RAIRO Rech. Opér.
(vol. 33, n° 1, 1999, pp. 69-86)

CONTRÔLE DYNAMIQUE DE FLUX DANS UN SYSTÈME D'ATTENTE AVEC PANNE (*)

par A. HAQIQ ([1]) et N. MIKOU ([2])

Communiqué par J.P. BRANS

Résumé. – *On considère deux files d'attente M/M/1 en parallèle. Une des files, notée 1, est assujettie à des pannes intermittentes. À l'aide de la théorie de la programmation dynamique, on montre l'existence d'une politique optimale à seuil, qui consiste à transférer, quand il le faut, les clients qui arrivent à la file 1 vers l'autre file afin de minimiser un coût instantané dépendant des longueurs des deux files.*

Mots clés : Système d'Attente, Programmation Dynamique, Processus Markovien de Décision, Politique Optimale.

Abstract. – *We consider two parallel M/M/1 queues. The server at one of the queues is subject to intermittent breakdowns. By the theory of dynamic programming, we determine a threshold optimal policy which consists to transfer, when it is necessary, the customers that arrive at the first queue towards the second queue in order to minimize an instantaneous cost depending of the two queue lengths.*

Keywords: Queuing System, Dynamic Programming, Markov Decision Processes, Optimal Policy.

1. INTRODUCTION

Dans les problèmes de contrôles de file d'attente, la recherche d'une politique optimale peut être relativement simple si on sait que cette politique possède une structure particulière, comme par exemple, la politique « bang-bang » ou la politique « à seuil » [11].

Les modèles de file d'attente constitués de deux files d'attente recevant des clients d'un ou plusieurs flux avec un temps de séjour contrôlé ont été étudiés dans plusieurs articles (*cf.* [2-5, 12, 15]).

(*) Reçu en mars 1996.
([1]) Faculté des Sciences et techniques, Département de Mathématiques et Informatique, Route de Casablanca, BP 577, Settat, Maroc, Tél. 03 40.07.36, Fax 03 40.09.69.
([2]) Faculté des Sciences, Département de Mathématiques et Informatique, avenue Ibn Batouta, BP 1014, Rabat, Maroc, Tél. 07 77.54.71, Fax 07 77.54.71.

Recherche opérationnelle/Operations Research, 0399-0559/99/01/$ 7.00

Dans [8], Lin et Kumar ont trouvé l'utilisation optimale, sans préemption, de deux services exponentiels dans une file d'attente avec des arrivées Poissoniennes. La politique optimale trouvée est à seuil.

Dans le cas des systèmes d'attente constitués de deux files pour lesquels on contrôle les flux d'arrivées, Foschini et Salz [3] ont montré l'efficacité des politiques « Feedback » dans le cas de forte charge.

En supposant que la distribution de temps de service est la même pour les deux files, Winston [15], Weber [12], Ephremides, Varaiya et Walrand [2] ont montré qu'envoyer les clients dans la file la moins chargée est une politique fortement optimale.

Ghonein [4] a établi l'optimalité des contrôles de commutation pour un modèle similaire dans lequel les arrivées hétérogènes sont contrôlées dans chacune des deux files, et les clients de l'une des deux entrent dans l'autre après leur service.

En 1984, Hajek [5] a étendu les études précédentes, en supposant que les distributions du temps de service sont différentes pour les deux files et en contrôlant les deux services avec possibilité de transition des clients d'une file à l'autre après leur service. Pour ce modèle, il a montré l'existence d'une politique optimale à seuil.

Dans ce papier, on considère le problème de contrôle de flux d'arrivée dans un système d'attente constitué de deux files comme dans le modèle de Foschini et Salz [3] mais auquel on ajoute les hypothèses suivantes :

1) un processus d'arrivée vers la deuxième file de clients non contrôlés ;

2) des processus de panne et de réparation dans le service de la première file ;

3) un coût supplémentaire quand un client du flux contrôlé passe dans la deuxième file.

On associe à ce modèle un problème de contrôle markovien d'horizon fini N et de coût actualisé total avec une fonction coût instantané dépendant des tailles des deux files, et on montre que la fonction valeur à l'étape N possède des propriétés structurelles. L'existence et l'unicité de la fonction valeur limite sont établies, cette dernière fonction conservant les mêmes propriétés structurelles, ce qui donne une caractérisation de la politique optimale.

Une motivation du système d'attente considéré est le problème de routage dynamique dans les systèmes informatiques ou dans les réseaux de télécommunication.

2. PROCESSUS MARKOVIEN DE DÉCISION (PMD)

Soit S un ensemble dénombrable et soit A un ensemble fini.

On considère le processus aléatoire $(Y_n)_{n\geq 1}$ à temps discret et à valeurs dans l'espace d'états S. Afin de contrôler ce processus, on dispose d'un jeu de manœuvres modélisé par l'espace A. À chaque instant $n = 1, 2, ...,$ après observation de l'état du processus, une décision est prise en appliquant une action représentée par la variable aléatoire A_n. L'impact du contrôle $(A_n)_{n\geq 1}$ sur l'évolution du système est illustré en définissant l'histoire du processus $(Y_n)_{n\geq 1}$ jusqu'à l'instant n par :

$$H_n = \{Y_1, A_1, ..., Y_{n-1}, A_{n-1}, Y_n\}, \qquad n \geq 2$$
$$H_1 = Y_1,$$

A_n est l'action choisie à l'instant n en fonction de H_n, après que Y_n ait été réalisé.

Nous réquérons que la dynamique du processus soit Markovienne en demandant que l'action de A_n agisse sur le prochain état Y_{n+1} seulement en fonction de l'état présent, indépendamment du passé. En d'autres termes, nous supposerons que la loi de transition est décrite par les probabilités :

$$P(Y_{n+1} = j | Y_n = i, A_n = a) = P(Y_{n+1} = j | H_n, Y_n = i, A_n = a). \quad (1)$$

On notera :

$$p_{ij}(a) = P(Y_{n+1} = j | Y_n = i, A_n = a).$$

DÉFINITION 1 : *On dit que* $(Y_n)_{n\geq 1}$ *forme un Processus Markovien de Décision si :*

• *à chaque instant de décision* n *un coût instantané est encouru, ce coût ne dépend que de l'état présent et de l'action choisie et est noté par* $C(Y_n, A_n)$*,*

• *la dynamique du système satisfait l'équation* (1) *et* $\sum_{j\in S} p_{ij}(a) = 1$ *pour tout* $i \in S$.

