

# Confection de tournées de livraison à partir de l'historique opérationnel

Pierre Baptiste<sup>1,2</sup>, Martin Trépanier<sup>1,2</sup>, Bruno Agard<sup>1,2</sup>

<sup>1</sup> École Polytechnique de Montréal, Département de mathématiques et de génie industriel, C.P. 6079, succursale Centre-Ville, Montréal, Québec, Canada, H3C 3A7

<sup>2</sup> Groupe Polygistique, Centre interuniversitaire de recherche sur les réseaux d'entreprise, le transport et la logistique (CIRRELT), Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec, Canada, H3C 3J7

[pierre.baptiste, mtrepanier, bruno.agard]@polymtl.ca

---

**RÉSUMÉ :** *Le problème de confection de tournées de livraison est un problème NP-difficile largement étudié dans la littérature. Bien que les chercheurs s'entendent à dire que ces problèmes soient quasi insolubles à l'optimal, il y a quand même des millions de livraisons qui sont effectuées chaque jour à travers le monde. Dans la plupart des cas, ces tournées permettent de livrer périodiquement à un ensemble de clients relativement stable. De plus, les tournées réalisées reposent sur la grande expérience des répartiteurs et des chauffeurs et tiennent compte de phénomènes qui sont souvent omis dans les algorithmes classiques de confection, comme le risque de congestion, les préférences des clients et des chauffeurs, ainsi que les configurations particulières des lieux de livraison. Cet article présente une démarche exploratoire visant à confectionner des tournées à partir de l'historique vécue, dans le cas d'une entreprise de distribution. Les travaux démontrent que cette avenue est possible en autant que les clients soient classifiés, et que le choix des tournées peut s'accompagner d'insertions ponctuelles visant à compléter l'historique.*

**MOTS-CLÉS :** *confection de tournées, historique, livraisons.*

---

## 1 Introduction

Chaque jour, des millions de livraisons par camion sont effectuées à travers le monde par des chauffeurs, sous la gouverne de répartiteurs qui fabriquent les tournées à effectuer dans le meilleur de leurs connaissances. Bien que, dans la plupart des cas, ces tournées ne soient pas optimales au sens mathématique du terme, elles sont graduellement améliorées grâce au retour sur expérience qui permet de prendre en compte des phénomènes et des événements survenus lors de l'opération. Par exemple, le risque de congestion est un phénomène difficilement palpable. En heure de pointe, certains répartiteurs préféreront éviter les autoroutes bondées, mais devront quand même les utiliser pour certains clients. Les acteurs du processus ont également chacun leurs préférences propres, ce qui teinte quelque peu la confection des tournées.

Des dizaines de logiciels commerciaux de confection de tournées existent pour les entreprises. Bardés d'algorithmes de recherche opérationnelle souvent extrêmement puissants, ces logiciels proposent des tournées qui ne sont presque jamais appliquées en entier en pratique. Pis encore, il semble que l'emploi en continu de ces logiciels, qui, pour la plupart, font table rase à chaque confection, complexifie le processus en créant des tournées toujours différentes et en risquant « d'infantiliser » les répartiteurs qui perdent une connaissance essentielle de l'opération. En outre, ces logiciels reposent sur des systèmes d'information géographique extrêmement lourds, dont la mise à jour est cruciale à l'obtention de tournées « justes ».

Loin de vouloir discrétiser l'emploi des algorithmes et des logiciels, cet article propose d'explorer une nouvelle méthode de confection de tournées basée sur l'historique

opérationnel de l'entreprise. Est-il possible de desservir un ensemble de clients en n'utilisant que des tournées déjà opérées ? Que peut-on tirer de l'expérience des tournées opérées, où des indicateurs de retour sur expérience sont disponibles au planificateur ? Qu'est-ce qui manque à l'historique ? C'est ce genre de question qui est abordé ici.

Dans un premier temps, une revue de littérature présente certains travaux touchant la confection de tournée s'inscrivant dans un souci de représenter au mieux la réalité. Puis, le contexte opérationnel des données utilisées est présenté. Des éléments méthodologiques touchant entre autres la théorie des ensembles sont ensuite exposés. L'article se clôt par une présentation de quelques résultats portant à la discussion.

## 2 Fondements

La littérature scientifique sur la confection des tournées de véhicules est majoritairement fondée, avec raison, sur les modèles mathématiques sous-jacents, vu la grande complexité de résolution de ce genre de problème. Dans une intéressante revue des travaux scientifiques touchant la gestion des opérations (Operations Management, OM) et la recherche opérationnelle (Operational Research, OR), Bertrand et Fransoo (2002) rapportent que la recherche visant la résolution des problèmes réels se dirige de plus en plus vers l'étude des phénomènes quantitatifs associés aux processus, voire des recherches empiriques.

