ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
*Série Mathématiques*

MOHAMED AHMED

**Modèles probabilistes pour choisir un itinéraire dans un réseau de trafic**

*Annales scientifiques de l'Université de Clermont-Ferrand 2*, tome 71, série *Mathématiques*, n° 20 (1982), p. 1-10
<http://www.numdam.org/item?id=ASCFM_1982__71_20_1_0>

# MODELES PROBABILISTES POUR CHOISIR UN ITINERAIRE DANS UN RESEAU DE TRAFIC

Mohamed AHMED
Université de PARIS VI

La probabilité de choisir certain itinéraire dans un réseau joue un rôle essentiel dans les modèles d'affectation stochastique du trafic, et celà provient du fait que ces modèles affectent les trafics aux itinéraires par :

$$M_k = P_k . M$$

où $M_k$ est le trafic affecté à l'itinéraire k, $P_k$ désigne la probabilité de choisir cet itinéraire , M est le trafic total à affecter en supposant que le réseau se compose d'une zone d'émission et d'une zone de destination de trafic et que ces zones sont liées par des itinéraires différents qui peuvent être disjoints ou chevauchés.

A cet effet, plusieurs modèles proposent différentes formes de la probabilité $P_k$.DIAL (1971) propose un modèle qui ne tient pas compte de la structure du réseau, par conséquent son modèle peut aboutir à des résultats érronés ; alors que DAGANZO (1977) tient compte de ce phénomène mais son modèle est limité par l'hypothèse essentielle que l'erreur d'estimation du temps de parcours jouit de la distribution normale.

MARTIN (1965) a proposé un modèle établissant la probabilité qu'un chemin particulier soit critique dans un graphe PERT, mais il a commis une erreur statistique lors de la dérivation du modèle.

Dans ce qui suit, on va introduire un modèle de trafic basé sur la technique de MARTIN après la modification de son modèle afin qu'il soit adapté aux conditions de trafic. Les hypothèses et les définitions, citées par DAGANZO (1971), seront employées en dérivant ce modèle. Ensuite, on va introduire un modèle général qui tient compte de la structure de réseau du trafic (défaut de modèle DIAL), et qui n'exige pas de distribution particulière concernant le temps de déplacement (défaut du modèle DAGANZO et de notre modèle précédent).

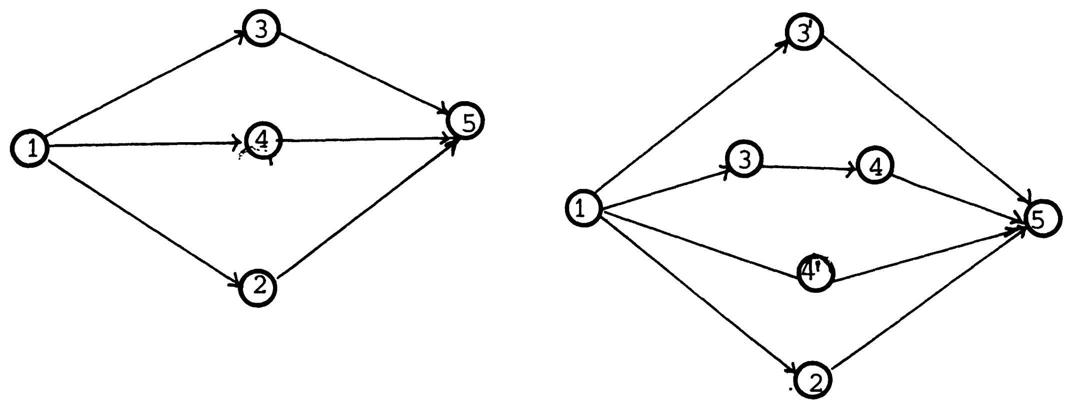
## I - DEFINITIONS ET NOTIONS :

On appelle réseau simple le réseau qui est composé de deux zones principales, l'une d'où part tout le trafic et l'autre d'où il parvient. En effet, les itinéraires liant ces zones sont de deux genres (AHMED 1979) :

1. Itinéraires simples dont les arcs ne sont pas inclus dans d'autres itinéraires.
2. Itinéraires complexes dont un ou plusieurs arcs sont inclus dans d'autres itinéraires.

