TABLE DES MATIÈRES

INTRODUCTION 1

CHAPITRE 1 REVUE DE LITTÉRATURE 5
1.1 La visualisation de données spatiales et spatio-temporelles 5
1.1.1 Les cubes spatio-temporels 5
1.1.2 Trip Vista 8
1.1.3 Technique « Focus + Context » appliquée à des réseaux routiers 9
1.1.4 Simplification des données 10
1.1.5 Les techniques de visualisation radiales 14
1.1.6 Résumé 16
1.2 La visualisation de graphes 16
1.3 Techniques de visualisations hybrides 18
1.4 Techniques de visualisations de liens et arcs 20
1.4.1 Techniques de visualisations d'arcs 20
1.4.2 Agrégation de données 22
1.4.3 Techniques de visualisation radiales 23

CHAPITRE 2 PROBLÉMATIQUE 27

CHAPITRE 3 PRÉSENTATION DE LA SOLUTION 31
3.1 Présentation générale 31
3.2 La visualisation géospatiale 32
3.3 La visualisation en mode topologique 33
3.4 La visualisation hybride 35
3.5 Résumé 37

CHAPITRE 4 CHOIX DE CONCEPTION 39
4.1 L'agencement 39
4.2 Type de modèle topologique 40
4.3 Les liens 44
4.3.1 Les liens de type « liens droits » 45
4.3.2 Les liens de type « lignes brisées » 45
4.3.3 Les liens de type « Courbe de Bézier » 46
4.3.4 Comparatif des différents types de liens 50
4.4 Les arrêtes d'adjacences 51
4.5 Les trajectoires 52
4.5.1 Dessin des trajectoires en mode géospatial 52
4.5.2 Dessin des trajectoires en mode topologique 53
4.5.3 Exemple concret 58

CHAPITRE 5 GÉNÉRATION DES DONNÉES 59
5.1 Génération aléatoire 59
5.2 Génération semi aléatoire: Principe de graphe à deux niveaux 62

5.3 Bruitage des données 64
5.4 Application de la méthode pour générer des données urbaines 66

CHAPITRE 6 EXPÉRIENCE CONTRÔLÉE 73
6.1 Protocole expérimental 73
6.2 Réalisation de l'expérience 75
6.3 Résultats 77

CHAPITRE 7 DISCUSSIONS 83
7.1 Application à des données géographiques 83
7.2 Agrégation des données 87
7.3 Visualisation hybride 87

CONCLUSION 89

BIBLIOGRAPHIE 91

LISTE DES TABLEAUX

Page

Tableau 1 Comparatif des caractéristiques des différentes techniques de visualisation38

Tableau 2 Comparatif du nombre de liens pour chaque visualisation.....44

Tableau 3 Comparatif des différents types de liens selon trois critères (Présence de croisements, visibilité des liens multiples, occlusion).....50

Tableau 4 Résultats de l'analyse de variance ANOVA sur le temps 78

Tableau 5 Résultats de l'analyse de variance ANOVA sur le taux d'erreur.....78

Tableau 6 Taux d'erreur moyen par type d'interface et par type de question 79

LISTE DES FIGURES

Page

Figure 0.1 Technique de visualisation géospatiale 2

Figure 0.2 Visualisation topologique avec l'affichage de trois trajectoires 3

Figure 1.1 Capture d'écran tirée du logiciel GeoTime 6

Figure 1.2 Capture d'écran du logiciel TripVista 8

Figure 1.3 À gauche, la représentation géospatiale des données, à droite, l'affichage des distances par rapport à un point fixe 11

Figure 1.4 Première représentation topologique du métro de Londres 12

Figure 1.5 Technique "Focus + Context" appliqué au plan du métro de Sidney 13

Figure 1.6 Capture d'écran de "Visits", montrant des données spatio-temporelles sous forme de "small-multiples" 14

Figure 1.7 Vue "Trip View" permettant de visualiser des trajectoires, la période du jour à laquelle elles apparaissent, ainsi que leur vitesse 15

Figure 1.8 Exemple explicatif de l'application de Circos 16

Figure 1.9 Capture d'écran montrant la technique *TreeMatrix* 19

Figure 1.10 Intégration des 4 techniques interactives pour contrôler les courbes dans un réseau biologique : (a) interactive bundling, (b) interactive link fanning, (c) link magnets, (d) interactive legends 20

Figure 1.11 Construction de l'agencement en palme de ventilateur, exemple d'interaction multitouche 21

Figure 1.12 (a) La technique *Hierarchical Edge Bundles* appliquée à un arbre avec un agencement de type *Baloon Layout*, (b) les mêmes données avec une représentation de type radiale 22

Figure 1.13 Technique de *Force-Directed Edge Bundling* appliquée à un graphe nœuds-liens représentant des flux migratoires aux États-Unis 23

Figure 1.14 (a) L'approche BubbleSet, (b) l'approche LineSets, (c) l'approche diagramme de Kelp 24

Figure 1.15 Diagramme de Kelp 25

Figure 1.16 Comparaison de 5 génomes (A, B, C, D, E), composés de 23 chromosomes avec *Circos* 25

Figure 2.1 Données GPS correspondant aux déplacements sur une période d'un mois (Visualisées à l'aide de Google Latitude) 27

Figure 2.2 Détails de la Figure 2.1 29

Figure 2.3 Exemple de technique de visualisation topologique.....30

Figure 2.4 Croquis initial montrant les différentes techniques de visualisation et un hybride entre les deux techniques.....30

Figure 3.1 Contrôleur de temps.....31

Figure 3.2 Technique de visualisation géospatiale. Ici, les trois losanges rouge vert et bleu montrent la position actuelle de trois personnes. 33

Figure 3.3 Vue topologique avec relations de voisinage et d'adjacence. 34

Figure 3.4 Visualisation topologique avec l'affichage de trois trajectoires.....35

Figure 3.5 Visualisation hybride 36

Figure 3.6 Visualisation hybride après réagencement manuel des pièces 37

Figure 4.1 Différentes manières de modéliser le plan. *a*: Le plan complet. *b*: Graphe nœuds-liens, avec les salles qui sont représentées par des nœuds, et les portes par des liens. *c*: Le graphe dual géométrique associé au graphe nœuds-liens. *d :* Le *line-graphe* associé au graphe nœuds-liens. *e :* Version simplifiée du *line-graphe*.....41

Figure 4.2 Un graphe (en noir) et son line graph associé (en rouge) 43

Figure 4.3 Mode topologique avec les liens droits.....45

Figure 4.4 Mode topologique avec des lignes brisées.....46

Figure 4.5 Mode topologique avec des courbes de Bézier.....47

Figure 4.6 Variation du facteur et effet sur les courbes de Bézier 48

Figure 4.7 Courbe de Bézier relationnelle avec un facteur de 10 49

Figure 4.8 Courbe de Bézier relationnelle avec un facteur de 0.1 49

Figure 4.9 Vue topologique avec arrêtes d'adjacences (montrées par des pointillés) .51

Figure 4.10 Affichage des trajectoires en mode géospatiale 53

Figure 4.11 Les trajectoires avec la technique de visualisation topologique 54

Figure 4.12 Manière alternative de dessiner les trajectoires autour d'un nœud 55

Figure 4.13 Affichage des trajectoires avec le mode de visualisation hybride 56

Figure 4.14 Exemple de représentation d'une trajectoire en mode topologique.....58

Figure 5.1 Schéma explicatif du *steering behavior*.....60

Figure 5.2 Quatre trajectoires générées totalement aléatoirement avec un comportement de type *wandering behaviour* 62

Figure 5.3 Superposition du plan avec le graphe interne 63

Figure 5.4 Exemple de trajectoires générées par la méthode expliquée ci-dessus, en superposition du graphe interne 65

Figure 5.5 Graphe nœuds-liens généré à partir d'un plan de Dublin.....67

Figure 5.6 Détails d'une partie du plan de Dublin.....68

Figure 5.7 Visualisation des données générées dans le prototype "Spatio-Temporal Visualizer". 70

Figure 6.1 Interface utilisée pour l'expérience, avec la vue géospatiale 76

Figure 6.2 Interface utilisée pour l'expérience, avec la vue topologique 77

Figure 6.3 Temps écoulé moyen pour chaque type de question et chaque interface en millisecondes.....79

Figure 6.4 Décomposition du temps écoulé médian en millisecondes pour chaque type de question 81

Figure 7.1 Représentation géographique de la carte et de la trajectoire.....84

Figure 7.2 Représentation topologique de la carte et de la trajectoire 85

Figure 7.3 Représentation hybride de la carte et de la trajectoire 86

Rapport-Gratuit.com

INTRODUCTION

L'avènement des téléphones intelligents, pour la plupart équipés de GPS, a rendu la géolocalisation accessible au plus grand nombre. Un système de géolocalisation permet de fournir à intervalle régulier des données spatio-temporelles (i.e. des données complexes ayant une composante spatiale, et une composante temporelle).

Il est possible d'observer qu'au sein d'un bâtiment ou en extérieur, les cibles suivies ont tendance à emprunter les mêmes chemins. Pour des données extérieures, les trajets suivent les routes ou les réseaux de transport (voir Figure 2.1, montrant des données GPS correspondant aux déplacements sur une période d'un mois où les trajets qui suivent les routes, et les liaisons entre stations de métro, sont reconnaissables). En ce qui concerne les données en intérieur, les chemins sont façonnés par les meubles et objets présents dans les pièces. Par exemple, dans un environnement de travail, une personne se déplacera entre son ou ses postes de travail, la salle de pause, la salle de réunion, la salle de bain, etc., en empruntant la plupart du temps le chemin le plus direct.

Cette particularité des données spatio-temporelles cause un important effet de superposition et donc d'occlusion dès lors que l'on désire afficher un certain nombre de trajets. Le problème soulevé est donc de trouver une nouvelle technique de visualisation qui permettrait de visualiser les données spatio-temporelles, en diminuant les effets d'occlusion.

L'une des solutions possibles est d'agréger les données afin de réduire les informations à afficher. La question préalable qui a été posée est : « Quelles données sont importantes dans le cadre de la visualisation de données spatio-temporelles au sein d'un bâtiment ?». Deux types de données ont semblé importants :

- Les déplacements à l'intérieur d'une pièce (intra)
- Les déplacements entre deux pièces (inter)

Les données de géolocalisation sont des données complexes, ayant plusieurs composantes. Aussi, les utilisateurs de logiciels permettant de visualiser ces données ont divers types de

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

tâches à réaliser, pouvant s'intéresser à la composante spatiale, la composante temporelle, la composante « cible », et ce, que ce soit de manière simple (par exemple, en s'intéressant à un lieux précis), ou de manière plus globale (on s'intéresse à une zone).

De par la diversité des tâches pouvant être réalisées, et de par la complexité des données, il est difficile de développer une interface idéale dans tous les cas. Ainsi, une interface de visualisation hybride pourrait permettre de tirer parti des avantages de deux types de visualisations, en fonction des tâches et des données.

Les déplacements au sein d'une pièce, les déplacements intra, peuvent être visualisés sous leur forme physique, sous la forme d'un plan, qui sera appelée technique de visualisation géospatiale (Figure 0.1).

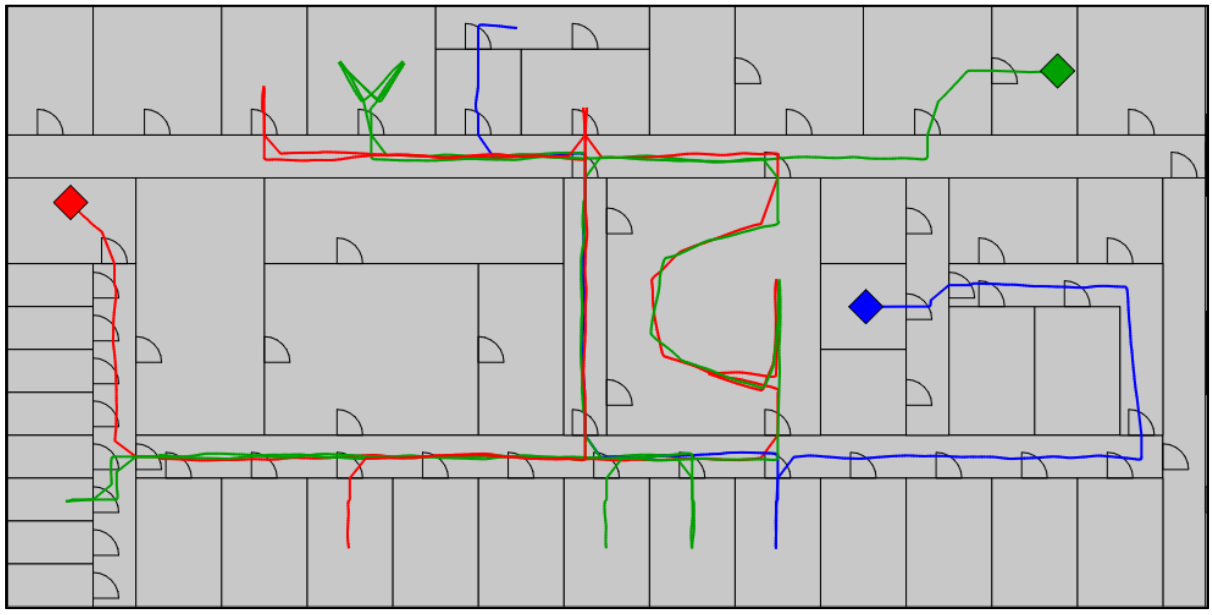

Figure 0.1 Technique de visualisation géospatiale

Afin de visualiser les déplacements entre deux pièces, le bâtiment sera visualisé sous la forme d'un graphe nœuds-liens – qui sera appelée technique de visualisation topologique – où les nœuds correspondront aux pièces, et les portes aux liens. Ainsi, seules les transitions d'une pièce à une autre, ainsi que la présence ou non dans une pièce seront visualisées

(Figure 0.2). Nous supposons que cette visualisation topologique, montrant le bâtiment sous forme d'un graphe nœuds-liens, permettra de mieux appréhender les déplacements entre les pièces.

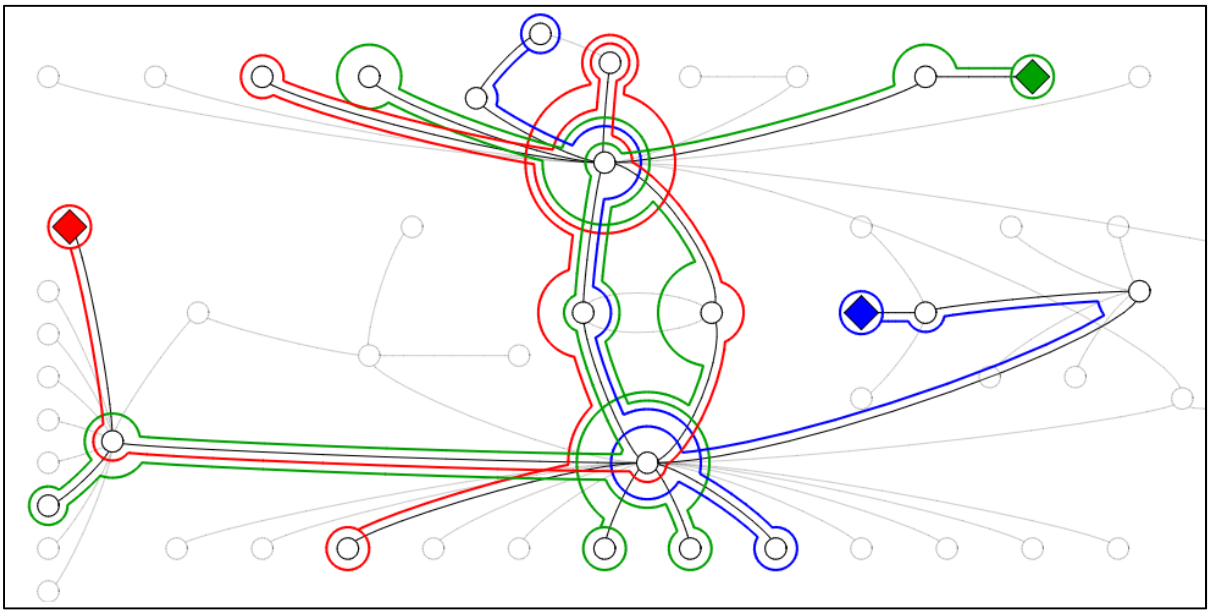

Figure 0.2 Visualisation topologique avec l'affichage de trois trajectoires

De plus, nous allons présenter une visualisation hybride permettant de tirer parti des avantages de la visualisation topologique ainsi que de ceux de la visualisation géospatiale.

Les contributions apportées par ce mémoire sont:

- La technique de visualisation topologique, ainsi que le décalage des trajets le long des arêtes et l'utilisation d'arcs circulaires autour des nœuds, pour désambiguïser les multiples trajets dans les données;
- Une discussion sur les différents modèles topologiques possibles, ainsi que différentes manières de dessiner les arêtes et les trajets, dans la visualisation topologique;
- Une technique de visualisation hybride combinant la visualisation topologique avec la visualisation géospatiale;

- Une évaluation expérimentale de l'efficacité de la technique de visualisation topologique.

CHAPITRE 1

REVUE DE LITTÉRATURE

Ce chapitre correspond à un survol de la littérature de plusieurs sujets abordés dans ce mémoire.

Dans l'introduction, deux techniques de visualisation sont mentionnées : la technique de visualisation géospatiale, ainsi que la technique de visualisation hybride. La première partie de cette revue de littérature sera consacrée à la visualisation de données spatiales et spatio-temporelles, qui servira pour le développement de la technique de visualisation géospatiale.

Dans un second temps, la revue de littérature se focalisera sur les techniques de visualisation de graphes nœuds-liens, qui seront utilisées pour la vue topologique.

Ensuite, la technique de visualisation proposée dans le mémoire étant une technique de visualisation hybride, une section sera consacrée à ce type de visualisation et aux vues coordonnées.

Pour finir, la dernière partie de cette revue de littérature sera consacrée aux différents moyens pouvant être utilisées pour représenter les trajectoires.

1.1 La visualisation de données spatiales et spatio-temporelles

1.1.1 Les cubes spatio-temporels

En 2005, Kapler et Wright (Kapler et Wright, 2005), proposent une implémentation des cubes spatio-temporels (*Space-Time Cube*). Dans ce type de visualisation les données issues de géolocalisation sont placées dans un cube (au lieu de les mettre sur un plan comme cela est fait traditionnellement), où les coordonnées x et y correspondent aux coordonnées GPS, et

où la composante z, la hauteur, est définie en fonction du temps. Le plan géographique est placé à la valeur $z = 0$, et correspond au temps courant dans le logiciel, alors que les valeurs de z positives représentent le future, et les valeurs négatives représentent le passé.

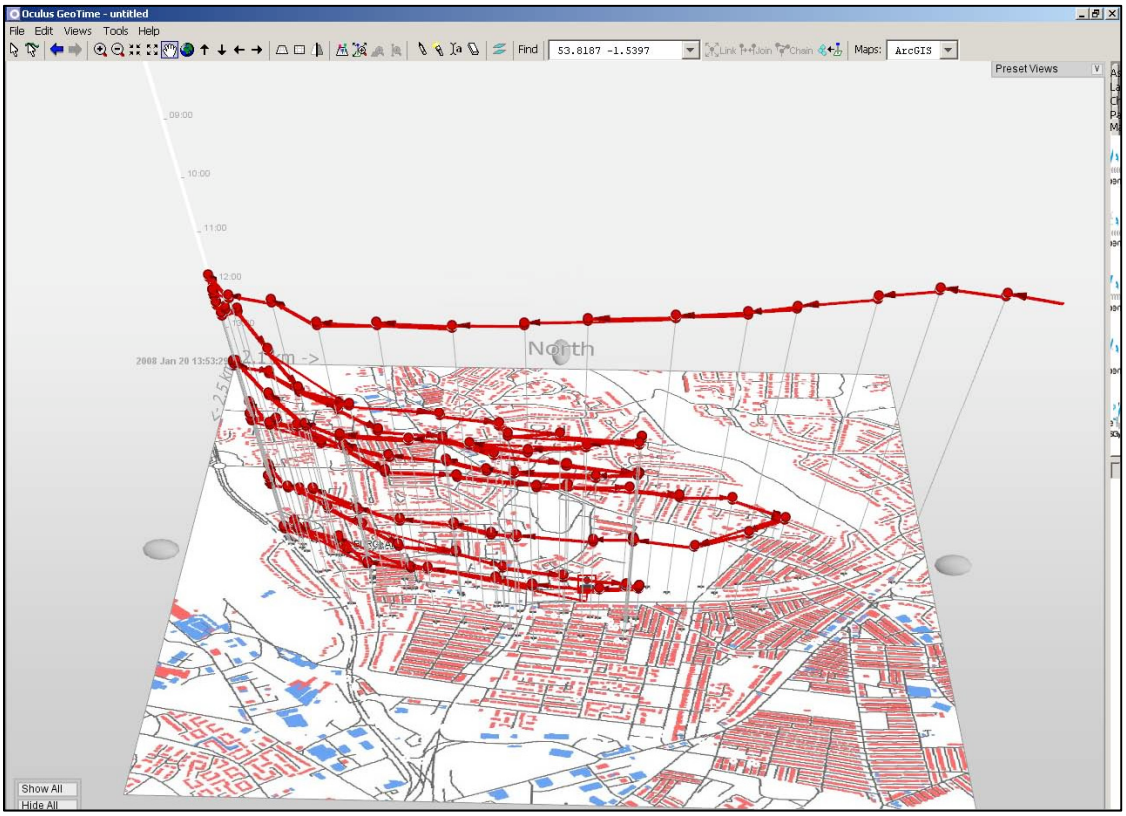

Figure 1.1 Capture d'écran tirée du logiciel GeoTime (Implémentation du Space Time Cube (Kapler et Wright, 2005))

Dans leur article de 2009, « *An Evaluation of Space Time Cube Representation of Spatiotemporal Patterns* » (Kristensson et al., 2009), Kristensson et al. présentent une manière d'évaluer l'efficacité du *Space-Time Cube*. Les questions fondamentales posées dans cet article sont les suivantes :

- Est-ce qu'un utilisateur débutant peut comprendre et utiliser le système après une courte période d'entrainement?

- Est-ce qu'il y'a des différences de performances en terme de temps de réponse entre l'interface *Space-Time Cube* et une interface de visualisation 2D?
- Est-ce qu'il y'a des différences de performances en terme de taux d'erreur entre l'interface *Space-Time Cube* et une interface de visualisation 2D, pour des catégories de question spécifiques?

Dans la pratique, les questions posées aux utilisateurs lors de l'expérience étaient classées selon les quatre catégories suivantes:

- Simple « Quand », et simple « Quoi et où » : Où est la personne rouge à 14h00 ?
- Simple « quand » et étendu «Quoi et où » : Est-ce qu'il y'a deux personnes au même endroit à 9h ?
- Étendu « quand » et simple « Quoi et où » : Quel bâtiment a été visité par la personne rouge durant la journée ?
- Étendu « quand » et étendu « quoi et ou » : Qui reste dans la zone du campus pendant la plus longue période ?

Un test statistique ANOVA (analyse de variance) a ensuite été appliqué pour valider la signification des résultats. Le résultat de l'expérience a montré qu'un utilisateur novice pouvait prendre en main l'interface *Space-Time Cube* avec une courte période d'entrainement, sans qu'il n'y'ait de différence significative sur l'ensemble des quatre catégories de question, que ce soit en terme de taux d'erreur ou en terme de temps de réponse. Néanmoins, pour la catégorie de question 4, le temps de réponse s'est trouvé être inférieur pour l'interface *Space-Time Cube.*

Dans une étude complémentaire (Kristensson et al., 2009), Kristensson et al. démontrent que les *Space-Time Cube* sont plus efficaces pour répondre aux questions concernant des patrons de déplacement complexes, tout en conseillant aux développeurs d'ajouter un mode de

visualisation classique (en deux dimensions) afin de permettre aux utilisateur de répondre aux questions simples plus aisément.

1.1.2 Trip Vista

TripVista, un logiciel de visualisation développé par Guo et al. en 2011, permet de visualiser les données relatives aux véhicules traversant un carrefour, et ce au travers de plusieurs vues coordonnées (Guo et al., 2011).

Les vues coordonnées sont un type de techniques de visualisation où plusieurs vues représentent des données en utilisant diverses techniques de visualisation (ou possiblement les mêmes techniques, avec des paramètres de caméra différents par exemple). Le modèle de données étant le même sur chacune des vues, si l'on applique un filtre ou une surbrillance sur l'une de celles-ci, le même effet est appliqué sur chacune des vues (Wang Baldonado, Woodruff et Kuchinsky, 2000) (Roberts, 2007).

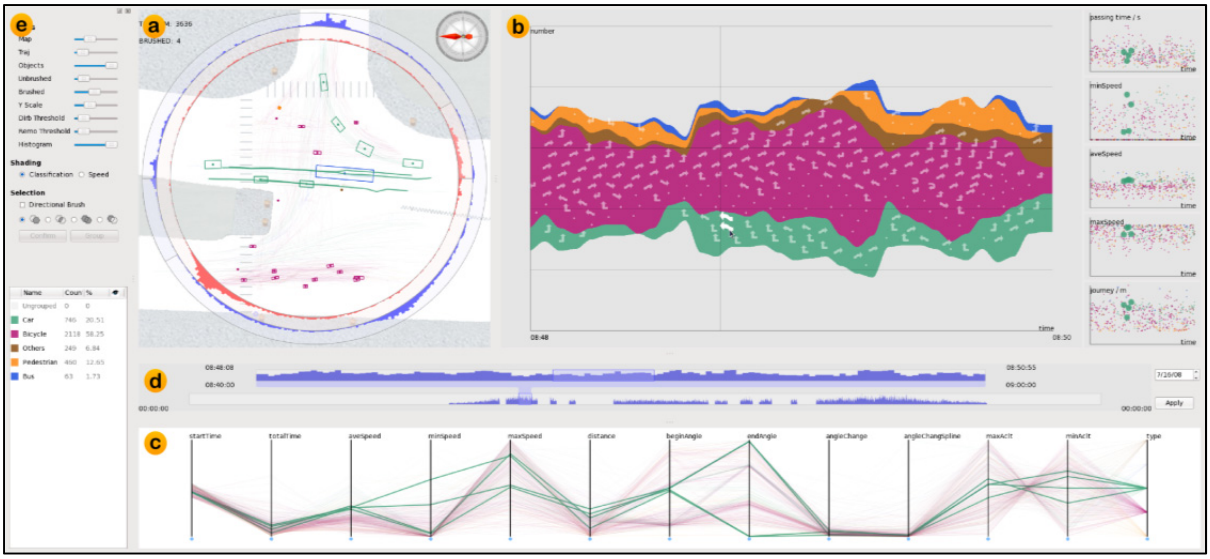

Figure 1.2 Capture d'écran du logiciel TripVista
Tirée de Guo et al. (2011)

La première vue correspond à un plan, la perspective spatiale, sur lequel sont dessinées les trajectoires. Les véhicules sont eux classés en cinq catégories (Bus, voitures, piétons, vélos, et autres) et sont représentés par des glyphes. Les deux anneaux entourant ce plan permettent de visualiser la densité d'entrées et de sorties à chacun des points, sous la forme d'histogramme. Le cercle rouge représente les entrées, et le cercle bleu les sorties.

