

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université des Sciences et de la Technologie Houari Boumediene
Faculté des Sciences Mathématiques Pures et Appliquées



**MEMOIRE PRESENTE POUR L'OBTENTION DU GRADE DE
MAGISTER**

EN : MATHEMATIQUES

SPECIALITE: RECHERCHE OPERATIONNELLE (METHODES STOCHASTIQUES)

par : **KABRI FADHILA**

THEME

**ETUDE DE MODELE D'ATTENTE M/M/C
AVEC RAPPELS**

Soutenue publiquement le :
Devant le jury composé de :

M ^r M. DJEDOUR	Professeur	U.S.T.H.B	Président
M ^r A.AISSANI	Professeur	U.S.T.H.B	Directeur de thèse
M ^r M.AIDER	Maître de conférences	U.S.T.H.B	Examinateur
Mme I.BOUCMAKH	Maître de conférences	U.S.T.H.B	Examinatrice
Mme H.GUERBIENNE	Chargée de cours	U.S.T.H.B	Examinatrice

Remerciements :

Ce travail a été réalisé sous la direction du professeur Amar Aissani auquel je voudrai exprimer toute ma gratitude et respects.

Je tiens que Mme. Guerbienne trouve dans ces quelques phrases l'expression de ma gratitude pour ses directives précieuses durant l'année théorique.

Je souhaite exprimer ma gratitude au professeur Djedour (USTHB) pour l'honneur qu'il me fait de présider le jury de ma thèse.

J'exprime ma reconnaissance à Mme.Bouchmakh (USTHB), Mme. Guerbienne (USTHB), et Mr.Aider (USTHB) d'avoir accepté d'examiner mon travail.

Je remercie également Mr.M.F.Neuts de « university of Arizona USA », Mr.J.R Artalejo de « university of complutense of Madrid » d'avoir répondu à mes correspondances, je leur suis très reconnaissante.

Je n'oublierai surtout pas de saluer vivement tous mes enseignants, particulièrement Mr.Bentarzi.

Je remercie mon père, pour m'avoir soutenu et avoir cru en moi.

Enfin, que toute personne qui m'a aidé de près ou de loin, trouve ici le témoignage de ma profonde reconnaissance.

SOMMAIRE

INTRODUCTION GENERALE -----	1
1. Phénomènes d'attente et processus stochastiques -----	3
1.0.Introduction -----	4
1.1.Structure de base d'un système de files d'attente -----	5
1.2.Classification des systèmes d'attente-----	5
1.3.Analyse mathématique -----	6
1.4.Processus des arrivées -----	7
1.5.Processus des Services -----	7
1.5.1.Le temps de service résiduel -----	8
1.6.Processus de Poisson et la loi exponentielle -----	8
1.6.1.Processus de comptage -----	8
1.6.2.Processus de poisson -----	9
1.6.3. La loi exponentielle -----	10
1.7.Processus de Naissance et de Mort -----	11
1.7.1.La notion d'état -----	11
1.7.2.Processus de Markov -----	11
1.7.3.Les processus de naissance et de mort -----	12
1.7.4.Solution stationnaire obtenue par l'écriture matricielle -----	13
1.8. Etude de quelques modèles classiques markoviens -----	14
1.8.1.La file M/M/1-----	14
1.8.2.La file M/M/c -----	16
1.8.3.Modèles à capacité limitée (Queues with truncation) -----	17
1.8.4. Modèle M/M/c/c (modèle d'Erlang) -----	18
2.Système de file d'attente M/M/c avec rappels -----	20
2.0.Introduction -----	21
2.1.Files d'attente avec rappel -----	22

2.2.Système de file d'attente M/M/1 avec rappel constant	23
2.2.1.Description du modèle	23
2.2.2.Calcul de la distribution stationnaire	23
2.2.3.Nombre moyen de clients dans le système	26
2.2.4. Probabilité de blocage	26
2.3.M/M/1 avec taux de rappel linéaire	27
2.3.1. Nombre moyen de clients dans l'orbite	30
2.3.2. Probabilité de blocage	30
2.4.Système M/M/2	30
2.5.Système M/M/c	32
2.5.1.Formules explicites de caractéristiques de performance	33
2.5.2. Remarque sur la nature numérique de la distribution stationnaire	35
2.6.Conclusion	36
3.Approximation du système M/M/c avec rappels	37
3.0.Introduction	38
3.1.Modèle tronqué	39
3.1.1.Description du modèle	39
3.1.2.Calcul de la distribution stationnaire	40
3.1.3. Caractéristiques de performance du système tronqué	44
3.1.4.Relation entre le système tronqué et le système initial	45
3.1.5.Convergence de solutions du modèle tronqué vers celles du modèle	48
original	
3.1.6. Choix du niveau de troncature M	49
3.2.Modèle tronqué généralisé	49
3.2.1.Description de quelques modèles tronqués généralisés	50
3.2.2.Solutions matricielles géométriques.	52
3.2.3. Mesures de performance	57
3.2.4.Solution algorithmique du modèle X^F	58
3.2.5.Solution algorithmique du modèle $X^{(NR)}$	63
3.3.Conclusion	68
3.4.Résultats numériques	69
3.4.1.Différentes mesures de performance obtenues par la méthode de troncature directe	69

3.4.2.Résultats numériques par le modèle $X^{(F)}$	70
3.4.3.Choix du niveau de troncature	71
3.4.4.Résultats numériques par le modèle $X^{(NR)}$	72
3.4.4.1.Choix du niveau de troncature	72
3.4.5.Comparaison numérique des résultats	74
3.4.6.Conclusion	75
Conclusion générale	76
Annexe	77
Références	78

INTRODUCTION GENERALE

Ce travail s'inscrit dans la continuité des études intensives menées dans le but de comprendre et de rendre compte des comportements stochastiques des modèles markoviens de type M/M/c et rappels exponentiels.

La théorie des files d'attente est bien adaptée à l'étude des performance des systèmes concrets (systèmes informatiques[45, 40], systèmes de télécommunication[44],...).Elle fut introduite pour la première fois dans sa version classique pour l'analyse des systèmes téléphoniques [18, 19].

Cette théorie classique s'est très vite montrée inefficace face à des systèmes réels de plus en plus complexes .Dés la fin des années 1940 ,des chercheurs tels que Kosten [39] et Wilkinson [46] ont mis en évidence les limites de la théorie classique qui ne permettait pas d'expliquer le comportement stochastique des systèmes téléphoniques où les abonnés répétaient leur appels en recomposant le numéro plusieurs fois jusqu'à l'obtention de la communication.

Dans ce contexte ,chaque appel bloqué génère un flot de demandes répétées indépendamment du reste des clients dans le groupe de rappel. Ainsi, dans la politique classique de rappel, les intervalles entre deux tentatives successives sont exponentiellement distribuées avec taux $\mu_j = j\mu$, quand la taille de l'orbite est j . Cependant, des applications récentes aux protocoles de communication (protocoles CSMA, disciplines AUTO-REPEAT) et aux réseaux locaux, montrent qu'il existe des situations de files d'attente dans lesquelles le taux de rappel est indépendant du nombre de clients(s'il y en a) dans l'orbite, c'est à dire , le taux de rappel est $\mu_j = \alpha (1-\delta_{0j})$, δ_{0j} est la fonction de Kronecker. Cette deuxième politique est appelée discipline de rappel constant et fut introduite par Fayolle[29] pour modéliser un système d'échange de téléphone. Cette politique de rappel est utile pour la modélisation du phénomène de rappel en communication et en réseaux informatiques où des tentative répétées sont faites par des unités du processeur indépendamment du nombre du messages fournis en chaque nœud du réseau. Artalejo et Gomez Corral [4] combinèrent les deux politiques en définissant une politique de rappel linéaire avec le taux de rappel $\mu_j = \alpha (1- \delta_{0j}) + j\mu$; Cette dernière peut être considérée pour la modélisation du système de réseau informatique qui consiste en un groupe de processeurs connectés avec une unité centrale de transmission (UCT). Si un processeur émet un message , il l'envoie d'abord à l'UCT. si la transmission médiane est possible, alors L'UCT envoie immédiatement le message, si non ce dernier sera stocké dans un espace –Mémoire appelé « buffer » et après un certain temps aléatoire, l'UCT doit réessayer la transmission. En supposant que ce temps aléatoire est exponentiellement distribué, on peut considérer un taux de rappel le plus simple possible en supposant qu'il y a deux contributions à l'intensité des rappels .La première est une constante α intrinsèque au réseau , alors que la deuxième $j\mu$ dépend de nombre de messages se trouvant dans le « buffer » . De plus la supposition que les intervalles de temps séparant les appels successifs sont exponentiellement distribués permet de tenir en compte de la présence simultanée des deux types de rappel(constant et classique)avec un taux $\alpha (1- \delta_{0j}) + j\mu$, quand la taille de l'orbite (buffer) est j .