La façon dont on choisit une action doit suivre une règle précise.

On définit la règle de décision à l'instant n comme étant une loi de probabilité conditionnelle $\pi(\,.\,|H_n)$ définie sur A, adaptée à la tribu $\boldsymbol{H}_n$ engendrée par H_n.

Une politique admissible (ou tout simplement une politique) π est une suite de règle de décision : $\pi = (\pi_1, \pi_2, ..., \pi_n, ...)$.

On désignera par Π l'ensemble des politiques admissibles.

DÉFINITION 2 : *Une politique est dite Markovienne ou sans mémoire si :*

$$\pi_n\,(\,.\,|H_n) = \pi_n\,(\,.\,|Y_n), \qquad \textit{pour tout} \quad n \geq 1$$

DÉFINITION 3 : *Une politique sans mémoire est dite stationnaire si* $\pi_n\,(\,.\,|Y_n)$ *est indépendante de* n.

On pose :

$$\pi_n\,(\,.\,|Y_n = i) = m_f(\,.\,|i) \qquad \text{pour tout} \quad i \in S.$$

Une politique stationnaire f s'écrit donc : $f = (m_f,\, m_f, ...)$.

3. DESCRIPTION DU MODÈLE

On considère un système d'attente constitué de deux files en parallèle. Deux flux de clients arrivent à ces files suivant deux processus de Poisson de paramètres distincts $\lambda_i\ (i = 1,\, 2)$. Lorsque les clients du flux 1 arrivent au système, ils trouvent devant eux un contrôleur qui les envoie soit à la file 1 soit à la file 2 suivant une certaine règle de décision, tandis que les clients du flux 2, lorsqu'ils arrivent, entrent directement dans la file 2. On remarque donc que seuls les clients du flux 1 sont contrôlés. Les temps de service des clients dans chacune des deux files sont des variables aléatoires indépendantes distribuées exponentiellement, avec un taux $\mu_i\ (i = 1,\, 2)$.

Les durées de fonctionnement normal et de réparation dans le service de la première file forment un processus de renouvellement alterné, les lois de ces durées étant exponentielles de paramètre respectivement α et β.

3.1. Formulation du problème

• L'état du système est décrit par le triplet de variables aléatoires $(X_t^1,\, X_t^2,\, Z_t)$, où X_t^i est le nombre de clients dans la file i à l'instant t, $i = 1, 2$ et

$$Z_t = \begin{cases} 0 & \text{lorsqu'il y a panne} \\ 1 & \text{sinon} \end{cases}$$

• l'espace d'états S est défini par :

$$S = \mathbb{N}^2 \times \{0,\, 1\},$$

• l'espace d'actions est :

$$A = \{0,\, 1\}.$$

Si l'élément a de A est nul (respectivement égal à 1) alors le contrôleur envoie le client du flux 1 qui arrive vers la file 1 (respectivement vers la file 2).

• Soit $K = \{(x, a), x \in S, a \in A(x)\}$ où $A(x) \subset A$ désigne l'espace des actions disponibles au contrôleur lorsque le système est dans l'état x.

• Soit la fonction $\lambda : K \mapsto \mathbb{R}$ où $\lambda(x_1, x_2, z; a)$ représente le taux de sortie de l'état (x_1, x_2, z) quand l'action a est choisie.

λ est donné par :

$$\lambda[(0, 0, 0); a] = \lambda_1 + \lambda_2 + \beta;$$
$$\lambda[(0, 0, 1); a] = \lambda_1 + \lambda_2 + \alpha;$$

pour $x_1 \geq 1$ on a :

$$\lambda[(x_1, 0, 0); a] = \lambda_1 + \lambda_2 + \beta;$$
$$\lambda[(x_1, 0, 1); a] = \lambda_1 + \lambda_2 + \mu_1 + \alpha;$$

pour $x_2 \geq 1$ on a :

$$\lambda[(0, x_2, 0); a] = \lambda_1 + \lambda_2 + \mu_2 + \beta;$$
$$\lambda[(0, x_2, 1); a] = \lambda_1 + \lambda_2 + \mu_2 + \alpha;$$

pour $x_1, x_2 \geq 1$ on a :

$$\lambda[(x_1, x_2, 0); a] = \lambda_1 + \lambda_2 + \mu_2 + \beta;$$
$$\lambda[(x_1, x_2, 1); a] = \lambda_1 + \lambda_2 + \mu_1 + \mu_2 + \alpha.$$

• Soit P le noyau stochastique défini sur S étant donné K.

Pour $x = (x_1, x_2)$, $y = (y_1, y_2)$ dans $\mathbb{N}^2$ et pour z, z_1 dans $\{0, 1\}$, P est donné par :

$$\begin{aligned} P[(y, z_1)|(x, z); a] \times \lambda[(x, z); a] = {} & \lambda_1 \cdot \mathbf{1}_{\{y=A_1 x, z_1=z\}} \\ & + \lambda_1 . \mathbf{1}_{\{y=A_2 x, z_1=z, a=1\}} + \lambda_2 . \mathbf{1}_{\{y=A_2 x, z_1=z\}} \\ & + \mu_1 . \mathbf{1}_{\{y=D_1 x, z_1=z=1, x_1 \neq 0\}} + \mu_2 . \mathbf{1}_{\{y=D_2 x, z_1=z, x_2 \neq 0\}} \\ & + [\alpha . z + \beta . (1 - z)] \times \mathbf{1}_{\{y=x, z_1 \neq z\}} \end{aligned}$$

où

$$D_1\,x=((x_1-1)^+,\,x_2);\quad A_1\,x=(x_1+1,\,x_2)\,;$$
$$D_2\,x=(x_1,\,(x_2-1)^+);\quad A_2\,x=(x_1,\,x_2+1)$$

et

$$x^+=\max(0,\,x).$$

• La fonction coût instantané $C(\,.\,,\,.\,)$ est donnée par :

$$C\,(X_t^1,\,X_t^2,\,Z_t;\,A_t)=c_1\,.\,X_t^1+c_2\,.\,X_t^2+c_3\,.\,\mathbf{1}_{\{A_t=1\}}$$

où A_t est la variable aléatoire donnant l'action à l'instant t et les c_i $(i=1,\,2,\,3)$ sont des constantes positives.

c_3 est un coût supplémentaire que doit payer un client du flux 1 lorsqu'il passe à la file 2.

