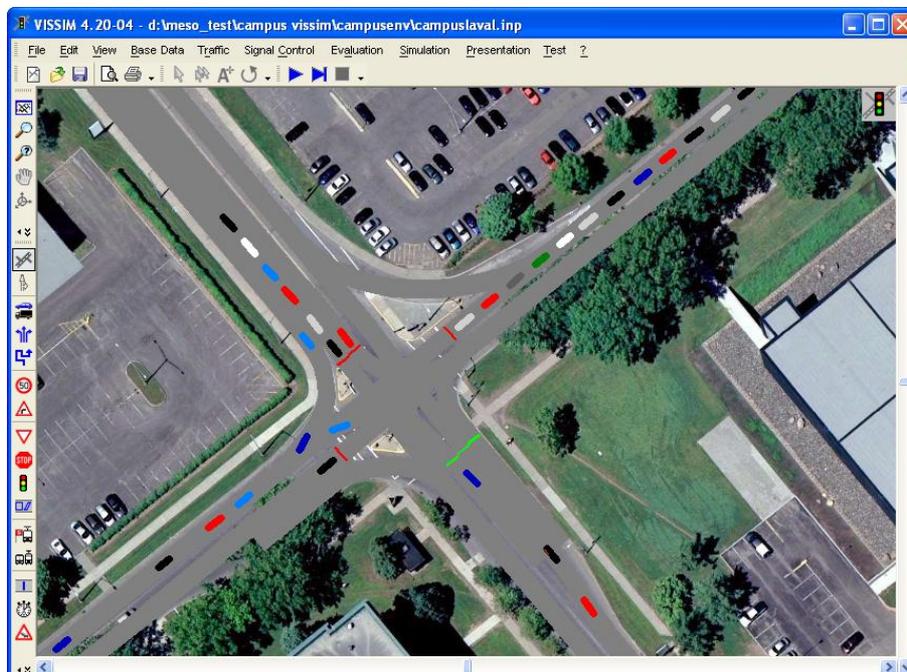


Figure 8.b: Interface du micro-simulateur Vissim piloté par TransNetSim pour la simulation détaillée du trafic sur le campus de l'université Laval.

4.3. Analyses et validations à l'aide du SOLAP

L'inspection et l'analyse des données entrantes et sortantes des simulations de la mobilité dans un EUP nécessitent l'usage d'outils. Des avancées récentes dans le domaine de l'analyse spatiale décisionnelle nous permettent d'utiliser l'approche SOLAP (*Spatial On-Line Analytical Processing*) pour effectuer ces analyses de façon simple et efficace.

4.3.1. L'analyse géospatiale décisionnelle

Afin d'optimiser l'analyse géospatiale, des systèmes spécifiques sont construits autour du paradigme multidimensionnel tel que défini dans le monde de l'informatique décisionnelle (i.e. cubes de données). Les structures basées sur ce paradigme ont démontré, depuis plus de 30 ans, qu'elles étaient plus faciles à comprendre par les utilisateurs que les structures transactionnelles et donc plus faciles à interroger. La structure multidimensionnelle considère les attributs comme une dimension, un peu à la manière des tableurs, et chaque donnée est déduite à partir de l'intersection de ces dimensions. On peut facilement visualiser deux dimensions dans une matrice, mais il est plus difficile de s'imaginer une structure à trois dimensions (cube) ou plus (hypercube). Cependant, c'est lorsque la structure supporte plusieurs dimensions que l'on peut démontrer la puissance de telles analyses. L'approche multidimensionnelle introduit de nouveaux concepts jusqu'alors inconnus du monde transactionnel comme les «dimensions», les «mesures», les «faits» et les «cubes» (Bédard *et al.*, 2007). Dans les cubes traditionnels, les *dimensions* sont des thèmes d'analyse (ex. période de temps, type de déplacement, motif) organisés selon une certaine hiérarchie (ex. année, mois, jour pour une dimension temporelle). Les membres du niveau inférieur (ex. les jours) s'agrègent ensemble pour composer les membres du niveau de détail supérieur dans la hiérarchie (ex. les mois). Une dimension peut avoir un caractère spatial comme la dimension territoire. On parle alors de dimension spatiale nominale puisque l'on utilise le nom du territoire pour définir les membres. Les *mesures* sont les valeurs numériques analysées pour les différentes dimensions (ex. le nombre de déplacements). Un ensemble de mesures agrégées selon un ensemble de dimensions forment un «cube». À l'intérieur d'un cube, les différentes combinaisons de dimensions et de mesures forment les «faits» (ex. le nombre de déplacements (mesure) pour le travail (dimension-type de déplacement) pour la région 1 (dimension-territoire) en matinée (dimension-temps) s'élèvent à 123 456 déplacements). En fait, la gestion spécifique de ces concepts dans les cubes permet ensuite à des outils spécialisés (ex. les outils OLAP) de présenter les analyses sous la forme de tableaux croisés et de diagrammes. Ces concepts doivent être ajustés au domaine spatial lorsque vient le temps d'introduire la gestion des géométries des objets. En fait, la gestion spécifique de la composante spatiale dans les cubes

