TD 10 : Réseaux de Jackson

lionel.rieg@ens-lyon.fr

Exercice 1 (File M/M/c)

On s'intéresse à une file d'attente avec $c > 1$ serveurs, mais les arrivées et les services restent des processus de Poisson, de paramètres respectifs λ et μ. Dans ce contexte, on définit la charge du système par $r = \lambda/\mu$ et le taux d'utilisation par $\rho = r/c$.

1. Donner le graphe de la chaîne induite et le générateur infinitésimal.
2. Donner et démontrer la condition de stabilité de la file.
3. Calculer la distribution stationnaire.
4. Déterminer L_q le nombre moyen de personne en attente dans le système.
5. Dans le cas d'une file G/M/c, donner la relation entre W le temps de séjour moyen et W_q le temps d'attente moyen.
6. En déduire L, W et W_q.
7. Donner la probabilité d'être servi immédiatement à l'arrivée dans la file.

Exercice 2 (Augmentation de capacité)

on reprend la file M/M/c de l'exercice précédent et, pour fixer les idées, on prend $\mu = 1$, $\rho = 0,75$ et $c = 12$. On a donc $\lambda = r = 9$ arrivées par unité de temps en moyenne, ce qui fait $\Delta = c - r = 3$ serveurs pour absorber les variations de trafic. Supposons que les arrivées quadruplent : $\lambda' = 4\lambda = 36$, comment choisir c' ?

1. Donner différentes grandeurs qu'on peut vouloir garder constantes pour la file d'attente.
2. Pour chacune, calculer le nouveau nombre de serveurs nécessaires.

Exercice 3 (Réseaux de Jackson)

Un réseau de Jackson est un ensemble de N files d'attente de types G/M/c$_p$ dont les taux de service sont caractérisés par $\mu_p(n_p) = \mu_p \cdot min(c_p, n_p)$ où n_p est le nombre de personnes dans la file p et c_p son nombre de serveurs en parallèle. Ces files d'attente sont branchées les unes sur les autres de sorte qu'à la sortie de la file p, on peut :

- entrer dans la file q avec probabilité $r_{p,q}$;
- définitivement sortir du système avec probabilité $r_{p,N+1}$.

Dans la suite, on suppose $\mu_p(n_p) > 0$ dès que $n_p > 0$ ainsi que $r_{p,p} = 0$ pour simplifier. Les arrivées exogènes (de l'extérieur du système) dans la file p forment un processus de Poisson d'intensité $\overline{\lambda}_p$. On admet que le vecteur $X(t)$ du nombre de personnes dans chaque file est une chaîne de Markov à temps continu.

1. Schématiser ce système et décrire le générateur infinitésimal de $X(t)$.
2. On suppose le système en régime stationnaire. Donner les équations d'équilibre de l'espérance des files.
3. Montrer que si le réseau est sans capture, *i.e.* si tout paquet a une probabilité non nulle de sortir du système, alors le système précédent possède une unique solution positive et finie.
4. Démontrer le théorème suivant :

Théorème

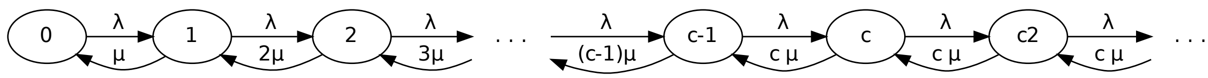
Soit une chaîne de Markov à temps continu irréductible de générateur infinitésimal $Q = (q_{ij})$.
Si on peut trouver une distribution π et des nombre positifs $\tilde{q}_{ij}$ (pour $i \neq j$) tels que $\pi(i)\tilde{q}_{ij} = \pi(j)q_{ji}$ et $\sum_{j \neq i} \tilde{q}_{ij} = -q_{ii}$,
alors π est la distribution stationnaire de la chaîne.

5. Supposons que chaque file soit M/M/c indépendante des autres. Quelle est la distribution stationnaire attendue pour $X(t)$?
6. Montrer que c'est effectivement la distribution stationnaire du système (alors que les chaînes ne sont pas indépendantes).

