Contrôle de connaissances FAT - MPRO

Laurent Decreusefond Anaïs Vergne

14 janvier 2019

Pour tout entier n et tout réel positif β, on rappelle que :

$$\int_0^{+\infty} x^n \beta e^{-\beta x} dx = \frac{n!}{\beta^n}.$$

Soit X_1, X_2 et Y des variables aléatoires indépendantes. La variable X_1 (respectivement X_2) suit une loi exponentielle de paramètre μ_1 (respectivement μ_2) et

$$\mathbf{P}(Y = 1) = p_1 \text{ et } \mathbf{P}(Y = 2) = p_2 = 1 - p_1.$$

Une variable aléatoire est dite de type *phase* lorsqu'elle a la loi de X_Y :

$$X_Y = \begin{cases} X_1 & \text{si } Y = 1, \\ X_2 & \text{si } Y = 2. \end{cases}$$

I. Préliminaires

1. Montrer que le temps moyen de service est donné par :

$$\mathbf{E}[X_Y] = p_1 \mathbf{E}[X_1] + p_2 \mathbf{E}[X_2] = p_1 \frac{1}{\mu_1} + p_2 \frac{1}{\mu_2}.$$

   On notera cette quantité par $1/\mu$.

2. Montrer que la variance du temps moyen de service est donnée par :

$$p_1 \frac{2}{\mu_1^2} + p_2 \frac{2}{\mu_2^2} - \frac{1}{\mu^2}. \tag{1}$$

3. Rappeler la probabilité stationnaire d'une file M/M/∞ alimentée par un processes de Poisson d'intensité α et avec des temps de service de loi exponentielle de paramètre β.

II. File M/PH/S/S

On considère d'abord un système avec **un nombre infini de serveurs**, des arrivées selon un processus de Poisson d'intensité λ et des arrivées de type *phase*. On note $X(t) = (X_1(t), X_2(t))$ le processus dont la première (respectivement la seconde) composante compte le nombre de clients de type 1 (respectivement de type 2) en service à l'instant t.

4. Quelle est la nature du processes des arrivées de clients qui seront finalement de type 1 ?

5. Ecrire le générateur infinitésimal de X.

6. Montrer que ce générateur est le même que celui de deux files M/M/∞ séparées, alimentée chacune par des processus de Poisson indépendants. On précisera les paramètre des deux files.

7. En déduire la probabilité stationnaire du processus X.

   On posera
   $$\rho_1 = \frac{\lambda p_1}{\mu_1},\ \rho_2 = \frac{\lambda p_2}{\mu_2}.$$

8. En utilisant le théorème de Kelly, montrer que la probabilité stationnaire de la file M/PH/S/S est telle que
   $$\pi(n) \propto \sum_{n_1+n_2=n} \frac{\rho_1^{n_1}}{n_1!}\frac{\rho_2^{n_2}}{n_2!}$$
   où le signe $\propto$ signifie « proportionnelle à ».

9. En déduire que la probabilité π est la même que celle d'une file M/M/S/S dont on précisera les paramètres.

III. File M/PH/1/∞

On considère maintenant la situation où l'on a un buffer. Le résultat d'insensibilité précédent n'est alors plus valable. On pourra utiliser le résultat suivant valable pour toute file M/GI/1 : le nombre moyen de clients à l'état stationnaire est donnée par (cf [1])

$$\bar{X} = \rho + \frac{\rho^2}{2(1-\rho)}\Big(1 + \frac{\mathrm{var}(\sigma)}{\mathbf{E}\left[\sigma\right]^2}\Big). \tag{2}$$

où $\rho = \lambda/\mu < 1$ avec λ le nombre moyen d'arrivées par unité de temps, σ est une variable aléatoire qui a la loi commune des temps de service et comme espérance $1/\mu$.

Les arrivées forment un processus de Poisson d'intensité λ. Il y a un seul serveur et un buffer de taille infinie. Afin d'étudier les performances de cette file d'attente, on considère le processus de Markov X qui représente le nombre de clients dans le système et la phase du client en service. L'espace d'état est donc :

$$E = \{0\} \cup \{(i, \gamma),\ i \geq 1,\ \gamma \in \{1, 2\}\}.$$

Par exemple, lorsque X est dans l'état $(5, 1)$, cela signifie qu'il y a quatre clients dans le buffer et que le client en service est en phase 1 donc que son temps de service est distribué comme une exponentielle de paramètre μ_1.

10. Quel est temps moyen de séjour le système à l'état stationnaire ?

11. Si $\mu_1 = 1$ et μ est fixé, exprimer $1/\mu_2$ en fonction de p.

12. Montrer que $\bar{X} \sim c(1-p)^{-1}$ quand p tend vers 1 où c est une constante que l'on ne cherchera pas à calculer.

13. En quoi ce comportement est-il différent de celui de la M/M/1 ?

14. Ecrire les coefficients non nuls du générateur infinitésimal de X.

15. Soit π le vecteur représentant la probabilité stationnaire, on pose $x_0 = \pi(0),\ x_i = (\pi(i, 1), \pi(i, 2))$. Ecrire les équations satisfaites par les x_i en utilisant des produits matriciels par blocs.