3.2. Objectif

Notre objectif est de décider à quelle file chaque client arrivant (du flux 1) devra être envoyé afin de minimiser le coût actualisé total suivant :

$$E_{X_0,\pi}\left[\int_0^\infty e^{-\delta\,.\,t}\,.\,(c_1\,.\,X_t^1+c_2\,.\,X_t^2+c_3\,.\,\mathbf{1}_{\{A_t=1\}})\,dt\right]$$

où $X_0=(x_0^1,\,x_0^2,\,z_0)$ est l'état initial du processus $X_t=(X_t^1,\,X_t^2,\,Z_t)$ et δ est le facteur d'actualisation $(0<\delta<\infty)$.

$E_{X_0,\pi}$ dénote l'espérance sous la politique π conditionnellement à l'état X_0.

Si on note par $(T_n)_{n\geq 1}$ la suite des instants de passage des clients du flux 1 à la file 2 alors

$$\int_0^\infty c_3\,.\,e^{-\delta\,.\,t}\,\mathbf{1}_{\{A_t=1\}}\,dt=\sum_{n=1}^\infty c_3\,.\,e^{-\delta\,.\,T_n}\,.\,\mathbf{1}_{\{A_{T_n}=1\}}$$

et le coût actualisé total à minimiser s'écrit alors sous la forme suivante :

$$E_{X_0,\pi}\left[\int_0^\infty e^{-\delta\,.\,t}\,.\,(c_1\,.\,X_t^1+c_2\,.\,X_t^2)\,dt+\sum_{n=1}^\infty c_3\,.\,e^{-\delta\,.\,T_n}\,.\,\mathbf{1}_{\{A_{T_n}=1\}}\right].$$

Dans toute la suite de l'article, E_{X_0} désignera $E_{X_0,\,\pi}$.

4. EXISTENCE DE POLITIQUES OPTIMALES À SEUIL

Proposition 1 : *Il existe une fonction non décroissante* $s^0(\,.\,)$ *(resp.* $s^1(\,.\,)$) *définie sur* $\{0, 1, 2, ...\}$ *et à valeurs dans* $\{0, 1, 2, ...\}$ *telle que la politique optimale est d'envoyer un client qui arrive, quand les longueurs des deux files sont* x_1 *et* x_2*, à la file 1 si* $x_2 \geq s^0(x_1)$ *dans le cas de panne (respectivement* $x_2 \geq s^1(x_1)$ *dans le cas de réparation de la panne) et à la file 2 sinon.*

Avant de démontrer la proposition 1, il est nécessaire de rappeler certains résultats, sur le problème de coût non actualisé.

Soit $(Y_n)_{n \geq 0}$ un PMD d'espace d'états S, et soit $(A_n)_{n \geq 0}$ une suite de variables aléatoires (v.a.) à valeurs dans A donnant les actions choisies par le contrôleur aux instants de décision n.

Le problème consiste à chercher une politique admissible $\pi \in \Pi$ qui minimise la fonctionnelle :

$$V_N(i, \pi) = E_\pi \left[\sum_{n=1}^{N} C(Y_n, A_n) | Y_1 = i \right]$$

où E_π dénote l'espérance sous la politique π, $N \in \mathbb{N}$ et $C(Y_n, A_n)$ est le coût instantané encouru par le PMD à l'instant n.

Soit

$$V_N(i) = \min_{\pi \in \Pi} V_N(i, \pi). \tag{2}$$

Proposition 2 : (i) *Pour tout* $i \in S$ *on a* :

$$\begin{cases} V_N(i) = \min_{a \in A(i)} \{C(i, a) + \sum_{j \in S} p_{ij}(a) \,.\, V_{N-1}(j)\}, \quad \forall N \geq 1 \\ avec \quad V_0(\,.\,) = 0 \end{cases} \tag{3}$$

(ii) *Si* $C(\,.\,,\,.\,) \geq 0$ *et* A *est fini, alors* :

$$\lim_{N \to +\infty} V_N(i) = V(i) \stackrel{\text{not}}{=} V_\infty(i) \tag{4}$$

et V *est l'unique solution bornée de l'équation de la Programmation Dynamique (E.P.D.) suivante* :

$$V(i) = \min_{a \in A(i)} \{C(i, a) + \sum_{j \in S} p_{ij}(a) \,.\, V(j)\} \tag{5}$$

(iii) *Sous les hypothèses de l'assertion* (ii), *la politique optimale existe et elle est stationnaire.*

Pour la démonstration de cette proposition, *voir* ([1, 7, 9, 14]).

Soit H l'ensemble des fonctions définies sur S et à valeurs dans $\mathbb{R}$, contenant la fonction identiquement nulle et défini de la manière suivante : si $f \in H$, alors

$$g(i) = \min_{a \in A(i)} \{C(i, a) + \sum_{j \in S} p_{ij}(a) \cdot f(j)\} \in H.$$

THÉORÈME 1 : *Si* $C(\,.\,,\,.\,) \geq 0$ *et* A *est fini et si l'ensemble* H *défini ci-dessus est fermé pour les limites ponctuelles (*i.e., *la limite ponctuelle de toute suite de fonctions* f_n *de* H *est une fonction de* H*) alors* $V \in H$.

Ce résultat se déduit immédiatement de (4), (*voir* [7, 13]).

Preuve de la proposition 1 : La preuve se fait en quatre étapes.