La prise en considération des appels répétés introduit de grandes difficultés analytiques. En effet des résultats analytiques explicites n'existent que pour un certain nombre de file d'attente avec rappel particulières, avec des hypothèses contraignantes sur certains paramètres tels que le nombre de serveurs (un seul serveur), les distributions de temps de rappel et d'arrivée (loi exponentielle) et l'état du système (régime stationnaire) par exemple, alors que pour beaucoup d'autres, les résultats obtenus sont extrêmement limités.

Ainsi, pour les files d'attente avec rappels et plusieurs serveurs qui ont été analysés sous la politique de rappel classique, la distribution d'équilibre de l'état du système s'exprime à l'aide des intégrales curvilignes [17] ou comme limite de fraction continues étendues [43]. D'un point de vue analytique, chacune des solutions sont des tentatives significatives, mais l'implémentation pratique exige une variété d'approximations et de méthodes de troncature.

Plusieurs articles intéressants ont été dévolus à l'approximation du système d'attente M/M/c avec rappel. tels que R. Wilkinson [46]. Artalejo, Gomez Corral et Neuts [7] utilisèrent des méthodes géométriques matricielles pour analyser la file M/M/c avec taux de rappel constant où l'analyse stationnaire n'exige pas une troncature de l'orbite. Falin et Artalejo [26] étudient une file à plusieurs serveurs différente dans la quelle les clients rejoignent une ligne d'attente classique et l'orbite dépend du nombre de clients dans la file. Choi et Pearce [12] étudient une politique de rappel indépendante de la taille de l'orbite pour une file M/M/1 avec temps de rappel généraux, c'est à dire à chaque instant où l'orbite n'est pas vide, on commence un processus de renouvellement de tentatives répétées.

L'objet de notre travail est de contribuer à cet intérêt particulier pour les files d'attente avec rappel et plus particulièrement aux systèmes M/M/c avec rappels exponentiels dans la politique classique de rappel. En effet un examen de la littérature montre que seulement le cas $c \leq 2$ a été étudié de manière analytique [17, 21, 34]. On propose d'étudier quelques méthodes d'approximation numériques. En l'absence de résultats théorique exacts, l'efficacité de telles approximations est testée à l'aide de quelques méthodes algorithmiques. Ce mémoire est organisé comme suit :

Le premier chapitre comprend une description d'un système d'attente ainsi que les définitions et notions mathématiques sur les processus stochastiques en usage dans la théorie des files d'attente. On présente également une synthèse sur les systèmes de files d'attente classiques de type M/M/1, M/M/c et M/M/c/c.

Le deuxième chapitre est consacré à l'analyse du modèle M/M/1 avec rappel [27]. Les deux types de rappel : constant et classique sont considérés. On présente ensuite une analyse du modèle M/M/c avec rappels.

Dans le troisième chapitre, on se penchera sur l'approximation des modèles markoviens de type M/M/c et rappels exponentiels moyennant deux méthodes :

La première est basée sur la méthode de troncature directe qui nous donne un système traiter à l'aide d'une solution algorithmique. La deuxième est basée sur des méthodes de troncature plus sophistiquées où en imposant une hypothèse simplifiée qui donne un modèle de file d'attente auxiliaire avec un état du système infini. Ici la solution stationnaire est donnée par la théorie de Marcel. Neuts [41].

Dans la deuxième partie du troisième chapitre à caractère pratique, après un aperçu des programmes utilisés pour la validation des approximations. Une comparaison des résultats numériques permet de tirer la meilleure approximation.

Enfin, une conclusion générale résumera les différents résultats obtenus à travers ce travail.

PREMIER CHAPITRE

***PHENOMENES D'ATTENTE
ET
PROCESSUS STOCHASTIQUES***

1. Phénomènes d'attente et processus stochastiques

1.0. introduction

Dans ce chapitre, nous présentons les approches de base de la théorie des files d'attente. Cette dernière est considérée comme une branche de la théorie des probabilités, dont l'objectif est l'étude mathématique des queues. La formation d'une file est un phénomène commun qui a lieu chaque fois que les demandes de services auprès d'un serveur excèdent la capacité de ce serveur. Cette capacité est parfois imposée .

La théorie considère des *clients*, qui se déplacent dans un réseau de serveurs, qui les traitent. Lorsque plusieurs clients tentent simultanément d'obtenir un service, certains doivent patienter et attendre dans des *files d'attente*, ou *tampons*. Le vocabulaire est fixé et conventionnel. Les clients peuvent être des humains faisant la queue à un guichet pour obtenir un ticket de train (le serveur est l'employé qui délivre les tickets). Ils peuvent représenter les paquets qui sont aiguillés vers une ligne de transmission et qui attendent la disponibilité de la ligne (deux paquets pouvant être en contention). Le serveur est alors l'automate d'émission, et le temps de service est le temps d'émission.

Les demandes de prise de circuit dans un système téléphonique sont un autre type de clients. Dans ce cas, les serveurs forment un groupe: ce sont les circuits qui desservent la direction demandée.

Dans les différents cas, la file possède des constituants qu'on se propose de rappeler dans ce chapitre.

Les deux aspects principaux développés dans ce chapitre sont l'étude de certains processus stochastiques (processus de poisson et processus de Markov..) et l'analyse de système d'attente classique M/M/1, M/M/c et M/M/c/c.

1.1. Structure de base d'un système de files d'attente

Un système de file d'attente peut être schématisé par la figure ci - dessous :

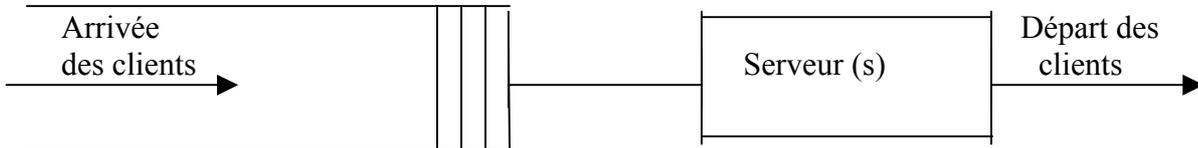


Figure 1.1.

L'unité du flot considéré est dite client . une séquence de clients arrive chez un serveur ou un centre de service (dite encore canal), si le client arrivant trouve le serveur occupé, il joint la file d'attente associée à ce serveur. A un certain moment, le client est sélectionné pour être servi suivant une règle dite politique ou discipline de service. Le service attendu est alors fourni, et le client quitte le système.

1.2. Classification des systèmes d'attente

Pour décrire ce système de façon précise, il faut pouvoir spécifier:

- La nature stochastique du processus des arrivées (flux d'entrée) qui est défini par la loi des intervalles séparant deux arrivées consécutives.
- La distribution du temps de service .
- Le nombre S de stations de service qui sont montés en parallèle. On admet généralement que les temps de service correspondants suivent la même distribution et que les clients qui arrivent forment une seule file d'attente.
- La discipline de service (quand un serveur se libère, quel client choisit-il) .
- La capacité N du système. Si $N < \infty$, certains clients qui arrivent vers le système n'ont pas la possibilité d'y entrer.

Et pour identifier un système d'attente, le formalisme suivant a été proposé et est unanimement adopté:

$$A / B / S / N / Z$$

La première lettre identifie la loi du processus des arrivées, la seconde le processus des services, avec dans les deux cas les conventions:

- M : loi "sans mémoire" (arrivées poissonniennes, service exponentiel);
- D : loi constante (déterministe);
- E_K : loi "Erlang-k"
- H_K : loi "hyper exponentielle" d'ordre k ;
- GI : loi générale, les variables successives étant indépendantes;
- G : loi générale, sans hypothèse d'indépendance.

