

Contrôle de système à événements discrets stochastiques

File $M/M/1$ batch avec échéances exponentielles

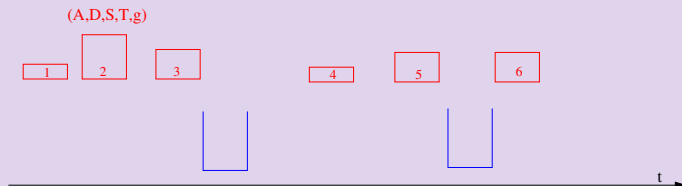
E. Hyon, A. Jean-Marie

LIP6 – INRIA/LIRMM Montpellier

Juillet 2008

Motivation première : problème générique

Bin Packing Stochastique, MultiKnapsack, File Attente



Tâches (items)

- **Date d'arrivée** (Flux Poisson)
- **Échéance** (exponentielle)
- Temps de Service (expo)
- Revenu (gain ou **pénalité**), Taille

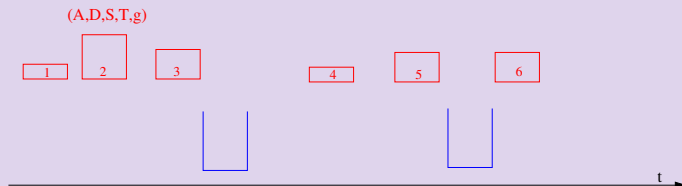
Serveur (bin)

- **Capacité** (fixée)
- **Temps de service** (max ou $\sum$)
- **Coût**

Minimisation de la moyenne des coûts (pertes et batches)

Motivation première : problème générique

Bin Packing Stochastique, MultiKnapsack, File Attente



Tâches (items)

- **Date d'arrivée** (Flux Poisson)
- **Échéance** (exponentielle)
- Temps de Service (expo)
- Revenu (gain ou **pénalité**), Taille

Serveur (bin)

- **Capacité** (fixée)
- **Temps de service** (max ou $\sum$)
- **Coût**

Minimisation de la moyenne des coûts (pertes et batches)

Techniques usuelles issues optimisation déterministe

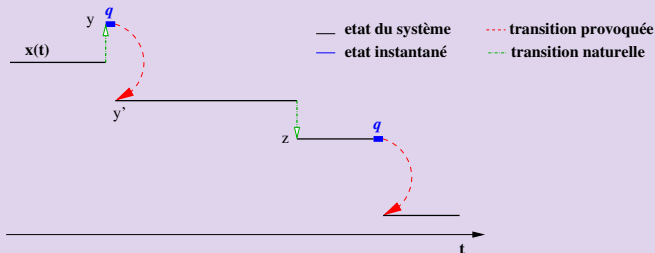
- MultiKnapsack (sac à dos) : adapté horizon fini
 - sac taille limitée
 - à remplir d'objets à **encombrement** et **coûts** variables
- Bin packing stochastique : occupation des bins

Processus de décisions Markoviens

- En temps discret : Chaîne de Markov $X_n = f(X_{n-1}, Z_n)$
MDP ajout d'un contrôle sur les transitions $X_n = f(X_{n-1}, Q_n, Z_n)$.
- MDP décision influe sur état.
- En temps continu : Processus de Markov $X_{T_n} = f(X_{T_{n-1}}, Z_n)$
SMDP ajout d'un contrôle $X_n = f(X_{n-1}, Q_n, Z_n)$.
- Dans SMDP la décision joue sur état et durée de transition.

Processus de Decision à temps continu

Evolution du processus

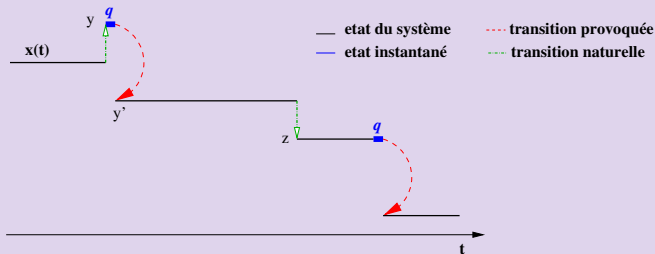


Détail des sauts

- Transition contrôlée : De y suivant décision q saut instantané en y' .
- Transition naturelle : De y'
 - ▶ Temporellement : saut au bout de t durée aléatoire dépend de y' .
 - ▶ Spatialement : saut en état z qui dépend de l'occurrence de l'événement e .
 - ▶ Modélisation : Etat y' apparaît dans dynamique mais pas dans modèle SMDP.