permet aux outils spécialisés (ex. les outils Spatial OLAP) d'exploiter cette composante et de représenter l'analyse sous la forme d'une carte. Pour y parvenir, il est nécessaire d'associer une géométrie à chacun des membres de la dimension que l'on veut spatialiser (ex. région, route). On parle alors de dimension spatiale géométrique. Lorsque certains membres de la dimension spatiale n'ont pas de géométrie, il s'agit d'une dimension spatiale mixte. La mesure peut aussi être spatialisée. Elle permet alors de dériver une valeur numérique à partir d'un opérateur métrique appliqué sur une géométrie (ex. le calcul d'une superficie appliqué sur les territoires) (Rivest *et al.*, 2001) ou mieux encore le décompte d'éléments spatiaux inclus dans un autre (ex. le nombre d'arrêts d'autobus inclus dans une région). Une mesure spatiale peut aussi être géométrique. Elle peut regrouper simplement les pointeurs vers une ou plusieurs géométries (Rivest *et al.*, 2001). Elle peut aussi être la géométrie résultante d'un opérateur d'analyse spatiale, comme par exemple d'une intersection entre deux éléments géométriques ou encore le résultat d'un opérateur de corridor sur un élément (Rivest *et al.*, 2001; Marchand *et al.*, 2004).

La gestion spécifique des concepts d'origine et de destination de ce projet a mené à l'implantation du concept de mesure spatiale. Une problématique particulière doit être traitée lorsque le cube possède plus d'une dimension spatiale et que l'on souhaite gérer une nouvelle géométrie résultant de leur combinaison. C'est le cas dans notre application où une dimension spatiale tient de rôle d'origine (O) et la seconde de destination (D). Contrairement aux dimensions spatiales régulières où la géométrie est associée seulement aux membres des dimensions spatiales (ex. O1, O2, D1, D2), il était d'intérêt dans ce projet d'associer une géométrie à chaque couple OD. C'est le cas précisément, lorsque l'on souhaite représenter l'origine-destination sous la forme d'un flux (i.e. ligne) entre les deux géométries ponctuelles (O et D). La quantité est exprimée sur la géométrie du flux par son épaisseur. Dans ce cas précis, le système doit supporter l'ensemble des géométries possibles de chaque origine vers chaque destination, ce qui représente par exemple 100 flux potentiels pour un ensemble de 10 origines et 10 destinations. Les géométries de ces flux (ou l'algorithme utile à leur création dynamique) doivent être prédéfinies et stockées dans le cube. Une adaptation a été faite dans la technologie *JMAP Spatial OLAP* afin de permettre l'utilisation de cette mesure spatiale dans l'interface-client.

4.3.2. Analyse de la population synthétique

Un cube de données a été produit afin de présenter la population synthétique et sa visualisation a été possible dans l'application SOLAP. Ce cube contenait le dénombrement des ménages, des personnes, des travailleurs et des véhicules composant les ménages. Les ménages étaient présentés par types (type de ménage) localisés sur une grille hexagonale. Cette grille hexagonale divisait uniformément le territoire en parties égales de 250m carrés. Il était possible à l'aide de l'application SOLAP de visualiser facilement, par exemple, si la distribution des

ménages par type (avec ou sans enfant) était représentative de la population réelle de la région de Québec (figure 9).