La seconde vue, la perspective temporelle, permet de visualiser certaines données relativement au temps. Cette perspective est en fait composée de six vues. Cinq nuages de point, pour lesquelles la coordonnée x correspond au temps et la coordonnée y peut être la vitesse minimale, la vitesse maximale, la vitesse moyenne, le temps passé sur le carrefour ou la distance parcourue, et une vue de type *Theme River (Havre, Hetzler et Nowell, 1999)*. Dans cette dernière, la coordonnée x correspond toujours au temps, mais la coordonnée y correspond cette fois au nombre de véhicules de chaque type. Les glyphes correspondent à une schématisation des trajectoires empruntées par les véhicules.

Enfin la dernière vue, la perspective multidimensionnelle, permet de visualiser différents types de données, sous forme de coordonnées parallèles.

Toutes ces vues sont coordonnées, et permettent de mettre en surbrillance, ou de filtrer les données. Ainsi, les anneaux sur la perspective spatiale peuvent aussi servir de filtre pour, par exemple, ne visualiser que les trajectoires sortant en passant à travers certaines zones. Dans ce cas, les données sont aussi filtrées sur toutes les autres vues.

1.1.3 Technique « Focus + Context » appliquée à des réseaux routiers

Les techniques de visualisation Focus + Context sont des techniques de visualisations permettant de mettre en évidence une partie des données (Le « focus ») tout en remettant des données dans leur contexte général (Cockburn, Karlson et Bederson, 2009).

Rapport-Gratuit.com

Haunert et Sering propose une technique de visualisation *Focus + Context* permettant d'effectuer un zoom sur une partie d'un plan représentant un réseau routier (*Focus*), tout en continuant à visualiser le plan dans son ensemble (Haunert et Sering, 2011). En effet, la particularité principale de cette technique, par rapport aux techniques de *Focus + Context* classiques, comme les lentilles *Fish Eye* par exemple, est que l'ensemble des données reste ici visible (alors qu'avec une lentille *Fish Eye* certaines données peuvent sortir de l'espace visible). La seconde particularité est que les déformations sont limitées, l'algorithme va procéder à une réduction des distances entre les intersections dans le *context*, de manière à préserver au maximum les positions relatives de ces intersections les unes par rapport aux autres.

1.1.4 Simplification des données

Plusieurs techniques de visualisation se basent sur le principe de la simplification des données, que ce soit en réduisant leur nombre, en réduisant les dimensions de ces données, ou en effectuant des regroupements, en fonction des caractéristiques des données, de leur position etc. (*Clustering*).

Dans un article datant de 2009, Crnovrsanin et al. proposent d'effectuer une réduction de la dimension spatiale des données (Crnovrsanin et al., 2009). En considérant la distance entre les données et une position de référence (fixe ou mobile), on peut alors visualiser cette donnée sous la forme d'un graphique, avec la distance en ordonnées, et le temps en abscisses.

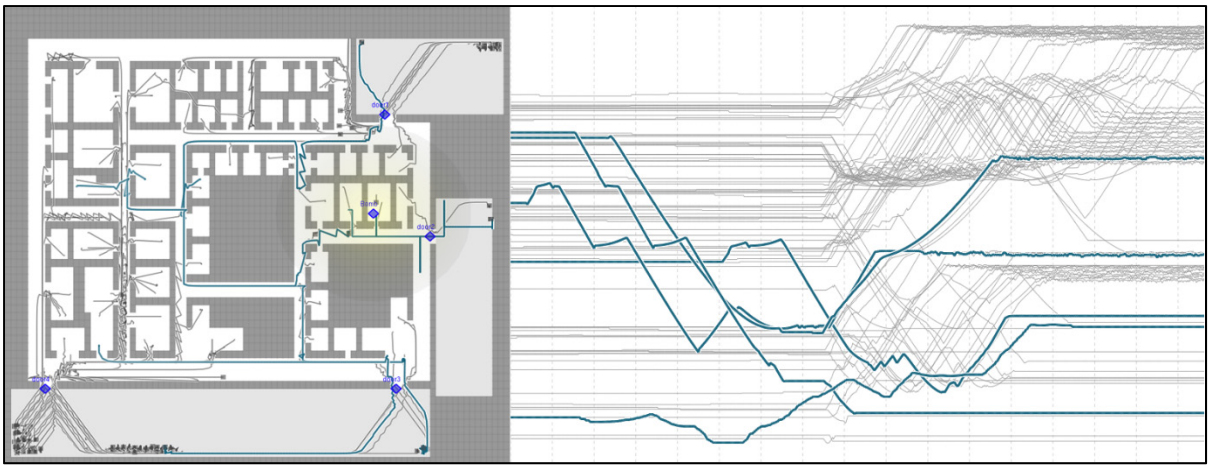

Figure 1.3 À gauche, la représentation géospatiale des données, à droite, l'affichage des distances par rapport à un point fixe
Tirée de Crnovrsanin et al. (2009)

Dans l'exemple ci-dessus (Figure 1.3), les données représentent les déplacements de personnes au sein d'un bâtiment dans le cas d'une alerte à la bombe. La position de référence est ici la position de la bombe, et on peut donc visualiser l'évolution de la distance de chaque personne par rapport à la bombe.

Les cartes de métro sont un type de technique de visualisation spatiale où les données sont représentées de manière abstraite. La première représentation de ce type est celle de la carte du métro de Londres, apparue en 1938, et conçue par Harry Beck, où la position des stations n'est pas exactement leur position géographique, mais une position calculée de manière à garder une position relative aux autres stations cohérentes, tout en respectant au maximum certaines contraintes :

- Garder des pentes de 0, 45 ou 90 degrés entre les stations
- Pas de superposition des labels avec trajets
- Garder des distances égales entre les stations

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

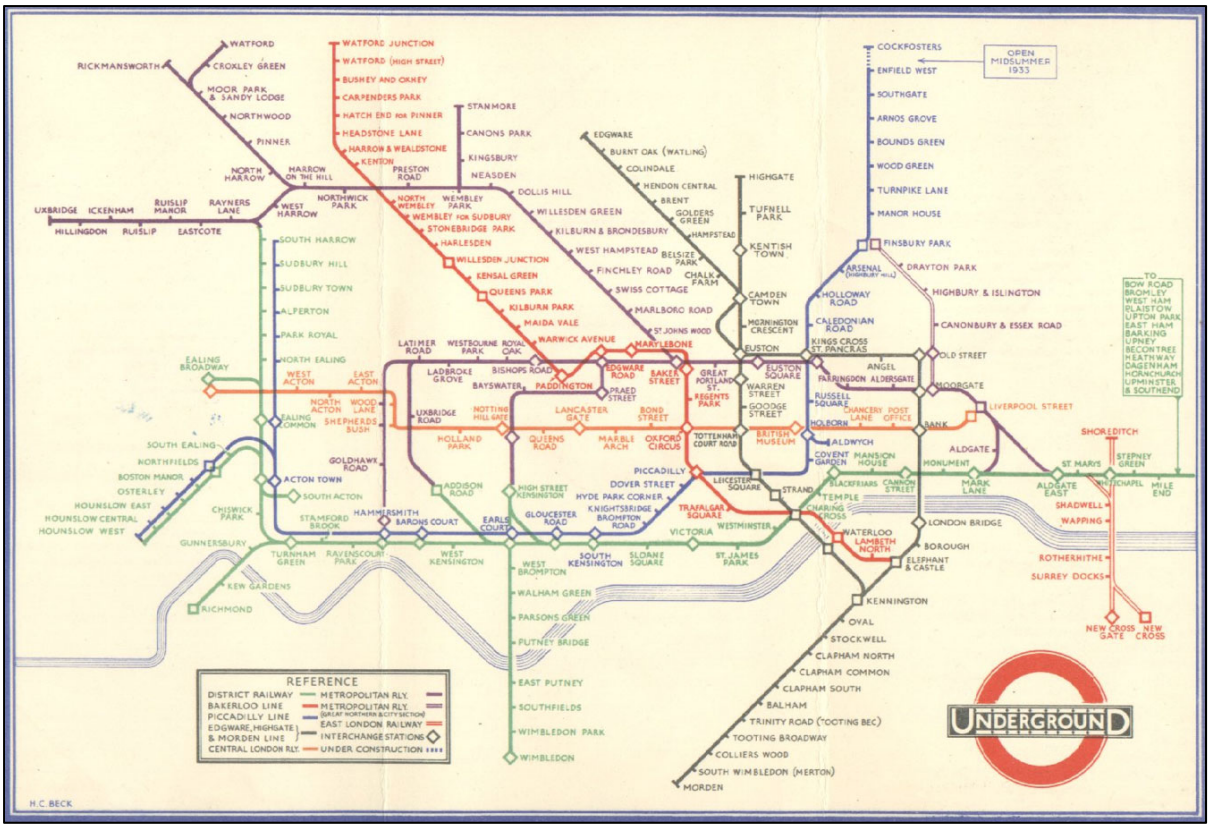

Figure 1.4 Première représentation topologique du métro de Londres
Par Harry Beck (1931)

En 2011, Wang et Chi proposent une technique *Focus + Context* à appliquer à ce genre de représentation topologique (Wang et Chi, 2011). Le principe de cette technique est de choisir une station de départ, ainsi qu'une station d'arrivée, et le trajet le plus court entre ces deux stations est alors mis en surbrillance. La représentation est réagencée, le focus, qui est sur le trajet, a pour effet d'augmenter l'espacement entre les stations qui le composent. Le contexte, qui représente les autres stations du réseau, voit la distance entre ses stations réduites, son opacité baissée de moitié, et les labels ne sont plus affichés. En procédant ainsi, la portion de l'espace utilisable occupée par les données « importantes » est plus élevée.

Figure 1.5 Technique "Focus + Context" appliqué au plan du métro de Sidney
Tirée de Wang et Chi (2011)

Thudt et al. utilisent dans « *Visits* » (Thudt, Baur et Carpendale, 2013) une technique hybride mêlant une visualisation de type plan avec une visualisation basée sur les *small-multiples* (la technique des *small-multiples* consiste à visualiser un ensemble de données sous forme d'une série de petits aperçus (Tufte, 1983)). Ici les aperçus représentent une certaine partie de l'ensemble de données, délimitée spatialement (une zone de la carte) ou temporellement (une période).

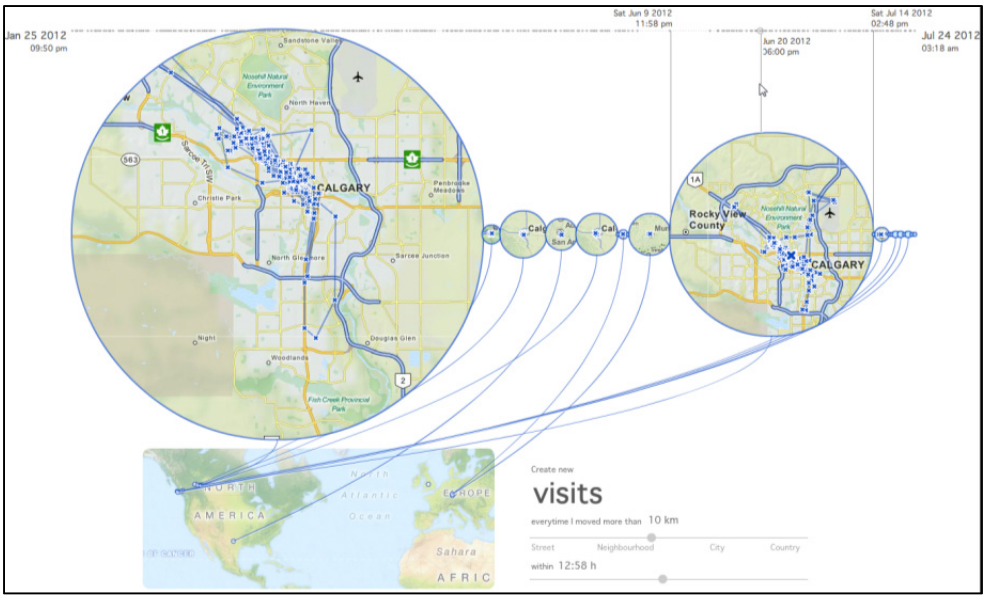

Figure 1.6 Capture d'écran de "Visits", montrant des données spatio-temporelles sous forme de "small-multiples"
Tirée de Thudt, Baur et Carpendale (2013)

1.1.5 Les techniques de visualisation radiale

Dans un article de 2011, Liu et al. utilisent aussi une visualisation de type plan, entourée d'anneaux concentriques (Liu et al., 2011). Dans ce cas-ci, les anneaux représentent une journée dans le temps, avec un histogramme permettant de voir la densité des trajets, par heures, et un des arcs d'anneaux permettant de visualiser l'aspect temporel des trajets.

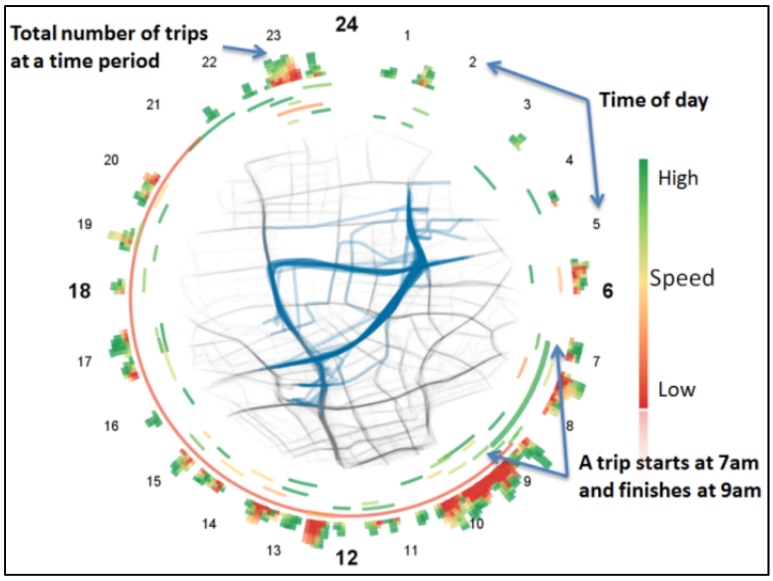

Figure 1.7 Vue "Trip View" permettant de visualiser des trajectoires, la période du jour à laquelle elles apparaissent, ainsi que leur vitesse
Tirée de Liu et al. (2011)

En se basant sur *Circos* (Une technique décrite un peu plus loin, voir la section 1.4.3 « Techniques de visualisation radiales »), Zeng propose une technique permettant de visualiser des trajectoires à des intersections de plusieurs chemins, en appliquant une technique proche du *Edge-Bundling*.

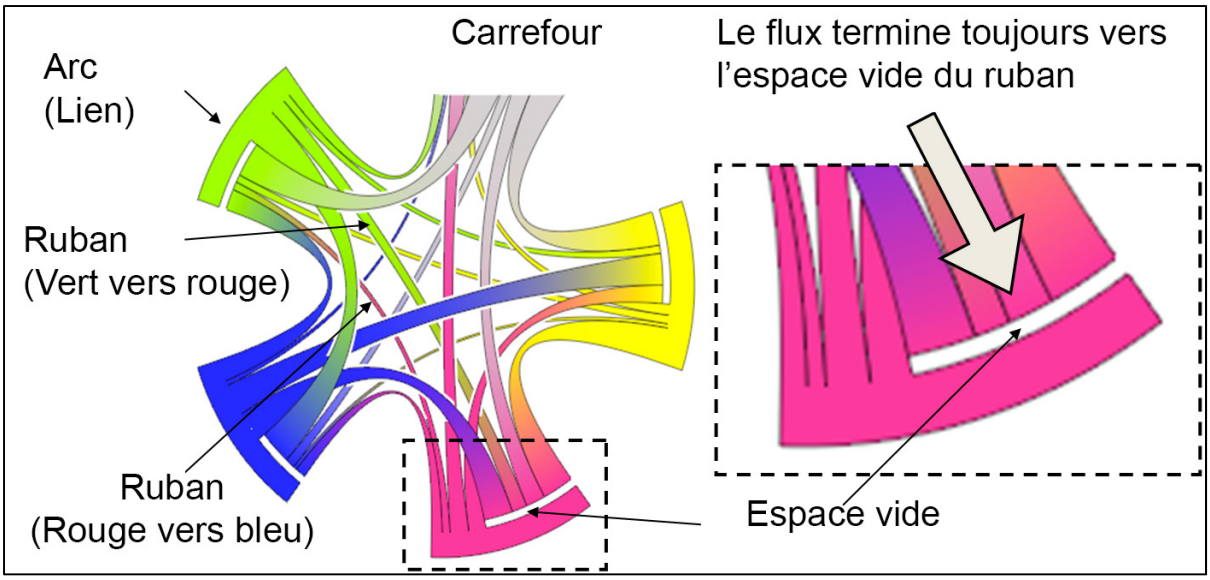

Figure 1.8 Exemple explicatif de l'application de Circos
Adapté de Zeng et al. (2013))

En regroupant les données ayant les mêmes provenances et les mêmes destinations, les données à afficher se trouvent réduites, et la largeur du ruban permet de mesurer la quantité de trajets qu'il représente. De plus, en différenciant les points d'entrée et de sortie des groupes de trajectoires dans les carrefours, cette technique permet de distinguer le sens des trajectoires.

1.1.6 Résumé

Certaines tendances similaires peuvent être distinguées dans cet ensemble d'article. En effet, plusieurs techniques proposent de réduire le nombre de données à afficher en regroupant celles-ci, alors que d'autres techniques ont pour contraintes de préserver les positions relatives des données pour diminuer la confusion.

1.2 La visualisation de graphes

Ivan Herman et al. présentaient en 2000 une vue d'ensemble des techniques de visualisation de graphes et des interactions (Herman, Melançon et Marshall, 2000). Selon lui, les

problèmes principalement rencontrés dans ce domaine sont liés à la taille du jeu de données à afficher. Pour les grands ensembles (la notion de grosse quantité de données ayant évolué depuis les années 2000), plusieurs problèmes peuvent être rencontrés :

- La limite matérielle pour le traitement des données (on omet le système d'affichage) peut être atteinte.
- La compréhension générale du jeu de données n'est plus efficace.
- Même si l'ensemble des données peut être représenté dans l'espace d'affichage, il est impossible d'observer distinctement les nœuds et les arrêtes.

Au niveau des interactions, ils considèrent le zoom et le *pan* comme des outils indispensables, tout en mentionnant que le zoom occasionne une perte des informations contextuelles. Il présente ensuite les techniques de *Focus* + *Context* comme une solution à ce problème.

Enfin, le dernier type de technique mentionné dans la vue d'ensemble d'Herman et al. est le regroupement de données (*Clustering*). Ils mentionnent dans son article deux types de regroupement :

- Les regroupements basés sur la structure du graphe
- Les regroupements basés sur la sémantique des nœuds et arêtes

En 2006, Lee et al. définissent les objets composant un graphe (le graphe lui-même, les nœuds, les liens, les chemins, les clusters, les groupes, et les composants connectés), et présentent une taxonomie des tâches pouvant être effectuées avec des interfaces de visualisation de données spatio-temporelles (Lee et al., 2006). D'abord, les tâches de bas niveau sont décrites puis les tâches de plus haut niveau sont ensuite définies comme une composition de tâches de bas niveau. Dans la même année, Andrienko et Andrienko proposent eux une grammaire complète pour représenter ces tâches (Andrienko et Andrienko, 2006).

1.3 Techniques de visualisations hybrides

Le principe des techniques de visualisation dites hybrides est de combiner plusieurs techniques de visualisation, afin de tirer profit des avantages de chacune de celles-ci, tout en minimisant leurs inconvénients.

Les hiérarchies élastiques (*Elastic Hierarchy*) sont des techniques de visualisation hybrides pour les données hiérarchiques, mêlant les *TreeMaps* et les graphes nœuds-liens, afin de tirer parti des avantages de ces deux techniques (Zhao, McGuffin et Chignell, 2005). Les *TreeMaps* sont décrits comme ayant pour principal avantage leur efficacité spatiale, alors que les graphes nœuds-liens sont plus clairs pour visualiser la structure des données (Johnson et Shneiderman, 1991) (Shneiderman, 1992).

NodeTrix est une autre technique de visualisation, mélangeant cette fois les matrices d'adjacences et les graphes nœuds-liens (Henry, Fekete et McGuffin, 2007). *NodeTrix* est utilisé dans l'article pour visualiser des données issues des réseaux sociaux, celles-ci présentant naturellement des regroupements. L'intérêt de cette visualisation est donc par exemple de montrer les données denses sous forme de matrice, et les autres relations sous forme de graphe nœuds-liens.

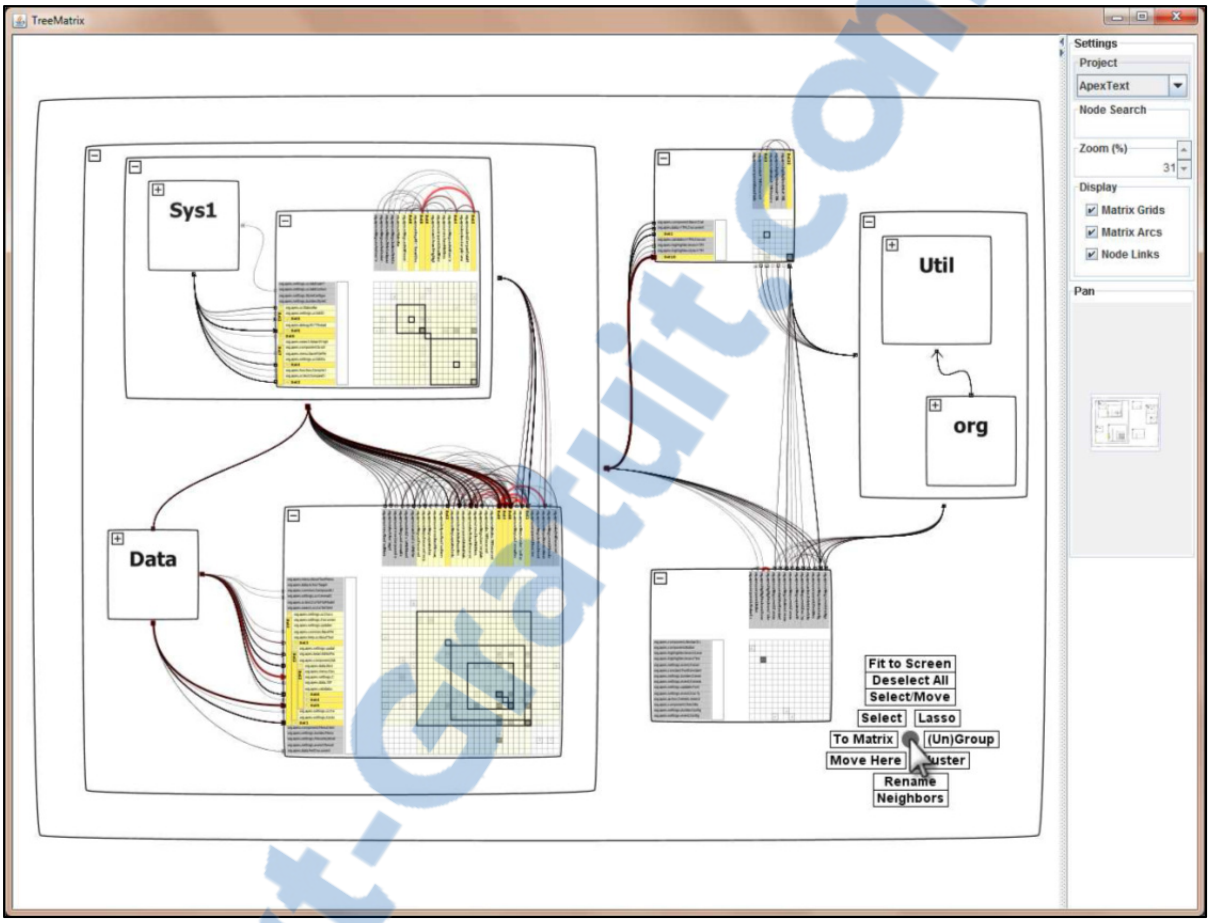

Figure 1.9 Capture d'écran montrant la technique *TreeMatrix*
Tirée de Rufiange, McGuffin et Fuhrman (2012)

TreeMatrix (Rufiange, McGuffin et Fuhrman, 2012) utilise les même techniques que celles utilisées dans la combinaison citée précédemment, en ajoutant les *TreeMaps* et les diagrammes en arc (Wattenberg, 2002), cette-fois ci pour afficher un *Compound Graph* (Un graphe où les nœuds ont aussi un aspect hiérarchique).

TreeMatrix, dans l'exemple ci-dessus (Figure 1.9 Capture d'écran montrant la technique *TreeMatrix*), montre la structure d'un ensemble de fichiers de code source (Classes, imports, etc.). Les rectangles imbriqués, à la façon d'un *TreeMap*, permettent de visualiser l'aspect hiérarchique des classes, alors que les matrices et les liens permettent de visualiser les imports. De la même manière que précédemment, certaines classes sont fortement liées et les

Rapport-Gratuit.com

relations denses sont donc intéressantes à visualiser sous forme de matrices, alors que d'autres relations sont plus intéressantes à visualiser sous forme de graphe-nœuds-liens.

1.4 Techniques de visualisations de liens et arcs

1.4.1 Techniques de visualisations d'arcs

Riche et al., dans leur article sur les liens courbés dans les graphes, (Riche et al., 2012), proposent 4 techniques interactives ayant pour objectif la manipulation des liens d'un graphe.