Dans la communauté de recherche opérationnelle, de grands efforts sont faits pour rendre les modèles et les algorithmes plus « collés » à la réalité. La notion de graphe de faisabilité (feasibility graph) présentée par Beasley et Cristofides (1997) réfère à une série de solutions faisables pour la livraison des clients, plutôt que de restreindre à des séquences fixes. Ruiz et al. (2004) présente un modèle exact reprenant la notion de faisabilité, en y ajoutant une série de contraintes visant à reproduire la réalité du terrain. Les chercheurs montrent les gains économiques potentiels de leur approche, mais n'indiquent pas si les tournées générées ont été effectivement réalisées. Marzolf et al. (2006) ont montré que des problématiques semblables existent pour la confection de tournées de monitoring routier et ont proposé une méthode tenant compte de l'expérience des planificateurs.

La recherche portant sur des cas réels n'est pas dénuée d'intérêt car elle permet de saisir toutes les nuances pouvant être apportées par l'opération. Eibl et al. (1994) rapportent une série de contraintes opérationnelles rencontrées dans l'industrie de livraison de la bière : clients avec fenêtres de temps restreintes, accès difficiles à certaines installations, réseau urbain truffé de rues à sens uniques, capacité des camions, règles non écrites concernant le balancement des heures de travail des chauffeurs, etc. Les auteurs présentent une implantation d'un logiciel de confection de tournées en précisant qu'il est difficile d'en démontrer les bénéfices sans considérer l'ensemble des coûts impliqués, comme ceux touchant la gestion du changement et la modification des façons de faire auprès des clients. Dans une étude de cas touchant des livraisons entre les bibliothèques de San Francisco, Apte et Mason (2006) insistent sur la nécessité de réviser les processus d'affaires préalables à la livraison (dans ce cas-ci, ajout de transbordement et de balancement) avant même de confectionner les tournées. Les gains sont ainsi beaucoup plus substantiels.

Dans les dernières années, le développement des solutions informatiques intégrées, l'avènement de logiciels de calculs de chemins peu coûteux et le développement de sites Internet cartographiques bouleverse considérablement le milieu de la confection de tournées. Rappelons que le nerf de la guerre dans le calcul de chemins routiers est le temps de parcours, qui dépend des conditions de congestion, qui elles dépendent de l'heure, la journée, voire la saison. Eglese et al. (2006) présentent la constitution de tables de temps de marche (timetable) et insistent sur l'importance d'en tenir compte dans l'élaboration des tournées. Les résultats démontrent de grandes variabilités résultant de l'utilisation ou non d'un temps de

parcours collant à la réalité. D’ailleurs, les temps de marche ne peuvent habituellement être obtenus qu’en liant un réseau provenant d’un système d’information géographique à des considérations opérationnelles, ce qui augmente d’autant la complexité de l’opération (Trépanier et Chapleau 2001).

Vu l’importance stratégique accordée aux opérations de transport dans les entreprises, on tente actuellement d’intégrer les systèmes de gestion du transport dans les systèmes de gestion de la chaîne logistique, ce qui permet entre autres l’emploi de systèmes d’information géographiques (Caputo et al. 2003). Gayialis et Tatsiopoulos (2004) ont présenté un système de confection de tournées basé sur les technologies informationnelles (« IT-driven ») essayant de tirer profit de l’information dans l’utilisation des algorithmes de recherche opérationnel sous-jacents. Par l’intégration aux autres systèmes de l’entreprise, ils peuvent évaluer immédiatement les effets de la planification des tournées sur les activités connexes tels que la prise de commande, le service à la clientèle, etc.

### 3 Méthodologie

La méthodologie de recherche que nous avons développée est liée à la nature des informations susceptibles d’être recueillies. Nous cherchons à montrer la faisabilité du passage d’un paradigme de calcul de tournées par une optimisation à une approche par apprentissage (Figure 1). Les deux approches sont d’ailleurs plus complémentaires qu’opposées.