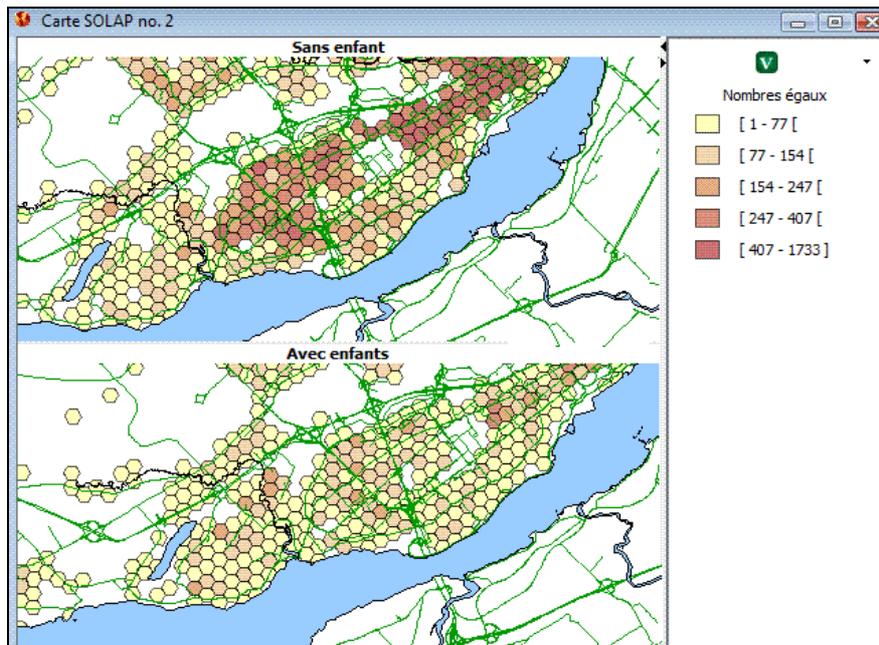


Figure 9. La distribution des ménages par type (avec ou sans enfant) sur les cellules hexagonales était représentative de la population réelle.

D'un autre côté, l'application nous a permis de mettre en évidence un problème relatif à la validité de l'échantillon de la population utilisée en considération de certaines caractéristiques, soit le nombre de véhicules et le nombre de travailleurs par ménage. En fait, l'analyse SOLAP a permis de souligner que l'échantillon de population utilisé a propagé uniformément le nombre de véhicules par ménage par rapport au nombre de travailleurs dans le ménage sur l'ensemble de la population sans tenir compte de facteurs économiques liés à la région (figure 10). Il apparaît que dans certains secteurs le nombre de véhicules par ménage était surestimé par rapport au nombre de travailleurs. Toutefois, cette particularité n'a pas engendré de problème particulier dans la simulation, car les déplacements étaient produits à partir du nombre de personnes possédant un permis de conduire. Cependant, la visualisation des données de la population synthétique dans le SOLAP a permis de soulever cette particularité dans les données de l'enquête.