S donne le nombre de serveurs.

N identifie la taille du système d'attente (si cette valeur est omise, elle est supposée infinie).
Z indique la discipline de service.

La discipline de service détermine l'ordre dans lequel les clients sont rangés dans la file et y sont retirés pour recevoir un service. Les disciplines les plus courantes sont :

a) FIFO (First In First Out) : les clients sont servis dans leur ordre d'arrivée.

b) LIFO (Last In First Out): cela correspond à une pile, dans laquelle le dernier client arrivé (posé sur la file) sera le premier traité (retiré de la pile).

c) RANDOM (Aléatoire) : le prochain client qui sera servi est choisi aléatoirement dans la file d'attente.

Remarque : Lorsque le dernier symbole de la notation n'est pas précisé, il est sous-entendu que $Z = \text{FIFO}$

Par exemple, M/M/1 fait référence au modèle de base: arrivées poissonniennes, service exponentiel, un seul serveur, file d'attente non bornée.

Dans une file d'attente, un paramètre de première importance est le taux d'utilisation, noté traditionnellement ρ . C'est le produit du taux d'arrivée des clients par le temps moyen de service:

$$\rho = \lambda \cdot E(s).$$

Où s est la variable aléatoire du temps de service.

1.3. Analyse mathématique

L'étude mathématique d'un système d'attente se fait par l'introduction d'un processus stochastique défini de façon appropriée. En premier lieu, on s'intéresse au nombre $N(t)$ de clients se trouvant dans le système à l'instant t ($t \geq 0$).

En fonction des quantités qui définissent la structure de système, on cherche à calculer :

- Les probabilités d'état $p_k(t) = \text{pr}(N(t) = k)$ qui définissent le régime transitoire du processus stochastique $\{N(t), t \geq 0\}$.
- Le régime stationnaire du processus stochastique est défini par :

$$P_k = \lim_{t \rightarrow \infty} P_k(t) = \text{Pr}(N = k)$$

($k = 0, 1, 2, \dots$). On écrira N au lieu de $N(+\infty)$.

A partir de la distribution stationnaire du processus $\{N(t), t \geq 0\}$, on pourra obtenir d'autres caractéristiques d'exploitation du système telles que :

$\overline{N_s} = E(N_s)$: le nombre moyen de client dans le système.

. $\overline{W}_{s(q)} = E(W_{s(q)})$: la durée d'attente d'un client dans le système (la file).

. B : la probabilité de blocage ; c'est la probabilité que tout les serveurs sont occupés.

On se contentera par la suite de déterminer le régime stationnaire d'un phénomène d'attente .

Il existe des systèmes d'attente dont l'évolution temporelle n'est pas déterminée par le processus $\{N(t), t \geq 0\}$, ces systèmes peuvent être décrits par des processus bivariés dans la description de l'état du système, comme c'est le cas des files d'attente avec rappels. Il est difficile de déterminer des formules explicites des distributions limites pour de tels modèles.

1.4.Processus des arrivées

Pour aller plus avant dans l'étude des propriétés du trafic, il va être nécessaire de passer en revue les deux composantes dont il dépend. A savoir, les arrivées de clients et leur service.

On observe les arrivées de clients à l'entrée du système qui sont générées d'une source fini ou infinie. Pour décrire le phénomène, la première idée venant à l'esprit sera d'utiliser l'intervalle du temps entre arrivées successives, ou bien le nombre des arrivées dans un intervalle donné.

Pendant la durée t , $N(t)$ arrivées se produisent. On chiffre le volume du flux arrivant par le taux d'arrivée dont la définition intuitive est:

$$\lambda = \lim_{t \rightarrow \infty} \frac{N(t)}{t} .$$

On pourra aussi estimer l'intervalle moyen entre arrivées: c'est l'inverse de la quantité précédente.

Le plus simple et le plus courant modèle est celui des arrivées complètement aléatoires, qui est caractérisé par le processus poissonien, dont les propriétés sont bien connues et développées dans maints ouvrages.

1.5.Processus des Services

Le processus de service pourra être d'une complexité extrême, mais on se borne le plus souvent à supposer que chaque durée de service est indépendante des autres, et qu'elles obéissent toutes à une même loi de distribution dite distribution de service: on parle de variables indépendantes et identiquement distribuées (i.i.d.). On décrira cette loi par sa distribution de probabilité :

$$B(x) = P\{\text{temps de service} \leq x\}$$

Bien qu'ils y aient quelques résultats théoriques concernant le processus de service dans les systèmes de files d'attente et qui peut être décrit d'une façon efficace par des distributions générales, en pratique, il est utile de supposer que la distribution est d'un type donné, caractérisée par un ensemble de paramètres.

On va à présent, décrire la distribution de service exponentielle (on se restreindra aux distributions utilisées dans notre étude) .

1.5.1. Le temps de service résiduel

Supposons qu'on observe un service de durée X qui a déjà atteint la durée y . On s'interroge sur la distribution de la durée restant à courir jusqu'à la fin de ce service (notons X_y la variable aléatoire correspondante. Dire que $X_y > x$ revient à dire que la durée totale sera supérieure à $y+x$, c'est une probabilité conditionnelle. On a alors pour la distribution de X_y :

$$P(X_y > x | X > y) = P(X > x + y | X > y)$$

$$= \frac{P(X > x + y)}{P(X > y)}$$

$$= \frac{1 - B(x + y)}{1 - B(y)}$$

De façon équivalente, la densité de X_y sera

$$b(x | y) = \frac{b(x + y) dx}{1 - B(y)} \quad (1.1)$$

1.6. Processus de Poisson et la loi exponentielle

1.6.1. Processus de comptage

La description mathématique d'un flux d'événements aléatoires peut se faire de deux manières différentes :

- On considère le nombre d'événements $N(t)$ se produisant dans l'intervalle de temps $[0, t]$ et on détermine la distribution de cette variable aléatoire discrète. Le processus stochastique

$\{N(t), t \geq 0\}$ est appelé processus de comptage.

- On considère les intervalles de temps qui séparent les instants d'apparitions de deux événements consécutifs. Ce sont des variables aléatoires continues et positives dont on admettra généralement qu'elles sont indépendantes et identiquement distribuées. La connaissance de leur distribution commune permettra alors de déterminer les propriétés caractéristiques du processus de comptage correspondants.

1.6.2. Processus de Poisson

Définition

On dit qu'un Processus de comptage $\{N(t), t \geq 0\}$ est un processus de Poisson s'il satisfait aux trois conditions suivantes :

•C1. Le Processus $N(t)$ est homogène dans le temps. C'est à dire que la probabilité d'avoir k événements dans un intervalle de longueur donné t ne dépend que de t et non pas de la position de l'intervalle par rapport à l'axe du temps. En d'autres termes :

$$\Pr(N(t+s) - N(s) = k) = \Pr(N(t) = k) = p_k(t) \text{ pour tout } s > 0, t > 0 \text{ et } k = 0,1,2\dots$$

•C2. Le Processus $N(t)$ est à accroissements indépendants. Ce que signifie que pour tout système d'intervalles disjoints, les nombres d'événements s'y produisant sont des variables aléatoires indépendants. En particulier :

$$\Pr(N(t+s) - N(s) = k, N(s) = j) = \Pr(N(t+s) - N(s) = k) \cdot \Pr(N(s) = j) = p_k(t) \cdot p_k(s) \text{ pour tout } s > 0, t > 0.$$

•C3. la probabilité que deux événements ou plus se produisent dans un petit intervalle de longueur Δt est négligeable par rapport à la probabilité qu'il n'y ait qu'un seul événement. En termes plus clairs :

$$p_k(\Delta t) = \begin{cases} o(\Delta t) & k \geq 2 \\ \lambda \cdot \Delta t + o(\Delta t) & k = 1 \\ 1 - \lambda \cdot \Delta t + o(\Delta t) & k = 0 \end{cases}$$

λ est appelé intensité du processus Poissonien où le nombre moyen d'événements se produisant par une unité de temps.