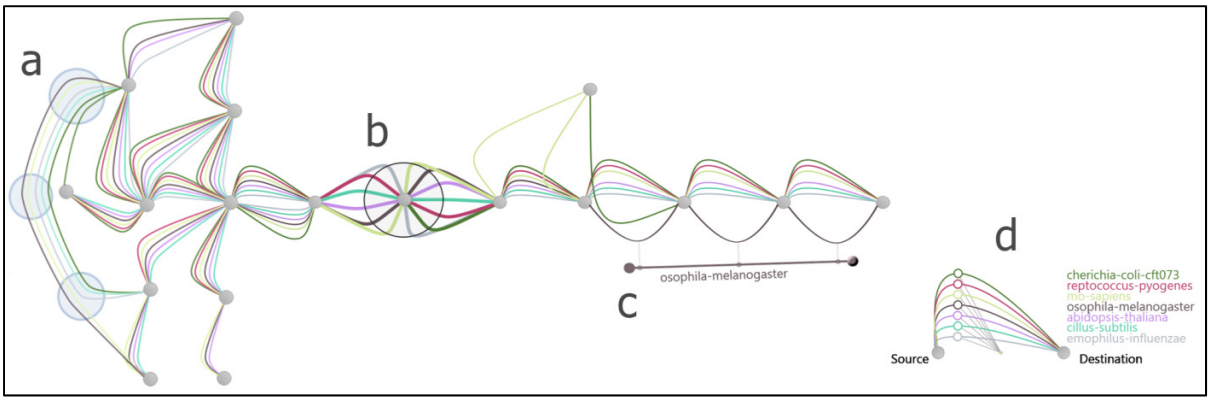

Figure 1.10 Intégration des 4 techniques interactives pour contrôler les courbes dans un réseau biologique : (a) interactive bundling, (b) interactive link fanning, (c) link magnets, (d) interactive legends
Tirée de Riche et al. (2012)

La seconde technique d'interaction proposée permet d'agencer les liens comme des palmes de ventilateur (*Interactive link fanning*). Deux points forment le diamètre du cercle utilisé pour le regroupement et tous les liens passant au travers du cercle sont affectés par l'interaction de cette manière:

- Chaque lien passe par le centre du cercle
- Des points de contrôle sont créés tout autour du cercle, à intervalle régulier. Deux points de contrôle pour chaque lien traversant le cercle de part en part, 1 point de contrôle pour les liens qui prennent leur origine à l'intérieur du cercle.

- Les liens sont ensuite déformés afin de suivre une courbe passant par les points précédemment cités.

Cette technique permet de minimiser l'occlusion dans la zone générée par le cercle, en repoussant au maximum les liens les uns des autres.

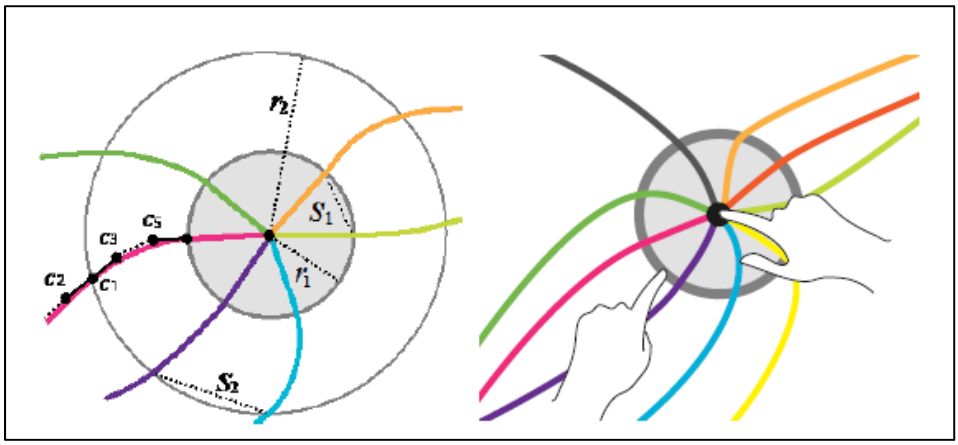

Figure 1.11 Construction de l'agencement en palme de ventilateur, exemple d'interaction multitouche
Tiré de Riche et al. (2012)

Dans un article paru en 2009, Blaas et al. proposent d'utiliser des courbes pour représenter des enchaînements d'états dans un graphe à état (Blaas et al., 2009). Dans ce type de graphe, chaque nœud représente un état distinct, et les liens les transitions possibles. L'idée développée dans l'article est que, plutôt que d'utiliser des liens droits pour représenter un enchaînement de transitions (ce qui causerait de la superposition), on utiliserait des courbes de Catmull-Rom, qui ont pour avantage de passer par chacun des points de contrôle, tout en étant influencé par tous les points de la courbe. Ainsi, seuls les enchaînements d'états identiques sont superposés.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

1.4.2 Agrégation de données

Danny Holten présente dans son article de 2006 une technique appelée *Hierarchical Edge Bundles* (Holten, 2006). Cette technique s'applique à des graphes particuliers, ayant un aspect hiérarchique, et des relations non-hiérarchiques. L'un des exemples de données présentées dans l'article sont les données d'ingénierie logicielles, où il y'a une hiérarchie de packages et classes ainsi que des relations entre les classes (les imports).

Le principe de base utilisé dans cette technique est que si l'utilisateur souhaite montrer une relation entre deux feuilles dans une représentation hiérarchique, plutôt que de le montrer sous forme d'un lien droit, il est possible de le représenter sous forme d'un chemin. Le second principe est que, plutôt que de décrire un chemin passant par chacun des parents reliant deux classes, on va se servir de la position de ces parents comme point de contrôle pour dessiner une courbe. Ainsi, une courbe de type *spline* étant générée à partir de l'ensemble des points de contrôles, le phénomène d'occlusion se trouve diminué.

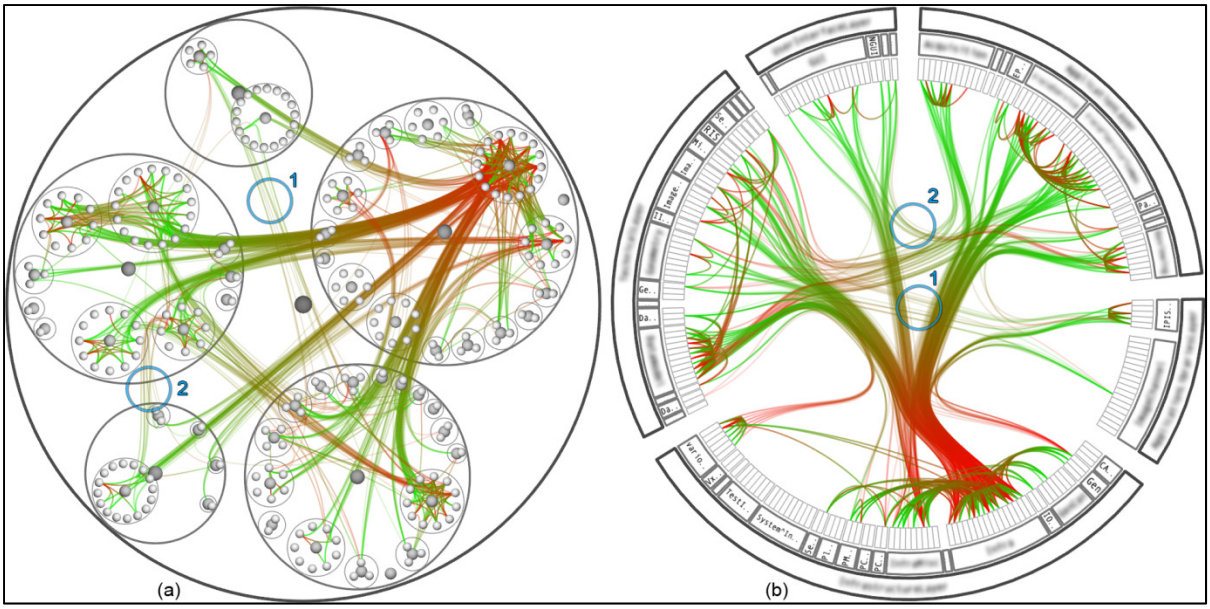

Figure 1.12 (a) La technique *Hierarchical Edge Bundles* appliquée à un arbre avec un agencement de type *Baloon Layout*, (b) les mêmes données avec une représentation de type radiale. Les couleurs donnent une information sur le sens de la relation (le nœud du côté

vert du lien représente la classe qui importe celle du côté rouge du lien)
Tiré de Holten (2006)

Holten étend ensuite sa technique à des données non hiérarchiques, avec la technique *Force Directed Edge Bundling* (Holten et Van Wijk, 2009).

L'Edge *Bundling* a aussi été utilisé pour améliorer l'efficacité des graphes *Layered Graph Layout*, où les nœuds sont organisés de manière verticale sur différentes couches. (Pupyrev, Nachmanson et Kaufmann, 2011).

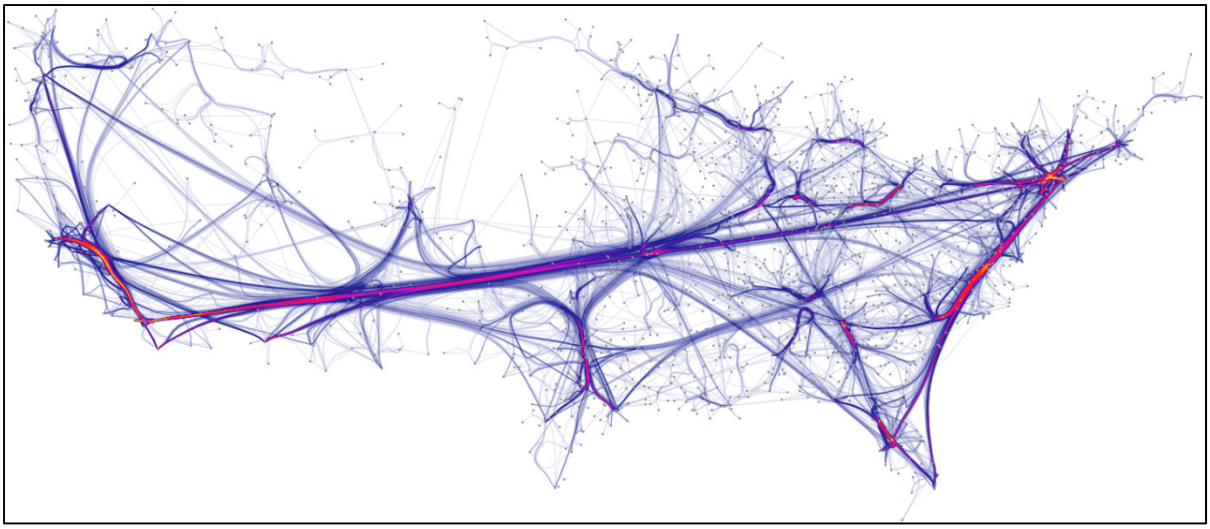

Figure 1.13 Technique de *Force-Directed Edge Bundling* appliquée à un graphe nœuds-liens représentant des flux migratoires aux États-Unis
Tiré de Holten et Van Wijk (2009)

1.4.3 Techniques de visualisation radiales

Le diagramme de Kelp (Dinkla et al., 2012) est une technique dont le principe est proche des *LineSets Diagrams*, utilisant l'agencement circulaire.

Le *LineSets Diagram* (Alper et al., 2011) permet de visualiser des sous-ensembles de nœuds ou de points ayant une position fixe. Chacun des nœuds faisant parti d'un sous-ensemble est

lié à un ou deux nœuds de ce même sous-ensemble, formant ainsi une chaine. Le sous-ensemble est ensuite représenté par une courbe reliant chacun de ces nœuds, d'un seul trait.

Figure 1.14 (a) L'approche BubbleSet, (b) l'approche LineSets,
(c) l'approche diagramme de Kelp
Tirée de Dinkla et al. (2012)

Le diagramme de Kelp se base sur un principe semblable au LineSets, à la différence ici que les lignes ne sont pas dessinées d'un seul trait, mais de manière à ce que le diagramme soit de taille minimum tout en diminuant les intersections.

Ces deux techniques ont en commun d'utiliser des cercles concentriques pour différencier les cas où un élément fait partie de plusieurs ensembles. Chaque cercle concentrique représentant l'appartenance à un ensemble.

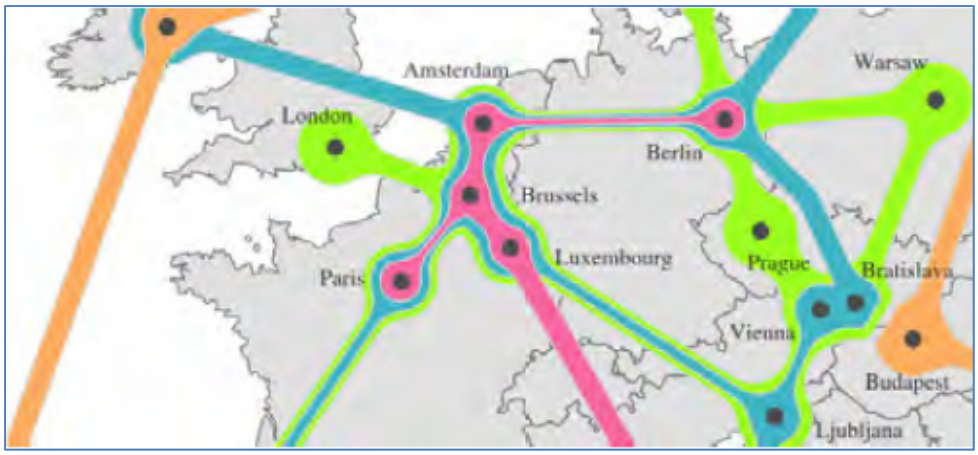

Figure 1.15 Diagramme de Kelp
Tirée de Dinkla et al. (2012)

Circos est une technique de visualisation radiale permettant de visualiser différents génomes (Krzywinski et al., 2009). Les données que l'on souhaite comparer sont affichées de manière concentrique, où chaque piste correspond à un échantillon, et chaque partie de piste correspond à un chromosome.

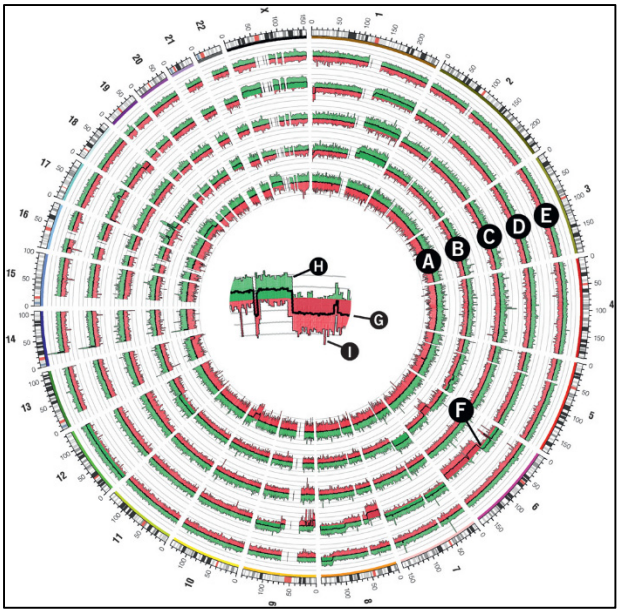

Figure 1.16 Comparaison de 5 génomes (A, B, C, D, E), composés de 23 chromosomes avec *Circos*
Tirée de Krzywinski et al. (2009)

CHAPITRE 2

PROBLÉMATIQUE

Un système de géolocalisation nécessite principalement deux dispositifs. Un dispositif d'acquisition, pour récupérer les données, et un dispositif permettant la visualisation. Les téléphones intelligents sont de nos jours largement popularisés, et ceux-ci disposant pour la plupart d'un GPS, ils peuvent servir de dispositif d'acquisition.

L'un des dispositifs de visualisation grand public jusqu'à récemment est Google Latitude (voir Figure 2.1 ci-dessous, le service a fermé au courant du mois d'août 2013). Cette interface permet de visualiser les données sur un laps de temps désiré, en représentant chaque position par un rond, et en les reliant avec la précédente par un trait droit.

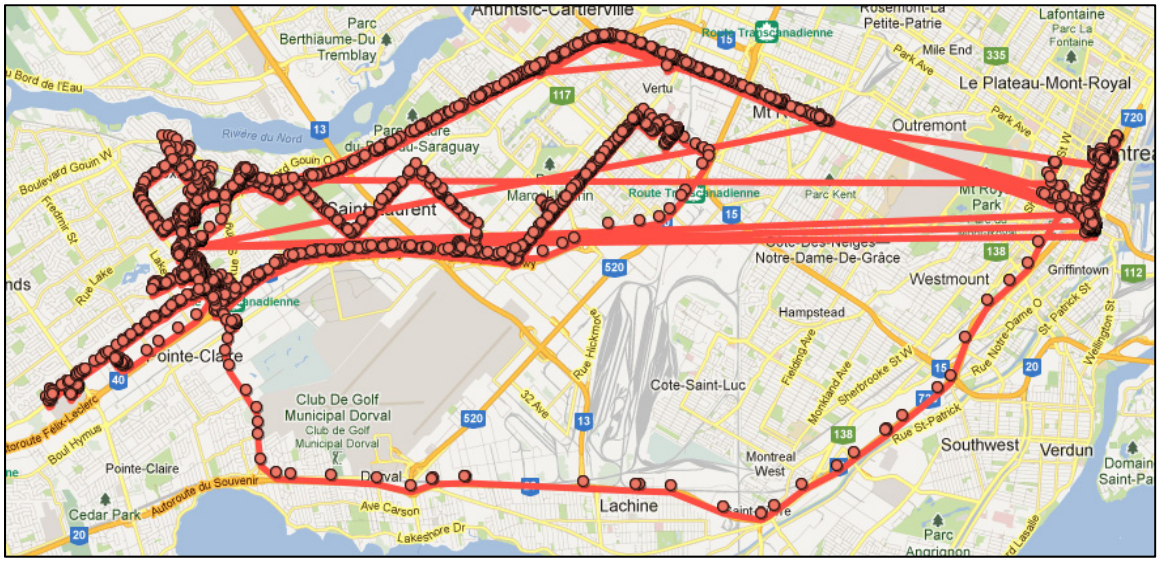

Figure 2.1 Données GPS correspondant aux déplacements sur une période d'un mois (Visualisées à l'aide de Google Latitude - https://latitude.google.com/)

On observe sur la capture d'écran ci-dessus, qui représente des données sur un mois à intervalle de 3 minutes, beaucoup de superposition, et un phénomène d'occlusion assez important. Il est ainsi impossible de déterminer le nombre de trajets distincts.

Comme mentionné précédemment, il est possible de classer les tâches principalement effectuées sur ce type d'interface selon plusieurs critères. La description ci-dessous ne se veut pas exhaustive, le but étant juste de décrire plusieurs types de questions. Andrienko et Andrienko ont proposé une théorisation des tâches relatives aux données spatio-temporelles (Andrienko et Andrienko, 2006).

Tout d'abord, les tâches peuvent être classées selon le type de données sur lequel porte la question. Les trois types sont :

- La localisation, la position d'un objet est alors recherchée. Exemple : « À quel endroit se trouve l'objet vert au moment *t* ?»
- La temporalité, un moment dans le temps est recherché. Exemple: « À quel heure l'objet vert se trouve-t-il à la position *x* ?»
- La cible, un objet cible est recherché. Exemple : « Quel objet se trouve à la position *x* au moment *t* ?»

Les tâches ont aussi des portées différentes. Dans les questions précédentes, la portée est dite « simple », on s'intéressait à une position, à un moment dans le temps, ou à une cible en particulier. La portée peut être plus globale, et s'intéresser à un ensemble : à une zone, à un intervalle de temps, ou à un sous ensemble de cibles. On pourrait donc avoir des questions de ce type :

- « Quelles positions sont visitées par l'objet vert durant l'intervalle de temps *i* ?»
- « Quelles positions sont visitées par les objets du sous-ensemble *A* durant l'intervalle de temps *i* ?»
- « À quelle heure l'objet vert se trouve-t-il dans la zone *z* ?»
- « Quelles cibles se trouvent dans la zone *z* durant l'intervalle de temps *i* ?»

Enfin, dans les exemples précédents, on s'intéresse particulièrement à sélectionner des données. Des variantes de tâches peuvent aussi exister, où les questions sont ciblées sur la vitesse ou l'ordonnancement :

- « Quelle cible est arrivée la première dans la zone *z* ?»
- « Quelle cible s'est déplacée le plus rapidement durant l'intervalle de temps *i* ?»
- « Dans quel ordre la cible verte a-t-elle traversé les zones *u v* et *w* ?»

Figure 2.2 Détails de la
Figure 2.1

Ci-dessus, une capture d'écran zoomée de la capture précédente où l'on peut voir une forte densité de points sur une petite zone. Il est alors très compliqué de répondre aux questions précédentes. Il n'est, par exemple, pas possible de répondre à une question de type « Dans quel ordre la cible verte a-t-elle traversé les zones *u*, *v* et *w* ?».

L'une des méthodes qui peut être appliquée pour améliorer la lisibilité des données est d'appliquer une étape de *clustering*, et donc de rajouter une couche d'abstraction aux données. L'une des idées envisagée est de réduire le plan sous la forme d'un graphe nœuds-liens, où les nœuds représenteraient des lieux précis, et les liens des chemins entre deux lieux précis. Cette méthode permettrait de simplifier les trajectoires, en réduisant la quantité de données visualisées, tout en permettant de visualiser les allers-retours et les trajectoires multiples, en ajoutant des décalages entres les courbes, de manière à éviter les occlusions. Cette méthode sera appelée dans ce document, la technique de visualisation topologique.

Rapport-Gratuit.com

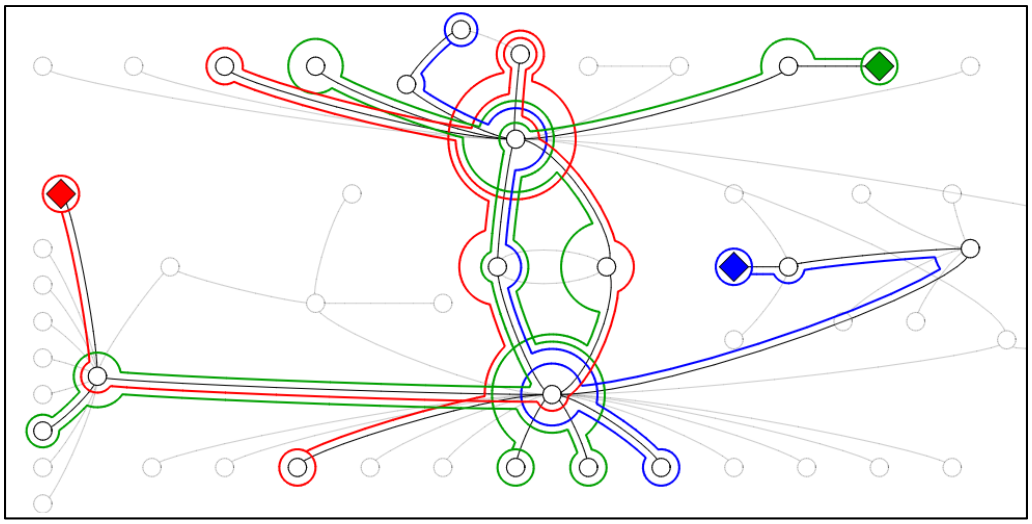

Figure 2.3 Exemple de technique de visualisation topologique

Néanmoins, si cette méthode peut permettre de différencier les trajectoires entre les différentes zones, cela ne permet pas de les distinguer au sein d'une zone. Pour cela, la visualisation géospatiale reste plus efficace. En effet, avec ce type de visualisation, on peut voir les détails à l'intérieur d'une pièce. Une visualisation hybride pourrait donc permettre de tirer profit des deux types de visualisation afin de permettre, en fonction des tâches que l'on souhaite réaliser, et de la densité de données dans certaines zones, de choisir la meilleure des deux visualisations pour obtenir un compromis plus efficace.

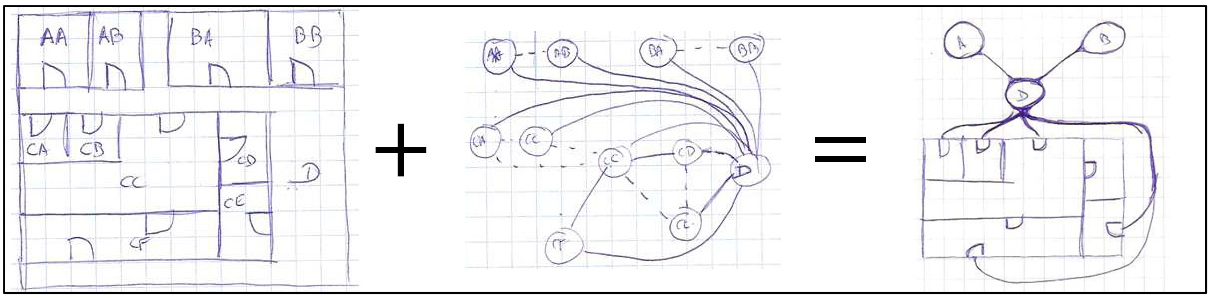

Figure 2.4 Croquis initial montrant les différentes techniques de visualisation et un hybride entre les deux techniques

CHAPITRE 3

PRÉSENTATION DE LA SOLUTION

Afin de tester les techniques de visualisation présentées dans le chapitre précédent, un prototype a été réalisé, dans le but d'effectuer une expérience contrôlée. Dans ce chapitre, les techniques de visualisation géospatiales et topologiques seront présentées, au travers de différentes captures d'écran. Les fonctionnalités implémentées seront aussi décrites, alors que des fonctionnalités envisagées mais non implémentées dans le prototype seront aussi mentionnées. Enfin, une section de ce paragraphe présentera la technique de visualisation hybride, avant de conclure par un comparatif succinct de ces trois techniques.

3.1 Présentation général

Certaines caractéristiques sont communes aux trois techniques de visualisation présentées ci-dessus :

- Il est possible de sélectionner une ou plusieurs pièces, qui seront alors visualisées avec un contour vert.
- Au passage de la souris, les pièces sont affichées en surbrillance, par un contour, cette fois ci en pointillé.
- Enfin, un glisseur permet de contrôler l'aspect temporel du prototype (Figure 3.1). Il s'agit du seul moyen d'interagir avec le temps dans le prototype. Un mode de lecture automatique a été implémenté mais non utilisé dans l'expérience contrôlée.

Figure 3.1 Contrôleur de temps

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

3.2 La visualisation géospatiale

La visualisation géospatiale est le type de visualisation de trajectoires classique, où la représentation du plan est proche de la réalité. On le retrouve par exemple sur GoogleMap, pour afficher des itinéraires et dans TripVista (Guo et al., 2011). Dans le cas d'un bâtiment, la visualisation en mode plan est proche d'un plan d'architecte.

Plusieurs choses sont à noter sur l'implémentation de cette visualisation :

- Chaque pièce est dessinée avec une texture, représentant l'agencement de la salle (meubles, murs, etc.)
- Les portes sont représentées par des quarts de cercles
- Au survol de la souris, les pièces sont placées en surbrillances (un contour en pointillé vert est dessiné)
- Les pièces sélectionnées sont affichées avec un contour vert.

Au niveau des trajectoires, celles-ci sont dessinées comme des lignes brisées où chaque point successif est relié par une ligne droite. La position actuelle de la cible est représentée par un carré de la même couleur que la trajectoire (Figure 3.2).

