Chapitre 4.
Modèles de files d'Attente.

4.5. Exercices.

Exercice 1 : cf. sérieTD

Exercice 2. Reprendre l'exercice 1 avec d'autres problèmes de contrôle: sur quelles variables peut-on agir ? Ecrire la fonction coût ainsi que le programme d'optimisation correspondant.

Exercice 3: Pour chacun des modèles à discipline FIFO ci-dessous,

1.Représenter le graphe de Markov et en déduire les équations de Chapman-Kolmogorov.
2.En déduire la solution (par récurrence ou en utilisant la méthode de la fonction génératrice).
3.Retrouver les principales mesures de performance données dans les tableaux 1 et 2.
4.Représenter graphiquement quelque unes de ces mesures de performances. Discuter.
5. Discuter les problèmes d'optimisation associés à ce modèle.

Les solutions sont résumées dans les tableaux 11 et 12 et peuvent être obtenus à l'aide des codes Mathematica en annexe.

(i) **Modèle M/M/1.** $\lambda_k = \lambda \ \forall k;\ \mu_k = \mu \ \forall k$ **;** $K = L = \infty$**.** Montrer que

(a) la loi stationnaire du nombre de clients dans le système est une loi géométrique de paramètre ρ i.e. $\pi_k = \rho^k(1-\rho),\ k = 0,1,2,...$

(b) la loi stationnaire du temps d'attente d'un client quelconque sachant que le serveur est occupé (avec une probabilité ρ) est une loi exponentielle de paramètre $\mu(1-\rho)$. Elle est égale à zéro (avec une probabilité 1-ρ i.e. si le serveur est libre, par conséquent la fonction de répartition du temps d'attente s'écrit $F_w(x) = 1 - \rho e^{-\mu(1-\rho)x},\ x \geq 0$

(c) La loi stationnaire du temps de séjour dans le système (attente+service) est exponentielle de paramètre $\mu(1-\rho)$ i.e. $F_V(x) = 1 - e^{-\mu(1-\rho)x},\ x \geq 0$

(d) Pour ce modèle, la minimisation du temps d'attente et la maximisation de l'utilisation du serveur sont-ils des objectifs compatibles ?

(e) Dans la perspective d'améliorer le temps de service, on étudie les deux scénarios suivants :

Scénario 1 : Acquérir un serveur deux fois plus rapide (un taux de service égal à 2μ .

Scénario 2 : Acquérir un second serveur de même taux μ fonctionnant en parallèle avec l'ancien.

Quelle est la meilleure variante ?

(ii) **Modèle M/M/1/K** La capacité du système est finie : $\lambda_k = \lambda$ si $k \le K$; $\lambda_k = 0$ si $k > K$; $\mu_k = \mu, \quad k = 1,2,...,K.$

Exemples: (1) Système constitué d'un processeur et d'une mémoire de capacité finie, (2) Une machine de production et une surface capacité de stockage limitée. (3) Un service de maintenance : les clients sont les machines en panne ; le serveur est le réparateur.

(ii) Modèle M/M/1/∞/L (noté habituellement M/M/1//L). La source des clients est finie : $\lambda_k = \lambda(L-k)$ si $k \le L$; $\lambda_k = 0$ λk= si $k > L$; $\mu_k = \mu, \quad k = 1,2,...$

Exemples :

- Système constitué de L machines (les clients) et d'un réparateur (le serveur) ; la capacité d'attente est illimitée (*Repairman problem*).
- Une machine de production et une surface à capacité de stockage limitée
- Réseau d'ordinateur : Un serveur central connecté à L terminaux.

Exercice 4 : On considère deux ordinateurs qui sont reliés par une ligne de $64\,kbits/\sec$ et 8 applications parallèles se partagent cette ligne. Chaque application génère un trafic poissonien de 2 paquets/seconde en moyenne. Le concepteur doit choisir entre deux solutions : (i) La première est de dédier une bande de base de 8 kbits/seconde à chaque application. Dans ce cas, chaque ligne de 8 $kbits/\sec$ agit comme une file d'attente indépendante (illimitée, FIFO) de taux de service $\mu = 4$ paquets/seconde. (ii) La seconde solution est d'utiliser un accès multiple à la même ligne de transmission de $64\,kbits/\sec$. Cela revient à un seul système de taux de service μ=4x8=32 paquets/seconde et un taux d'arrivées de $\lambda = 2 \times 8 = 16$ paquets/s. Quelle est la meilleure variante ?