Nous distinguerons ici trois types de tournées :

1. Les tournées **théoriques**, c’est-à-dire les tournées conçues à l’aide d’un algorithme visant à minimiser les ressources consommées (temps, distance) à partir d’une liste de clients déjà affectés. Cette minimisation prend en compte des temps de parcours attribués aux liens du réseau routier.
2. Les tournées **corrigées**, c’est-à-dire des tournées basées sur les tournées théoriques, mais qui ont été modifiées par le répartiteur en tirant partie de son expérience. Les changements apportés peuvent toucher aussi bien la liste des clients desservis que l’ordre dans laquelle ils sont desservis.
3. Les tournées **réalisées**, c’est-à-dire les tournées qui ont été réellement effectuées sur le terrain. Même dans les meilleures situations, les tournées corrigées peuvent ne pas être exécutées telles quelles, à cause des imprévus de dernière minute et des aléas de la congestion routière.

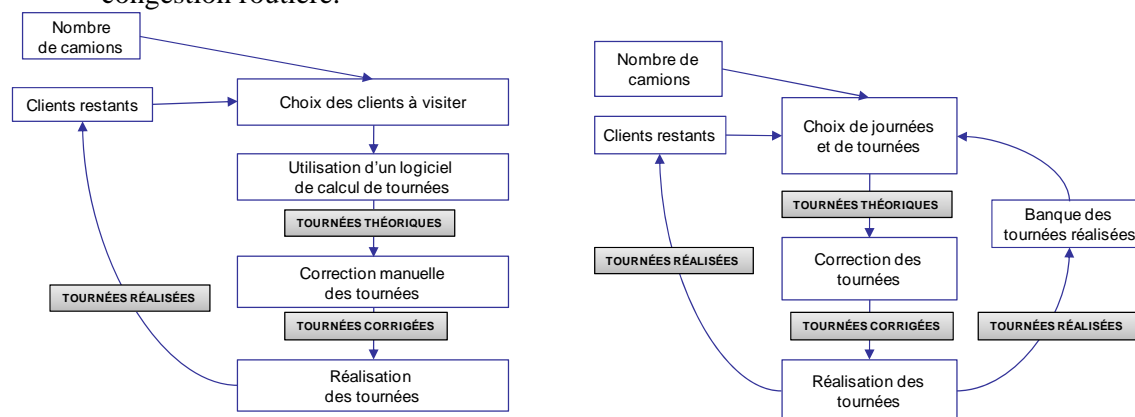


Figure 1: modèles de fonctionnement étudiés

Notre hypothèse est que l’effort de correction des tournées sera plus faible avec une qualité de tournées comparable. Les raisons des corrections sont : la connaissance des choix des clients, les problèmes liés au trafic, les problèmes liés à la manipulation des produits, les problèmes liés à la structure du réseau de transport, et les préférences des chauffeurs.

Pour étudier cette faisabilité, nous proposons une approche en 3 temps :

- collecte d'un ensemble de données réelles ;
- recherche des meilleures stratégies d'apprentissage à partir des données historiques et
- validation de la faisabilité de l'apprentissage.

La recherche de la meilleure stratégie se fera en utilisant les données collectées comme un ensemble d'apprentissage. En retranchant soit une tournée, soit une journée, soit une semaine complète à l'ensemble des tournées collectées, on dispose d'une part d'un item nouveau (l'objet retranché) et d'autre part d'un ensemble d'apprentissage (le reste). On peut alors utiliser les méthodes proposées pour obtenir les tournées de la semaine ou de la journée que l'on peut comparer aux tournées réalisées. La Figure 2 donne une idée du schéma de validation. Une fois ces tournées obtenues, on peut les comparer aux tournées calculées par le système classique, puis corrigées. La phase de comparaison doit aussi inclure une évaluation par l'expert des solutions proposées par le système.

Finalement, la validation réelle demandera d'utiliser non pas les tournées légèrement corrigées dont nous disposons ici, mais les données sur les tournées réalisées (avec retour sur expérience), non disponibles à ce jour.