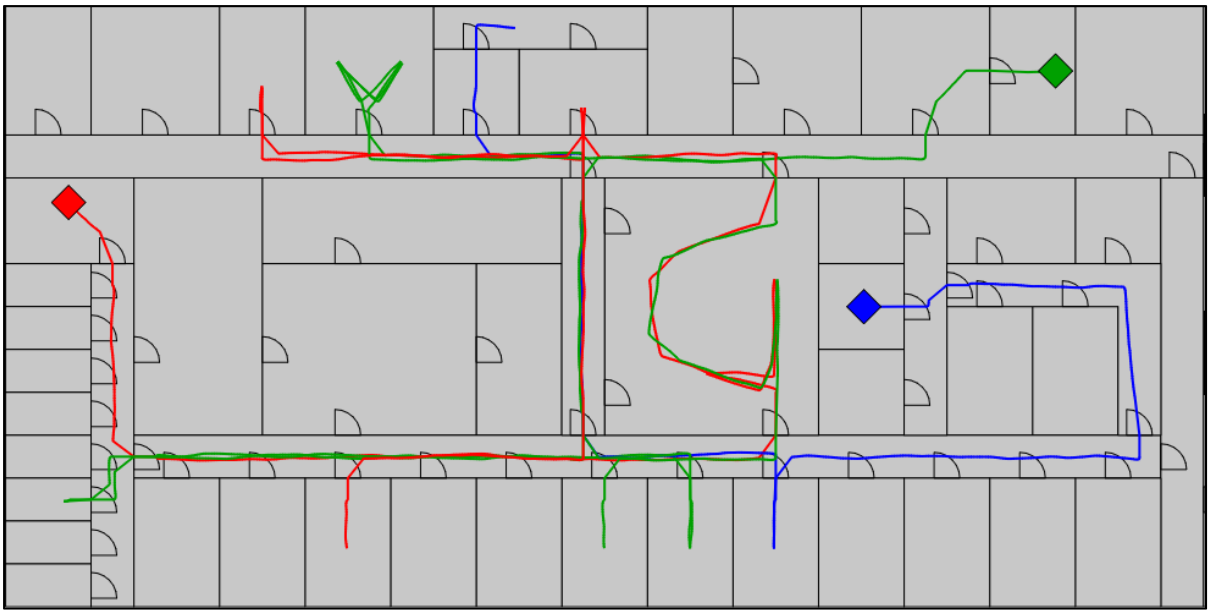

Figure 3.2 Technique de visualisation géospatiale. Ici, les trois losanges rouge vert et bleu montrent la position actuelle de trois personnes.

3.3 La visualisation en mode topologique

Assez rapidement, si on affiche plusieurs trajectoires d'une durée assez conséquente, on ne parvient plus à visualiser les trajectoires distinctement. Le principe que l'on a cherché à appliquer ici est de simplifier les détails des informations afin de rendre les trajets plus faciles à distinguer.

Dans le cas de la vue topologique, on va chercher à montrer un aperçu abstrait du plan, en montrant les pièces et les relations entre elles. On remarque que deux types de relations peuvent exister entre deux pièces :

- Une relation d'adjacence (Si deux pièces partagent un mur)
- Une relation de voisinage (Si deux pièces partagent une porte)

Le principe de la vue topologique est alors de n'afficher que ces relations dans un graphe nœuds-liens. Les nœuds représentant les pièces, et les liens représentant les différents types

de relations. Les relations de voisinage (les portes donc) sont affichées par des liens pleins, alors que les relations d'adjacence sont affichées par des liens en pointillés (Figure 3.3).

L'utilisation d'un algorithme d'agencement basé sur les forces a été envisagée. Néanmoins, la position des nœuds étant la même que celle des pièces, l'application de l'algorithme aurait permis une meilleure lisibilité du graphe en tant que tel (en évitant les croisements d'arcs) mais la compréhension de l'aspect géographique du plan aurait été plus compliqué. Le choix a donc été fait de ne pas appliquer un tel algorithme.

De la même manière que pour la visualisation géospatiale, il est possible de mettre une pièce en surbrillance, ou d'en sélectionner une ou plusieurs.

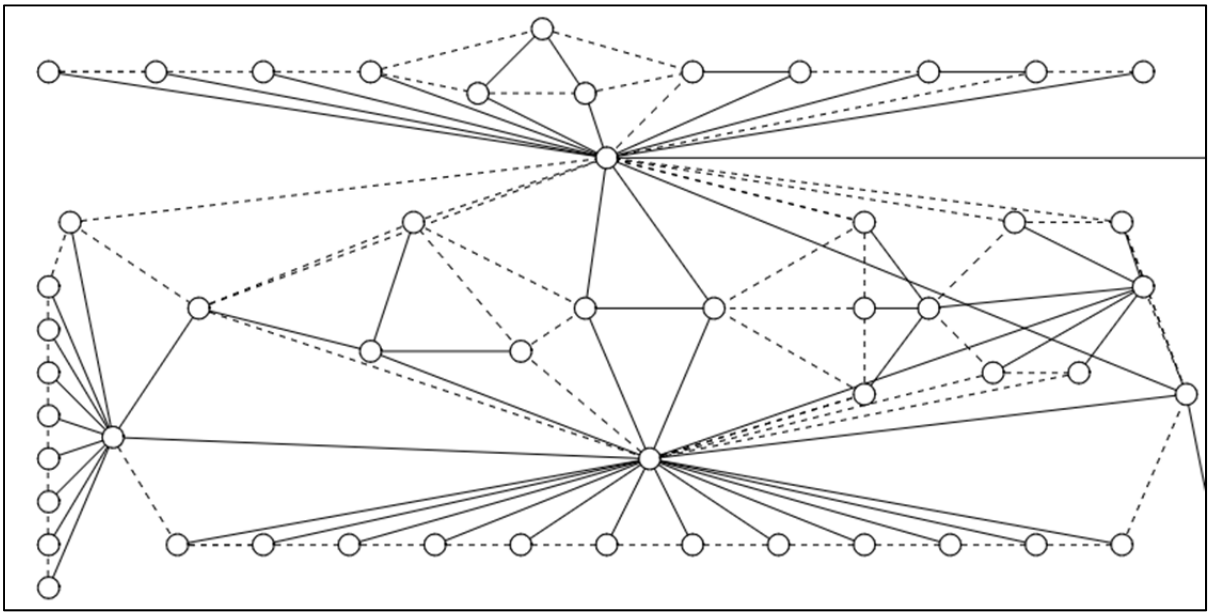

Figure 3.3 Vue topologique avec relations de voisinage et d'adjacence. Les traits pleins représentent ici les portes, alors que les traits en pointillés représentent les « adjacences », c'est-à-dire deux pièces séparées par un mur.

L'affichage des relations d'adjacence a finalement été abandonné, celui-ci rendant la vue topologique confuse, sans rajouter d'information importante (Figure 3.4).

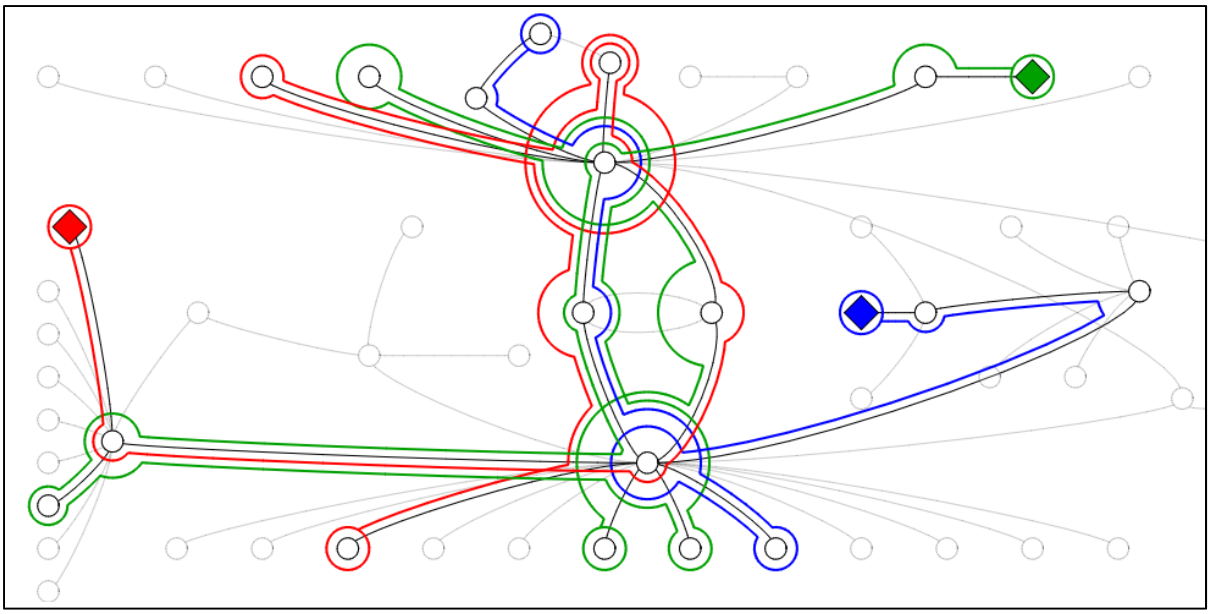

Figure 3.4 Visualisation topologique avec l'affichage de trois trajectoires

En ajoutant une couche d'abstraction à notre modèle, on ne visualisera plus les trajectoires dans leur totalité. Ce type de vue offre un aperçu global des données, permettant de visualiser :

- L'ordre de visite des pièces
- Les allers retours
- L'ordre de passage dans des pièces

3.4 La visualisation hybride

Chacune des vues précédemment mentionnées a des avantages et des inconvénients. L'idée de la vue hybride est de visualiser chaque pièce individuellement sous forme topologique ou sous forme géospatiale.

Dans l'implémentation de la vue hybride, il est possible de changer la forme d'une pièce en cliquant dessus (Passant d'une vue géospatiale à une vue topologique, et inversement). Il est aussi possible de changer la forme de plusieurs pièces, à l'aide d'un lasso.

Deux approches peuvent être envisagées :

- Laisser à l'utilisateur l'entière liberté de choisir le type de visualisation
- Proposer algorithmiquement un agencement où la forme de chaque pièce sera choisie en fonction des avantages et inconvénients de chacune des vues en fonction des données. Par exemple, l'algorithme pourrait choisir de mettre sous forme géospatiale toutes les chambres où il y a une rencontre entre deux ou plusieurs personnes, et mettre toutes les autres chambres sous forme topologique.

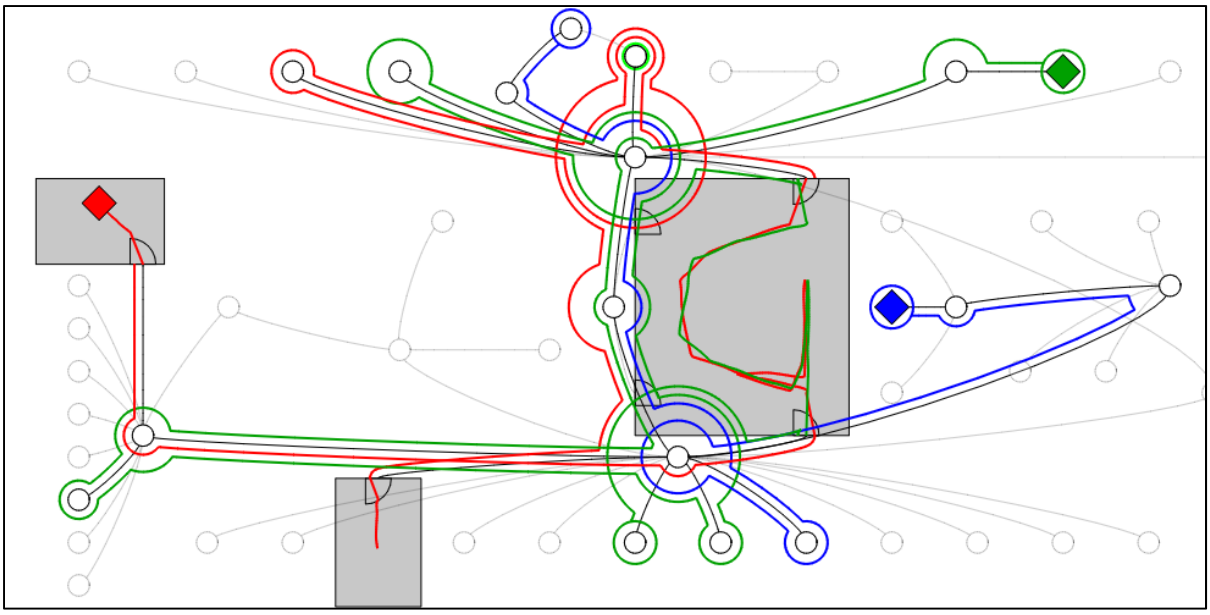

Figure 3.5Visualisation hybride

Une fonctionnalité permettant de déplacer les pièces a aussi été implémentée. Il est possible de déplacer les pièces une à une, ou de déplacer un ensemble de pièces sélectionnées au préalable.

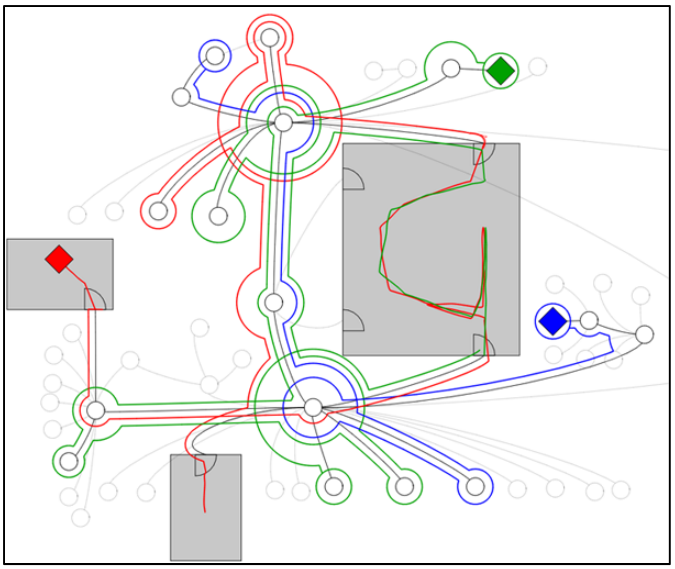

Figure 3.6 Visualisation hybride après réagencement manuel des pièces

Dans l'exemple ci-dessus, les pièces sont réagencées afin de réduire l'espace utilisé pour représenter le bâtiment, tout en gardant une cohérence dans la position relative des pièces. La technique possède ici des caractéristiques « focus-in-context » et « multi-focales », car elle permet de visualiser certaines parties du plan en détail, tout en gardant le contexte visible.

3.5 Résumé

Le tableau ci-dessous résume les avantages et inconvénients de chacun des trois types de visualisation en fonction de différentes caractéristiques pouvant influencer la qualité et la facilité de visualisation des données.

Tableau 1 Comparatif des caractéristiques des différentes techniques de visualisation

	Vue plan	**Vue topologique**	**Vue hybride**
Utilisation du contrôleur de temps	Oui	Non	Non
Permet de voir les détails à l'intérieur d'une pièce	Oui	Non	Oui
Permet de comprendre l'ordre des visites d'une pièce	Oui (En utilisant le contrôleur de temps)	Oui	Oui
Permet d'évaluer la fréquence des passages	Oui (En utilisant le contrôleur de temps)	Oui	Oui
Facilité d'utilisation pour un débutant	Facile	Nécessite une période d'entrainement	Nécessite une période d'entrainement
Phénomène d'occlusion	Important	Faible	Faible
Espace utilisé	100%	100% (Peut être réduit grâce au réagencement)	100% (Peut être réduit grâce au réagencement)

CHAPITRE 4

CHOIX DE CONCEPTION

Ce chapitre va aborder les différents choix de conception qui ont été effectués lors de la réalisation du prototype. Dans un premier temps, il s'agira des choix liés à l'agencement du plan / graphe, puis des choix quant au modèle topologique utilisé. Ensuite, les choix relatifs aux arêtes d'adjacences seront évoqués, avant de terminer sur la représentation des liens et des trajectoires.

4.1 L'agencement

Comme vu dans l'article de Herman et al., « *Graph Visualization and Navigation in information Visualization : A Survey* » (Herman, Melançon et Marshall, 2000), l'aspect esthétique le plus important dans le domaine de la visualisation de graphe est d'éviter au maximum les croisements d'arrêtes. Les bonnes pratiques qu'il mentionne sont les suivantes :

- Vérifier que le graphe est un graphe planaire. Il s'agit d'une condition nécessaire et suffisante pour qu'un graphe puisse être dessiné sans croisement d'arcs.
- Appliquer un algorithme d'agencement, pour modifier la position des nœuds et ainsi éviter les croisements

Ici, de par la nature des données représentées par le graphe nœuds-liens, la position des pièces est un aspect cognitif très important pour la compréhension générale du plan. Ainsi, le choix a été fait de ne pas appliquer d'algorithme d'agencement automatique, afin de garder la position physique réelle des pièces.

Rapport-Gratuit.com

Néanmoins, une fonctionnalité permettant de déplacer les pièces / nœuds manuellement a été implémentée. L'avantage de ce réagencement manuel par rapport au réagencement automatique est que l'utilisateur prend la décision de déplacer les pièces, et il peut donc les déplacer pour garder une position qu'il jugera non-confondante pour lui. Dans la figure présentée précédemment (Figure 3.6), les pièces sont réagencées de manière à ce que les pièces voisines dans la réalité le restent aussi dans la vue topologique. L'idée appliquée dans cet exemple est de réduire les distances entre les pièces les moins importantes dans la visualisation (ici, les pièces qui ne sont traversées par aucune trajectoire) afin que les données les plus utiles occupent une proportion de l'espace visible plus importante.

4.2 Type de modèle topologique

Le paragraphe suivant abordera d'une manière succincte certaines notions de base de la théorie des graphes, et les notations suivantes seront utilisées:

- N, le nombre de nœuds. N_i représentera le nombre de nœuds de degré i (connectés à i liens). N_{i+} représentera le nombre de nœuds de degré i ou plus.
- A, le nombre d'arcs.
- F, le nombre de faces ou régions. Une région, dans un graphe planaire, correspond à une zone délimitée par des arrêtes, ne contenant pas elle-même de sous-région.
- $G(N,A,F)$, le graphe composé de N nœuds, A arcs et F faces. Si le nombre de faces n'est pas représentatif (dans le cas des hypergraphes) on notera juste $G(N,A)$.

Certaines règles sont vérifiées vis-à-vis du nombre de nœuds et d'arcs :

- Si le graphe est connexe (il existe une chaîne passant par tous les nœuds), alors $A \geq N - 1$
- Si le graphe est planaire (pas de croisement d'arcs) alors $F = A - N + 2$

Plusieurs méthodes peuvent être utilisées pour modéliser un plan sous une forme topologique. Dans le chapitre précédent, la méthode décrite est de considérer les pièces comme des nœuds, et les portes comme des liens. La figure ci-dessous (Figure 4.1) représente les différentes visualisations qui ont été évaluées.

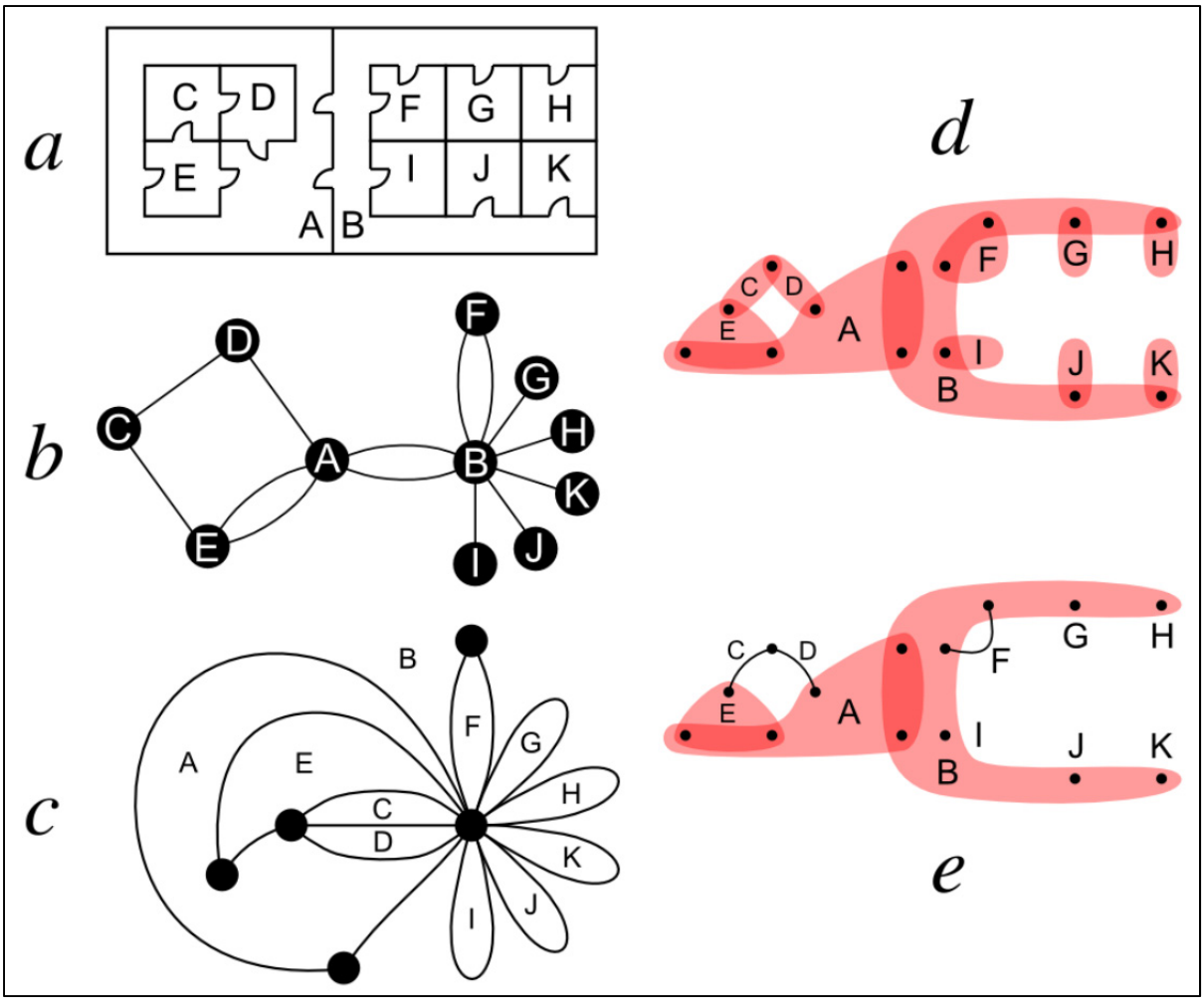

Figure 4.1 Différentes manières de modéliser le plan. *a*: Le plan complet. *b*: Graphe nœuds-liens, avec les salles qui sont représentées par des nœuds, et les portes par des liens. *c*: Le graphe dual géométrique associé au graphe nœuds-liens. *d :* Le *line-graphe* associé au graphe nœuds-liens. *e :* Version simplifiée du *line-graphe*

Les Figure 4.1.*a* et Figure 4.1.*b* représentent respectivement les vues géospatiale et topologique comme décrites dans le chapitre précédemment. Soit *n* le nombre de pièces, et *a* le nombre de portes, le nombre de nœuds *N* dans le graphe est donc égal à *n* et le nombre

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

d'arcs A est égal à a. En partant du principe que la topologie du bâtiment représente un graphe connexe (Pour n'importe quelle pièce il existe au moins un chemin pour rejoindre toutes les autres pièces du bâtiment), l'inégalité $a \geq n - 1$ est vérifiée. Si le graphe est planaire (Ce qui est le cas pour un plan représentant un étage), en appliquant la formule $F = A - N + 2$, le nombre de faces f du graphe peut être calculé, et est ici égal à 5 (En incluant la face externe). Le graphe représenté par le schéma b, noté G_t, (t pour *topologique*) peut donc être décrit avec les propriétés $G_t\,(11,14,5)$

La Figure 4.1.*c* représente une vue topologique en utilisant la notion de graphe dual. Le graphe dual est construit en utilisant les deux règles suivantes :

- À chaque face correspond un nœud
- Pour toutes faces adjacentes on ajoute un lien entre les deux nœuds. Les adjacences étant les arrêtes du graphe de base, le nombre d'arrêtes du graphe et de son duel sont donc égales.

Le graphe dual est appelé ainsi car sa définition est symétrique. En effet, si G_1 est le graphe dual de G_2, alors G_2 est aussi le graphe dual de G_1.

Selon la définition, le graphe dual G_2 de $G_1(n, a, f)$, est défini avec les propriétés suivantes $G_2(f,a,n)$.Pour le schéma c, qui représente le graphe dual du plan G_d (d pour *dual*), le graphe est définie par les propriétés $G_d(5,14,11)$.

Les Figure 4.1.*e* et Figure 4.1.*f* représentent les *line-graphes* correspondant au plan. Un *line-graphe* est construit en suivant les règles suivantes :

- À chaque lien du graphe source correspond un nœud dans le *line-graphe.*
- Si deux liens ont un noeud en commun dans le graphe source, alors on ajoute un lien dans le *line-graphe*.

Ainsi, selon les deux règles, et comme on peut le voir sur la figure ci-dessous (Figure 4.2), un nœud de degré *x* dans le graphe source, donnera une clique dans son *line-graphe* associé (une clique étant un ensemble de nœuds tous connectés entre eux).

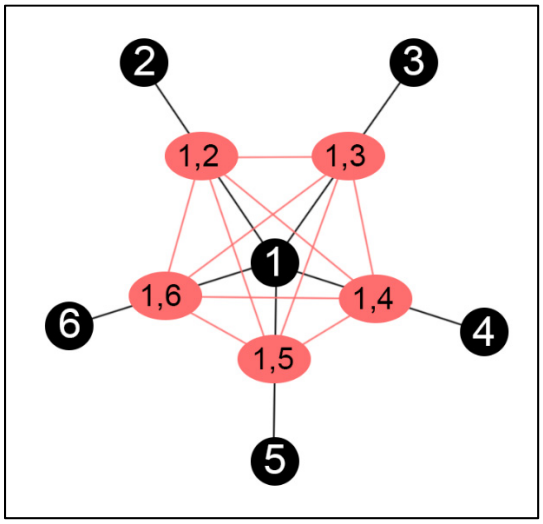

Figure 4.2 Un graphe (en noir) et son
line graph associé (en rouge)

Un hypergraphe est un type de graphe où les arcs connectent un ensemble de nœuds, avec une représentation similaire aux diagrammes de Venn.

Les Figure 4.1.*d* et Figure 4.1.*e* nous montrent donc le *line-graphe* associé au plan. L'ensemble noté A, par exemple, correspond à une clique, englobant toutes les portes connectées entre elles par la salle A. Dans la Figure 4.1.*b*, le nœud A est de degré 5, l'ensemble A des Figure 4.1.*d* et Figure 4.1.*e* est donc composé de 5 nœuds.

La Figure 4.1.*e* est une simplification de la Figure 4.1.*d* :

- Quand un nœud ne fait partie que d'un et un seul ensemble (Les portes menant à une salle n'ayant qu'une porte, comme le nœud *H* par exemple), les ensembles unitaires ne sont plus représentés.
- Les cliques ne contenant que deux nœuds sont représentées par des liens. (Exemple, le lien correspondant à la salle *C*)

L'avantage de ces deux dernières représentations est donc de réduire le nombre de liens affichés.