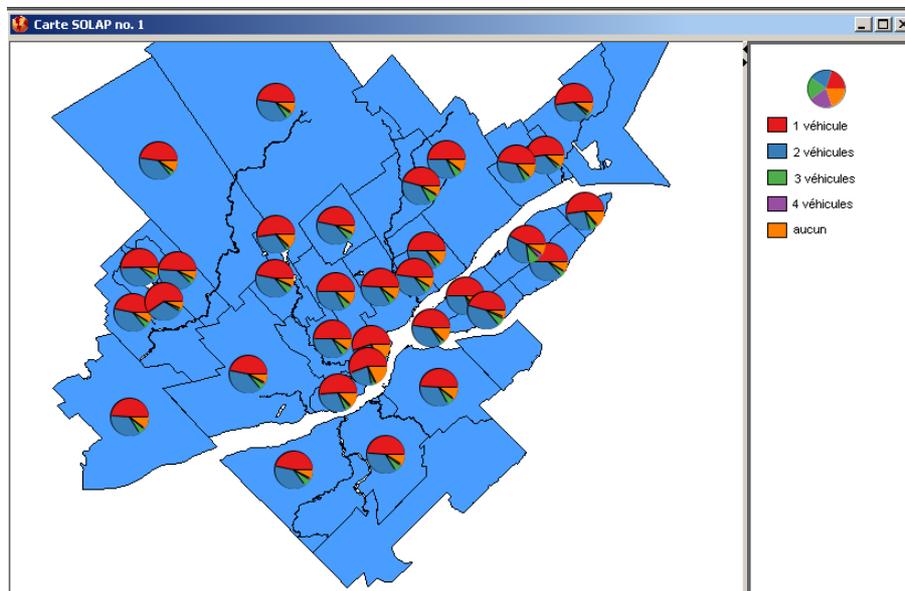


Figure 10. La distribution uniforme du nombre de véhicules par ménage par rapport au nombre de travailleurs n'était pas fidèle à la réalité. Les camemberts superposés présentent le nombre de travailleurs (valeurs) par nombre de véhicules.

4.3.3. Analyse des résultats de simulation

Un aspect très intéressant de l'utilisation de l'outil SOLAP est la possibilité de visualiser les résultats de la simulation sur différentes cartes. Pour y parvenir, deux cubes ont été produits. Le premier a permis de présenter la demande en transports selon l'origine et la destination selon la période de départ, le motif de déplacement et le mode de transport. Celui-ci a permis de faire ressortir facilement les principaux secteurs de destinations des travailleurs sur le territoire d'étude par une demande en transport très forte pour certains points de destination. La distribution spatiale des points de destination a permis d'identifier des zones génératrices de déplacements qui ont été ensuite validées sur le territoire par l'existence de pôles économiques et touristiques importants (tels que la Colline parlementaire et le Vieux-Québec) (figure 11). D'une façon intéressante, l'application nous a permis d'identifier une nouvelle zone de destination au nord-ouest soit celle correspondant au nouveau développement technologique et académique Saint-Roch (i.e. École des arts visuelles, I.N.R.S., Ubisoft). Ces zones ont ensuite pu être ajoutées à l'application comme information supplémentaire afin de définir l'origine des travailleurs vers ces zones spécifiques. Nous avons pu observer que les travailleurs de la zone étudiée proviennent en majorité des municipalités centrales du territoire.

Le second cube permet aussi d'illustrer les déplacements selon l'origine et la destination, la période de départ, le motif de déplacement et le mode de transport mais cette fois-ci sur le réseau routier de la grande région de Québec. La

visualisation des déplacements par période de départ a permis de souligner l'étranglement du réseau routier qui se localise au niveau des ponts traversant le Fleuve Saint-Laurent ainsi que l'achalandage important sur les grandes autoroutes en direction est-ouest. Une comparaison avec le comptage réel des véhicules sur le territoire a permis de constater une sous-estimation du trafic sur les artères principales du fait que nous ne simulons ni les déplacements interurbains ni le transport en commun. Cependant, la comparaison des comptages réels et simulés a permis de constater que les proportions étaient bien respectées (figure 13).



Figure 11. La distribution spatiale de points de destination où le nombre de déplacements est élevé a permis de définir de grandes zones génératrices de déplacements dans la Ville de Québec qui correspondent à des pôles économiques importants.