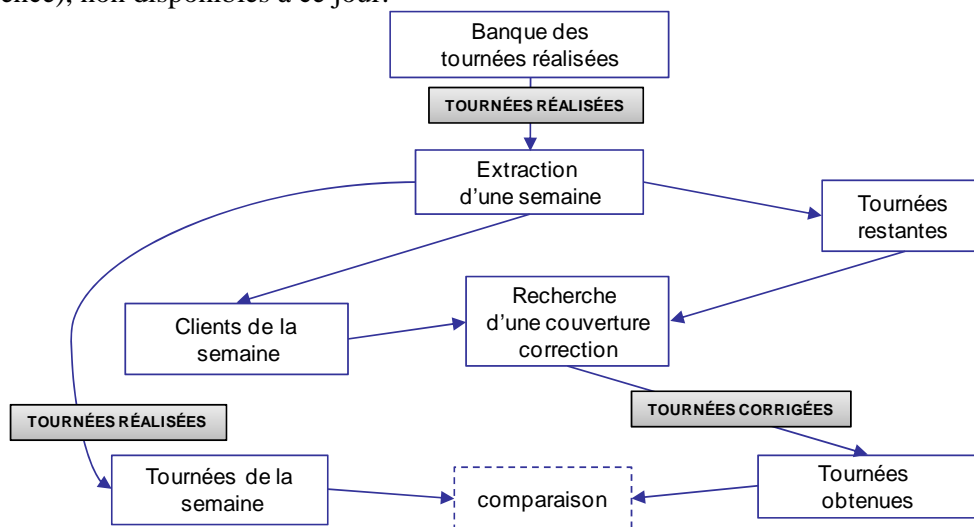


Figure 2: schéma de base de la phase de validation

### 3.1 Collecte des données

Une entreprise de la région montréalaise a été utilisée pour la cueillette de données. Elle a effectué 951 tournées de livraison sur 27 semaines, à 5 jours par semaine, en utilisant entre 7 et 8 camions par jour. Elle a effectué 16 384 livraisons, soit environ 17 arrêts par tournée. Elle a visité 2150 clients différents (800 n'ont été visités qu'une fois, et 375 deux fois). L'entreprise fonctionne suivant le schéma donné à la gauche de la Figure 1. Chaque semaine, il faut visiter un ensemble de clients (en moyenne, 330 clients, écart type de 100). Le répartiteur choisit chaque jour la liste des clients à visiter et le nombre de camions. Il utilise un logiciel pour établir les 7 ou 8 tournées. Le logiciel sectorise de manière radiale les clients à visiter puis établit un algorithme de plus court chemin. Les tournées sont ensuite manuellement corrigées, pour en faire des tournées « acceptables » par l'utilisateur, du point de vue du répartiteur. Ensuite, les tournées sont effectuées. Nous n'aurons accès ici qu'aux tournées légèrement corrigées.

La première analyse consiste à observer la répartition des clients visités dans l'espace. La Figure 3 montre le résultat de deux agrégations spatiales. Dans le découpage de gauche, les cellules ont 300 m de côté, tandis que le découpage de droite compte des cellules de 3 km de côté. Les cellules noires ont été visitées plus de 100 fois, les grises foncées plus de 50 fois,

les grises claires entre 1 et 50. Dans l'agrégation de gauche, il y a 450 cellules non vides, dont seulement 115 ont été visitées une seule fois. Dans celle de droite, seulement 95 cellules ont été visitées. Le lecteur avisé reconnaîtra les villes de Montréal et Laval (au nord-ouest).

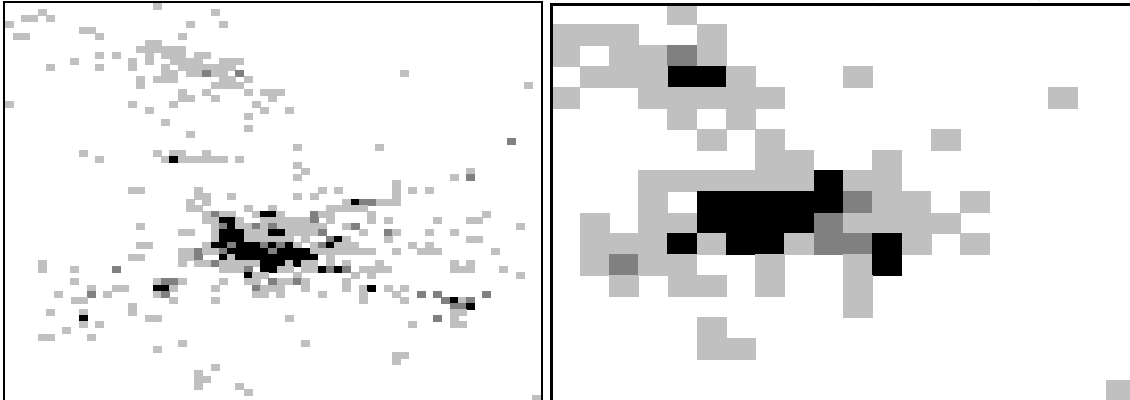


Figure 3: cellules agrégées

### 3.2 Recherche d'une méthode d'apprentissage

L'apprentissage revient dans ce cas à rechercher un ensemble de tournées existantes (déjà faites et validées) qui peut permettre de « réaliser » en tout ou en partie l'ensemble des livraisons d'une journée ou d'une semaine donnée.