Tableau 2 Comparatif du nombre de liens pour chaque visualisation

	Nombre de liens (Boucle non incluses)	**Nombre de boucles ou sous-ensembles**	**Total**
Graphe *G(r,d,f)*	a	0	a
Graphe dual *G(f,d,r)*	$a - n_1$	n_1	a
Line-graphe *G(d,r)*	0	n	n
Line graphe simplifié *G(d,r-r₁)*	n_2	n_{3+}	$n - n_1$

En utilisant l'inégalité décrite précédemment ($a \geq n - 1$), on peut déduire que le *line-graphe* simplifié est la représentation qui nécessite le moins d'arcs. De plus, dans le cas d'un bâtiment, où beaucoup de pièces n'ont qu'une seule porte (i.e. un nombre de nœud de degré 1 élevé), le *line graphe* simplifié serait encore plus intéressant dans le but de réduire le nombre d'arcs.

Néanmoins, pour le prototype, le choix a été fait d'utiliser la représentation du graphe nœuds-liens classique (Figure 4.1.*b*), dans le but de faciliter la compréhension par l'utilisateur.

4.3 Les liens

Plusieurs méthodes ont été testées pour l'affichage des liens dans la vue topologique. Ci-dessous, les différentes méthodes seront détaillées, avec les avantages et inconvénients de chacune.

4.3.1 Les liens de type « liens droits »

Tout d'abord, la première méthode implémentée a été d'utiliser des liens droits entre chaque nœud.

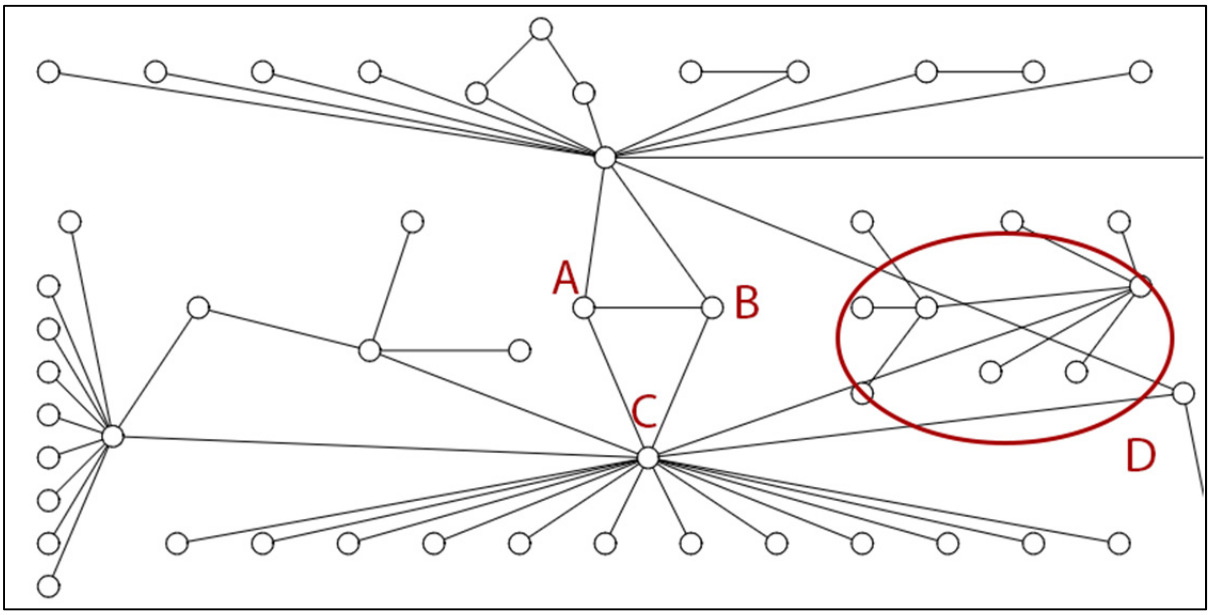

Figure 4.3 Mode topologique avec les liens droits

Comme on peut le voir avec la Figure 4.3 ci-dessus, les liens droits ne permettent pas de différencier les liens multiples entre deux nœuds, comme entre les nœuds A et B (Il y'a deux portes entre les pièces notées A et B, mais un seul lien est visible). Le second problème qui apparait avec ce type de lien est un phénomène de croisement de lien (*Edge-crossing*), dans la zone D par exemple. Néanmoins, ces liens ont l'avantage de présenter un phénomène d'occlusion limité pour les nœuds ayant beaucoup de liens, comme le nœud C.

4.3.2 Les liens de type « lignes brisées »

Dans un premier temps, on essaye d'utiliser des lignes brisées afin de régler les deux problèmes cités ci-dessus. Les lignes brisées sont définies par 5 points. Le centre de chaque

nœud concerné (*ef1* et *ef2* sur la Figure 4.4), le point d'ancrage de la porte (*ef4*) et les deux points perpendiculaires au mur contenant la porte (*ef3* et *ef5*).

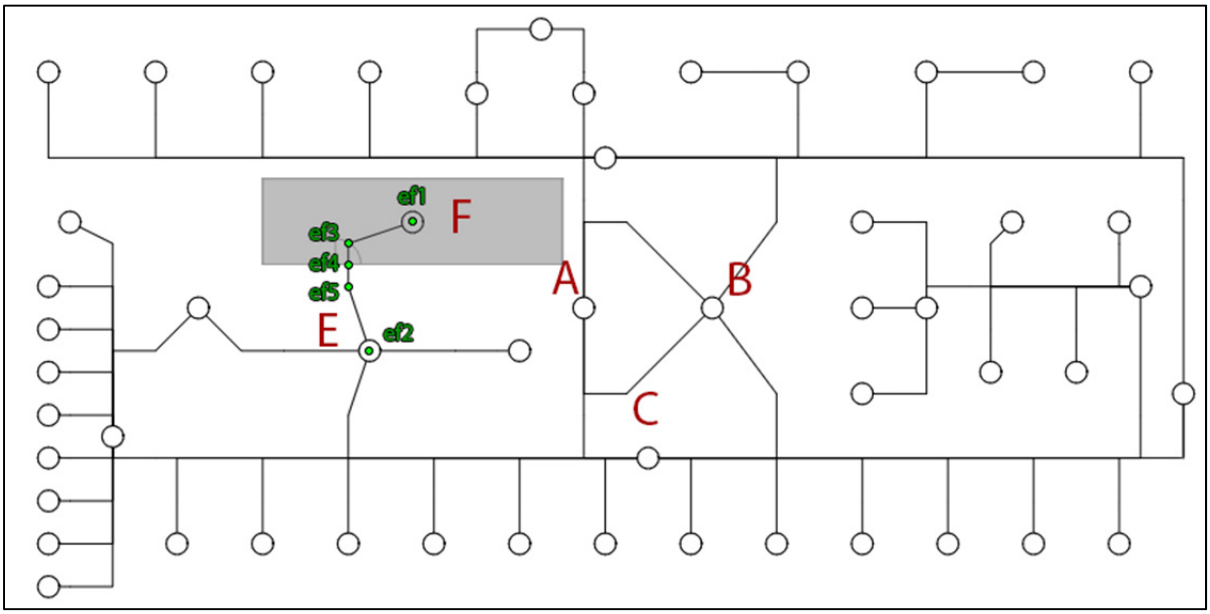

Figure 4.4 Mode topologique avec des lignes brisées

Cette fois, on peut bien voir les deux liens entre A et B de manière distincte, et on n'observe plus de phénomène de croisement d'arcs, là où on en observait sur la Figure 4.3. Néanmoins, on a rajouté de l'occlusion autour des nœuds ayant plusieurs voisins (Nœuds C par exemple).

4.3.3 Les liens de type « Courbe de Bézier »

La dernière solution implémentée est d'utiliser des courbes de Bézier, afin de garder les avantages des lignes brisées, tout en évitant les occlusions. On utilise les mêmes points que pour les lignes brisées comme points de contrôle pour la courbe.

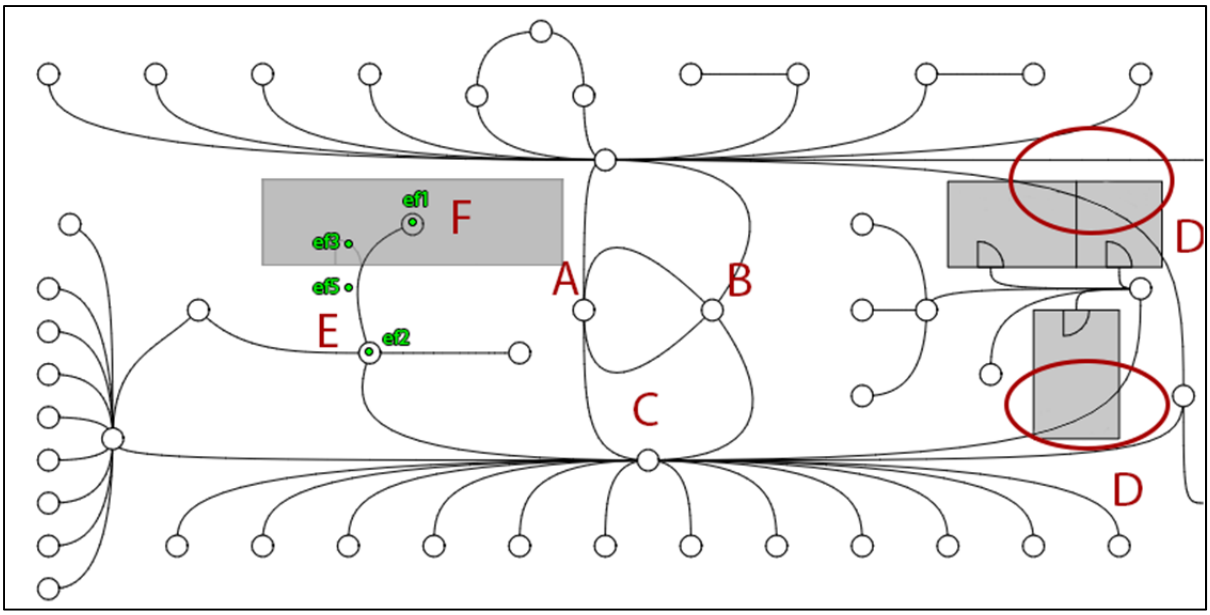

Figure 4.5 Mode topologique avec des courbes de Bézier

L'un des avantages des lignes brisées est ici conservé. En effet, les liens multiples entre A et B sont bien différenciés. Contrairement aux lignes brisées, il n'y a pas cette fois d'occlusion totale pour les liens avec de multiples arrêtes (Nœud C). Néanmoins, on observe à nouveau le phénomène de croisement d'arcs, notamment aux zones délimitées, D.

Afin d'éviter les croisements d'arcs et les intersections avec les pièces en mode géospatiale, on a implémenté les courbes de Bézier relationnelles. L'équation générale d'une courbe de Bézier passant proche des points de contrôles P_i est

$$B(t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} \; P_i \tag{4.1}$$

où t varie entre 0 et 1.

Dans notre cas, on va générer notre courbe à partir de 4 points de contrôles. On a alors l'équation suivante:

$$B(t) = \sum_{i=0}^{3} \binom{3}{i} t^i (1-t)^{3-i} \; P_i \tag{4.2}$$

Et chaque point est donc défini par :

$$P(t) = P_0(1-t)^3 + 3P_1 t(1-t)^2 + \; 3P_2 t^2(1-t) + P_3 t^3 \tag{4.3}$$

La courbe de Bézier relationnelle est obtenue à partir de la même fonction, à la différence que l'on multiplie chaque point par un facteur w_i. Plus le facteur sera élevé, et plus la courbe sera proche des points de contrôle.

L'équation générale d'une courbe de Bézier relationnelle est la suivante :

$$B(t) = \frac{\sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} \, P_i w_i}{\sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} \, w_i} \tag{4.4}$$

Ou dans notre cas, avec 4 points, et donc $n = 3$, nous posons $w_o = w_3 = 1$ et $w_1 = w_2 = w_{12}$:

$$P(t) = \frac{P_0(1-t)^3 + 3w_{12}P_1 t(1-t)^2 + 3w_{12}P_2 t^2(1-t) + P_3 t^3}{(1-t)^3 + 3w_{12} t(1-t)^2 + 3w_{12} t^2(1-t) + t^3} \tag{4.5}$$

Ci-dessus, on peut voir l'effet des différents facteurs sur les courbes de Bézier relationnelle.

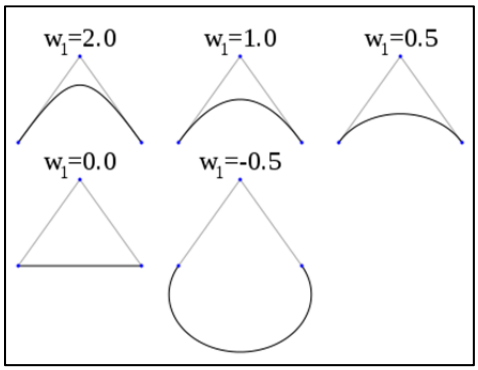

Figure 4.6 Variation du facteur et effet sur les courbes de Bézier (Tirée de Wikipedia.org)

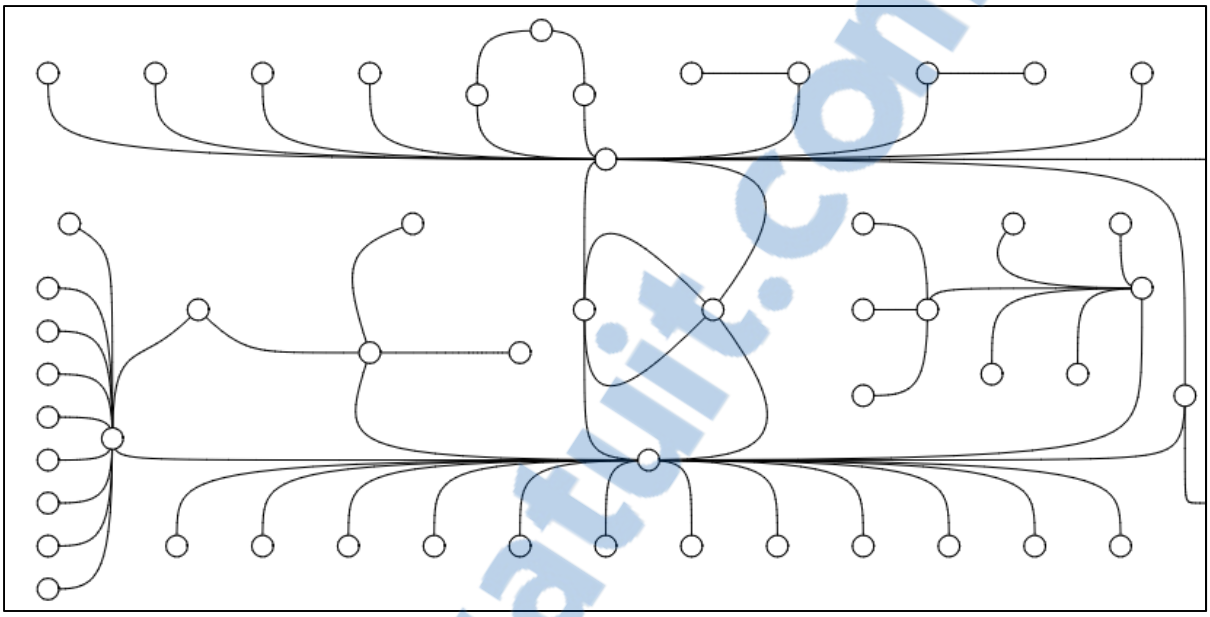

Figure 4.7 Courbe de Bézier relationnelle avec un facteur de 10

Avec un facteur élevé, de même que pour les liens en forme de lignes brisées, on n'observe pas de croisement d'arcs, et on peut voir distinctement les liens multiples. Néanmoins, on observe le même phénomène d'occlusion que pour les lignes brisées lorsqu'un nœud est de degré élevé.

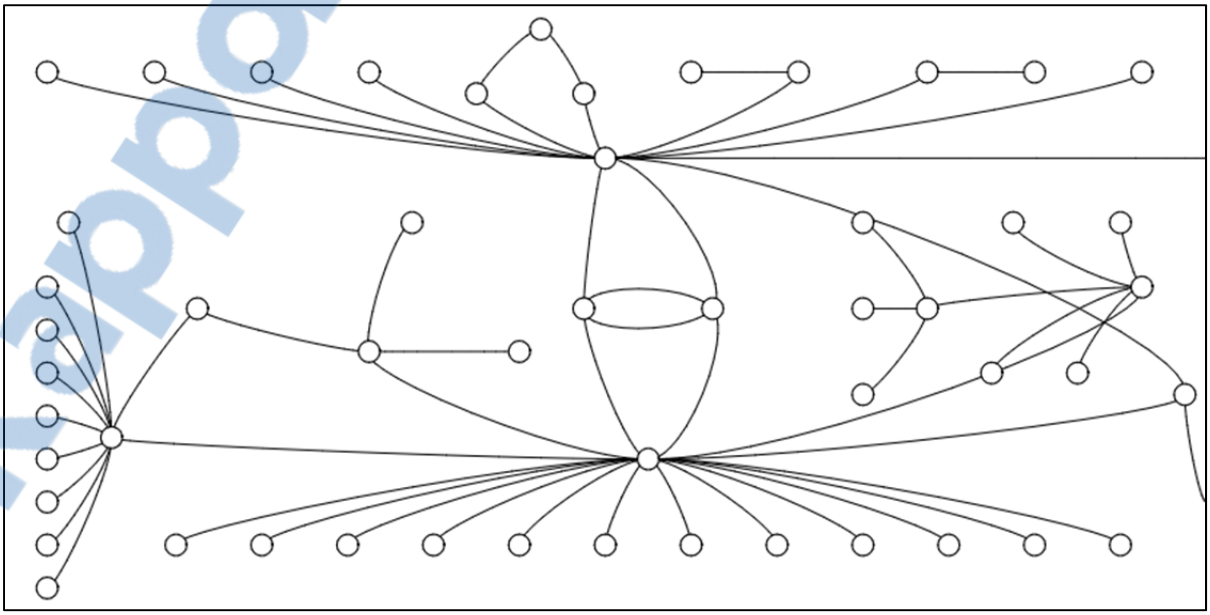

Figure 4.8 Courbe de Bézier relationnelle avec un facteur de 0.1

Rapport-Gratuit.com

En prenant un facteur plus faible, on différencie toujours les liens multiples, mais on observe cette fois à nouveau des croisements d'arcs, alors que les informations sont toujours visibles pour les nœuds à haut degré.

4.3.4 Comparatif des différents types de liens

Les différents types de liens sont ici évalués selon trois critères :

- La présence de croisement, entre les liens, ou entre les liens et les pièces.
- La possibilité de voir distinctement les liens multiples entre deux nœuds
- Le niveau d'occlusion autour des nœuds ayant un degré élevé.

Tableau 3 Comparatif des différents types de liens selon trois critères (Présence de croisements, visibilité des liens multiples, occlusion)

	Croisements	**Liens multiples distincts**	**Occlusion autour des nœuds à haut degré**
Liens droits	Oui	Non	Limité
Lignes brisées	Non	Oui	Très élevée
Courbes de Bézier	Oui	Oui	Élevée
Courbes de Bézier fortement pondérées (facteur 10)	Non	Oui	Très élevée
Courbes de Bézier légèrement pondérées (facteur 0.1)	Oui	Oui	Limitée

Dans le cas du prototype, la possibilité de voir les liens multiples est une fonctionnalité essentielle (Pour la compréhension de la topologie du plan), les liens droits n'ont donc pas été retenus.

Parmi les autres critères, la présence d'occlusion a été jugée plus critique que la présence de croisement, car gênant plus la visibilité. Le type de lien choisi pour le prototype a donc été les courbes de Bézier avec une légère pondération, car ils produisent le moins d'occlusion.

4.4 Les arêtes d'adjacences

Il a tout d'abord été envisagé de dessiner les arêtes d'adjacences. Ainsi, celles-ci étaient représentées sous forme de lignes droites en pointillé.

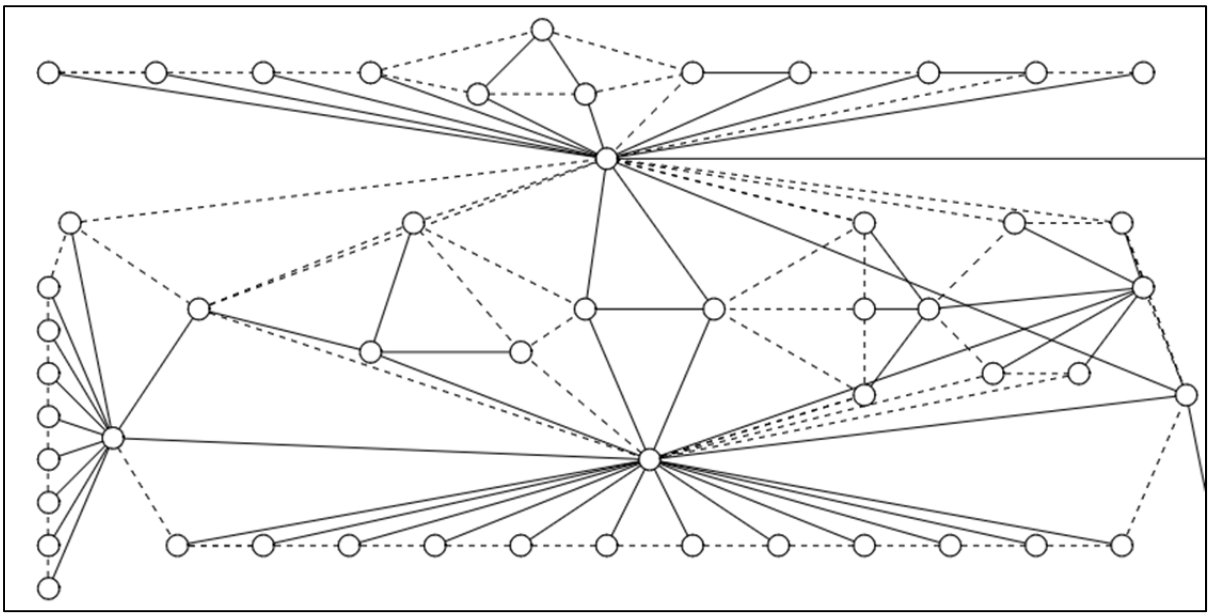

Figure 4.9 Vue topologique avec arrêtes d'adjacences (montrées par des pointillés)

Ce type d'affichage a dans un premier temps été implémenté dans l'hypothèse où un algorithme d'agencement serait appliqué au graphe nœuds-liens. En effet, dans ce cas-ci, la position physique des pièces aurait été altérée, et les liens d'adjacences auraient permis de diminuer la confusion liée à cette altération, en donnant une information sur la position relative des pièces, une à une.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

Néanmoins, l'algorithme d'agencement n'a pas été implémenté, et l'affichage des arrêtes d'adjacence augmentant la quantité d'information à visualiser, et donc le phénomène d'occlusion et les croisements d'arcs, cet aspect n'a pas été maintenu.

4.5 Les trajectoires

Tout d'abord, une trajectoire est composée d'une série de points à trois dimensions, les coordonnées x et y pour la position, et une composante temporelle.

Plusieurs manières de dessiner les trajectoires sont à distinguer : Les trajectoires en mode géospatial, et les trajectoires en mode topologique, où on différenciera les trajectoires au sein des pièces de celles des transitions.

Dans le prototype, chaque trajectoire va être découpée en sous-trajectoires, de manière à ce que chaque sous-trajectoire corresponde à l'ensemble des points entre deux passages successifs entre deux portes (pouvant être la même porte). Lorsqu'une trajectoire sera affichée, chaque sous-trajectoire sera donc dessinée selon l'état de la pièce la contenant.

4.5.1 Dessin des trajectoires en mode géospatial

En mode géospatial, le détail de chaque sous-trajectoire est affiché. Ainsi, notre trajectoire étant composée de points à deux dimensions pour la composante spatiale, elles seront dessinées comme une ligne brisée passant par chacun des points définissant la sous trajectoire.

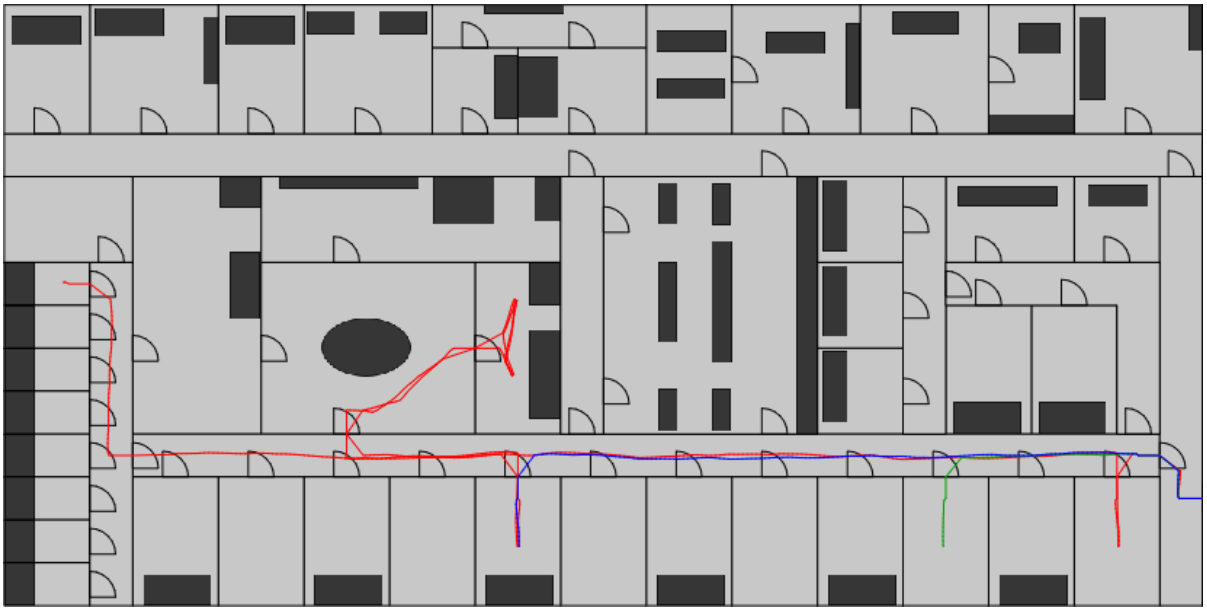

Figure 4.10 Affichage des trajectoires en mode géospatial

4.5.2 Dessin des trajectoires en mode topologique

Comme mentionné précédemment, en mode topologique, l'emphase n'est plus mise sur les trajectoires à l'intérieur des pièces mais dans ce type de visualisation, on essaye de montrer l'aspect général de la trajectoire, et plus particulièrement :

- L'ordre dans lequel une cible a traversé les différentes pièces.
- L'ordonnancement relativement aux autres trajectoires (i.e., si une sous-trajectoire se situe temporellement avant ou après une autre sous-trajectoire).