Processus de Decision à temps continu

Evolution du processus



Etat $x = (n, \beta)$

- n nombre clients dans file (pas dans système)
- β occupation du serveur (0 ou 1)

Transitions

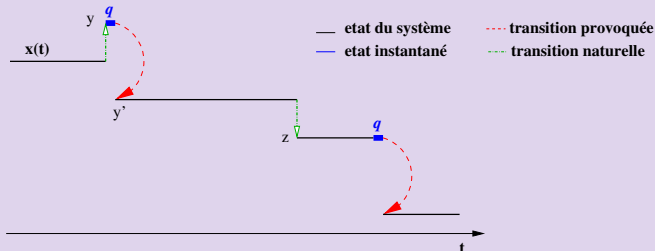
- Proba transition $p(z|(y, q))$
- Proba durée transition : $\mathcal{E}xp \approx \Lambda(y, q)$.

Coûts

- Coût transition provoquée $c^P(y, q)$
- Coût transition naturelle $c^L(e|(y, q))$

Processus de Decision à temps continu

Evolution du processus



Instants de décisions SMDP

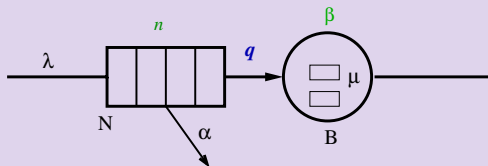
Dans SMDP : prise de décision (contrôle) **uniquement** aux **instants de transitions**.

L'ensemble $\{q\}$ des décisions successives est la **politique**.

Si holding costs (coûts dépendant durée) SMDP pas forcément optimal : instants décisions font partie du pb d'optimisation. Cadre du contrôle optimal "classique" avec ODE.

Modèle File d'attente contr lée avec services par Batch

Ex : Modèle de file d'attente contrôlée



Contrôle optimal : Trouver la politique d'admission des clients

- Politique c'est la détermination de l'action à effectuer dans état x .
- Déterminer $q(x)$.

Fonction de coût

Cout discounté

$$v_{\theta}^{\pi}(x) = \mathbb{E}_x^{\pi} \left[\sum_{n=0}^{\infty} \exp^{-\theta T_n^{\pi}} c^P(x_{T_n^{\pi}}, q_{T_n^{\pi}}) + \exp^{-\theta T_{n+1}^{\pi}} c^L(x_{T_{n+1}^{\pi}} | (x_{T_n^{\pi}}, q_{T_n^{\pi}})) \right]. \quad (1)$$

Cout Moyen

$$v_{\theta}^{\pi}(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_x^{\pi} \left[\sum_{n=0}^N c^P(x_{T_n^{\pi}}, q_{T_n^{\pi}}) + c^L(x_{T_{n+1}^{\pi}} | (x_{T_n^{\pi}}, q_{T_n^{\pi}})) \right]. \quad (2)$$

Fonction de valeur

On cherche la politique π (déterministe, stationnaire, markovienne) telle que

$$v_{\theta}^*(x) = \min_{\pi} v_{\theta}^{\pi}(x) \quad (3)$$

où v_{θ}^{π} solution de l'équation fonctionnelle (issue de (1)) : $v_{\theta}^{\pi} = T v_{\theta}^{\pi}$

Fonction de coût

Cout discounté

$$v_{\theta}^{\pi}(x) = \mathbb{E}_x^{\pi} \left[\sum_{n=0}^{\infty} \exp^{-\theta T_n^{\pi}} c^P(x_{T_n^{\pi}}, q_{T_n^{\pi}}) + \exp^{-\theta T_{n+1}^{\pi}} c^L(x_{T_{n+1}^{\pi}} | (x_{T_n^{\pi}}, q_{T_n^{\pi}})) \right]. \quad (1)$$

Cout Moyen

$$v_{\theta}^{\pi}(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_x^{\pi} \left[\sum_{n=0}^N c^P(x_{T_n^{\pi}}, q_{T_n^{\pi}}) + c^L(x_{T_{n+1}^{\pi}} | (x_{T_n^{\pi}}, q_{T_n^{\pi}})) \right]. \quad (2)$$

Fonction de valeur

On cherche la politique π (déterministe, stationnaire, markovienne) telle que

$$v_{\theta}^*(x) = \min_{\pi} v_{\theta}^{\pi}(x) \quad (3)$$

où v_{θ}^{π} solution de l'équation fonctionnelle (issue de (1)) : $v_{\theta}^{\pi} = T v_{\theta}^{\pi}$

Expression du problème de file d'attente

Transitions en état $x = (n, \beta)$

- Taux global

$$\Lambda(x, q) = \mathbf{1}_{\{\beta=1\}} \cdot (n\alpha + \lambda + \mu) + \mathbf{1}_{\{\beta=0\}} \cdot ((n - q)^+ \alpha + \lambda + \mu \mathbf{1}_{\{q > 0\}})$$

- Taux événements :

- ▶ Arrivée client : λ .