• 1[re] *étape* : **Réduction à un problème de contrôle à temps discret et détermination de la fonction valeur à l'étape *n*.**

1. Réduction à un problème de contrôle Markovien à temps discret :

Soit un temps aléatoire T de loi exponentielle de paramètre δ indépendant du processus $(X_t, A_t)_t$.

On considère le processus $Y_t = (Y_t^1, Y_t^2, Z_t)$ défini par :

$$\begin{cases} Y_t^1 = X_t^1, \quad Y_t^2 = X_t^2 \qquad \text{pour} \quad t < T. \\ Y_t^1 = Y_t^2 = 0 \qquad \text{pour} \quad t \geq T \end{cases}$$

et $(t_n)_n$ la suite des instants de saut du processus $(Y_t)_t$.

On pose : $t_0 = 0$ et on note : $Y_{t_n} = Y_n$.

Il est évident que $(Y_n)_n$ est un processus de Poisson de paramètre $\delta + \gamma$, où

$$\gamma = \lambda_1 + \lambda_2 + \mu_1 + \mu_2 + \alpha + \beta.$$

On suppose que les fonctions coûts instantanés associées aux processus X_t et Y_t sont égales pour tout $t < T$, et la fonction coût instantané associée au processus Y_t est identiquement nulle pour tout $t \geq T$.

On a :

$$\begin{aligned} E_{X_0}\left[\sum_{n=1}^{\infty} c_3 \cdot \mathbf{1}_{\{A_{t_n}=1\}}\right] &= E_{X_0}\left[\sum_{n=1}^{\infty} c_3 \cdot \mathbf{1}_{\{A_{t_n}=1\}} \cdot \mathbf{1}_{\{T>t_n\}}\right] \\ &= E_{X_0}\left[E\left\{\sum_{n=1}^{\infty} c_3 \cdot \mathbf{1}_{\{A_{t_n}=1\}} \cdot \mathbf{1}_{\{T>t_n\}} | (X_t, A_t), t \leq t_n\right\}\right] \\ &= E_{X_0}\left[\sum_{n=1}^{\infty} c_3 \cdot \mathbf{1}_{\{A_{t_n}=1\}} \cdot E\left(\mathbf{1}_{\{T>t_n\}} | (X_t, A_t), t \leq t_n\right)\right]. \end{aligned}$$

Or T est indépendant de $(X_t,\, A_t)_t$, donc

$$E\,(\mathbf{1}_{\{T>t_n\}}|(X_t,\, A_t),\, t \le t_n) = E\,(\mathbf{1}_{\{T>t_n\}}) = P\,(T > t_n) = e^{-\delta\,.\,t_n}$$

donc

$$\begin{aligned} E_{X_0}\left[\sum_{n=1}^{\infty} c_3\,.\,\mathbf{1}_{\{A_{t_n}=1\}}\right] &= E_{X_0}\left[\sum_{n=1}^{\infty} c_3\,.\,\mathbf{1}_{\{A_{t_n}=1\}}\,.\,e^{-\delta\,.\,t_n}\right] \\ &= E_{X_0}\left[\sum_{n=1}^{\infty} c_3\,.\,e^{-\delta\,.\,T_n}\right]. \end{aligned}$$

On a aussi :

$$\begin{aligned} & E_{X_0}\left[\int_0^T (c_1\,.\,Y_t^1 + c_2\,.\,Y_t^2)\,dt\right] \\ &\quad = E_{X_0}\left[\int_0^{\infty} (c_1\,.\,X_t^1 + c_2\,.\,X_t^2)\,.\,\mathbf{1}_{\{t<T\}}\,dt\right] \\ &\quad = \int_0^{\infty} E_{X_0}\,[(c_1\,.\,X_t^1 + c_2\,.\,X_t^2)\,.\,\mathbf{1}_{\{t<T\}}]\,dt \\ &\quad = \int_0^{\infty} E_{X_0}\,(c_1\,.\,X_t^1 + c_2\,.\,X_t^2)\,.\,E_{X_0}\,(\mathbf{1}_{\{t<T\}})\,dt \\ &\quad = \int_0^{\infty} E_{X_0}\,(c_1\,.\,X_t^1 + c_2\,.\,X_t^2)\,.\,E\,(\mathbf{1}_{\{t<T\}})\,dt \\ &\quad = \int_0^{\infty} E_{X_0}\,(c_1\,.\,X_t^1 + c_2\,.\,X_t^2)\,.\,e^{-\delta\,.\,t}\,\,dt \\ &\quad = \int_0^{\infty} E_{X_0}\,[(c_1\,.\,X_t^1 + c_2\,.\,X_t^2)\,.\,e^{-\delta\,.\,t}]\,\,dt \\ &\quad = E_{X_0}\left[\int_0^{\infty} (c_1\,.\,X_t^1 + c_2\,.\,X_t^2)\,.\,e^{-\delta\,.\,t}\,dt\right] \end{aligned}$$

le passage de la deuxième égalité à la troisième et de la troisième à la quatrième étant justifié par le fait que T est indépendant de $(X_t)_t$. D'où :

$$\begin{aligned} & E_{X_0}\left[\int_0^{\infty} (c_1\,.\,X_t^1 + c_2\,.\,X_t^2)\,.\,e^{-\delta\,.\,t}\,dt + \sum_{n=1}^{\infty} c_3\,.\,e^{-\delta\,.\,T_n}\right] \\ &\quad = E_{X_0}\left[\int_0^T (c_1\,.\,Y_t^1 + c_2\,.\,Y_t^2)\,dt + \sum_{n=1}^{\infty} c_3\,.\,\mathbf{1}_{\{A_{t_n}=1\}}\right]. \end{aligned}$$

Ce résultat montre que le problème de contrôle initial en temps continu est transformé en un autre problème de contrôle équivalent en temps continu, mais non actualisé.