La manière de dessiner les trajectoires est inspirée de la circulation routière. Ainsi, si l'on considère les arêtes comme des segments de route, et les nœuds comme des ronds-points, les trajectoires suivront ces routes dans le sens contraire des aiguilles d'une montre.

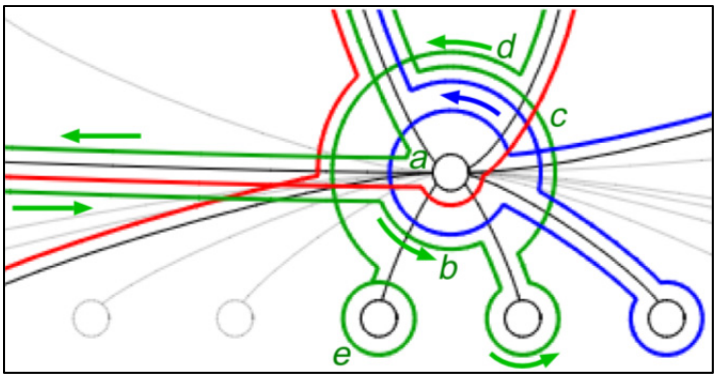

Figure 4.11 Les trajectoires avec la technique de visualisation topologique

4.5.2.1 Trajectoires autour d'un nœud

À la manière d'un rond-point, pour lequel les arêtes représenteraient les différentes routes, les trajectoires contournent ici le nœud dans le sens inverse des aiguilles d'une montre.

Les arcs de cercle sont dessinés selon l'équation paramétrique d'un cercle qui est la suivante :

$$P(t) = \big(x(t), y(t)\big)\ avec \begin{cases} x(t) = a + r\cos t \\ y(t) = b + r\sin t \end{cases} \tag{4.6}$$

Avec a et b respectivement les ordonnées et abscisses de C, centre du cercle, et r son rayon. Dans le cas d'un cercle complet, on a: $t \in [0, 2\pi[$. Pour un arc de cercle, rejoignant P_1 et P_2 deux points du cercle, on aura donc $t \in [a_1, a_2[$, avec a_1 angle du vecteur $\overrightarrow{CP_1}$ et a_2 angle du vecteur $\overrightarrow{CP_2}$.

Pour les sous-trajectoires, les points P_1 et P_2sont respectivement définis comme les points d'intersection entre le cercle et la sous-trajectoire de transition précédente, et entre le cercle et la sous-trajectoire de transition suivante.

Ces arcs de cercles concentriques sont ainsi dessinés de manière chronologique. Si un arc de cercle, représentant une sous-trajectoire, a un rayon plus petit qu'un second quart de cercle, alors le premier est rentré dans la pièce avant le second. Ceci n'apporte néanmoins pas d'indication sur la durée passée dans la pièce, ni sur le moment de sortie de la pièce. Pour pallier à ce problème, une autre solution a été envisagée, qui consistait à dessiner les arcs de cercles avec un rayon de début, et un rayon de fin différents, permettant de voir l'ordre de sortie des cibles (Figure 4.12).

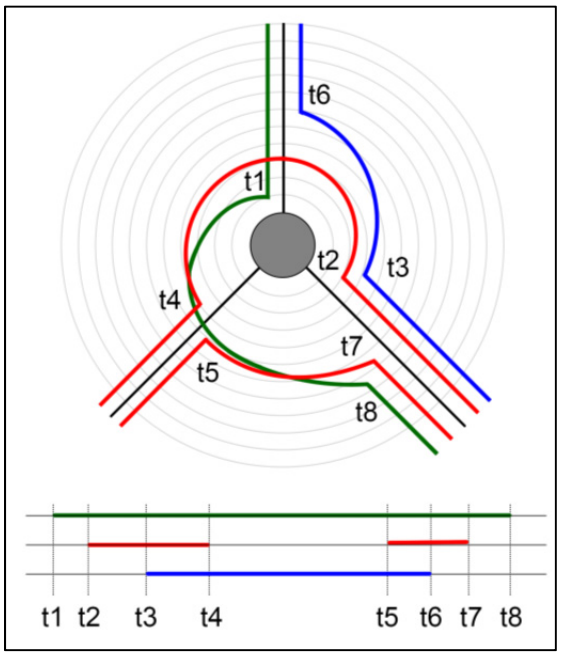

Figure 4.12 Manière alternative de dessiner les trajectoires autour d'un nœud

Dans l'exemple ci-dessus (Figure 4.12), la trajectoire verte est la première à rentrer dans la salle (t_1), et la dernière à en sortir (t_8), ainsi, l'arc de cercle sera interpolé entre un arc de cercle de rayon 1 et de rayon 8. La cible rouge rentre ensuite (t_2), suivie de la bleue (t_3), puis en sort (t_4), re-rentre dans la pièce plus tard (t_5), avant que la bleue n'en sorte (t_6) puis la rouge (t_7). Il est donc possible avec cette méthode de voir les ordres d'entrée et de sortie des cibles dans une pièce, néanmoins, les croisements des trajectoires peuvent être confondants. Le choix a donc été fait de ne pas utiliser cette méthode.

Pour déterminer l'espacement maximum entre deux cercles concentriques, on commence par calculer le rayon maximal utilisable comme la distance minimale entre le barycentre de la pièce et chacun des murs de la pièce, que l'on divise ensuite par le nombre de sous-trajectoires passant par cette pièce. On calcule ensuite l'espacement entre deux cercles concentriques. Soit e l'espacement entre deux sous-trajectoires successives, r_m le rayon maximal, et n le nombre de sous-trajectoires propre à la pièce, on obtient $e = \frac{r_m}{n+1}$

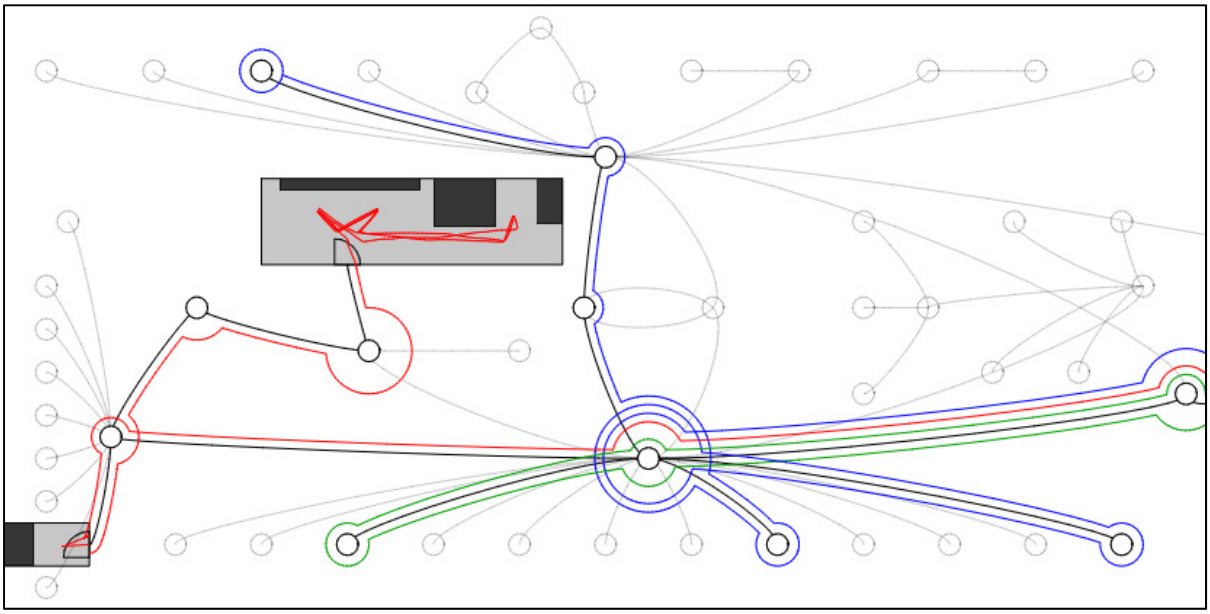

Figure 4.13 Affichage des trajectoires avec le mode de visualisation hybride

4.5.2.2 Trajectoires de transition entre deux nœuds

Les transitions, qui sont ponctuelles en mode géospatial, ont en mode topologique un aspect particulier, en effet, elles sont ponctuelles temporellement (elles correspondent aux passages d'une porte) mais ne le sont pas géométriquement (elles correspondent aux liens entre deux pièces)

Pour le dessin des sous-trajectoires correspondant aux transitions, on utilise les mêmes points que ceux générés pour le dessin des liens, auxquels on applique un *offset*. Dans le cas où l'on utilise des lignes droites pour représenter les liens, une même translation appliquée à chaque point suffit à obtenir la courbe parallèle. Pour les courbes de Bézier, appliquer une unique translation ne suffit pas, il faut, pour chaque point, calculer la translation en prenant en compte le point précédent et suivant.

Une courbe à laquelle on a appliqué un offset est définie par l'ensemble des points qui sont situés à une distance perpendiculaire de chacun des points de la courbe de base. Pour une courbe ayant pour équation paramétrique $P(t) = (x(t), y(t))$

La courbe avec offset est alors définie par :

$$\Omega\big(\rho, P(t)\big) = P(t) + \rho \times \frac{\big(y'(t), -x'(t)\big)}{\sqrt{x^2(t) + y^2(t)}} \tag{4.7}$$

La valeur ρ est constante pour l'ensemble des trajectoires. Avec une valeur positive pour ρ, les courbes seront décalées « vers la droite », ce qui correspond à la métaphore de la circulation routière.

Dans le prototype, les trajectoires sont dessinées de manière chronologique. Ainsi, pour deux sous trajectoires T_1 et T_2 successives, si T_1 est plus proche du lien que T_2, alors la transition correspondante à T_1 sera chronologiquement avant la transition correspondante à T_2. Ceci n'étant valable que pour deux transitions ayant eu lieu dans le même sens.

4.5.3 Exemple concret

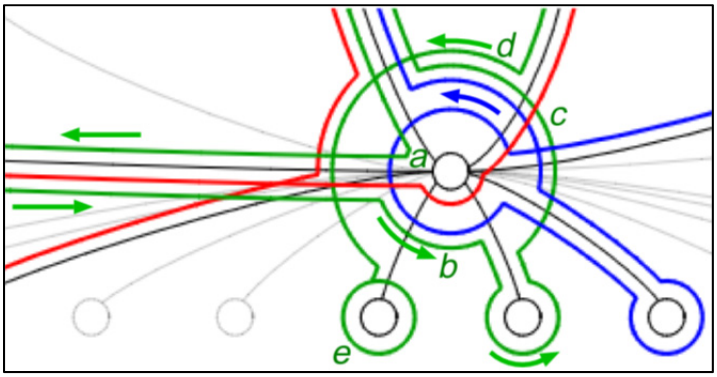

Figure 4.14 Exemple de représentation d'une trajectoire en mode topologique

La Figure 4.14 ci-dessus montre un exemple de dessin de trajectoire en mode topologique. En s'appuyant sur la métaphore décrite précédemment, il est possible d'observer que la trajectoire verte accède d'abord à la salle (*a*), en sort vers la gauche, y repasse (*b*), en sort vers le haut (*c*), y revient (*d*), avant de terminer dans la salle notée *e*. Cette représentation permet donc de voir l'ordre de déplacement de la cible, et ce de manière statique.

CHAPITRE 5

GÉNÉRATION DES DONNÉES

Ne bénéficiant pas de données réelles, il a fallu générer des données de la manière la plus réaliste possible. Pour ce faire, une méthode de génération semi aléatoire a été développée, en se basant sur des techniques utilisées, par exemple, en robotique. Un sous graphe est généré pour chaque pièce, incluant les nœuds de passages et les nœuds d'arrêts possibles pour contrôler la génération des trajectoires (en excluant par exemple le passage dans des zones qui correspondraient à des meubles).

Dans un premier temps, la méthode aléatoire sera décrite, elle-même utilisée en partie pour la méthode de génération semi-aléatoire. Ensuite, le principe de bruit des données sera évoqué, avant de présenter cette méthode semi aléatoire utilisée dans le cadre d'un autre projet de recherche.

Lors de la génération des données, certaines contraintes spatiales et temporelles, que l'on nommera « nœud de contrainte » sont utilisées et sont définies de la manière suivante :

- Un nœud interne pour la composante spatiale
- Une durée d'arrêt, pour la composante temporelle

5.1 Génération aléatoire

La première méthode a été de générer les trajectoires de manière totalement aléatoire Une liste de contraintes était définie:

- Un nœud de contrainte de départ
- Un nœud de contrainte de fin
- Des nœuds de contrainte intermédiaires

Rapport-Gratuit.com

Pour chacune des contraintes, la salle était choisie aléatoirement, en vérifiant qu'elle était différente de la précédente. Les intervalles de temps étaient eux aussi choisis aléatoirement, en s'assurant une durée minimum entre deux intervalles de temps successifs.

Tout d'abord, un algorithme de plus court chemin était utilisé afin de répondre à toutes les contraintes spatiales. La seconde étape était de générer des données aléatoires au sein d'une pièce grâce à un algorithme de génération de « comportement directionnel » (*steering behavior*) (Reynolds, 1999), de type « flâneur » (*wandering*).

Le principe de cet algorithme est de générer les points successifs aléatoirement en fixant certains paramètres:

- Le cercle A détermine l'espacement maximal successif entre deux points.
- L'arc de cercle formé par les points b1 et b2 sur le cercle B détermine l'angle maximum sur lequel peut se déplacer le cercle A.

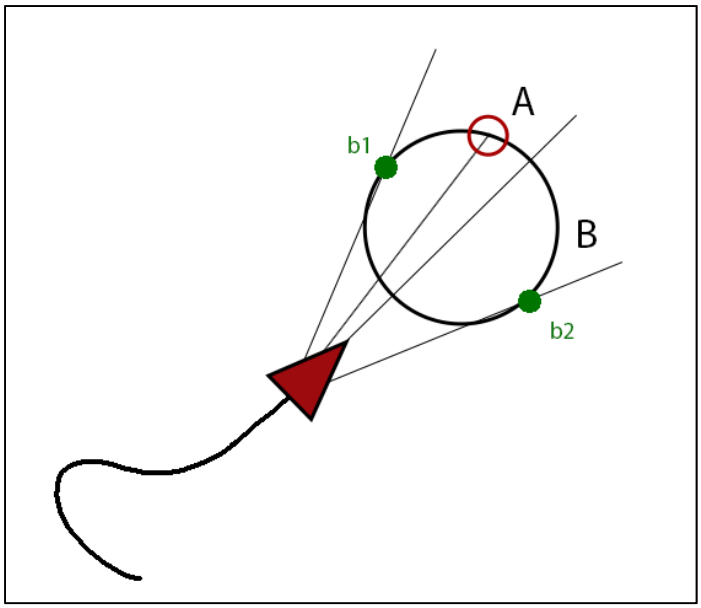

Figure 5.1 Schéma explicatif du *steering behavior*

Ainsi, en variant les rayons des cercles A et B l'aspect des courbes générées va varier :

- Plus le rayon du cercle A sera petit, plus l'aspect de la courbe sera doux, avec des petits virages. Au contraire, si le rayon augmente, les virages entre deux points seront plus marqués.
- La variation du diamètre du cercle B va quant à elle modifier l'amplitude des virages de la courbe de manière plus globale. Ainsi, plus le diamètre sera réduit, plus la courbe aura une allure rectiligne, alors que s'il augmente, la courbe effectuera des virages plus amples.

L'algorithme général vient d'être décrit ici. Deux autres aspects doivent être pris en compte.

Tout d'abord, les points générés doivent rester au sein d'une pièce. Pour ce faire, la méthode du calcul d'intersection est utilisée dans le processus de validation.

Soit P_0 le dernier point généré et validé, et P_1 le point à valider. $v_{init} = \overline{P_0P_1}$ est le vecteur généré par ces deux points. Celui-ci est ensuite normalisé (v_{norm}), et multiplié par un facteur k arbitrairement choisi, de manière à obtenir $v = k \times v_{norm}$. Dans l'exemple à la page suivante, le facteur est de 10. Plus ce facteur sera élevé (respectivement faible) plus la distance entre chaque point de la trajectoire et chacun des murs de la pièce sera élevée (respectivement faible). L'intersection entre le rayon v et chaque segment correspondant aux murs est ensuite calculée, et le point P_1est validé si et seulement si il n'y a aucune intersection.

Le second aspect à prendre en compte est que les trajectoires doivent passer aux travers des portes. Ainsi, sur la fin de l'intervalle de temps, un vecteur force est ajouté au comportement, afin de diriger la trajectoire vers le centre de la porte. Ce principe est plus précisément décrit dans le paragraphe suivant, puisqu'il est à la base de la génération des données semi-aléatoires.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

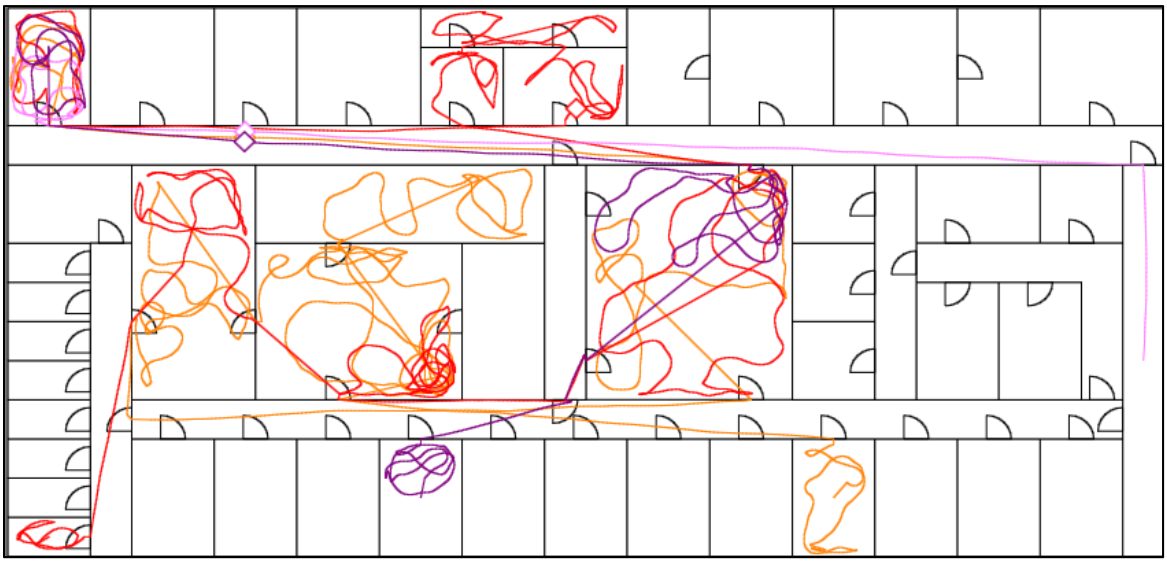

Figure 5.2 Quatre trajectoires générées totalement aléatoirement avec un comportement de type *wandering behaviour*

5.2 Génération semi aléatoire: Principe de graphe à deux niveaux

Dans le domaine de l'intelligence artificielle, appliquée à la robotique par exemple, les comportements de type « humain » sont généralement modélisés à l'aide de graphe. En observant le comportement humain, certaines règles dans les déplacements peuvent être notées :

- Une personne ne reste immobile qu'à certains endroits (Une chaise, un bureau, une table …)
- Certains endroits seront des « points de passage » (Par exemple, dans un laboratoire, un couloir naturel est formé par les bureaux), où les cibles, en général, ne resteront pas statiques.
- Certains endroits ne sont jamais visités : Les endroits où l'on trouve des meubles, ou les coins des pièces par exemple

De manière plus générale, les trajectoires peuvent être décrites de la manière suivante: Sur un intervalle de temps, une cible va se rendre dans plusieurs pièces, et dans chacune de ces pièces, va rester un certain nombre de fois immobile pour des durées plus ou moins longues.

La première étape est donc de définir ces différents points de passage et points d'arrêt. Pour chaque pièce, le graphe interne est le graphe défini par ces points.

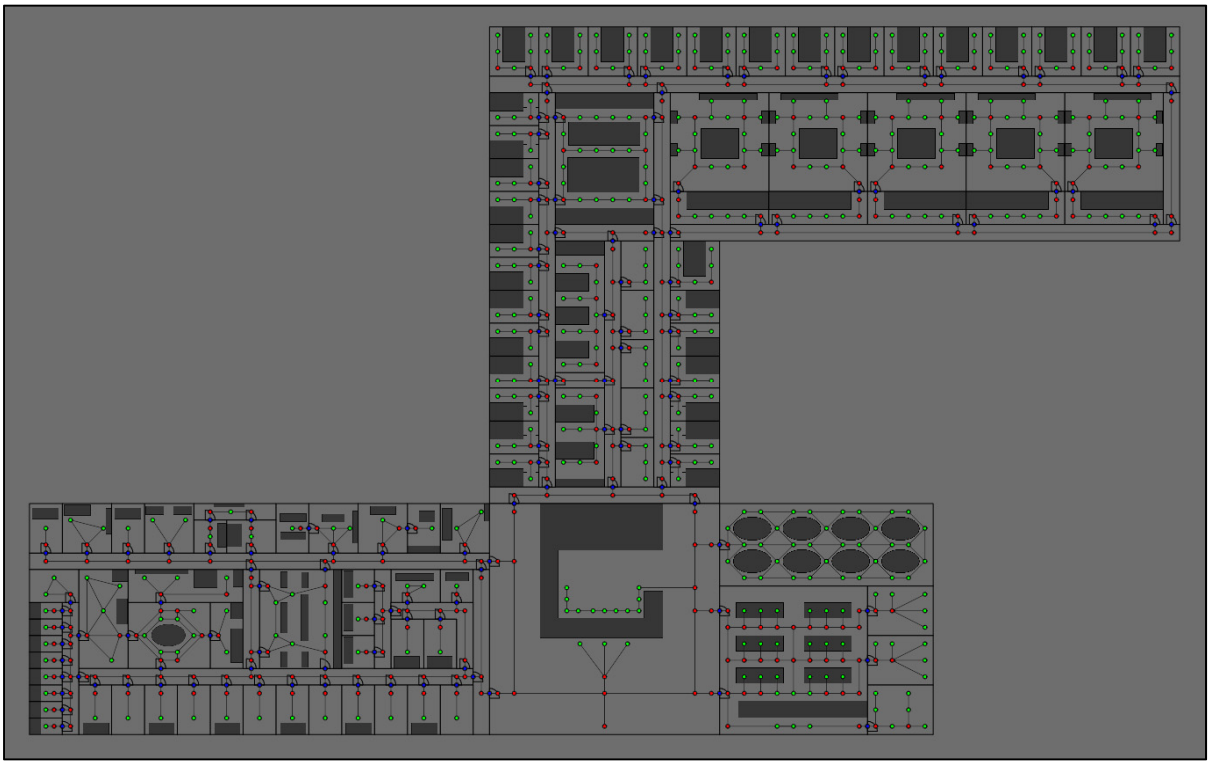

Figure 5.3 Superposition du plan avec le graphe interne

Pour réaliser le graphe interne, il a fallu imaginer des patrons pour chaque pièce, en leur donnant à chacune une utilité, et en représentant leur agencement. Sur la Figure 5.3, on peut voir une superposition du plan avec chaque graphe interne. Les points verts représentent les points d'arrêt, les points rouges représentent les points de passage, et les points bleus représentent les portes.

La seconde étape est de générer les contraintes qui seront appliquées :

Pour chaque trajectoire t :

Soit I un nombre de pièces aléatoires ($1 < I <= 10$)

Soit R l'ensemble des I pièces tirées aléatoirement, et R_i chacune de ces pièces (avec $0 <= i < I$). On vérifie que deux pièces successivement tirées sont toujours différentes.

Pour chaque pièce R_i

Soit N, un nombre de points d'arrêts aléatoires ($0 < N <= 5$). S'il n'y a qu'un point d'arrêt dans une pièce, on a $N = 1$.

Soit P l'ensemble des N points tirés aléatoirement, et P_n chacun de ces points (avec $0 <= n$). On vérifie que deux points successivement tirés sont toujours différents.

Pour chaque point P_n

Soit T, une durée d'arrêt ($15 <= T <= 35$)

On ajoute aux contraintes le point d'arrêt défini par les composantes:

- Spatiales, avec le point d'arrêt P_n et la pièce R_i
- Temporelle, avec la durée T

Algorithme 1: Pseudo algorithme de génération de contraintes

Le plus court chemin passant par chacun des nœuds de contrainte est alors calculé, en utilisant l'algorithme de Dijkstra.

5.3 Bruitage des données

Les données générées seraient actuellement rectilignes. Dans la pratique, deux types de bruit peuvent survenir :

- Un bruit dû aux imprécisions des systèmes de traçage
- Un bruit dû au fait que les individus ne se déplacent pas de manière totalement rectiligne.

Afin d'avoir ce même type de bruit, on applique une méthode proche de celle vue précédemment, à la différence que cette fois, afin de diriger la trajectoire vers le prochain point de passage, on calcule le point intermédiaire suivant en utilisant la méthode de *wandering-behaviour*, puis, on lui additionne un vecteur $v_{d,}$ avant de normaliser le résultat de manière à avoir des points intermédiaires espacés régulièrement.

Le vecteur v_d est égal au vecteur $\overrightarrow{AB}$ (Avec A le point intermédiaire précédent, et B le point d'arrêt destination), normalisé, puis multiplié par un facteur (Dans le cas du prototype, le facteur utilisé était 2).