    On pourra introduire les matrices

    $$D = \begin{pmatrix} -(\lambda+\mu_1) & 0 \\ 0 & -(\lambda+\mu_2) \end{pmatrix}\ U = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\ \textit{et}\ L = \begin{pmatrix} p_1\mu_1 & p_2\mu_1 \\ p_1\mu_2 & p_2\mu_2 \end{pmatrix}$$

    ainsi que les vecteurs

    $$\Lambda = \begin{pmatrix} \lambda p_1 & \lambda p_2 \end{pmatrix}\ \textit{et}\ M = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}.$$

On cherche une solution de ces équations sous la forme $x_i = x_0 R^i$ où R est une matrice 2×2 à déterminer.

16. Trouver l'équation satisfaite par R.

Admetttons que cette équation détermine R de manière unique.

17. Comment déterminer x_0 ?

IV. Blockchain

IV.1. Sans délai de validation

Dans la blockchain, des opérateurs appelés *mineurs* font de savants et longs calculs cryptographiques pour valider un *bloc* de transactions. On suppose que la durée de cette opération, appelée *minage*, suit une loi exponentielle. Il y a toujours des blocs à traiter pour tout le monde.

Dans un premier temps, considérons qu'un mineur informe instantanément toute la communauté de son succès dans le minage d'un bloc. C'est alors lui qui empoche la prime de minage et c'est son bloc qui est rajouté à la chaîne des blocs (*la blockchain*).

On suppose que l'on a deux catégories de mineurs, ceux qui minent à un débit de λ_1 blocs par heure et ceux qui minent à un rythme de λ_2 blocs par heure. On note $X_1(t)$ et $X_2(t)$ le nombre total de blocs minés par chacun à l'instant t.

18. Quelle est la nature des processus X_1 et X_2 ?

19. Ecrire le générateur infinitésimal de $Y = X_1 - X_2$.

20. Montrer que la chaîne de Markov incluse $\hat{Y}$ saute de $+1$ avec probabilité $\lambda_1/(\lambda_1+\lambda_2)$ et de -1 avec la probabilité complémentaire.

21. En utilisant la loi forte des grands nombres, montrer que quand n tend vers l'infini, $\hat{Y}_n$ tend presque-sûrement vers $+\infty$ (respectivement $-\infty$) quand $\lambda_1 > \lambda_2$ (respectivement quand $\lambda_2 < \lambda_1$).

22. Que peut-on en déduire sur la nature de Y (récurrence, transience) quand $\lambda_1 \neq \lambda_2$?

23. Quel est l'intérêt d'avoir toujours plus de puissance de calcul (sous-entendu plus que les autres) ?

24. Donner l'algorithme de simulation du processus Y.

IV.2. Avec délai de validation

Pour tenir compte des délais de transmission, on considère maintenant que lorsqu'un mineur a validé un bloc, la communauté n'en prend connaissance qu'au bout d'une durée de loi exponentielle de paramètre μ. Dans le cas de nos deux mineurs, si les deux chaînes ont des longueurs différentes, la plus longue l'emporte (et le mineur à laquelle elle appartient empoche la prime). Tout le monde se met à travailler à partir de cette nouvelle chaîne. On note X_1 la longueur de la chaîne privée du mineur 1 et X_2 celle du mineur 2. On part d'une situation où les deux mineurs sont d'accord sur la chaîne publique et donc $X(0) = (X_1(0), X_2(0)) = (0, 0)$. Au fur et à mesure que l'un et l'autre mine, X augmente de $(1, 0)$ ou de $(0, 1)$. Dès qu'un bloc est miné, il est envoyé pour communication à l'autre

groupe. Cet envoi n'est pris en compte qu'au bout d'un temps de loi exponentielle de paramètre μ. On suppose que μ est très grand devant λ_1 et λ_2. Si les deux chaînes ont la même longueur, rien ne se passe et les deux mineurs continuent de travailler sur leurs chaînes respectives. Si les deux chaînes sont de longeur différente, celui qui a la plus longue empoche et sa chaîne est ajoutée à la chaîne publique : le processus X retombe alors à l'état $(0,0)$.

Le papier duquel ce modèle est tiré [2] considère que les coefficients non nuls (hors ceux de la diagonale) du générateur de X sont donnés par

Etat initial	Etat final	Taux
(n_1, n_2)	(n_1+1, n_2)	λ_1
	(n_1, n_2+1)	λ_2
	$(0,0)$	μ si $n_1 \neq n_2$
	(n_1, n_2-1)	$\mu_2 n_2$

TABLE 1 – Taux de transition dans un modèle de blockchain.

25. Quelles sont les hypothèses qui manquent dans la description précédente pour que ce modèle corresponde à peu près à la réalité ?

    On s'intéressera au fonctionnement de l'algorithme pas à savoir si les hypothèses sur les lois des variables aléatoires sont pertinentes.

