Optimisation de réseaux de transport

7 février 2008

Serigne Gueye

Université du Havre

Etant donné plusieurs localisations (ville, quartier, zonages issus d'un découpage plus sophistiqué, etc.), d'une zone géographique, des demandes (origine-destination) estimées entre chacune des localisations et un ensemble de liens les reliant, optimiser un réseau de transport pour ces demandes consiste à déterminer le dimensionnement (capacité) optimal des liens permettant de satisfaire toutes les demandes au moindre coût. Ce problème prend le nom de synthèse de réseau dans la littérature du domaine [2].

Le réseau permettant de satisfaire les demandes peut être vu comme un graphe dont les noeuds $i \in V = \{1, 2, ..., n\}$ représentent les localisations et les arcs $e = (i, j) \subset E = V \times V$ les liens de transport auxquels sont associés des fonctions de coût. Nous notons $c_e(x)$ ces fonctions. Elles dépendent du flot x_e (à déterminer) de trafic écoulé sur le lien. x représente les vecteurs des flux sur tous les arcs (i.e $x = (x_e)_{e \in E}$). Dans l'absolu, comme nous venons d'écrire que la fonction de coût c_e dépend uniquement de x_e il conviendrait d'écrire $c_e(x_e)$. La notation plus général $c_e(x)$ permet de tenir compte des cas où le coût sur un arc dépend également des coûts sur les autres arcs.

Optimiser ce réseau consiste à déterminer les capacités de chacun de ces liens de manière à ce que toutes les demandes soient satisfaites, tout en minimisant une fonction objectif représentant le coût du réseau. On prend en général la somme des coûts de chaque arc, et quand la décision conduit à ce qu'une capacité nulle soit affectée à un lien cela revient physiquement à le supprimer du réseau.

Les capacités des liens dépendant des flots qu'ils écoulent, il est nécessaire de déterminer les routes qu'empruntent les demandes.

Concrétement, dans un réseau de transport urbain de personnes les liens se ramènent aux axes routiers, le dimensionnement à leurs capacités (nombre de voies ou longueur nécessaire à écouler x_e véhicules, etc.), et le coût sur chaque arc peut être (par exemple) le temps moyen unitaire (t_e, fixé à priori) estimé pour parcourir l'arc (se rendre de l'origine de l'arc vers sa destination). L'interprétation des noeuds du réseau peut être macroscopique, dans ce cas il s'agit de quartiers entiers ou de villes d'une agglomération, ou microscopique, dans ce cas on s'intéressera aux ronds-points ou aux carrefours.

Si on suppose que t_e est connue, un exemple de fonction de coût pourrait être :

$$c_e(x_e) = t_e * x_e$$

On peut cependant relever que t_e devrait dépendre de la décision sur la capacité à allouer à l'arc e. En effet, si le dimensionnement de la capacité consiste à décider de la longueur de la route séparant une origine et une destination, plus la route sera longue et plus le temps pour la parcourir le sera. Comme la décision sur le dimensionnement de la capacité est liée à celle du flot x_e, t_e devrait donc dépendre aussi de x_e. La fonction de coût devient alors non linéaire :

$$c_e(x_e) = t_e(x_e) * x_e$$

Selon la nature de la fonction de coût considérée (linéaire ou non), la synthèse du réseau peut se révéler être un exercice simple (cas linéaire) ou extrêment difficile (cas non linéaire, se ramenant à un problème NP-complet).

L'estimation de la demande est réalisée par le calcul d'une matrice dite "origine-destination" fournissant entre chaque noeud du réseau la quantité totale estimée du flot entre l'origine et la destination. La détermination de cette matrice est une application mathématique très riche ayant fait l'objet de plusieurs études (voir Bierlaire [1]). Les techniques permettant de calculer cette matrice origine-destination peuvent se classifier en méthodes statiques ou dynamiques selon que l'on calcule une seule matrice ou une matrice dépendant du temps. Parmis les méthodes statiques, on peut citer :

1. La résolution d'un problème de "*maximisation d'entropie*" (voir Bierlaire [1]) : simple car il s'agit de la maximisation d'une fonction concave dans un domaine convexe. Nous la détaillerons en section 1
2. La modélisation sous la forme d'un problème de moindres carrées avec contraintes sur les bornes des variables représentant la distance entre des observations (comptages par exemple) sur le terrain et la matrice à estimer. Dans ce cas on cherche donc à minimiser cette distance.
3. Les modèles entrées-sortie en transport de marchandises

Une matrice origine-destination statique ne contient que des valeurs fixes de flux, et peut donc être vue comme une photographie approximative sur une période donnée. La dynamique urbaine ou celle de la demande de produits (en transport de marchandises), susceptibles de modifier de manière significative sur une autre période les informations exploitées pour la construction du réseau, ne sont pas considérées. D'où la nécessité de modèles plus dynamiques parmi lesquels on trouve la méthode du filtre de Kalman appliqué au transport urbain ([1]).