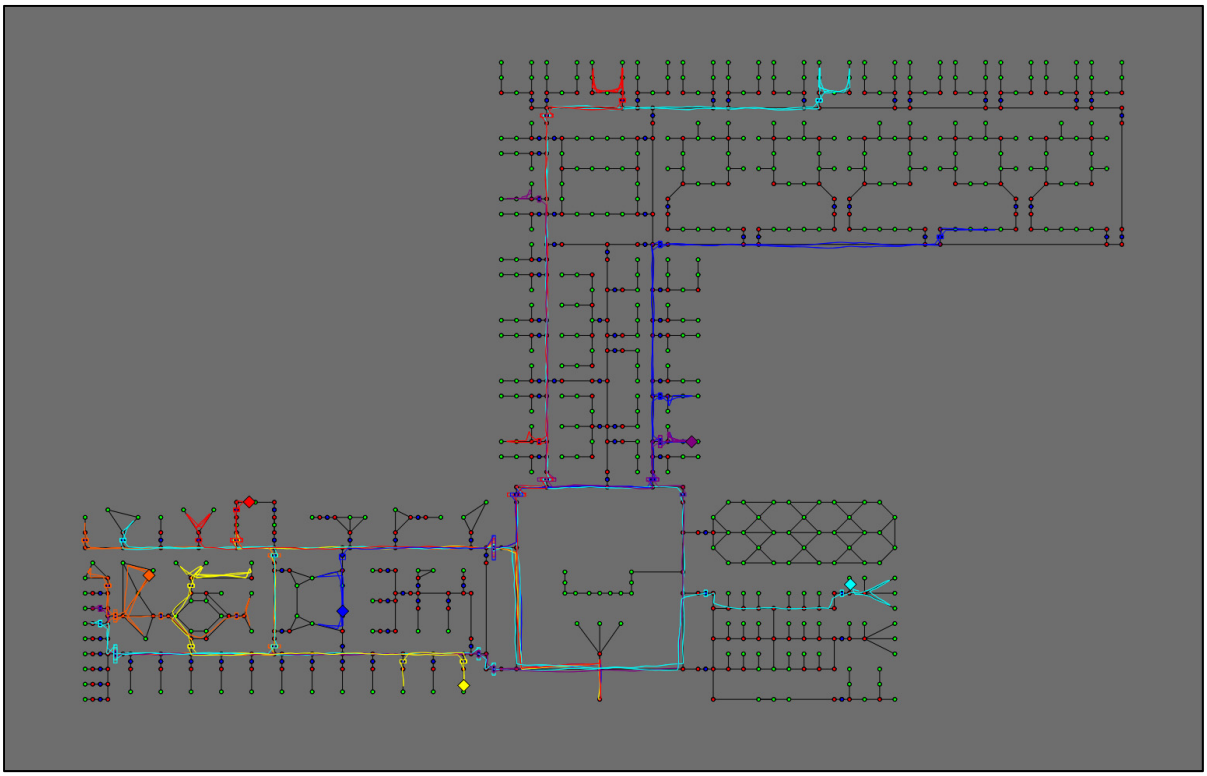

Figure 5.4 Exemple de trajectoires générées par la méthode expliquée ci-dessus, en superposition du graphe interne

5.4 Application de la méthode pour générer des données urbaines

L'un des projets du laboratoire nécessitait d'obtenir des données de déplacement d'objet, mais cette fois-ci en milieu urbain. Pour ce projet, le focus était mis sur des cibles ayant des rendez-vous régulier.

Comme dans les paragraphes précédents, on utilise, pour générer les données, un graphe nœuds-liens. Pour le choix de la ville, deux critères ont été importants:

- Les données générées doivent être utilisées dans une expérience contrôlée effectuée par une équipe de l'Université de Manitoba. Pour éviter les confusions liées à la langue, la ville doit être anglophone et les noms des rues en anglais.
- Les villes modernes comme New York et Montréal sont construites autour d'axes horizontaux et verticaux. Afin d'éviter d'obtenir un graphe de type quadrillage, le choix a été fait sur une ville « ancienne ».

En suivant ces critères, la ville de Dublin (Irlande) a été choisie pour générer les données.

Figure 5.5 Graphe nœuds-liens généré à partir d'un plan de Dublin

Le graphe nœuds-liens généré à partir du plan de Dublin contient un total de 2695 nœuds et 3464 liens.

Dans un premier temps, les nœuds correspondaient aux intersections. Néanmoins, en procédant ainsi, les données générées ne suivaient pas la courbe des routes. Dans un second temps, des points intermédiaires ont donc été rajoutés afin cette fois d'avoir des arrêtes entre deux nœuds qui suivaient les routes. Les trajectoires générées étaient ainsi proches de trajectoires réelles.

Figure 5.6 Détails d'une partie du plan de Dublin

Les données qui devaient être générées devaient montrer les trajectoires suivies par des cibles qui participaient à des rendez-vous. Certaines cibles s'y rendant, d'autre non, et certaines ayant du retard.

Le constat de base effectué pour la génération des données est que, de manière très réductrice, un individu va au cours d'une journée, se rendre dans un certain lieu le matin, un autre le midi, un autre l'après-midi (possiblement le même que le matin) un autre le soir, et

toujours le même durant la nuit. Un système de contraintes a donc été implémenté pour générer des données pseudo-aléatoires.

Les contraintes utilisées pour générer les données étaient les suivantes :

- Une durée totale en nombre de jours est définie
- De manière générale, toutes les journées sont découpées en créneaux. Par exemple :
    - 8h-10h
    - 10h30-13h
    - 14h-19h
    - 20h-21h
    - 21h30-7h (Créneau nuit)
- Plusieurs groupes de trajectoires peuvent être définies. Chaque groupe se rencontre un certain nombre de fois à un point de rendez-vous défini. Certaines trajectoires ne font partie d'aucun groupe. (Un groupe contient au moins 2 cibles)
- Chaque cible, membre d'un groupe ou non, a un nœud « maison ». La cible y reste chaque nuit (Pendant le créneau correspondant)
- Un « agenda » est défini pour chaque groupe. Un agenda correspond à une liste de points de rendez-vous, ayant lieu zéro, une, ou plusieurs fois par jour.
- Chaque cible membre d'un groupe, a son propre agenda, avec des propriétés définies : La présence ou non à chaque rendez-vous, ou le temps de retard.

On génère ensuite les données de manière à ce que chaque cible reste « immobile » à un nœud aléatoire pendant la durée d'un créneau (avec plus ou moins 10mn de retard) tout en respectant les contraintes liées aux groupes et aux agendas.

Rapport-Gratuit.com

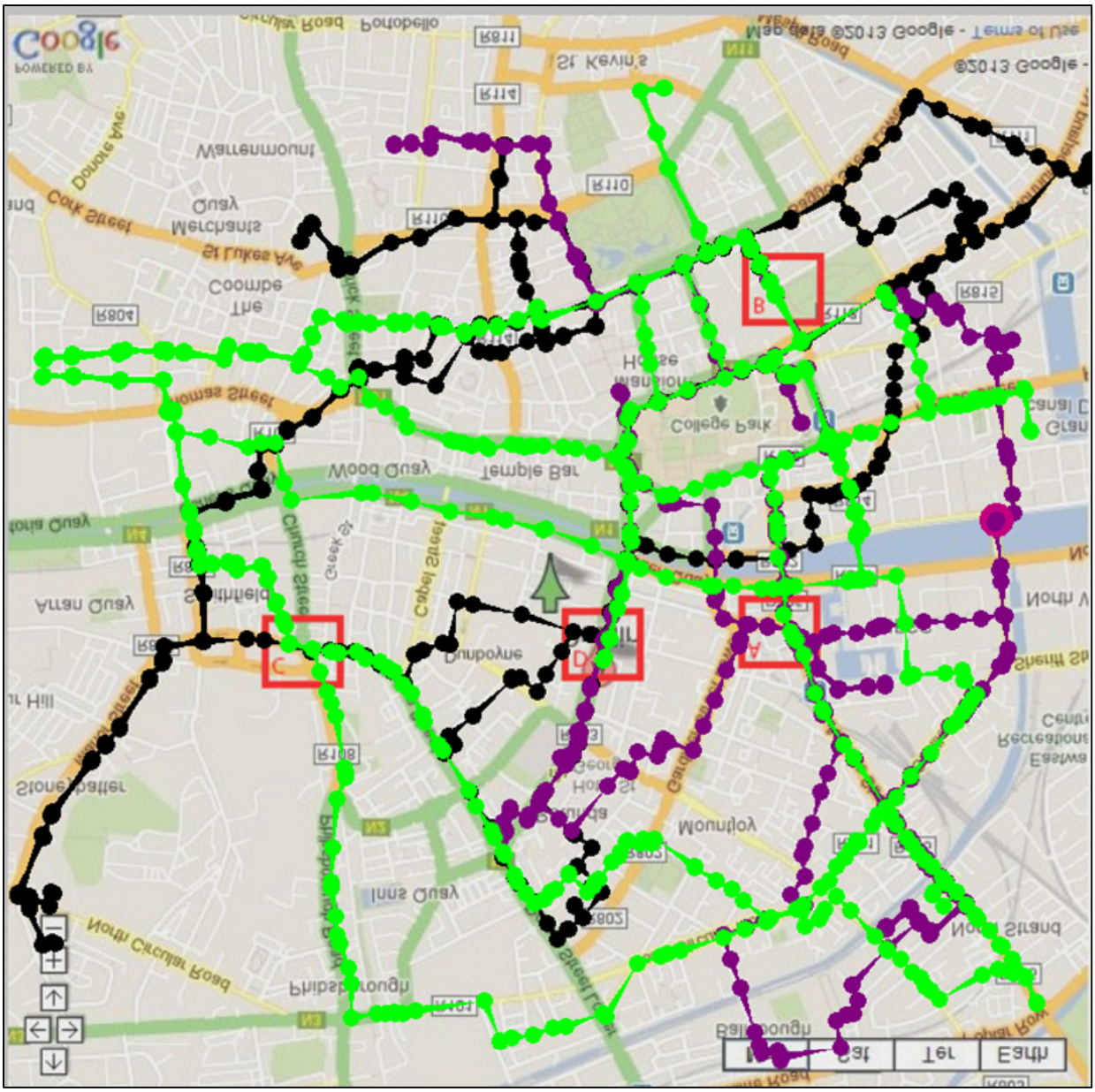

Figure 5.7 Visualisation des données générées dans le prototype "Spatio-Temporal Visualizer". Les données ci-dessus ne concernent que trois trajectoires faisant partie d'un groupe.

L'algorithme de recherche de plus court chemin aurait cette fois dû être modifié, et deux notions rajoutées :

- La notion de « sens interdit », en ajoutant des orientations aux arcs.

- La notion de « poids ». Il s'agit en fait d'un facteur qui va être appliqué à la longueur de chaque arc au moment du parcours, ainsi, une route principale aura un faible facteur, alors qu'une route secondaire un facteur un peu plus élevé. Si la longueur entre 2 chemins est sensiblement différente, l'algorithme choisit alors automatiquement celui qui va passer par une route principale.

Néanmoins, au vu de la taille du graphe généré (En terme de nombre de nœuds et de nombre d'arrêtes), la fonctionnalité n'a pas été implémentée dans les faits (Tous les arcs sont à double sens, et le poids est fixé à 1).

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

CHAPITRE 6

EXPÉRIENCE CONTRÔLÉE

6.1 Protocole expérimental

Afin de valider l'efficacité ou non de la vue topologique, une expérience contrôlée a été réalisée. Un protocole similaire à celui utilisé par Per Ola Kristensson (Kristensson et al., 2009) décrit dans la revue de littérature a été employé. Deux interfaces, ont été comparées, I_{topo} l'interface topologique, et I_{geo} l'interface géospatiale, et ce pour quatre catégories de tâches.

Deux jeux de données ont été utilisés lors de l'expérience, que l'on nommera D_1 et D_2. Chaque jeu de données était composé de 3 ensembles de trajectoires. Chacun de ces ensembles était composé de 6 trajectoires, et le participant devait effectuer les 4 tâches pour chacun d'entre eux. Avant chaque phase de l'expérience, le participant effectuait une phase d'entrainement non mesurée, pour lui permettre de se familiariser avec les interfaces et les types de contrôles. Au total, le participant devait donc effectuer 24 tâches qui étaient mesurées, et 8 tâches d'entrainement.

Afin de s'assurer que l'ordonnancement ne biaisait pas les résultats, un carré latin était appliqué. On obtenait alors les 4 ordonnancements suivants :

- $O_1 = (I_{topo}, D_1)$, (I_{geo}, D_2)
- $O_2 = (I_{topo}, D_2)$, (I_{geo}, D_1)
- $O_3 = (I_{geo}, D_1)$, (I_{topo}, D_2)
- $O_4 = (I_{geo}, D_2)$, (I_{topo}, D_1)

Il est alors possible dans un second temps, en effectuant une analyse statistique, de vérifier que l'ordonnancement n'influe pas les résultats. Pour avoir suffisamment de données pour que celles-ci soient représentatives, chaque ordonnancement va être effectué 3 fois, pour un total de 12 participants.

Les 4 tâches que les participants devaient effectuer étaient des questions à choix multiples, avec 6 choix possibles pour chaque question. Les tâches et les réponses possibles étaient les suivantes:

T1: Combien de fois la cible *A* traverse-t-elle la salle *U*?
Réponses possibles :

- 1
- 2
- 3
- 4
- 5
- 6

T2: Dans quel ordre la cible *A* a-t-elle parcouru les salles *U*, *V* et *W*?

- *U -> V -> W*
- *U -> W -> V*
- *V -> U -> W*
- *V -> W -> U*
- *W -> U -> V*
- *W -> V -> U*

T3: Quelle cible est arrivée en premier dans la salle *U*?

- *A*
- *B*
- *C*
- *D*
- *E*
- *F*

T4: Combien de cibles ont été présentes à la fois dans la salle *U* et dans la salle *V*?

- 1
- 2
- 3
- 4
- 5
- 6

6.2 Réalisation de l'expérience

L'expérience a été réalisée par huit participants, sept de sexe masculin, une de sexe féminin, âgés de 20 à 36 ans, et tous avaient une formation en ingénierie. Chacun de ces participants ont réalisé 32 tâches, dont 24 mesurées, ce qui donne donc un total de $24 \times 8 = 192$ essais mesurés.

L'interface était composée de trois parties :

- Un *slider*, pour contrôler le temps et déplacer les cibles le long des trajectoires correspondantes.
- Un panneau, où défilent les questions, et le choix de réponses.
- La vue principale.

Le fonctionnement suivant a été respecté pour chaque question de l'expérience :

- Dans un premier temps, la question était affichée sur le panneau de questions. Dans la vue principale, le plan était affiché, mais aucune trajectoire ne l'était. Enfin, les pièces importantes pour répondre aux questions étaient indiquées par un label.
- Afin de passer à la question, l'utilisateur devait cliquer sur le panneau.
- A partir de ce moment, les labels étaient partiellement cachés (L'opacité passait à 25% pour éviter l'occlusion), les trajectoires étaient affichées, et les choix de réponses étaient montrés. Le temps et les interactions étaient

alors mesurés à partir de ce moment-là, et ce jusqu'à la réponse de l'utilisateur.

- Une fois que l'utilisateur a répondu à la question, une surbrillance rouge était appliquée à sa réponse si celle-ci était fausse, et une surbrillance verte à la réponse correcte. Ceci permettait alors à l'utilisateur de regarder de nouveau les données, pour vérifier son erreur, le cas échéant.
- Entre chaque question, la vue était réinitialisée, la camera était recentrée, le zoom était mis à 0, et le *slider* remis à sa position initiale.

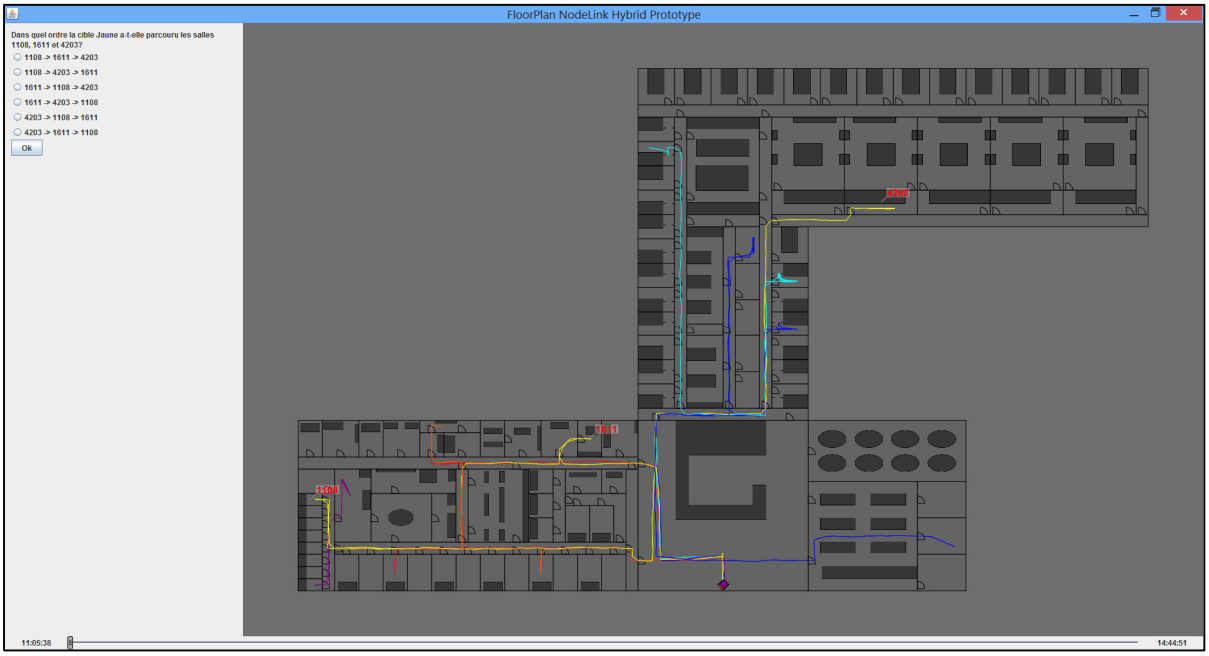

Figure 6.1 Interface utilisée pour l'expérience, avec la vue géospatiale

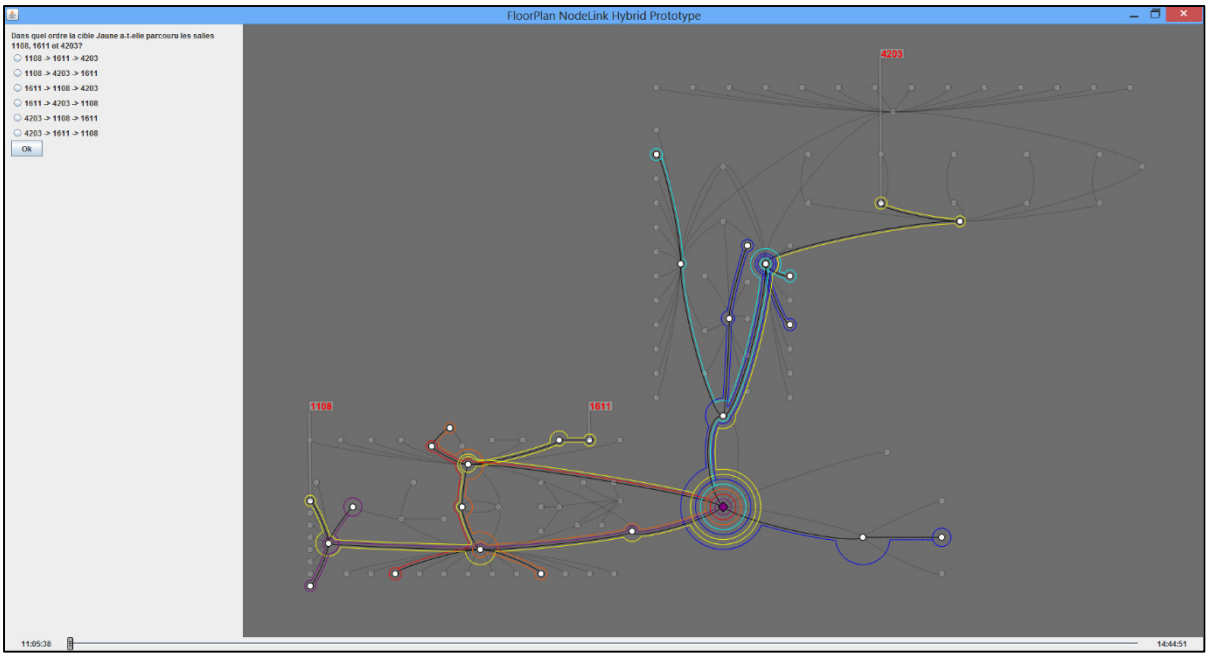

Figure 6.2 Interface utilisée pour l'expérience, avec la vue topologique

6.3 Résultats

Une fois que les huit participants ont réalisé l'évaluation, une analyse de variance (ANOVA) a été effectuée sur l'échantillon. Cette analyse permet de vérifier quels facteurs ont une influence significative sur les résultats, dans notre cas, sur le temps pris pour répondre à chaque question, ainsi que sur le taux d'erreur.

Les facteurs étaient les suivants :

- Le type d'interface
- Le type de question
- L'ordonnancement (l'ordre d'expérimentation des interfaces, ainsi que l'ordre des jeux de données)

L'analyse nous donnait les résultats ci-dessous :

Tableau 4 Résultats de l'analyse de variance ANOVA sur le temps

Effet	**Somme du carré des écarts**	**Degré de liberté d**	**F**	**p-value**
Interface	1	7	31.043	8.406e-04
Type de question	3	21	19.206	3.145e-06
Interface ×type de Question	3	21	17.335	6.761e-06

Tableau 5 Résultats de l'analyse de variance ANOVA sur le taux d'erreur

Effet	**Degré de liberté d**	**p-value**
Interface	1	0,601
Type de question	3	0,776
Interface ×type de Question	3	0,078

Les résultats des analyses nous permettent de conclure les points suivants :

- La technique de visualisation utilisée a un effet sur le temps ($F_{1,7} = 31.0, p < 0.001$)
- Le type de question à un effet sur le temps ($F_{3,21} = 19.206, p < 0.001$)
- L'ordonnancement n'a pas d'effet sur le temps ($p > 0.05$)
- Ni le type de visualisation ni le type de question n'ont d'effet sur le taux d'erreur ($p > 0.05$)

Si l'analyse nous permet d'évaluer si les conditions ont des effets sur les variables, ANOVA ne nous permet pas d'évaluer l'influence de ces variables.

Tableau 6 Taux d'erreur moyen par type d'interface et par type de question

	T1	**T2**	**T3**	**T4**
Interface géospatiale	8.33%	0%	16.67%	12.5%
Interface topologique	4.17%	12.5%	4.17%	8.33%

Le tableau ci-dessus (Tableau 6), nous montre le taux d'erreur moyen par type de question et par type d'interface. Ces taux, ne permettent pas de conclure qu'une interface est meilleure qu'une autre en termes de fiabilité des réponses.

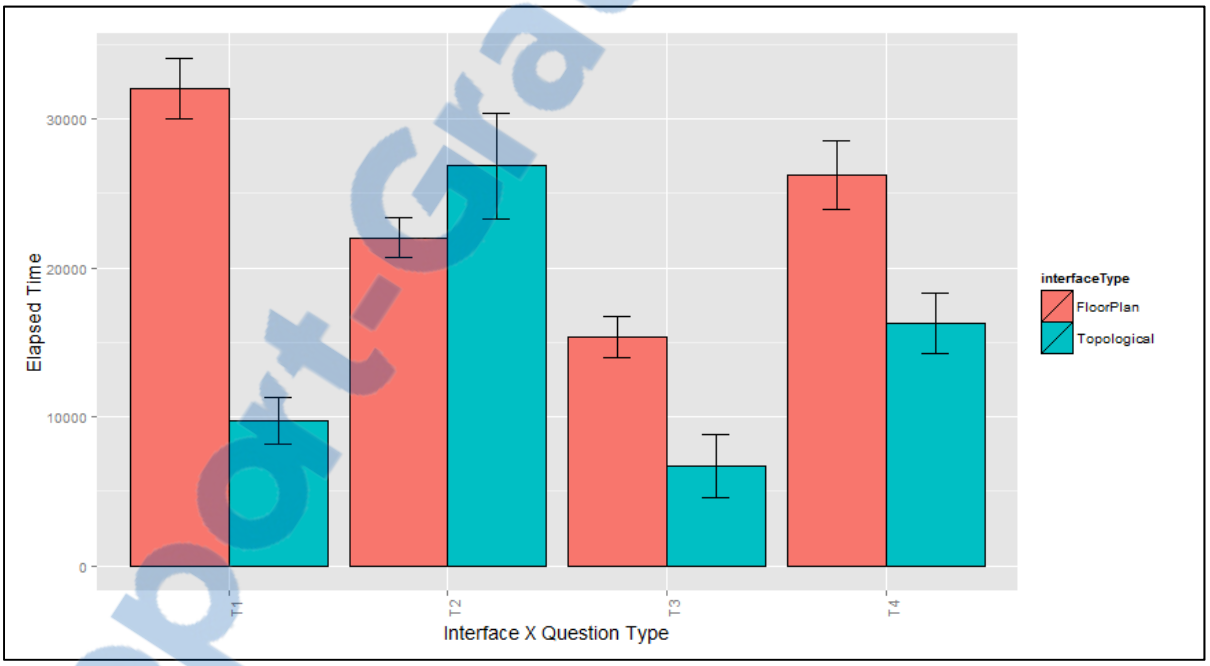

Figure 6.3 Temps écoulé moyen pour chaque type de question et chaque interface en millisecondes

Le test ANOVA (Analyse de variance) dont les résultats se trouvent ci-dessus permettent de mettre en évidence que les facteurs « technique de visualisation » et « type de question » semblent avoir un effet sur le temps écoulé. Le graphique ci-dessus (Figure 6.3) représente le temps écoulé moyen pour répondre à chaque type de question, pour chaque interface. On observe que l'interface topologique est significativement plus rapide pour répondre aux

Rapport-Gratuit.com

questions *T1*, *T3* et *T4*, alors que l'interface géospatiale est significativement plus rapide pour répondre aux questions *T2*.

Lors de l'expérience, le temps écoulé total était mesuré, mais chaque interaction était aussi chronométrée, permettant ainsi d'avoir la décomposition du temps en quatre catégories :

- Le temps passé à translater la camera
- Le temps passé à effectuer un zoom
- Le temps passé à utiliser le contrôleur de temps
- Le temps passé sans effectuer d'interaction, considéré comme le temps de réflexion

Le graphique ci-dessus (Figure 6.4) représente les données issues de cette décomposition. Les graphiques ci-dessus montrent que la plupart des utilisateurs n'ont pas effectué d'interaction avec l'interface topologique. En effet, il était possible de répondre aux questions sans effectuer de zoom, de pan, ni interagir avec le temps. Cette observation permet de conclure que l'interface topologique permet de visualiser un ensemble de données spatiotemporelles de manière statique, tout en restant compréhensible.

En ce qui concerne l'interface géospatiale, les données montrent que pour les questions de type T1, T2 ou T3, il est nécessaire d'utiliser le contrôleur de temps.