Solutions

▸ Exercice 1

1. Voici le graphe de la chaîne :



   et le générateur infinitésimal est donné par

$$q_{i,i+1} = \lambda \qquad q_{i+1,i} = i\mu \quad \text{pour } i < c-1 \qquad q_{i+1,i} = c\mu \quad \text{pour } i \geqslant c-1$$

2. Comme toujours, la condition est $\rho < 1$. Elle se démontre à nouveau en utilisant le premier théorème de Foster, en choisissant $F = \{0, \ldots, c-1\}$ et $\varepsilon = \lambda - c\mu$.
3. Notons π la distribution stationnaire. En écrivant l'équilibre de flux pour les i premiers états, on obtient les équations

$$\lambda\pi_i = i\mu\pi_{i+1} \qquad \text{si } i < c-1$$
$$\lambda\pi_i = c\mu\pi_{i+1} \qquad \text{si } c \geqslant c-1$$

   ce qui donne $\pi_i = \frac{r^i}{i!}\pi_0$ pour $i \leqslant c$ et $\pi_i = \frac{r^i}{c!c^{i-c}}\pi_0 = \frac{r^c}{c!}\rho^{i-c}\pi_0$ si $i \geqslant c$. Plus généralement, si on note $\mu(i)$ le taux de service lorsqu'il y a i personnes dans la file, l'équilibre des flux donne $\pi_i = \frac{\lambda^i}{\prod_{n=1}^{i}\mu(n)}\pi_0$. Comme à chaque fois, la valeur des π_0 est donnée par la condition $\sum_{i\geqslant 0}\pi_i = 1$. Malheureusement, ici π_0 n'a pas de forme simple :

$$\pi_0 = \left(\frac{r^c}{c!(1-\rho)} + \sum_{n=0}^{c-1}\frac{r^n}{n!}\right)^{-1}$$

4. On choisit de calculer en premier $L_q = \sum_{n\geqslant c}(n-c)\pi_n$ car il permet d'éviter la forme particulière des c premières valeurs de π.

$$L_q = \sum_{n\geqslant c+1}(n-c)\pi_n = \sum_{n\geqslant 1} n\pi_{n+c} = \sum_{n\geqslant 1} n\frac{r^c}{c!}\rho^n\pi_0 = \frac{r^c\rho}{c!}\pi_0\sum_{n\geqslant 1} n\rho^{n-1} = \frac{r^c\rho}{c!}\pi_0\left(\frac{1}{1-\rho}\right)' = \frac{r^c\rho}{c!(1-\rho)^2}\pi_0$$

5. Le temps de séjour moyen est le temps d'attente moyen plus le temps de service moyen : $W = W_q + \frac{1}{\mu}$.
6. On utilise la loi de Little et la question précédente. On obtient en particulier l'égalité $L = L_q + r$, vraie pour toute file G/G/c.

$$W_q = \frac{L_q}{\lambda} = \frac{r^c}{c!(c\mu)(1-\rho)^2}\pi_0 \qquad W = \frac{1}{\mu} + \frac{r^c}{c!(c\mu)(1-\rho)^2}\pi_0 \qquad L = r + \frac{r^c\rho}{c!(1-\rho)^2}\pi_0$$

7. $$W_q(0) = \mathbb{P}(X \leqslant c-1) = \pi_0\sum_{n=0}^{c-1}\frac{r^n}{n!} = 1 - \frac{r^c\pi_0}{c!(1-\rho)} \quad \text{en explicitant la valeur de } \pi_0.$$

▸ Exercice 2

1. Dans les grandeurs usuelles, on peut choisir de garder constant :
   (a) le taux de service : ρ,
   (b) la probabilité d'attendre : $1 - W_q(0)$ (mesure de congestion),
   (c) le nombre de serveur pour absorber les variations de trafic : $\Delta = c - r$.
2. Dans les trois cas, le nouveau nombre de serveur est :
   (a) $\rho' = \rho \iff c'\mu = 4c\mu$ donc $c' = 4c = 48$,
   (b) on calcule successivement $W_q(0)$ pour différentes valeurs de c' avant de trouver $c' = 42$,
   (c) $c' = \Delta + r' = 3 + 4r = 39$.

   Comme ce n'est pas le même nombre, il faut faire des compromis entre la qualité et l'efficacité. le premier choix est appelé *domaine de qualité*, le second *domaine de qualité et d'efficacité* et le dernier *domaine d'efficacité*, ce qui illustre la motivation de chacun de ces choix.

▸ Exercice 3

Une des difficultés de cet exercice est que l'espace d'état n'est pas $\mathbb{N}$ mais $\mathbb{N}^N$, donc que les indices sont des vecteurs de taille N. Pour conserver des notations faciles à lire, on prend $\vec{n}$ pour représenter les états, n_p pour la p^e composante. Afin de ne pas mélanger ces notations avec celle des distributions, si π est une distribution sur le système global (donc sur des N-uplet), π_p représente la loi marginale de la p^e file et les probabilités d'être dans un état donné seront dénoté par $\pi(\vec{n})$ et $\pi_p(n_p)$.