Ce sont là les hypothèses classiques qui conduisent au processus de poisson.

Théorème.1.6.1.[1]

Si un processus de comptage $\{N(t), t \geq 0\}$ satisfait aux trois conditions précédentes, alors

$$\Pr(N(t) = j) = P_j(t) = e^{-\lambda t} \frac{(\lambda t)^j}{j!}$$

pour $t > 0$ et $k = 0,1,2\dots$ et par conséquent on a :

$$E[N(t)] = \text{Var}[N(t)] = \lambda t.$$

Preuve

D'après le théorème de probabilités totales et en tenant compte des conditions C2 et C3, on a pour $n = 1, 2, \dots$ l'équation suivante :

$$\begin{aligned} P_n(t+\Delta t) &= P_n(t)P_0(\Delta t) + P_{n-1}(t)P_1(\Delta t) + \sum_{i=2}^n P_{n-1}(t)P_i(\Delta t) \\ &= P_n(t)(1-\lambda \cdot \Delta t) + P_{n-1}(t)\lambda \cdot \Delta t + o(\Delta t). \end{aligned}$$

D'où ;

$$\frac{P_n(t+\Delta t) - P_n(t)}{\Delta t} = \lambda [P_{n-1}(t) - P_n(t)] + o(\Delta t).$$

Lorsque Δt tend vers zéro, on obtient

$$P_n'(t) = \lambda [P_{n-1}(t) - P_n(t)],$$

pour $n = 0$, on a

$$P_0'(t) = -\lambda P_0(t),$$

ce système d'équations est connu sous le nom d'équations de Kolmogorov ; il peut être résolu par récurrence. La condition initiale étant $\Pr(N(0) = 0) = P_0(0) = 1$ et $\Pr(N(0) = j) = P_j(0) = 0$ par suite on a $P_0(t) = e^{-\lambda t}$. Pour $n \geq 1$, posons $P_n(t) = q_n(t) e^{-\lambda t}$, ce qui implique

$$q_n'(t) = \lambda q_{n-1}(t), \text{ et par conséquent } q_n(t) = \frac{(\lambda t)^n}{n!} \text{ puisque } q_n(0) = 0 \text{ et } q_0(t) = 1.$$

On a donc bien $P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$.

1.6.3. La loi exponentielle

La loi du temps de service et du temps des inter arrivées la plus populaire est la loi exponentielle.

On va montrer par la suite que pour un processus d'arrivée poissonien, le processus de temps des inter arrivées associé suit une loi exponentielle :

Soit T la variable aléatoire qui définit l'intervalle entre deux arrivées successives (intèr arrivée), on écrit :

$$\Pr(T \geq t) = \Pr(\text{il y a pas d'arrivée dans le temps } t) = p_0(t) = e^{-\lambda t}.$$

On note par $A(t)$ la loi de probabilité de T ,

$$A(x) = P(T \leq t) = 1 - e^{-\lambda t} \text{ pour } t \geq 0$$

La densité correspondante est $a(t) = \lambda e^{-\lambda t}$. d'où T a une distribution exponentielle.

La loi exponentielle est très utilisée grâce à la propriété "sans mémoire", et qui pourrait ici s'énoncer ainsi: savoir que le service a déjà duré un certain temps n'apporte aucun renseignement sur sa fin prochaine. On remarquera en effet que, pour la loi exponentielle, la densité calculée $b(x | y)$ de (1.1), vaut:

$$b(x | y) = \frac{b(x + y)}{1 - B(y)} = \frac{v \cdot e^{-v \cdot (x + y)}}{e^{-v \cdot y}} = v \cdot e^{-v \cdot x} = b(x), \quad (1.2)$$

indépendamment de y . L'application de cette "absence de mémoire" montre que la probabilité d'une fin de service dans l'instant qui vient (dans l'intervalle $[t, t+dt]$) est $v dt$, quel que soit la durée du service. (v est appelé le taux de service)

La loi exponentielle se caractérise par ses moments:

- Moyenne de la variable : $1/v$.
- Variance de la variable : $1/v^2$.

1.7. Processus de Naissance et de Mort

1.7.1. La notion d'état

La notion d'état a une base intuitive: décrire l'état d'un système, c'est donner la liste des caractéristiques qu'il possède, et aussi donner les éléments qui permettent de prévoir son évolution. En mécanique, par exemple, on décrit l'état d'un point matériel par ses coordonnées x, y, z et par sa vitesse v_x, v_y, v_z . Ces quantités une fois connues, on peut décrire la position du point, et via les équations de la mécanique, la trajectoire qu'il doit suivre.

Il en est de même pour un système stochastique, à la différence que la prévision du futur y prend un caractère probabiliste: on ne saura prévoir l'état exact dans un certain temps, mais seulement sa distribution de probabilité.

1.7.2. Processus de Markov

Considérons une variable aléatoire X , à espace d'état discret (X ne prend ses valeurs que dans un espace discret. Pour fixer les idées, X peut représenter le nombre des clients dans un tampon: $X=0, 1, 2, \dots$). On peut noter les états $E_1, E_2, \dots, E_k \dots$. Notons t_1, t_2, \dots, t_n les instants successifs des changements d'état de X , et x_1, \dots, x_n la suite des états visités. Les t_n pourraient être, par exemple, les instants d'arrivée et de départ dans le tampon. Dans le cas général, la valeur de X à l'instant t_{n+1} dépend de tout le passé, c'est à dire de $X(t_1), X(t_2), \dots$. On dira que X obéit à la propriété de Markov si et seulement si:

$$\begin{aligned} P[X(t_{n+1}) = x_{n+1} \mid X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, \dots, X(t_0) = x_0] \\ = P[X(t_{n+1}) = x_{n+1} \mid X(t_n) = x_n], \end{aligned} \quad (1.3)$$

pour tout $t_0 = 0 < t_1 < t_2 < \dots < t_n$.

En bref, seul l'état courant (à t_n) influence l'évolution prochaine du processus. C'est encore la propriété sans mémoire.

Supposons de plus le processus homogène, c'est à dire invariable par translation dans le temps. Ceci autorise à introduire la notation:

$$P_{ij} = P[X(t_{n+1}) = E_j \mid X(t_n) = E_i].$$

La propriété de Markov se traduit par:

$$P[X(t_{n+1}) = E_k] = \sum_j p_{kj} P[X(t_n) = E_j]. \quad (1.4)$$

C'est là la forme élémentaire d'un résultat fondamental, connu sous le nom d'Equation de Chapman - Kolmogorov.

1.7.3. Les processus de naissance et de mort

L'étude des processus de Markov les plus généraux est un des grands chapitres du Calcul des Probabilités. Pour les besoins très limités de ce rappel, nous limitons au cas très particulier - et très utile - où le processus ne peut faire de "sauts" que vers ses voisins les plus proches.

Les seuls mouvements autorisés sont de E_k vers E_{k+1} et E_{k-1} . Le taux de passage de E_k à E_{k+1} est traditionnellement noté λ_k : c'est le taux de naissance. Le taux du passage de E_k vers E_{k-1} est le taux de mort ν_k . Le processus de naissance et de mort a été introduit dans l'étude de l'évolution des populations: l'état E_k fait référence à la taille. En accord avec cette origine, et pour simplifier l'écriture, on nommera les états plus simplement 1, 2, 3, ..., sans perte de généralité.

Nous cherchons à dériver la distribution de probabilité de l'état, c'est à dire la fonction

$$P_k(t) = P[X(t) = k]$$

Nous allons procéder en utilisant la dynamique des équation de Chapman - Kolmogorov. Plaçons-nous à un instant t , et examinons les possibilités. A l'instant $t+\Delta t$, l'état sera k si:

A t , l'état était k et rien ne s'est produit.

A t , l'état était $k-1$ et une naissance s'est produit durant l'intervalle $(t, t+\Delta t)$.

A t , l'état était $k+1$ et une mort s'est produit durant l'intervalle $(t, t+\Delta t)$.