Etat final $(\min(n + 1, N_{\max}), \beta)$ ou $(\min((n - q)^+ + 1, N_{\max}), 1)$

- ▶ Départ (fin d'un batch) : μ .

Etat final $(n, 0)$ ou $((n - q)^+, 0)$.

- ▶ Perte client : $n\alpha$ ou $(n - q)^+ \alpha$.

Etat final $(n - 1, \beta)$ ou $((n - q)^+ - 1, 1)$

La probabilité d'une transition de (x, q) vers y égal taux événement vers y divisé par $\Lambda(x, q)$.

Coûts

- Coût instantané d'une décision $c^P(x, q) = k_B \mathbf{1}_{\{q > 0\}}$

- Coût d'une transition $c^N(e|(x, q)) = k_L \mathbf{1}_{\{e=perte\}} + k_f \mathbf{1}_{\{e=file\ pleine\}}$

Expression du problème de file d'attente

Transitions en état $x = (n, \beta)$

- Taux global

$$\Lambda(x, q) = \mathbf{1}_{\{\beta=1\}} \cdot (n\alpha + \lambda + \mu) + \mathbf{1}_{\{\beta=0\}} \cdot ((n - q)^+ \alpha + \lambda + \mu \mathbf{1}_{\{q > 0\}})$$

- Taux événements :

- ▶ Arrivée client : λ .

Etat final $(\min(n + 1, N_{\max}), \beta)$ ou $(\min((n - q)^+ + 1, N_{\max}), 1)$

- ▶ Départ (fin d'un batch) : μ .

Etat final $(n, 0)$ ou $((n - q)^+, 0)$.

- ▶ Perte client : $n\alpha$ ou $(n - q)^+ \alpha$.

Etat final $(n - 1, \beta)$ ou $((n - q)^+ - 1, 1)$

La probabilité d'une transition de (x, q) vers y égal taux événement vers y divisé par $\Lambda(x, q)$.

Coûts

- Coût instantané d'une décision $c^P(x, q) = k_B \mathbf{1}_{\{q > 0\}}$
- Coût d'une transition $c^N(e|(x, q)) = k_L \mathbf{1}_{\{e=perte\}} + k_f \mathbf{1}_{\{e=file\ pleine\}}$

SMDP : Méthodes de résolution

Résolution SMDP classique

Actualisation dépend de la loi de durée de la transition (paramètre $\Lambda(x, q)$).
Actualisation introduite dans la fonction de valeur ($v = Tv$) qui devient

$$v_{\theta}^{\pi}(x) = \min_q c^P(x, q) + \frac{\Lambda(x, q)}{\Lambda(x, q) + \theta} \sum_{e \in \mathcal{E}} c^N(e|(x, q)) p(e|(x, q)) + \frac{\Lambda(x, q)}{\Lambda(x, q) + \theta} \sum_{y \in \mathcal{X}} p(y|(x, q)) v_{\theta}^{\pi}(y) \quad (4)$$

Solution est trouvée en trouvant le point fixe.

Uniformisation (politiques déterministes)

Ajout d'événements virtuels d'un état vers lui-même. Transforme le processus des événements.

- Toutes les durées des transitions ont la même loi.
- Uniformisation modifie les coûts et les probabilités de transitions

SMDP : Méthodes de résolution

Résolution SMDP classique

Actualisation dépend de la loi de durée de la transition (paramètre $\Lambda(x, q)$).
Actualisation introduite dans la fonction de valeur ($v = Tv$) qui devient

$$v_{\theta}^{\pi}(x) = \min_q c^P(x, q) + \frac{\Lambda(x, q)}{\Lambda(x, q) + \theta} \sum_{e \in \mathcal{E}} c^N(e|(x, q)) p(e|(x, q)) + \frac{\Lambda(x, q)}{\Lambda(x, q) + \theta} \sum_{y \in \mathcal{X}} p(y|(x, q)) v_{\theta}^{\pi}(y) \quad (4)$$

Solution est trouvée en trouvant le point fixe.

Uniformisation (politiques déterministes)

Ajout d'événements virtuels d'un état vers lui-même. Transforme le processus des événements.

- Toutes les durées des transitions ont la même loi.
- Uniformisation modifie les coûts et les probabilités de transitions

SMDP : Calcul de la politique optimale

Itération valeur

On calcule le couple politique optimale-valeur optimale.