Les méthodes statiques ont semble-t-il fait l'objet de travaux (et d'applications) plus importants que les modèles dynamiques. On constate par exemple que l'utilisation de modèle de dynamique de populations au problème de dimensionnement de réseau n'a pas (à notre connaissance) fait l'objet d'études.

Un champ de recherche que nous proposons donc d'aborder est l'étude de l'intégration des modèles de dynamique urbaine dans ceux plus opérationnels de synthèse de réseaux.

La section 3 est consacrée à détailler un procédé intégrant les deux aspects. Nous nous intéresserons particulièrement aux modèles de dynamique urbaine qui ont fait l'objet d'une thèse récente ("Modélisation et étude mathématique et informatique de modèles en dynamique urbaine : Application à la ville de Marrakech", Jalila El Ghordaf, Unité de Recherche GEODES - IRD, Novembre 2007). Nous présentons en section 3 une modèlisation dynamique possible tirée de l'article [3].

Quand les demandes origine/destination ont été calculées, le problème proprement dit de l'optimisation des capacités se modèlisent comme un problème de minimisation d'une fonction représentant le coût global du réseau sujet à une série de contrainte que nous détaillerons section 2.

1 Calcul d'une matrice Origine-Destination (O-D) statique

Le calcul d'une matrice O-D se fait en déterminant au préalable pour chaque zone $i \in \{1, 2, ..., n\}$ le nombre de voyageurs O_i qui en partent et le nombre de voyageurs D_i qui y arrivent.

Si $T = (T_{ij})_{1 \leq i,j \leq n}$ désigne la matrice O-D à trouver, celle-ci doit alors vérifier :

$$\sum_{j=1}^{n} T_{ij} = O_i \ \forall \ i = 1, 2, ..., n \tag{1}$$

$$\sum_{i=1}^{n} T_{ij} = D_j \ \forall \ j = 1, 2, ..., n \tag{2}$$

L'équation 1 (resp. 2) permet de distribuer les voyageurs partant (resp. arrivant) en i sur les valeurs T_{ij}. Comme un très grand nombre de matrices peuvent vérifier ces deux équations, la matrice qui sera retenue est celle qui minimisera un certain critère. L'un des critères utilisé dans les méthodes d'entropie est celui de la maximisation de la vraisemblance de la matrice. Le problème d'optimisation résultant est alors le suivant :

$$\begin{array}{llll} (DEMANDE) & Max & \frac{(\sum\limits_{i,j} T_{ij})!}{\prod\limits_{i,j} T_{ij}!} & \\ & s - \text{à} : & \sum\limits_{j=1}^{n} T_{ij} & = O_i \\ & & \sum\limits_{i=1}^{n} T_{ij} & = D_j \\ & & T_{ij} & \in \mathbb{R}^+ \end{array}$$

La fonction objectif de ce problème est une mesure de la vraisemblance de la matrice. Elle est obtenue en comptant le nombre de descriptions détaillées possibles associées à la matrice et en supposant qu'elles sont toutes équiprobables. Une description détaillée consiste à considérer pour chaque individu son point de départ et d'arrivée. De manière à simplifier le problème, un logarithme est appliquée à la fonction objectif suivi d'approximations des factorielles conduisant au problème équivalent suivant

$$\begin{array}{llll} (DEMANDE) & Max & -T_{ij}log(T_{ij}) - T_{ij} & \\ & s - \text{à} : & \sum\limits_{j=1}^{n} T_{ij} & = O_i \\ & & \sum\limits_{i=1}^{n} T_{ij} & = D_j \\ & & T_{ij} & \in \mathbb{R}^+ \end{array}$$

Il s'agit de la maximisation d'une fonction concave dans un domaine convexe pour lequel de nombreuses méthodes de complexité polynomiale existent.

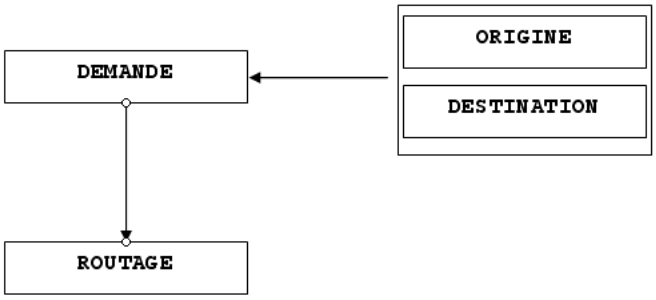

FIG. 1 – Routage statique

2 Synthèse statique du réseau

Avec la matrice O-D la synthèse statique du réseau revient à déterminer le routage des volumes T_{ij} permettant de minimiser une fonction de coût global du réseau.