On a :

$$E_{X_0}\left[\int_0^T (c_1 . Y_t^1 + c_2 . Y_t^2)\, dt\right] = E_{X_0}\left[\sum_{n=0}^{\infty} \int_{t_n}^{t_{n+1}} (c_1 . Y_n^1 + c_2 . Y_n^2)\, dt\right]$$
$$= E_{X_0}\left[\sum_{n=0}^{\infty} (c_1 . Y_n^1 + c_2 . Y_n^2)\,(t_{n+1} - t_n)\right].$$

Le terme entre crochets dans cette dernière égalité est positif, donc on peut utiliser le théorème de convergence monotone en permutant E_{X_0} avec la somme. De plus, puisque $(t_{n+1} - t_n)$ est indépendant de $(Y_n^i)_{i=1,\,2}$ alors on a :

$$\begin{aligned} & E_{X_0}\left[\int_0^T (c_1 . Y_t^1 + c_2 . Y_t^2)\, dt\right] \\ & = \sum_{n=0}^{\infty} E_{X_0}\,(c_1 . Y_n^1 + c_2 . Y_n^2) \,.\, E\,(t_{n+1} - t_n) \\ & = \frac{1}{\delta + \gamma}\, E_{X_0}\left[\sum_{n=0}^{\infty} (c_1 . Y_n^1 + c_2 . Y_n^2)\right] \end{aligned}$$

d'où :

$$\begin{aligned} & E_{X_0}\left[\int_0^T (c_1 . Y_t^1 + c_2 . Y_t^2)\, dt + \sum_{n=1}^{\infty} c_3 . \mathbf{1}_{\{A_{t_n}=1\}}\right] \\ & = \frac{1}{\delta + \gamma}\, E_{X_0}\left[\sum_{n=0}^{\infty} (c_1 . Y_n^1 + c_2 . Y_n^2)\right] + E_{X_0}\left[\sum_{n=1}^{\infty} c_3 . \mathbf{1}_{\{A_{t_n}=1\}}\right]. \end{aligned}$$

Le problème de contrôle équivalent en temps continu est lui-même transformé en un autre problème de contrôle équivalent mais en temps discret. Le processus $(Y_n)_{n \geq 0}$ est une chaîne de Markov contrôlée qui n'est rien d'autre que celle obtenue par uniformisation, comme dans [10], du processus $(Y_t)_t$ et dont la matrice de transition P_1 se déduit facilement de la matrice P en supprimant aux dénominateurs des valeurs de P les taux $\lambda\,(i,\, a)$ et en les remplaçant par la valeur $\delta + \gamma$.

Il reste uniquement à définir les éléments de P_1 pour $j = i$ (saut dans le même état) et pour $j = m$ (l'état du processus $(X_t)_t$ à l'instant T).

On note $m = (m_1, m_2, z)$ pour $z \in \{0, 1\}$.

On a alors

$$P_1 [m_1, m_2, z | x_1, m_2, z; a] = P_1 [m_1, m_2, z | m_1, x_2, z; a] = \frac{\delta}{\delta + \gamma}$$

$\forall x_1 \neq m_1, \forall x_2 \neq m_2, \forall z \in \{0, 1\}, \forall a \in \{0, 1\}$ et

$$P_1 [x_1, x_2, z | x_1, x_2, z; a] = 1 - \frac{\lambda' [(x_1, x_2, z); a]}{\delta + \gamma}.$$

$\forall x_1, x_2 \in \mathbb{N}, \forall z \in \{0, 1\}$ et $\forall a \in \{0, 1\}$ où $\lambda' [(x_1, x_2, z); a]$ est le taux de sortie de l'état (x_1, x_2, z) quand l'action a est choisie pour le processus équivalent $(Y_t)_t$.

2. Détermination de la fonction valeur sur n étapes de la chaîne de Markov contrôlée $(Y_n)_{n \geq 0}$: Posons $\delta + \gamma = 1$.

On note par V_n^0 (respectivement V_n^1) la fonction valeur à l'étape n lorsqu'il y a panne (respectivement lorsqu'il n'y a pas panne).

V_n^i $(i = 1, 2)$ est caractérisée par l'équation d'optimalité de programmation dynamique suivante :

$$V_{n+1}^i = \phi V_n^i \qquad (i = 0, 1), \quad (n = 0, 1, ...)$$

où ϕ est l'opérateur défini comme dans l'assertion (i) de la proposition 2 par :

$$\begin{cases} \phi V_n^i (x) = \min\limits_{a \in A(x)} \{C(x, a) + \sum\limits_{y \in \mathbb{N}^2, j \in \{0, 1\}} p_{xy}(a) \,.\, V_n^j (y)\}, & \forall n \geq 1 \\ \text{avec} \quad V_0^i (\,.\,) = 0 \end{cases}$$

où $x = (x_1, x_2) \in \mathbb{N}^2$.

En introduisant dans cette équation les probabilités de transition $p_{xy}(a)$ de la matrice P_1, on trouve :

$$\begin{aligned} \phi V_n^0 (x) = & \, c \,.\, x + (\mu_1 + \alpha) \,.\, V_n^0 (x) + \mu_2 \,.\, V_n^0 (D_2 x) \\ & + \lambda_2 \,.\, V_n^0 (A_2 x) + \beta \,.\, V_n^1 (x) \\ & + \lambda_1 \,.\, \min \{V_n^0 (A_1 x); V_n^0 (A_2 x) + c_3\} \end{aligned}$$

$$\begin{aligned} \phi V_n^1 (x) = & \, c \,.\, x + \mu_1 \,.\, V_n^1 (D_1 x) + \mu_2 \,.\, V_n^1 (D_2 x) \\ & + \alpha \,.\, V_n^0 (x) + \lambda_2 \,.\, V_n^1 (A_2 x) \\ & + \beta \,.\, V_n^1 (x) + \lambda_1 \,.\, \min \{V_n^1 (A_1 x); V_n^1 (A_2 x) + c_3\} \end{aligned}$$

où $c \,.\, x = c_1 \,.\, x_1 + c_2 \,.\, x_2$.

• 2e *étape* : **Propriétés structurelles**.

On va déterminer les propriétés structurelles caractérisant la politique optimale (si elle existe) en suivant la même démarche que celle utilisée dans [5].