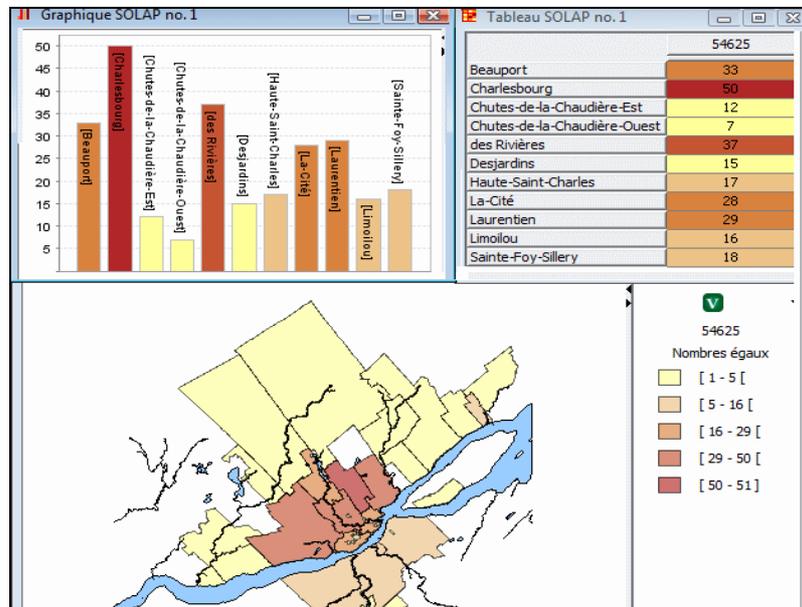


Figure 12. Carte choroplèthe et diagrammes illustrant les municipalités d'origine des travailleurs du secteur de destination étudié.

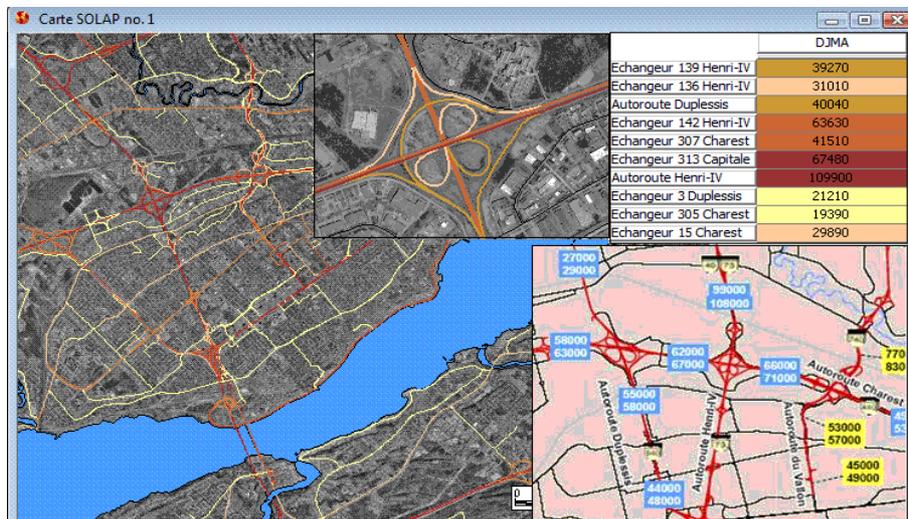


Figure 13. Résultat de la simulation des déplacements sur le réseau routier avec en médaillon (en haut, le comptage des véhicules dans un secteur dans la simulation) et (en bas le comptage réel)

5. Conclusion

Dans cet article nous avons présenté une nouvelle approche de conception d'un *Environnement Urbain Virtuel* (EUV) permettant une modélisation multi-échelle grâce à l'utilisation d'un patron de conception qui met en relation les lieux d'activité et les différents réseaux de transport que peuvent utiliser les personnes. Dans ce patron, une notion simple de desserte nous permet de modéliser à la fois l'accès à des lieux et les transferts entre différents réseaux, ce qui permet de simuler la multi-modalité. Ce patron peut être spécialisé dans diverses structures à diverses échelles (macro, méso et micro). Nous avons aussi formalisé l'EUV en utilisant la théorie des graphes, ce qui permet une implantation rigoureuse de ces concepts, tout en tenant compte des principales fonctions classiquement utilisées pour la simulation des parcours dans les réseaux de transports. De plus, nous avons proposé une méthode pour enrichir l'EUV avec les données de la population qui l'occupe afin d'aboutir à l'*Environnement Urbain Peuplé* (EUP). Cela nous a permis de créer systématiquement les individus et les ménages d'une population synthétique et de les associer aux lieux qu'ils occupent et utilisent dans l'environnement urbain. Ainsi, notre approche permet de modéliser à la fois la demande et l'offre en transports. Cependant, le succès d'une telle approche repose sur une démarche rigoureuse permettant de bien estimer ces deux aspects. Selon le principe « diviser pour régner », nous avons proposé une démarche qui réduit le problème en deux sous-problèmes similaires mais qui se traitent différemment : 1) l'allocation des attributs non spatiaux en combinant une approche de reconstruction synthétique avec un algorithme d'apprentissage non supervisé et 2) l'allocation des attributs spatiaux en combinant une approche de prédiction des temps de parcours avec un algorithme d'appariement dans un graphe biparti et pondéré. Au niveau de la validation, en plus des méthodes statistiques qui existent pour mesurer la vraisemblance d'une population synthétique, plusieurs validations unitaires peuvent être faites au moment même de la création de l'EUP. La plupart de ces validations et mesures de vraisemblance viennent naturellement avec les algorithmes utilisés (prédiction, appariement, etc.) et c'est tout l'avantage de cette réutilisation logicielle.