L'intervalle Δt est petit. D'autres événements plus complexes (2 arrivées, etc.) auront lieu avec des probabilités en $o(\Delta t)$. On écrira donc:

$$P_k(t+\Delta t) = P_k(t)(1 - \lambda_k \Delta t - \nu_k \Delta t) + P_{k-1}(t) \lambda_{k-1} + P_{k+1}(t) \nu_{k+1} \Delta t + o(\Delta t) .$$

Cette équation transcrit fidèlement la description des possibilités donnée ci-dessus. A partir de cette forme, on écrit une équation différentielle par les traitements classiques:

$$\frac{P_k(t+\Delta t) - P_k(t)}{\Delta t} = -(\lambda_k + \nu_k) \cdot P_k(t) + P_{k-1}(t) \cdot \lambda_{k-1} + P_{k+1}(t) \cdot \nu_{k+1} + \frac{o(\Delta t)}{\Delta t} . \quad (1.5)$$

Cas particulier, en 0 la transition de "mort" n'existera pas. Finalement, le passage à la limite conduit au système différentiel fondamental:

$$\frac{dP_k(t)}{dt} = -(\lambda_k + \nu_k) \cdot P_k(t) + P_{k-1}(t) \cdot \lambda_{k-1} + P_{k+1}(t) \cdot \nu_{k+1} \quad (1.6)$$

$$\frac{dP_0(t)}{dt} = -\lambda_0 \cdot P_0(t) + P_1(t) \cdot \nu_1 \quad (1.7)$$

Nous nous limiterons au cas de systèmes à l'équilibre. Pour un système dynamique à l'équilibre, un régime permanent existe, tel que les probabilités d'état ne dépendent pas du temps. On réécrira simplement le système ci-dessus, en y annulant toutes les dérivées. La distribution stationnaire sera notée P_k .

$$\begin{aligned} (\lambda_k + \nu_k) P_k &= \lambda_{k-1} P_{k-1} + \nu_{k+1} P_{k+1} \\ \lambda_0 P_0 &= \nu_1 P_1 \quad , \end{aligned} \quad (1.8)$$

pour résoudre ce système ; on pose $\alpha_n = \nu_n P_n - \lambda_{n-1} P_{n-1}$.
A l'aide des équations (1.8), on peut écrire :

$$\alpha_n = v_{n+1} P_{n+1} - \lambda_n P_n = \alpha_{n+1}$$

puisque $\alpha_1 = v_1 P_1 - \lambda_0 P_0 = 0$, alors on a $\alpha_n = 0 \quad \forall n$

d'où le système d'équation :

$$v_n P_n = \lambda_{n-1} P_{n-1}$$

ce que permet de déduire la solution :

$$P_n = \frac{\lambda_{n-1} \dots \lambda_0}{v_n \dots v_1} P_0 \quad (1.9)$$

avec

$$P_0 = \frac{1}{\sum_{n=0}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda_i}{v_{i+1}}} \quad (1.10)$$

l'existence de la solution (la condition d'ergodicité) est liée à la convergence de la série figurant au dénominateur de l'expression (1.10).

Remarque : le système d'équations (1.8) admet une interprétation très fructueuse. On peut interpréter le terme de gauche $(\lambda_k + v_k) P_k$ comme un flux de probabilité qui quitterait l'état k . Un terme tel que $\lambda_k P_k$ mesure le flux de probabilité montant de k vers $k+1$. Ce flux est le produit de la probabilité de se trouver dans l'état k par le taux de départ montant (sachant qu'on s'y trouve, si l'on veut). La même interprétation vaut pour les autres termes. L'équation dit alors que, pour que les probabilités gardent des valeurs constantes, il doit y avoir égalité entre les flux entrant et sortant des différents états.

1.7.4. La solution stationnaire obtenue par l'écriture matricielle

Définition

Un processus de Markov est dit ergodique si $\forall i, j; \lim_{t \rightarrow \infty} P_{ij}(t) = (\pi_j)_{j \in E}$ est l'unique solution du système $\Pi Q = 0$ et $\sum_{i \in E} \pi_i = 1$. (1.11)

Q est la matrices des taux de transitions, elle est appelée le générateur infinitésimal, E est l'ensemble des états. Tout solution qui vérifie (1.11) est appelée solution stationnaire du processus de Markov.

La solution stationnaire d'un processus de naissance et de mort est obtenue en résolvant le système (1.11) où ;

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdot & \cdot & \cdot & 0 \\ v_1 & -(\lambda_1 + v_1) & \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & v_2 & -(\lambda_2 + v_2) & \lambda_2 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & v_i & -(\lambda_i + v_i) & \lambda_i & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

$$\Pi = (\pi_0, \pi_1, \dots, \pi_i, \dots)$$

Les équations s'écrivent sous forme explicite :

$$\begin{aligned} \lambda_0 \pi_0 &= v_1 \pi_1 \\ (\lambda_i + v_i) \pi_i &= \lambda_{i-1} \pi_{i-1} + v_{i+1} \pi_{i+1}, \end{aligned}$$

qui est exactement le système obtenu en (1.8) d'où la solution.

1.8. Etude de quelques modèles classiques markoviens

1.8.1. La file M/M/1

C'est le modèle à file d'attente le plus célèbre - à juste titre, certainement. Il permet en effet d'illustrer les concepts fondamentaux liés à l'attente devant un serveur.

Le système est décrit par le processus des arrivées, poissonnien de taux λ , la loi exponentielle de service (taux v), la file d'attente de capacité infinie. Les clients pourront être supposés servis dans l'ordre de leur arrivée. Lorsque $n > 0$ clients sont dans le système (un en service, $n-1$ en attente) à un instant t , la fin de service du client en cours se produit dans l'intervalle $(t, t+\Delta t)$ avec une probabilité $v \cdot \Delta t + o(\Delta t)$: propriété fondamentale de la loi de service exponentielle. De même, une arrivée se produira dans l'intervalle, avec probabilité $\lambda \Delta t + o(\Delta t)$. On applique à ce système le formalisme et les résultats obtenus plus haut sur le processus de naissance et de mort. L'état est décrit par N , qui représente cette fois le nombre total des clients dans le système. Il évolue selon un processus de Naissance et de Mort très simple .

L'analyse du système, selon la procédure des équations de Chapman - Kolmogorov, aboutit aux équations suivantes:

$$\begin{aligned} \lambda P_0 &= P_1 v \\ (\lambda + v) P_k &= P_{k-1} \lambda + P_{k+1} v . \end{aligned}$$

Des équations (1.9) et (1.10) on trouve:

$$P_k = \rho^k P_0 \quad (1.12)$$

Avec

$$P_0 = \frac{1}{\sum_{i=0}^{\infty} \rho^i} = 1 - \rho \quad (1.13)$$

Où, $\rho = \frac{\lambda}{\nu}$ est appelé « l'intensité de trafic(ou taux de charge) de la file.

L'existence de la solution est liée à la convergence de la série figurant au dénominateur de l'expression (1.13). On a donc :

$\lambda < \nu$ ou $\rho < 1$: la file est stable.

$\lambda \geq \nu$ ou $\rho \geq 1$: la file est instable.

Les paramètres de performance de la file M/M/1 sont :

$$\overline{N}_s = \sum_{n=0}^{\infty} n P_n = \frac{\rho}{1-\rho} \quad (1.14)$$

$$\overline{W}_s = \frac{1}{\lambda} \cdot \frac{\rho}{1-\rho} \quad (1.15)$$

Ce dernier résultat est obtenu par la formule de Little ;

$$\overline{W}_q = \overline{W}_s - \frac{1}{\nu} = \frac{\rho^2}{\lambda(1-\rho)} = \frac{\rho}{\nu-\lambda} \quad (1.16)$$

$$\overline{N}_q = \lambda \overline{W}_q = \frac{\rho^2}{\lambda(1-\rho)} = \frac{\lambda\rho}{\nu-\lambda} \quad (1.17)$$

La file d'attente M/M/1 est largement utilisée dans la modélisation de système pour plusieurs raisons :

.Elle est extrêmement simple à traiter avec un ensemble de propriétés facilement exprimable par des formules aisées à manipuler. Elle modélise d'une manière générale tout système de type guichet.