Exercice 5 : (retour à l'exemple (ii).3). Les demandes provenant des terminaux sont traitées en FIFO et une seule à la fois. L'usager a un comportement cyclique avec une phase de réflexion (usager oisif), pendant laquelle le programmeur réfléchit et tape sa demande, suivie d'une phase d'attente de la réponse (usager actif). Le temps de réflexion suit une loi exponentielle de paramètre λ. Le temps de traitement de la demande suit une loi exponentielle de paramètre μ (ce temps de traitement comprend le temps de chargement du programme correspondant et le temps d'exécution). Soit $X(t)$ le nombre d'usagers actifs à l'instant t. Si i usagers sont actifs à l'instant t, alors L-i usagers sont susceptibles de passer à l'état actif pendant la période $(t, t+h)$.

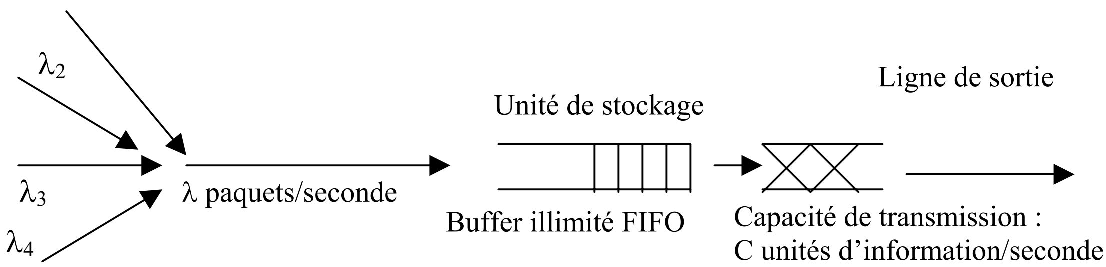
(1).Quelle est la condition d'existence d'un régime stationnaire ? (2) Quelle est la fraction de temps pendant laquelle le système est oisif i.e. sans usagers actifs. (3) Quel est le nombre moyen d'usagers actifs en régime stationnaire (4)Quel est le temps de réponse moyen (durée moyenne de la période d'activité d'un usager.

Application numérique : $M = 4, \quad 1/\lambda = 25 \quad \sec ondes \; 1/\mu = 100 \quad \sec ondes$.

Exercice 6 : On considère un système *M/M/m* de paramètres $\lambda=1$ et $\mu=2$ Déterminer le nombre optimal de serveurs qui minimise le temps d'attente dans le système.

Exercice 7: Même question pour un système M/M/1/K (λ,μ), lorsqu'on donne:c_1=coût unitaire d'exploitation du serveur, c_2=coût unitaire de séjour d'un client, c_3=coût unitaire d'une place dans la file d'attente, c_4=coût unitaire de refus d'un client qui demande service. (1) Trouver la valeur optimale de la capacité K* qui minimise les coûts unitaires d'exploitation du système. (2) Trouver les valeurs optimales de la capacité K* et du taux de service minimisant ces coûts.

Exercice 8 : Chaque nœud d'un réseau à commutation de paquets peut-être modélisé comme un système de files d'attente. On considère ici le processus de mémorisation et de transmission des informations dans un concentrateur muni de n lignes d'entrées (en provenance des autres terminaux et autres nœuds du réseau) et une ligne de sortie(figure 1). La longueur moyenne d'un paquet est de $\theta = 1\ 000\ \ bits$ et la capacité de la ligne de sortie est de $C = 9\ 600\ bits/\sec onde$, de telle sorte que le temps de service (transmission du paquet) suit une loi exponentielle de paramètre $\theta C = 9.6$ paquets/seconde. Chaque ligne d'entrée a une capacité de 4800 bits/seconde et délivre un trafic Poissonien de taux $\lambda_i = 2$ paquets/seconde. ($i = 1,2,...,n$). On prendra pour l'application numérique $n = 4$.
(1) Montrer que le flux des arrivées (de paquets) au concentrateur est poissonien de taux λ=8 paquets par seconde (Justifier votre réponse sans démonstration).(2) Quel est le nombre moyen de paquets dans le buffer et dans le concentrateur ? (3) Quel est le temps moyen d'attente (dans le buffer) et de séjour d'un paquet dans le concentrateur (attente dans le buffer+temps de transmission) ? (4) On cherche à dimensionner la capacité de la ligne de manière à minimiser le temps de séjour d'un paquet dans le concentrateur. Quelle est la capacité optimale ?