Si l'on considère l'ensemble  $C$  des clients et l'ensemble  $T$  des tournées existantes, on peut considérer le graphe bipartite  $G(C,T,A)$  où un arc «  $a$  » est défini entre le client «  $c$  » et une tournée «  $t$  » si ce client appartient à cette tournée. Le problème consiste à chercher pour une semaine  $S$  donnée correspondant à un sous-ensemble «  $C_s$  » de  $C$  (les clients à parcourir), une couverture minimale en tournées (avec plusieurs contraintes additionnelles).

Cette approche idéale n'est cependant pas faisable directement. Dans notre cas, les données collectées montrent que si 20% des clients représentent 80% des arrêts, 40% des clients ne sont visités qu'une seule fois sur toute la période. Il s'ensuit que l'apprentissage de tournées complètes n'est pas possible. C'est pourquoi nous nous sommes tournés vers une agrégation spatiale des données. Il n'en reste pas moins que le travail avec des données réelles montre qu'il restera une phase de correction en complément de l'apprentissage. La recherche d'une couverture minimale doit se fonder dans une démarche plus globale de filtrage et de correction.

Pour un tel problème, l'apprentissage peut se concevoir dans un espace à deux dimensions (voir Figure 4) : l'agrégation spatiale (regroupement de clients en cellule de taille plus ou moins grande) et l'objet appris (une semaine, une journée, une tournée). Idéalement, l'apprentissage permettra de trouver d'abord, pour chaque semaine, une autre semaine ayant déjà contenu « exactement » ces clients (ensemble d'apprentissage très petit, besoin d'une forte répétitivité). On reprendrait alors le découpage en journée et les tournées de chaque journée. À l'autre extrémité, on agrège les clients en grandes zones et l'on apprend en utilisant des tournées effectuées des jours différents et des semaines différentes. L'ensemble d'apprentissage est très grand et l'objet appris n'est pas exactement celui cherché (agrégation).

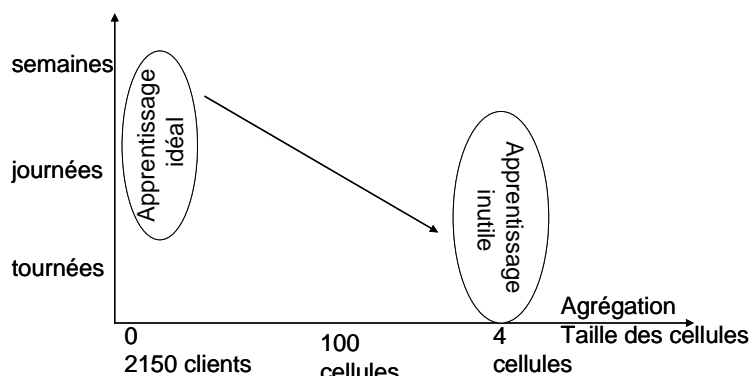


Figure 4: espace des apprentissages possibles

### 3.3 Tournée utilisant les clients (pas d'agrégation)

La première hypothèse consiste à chercher à générer pour une journée donnée (environ 7 tournées à effectuer) toutes les tournées en utilisant une journée du passé.

L'analyse des données fournies montre que les livraisons hebdomadaires ne sont pas réparties dans la semaine en fonction de critères géographiques mais qu'au contraire, les livraisons de chaque journée couvrent tout le territoire. La Figure 5 montre que deux journées différentes de 7 tournées chacune couvrent toutes deux l'ensemble du territoire et qu'en fait, le logiciel traite de la même manière les deux journées. La tournée J1-1 correspond à la tournée J2-1, J1-2 à J2-3, J1-3 à J2-6, J1-4 à J2-4, J1-5 à J2-5, J1-6 à J2-2 et J1-7 à J2-7.

De cette constatation peut naître l'espoir de trouver, pour une journée donnée, une autre journée suffisamment semblable pour servir de modèle à l'élaboration de ses tournées. Cependant, en réalisant une classification hiérarchique entre l'ensemble des 135 journées, on constate que les différences entre journées sont grandes (Figure 6).