Nous avons aussi présenté le méso-simulateur de déplacements *TransNetSim* qui tire partie de tous les avantages de notre EUP. Son utilisation a été illustrée par la simulation du trafic de véhicules dans la ville de Québec. Notre approche a aussi permis de développer une micro simulation du trafic pour une portion de ce grand territoire. Cette micro-simulation a été réalisée avec le micro-simulateur commercial *Vissim* et intégrée sans problème dans *TransNetSim*. Finalement, nous avons montré comment l'utilisation d'outils SOLAP permet d'une façon simple et efficace d'analyser et de calibrer les résultats de nos simulations (génération de la population synthétique, données résultant de la simulation du trafic). Cette facilité d'inspection des données en entrée et en sortie des simulations, démontre l'intérêt de coupler des outils SOLAP avec des outils de simulation en transport pour accélérer la calibration et l'analyse des résultats obtenus.

Remerciements

Ce projet a été financé par GEOIDE (www.geoide.ulaval.ca), le réseau de centres d'excellence en géomatique du Canada (Projet MUSCAMAGS).

Bibliographie

- Albeverio, S., Andrey, D., Giordano, P., Vancheri, A., *The dynamics of complex urban systems: An interdisciplinary approach*, Berlin, Springer, 2008.
- Arentze, T., Timmermans, H., Hofman, F., « Creating Synthetic Household Populations: Problems and Approach », *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2014, 2007, p. 85-91.
- Ayed, H., Khadraoui, D., Habbas, Z., Bouvry, P., Merche, J.P., « Transfer Graph Approach for Multimodal Transport Problems », *Proceedings of the Second International Conference MCO 2008*, Metz, France - Luxembourg, 8-10 september, 2008.
- Barrett, C. L., Beckman, R. J., *TRANSIMS, Vol. 3*, Los Alamos Unclassified Reports 00-1725 and 00-1767. Los Alamos National Laboratory, Los Alamos, N. Mex., 2004.
- Batty M., Longley P., *Fractal Cities*, Londres, Academic Press, 1994.
- Beckman, R.J., Baggerly K.A., McKay, M.D. « Creating Synthetic Baseline Populations », *Transportation Research-A*, vol. 30, n° 6, 1996, p. 415-429.
- Bédard, Y., Rivest, S., Proulx, M.-J., « Spatial On-Line Analytical Processing (SOLAP): Concepts, Architectures and Solutions from a Geomatics Engineering Perspective », *Data Warehouses and OLAP*, Londres, IRM Press, 2007, p. 298-319.
- Benenson I., Torrens, P.M., *Geosimulation: Automata-Based Modeling of Urban Phenomena*, Londres, John Wiley & So, 2004.
- Chaker, W., Moulin, B., « Forecasting Travel Supply and Demand by Modeling Multiscale Urban Environments », *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2064, 2008, p. 65-72.
- Chaker, W., Moulin, B., Thériault, M. « Modeling Transport Networks with Design Pattern: Application to Hybrid Traffic Simulations », *19th IASTED International Conference on Modelling and Simulation*, Montreal, Quebec, Canada, 2007, ACTA Press, p. 568-573.
- Chaker, W., Moulin, B., Thériault, M., « Multiscale Modeling of Virtual Urban Environments and Associated Populations », *Jiang B. and Yao X. (éditeurs), Geospatial analysis and modeling of urban structure and dynamics*, Berlin, Springer, (à paraître en 2009).
- Donikian, S., « VUEMS: a virtual urban environment modeling system ». *Proceedings of Computer Graphics International conference*, 1997. p.84-92.
- Gamma, E., Helm, R., Johnson, R., Vlissides, J., *Design patterns: elements of reusable object-oriented software*, Addison-Wesley Longman Publishing Co., 1995