.Les hypothèses qu'elle utilise (processus markoviens en entrée et en sortie) sont classiquement utilisées dans la modélisation de divers systèmes, en l'absence de caractérisation plus précise et justifiée des paramètres.

1.8.2.La file M/M/c

Mêmes hypothèses et notations que ci-dessus, avec un nombre c de serveurs identiques en service.

Imaginons le système dans l'état $k < c$. Cela signifie que k clients sont présents, qui occupent k serveurs (les autres sont inactifs). Si $k > c$, il y a au moins un client en attente dans la file. Le nombre de clients dans le système au temps t noté N(t) est un processus de naissance et de mort . les taux de transition sont donnés par :

$$\lambda_k = \lambda \text{ pour tout } k ,$$

et

$$v_k = \begin{cases} k.v & \text{si } k \leq c, \\ c.v & \text{si } k > c, \end{cases}$$

en utilisant les équations (1.9) et (1.10) ; on trouve la solution stationnaire :

$$P_k = \begin{cases} \frac{\lambda^k}{k!.v^k} P_0 & \text{Si } k \leq c, \\ \frac{\lambda^k}{c^{k-c} c! v^k} P_0 & \text{Si } k > c, \end{cases} \quad (1.18)$$

et

$$P_0 = \left[\sum_{k=0}^{c-1} \frac{A^k}{k!} + \frac{A^c}{c!} \frac{1}{1-\rho} \right]^{-1} \quad (1.19)$$

$$A = \frac{\lambda}{v} \quad \text{et} \quad \rho = \frac{\lambda}{cv},$$

ce qui est réalisé si et seulement si $\rho = \frac{\lambda}{cv} < 1$.

La probabilité d'attendre est

$$P_W = \frac{A^c}{c!} \frac{1}{1-\rho} P_0. \quad (1.20)$$

Le nombre moyen de clients dans la file est donné par :

$$\begin{aligned} \bar{N}_q &= \sum_{k=0}^{\infty} (k-c) P_k \\ &= \left[\frac{\frac{A^{c+1}}{c}}{c! \left(1 - \frac{A}{c}\right)^2} \right] P_0. \end{aligned} \quad (1.21)$$

Le temps moyen d'attente se déduit de ce qui précède, à partir de la formule de Little:

$$\overline{W}_q = \left[\frac{A^c v}{(c-1)!(cv - \lambda)^2} \right] P_0 \quad (1.22)$$

1.8.3. Modèles à capacité limitée (Queues with truncation)

Dans la pratique et pour tenir compte de la capacité limitée des systèmes (mémoire limitée de machines, guichets de service,..), on est souvent conduit à prendre l'hypothèse de file d'attente à capacité limitée du type M / M / 1 / N.

Dans ce cas, un client qui arrive alors que le système soit plein est rejeté. D'une manière équivalente, on peut dire que le taux d'arrivée dépend de l'état de la file ; il vaut λ si k est strictement inférieur à N , et 0 si non.

Les équations différentielles du modèle M / M / 1 / ∞ restent valides pour ce modèle pour $k < N$.

Pour $k=N$; la probabilité qu'une arrivée entre au système vaut 0; d'où l'équation différentielle,

$$P_N(t + \Delta t) = P_N(t)(1 - v \cdot \Delta t) + P_{N-1}(t)(\lambda \cdot \Delta t).$$

$$\text{A l'état d'équilibre, on a : } P_N = \frac{\lambda}{v} P_{N-1}.$$

$$\text{Et par conséquent, } P_k = \rho^k P_0 \quad \text{si } k \leq N \quad (1.23)$$

$$P_0 = \frac{1}{\sum_{k=0}^N \rho^k} = \frac{1 - \rho}{1 - \rho^{N+1}} \quad (1.24)$$

La probabilité de rejet :

$$P_N = \frac{\rho^N (1 - \rho)}{1 - \rho^{N+1}} \quad (1.25)$$

Quand $N \rightarrow \infty$ (capacité illimitée), on aboutit au résultat du paragraphe (1.20).

La théorie générale des processus de naissance et de mort applicable à ce système permet d'accéder aux paramètres de performance :

$$\overline{N}_s = \frac{\rho \left[- (N+1) \rho^N + N \rho^{N+1} \right]}{(1-\rho)(1-\rho^{N+1})} \quad (1.26)$$

$$\overline{W}_s = \frac{[1 - (N + 1)\rho^N + N\rho^{N+1}]}{\lambda(1 - \rho)(1 - \rho^N)} \quad (1.27)$$

Même analyse est faite pour le modèle M/M/c/N où,

$$\lambda_k = \begin{cases} \lambda & \text{si } k \leq N, \\ 0 & \text{si } k > N, \end{cases}$$

on trouve :

$$P_k = \begin{cases} \frac{1}{k!} \left(\frac{\lambda}{v}\right)^k P_0 & \text{si } 0 \leq k < c, \\ \frac{\lambda^k}{c^{k-c} c! v^k} P_0 & \text{si } c \leq k \leq N, \end{cases} \quad (1.28)$$

Avec :

$$P_0 = \left[\sum_{k=0}^{c-1} \frac{1}{k!} \left(\frac{\lambda}{v}\right)^k + \sum_{k=c}^N \frac{1}{c^{k-c} c!} \left(\frac{\lambda}{v}\right)^k \right]^{-1} \quad (1.29)$$

$$= \left[\sum_{k=0}^{c-1} \frac{1}{k!} \left(\frac{\lambda}{v}\right)^k + \frac{\left(\frac{\lambda}{v}\right)^c}{c!} \frac{1 - \left(\frac{\lambda}{cv}\right)^{N-c+1}}{1 - \frac{\lambda}{cv}} \right]^{-1}$$

1.8.4. Le modèle M/M/c/c (modèle d'Erlang)

A leur arrivée, les clients sont servis immédiatement tant qu'un serveur au moins est libre; si les serveurs sont tous occupés, le client qui arrive est rejeté, et est supposé disparaître définitivement. On parle de modèle "à perte".

Le processus N_t est un processus de naissance et de mort avec :

$$\begin{aligned} v(i) &= i v. & \text{Si } i \leq c \\ \lambda(i) &= \lambda & \text{Si } i < c \text{ et } \lambda(i) = 0 \text{ si } i \geq c. \end{aligned}$$

D'où, la matrice de générateur infinitésimal associée s'écrit sous la forme suivante :

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \dots & 0 \\ v & -(\lambda+v) & \lambda & 0 & \dots & \dots & 0 \\ 0 & 2v & -(\lambda+2v) & \lambda & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & cv & -cv \end{pmatrix}$$

C'est un cas particulier du modèle M/M/c/N avec N = c. La solution stationnaire est obtenue par les équations (1.23) et (1.24) en posant N = c :

$$P_k = \begin{cases} P_0 \frac{\lambda^k}{k! v^k} & \text{si } k \leq c, \\ 0 & \text{si non.} \end{cases} \quad (1.30)$$

Et

$$P_0 = \left[\sum_{i=0}^c \left(\frac{\lambda}{v} \right)^i \frac{1}{i!} \right]^{-1} \quad (1.31)$$

Ce qui entraîne :

$$P_k = \frac{\frac{A^k}{k!}}{\sum_{i=0}^k \frac{A^i}{i!}} \quad (1.32)$$

Et par suite,

$$P_c = \frac{\frac{A^c}{c!}}{\sum_{i=0}^c \frac{A^i}{i!}} \quad (1.33)$$

comme probabilité de trouver le système plein; compte tenu de la nature des arrivées, c'est aussi la probabilité de rejet.

Ce système particulier a beaucoup été utilisé, surtout en téléphonie, et la solution stationnaire est encore appelée formule d'Erlang. Elle a été démontrée par Erlang en 1917.

DEUXIEME CHAPITRE

SYSTEME DE FILE D'ATTENTE M/M/C AVEC RAPPEL

Deuxième Chapitre

Système de file d'attente M/M/c avec rappel

2.0.Introduction

Dans ce chapitre , on présente une description générale de système avec rappel en résumant les principaux résultats sur les files markoviennes avec discipline de rappel constant et classique M/M/c avec $c \leq 2$. on met le point sur l'existence et le calcul de la distribution stationnaire et quelques mesures de performance associées . On présente aussi quelques résultats analytiques de système markovien avec $c > 2$ à l'aide de la fonction génératrice.