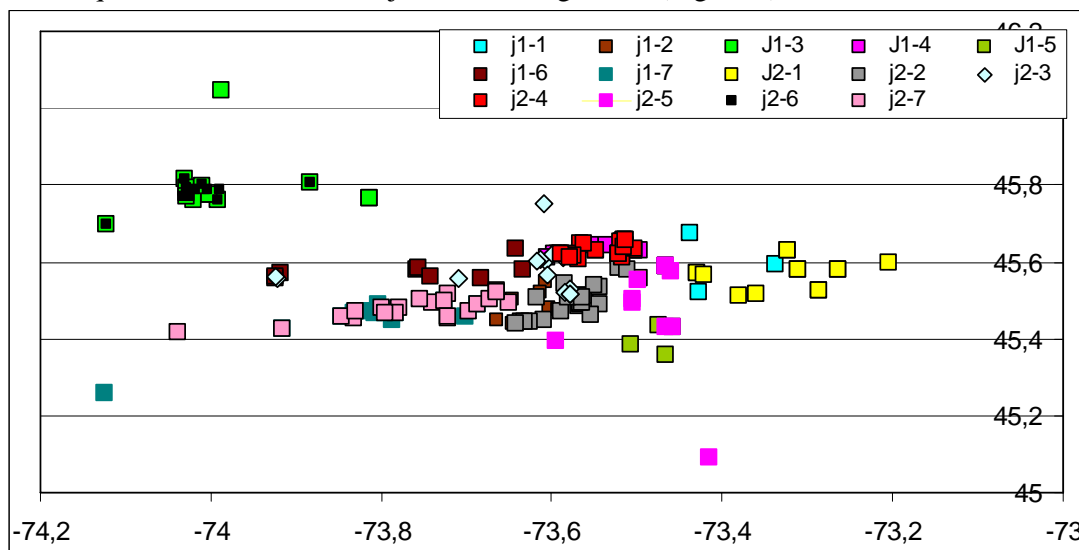
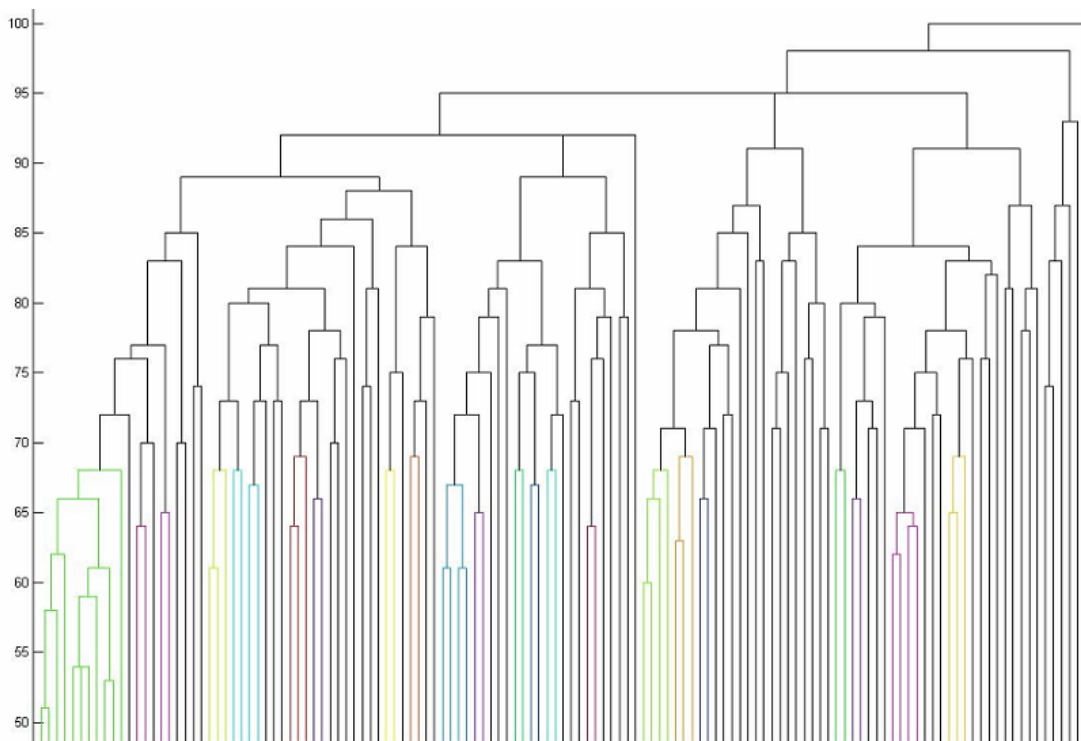


Figure 5: les 7 tournées de 2 journées, J1 et J2



**Figure 6 : dendrogramme de classification des journées**

La classification est faite en utilisant comme indice de dissimilitude le nombre de clients différents d'une journée à l'autre, puis comme distance l'ultra métrique supérieure minimale. Il ressort de cette analyse qu'il est illusoire de chercher à apprendre seulement à partir des tournées, des journées ou des semaines à partir de cet ensemble de données. Ceci est principalement dû aux 1750 clients sur 2100 qui ne sont visités qu'une ou deux fois.