- Goldberg A V., « An Efficient Implementation of a Scaling Minimum-Cost Flow Algorithm », *Journal of Algorithms*, vol. 22, n° 1, janvier 1997, p. 1-29
- Gong, J., Hui, L., « Virtual Geographical Environments: Concept, Design, and Applications » *Proceedings of the International Symposium on Digital Earth (ISDE)*, 1999, Beijing, Chine, p. 369-375.
- Guo, J.Y., Bhat, C.R., « Population Synthesis for Microsimulating Travel Behavior », *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2014, 2007, p. 92-101.
- Ishida, T., Gasser, L., Nakashima, H., « Massively Multi-Agent Systems I », *Lecture Notes in Artificial Intelligence*, Vol. 3446, 2005, Springer.
- Jiang, B., « SimPed: simulating pedestrian flows in a virtual urban environment », *Journal of Geographic Information and Decision Analysis*, 1999, vol. 3, n° 1, p. 21-30
- Lozano, A., Storchi, G., « Shortest viable hyperpath in multimodal networks », *Transportation Research Part B: Methodological*, vol. 36, n° 10, 2002, p. 853-874
- Mackaness, W.A., Ruas, A., Sarjakoski, L.T., *Generalization of geographic information cartographic modeling and applications*, Holland, Elsevier, 2007.
- Malaikrisanachalee, S. and T. M. Adams. "Lane-based Model for Transportation Network Flow Analysis and Inventory Management." TRR Journal of the Transportation Research Board. No. 1935, 2005, p. 101-110
- Marchand, P., Brisebois, A., Bédard, Y., Edwards G., « Implementation and evaluation of a hypercubebased method for spatio-temporal exploration and analysis », *Journal of the International Society of Photogrammetry and Remote Sensing*, vol. 59, Elsevier, 2004, p. 6-20
- Mathis, P., *Graphes et réseaux. Modélisation multiniveau*, Hermes, 2003.
- Moulin B., Chaker W., Perron J., Pelletier P., « MAGS Project: Multi-agent Geosimulation and Crowd Simulation », *conférence COSIT 2003*, 24-28 September 2003, Springer, Vol. 2825, 2003, p. 151-168.
- Rivest, S., Bédard, Y., Marchand P., « Toward Better Support for Spatial Decision Making: Defining the Characteristics of Spatial On-Line Analytical Processing (SOLAP) », *Geomatica*, vol. 55, n°4, 2001, p. 539-555
- Timmermans, H.J.P., Zhang, J., « A need-based model of multi-day, multi-person activity generation. » *Transportation Research. Part B*. vol. 43. no. 2. 251-265, 2009
- Vanajakshi, L., Rilett, L.R., « Support vector machine technique for the short term prediction of travel time », *IEEE Intelligent Vehicles Symposium*, Proceedings 2007, p. 600-605
- Waddell, P. « UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning » *Journal of the American Planning Association*, Vol. 68, No. 3, 2002, p. 297-314

Williamson P., Birkin M., Rees P. « The Estimation of Population Microdata by Using Data From Small Area Statistics and Samples of Anonymised records », *Environment and Planning A*, 1998, vol. 30, p. 785-816.

Wu, C.-H., Ho, J.-M., Lee, D.T., « Travel-time prediction with support vector regression », *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, n° 4, 2004, p. 276-281