Il est possible de transformer un réseau quelconque en un réseau simple en utilisant ce que l'on appelle "des arcs duaux" à la place de ceux qui sont contenus dans plusieurs itinéraires.

A titre d'exemple, étant donné un réseau comme ci-après, on peut le transformer en réseau simple en utilisant les arcs duaux (1,3') et (4',5).




soit,

- K l'ensemble d'itinéraires simples ;
- $t_k$ , $\forall k \in K$, v.a correspondent au temps de déplacement prévu par un usager au moment de choisir son itinéraire. L'usager sera supposé avoir des connaissances préalables sur tous autres itinéraires ;
- $T'_k$ , $\forall k \in K$, le temps mesuré de parcours d'itinéraire k, autrement dit le temps réalisé par l'usager. Bien entendu que, même sous des conditions identiques de trafic, $T'_k$ peut varier de temps à autre, et par conséquent on choisit la moyenne.

## II - CONSTITUTION DU MODELE :

Etant donné un réseau simple, muni de l'ensemble des arcs duaux 1 ; on s'intéresse actuellement à établir la probabilité de choisir un itinéraire particulier dans ce réseau.

Il découle de ce qui a été dit aux définitions que :

$$t_k = T'_k + \varepsilon_k \quad , \quad \forall k \in K \tag{1}$$

où $\varepsilon_k$ est l'erreur d'estimation, supposée avoir une distribution normale avec

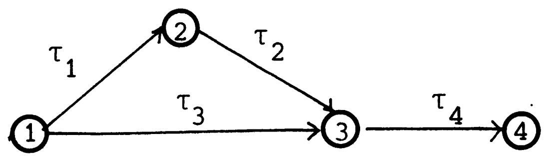
$$E(\varepsilon_k) = 0 \quad \text{et} \quad V(\varepsilon_k) << \infty \quad , \quad \forall k \in K.$$

Dans ce modèle on va adopter la façon de choisir un itinéraire particulier citée par DAGANZO (1977). Il s'agit que chaque usager apprécie d'abord le temps de chaque itinéraire puis qu'il choisisse le plus court (le moins coûteux). Il est hors de notre but d'étudier les formules mesurant le temps (le coût) de déplacement mais ces formules seront supposées données. Soit $t_k$ le temps prévu de l'itinéraire k, donc en vertu de cette hypothèse on a que

$$t_k \le t_i \quad , \quad \forall i \ne k.$$

Soit $t = (t_1, \ldots\ldots, t_K)$. Il faut signaler que c'est à cause de chevauchement d'itinéraires que t est un vecteur de multivariables dépendantes, c'est-à-dire que les arcs duaux font $t_i$ dépendantes.

Pour démontrer cette idée, considérons le réseau ci-desous composé de deux itinéraires (1-2-3-4) et (1-3-4) qui sont réellement chevauchés par le tronçon (3-4).




$\tau_1, \ldots, \tau_4$ correspondent aux variables du temps des tronçons et qu'elles sont supposées indépendantes. En effet, les variables $t_1$ et $t_2$ où

$$t_1 = \tau_1 + \tau_2 + \tau_4 \quad \text{et} \quad t_2 = \tau_3 + \tau_4$$

sont dépendantes à cause de l'existance de $\tau_4$.

En retour à notre modèle, soit $X = (x_1, \ldots, x_l)$ le vecteur de variables du temps pour l'ensemble l des arcs duaux et f(X) la fonction de la densité de probabilité correspondant.

Le fait que chaque itinéraire peut se composer d'arcs duaux et d'autres arcs non duaux permet de remplacer (1) par

$$t_k = (T_k + x'_k) + \varepsilon_k \quad , \quad \forall k \in K$$

où,

$T_k$ correspond au temps mesuré sur les arcs non duaux figurant dans l'itinéraire k ;

$x'_k$ correspond au temps mesuré sur les arcs duaux figurant dans l'itinéraire k.