### 3.4 Recherche du niveau d'agrégation nécessaire

Ce qui rend impossible l'apprentissage de tournées à base de clients, c'est le trop grand nombre de clients visités une seule fois. L'agrégation doit donc d'une part être un découpage suffisamment fin pour que l'on puisse considérer que deux clients de la cellule sont virtuellement équivalents, et d'autre part suffisamment grande pour minimiser les cellules visitées une seule fois. En mettant la barre à 10% au plus de cellules visitées une seule fois, on obtient un découpage de la zone en 400 cellules de 3 km sur 3, dont 95 ont été visitées et seulement 8 visitées une seule fois. Pour étudier la faisabilité de l'apprentissage avec cette agrégation, on peut étudier la matrice de similarité entre les journées.

Une journée  $J_j = (n_{i,k})_{k=1..K}$  est un vecteur d'entier,  $n_{i,k}$  étant le nombre de client de la cellule  $k$  visité le jour «  $i$  ». On définit la similarité  $D(J_i, J_j)$  par :

$$D(J_i, J_j) = \frac{\sum_k \min(n_{i,k}, n_{j,k})}{\sum_k n_{i,k}}$$

Cet indice est compris entre 0 et 1, et il est asymétrique car le dénominateur dépend de la première journée. Si l'indice est à 1, c'est que la journée  $J_i$  est totalement incluse dans la journée  $J_j$  et donc que toutes les tournées faites la journée  $J_j$  peuvent servir telles quelles pour la journée  $J_i$ .

Sur l'ensemble des 130 journées, la similarité minimale entre une journée  $J_i$  et le reste des journées est de 0,84 et la moyenne 0,96. La Figure 7 montre sur les 40 premières journées (extraites des 130, pour demeurer lisible) la similarité maximale d'une journée avec une autre

journée de l'ensemble d'apprentissage. Les 40 premières journées ont été ordonnées en fonction du nombre de clients visités. Il est clair que l'indice de similarité est non symétrique et qu'une journée composée d'un nombre très élevé de clients n'a pas son équivalent dans l'ensemble d'apprentissage.

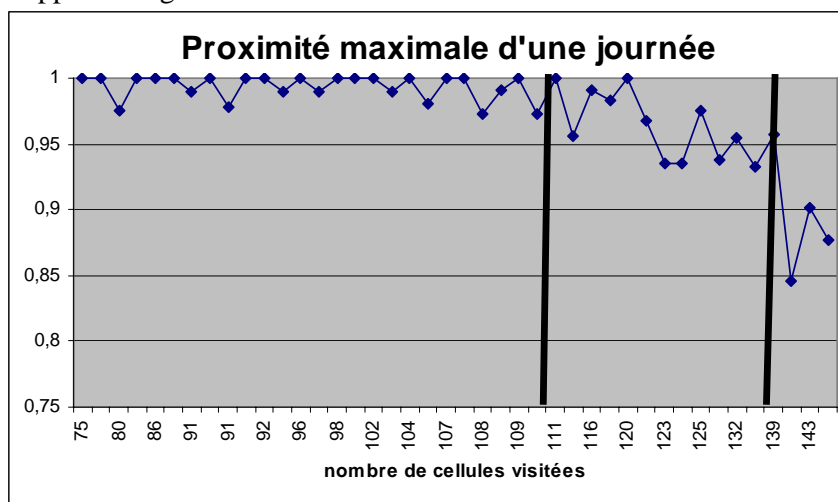


Figure 7: distance entre journées

Cette figure montre clairement que pour la majeure partie des journées, il existe dans l'ensemble historique une journée très voisine.

On voit donc que l'on peut définir 3 zones, en fonction du nombre de clients visités :

- Inférieur à 110 : aucun apprentissage n'est nécessaire, il y a dans l'historique des journées une journée tellement similaire que toutes les tournées effectuées cette journée peuvent être reprises telles quelles.
- Entre 110 et 140 clients, aucune journée ne correspond exactement, on doit utiliser un algorithme de couverture minimale en utilisant non pas les journées mais les tournées effectuées.
- Au delà, il faut utiliser la méthode actuelle de génération de tournées.