2.1. Files d'attente avec rappel

Terminologie et notation

Par analogie à la notation établie pour la file d'attente classique, celle du modèle de file d'attente avec rappel s'écrit par :

$$A / B / S / Z / O / H$$

Avec :

O : capacité de l'orbite.

H : fonction de persévérance.

Où H permet de définir le comportement du client devant une situation de blocage (serveurs occupés). Elle peut être décrite par une séquence $\{H_0, H_1, \dots, H_K, \dots\}$ de probabilités de rappel après un échec.

1) Quand $H_K = 1$ pour $K \geq 0$, Le système devient un système sans perte. ainsi, chaque client reçoit éventuellement le service si O est infini. Dans ce cas, $H = NL$ (No Loss).

2) quand $H_K = \alpha$ ($\alpha \leq 1$) pour $K \geq 0$, Le système devient un système à perte géométrique et $H = GL$ (Geometric Loss).

Remarques :

1) Si la capacité de l'orbite est infinie, elle est omise dans la notation. H est aussi omise dans le cas d'un système sans perte.

2) La distribution des temps de rappel est supposée exponentielle de taux μ , $1/\mu$ étant la durée moyenne des intervalles de rappel. C'est la raison pour laquelle elle est omise dans la notation.

3) Lorsque $\mu \rightarrow \infty$, le système d'attente avec rappel se ramène à un système d'attente classique, le temps de rappel étant nul. Ainsi, tout appel primaire bloqué (qui trouve tous les serveurs occupés) restera dans le système jusqu'à ce qu'un serveur soit libéré.

3) Lorsque $\mu \rightarrow 0$, le système d'attente avec rappel est un système d'Erlang avec perte. Cette hypothèse signifie que l'intervalle de temps entre deux rappels successifs du même client est infini. Ainsi, tout appel primaire bloqué est automatiquement perdu.

4) Lorsque la capacité de l'orbite est nulle ($O = 0$), il n'y a aucune position d'attente où de rappel. Ainsi tout appel bloqué est perdu pour toujours. Ce modèle correspond aussi au système d'Erlang avec perte.

2.2. Système de file d'attente M/M/1 avec rappel constant

2.2.1. Description du modèle

On considère un système de file d'attente à un seul serveur dans laquelle les clients primaires arrivent selon un flot poissonien de taux $\lambda > 0$. Les durées de service d'un client sont indépendantes et exponentiellement distribuées de moyenne $1/v$; un client qui trouve le serveur libre est immédiatement servi. Les clients qui trouvent le serveur occupé vont directement en orbite et rappellent ultérieurement pour service. Les intervalles de temps séparant les rappels successifs sont supposés indépendants et exponentiellement distribués avec un taux $\mu_j = \mu(1 - \delta_{0j})$, qui est indépendant du nombre de clients dans l'orbite. Les temps des inter arrivées et les temps de service sont supposés mutuellement indépendants. L'état du système à l'instant t peut être décrit au moyen d'un processus bidimensionnel $X(t) = (C(t), N(t))$, où $N(t)$ est le nombre de clients dans l'orbite et $C(t)$ est le nombre de serveurs occupés à l'instant t .

On note que le processus X est un processus de Markov qui prend ses valeurs dans l'ensemble des états $S = \{0, 1\} \times \mathbb{N}$.

L'espace des états et les transitions sont donnés par la figure ci-dessous :

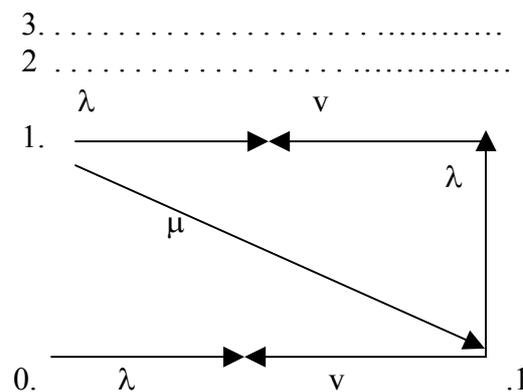


Fig.2.1

2.2.2. Calcul de la distribution stationnaire

On définit la distribution limite $P_{ij} = \lim_{t \rightarrow \infty} \Pr \{ (C(t), N(t)) = (i, j) \}; (i, j) \in S$ si elle existe.

On considère maintenant les équations de Kolmogorov pour les probabilités p_{ij} données à partir des transitions suivantes :

Au temps $(t + \Delta t)$ l'état $(1, j) \in S$ est atteint de l'un des états suivants :

- $(0, j)$: arrivée de client primaire avec une probabilité égale à $\lambda \Delta t + o(\Delta t)$.
- $(0, j+1)$: rappel d'un client de l'orbite pour service avec une probabilité égale à $\mu(1 - \delta_{0j}) \Delta t + o(\Delta t)$.
- $(1, j-1)$: arrivée de client primaire vers l'orbite avec une probabilité égale à $\lambda \Delta t + o(\Delta t)$.
- $(1, j)$: ni arrivée, ni service complété, avec une probabilité égale à $1 - (\lambda + v) \Delta t + o(\Delta t)$.

Au temps $(t + \Delta t)$ l'état $(0, j) \in S$ est atteint de l'un des états suivants :

•(1,j) : complétion de service une probabilité égale à $v\Delta t + o(\Delta t)$.

•(0, j) : ni arrivée, ni rappel avec une probabilité égale à $1-(\lambda + \mu (1-\delta_{0j})) \Delta t + o(\Delta t)$.

D'où les équations des états sont données par :

$$P_{0j}(t+\Delta t) = v P_{1j}(t) \Delta t + (1- (\lambda + \mu (1- \delta_{0j})) \Delta t) P_{0j}(t) + o(\Delta t)$$

$$P_{1j}(t+\Delta t) = \lambda P_{0j}(t) \Delta t + \mu (1-\delta_{0j}) P_{0j+1}(t) \Delta t + \lambda P_{1j-1}(t) \Delta t + (1-(\lambda+v)) P_{1j}(t) \Delta t + o(\Delta t)$$

On divise les deux équations par Δt et on prends la limite lorsque $\Delta t \rightarrow 0$ pour obtenir :

$$dP_{0j}(t) / dt = - (\lambda + \mu (1-\delta_{0j})) P_{0j} + v P_{1j} ;$$

$$dP_{1j}(t) / dt = (\lambda + v) P_{1j} - \lambda P_{0j} - \mu (1-\delta_{0j}) P_{0j+1} - \lambda P_{1j-1} ; i = 0, 1, j \geq 0$$

A l'état stationnaire, on a que P_{ij} existe et est indépendante de t (c'est à dire le comportement de processus est stable) et donc $dP_{ij}(t)/dt = 0$

D'où les équations suivantes :

$$(\lambda + \mu (1-\delta_{0j})) P_{0j} = v P_{1j} ; \quad (2.1)$$

$$(\lambda + v) P_{1j} = \lambda P_{0j} + \mu(1-\delta_{0j}) P_{0j+1} + \lambda P_{1j-1} ; \quad (2.2)$$

Une combinaison des deux équations (2.1) et (2.2) donne :

$$\mu(1-\delta_{0j}) v P_{0j+1} - \lambda (\lambda + \mu(1-\delta_{0j})) P_{0j} = \mu (1-\delta_{0j}) v P_{0j} - \lambda (\lambda + \mu(1-\delta_{0j})) P_{0j-1}, \forall j \geq 1$$

D'où le système d'équations suivant :

$$(\lambda + \mu (1-\delta_{0j})) P_{0j} = v P_{1j} ;$$

$$(\mu(1-\delta_{0j}) v + \lambda (\lambda + \mu (1-\delta_{0j}))) P_{0j} = \lambda (\lambda + \mu(1-\delta_{0j})) P_{0j-1} + \mu (1-\delta_{0j}) v P_{0j+1}, j \geq 1$$