Si $e \in E$ est un arc du réseau, rappelons que x_e est le flot de trafic sur e et $c_e(x)$ le coût. La mesure du coût global est en général pris comme étant la somme des coûts sur tous les arcs :

$$g(x) = \sum_{e \in E} c_e(x) x_e$$

Désignons par :
- P : l'ensemble des paires origine-destination dans le réseau.
- K_p : $(p \in P)$ l'ensemble. des chemins possibles entre l'origine et la destination de la paire p.
- K : l'ensemble des chemins possibles pour toutes les paires.
- h_k : le flot de trafic écoulé sur le chemin $k \in K$.

Avec ces notations, le routage des flots minimisant le coût du réseau est :

$$\begin{array}{lrlll} (ROUTAGE) & Min & g(x) & & \\ & s - \grave{a} : & \sum_{k \in K_p} h_k & = T_p & \forall\, p \in P \\ & & \sum_{p \in P} \sum_{k \in K_p} \delta_{ek} h_k & = x_e & \forall\, e \in E \\ & & x_e & \in \mathbb{R}^+ & \end{array}$$

où $\delta_{ek} = 1$ si l'arc e appartient au chemin k et 0 sinon.

Les flots x_e calculés dans ce modèle servent à dimensionner la capacité des liens. Ainsi, dans un réseau routier on dimensionnera la route de manière à ce qu'elle puisse absorber "sans congestion" un maximum de x_e véhicules.

Pour résumer schématiquement le processus, désignons par $(ORIGINES)$ et $(DESTINATIONS)$ les méthodes statiques permettant de générer O_i et D_j, la chaine de modélisation est représentée figure 1.

Le processus linéaire de la figure se complique quelque peu si l'on tient en compte le fait qu'en pratique les valeurs O-D T_p sont influencées par les flots x_e sur le réseau.

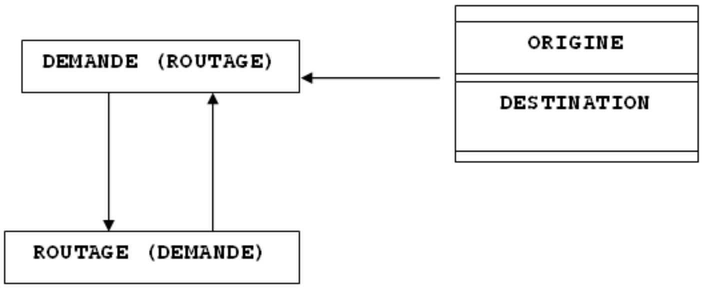

FIG. 2 –

Dans un réseau routier, pour une origine i et une destination j données, des véhicules n'ayant ni la même origine i ni la même destination j peuvent créer des congestions au niveau d'arcs se trouvant sur des chemins K_{ij} ; dissuadant ainsi les individus en i de se rendre en j. Le modèle statique de calcul des valeurs T_p devrait donc se raffiner en tenant compte du coût, noté u_p, pour se rendre de l'origine de p à sa destination. Comme ce coût dépend lui même des flots calculés, celui-ci est estimé de la manière suivante :

$$u_p = \min_{k \in K_p} \sum_{e \in E} \delta_{ek} c_e(x) \tag{3}$$

u_p est le coût minimal sur l'ensemble des chemins possibles pour la paire origine-destination p.

Tenant compte de ce coût, le calcul des valeurs O-D et la synthèse de réseau utilisent mutuellement leurs résultats de manière à déterminer la demande T_p et la synthèse optimale. Les deux problèmes liés l'un à l'autre sont :

$$\begin{array}{llll}
(DEMANDE(ROUTAGE)) & Max & -T_{ij} log(T_{ij}) - T_{ij} & \\
& s-\grave{a}: & \sum_{j=1}^{n} T_{ij} & = O_i \\
& & \sum_{i=1}^{n} T_{ij} & = D_j \\
& & \sum_{i,j} T_{ij} u_{ij} & = C \\
& & T_{ij} & \in \mathbb{R}^+
\end{array}$$

où C est un coût total observé.

$$\begin{array}{lllll}
(ROUTAGE(DEMANDE)) & Min & g(x) & & \\
& s-\grave{a}: & \sum_{k \in K_p} h_k & = T_p(u) & \forall\, p \in P \\
& & \sum_{p \in P} \sum_{k \in K_p} \delta_{ek} h_k & = x_e & \forall\, e \in E \\
& & x_e & \in \mathbb{R}^+ &
\end{array}$$

où $T_p(u)$ est obtenu en résolvant $(DEMANDE(ROUTAGE))$.