1. Pour déterminer le générateur infinitésimal, il faut bien décomposer les types de transitions possibles pour un file donnée p :
   – les arrivées exogènes (en vert sur le schéma) : $\overline{\lambda}_p$,
   – les sorties du réseau (en rouge sur le schéma) : $\mu_p(n_p) \cdot r_{p,N+1}$,
   – les routages de la file p à la file $q \neq p$: $\mu_p(n_p) \cdot r_{p,q}$.
   On en déduit les termes non nuls du générateurs infinitésimal :
   $$q_{\vec{n},\vec{n}+e_p} = \overline{\lambda}_p \qquad q_{\vec{n}+e_p,\vec{n}} = \mu_p(n_p)\, r_{p,N+1} \qquad q_{\vec{n}+e_p,\vec{n}+e_q} = \mu_p(n_p)\, r_{p,q}$$
2. Notons λ_p l'espérance de l'entrée dans la file p, qui combine les arrivées exogènes et les routages depuis les autres files. Comme les files sont en équilibre, tout ce qui entre en sort donc l'espérance de la sortie de la file p est également λ_p. Une fraction $r_{p,q}$ de cette quantité est envoyé vers la file q. Faisant la somme des différentes types d'entrées, on trouve les équations suivantes :
   $$\forall p \in [\![1, N]\!], \qquad \lambda_p = \overline{\lambda}_p + \sum_{q\neq p} \lambda_q r_{q,p}$$
   C'est un système linéaire d'inconnues les λ_p qu'on appelle les *équations de trafic*.
3. La propriété « être sans capture » se traduit formellement ainsi : pour tout p, il existe une suite $p_1, \dots, p_n$ d'éléments de $[\![1, N]\!]$ telle que $r_{p,p_1} r_{p_1,p_2} \cdots r_{p_{n-1},p_n} r_{p_n,N+1} > 0$. Si on note $R = (r_{p,q})_{1\leqslant p,q\leqslant N}$ la matrice de routage entre les files, on cherche une solution au système $\lambda = \overline{\lambda} + \lambda R \iff \lambda(I-R) = \overline{\lambda}$ qui admet une solution si et seulement si la série $K = I + R + R^2 + \cdots$ converge (et dans ce cas, elle vaut $(I-R)^{-1}$). L'unique solution positive et finie et alors donnée par $\lambda = \overline{\lambda} K$. Montrons donc que K est finie. La matrice de transition entre les files (en comptant les sorties du système) est :
   $$\begin{pmatrix} & & & r_{1,N+1} \\ & R & & \vdots \\ & & & r_{N,N+1} \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$
   L'hypothèse de « non capture » signifie qu'on ne peut rester indéfiniment dans R, donc que les états $i \in [\![1, N]\!]$ sont transcients. Cela signifie en particulier que pour tout couple $(i, j) \in [\![1, N]\!]^2$, $\sum_{n\geqslant 0} p_{ij}(n) < +\infty$. Or $p_{ij}(n) = (R^n)_{ij}$ donc cette condition exprime bien le fait que K est finie.
4. Supposons avoir π et des $\tilde{q}_{ij}$ tels que $\forall ij \in \mathcal{E}, \pi(i)\,\tilde{q}_{ij} = \pi(j)\,q_{ji}$ et $\sum_{j\neq i} \tilde{q}_{ij} = -q_{ii}$. Par définition, π est une distribution stationnaire si et seulement si $\pi Q = 0$. Calculons donc $(\pi Q)_i$:
   $$(\pi Q)i = \sum_{j\in\mathcal{E}} \pi(j)\,q_{ji} = \pi(i)\,q_{ii} + \sum_{j\neq i} \pi(j)\,q_{ji} = \pi(i)\,q_{ii} + \pi(i) \sum_{j\neq i} \tilde{q}_{ij} = \pi(i)\,q_{ii} + \pi(i)\,(-q_{ii}) = 0$$
   Ainsi, π est bien la distribution stationnaire de la chaîne.
5. Dans le cas, d'une file M/M/c, la distribution stationnaire est donnée par
   $$\pi(i) = \frac{\lambda^i}{\prod_{n=1}^{i} \mu(n)} \pi(0).$$
   Comme les files sont indépendantes, la distribution jointe $\pi(\vec{n})$ est le produit des distributions marginales $\pi_p(n_p)$, d'où
   $$\pi(\vec{n}) = \prod_{p=1}^{N} \pi_p(n_p) = \prod_{p=1}^{N} \frac{\lambda^{n_p}}{\prod_{r=1}^{n_p} \mu_p(r)} \pi_p(0).$$
6. D'après le théorème démontré précédemment, il suffit de trouver des $\tilde{q}_{\vec{n},\vec{m}}$ tels que
   $$\forall(\vec{n},\vec{m}) \in \mathcal{E}^2,\ \pi(\vec{n})\,\tilde{q}_{\vec{n},\vec{m}} = \pi(\vec{m})\,q_{\vec{m},\vec{n}} \qquad \text{et} \qquad \sum_{\vec{m}\neq\vec{n}} \tilde{q}_{\vec{n},\vec{m}} = -q_{\vec{n},\vec{n}} = \sum_{\vec{m}\neq\vec{n}} q_{\vec{n},\vec{m}}.$$