Exercice 9 : Un atelier comprend 3 machines automatiques entretenues par un réparateur La durée de bon fonctionnement de chaque machine est une variable aléatoire de loi exponentielle de moyenne $1/\lambda = 9$ heures. Le réparateur ne peut réparer qu'une machine à la fois et le temps de réparation est une variable aléatoire de loi exponentielle de moyenne $1/\mu = 2$ heures. Soit $X(t)$ le nombre de machines en panne à l'instant t.
(i) Calculer les probabilités d'état stationnaires et le nombre moyen de machines en panne. Quel est le temps moyen d'oisiveté du réparateur ? Quel est le temps moyen d'immobilisation d'une machine en panne ? Interpréter les débits absolus et relatifs en

termes de productivité du réparateur. Quelle est la productivité de l'atelier (compte tenu des pannes) si chaque machine produit $n = 5$ articles par unités de temps

(ii) On suggère en guise d'approximation grossière que le parc de machines est infini (grand) de telle sorte que le processus de pannes soit poissonien avec un taux de pannes de 3 pour chaque période de 9 heures. Comparer les résultats avec ceux du point (i).

(iii) On se place dans le cas (ii) (parc infini) avec les même données. Trouver le nombre optimal de réparateurs à affecter à la maintenance des machines si le salaire (unitaire) d'un réparateur est $C_1 = 10\ D.A./heure$ et le coût unitaire d'immobilisation d'une machine est $C_2 = 20\ D.A./heure$.

Exercice 10: Les mesures effectuées sur le portail d'un réseau montrent que les paquets de données arrivent selon un processus de Poisson de moyenne 125 paquets par seconde. Le portail met environ 2 millisecondes pour les transmettre. Afin d'analyser les performances du portail, on opte pour un modèle M/M/1/K. Le paramètre K=8 représente le buffer (zone de stockage des paquets en attente de transmission). (i) Peut-on considérer ce système comme stable ? Eventuellement sous quelle condition ? (ii) Quelle est la charge du système ? (iii) Déterminer la probabilité de perte d'un paquet ? (iv) Quel est le taux d'arrivée réel des paquets (i.e. les paquets admis pour transmission) ? (v) En déduire le taux de perte des paquets (overflow) ; (vi) Quel est le débit relatif ? absolu ? (qui mesurent en quelque sorte les proportion et nombre espéré de paquets réellement transmis). (vii) Quel est le temps moyen de séjour d'un paquet dans le portail ? (viii) Comment choisir la capacité du buffer pour limiter la probabilité de perte à une valeur inférieure à 10^{-4} ? (ix) Utiliser le modèle M/M/1 pour répondre aux questions ci-dessus ? Que peut-on suggérer ? Avez-vous une explication ?

Exercice 11. Une clinique dispose d'un service d'urgence tenu par un seul médecin. Les malades se présentent selon un processus de Poisson de taux 96 malades par jour (24 heures) et les durées des soins sont indépendantes et suivent une loi exponentielle de moyenne 12 minutes pour chaque malade. Les malades sont soignés dans le cabinet du médecin suivant leur ordre d'arrivée et il n'y a pas de limitation de places dans le service d'urgence.

A. Donner la notation de Kendall de cette file d'attente ; préciser ses paramètres λ et μ. (1) Montrer que la condition d'ergodicité est vérifiée. (2) Donner la loi du nombre de malades dans le système en régime stationnaire et en déduire la probabilité qu'il y ait 4 malades en attente. (3) Déterminer les mesures de performance de cette file d'attente.

B. On souhaite que le nombre moyen de malades en attente dans la salle soit $\leq 1/2$ (condition C). (a) Une première solution est d'agir sur le temps de soin. Quelle doit être la durée moyenne de soin minimale pour que la condition C soit vérifiée ? Quelle est alors la probabilité qu'un malade attende plus de *2* heures ? (b) Une solution alternative serait d'affecter d'autres serveurs identiques parallèles, chacun assurant le même temps moyen de service de *12* minutes. (i) Déterminer le nombre minimal m de médecins nécessaires pour assurer un régime stationnaire du système. (ii) *Expliquer* (sans calculs) comment déterminer le nombre de serveurs vérifiant la condition C.