Cela démontre que :

$$E(t_k) = T_k + x'_k \quad , \quad \forall k \in K \tag{2}$$

$$V(t_k) = V(\varepsilon_k) \quad , \quad \forall k \in K \tag{3}$$

En supposant les éléments du vecteur X fixés mais non spécifiés, donc t devient un vecteur de multivariables indépendantes (si on pose $\tau_4$ fixe dans l'exemple précédent, donc $t_1$ et $t_2$ deviennent indépendants). Afin de distinguer les notations, on désigne $t^X = (t_1, \ldots, t_K \mid X)$ dont l'espérance et la variance des composants sont données par (2) et (3) et que

$$\mathrm{Cov}\,(t_i, t_j \mid X) = 0.$$

A partir de cela, la probabilité cherchée est donnée par :

$$P_k = \Pr.(t_k \leq t_1, \ldots, t_k \leq t_K) = \int \Pr(t_k \leq t_1, \ldots, t_k \leq t_K \mid X).\ f(X)dX \quad (4).$$

Il faut avouer que nous utilisions la technique de MARTIN (1965) pour arriver à (4), mais nous avons quelques remarques à signaler avant de développer (4). MARTIN a établit un modèle qui donne la probabilité qu'un chemin particulier dans un graphe PERT soit critique. En effet nous modifions partiellement sa formule de probabilité pour qu'elle soit adaptée au problème de trafic. En général cette modification concerne deux choses particulières :

1. La dérivation de $\Pr.(t_k \leq t_i, \forall i \neq k)$ au lieu de $\Pr.(t_k \geq t_i, \forall i \neq k)$ et de rendre le modèle de MARTIN adéquate au problème de trafic.

2. La correction de l'erreur commise par MARTIN lors de la dérivation de son modèle ; car MARTIN considère vraie la proposition suivante :

$$\Pr.(t_k - t_1 \geq 0, \ldots, t_k - t_K \geq 0 \mid X) = \prod_{i=1;i\neq k}^{K} \Pr.(t_k - t_i \geq 0 \mid X)$$

mais une telle proposition est statistiquement inexacte parce que les évènements $(t_k - t_i \geq 0 \mid X \; ; \; \forall i \in K)$ sont dépendants même si les variables composant sont indépendantes.

Retournons à notre modèle, précisément à (4), et considérons la probabilité :

$$\Pr.(t_k \leq t_1, \ldots, t_k \leq t_K \mid X) \qquad (5)$$

Pour développer (5) nous utilisons la technique de DAGANZO (1977).

Soit : $t_k^X = (t_k - t_1, \ldots, t_k - t_K \mid X)$ ; et $L_k$ une matrice d'ordre $K \times (K-1)$ définie par :

$$L_k = \begin{bmatrix} -1 & 0 \ldots\ldots 0 & 0 \\ 0 & -1 \ldots\ldots 0 & 0 \\ \vdots & \vdots \quad\quad\;\; \vdots & \vdots \\ 1 & 1 \ldots\ldots 1 & 1 \\ \vdots & \vdots & \\ 0 & 0 \ldots\ldots -1 & 0 \\ 0 & 0 \ldots\ldots 0 & -1 \end{bmatrix} \quad \text{la ligne } k$$

Or le vecteur $t_k^X$ peut s'exprimer ainsi

$$t_k^X = t^X \, . \, L_k.$$

Il est évident que les éléments de $t_k^X$ sont des composants linéaires de variables normales, par conséquent $t_k^X$ jouit d'une distribution de multivariables normales caractérisée par :

$$E(t_k^X) = E(t^X) \, . \, L_k \; ;$$

$$V(t_k^X) = L_k' \, . \, V(t^X) \, . \, L_k$$

où $L_k'$ est la transposition de $L_k$ et que

$$V(t^X) = \begin{bmatrix} V(\varepsilon_1) & 0 \;\ldots\ldots & 0 \\ 0 & V(\varepsilon_2) & \\ \vdots & & \\ 0 & 0 \;\ldots\ldots & V(\varepsilon_K) \end{bmatrix}$$

Donc, (5) s'établit ainsi

$$\Pr.(t_k^X \leq (0,\ldots,0) = \int_{-\infty}^{0} (K-1) \text{ fois} \int_{-\infty}^{0} \frac{\exp\left[-\frac{1}{2}T'.V^{-1}(t_k^X).T\right]}{(2\Pi)^{\frac{K-1}{2}} . |V(t_k^X)|^{\frac{1}{2}}} dt_k^X \ ; \quad (6)$$

où

$$T = t_k^X - E(t_k^X).$$

On ne peut pas transformer $\left[T'.V^{-1}(t_k^X).T\right]$ , qui est actuellement en forme quadratique, en une forme canonique par l'emploi d'une transformation orthogonale, car une telle transformation n'aboutira pas à une séparation des domaines d'intégrales données en (6). La probabilité non conditionnée de choisir l'itinéraire k sera donnée par (4) en remplaçant

$\Pr.(t_k \leq t_1,\ldots, t_k \leq t_K \mid X)$ par (6).