Soit H l'ensemble des fonctions réelles h définies sur $\mathbb{N}^2$ vérifiant les propriétés suivantes :

a) $h(x)$ est croissante en x_1 et en x_2.

b) $h(x+y) - h(x)$ est croissante en x_1 et en x_2 pour chaque y fixé dans $\mathbb{N}^2$.

*c*1) $h(A_2\, x) - h(A_1\, x)$ est décroissante en x_1.

*c*2) $h(A_2\, x) - h(A_1\, x)$ est croissante en x_2.

On veut montrer que l'ensemble H est fermé pour les limites ponctuelles (*voir* théorème 1), pour cela on montre d'abord le résultat suivant :

$$si \quad h \in H \quad alors \quad \phi\, h \in H.$$

La preuve de ce résultat est une conséquence des trois lemmes cités en annexe.

• 3e *étape* : **Passage à la limite.**

Puisque la fonction coût instantané $C(\,.\,,\,.\,)$ est positive et l'espace d'actions A est fini, la politique optimale existe d'après l'assertion (ii) de la proposition 2 et $\lim_{n\to\infty} V_n^i(x)$ existe pour tout x dans $\mathbb{N}^2$ et pour $i = 0,\, 1$. Si on note par V_∞^i $(i = 0,\, 1)$ la fonction limite, alors V_∞^i sera l'unique solution bornée de l'équation de programmation dynamique suivante :

$$V_\infty^i = \phi\, V_\infty^i \tag{6}$$

(d'après le même résultat).

V_∞^i est appelée fonction valeur à l'infini.

De plus, V_n^i est dans H pour tout n, donc le théorème 1 implique que V_∞^i est aussi dans H, c'est-à-dire, que V_∞^i vérifie les propriétés *a*), *b*), *c*1) et *c*2) de la deuxième étape.

• 4e *étape* : **Optimalité de la politique à seuil.**

La politique optimale est l'une qui réalise le minimum dans (6) pour $i = 0$ dans le cas de panne (respectivement $i = 1$ dans le cas de réparation de la panne), c'est-à-dire, la décision optimale est d'envoyer le client de flux 1 lorsqu'il arrive à la file 1 si et seulement si :

$$V_\infty^i(A_1\, x) \le V_\infty^i(A_2\, x) + c_3, \qquad i = 0,\, 1. \tag{7}$$

5. CARACTÉRISATION DE LA POLITIQUE À SEUIL

Si on suppose qu'il est optimal d'envoyer un client arrivant du flux 1 à la file 1 quand l'état des deux files est $x = (x_1, x_2)$ alors d'après (7) on a :

$$V_\infty^i (A_1 x) - V_\infty^i (A_2 x) \leq c_3, \qquad i = 0, 1.$$

Or $V_\infty^i \in H$, donc $(V_\infty^i (A_1 x) - V_\infty^i (A_2 x))$ est décroissante en x_2, et par suite :

$$V_\infty^i (A_2 A_1 x) - V_\infty^i (A_2 A_2 x) \leq V_\infty^i (A_1 x) - V_\infty^i (A_2 x) \leq c_3, \qquad i = 0, 1,$$

ce qui implique que

$$V_\infty^i (A_1 (A_2 x)) \leq V_\infty^i (A_2 (A_2 x)) + c_3, \qquad i = 0, 1.$$

Ces inégalités montrent que si l'optimalité est atteinte lorsqu'on envoie un client du flux 1 à la file 1, alors toute décision qui consiste à ajouter après cela, un ou plusieurs clients dans la file 2 est optimale.

De façon identique, on peut montrer que si la décision optimale est d'envoyer un client arrivant du flux 1 à la file 2, alors toute décision qui consiste à ajouter après cela, un ou plusieurs clients dans la file 1 est optimale. On a donc bien l'optimalité d'une politique à seuil.

6. CONCLUSION

À l'aide de la théorie de la programmation dynamique, on a bien montré l'existence d'une politique optimale à seuil, dépendant de l'état du serveur de la première file, qui minimise le coût instantané défini dans la formulation du problème.

ANNEXE

Soient h une fonction définie sur $\mathbb{N}^2$ à valeurs dans $\mathbb{R}$ et $\tilde{h}$ une fonction définie sur $\mathbb{Z}^2$ à valeurs dans $\mathbb{R}$ telle que :

$$\tilde{h}(x_1, x_2) = h(x_1^+, x_2^+). \tag{8}$$

On considère les opérateurs

$$\hat{A}_1, \hat{A}_2, \hat{D}_1, \hat{D}_2 : \mathbb{Z}^2 \mapsto \mathbb{Z}^2$$

définis par :

$$\hat{A}_1\, x = (x_1 + 1,\, x_2), \qquad \hat{A}_2\, x = (x_1,\, x_2 + 1)$$
$$\hat{D}_1\, x = (x_1 - 1,\, x_2), \qquad \hat{D}_2\, x = (x_1,\, x_2 - 1)$$

où $x = (x_1,\, x_2)$.

Soit $\hat{\phi}$ l'opérateur agissant sur les fonctions définies sur $\mathbb{Z}^2$ et défini de la même façon que l'opérateur ϕ de la première étape avec $A_i,\, D_i$ remplacés par $\hat{A}_i,\, \hat{D}_i\, (i = 1,\, 2)$.

Soit $\tilde{H}$ l'ensemble des fonctions réelles $\tilde{h}$ définies sur $\mathbb{Z}^2$ par (8) et vérifiant les propriétés *a*), *b*), *c*1) et *c*2) de la deuxième étape.

Remarque 1 : *La propriété b) est équivalente aux propriétés suivantes :*

(i) $h\,(x_1 + 1,\, x_2) - h\,(x_1,\, x_2) \geq h\,(x_1,\, x_2) - h\,(x_1 - 1,\, x_2)$,

(ii) $h\,(x_1,\, x_2 + 1) - h\,(x_1,\, x_2) \geq h\,(x_1,\, x_2) - h\,(x_1,\, x_2 - 1)$,

(iii) $h\,(x + y) - h\,(x)$ *est croissante en* x_1 *pour* $y = (0,\, 1)$ *et croissante en* x_2 *pour* $y = (1,\, 0)$.