Des algorithmes itératifs ont été proposés pour résoudre $(DEMANDE(ROUTAGE))$.

Les inter-relations entre les différents modèles sont une nouvelle fois schématisées dans la figure 2.

3 Matrice O-D et synthèse dynamique

Les méthodes présentées précedemment peuvent s'envisager en considérant un aspect temporel. Les valeurs O_i et D_j donnant le nombre d'individus quittant ou arrivant à une zone i ou j sont dans la pratique fluctuant en fonction du temps. Une meilleure connaissance des valeurs temporelles permettrait alors une meilleure évaluation des demandes T_p et donc un meilleur dimensionnement du réseau.

Dans des travaux récents, J. El Ghordaf [3] s'est intéressée à modèliser la croissance de la population de deux régions sous la forme de deux équations différentielles ordinaires non linéaires du premier ordre.

$$\begin{cases} \frac{dx_1}{dt} & = k_1 x_1 (P_1 - x_1 - \beta x_2) \\ \frac{dx_2}{dt} & = k_2 x_2 (P_2 - x_2 - \beta x_1) \end{cases}$$

avec :
- x_i $(i = 1, 2)$: quantité de population dans les deux régions.
- k_i $(i = 1, 2)$: taux de naissance et d'immigration dans les deux régions.
- β : paramètre d'interdépendence entre les régions, typiquement les ressources d'une région utilisées par les habitants de l'autre région.
- P_i $(i = 1, 2)$: le potentiel de la région tenant compte du taux de mortalité et d'émigration du nombre maximal d'individus que la région peut abriter et de k_i.

Le modèle est ensuite étudié d'un point de vue théorique.

En prenant un nombre quelconque de régions (ou localisations) une telle modélisation reste encore possible et des solutions numériques pour une discrétisation donnée du temps peuvent être obtenues. Si i est une zone origine et j une zone destination une hypothèse (certes restrictive) que nous pouvons faire est que la quantité $x_i(t)$ de population est susceptible (dans sa totalité) de quitter la zone i et que la quantité $x_j(t)$ est susceptible d'y arriver. Sous ces hypothèses, les valeurs $x_i(t)$ et $x_j(t)$ apparaissent comme les données O_i et D_j mais variant en fonction du temps. Le problème de calcul de matrices O-D temporelles sur une période $t \in \{0, \overline{T}\}$ peut se reformuler alors de la manière suivante :

$$\begin{array}{llll} (DEMANDE(t)) & Max & -T_{ij}(t)log(T_{ij}(t)) - T_{ij}(t) & \\ & s - \grave{a} : & \sum_{j=1}^{n} T_{ij}(t) & = O_i(t) \\ & & \sum_{i=1}^{n} T_{ij}(t) & = D_j(t) \\ & & T_{ij}(t) & \in \mathbb{R}^+ \end{array}$$

et le calcul du routage se reformule en :

$$\begin{array}{lllll} (ROUTAGE(t)) & Min & g(x(t)) & & \\ & s - \grave{a} : & \sum_{k \in K_p} h_k(t) & = T_p(t) & \forall\, p \in P \\ & & \sum_{p \in P} \sum_{k \in K_p} \delta_{ek} h_k(t) & = x_e(t) & \forall\, e \in E \\ & & x_e(t) & \in \mathbb{R}^+ & \end{array}$$

Dans le problème de routage, les capacités attribuées au temps t, surtout dans le cas pratique d'un réseau routier, ne peuvent être complétement détruite au temps $t + 1$. Il convient donc d'ajouter une contrainte supplémentaire, $x_e(t) \geq x_e(t - 1)$ permettant l'extension du réseau et de transformer la première contrainte en inégalité de manière à conserver la réalisabilité du problème. Le problème serait donc pour chaque t :