Enfin, il faut avouer que l'importance de ce modèle soit plutôt théorique que pratique car la formule de probabilité conclue n'est pas facile à calculer au cas de l'application sur un réseau comportant d'arcs duaux.

## III - MODELE GENERAL

Dans le but d'établir un modèle général qui ne pose préalablement pas de distribution particulière au temps d'itinéraires, nous introduisons le lemme et la propostion suivants :

LEMME :

Soit x, y, z, variables aléatoires quelconques. Si x et y sont conditionnellement indépendantes et que la condition soit sur z, alors :

$$f(x,y|z) = f(x|z).f(y) \text{ (voir référence 7).}$$

PROPOSITION :

Soit $x_i$, $\forall i \in N$, v.a indépendantes dont $f_i$ sont les p.d.f. Si $x_j$, $j \in N$, est supposée fixée (égale à $\tau$ ), alors

$$\Pr.(x_1 \leq x_j,\ldots,x_N \leq x_j \mid x_j=\tau) = \prod_{i\neq j}^{N} \Pr.(x_i \leq x_j | x_j = \tau)$$

car les évènements sont conditionnellement indépendants.

Il est intéressant de relier les résultats précédents à la formule essentielle (4). A part, considérons la quantité :

$$Pr.(t_1 \geq t_k, \ldots, t_K \geq t_k \mid X)$$

où $t_i$, $\forall i \in K$, sont conditionnellement indépendantes mais les évènements $(t_i \geq t_k, \forall i \in K)$ ne le sont pas.

Du point de vue statistique on a que :

$$Pr.(t_1 \geq t_k, \ldots, t_K \geq t_k \mid X)$$

$$= \int Pr.(t_1 \geq t_k, \ldots, t_K \geq t_k \mid t_k = \tau, X) f_k(\tau \mid X) d\tau \qquad (7)$$

En effet les évènements inclus dans la quantité

$$Pr.(t_1 \geq t_k, \ldots, t_k \geq t_k \mid t_k = \tau, X) \qquad (8)$$

deviennent, grâce à la condition $t_k = \tau$ , conditionnellement indépendants, alors en vertu de la proposition précédente on peut exprimer (8) par

$$\prod_{i \neq k}^{K} Pr.(t_i \geq t_k \mid t_K = \tau, X) = \prod_{i \neq k}^{K} \left[1 - F_i(\tau \mid X)\right].$$

Finalement, la probabilité en question (4) s'établit ainsi :

$$P_k = Pr.(t_1 \geq t_k, \ldots, t_K \geq t_k)$$

$$= \int \{ \int \prod_{i \neq k}^{K} [1 - F_i(\tau \mid X)] \, . \, f_k(\tau \mid X) d\tau \} . f(X) dX \qquad (9)$$

où $f_k$ est la fonction de la densité de $t_k$.

Il est indispensable que (9) satisfasse à $\sum_{k=1}^{K} P_k = 1$ dont le raisonnement sera présenté dans l'appendice.

## IV - CONCLUSION :

De ce qui précède, on s'aperçoit que (9) est plus inclusive que la formule conclue en premier modèle, car (9) ne pose pas de conditions sur le type de la loi de probabilité employée. Mais, malheureusement, la comparaison entre les résultats de ces modèles ne peut pas être conclue même si ces modèles sont appliqués à un réseau simple. Les difficultés viennent du fait que, d'une part, le calcul du modèle général devient extrêmement difficile dans le cas où la distribution soit normale, car il s'agit de calculer $\int_{-\infty}^{\tau} f(x)dx$ où $f(x)$ est la fonction de la densité de la distribution normale et que $\tau$ est une réelle inconnue ;

d'autre part les mêmes difficultés peuvent être rencontrées lorsque l'on applique la formule du premier modèle à un réseau comportant les arcs duaux.

V - APPENDICE

On montrera que $\sum_{k=1}^{K} P_k = 1$.