Les propriétés (i) et (ii) proviennent du fait que $h\,(x + y) - h\,(x)$ est croissante en x_1 pour $y = (1,\, 0)$ et en x_2 pour $y = (0,\, 1)$, respectivement.

LEMME 1 : (i) *La restriction d'une fonction* h *de* $\tilde{H}$ *à* $\mathbb{N}^2$ *est dans* H ;

(ii) *Si* h *est dans* H *alors* $\tilde{h}$ *est dans* $\tilde{H}$;

(iii) *Si* h *est une fonction définie sur* $\mathbb{N}^2$*, croissante en* x_1 *et en* x_2*, alors la restriction de* $\hat{\phi}\tilde{h}$ *à* $\mathbb{N}^2$ *est égale à* $\phi\, h$ (i.e., $\phi h = \hat{\phi}\tilde{h}_{/\mathbb{N}^2}$).

Preuve : La preuve de l'assertion (i) est triviale.

En utilisant la remarque 1 et en distinguant tous les cas possibles, il est facile de démontrer l'assertion (ii) (*voir* [6]).

Montrons l'assertion (iii).

Soit h une fonction définie sur $\mathbb{N}^2$, montrons que

$$\hat{\phi}\tilde{h}_{/\mathbb{N}^2} = \phi h \tag{9}$$

on va montrer l'égalité (9) pour $h = V_n^0$ et la démonstration sera identique pour $h = V_n^1$.

Soit $x_1, x_2 \in \mathbb{N}$, on a

$$\begin{aligned}
\hat{\phi}\tilde{V}_n^0(x) &= c\,.\,x + (\mu_1 + \alpha)\,.\,\tilde{V}_n^0(x) + \mu_2\,.\,\tilde{V}_n^0(\hat{D}_2\,x) \\
&\quad + \lambda_2\,.\,\tilde{V}_n^0(\hat{A}_2\,x) + \beta\,.\,\tilde{V}_n^1(x) \\
&\quad + \lambda_1\,.\,\min\{\tilde{V}_n^0(\hat{A}_1\,x);\ \tilde{V}_n^0(\hat{A}_2\,x) + c_3\} \\
&= c\,.\,x + (\mu_1 + \alpha)\,.\,\tilde{V}_n^0(x_1, x_2) + \mu_2\,.\,\tilde{V}_n^0(x_1, x_2 - 1) \\
&\quad + \lambda_2\,.\,\tilde{V}_n^0(x_1, x_2 + 1) \\
&\quad + \beta\,.\,\tilde{V}_n^1(x_1, x_2) + \lambda_1\,.\,\min\{\tilde{V}_n^0(x_1 + 1, x_2); \\
&\qquad \tilde{V}_n^0(x_1, x_2 + 1) + c_3\} \\
&= c\,.\,x + (\mu_1 + \alpha)\,.\,V_n^0(x_1^+, x_2^+) + \mu_2\,.\,V_n^0(x_1^+, (x_2 - 1)^+) \\
&\quad + \lambda_2\,.\,V_2^0(x_1^+, (x_2 + 1)^+) + \beta\,.\,V_n^1(x_1^+, x_2^+) \\
&\quad + \lambda_1\,.\,\min\{V_n^0((x_1 + 1)^+, x_2^+),\ V_n^0(x_1^+, (x_2 + 1)^+) + c_3\},
\end{aligned}$$

et puisque $x_1, x_2 \in \mathbb{N}$ alors

$$\begin{aligned}
\hat{\phi}\tilde{V}_n^0(x) &= c\,.\,x + (\mu_1 + \alpha)\,.\,V_n^0(x_1, x_2) + \mu_2\,V_n^0(x_1, (x_2 - 1)^+) \\
&\quad + \lambda_2\,.\,V_n^0(x_1, x_2 + 1) \\
&\quad + \beta\,.\,V_n^1(x_1, x_2) + \lambda_1\,.\,\min\{V_n^0(x_1 + 1, x_2); \\
&\qquad V_n^0(x_1, x_2 + 1) + c_3\} \\
&= c\,.\,x + (\mu_1 + \alpha)\,.\,V_n^0(x) + \mu_2\,.\,V_n^0(D_2\,x) \\
&\quad + \lambda_2\,.\,V_n^0(A_2\,x) + \beta\,.\,V_n^1(x) \\
&\quad + \lambda_1\,.\,\min\{V_n^0(A_1\,x);\ V_n^0(A_2\,x) + c_3\} \\
&= \phi\,V_n^0,
\end{aligned}$$

ce qui termine la démonstration de l'assertion (iii).

LEMME 2 : *Si h est dans $\tilde{H}$ alors $\hat{\phi}\,h$ est dans $\tilde{H}$.*

Preuve : Soit $h \in \tilde{H}$.

D'après la définition de $\hat{\phi}$, il est facile de vérifier que $h\,\hat{A}_1$, $h\,\hat{A}_2$, $h\,\hat{D}_1$, et $h\,\hat{D}_2$ sont dans $\tilde{H}$, donc pour démontrer le lemme 2 il suffit de vérifier que :

$$f(x) = \min(h(A_1\,x),\ h(A_2\,x) + c_3) \in \tilde{H}.$$

Il est facile de montrer que la fonction f vérifie la propriété a) puisque la fonction h la vérifie également.

Pour x dans $\mathbb{Z}^2$, nous allons montrer les propriétés suivantes :

$$f(A_1^2 x) - f(A_1 x) \geq f(A_1 x) - f(x) \tag{10}$$

$$f(A_1 A_2 x) - f(A_2 x) \geq f(A_1 x) - f(x) \tag{11}$$

$$f(A_1 A_2 x) - f(A_1^2 x) \leq f(A_2 x) - f(A_1 x) \tag{12}$$

où A_i^n désigne l'arrivée de n clients dans la file i.