On admet [2] que la probabilité stationnaire existe et est donnée par la formule suivante :

$$\pi(k,l) = \pi(0,0)\,\lambda_1^k \lambda_2^l \sum_{i=0}^{\min(k,l)} \frac{(|k-l|+i)2^i \binom{k+l-i}{k}}{(k+l-i)(\lambda_1+\lambda_2)^i(\lambda_1+\lambda_2+\mu)^{k+l-i}}.$$

Pour $\lambda_1 = 0,6\,h^{-1}$, $\lambda_2 = 0,54\,h^{-1}$ et $\mu = 285\,h^{-1}$, les premières valeurs des composantes de π sont données par

(k,l)	0	1	2	3
0	0,9757	0,0181	0,0003	0,0000
1	0,0020	0,0037	0,0001	0,0000
2	0,0000	0,0000	0,0000	0,0000
3	0,0000	0,0000	0,0000	0,0000

TABLE 2 – Premières composantes de π.

26. Quel est le pourcentage du temps où le mineur 2 empoche la prime ? Combien de fois gagne-t'il plus que le mineur 1 ? Est-ce cohérent avec les vitesses de minage ?

Références

[1] L. Decreusefond et P. Moyal. *Stochastic Modeling and Analysis of Telecom Networks*. 00000. ISTE Ltd and John Wiley & Sons Inc, 2012.

[2] J. Göbel et al. "Bitcoin blockchain dynamics : The selfish-mine strategy in the presence of propagation delay". In : *Performance Evaluation* 104 (oct. 2016), p. 23-41. doi : 10.1016/j.peva.2016.07.001.

Fin du problème

Corrigé

3) C'est une loi de Poisson de paramètre α/β.

4) A l'arrivée d'un client, il est de type 1 avec probabilité p et de type 2 avec probabilité $1-p$. D'après les propriétés d'amincissement du processus de Poisson, les clients de type 1 forment un processus de Poisson d'intensité λp et d'intensité $\lambda(1-p)$ pour ceux de type 2.

5) Les transitions sont données par

Etat initial	Etat final	Taux
(n_1, n_2)	(n_1+1, n_2)	λp_1
	(n_1, n_2+1)	λp_2
	(n_1-1, n_2)	$\mu_1 n_1$
	(n_1, n_2-1)	$\mu_2 n_2$

TABLE 3 – Taux de transition de la file M/PH/∞

6) La première file a pour paramètres λp et μ_1, la deuxième $\lambda(1-p)$ et μ_2.

7) Comme X a le même générateur que le couple de file M/M/∞, il a la même loi donc la même distribution stationnaire. La loi stationnaire est donc le produit des lois stationnaires de chaque file :

$$\pi(n_1, n_2) = e^{-\rho_1} e^{-\rho_2} \frac{\rho_1^{n_1}}{n_1!} \frac{\rho_2^{n_2}}{n_2!}.$$

8) Comme les files M/M/∞ sont des processus réversibles, leur conjonction aussi. Par conséquent, X est aussi réversible. Donc le lemme de Kelly dit que la probabilité stationnaire du processus tronqué est la troncature de la probabilité stationnaire :

$$\pi(n)) \propto \sum_{n_1+n_2=n} \frac{\rho_1^{n_1}}{n_1!} \frac{\rho_2^{n_2}}{n_2!}.$$

9) Vu que l'on a

$$\begin{aligned}
\sum_{n_1+n_2=n} \frac{\rho_1^{n_1}}{n_1!} \frac{\rho_2^{n_2}}{n_2!} &= \frac{1}{n!} \sum_{n_1+n_2=n} n! \frac{\rho_1^{n_1}}{n_1!} \frac{\rho_2^{n_2}}{n_2!} \\
&= \frac{1}{n!} \sum_{n_1=0}^{n} \binom{n_1+n_2}{n_1} \frac{\rho_1^{n_1}}{n_1!} \frac{\rho_2^{n_2}}{(n-n_1)!} \\
&= \frac{\Big(\rho_1+\rho_2\Big)^n}{n!},
\end{aligned}$$

la loi stationnaire est une loi de Poisson tronquée, de paramètre $\rho_1 + \rho 2 = \lambda/\mu$. Ce qui correspond à la loi stationnaire d'une M/M/S/S de paramètres λ et μ.

10) Pour les états $(i,1)$ et $(i,2)$, le nombre de clients est i donc le nombre moyen de clients est donné par

$$\sum_{i=1}^{\infty} i\Big(\pi_{(i,1)} + \pi_{(i,2)}\Big).$$

11) On obtient

$$1/\mu_2 = \frac{1/\mu - p}{1-p}.$$

25) — Chaque mineur n'envoie pas de nouveaux paquets à valider tant que le premier n'a pas été accepté. Sinon on aurait des taux de descente de $k\mu$. Ou alors on admet que le système ne sera que très rarement dans un état avec $k \geq 2$ ou $l \geq 2$ parce que $\mu \gg \lambda_2$.

— La blockchain est un algorithme décentralisé et pourtant chaque mineur doit être capable de déterminer qui a la chaîne la plus longue. C'est impossible sans organe de contrôle central. Le pari est qu'il est rare d'avoir une transaction dans un état autre que les états $(1,0)$, $(1,1)$, $(0,1)$ parce que le taux de validation est très élevé devant les taux de minage.