Pour l'abréviation on note :

$$f_i(\tau|X) = f_i \qquad \text{et} \qquad F_i(\tau|X) = F_i.$$

On a manifestement que :

$$\sum P_k = \iint \Big[ \sum_{k=1}^{K} \int \prod_{i\neq k}^{K} (1-F_i).f_k d\tau \Big] . f(X)dX \tag{10}$$

Considérons d'abord la quantité :

$$\sum_{k=1}^{K} \int \prod_{i\neq k}^{K} (1-F_i).f_k \, d\tau \tag{11}$$

Cette quantité se partage en

$$\int \prod_{i=2}^{K} (1-F_i)f_1 \, d\tau \;+\; \sum_{k=2}^{K} \int \prod_{i\neq k}^{K} (1-F_i)f_k \, d\tau \tag{12}$$

L'intégration par partition de la première partie de (12) se ramène aux deux termes suivants, dont le premier s'annule

$$F_1 . \prod_{i=2}^{K} (1-F_i) \Big|_{-\infty}^{\infty} + \int F_1 . \sum_{k=2}^{K} f_k \prod_{\substack{\forall i\neq k \\ \neq 1}} (1-F_i)d\tau$$

$$= -\int (1-F_1) . \sum_{k=2}^{K} f_k \prod_{\substack{\forall i\neq k \\ \neq 1}} (1-F_i)d\tau \;+ \int \sum_{k=2}^{K} f_k . \prod_{\substack{\forall i\neq k \\ \neq 1}} (1-F_i)d\tau$$

$$= -\int \sum_{k=2}^{K} \prod_{i\neq k}^{K} (1-F_i).f_k \, d\tau + \int \sum_{k=2}^{K} \prod_{\substack{\forall i\neq k \\ \neq 1}}^{K} (1-F_i).f_k d\tau \tag{13}$$

En ajoutant la deuxième partie de (12) à (13), donc (11) sera égale à :

$$\sum_{k=2}^{K} \int f_k . \prod_{\substack{\forall i\neq k \\ \neq 1}} (1-F_i) \, d\tau$$

De la même façon on peut continuer par cet algorithme jusqu'à ce que (11) soit égale à :

$$\sum_{k=K-2}^{K} \int f_k . \prod_{\substack{\forall i\neq k \\ \neq 1,2,\dots,(K-3)}} (1-F_i) \qquad d\tau$$

ce qui est à son tour égale à :

$$\sum_{k=K-1}^{K} \int f_k \prod_{\substack{\forall i \neq k \\ \neq 1,2,\ldots,(k-2)}} (1-F_i) \, d\tau$$

$$= \int f_{K-1} \cdot (1-F_K) \, d\tau + \int f_K \cdot (1-F_{K-1}) d\tau$$

$$= F_{K-1} \cdot (1-F_K) \Big|_{-\infty}^{\infty} + \int F_{K-1} \cdot f_K \, d\tau + \int f_K d\tau - \int f_K \cdot F_{K-1} \, d\tau .$$

Le premier terme s'annule et le deuxième s'annule aussi avec le quatrième ; alors il ne reste que

$$\int f_K \, d\tau = 1.$$

Donc (10) sera égale à

$$\int 1 . f(X) dX = 1.$$

## BIBLIOGRAPHIE

[1] AHMED M. (1979), "Etude Critique de certains Modèles de Trafic Urbain avec diverses variants et application". Thèse de 3ème Cycle, Université Pierre et Marie Curie, Paris VI.

[2] DIAL R. (1971), "A probabilistic Multipath Traffic Assignment Model which obviates path enumeration", Transportation Research, Vol. 5.

[3] DIAL et autres (1971), "A multipath Traffic Assignment Model", Highway Research Record, n° 369.

[4] DAGANZO C. (1977), "On the traffic assignment problem with flow dependent costs - I", Transportation Researchs, vol II.

[5] DAGANZO C. et SHEFFI Y. (1977), "On stochastic models of traffic assignment", Transportation Science, vol II n° 3.

[6] MARTIN J. (1965), "Distribution of the time through a directed acyclic network", Operations Research, vol. 13, N°1.

[7] PAPOULIS A., "Probability, random variables and stochastic Processes", Mc Graw-Hill Kogakusha, Ch. 8, p. 238.

---------