Les inégalités (10) et (11) impliquent que la fonction f vérifie la propriété *b*) pour $y = (1, 0)$, et par symétrie f vérifie aussi la propriété *b*) pour $y = (0, 1)$.

D'après la remarque 1, f vérifie la propriété *b*) pour tout y dans $\mathbb{N}^2$.

L'inégalité (12) implique que f vérifie la propriété *c*1) et par symétrie elle vérifie aussi la propriété *c*2).

Si f vérifie les trois inégalités précédentes, alors f sera dans $\tilde{H}$, et par conséquent le lemme 2 sera prouvé.

Pour cela on pose :

$$a = h(A_2^2 x), \quad b = h(A_2^2 A_1 x), \quad c = h(A_2 x), \quad d = h(A_2 A_1 x),$$

$$e = h(A_2 A_1^2 x), \quad f = h(A_1 x), \quad g = h(A_1^2 x), \quad h = h(A_1^3 x),$$

où $x = (x_1, x_2) \in \mathbb{Z}^2$.

Les inégalités (10)-(12) s'écrivent respectivement de la façon suivante :

$$h \wedge (e + c_3) - g \wedge (d + c_3) \geq g \wedge (d + c_3) - f \wedge (c + c_3) \tag{13}$$

$$e \wedge (b + c_3) - d \wedge (a + c_3) \geq g \wedge (d + c_3) - f \wedge (c + c_3) \tag{14}$$

$$e \wedge (b + c_3) - h \wedge (e + c_3) \leq d \wedge (a + c_3) - g \wedge (d + c_3) \tag{15}$$

où $a \wedge b$ dénote le minimum entre a et b.

Pour faciliter la démonstration des inégalités précédentes, on utilise les inégalités (16)-(18) ci-dessous, qui sont faciles à vérifier.

$$\begin{cases} c-f \geq d-g \geq e-h \\ e-d \geq d-c \\ h-g \geq g-f; \quad e-g \geq d-f \end{cases} \tag{16}$$

$$\begin{cases} a-d \geq b-e \geq d-g \\ a-d \geq c-f \geq d-g \\ a-c \leq b-d; \quad d-f \leq e-g \\ d-c \leq e-d; \quad d-f \leq b-d \end{cases} \tag{17}$$

$$\begin{cases} a-d \geq b-e \geq e-h \\ a-d \geq d-g \geq e-h. \end{cases} \tag{18}$$

• L'inégalité (13) sera alors établie en utilisant (16) et en considérant séparément les 4 cas suivants :

$$-c_3 \geq c-f; \quad c-f \geq -c_3 \geq d-g; \quad d-g \geq -c_3 \geq e-h \quad \text{et} \quad e-h \geq -c_3.$$

• L'inégalité (14) sera établie en utilisant (17) et en considérant les 3 cas suivants :

$$-c_3 \geq a-d; \quad a-d \geq -c_3 \geq d-g \quad \text{et} \quad d-g \geq -c_3.$$

• Finalement l'inégalité (15) sera établie en utilisant (18) et en considérant les 3 cas suivants :

$$-c_3 \geq a-d, \quad a-d \geq -c_3 \geq e-h \quad \text{et} \quad e-h \geq -c_3.$$

LEMME 3 : *Si h est dans H alors ϕh est dans H.*

Preuve : Si h est dans H, alors d'après l'assertion (ii) du lemme 1, $\tilde{h}$ est dans $\tilde{H}$, et d'après le lemme 2, $\hat{\phi}\tilde{h}$ est dans $\tilde{H}$. D'autre part, d'après l'assertion (iii) du lemme 1, ϕh est la restriction à $\mathbb{N}^2$ de $\hat{\phi}\tilde{h}$, élément de $\tilde{H}$.

Finalement, par application de l'assertion (i) du lemme 1, on montre que ϕh est élément de H.

BIBLIOGRAPHIE

1. D. Bertsekas, Dynamic programming: deterministic and stochastic models. Prentice Hall, Englewood Cliffs, New Jersey, 1987.
2. A. Ephremides, P. Varaiya et J. Walrand, A simple dynamic routing problem, *IEEE Trans. Automat. Contr.*, 1980, *AC-25*, n° 4, p. 690-693.
3. G. J. Foschini et J. Salz, A basic dynamic routing problem and diffusion, *IEEE Trans. Commun.*, *COM-26*, p. 320-347.
4. H. Ghoneim, Optimal control of arrivals to a network of two queues in series. *Ph. D. disertation, Prog. Oper. Res.*, North Carolina State Univ., Releigh, Nc, 1980.
5. B. Hajek, Optimal control of two interacting service station, *IEEE Trans. Automat. Contr.*, 1984, *AC-29*, n° 6, p. 491-499.
6. A. Haqiq, Contrôle de flux dans des réseaux de télécommunications : cas du réseau ATM. Thèse de troisième cycle, n° 1279, Faculté des sciences de Rabat, 20 juillet 1995.
7. P. R. Kumar et P. Varaiya, Stochastic systems: estimation, identification and adaptative control. Prentice Hall, Englewood Cliffs, New Jersey, 1986.
8. W. Lin et P. R. Kumar, Optimal control of a queueing system with two heterogeneous services, Preprint, décembre 1982.
9. S. M. Ross, Introduction to stochastic dynamic programming. Academic Press, New York, 1983.
10. R. Serfozo, An equivalence between continuous and discrete time Markov decision processes. *Oper. Res.*, 1979, n° *27*, p. 616-620.
11. J. Walrand, An introduction to queueing networks. Prentice Hall International Editions, Englwood Cliffs, New Jersey 07632, 1988.
12. R. Weber, On the optimal assignment of customers to parallel servers, *J. Appl. Prob.*, 1978, n° *15*, p. 406-413.
13. P. Whittle, Optimisation over time: dynamic programming and stochastic control. *I*, John Wiley and Sons, New York, 1982.
14. P. Whittle, Optimisation over time: dynamic programming and stochastic control. *II*, John Wiley and Sons, New York, 1983.
15. W. Winston, Optimality of the shortest-processing-time discipline, *J. Appl. Prob.*, 1977, n° *14*, p. 181-189.