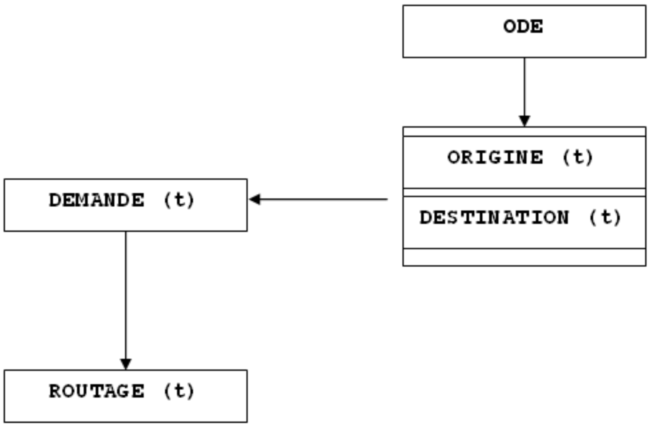

FIG. 3 – Matrice temporelle

$$
\begin{array}{lllll}
(ROUTAGE(t)) & Min & g(x(t)) & & \\
& s-a: & \sum\limits_{k \in K_p} h_k(t) & \geq T_p(t) & \forall\, p \in P \\
& & \sum\limits_{p \in P} \sum\limits_{k \in K_p} \delta_{ek} h_k(t) & = x_e(t) & \forall\, e \in E \\
& & x_e(t) & \geq x_e(t-1) & \\
& & x_e(t) & \in \mathbb{R}^+ &
\end{array}
$$

Comme illustré dans la figure 3, le processus ici reste linéaire.

Dans la pratique, il permettra de proposer une politique sur un horizon donné d'extension d'un réseau donné compte tenu d'une dynamique de population modélisée par ailleurs.

Le problème devient quelque peu plus difficile si l'on suppose que les valeurs $O_i(t)$ et $D_j(t)$ dépendent aussi du routage.

L'interdépendence β entre deux régions du modèle d'équation différentielle est certes fonction de l'intensité des services qu'une région offre à une autre mais est également lié à l'accessiblité à ces services. Si une région offre un nombre important de services mais qu'elle est difficilement accessible du fait de congestions régulières (dû à la configuration du réseau) ou de la distance pour l'atteindre cela influe nécessairement sur β. Dans le cas de congestion comme cela dépend du flot sur les chemins entre les régions, nous pouvons supposer que β est une fonction de x (flot issu du routage $(ROUTAGE(t))$) qu'il convient de choisir judicieusement. On peut par exemple considérer que pour deux régions i et j :

$$\beta(u) = \frac{nombre\ de\ services\ offerts}{u_{ij}}.$$

où u_{ij} est calculé grâce à la formule 3

Le sens de cette formule est que plus u_{ij} est petit (le coût des accès est faible), plus le nombre de service offerts est élevé, et plus les régions sont interdépendentes. Des régions s'offrant mutuellement beaucoup de services mais avec une valeur u_{ij} très grande (accès difficile à l'un et l'autre) seront moins interdépendentes.

Ceci donne lieu au système d'équation différentielle ordinaire suivant :

$$EDO(u) : \left\{ \begin{array}{ll} \frac{dx_1}{dt} & = k_1 x_1 (P_1 - x_1 - \beta(u) x_2) \\ \frac{dx_2}{dt} & = k_2 x_2 (P_2 - x_2 - \beta(u) x_1) \end{array} \right.$$

et donc au processus de synthèse de réseau représenté figure 4.

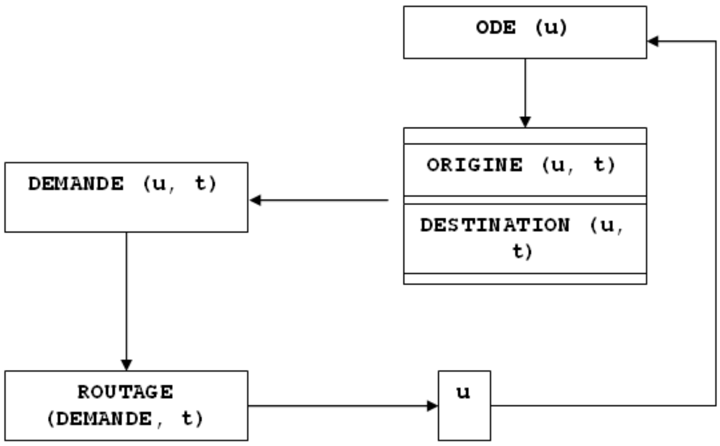

Fig. 4 – Matrice dynamique

Références

[1] M. Bierlaire. *Mathematical models for transportation demand analysis.* PhD thesis, Facultés Universitaires Notre-Dame de la Paix de Namur, 1996.

[2] G. Finke. *Recherche opérationnelle et réseaux : Méthodes d'analyse spatiale.* Aspects fondamentaux de l'analyse spatiale. Lavoisier, 2002.

[3] O. Arino J. El Ghordaf, M.H. Hbid. A mathematicla study of a two-regional population growth model. *C.R. Biologies*, 327 :977–982, 2004.