Comme tous les $\pi(\vec{n})$ sont non nuls, on peut prendre $\tilde{q}_{ij} = \frac{\pi(j)}{\pi(i)} q_{ji}$ et il ne reste qu'à vérifier que les deux sommes sont égales. Pour cela, on va les développer suivant les valeurs non nulles des $q_{\vec{n},\vec{m}}$. on rappelle qu'il y a trois types de variations pour chacune des N files du système : les arrivées exogènes, les sorties du système et les routage entre les files. On commence par calculer la seconde somme car elle est plus simple.

$$\sum_{\vec{m}\neq\vec{n}} q_{\vec{n},\vec{m}} = \sum_{p=1}^{N} \left(\overline{\lambda} \quad + \quad \mu_p(n_p)\, r_{p,N+1} \quad + \quad \sum_{q\neq p} \mu_p(n_p)\, r_{p,q} \right) = \sum_{p=1}^{N} \left(\overline{\lambda} + \mu_p(n_p) \sum_{q=1}^{N+1} r_{p,q} \right) = \sum_{p=1}^{N} \left(\overline{\lambda} + \mu_p(n_p) \right)$$

La première somme est plus technique à calculer car les rôles de $\vec{n}$ et $\vec{m}$ sont inversés (même si c'est toujours $\vec{m}$ qui varie et $\vec{n}$ qui est constant) et il faut veiller à ne pas se perdre dans les indices et leur signification.

$$\begin{aligned}
\sum_{\vec{m}\neq\vec{n}} \tilde{q}_{\vec{n},\vec{m}} &= \sum_{\vec{m}\neq\vec{n}} \frac{\pi(\vec{m})}{\pi(\vec{n})}\, q_{\vec{m},\vec{n}} \\
&= \sum_{p=1}^{N} \left(\underbrace{\frac{\pi(\vec{n}-e_p)}{\pi(\vec{n})} \overline{\lambda}_p}_{\text{arrivées exogènes}} \quad + \quad \underbrace{\frac{\pi(\vec{n}+e_p)}{\pi(\vec{n})} \mu_p(n_p+1)\, r_{p,N+1}}_{\text{sorties du système}} \quad + \quad \sum_{q\neq p} \underbrace{\frac{\pi(\vec{n}+e_q-e_p)}{\pi(\vec{n})} \mu_p(n_p+1)\, r_{q,p}}_{\text{routage de } q \text{ à } p} \right) \\
&= \sum_{p=1}^{N} \left(\frac{\mu_p(n_p)}{\lambda_p} \overline{\lambda}_p \quad + \quad \frac{\lambda_p}{\mu_p(n_p+1)} \mu_p(n_p+1)\, r_{p,N+1} \quad + \quad \sum_{q\neq p} \frac{\lambda_q\, \mu_p(n_p)}{\lambda_p\, \mu_q(n_q+1)} \mu_q(n_q+1)\, r_{q,p} \right) \\
&= \sum_{p=1}^{N} \left(\frac{\overline{\lambda}_p}{\lambda_p} \mu_p(n_p) \quad + \quad \lambda_p\, r_{p,N+1} \quad + \quad \frac{\mu_p(n_p)}{\lambda_p} \sum_{q\neq p} \lambda_q r_{q,p} \right) \\
&= \sum_{p=1}^{N} \left(\frac{\overline{\lambda}_p}{\lambda_p} \mu_p(n_p) \quad + \quad \lambda_p\, r_{p,N+1} \quad + \quad \frac{\mu_p(n_p)}{\lambda_p} (\lambda_p - \overline{\lambda}_p) \right) \qquad \text{en utilisant les équations de trafic (\textit{cf} question 2)} \\
&= \sum_{p=1}^{N} \left(\mu_p(n_p) + \lambda_p\, r_{p,N+1} \right)
\end{aligned}$$

L'égalité sera donc satisfaite si

$$\sum_{p=1}^{N} \overline{\lambda}_p = \sum_{p=1}^{N} \lambda_p r_{p,N+1}$$

qu'on obtient en sommant les N équations de trafic (car $\sum_{p\neq q} r_{q,p} = 1 - r_{q,N+1}$).

Au final, on a démontré que même si les deux hypothèses faites à la question 5, à savoir :

- l'indépendance des files entre elles et
- que les arrivées totales dans la file p suivent un processus de Poisson d'intensité λ_p

ne sont pas vérifiées, les résultats sont les mêmes. Ceci n'est vrai que pour les réseaux de Jackson et dès qu'on en sort, rapidement on n'arrive plus à obtenir des résultats exacts ou à justifier les hypothèses que l'on est obligé de faire.