Finalement, cette étude montre qu'avec un niveau d'agrégation correct, une majeure partie des journées ne sont que des recopies de journées existantes, que l'apprentissage concernera un sous-ensemble finalement assez restreint de journées et qu'il restera toujours des journées nécessitant un algorithme de création de tournées.

### 3.5 La validation

Nous avons effectué une étude de faisabilité théorique de notre approche. Il faudra dans un second temps faire une validation. Cette validation devra prendre en compte les points suivants :

- Il faut travailler avec les tournées réellement effectuées car la justification même de notre proposition est que les tournées réelles sont faisables et donneront une meilleure idée des contraintes de terrain que les tournées corrigées.
- Il faut prendre en compte la capacité des véhicules et les demandes des clients. Dans le cas étudié, la capacité des camions effectuant les tournées n'était pas un gros problème. Il faut néanmoins vérifier que ces capacités soient respectées.
- Il faut traiter de l'affectation. En effet, dans les zones 1 et 2, on sait qu'un ensemble donné de tournées peut faire le travail d'une journée donnée, mais il faut encore affecter les clients d'une cellule à une tournée (plusieurs tournées utilisent des clients de la même cellule).



Ce dernier degré de liberté devrait permettre de régler le problème de capacité. Idéalement, ce degré de liberté serait parfaitement utilisable par l'expert qui peut ainsi utiliser ses connaissances « locales » de préférence. Mais on crée une tâche jusqu'ici inconnue, qui n'a pas à ce jour de place dans l'organisation du travail. Ce point devra être éclairci.

## 4 Conclusion

Notre objectif était d'explorer une voie originale pour constituer des tournées de véhicules avec une méthode d'apprentissage utilisant une base de données historique. Nous présentons une méthode de recherche basée sur l'exploitation de données réelles, la recherche d'une méthode combinant un filtrage (clients exceptionnels), une couverture minimale et une correction finale. Malheureusement, nous ne disposons que des tournées corrigées. Nous n'avons pas accès aux tournées réalisées. Les résultats initiaux montrent qu'il y a de la place pour trois approches complémentaires : une zone de simple réutilisation de journées existantes, une zone permettant effectivement de réutiliser des tournées existantes (effectuées) pour constituer la base des tournées de la journée et une zone où l'algorithme actuel trouve tout son intérêt. Il reste à valider que le résultat obtenu intègre bien les éléments corrigés par l'expert et ne demandera pas autant de modification que les tournées initialement obtenue par le logiciel. Cette approche doit être vue comme un complément de l'approche classique, pas comme une alternative.

## 5 Références

- Apte U.M., Mason, F.M. (2006). « Analysis and improvement of delivery operations at the San Francisco Public Library », *Journal of Operations Management* 24 pp. 325–346.
- Beasley, J.E., Cristofides, N. (1997). « Vehicle routing with a sparse feasibility graph », *European Journal of Operational Research* 98 pp. 499-511.
- Bertrand, J.W.M., Fransoo, J.C. (2002). « Operations management research methodologies using quantitative modeling », *International Journal of Operations & Production Management* 22(2) pp. 241-264.
- Caputo, C., Pelagage, M., Scacchia, F. (2003). « Integrating transport systems in supply chain management tools », *Industrial Management & Data Systems* 103(7) pp. 503-515.
- Eglese, R., Madena, W., Slaterb, A. (2006). « A Road Timetable to aid vehicle routing and scheduling », *Computers & Operations Research* 33 pp. 3508–3519.
- Eibl, P.G., Mackenzie, R., Kidner, D.B. (1994). « Vehicle Routing and Scheduling in the Brewing Industry: A Case Study », *International Journal of Physical Distribution & Logistics Management*, 24(6), pp. 27-37.
- Gayialis, S.P., Tatsiopoulos, I.P. (2004). « Design of an IT-driven decision support system for vehicle routing and scheduling », *European Journal of Operational Research* 152 pp. 382–398.
- Marzolf, F., Trépanier, M., Langevin, A. (2006). Road Network Monitoring: Algorithms and a Case Study, *Computers & Operations Research*, 33(12), pp. 3494-3507.
- Ruiz, R., Maroto, C., Alcaraz, J. (2004). « A decision support system for a real vehicle routing problem », *European Journal of Operational Research* 153, pp. 593–606.
- Trépanier, M., Chapleau, R. (2001). Linking Transit Operational Data to Road Network with a Transportation Object-Oriented GIS, *Urban and Regional Information Systems Association Journal*, 13(2), pp. 23-27.