Détail itération

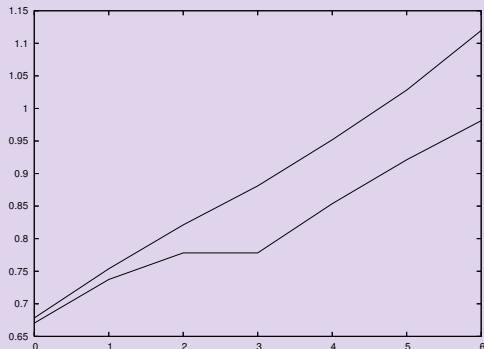
- 1 initialise $v_\theta(x)$ à 0.
- 2 $\forall x$, calcul de $v_\theta(x)$ pour toutes les décisions q
- 3 on pose : $q(x) = \arg \min$
- 4 $v_\theta^*(x) = v_\theta(\arg \min)$.

SMDP : Calcul de la politique optimale

Exemple :

Paramètres : File 6 ; Batch 3 ; $\lambda = \alpha = \mu = 1$; $k_B = k_F = k_L = 1$ et $\theta = 0.1$.

$$\begin{aligned}v_{\theta}^*(0, 0) &= 0, \\v_{\theta}^*(1, 0) &= 0, \\v_{\theta}^*(2, 0) &= 2, \\v_{\theta}^*(3, 0) &= 3, \\v_{\theta}^*(4, 0) &= 3, \\v_{\theta}^*(5, 0) &= 3, \\v_{\theta}^*(6, 0) &= 3, \\v_{\theta}^*(\cdot, 1) &= 0.\end{aligned}$$

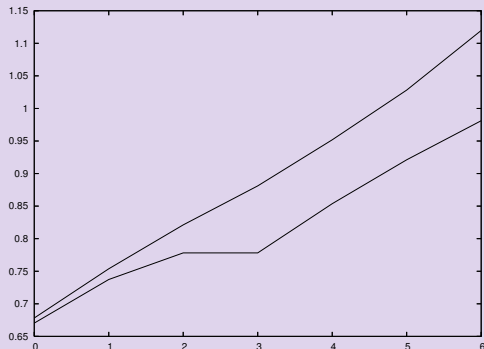


SMDP : Calcul de la politique optimale

Exemple :

Paramètres : File 6 ; Batch 3 ; $\lambda = \alpha = \mu = 1$; $k_B = k_F = k_L = 1$ et $\theta = 0.1$.

$$\begin{aligned}v_{\theta}^*(0, 0) &= 0, \\v_{\theta}^*(1, 0) &= 0, \\v_{\theta}^*(2, 0) &= 2, \\v_{\theta}^*(3, 0) &= 3, \\v_{\theta}^*(4, 0) &= 3, \\v_{\theta}^*(5, 0) &= 3, \\v_{\theta}^*(6, 0) &= 3, \\v_{\theta}^*(\cdot, 1) &= 0.\end{aligned}$$



Propriétés qualitatives

Caractériser plus finement les politiques optimales *i.e.* connaître des propriétés de la fonction $x \rightarrow q(x)$.

- Permet d'accélérer le calcul.
- Permet si la caractérisation est assez précise d'obtenir directement la politique (formule close ou par calcul).

Définition (Contrôle Monotone)

Soit $x \in \mathbb{N}^m$. On dit que le contrôle est monotone en e_i si $q(x) \leq q(x + e_i)$.

Définition (Politique à seuil)

On dit qu'un contrôle optimal est une politique à seuil s'il existe $K \in \mathbb{N}$ tel que

$$q^*(x) = \begin{cases} a_1 & \text{si } x < K \\ a_2 & \text{si } x \geq K \end{cases} \quad (5)$$

Politiques structurées

Définition (Fonction sous modulaire)

Soit x_1 et $x_2 \in \mathcal{L}$ et $q_1, q_2 \in \mathcal{Q}$ tel que $\mathcal{L} \times \mathcal{Q}$ soit un treillis.

La fonction f est sous modulaire si

$$f((x_1, q) \vee (x_1, q_1)) + f((x_1, q) \wedge (x_1, q_1)) \leq f(x_1, q_1) + f(x_2, q_2)$$

Théorème (G & Y ou Putt)

Si $f(x, q)$ est sousmodulaire sur $\mathcal{L} \times \mathcal{Q}$ alors il existe un contrôle optimal monotone croissant en x .

Théorème (Fonction de valeur convexe)

Si la fonction de valeur est convexe alors il existe un seuil.