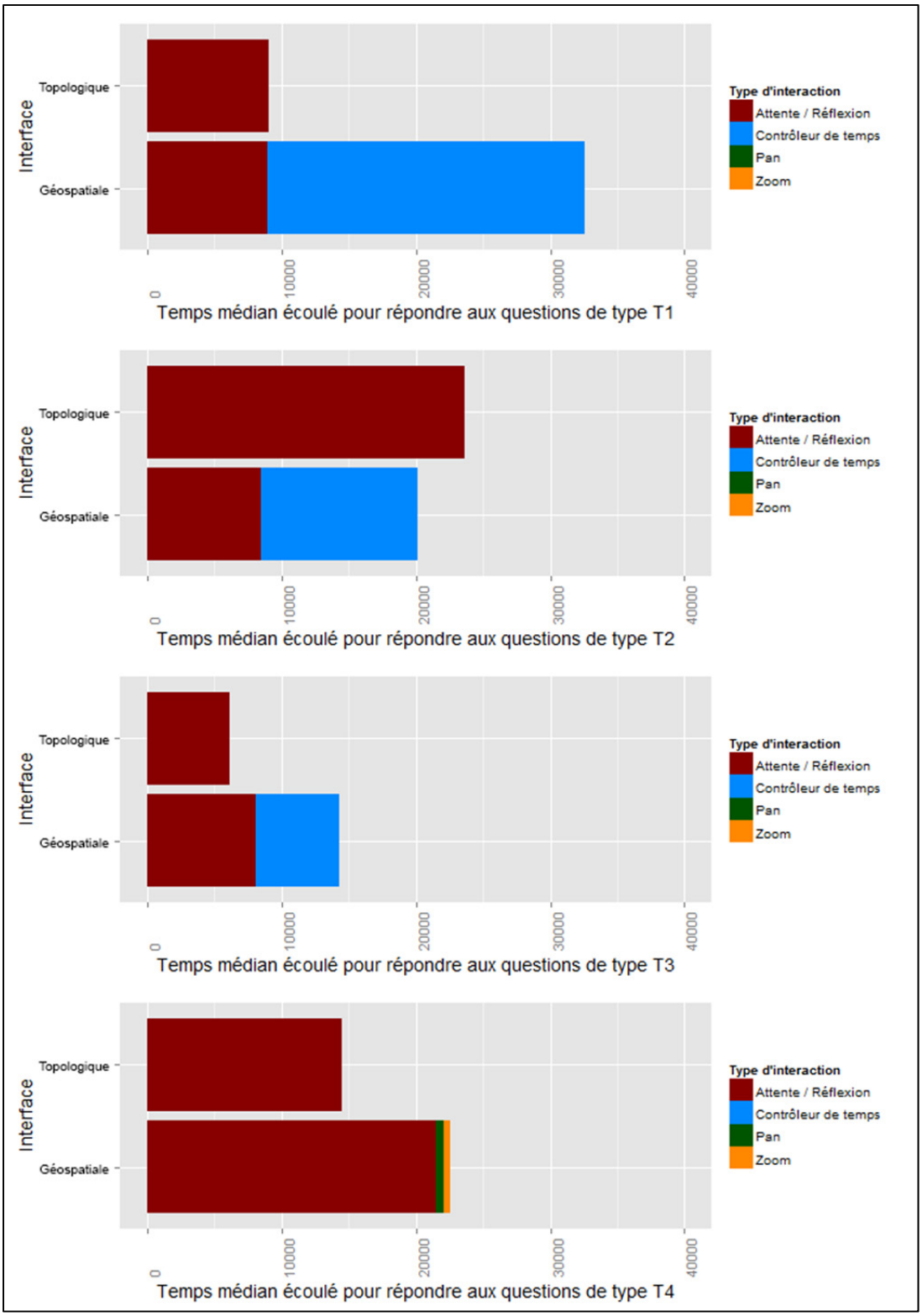

Figure 6.4 Décomposition du temps écoulé médian en millisecondes pour chaque type de question

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

CHAPITRE 7

DISCUSSIONS

Plusieurs idées ont été envisagées mais n'ont pas été implémentées. Parmi celles-ci, on va détailler l'application à des données géographiques, ainsi que la possibilité d'ajouter une étape de *clustering*. Aussi, la technique hybride présentée précédemment, qui n'a pas été évaluée par une expérience contrôlée, pourrait faire l'objet d'une seconde étude.

7.1 Application à des données géographiques

La possibilité d'étendre la méthode implémentée dans le prototype aux données géographiques a aussi été envisagée. Analogiquement, les pièces seraient remplacées par des pays (ou régions) et les liens correspondraient aux frontières.

On pourrait par exemple imaginer une utilisation pour une personne qui aurait récupéré les données GPS de son voyage. Ainsi, il pourrait visualiser les déplacements entre les pays sous forme topologique, et passer sous forme « géographique » s'il souhaite visualiser ses déplacements au sein d'un pays.

L'exemple ci-dessous nous montre le trajet d'une personne au sein de plusieurs pays de l'union européenne. La première figure (Figure 7.1) représente les données sous leur forme géographique. Un phénomène d'occlusion est présent, de par le fait qu'un voyageur est susceptible d'emprunter plusieurs fois le même chemin pour l'aller et le retour (par exemple, en empruntant des autoroutes, des ponts, des tunnels, etc.). Ainsi, certains points sont confus (Entre la France – FR, et la Suisse - CH), et ne permettent pas de suivre aisément le trajet emprunté.

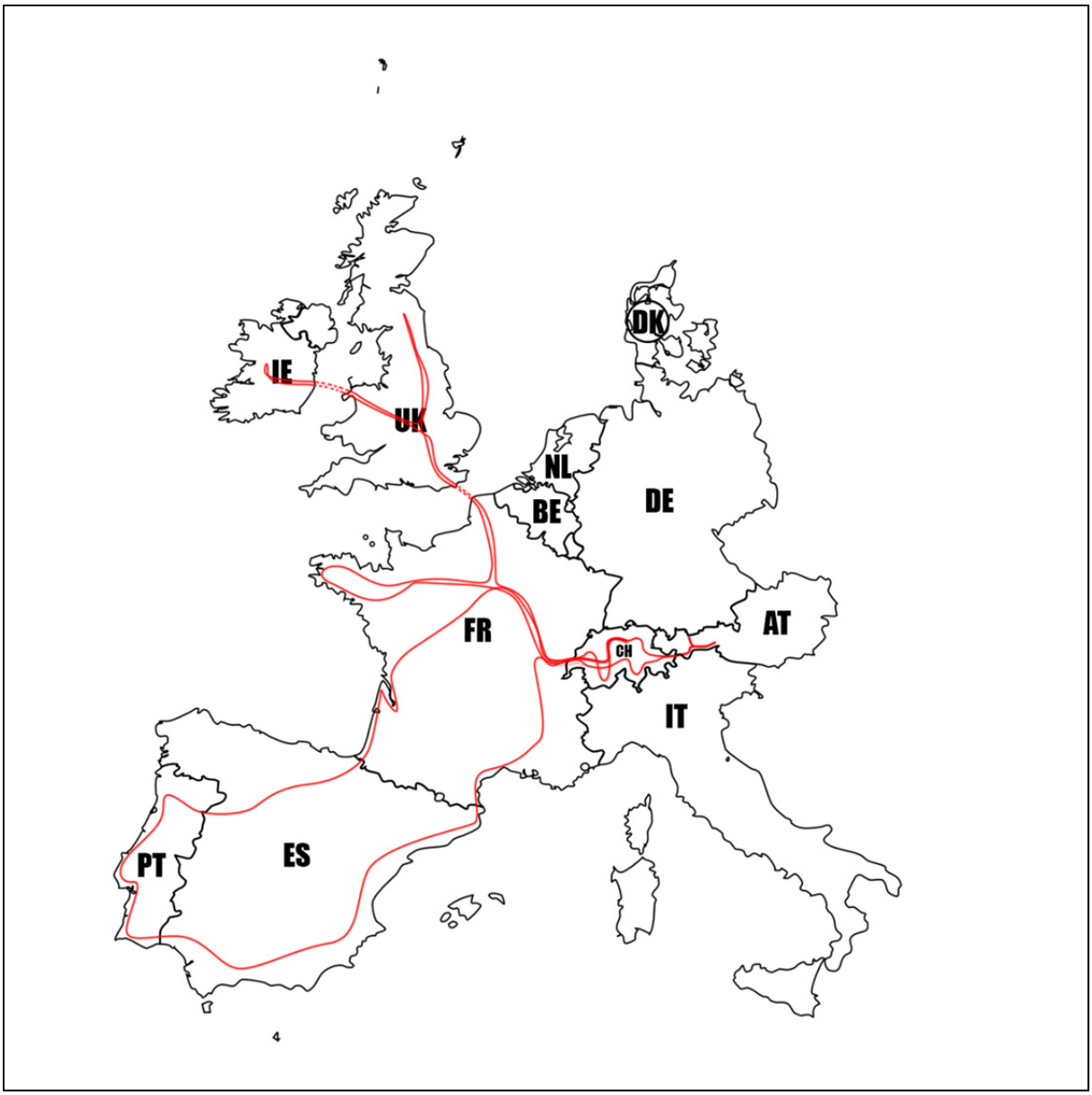

Figure 7.1 Représentation géographique de la carte et de la trajectoire

La seconde figure ci-dessous (Figure 7.2) nous montre les mêmes données sous la forme topologique. Les cercles de couleur pourpre (plus foncé et plus épais) représentent le point de départ et la fin de la trajectoire. Cette fois, on n'observe plus de phénomène d'occlusion, mais on ne peut plus voir les détails de la trajectoire au sein d'un pays. On remarque aussi que, contrairement à la figure précédente, le pourcentage de l'espace disponible utilisé est peu élevé.

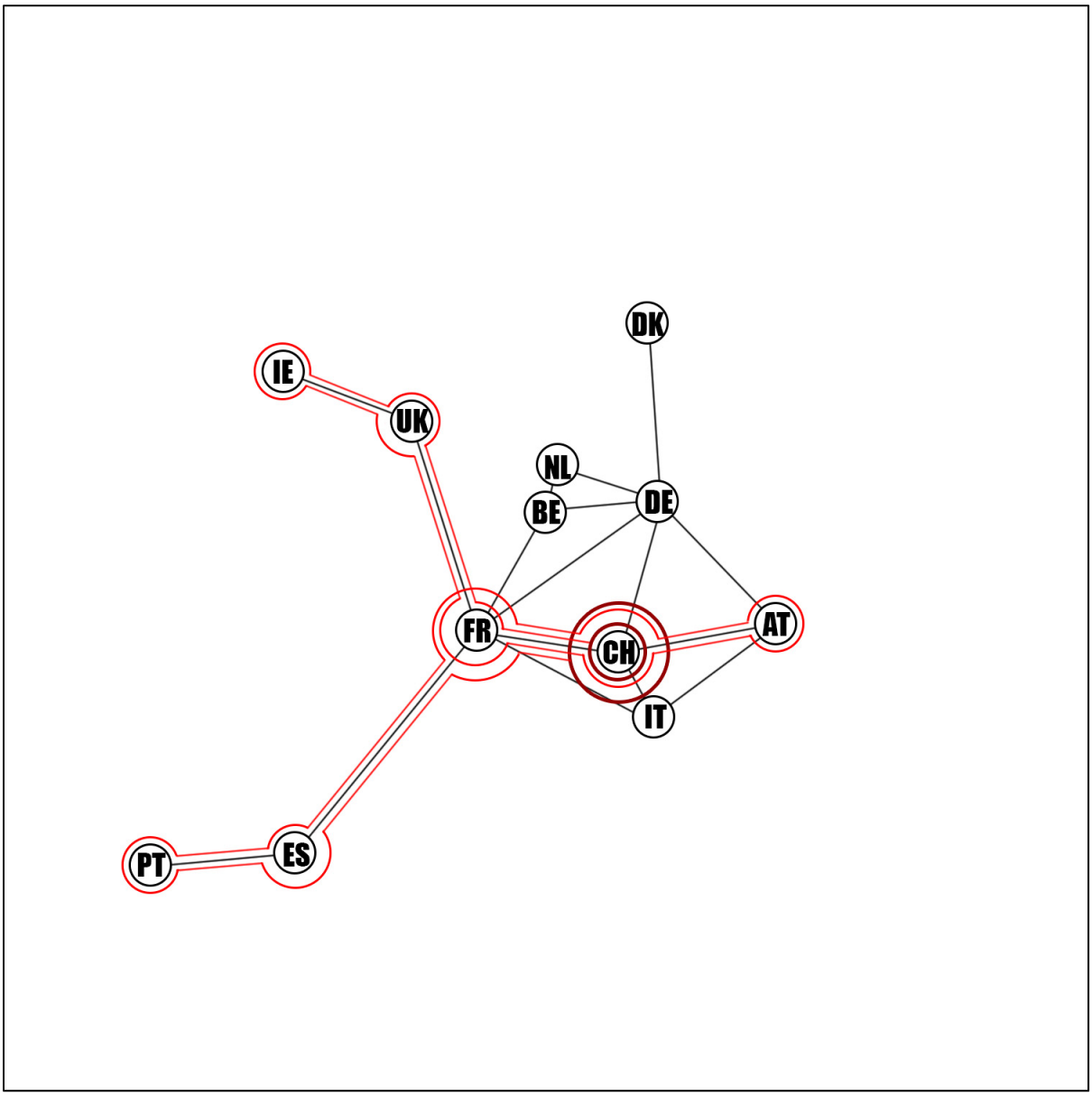

Figure 7.2 Représentation topologique de la carte et de la trajectoire

Enfin, la figure suivante (Figure 7.3) est une représentation hybride des mêmes données. On part de l'observation faite précédemment (par rapport au pourcentage de l'espace utilisé) et on applique un effet de loupe sur la zone qui nous intéresse, ici la Suisse (CH). Ainsi, on peut cette fois-ci observer que la trajectoire commence et finit au même endroit, en y passant une

troisième fois, et que la personne cible visite l'Irlande (IE), avant de visiter l'Autriche (AT), puis l'Espagne (ES) et le Portugal (PT) avant de retourner en Suisse.

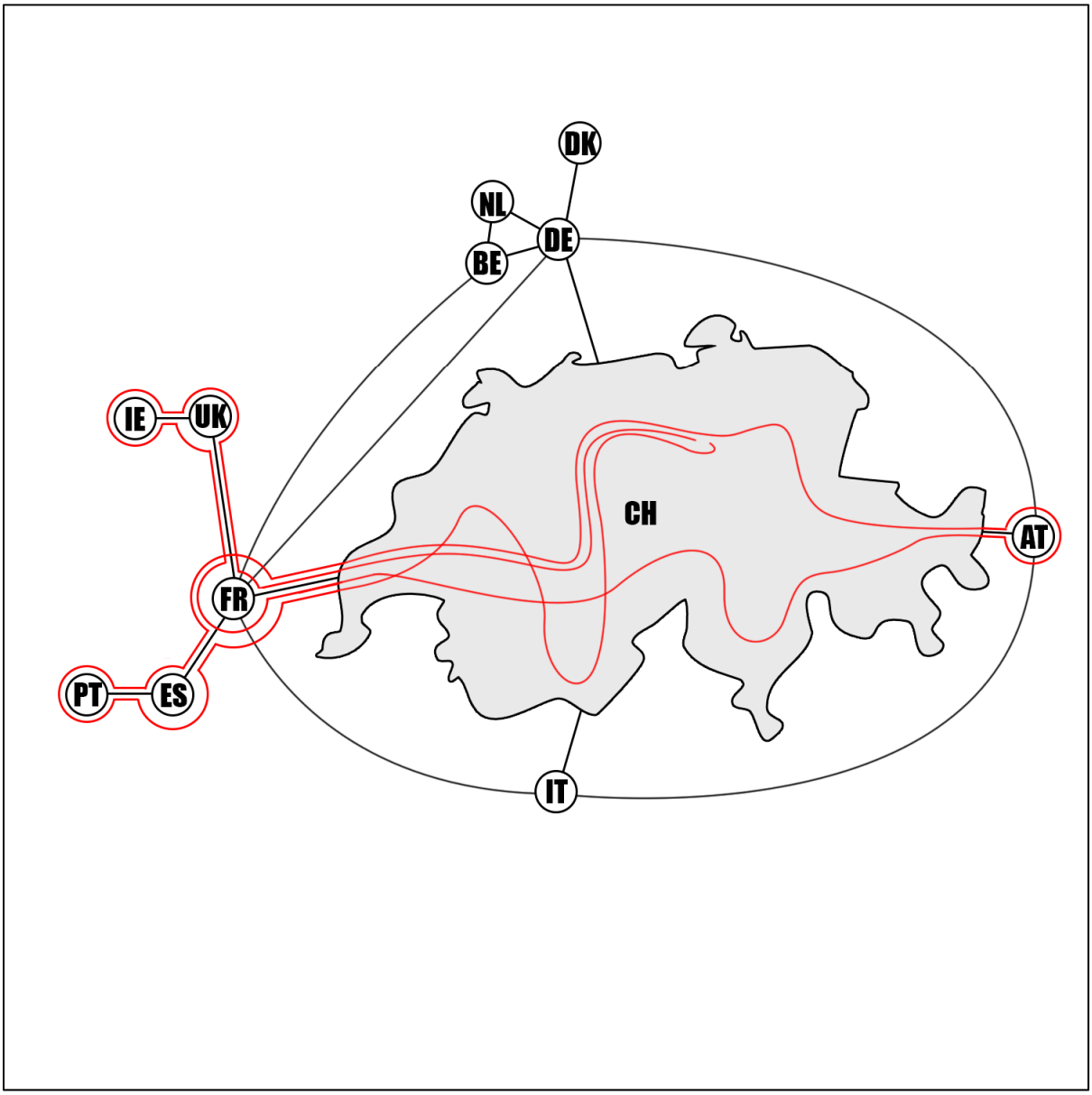

Figure 7.3 Représentation hybride de la carte et de la trajectoire

7.2 Agrégation des données

Dans le paragraphe sur la génération des données (5.2 Génération semi aléatoire: Principe de graphe à deux niveaux), on fait l'observation que les données sont naturellement agrégées. Dans cette optique, plusieurs évolutions seraient envisageables:

- Dans un premier temps, confirmer l'observation
- Appliquer un algorithme d'analyse qui agrégerait les données, et créerait le graphe nœuds-liens correspondant (Procéder de la manière inverse de celle utilisée pour générer les données donc)
- Afficher les trajectoires dans une vue topologique, où les données internes à chaque pièces seraient elles aussi affichées sous forme de graphe nœuds-liens.

L'une des faiblesses de la vue topologique est qu'elle ne permet pas de visualiser les données à l'intérieur d'une pièce. Avec cette modification, on pourrait avoir un niveau de détail supplémentaire sur ce qu'il se passe à l'intérieur d'une pièce.

7.3 Visualisation hybride

Enfin, il est montré dans ce mémoire que chacune des deux techniques présentées, la technique géospatiale et la technique topologique, a des avantages et des inconvénients en fonction de la tâche que l'utilisateur souhaite effectuer. Si la possibilité d'avoir une technique hybride a été présentée et implémentée, celle-ci n'a néanmoins pas été testée au cours d'une expérience contrôlée. Une étude complémentaire pourrait être faite sur l'évaluation de la technique de visualisation hybride par rapport à une interface avec deux vues coordonnées (Vue géospatiale d'un côté, et topologique de l'autre).

CONCLUSION

Plusieurs contributions sont présentées au cours de ce mémoire.

Tout d'abord, deux techniques de visualisation pouvant être utilisées pour visualiser des données au sein d'un bâtiment, la technique géospatiale et la technique topologique ont été présentées. La technique topologique permet de simplifier le plan d'un bâtiment et de montrer sans ambiguïté des trajectoires se superposant, tout en indiquant leur direction de manière implicite en utilisant une métaphore de la circulation routière.

Dans un second temps, une technique de visualisation hybride a été présentée, combinant les représentations géospatiales et topologiques, tout en permettant à l'utilisateur de configurer une visualisation multi-focale, pour montrer les détails des trajectoires à l'intérieur d'une salle sous la forme géospatiale, et pour montrer le contexte sous forme topologique.

Ces trois techniques de visualisation ont été implémentées dans un prototype fonctionnel, qui a ensuite été utilisé dans une expérience contrôlée qui nous a permis de mettre en évidence que chacune des techniques de visualisation a un avantage en termes de temps de réponses, en fonction des tâches demandées à l'utilisateur. Sur les quatre tâches choisies dans l'expérience, la technique topologique est significativement plus efficace pour trois d'entre elles.

Afin d'effectuer un choix logique quant au modèle de graphe à utiliser, une discussion des différents modèles possibles a été effectuée. Aussi, afin de choisir la façon de dessiner les courbes pour relier les nœuds de manière optimale, une comparaison des différents types de représentation de courbes a été faite.

Un outil permettant de générer des données de manière semi-aléatoire (basé sur un système de sous-graphe et de contraintes) a été développé afin d'obtenir des données spatio-

Rapport-Gratuit.com

temporelles réalistes. Cet outil a ensuite été utilisé dans le cadre d'un autre projet du laboratoire pour des données extérieures.

Enfin, il est envisagé d'étendre ces différentes techniques afin d'être appliquées à des données géographiques, plus larges, permettant de visualiser des voyages couvrant plusieurs pays.

BIBLIOGRAPHIE

Alper, Basak, Nathalie Riche, Gonzalo Ramos et Mary Czerwinski. 2011. « Design study of linesets, a novel set visualization technique ». *Visualization and Computer Graphics, IEEE Transactions on,* vol. 17, n° 12, p. 2259-2267.

Andrienko, N., et G. Andrienko. 2006. *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer-Verlag Berlin Heidelberg.

Blaas, Jorik, Charl Botha, Edward Grundy, Mark Jones, Robert Laramee et Frits Post. 2009. « Smooth graphs for visual exploration of higher-order state transitions ». *Visualization and Computer Graphics, IEEE Transactions on,* vol. 15, n° 6, p. 969-976.

Cockburn, Andy, Amy Karlson et Benjamin B. Bederson. 2009. « A review of overview+detail, zooming, and focus+context interfaces ». *ACM Comput. Surv.,* vol. 41, n° 1, p. 1-31.

Crnovrsanin, Tarik, Chris Muelder, Carlos Correa et Kwan-Liu Ma. 2009. « Proximity-based visualization of movement trace data ». In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on*. p. 11-18. IEEE.

Dinkla, Kasper, Marc J. van Kreveld, Bettina Speckmann et Michel A. Westenberg. 2012. « Kelp Diagrams: Point Set Membership Visualization ». *Computer Graphics Forum,* vol. 31, n° 3pt1, p. 875-884.

Guo, Hanqi, Zuchao Wang, Bowen Yu, Huijing Zhao et Xiaoru Yuan. 2011. « TripVista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection ». In *Pacific Visualization Symposium (PacificVis), 2011 IEEE*. p. 163-170. IEEE.

Haunert, J-H, et Leon Sering. 2011. « Drawing road networks with focus regions ». *Visualization and Computer Graphics, IEEE Transactions on,* vol. 17, n° 12, p. 2555-2562.

Havre, Susan, Beth Hetzler et Lucy Nowell. 1999. « ThemeRiverTM: In search of trends, patterns, and relationships ».

Henry, Nathalie, J-D Fekete et Michael J McGuffin. 2007. « NodeTrix: a hybrid visualization of social networks ». *Visualization and Computer Graphics, IEEE Transactions on,* vol. 13, n° 6, p. 1302-1309.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

Herman, Ivan, Guy Melançon et M. Scott Marshall. 2000. « Graph visualization and navigation in information visualization: A survey ». *Visualization and Computer Graphics, IEEE Transactions on,* vol. 6, n° 1, p. 24-43.

Holten, Danny. 2006. « Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data ». *Visualization and Computer Graphics, IEEE Transactions on,* vol. 12, n° 5, p. 741-748.

Holten, Danny, et Jarke J Van Wijk. 2009. « Force-Directed Edge Bundling for Graph Visualization ». In *Computer Graphics Forum*. Vol. 28, p. 983-990. Wiley Online Library.

Johnson, Brian, et Ben Shneiderman. 1991. « Tree-maps: A space-filling approach to the visualization of hierarchical information structures ». In *Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on*. p. 284-291. IEEE.

Kapler, Thomas, et William Wright. 2005. « Geotime information visualization ». *Information Visualization,* vol. 4, n° 2, p. 136-146.

Kristensson, Per Ola, N Dahlback, Daniel Anundi, Marius Bjornstad, Hanna Gillberg, Jonas Haraldsson, Ingrid Martensson, Mathias Nordvall et Josefine Stahl. 2009. « An evaluation of space time cube representation of spatiotemporal patterns ». *Visualization and Computer Graphics, IEEE Transactions on,* vol. 15, n° 4, p. 696-702.

Krzywinski, Martin, Jacqueline Schein, İnanç Birol, Joseph Connors, Randy Gascoyne, Doug Horsman, Steven J Jones et Marco A Marra. 2009. « Circos: an information aesthetic for comparative genomics ». *Genome research,* vol. 19, n° 9, p. 1639-1645.

Lee, Bongshin, Catherine Plaisant, Cynthia Sims Parr, Jean-Daniel Fekete et Nathalie Henry. 2006. « Task taxonomy for graph visualization ». In *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*. p. 1-5. ACM.

Liu, He, Yuan Gao, Lu Lu, Siyuan Liu, Huamin Qu et Lionel M Ni. 2011. « Visual analysis of route diversity ». In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on*. p. 171-180. IEEE.

Pupyrev, Sergey, Lev Nachmanson et Michael Kaufmann. 2011. « Improving layered graph layouts with edge bundling ». In *Graph Drawing*. p. 329-340. Springer.

Reynolds, Craig W. 1999. « Steering behaviors for autonomous characters ». In *Game Developers Conference*. Vol. 1999, p. 763-782.

Riche, Nathalie Henry, Tim Dwyer, Bongshin Lee et Sheelagh Carpendale. 2012. « Exploring the design space of interactive link curvature in network diagrams ». In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. p. 506-513. ACM.

Roberts, J. C. 2007. « State of the Art: Coordinated & Multiple Views in Exploratory Visualization ». In *Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV '07. Fifth International Conference on*. (2-2 July 2007), p. 61-71.

Rufiange, Sébastien, Michael J McGuffin et Christopher P Fuhrman. 2012. « TreeMatrix: A hybrid visualization of compound graphs ». In *Computer Graphics Forum*. Vol. 31, p. 89-101. Wiley Online Library.

Shneiderman, Ben. 1992. « Tree visualization with tree-maps: 2-d space-filling approach ». *ACM Transactions on graphics (TOG),* vol. 11, n° 1, p. 92-99.

Thudt, Alice, Dominikus Baur et Sheelagh Carpendale. 2013. « Visits: A Spatiotemporal Visualization of Location Histories ». *Eurographics Conference on Visualization (EuroVis) (2013)*

Tufte, Edward R. 1983. *The Visual Display of Quantitative Information*. Graphics Press.

Wang Baldonado, Michelle Q, Allison Woodruff et Allan Kuchinsky. 2000. « Guidelines for using multiple views in information visualization ». In *Proceedings of the working conference on Advanced visual interfaces*. p. 110-119. ACM.

Wang, Yu-Shuen, et Ming-Te Chi. 2011. « Focus+ context metro maps ». *Visualization and Computer Graphics, IEEE Transactions on,* vol. 17, n° 12, p. 2528-2535.

Wattenberg, Martin. 2002. « Arc diagrams: Visualizing structure in strings ». In *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*. p. 110-116. IEEE.

Wikipedia.org. « Bézier curve - Rational Bézier curves ». < http://en.wikipedia.org/wiki/B%C3%A9zier_curve >.

Zeng, Wei, Chi-Wing Fu, Stefan Müller Arisona et Huamin Qu. 2013. « Visualizing Interchange Patterns in Massive Movement Data ». *Eurographics Conference on Visualization (EuroVis) 2013,* vol. Volume 32 (2013), Number 3.

Zhao, Shengdong, Michael J McGuffin et Mark H Chignell. 2005. « Elastic hierarchies: Combining treemaps and node-link diagrams ». In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*. p. 57-64. IEEE.