Exercice 12. Un *Cyber-Espace* comporte m postes. Les clients arrivent selon un processus de Poisson de paramètre $\lambda=2$ clients/heure. L'analyse statistique des données a montré que la durée T de la connexion de chaque client suit une loi exponentielle de moyenne $E(T)=30$ minutes. Le client qui trouve tous les postes occupés va à la recherche d'un autre *Cyber*. (i) Y a t-il une condition d'existence d'un régime stationnaire ? Si oui, laquelle ? (ii) Dans l'affirmative, évaluer la distribution de probabilités stationnaire. (iii) Calculer le nombre moyen de postes occupés et le nombre moyen de clients dans le *Cyber*. (iv) Quelle est la proportion de clients satisfaits (servis)? Non satisfaits ? (v) Combien de postes devrait-on avoir pour satisfaire 95% de la clientèle ? (vi) Combien en faudrait-il pour satisfaire 80% de la clientèle lorsque la durée moyenne de connexion est heure.

Exercice 14: Dans une fabrique de semelle de chaussures on dispose de deux ateliers contenant plusieurs machines chacun (un atelier pour les chaussures d'hommes et l'autre pour les chaussures de femmes). La qualité du travail dépend du bon réglage des machines. Dès qu'une machine se dérègle (ou tombe en panne), on fait appel à l'équipe de maintenance (constituée de 2 personnes) pour réglage (réparation). Les études statistiques sur les dérèglements et les temps de remise en marche ont montré que les pannes se produisait à raison de 4 pannes l'heure en moyenne selon un processus de Poisson et que le temps de remise en marche était exponentiellement distribué de paramètres 6 réparations/heure.

(I) **Etude de la situation actuelle.** (1) Donner la notation de Kendall du modèle dont on précisera les paramètres ; (2) Quelle est la charge ou intensité du trafic ? que peut-on en déduire ? (3) En moyenne combien de machines restent inutilisées en permanence ? (4) Quel est le nombre moyen de machines en attente de réglage ? Combien chaque machine attend-elle en moyenne sa prise en charge ? (5) Si l'heure d'immobilisation d'une machine revient à 1000 DA, quelle est la perte (ou plutôt le manque à gagner) sur un mois ? (on suppose que la production est continue sur 30 jours à raison de 8 heures par jour). (6) Quel est le coût dû seulement à l'attente d'être réparé ? (7) Quelle est la proportion de temps où l'équipe de maintenance est inactive ?

(II) Variante d'amélioration. Pour améliorer le service on examine les différents scénarios suivants :

Scénario 1 : On renforce l'équipe existante qui passe à trois personnes de telle sorte à augmenter le taux de réparation à 7 réparations par heure. Mêmes questions qu'en (I). Y a-t-il une économie de réalisée, si oui de combien est-elle?

Scénario 2 : On met en place une seconde équipe de maintenance avec un coût fixe de 50 000 DA le mois. Quel est ce nouveau modèle ? Préciser ses paramètres. Même questions que précédemment. Y a-t-il une économie de réalisée, si oui de combien ?

Scénario 3 : On garde comme précédemment deux équipes (avec le même coût fixe de mise en place), mais chacune s'occupant d'un seul atelier. L'atelier 1 (homme) comprend 25% du parc machine contre 75% pour le second. Cette solution permet de réduire les déplacements des équipes de maintenance et augmenter ainsi le taux de réparation qui passe à 8 réparations de l'heure. Quel est ce nouveau modèle ? Calculer les mesures de performance associées et l'économie réalisée dans ce cas, s'il y en a une.

Enfin, conclure quel est le meilleur choix parmi les trois scénarios. Discuter.

Exercice 18 : La poste du quartier comprend deux (02) guichets, le premier G_1 spécialisé pour les retraits et envoi de mandats, chèques etc…et le second G_2 pour les affranchissements et vente de timbres divers. Les processus d'arrivées de clients à ces guichets sont poissoniens de taux $\lambda_1=4$ clients/heure et $\lambda_2=2$ clients/heure respectivement. On suppose que le temps de service au niveau de chaque guichet obéit à une loi exponentielle de moyenne $\mu=12$ minutes (la même pour chaque guichet). En vue d'étudier les possibilités d'amélioration des prestations on se propose de comparer les deux variantes :
V1 : conserver cette spécialisation des guichets (chacun avec sa propre file d'attente).
V2 : Regrouper les deux guichets en un seul avec une file commune, les temps de service de chacun des employés ayant la même loi exponentielle de moyenne 12 minutes.
(1).Comparer les deux variantes du point de vue du temps d'attente et de séjour de chaque client. (2) Quel est le nombre optimal de guichets qui minimisent les coûts unitaires d'exploitation si on connaît (a) le coût unitaire d'exploitation d'un serveur=10 unités monétaires/heure, (b) le coût unitaire de séjour (attente) d'un client dans le système (file)=1 unité/heure.