Qui est du même type que celui de (1.8), ce qui permet de déduire l'équation :

$$\mu(1-\delta_{0j}) v P_{0j} - \lambda (\lambda + \mu (1-\delta_{0j})) P_{0j-1} = 0 ,$$

et par suite :

$$P_{0j} = \frac{\lambda (\lambda + \mu (1 - \delta_{0,j}))}{\mu (1 - \delta_{0,j}) \cdot v} P_{0j-1} .$$

Nous obtenons par récurrence :

$$P_{0j} = \prod_{i=1}^j \frac{\lambda (\lambda + \mu (1 - \delta_{0,j}))}{\mu (1 - \delta_{0,j}) \cdot v} P_{00} .$$

D'où,

$$P_{0j} = \left(\frac{\lambda (\lambda + \mu)}{\mu v} \right)^j P_{00}$$

A partir de l'équation (2.1) on obtient :

$$P_{1j} = \frac{\lambda + \mu(1 - \delta_{0j})}{v} \left(\frac{\lambda (\lambda + \mu)}{\mu v} \right)^j P_{00}$$

Par normalisation on a :

$$\sum_{j=0}^{\infty} P_{0j} + \sum_{j=0}^{\infty} P_{1j} = 1$$

Par conséquent :

$$P_{00} \left[\sum_{j=0}^{\infty} \left(\frac{\lambda (\lambda + \mu)}{\mu v} \right)^j + \frac{\lambda + \mu}{v} \sum_{j=0}^{\infty} \left(\frac{\lambda (\lambda + \mu)}{\mu v} \right)^j \right] = 1$$

La condition d'ergodicité du processus devient :

$$\sum_{j=0}^{\infty} \left(\frac{\lambda (\lambda + \mu)}{\mu v} \right)^j < \infty$$

c'est à dire :

$$\frac{\lambda (\lambda + \mu)}{\mu v} < 1 \quad (2.3)$$

Enfin, les probabilités stationnaires pour ce processus markovien sont données par :

$$P_{00} = \left[\frac{\mu v}{\mu v - \lambda (\lambda + \mu)} + \frac{\lambda + \mu}{v} \frac{\mu v}{\mu v - \lambda (\lambda + \mu)} \right]^{-1}$$

$$P_{00} = \frac{\mu v - \lambda (\lambda + \mu)}{\mu (\lambda + \mu + v)} \quad (2.4)$$

et

$$P_{0j} = \left(\frac{\lambda (\lambda + \mu)}{\mu v} \right)^j P_{00} \quad (2.5)$$

$$P_{1j} = \frac{\lambda + \mu}{v} \left(\frac{\lambda (\lambda + \mu)}{\mu v} \right)^j P_{00} \quad (2.6)$$

On peut obtenir maintenant plusieurs caractéristiques de performance :

2.2.3. Nombre moyen de clients dans le système

Pour calculer le nombre moyen de clients dans le système en régime stationnaire, $E(N)$, on peut utiliser la relation :

$$E(N) = \lim_{z \rightarrow 1} P'(z) = P'(1) . \quad (2.7)$$

Où $P_0(z)$, $P_1(z)$ sont les fonctions génératrices correspondant aux P_{0j} et P_{1j} respectivement :

$$P_0(z) = \sum_{j=0}^{\infty} z^j P_{0j} \quad |z| \leq 1 . \quad (2.8)$$

$$P_1(z) = \sum_{j=0}^{\infty} z^j P_{1j} \quad |z| \leq 1 . \quad (2.9)$$

Soit q_j la distribution stationnaire du nombre de clients dans l'orbite, $q_j = \Pr(N = j)$ a pour fonction génératrice,

$$P(z) = P_0(z) + P_1(z) = \sum_{j=0}^{\infty} z^j (P_{0j} + P_{1j})$$

En appliquant cette transformée

$$\begin{aligned} P(z) &= \sum_{j=0}^{\infty} z^j P_{00} \left[\left(\frac{\lambda(\lambda + \mu)}{\mu\nu} \right)^j + \frac{\lambda + \mu}{\nu} \left(\frac{\lambda(\lambda + \mu)}{\mu\nu} \right)^j \right] \\ &= P_{00} (\nu + \lambda + \mu) \frac{\mu}{\mu\nu - z(\lambda(\lambda + \mu))} , \end{aligned}$$

après calcul on obtient :

$$P(z) = \frac{\mu \cdot \nu - \lambda(\lambda + \mu)}{\mu \cdot \nu - z(\lambda(\lambda + \mu))} . \quad (2.10)$$

D'où,

$$E(N) = \frac{\lambda(\lambda + \mu)}{\mu \cdot \nu - \lambda(\lambda + \mu)} \quad (2.11)$$

2.2.4. Probabilité de blocage

C'est la probabilité que le serveur soit occupé, elle est donnée par : $P_1 = P_1(1)$, où,

$$P_1(z) = \sum_{j=0}^{\infty} z^j P_{1j} = \frac{\lambda + \mu}{\nu} \cdot \frac{\mu \cdot \nu - \lambda(\lambda + \mu)}{\mu(\lambda + \mu + \nu)} \sum_{j=0}^{\infty} \left(\frac{\lambda(\lambda + \mu)}{\mu \cdot \nu} z \right)^j$$

$$= \frac{\lambda + \mu}{(v + \mu + v)} \cdot \frac{\mu \cdot v - \lambda (\lambda + \mu)}{(\mu \cdot v - z (\lambda (\lambda + \mu)))}$$

D'où,

$$P_1 = \frac{\lambda + \mu}{\lambda + \mu + v} \quad (2.12)$$

2.3.M/M/1 avec taux de rappel linéaire

Avec la même discipline, on considère la file M / M / 1 où le taux de rappel dépend de nombre de clients dans l'orbite ; c'est à dire $\mu_j = j\mu$.

Théorème[27]

Pour la file M/M/1 avec rappel linéaire, la distribution stationnaire de l'état du système ; $P_{ij} = \lim_{t \rightarrow \infty} \Pr \{C(t) = i, N(t) = j\}$ est donnée par :

$$P_{0j} = \left(\frac{\rho}{\mu} \right)^j \frac{1}{j!} \prod_{i=0}^{j-1} (\lambda + i\mu) (1 - \rho)^{\frac{\lambda}{\mu} + 1} \quad (2.13)$$

$$P_{1j} = \left(\frac{\rho}{\mu} \right)^j \frac{\rho}{j!} \prod_{i=1}^j (\lambda + i\mu) (1 - \rho)^{\frac{\lambda}{\mu} + 1} \quad (2.14)$$

avec $\rho = \frac{\lambda}{v} < 1$.

Les fonctions génératrices correspondantes sont données par :

$$P_0(z) = \sum_{j=0}^{\infty} z^j P_{0j} = (1 - \rho) \cdot \left(\frac{1 - \rho}{1 - \rho \cdot z} \right)^{\frac{\lambda}{\mu}} \quad (2.15)$$

$$P_1(z) = \sum_{j=0}^{\infty} z^j P_{1j} = \rho \cdot \left(\frac{1 - \rho}{1 - \rho \cdot z} \right)^{\frac{\lambda}{\mu} + 1} \quad (2.16)$$

Preuve:

Dans le cas où le temps de service a une loi exponentielle, le processus $(C(t), N(t))$ est un processus de Markov d'espace d'états $S = \{0, 1\} \times \mathbb{N}$

Au temps $(t + \Delta t)$ l'état $(1, j) \in S$ est atteint de l'un des états suivants :

- $(0, j)$: arrivée de client primaire avec une probabilité égale à $\lambda \Delta t + \theta(\Delta t)$

- (0, j+1) : rappel d'un client de l'orbite pour service avec une probabilité égale à $(j+1)\mu \Delta t + \theta(\Delta t)$
- (1, j-1) : arrivée de client vers l'orbite avec une probabilité égale à $\lambda \Delta t + \theta(\Delta t)$

- (1, j) : ni arrivée, ni service complété, avec une probabilité égale à $1 - (\lambda + \nu) \Delta t + \theta(\Delta t)$

Au temps $(t+\Delta t)$ l'état $(0, j) \in S$ est atteint de l'un des états suivants :

- (1, j) : complétion de service, avec une probabilité égale à $\nu \Delta t + \theta(\Delta t)$
- (0, j) : ni arrivée, ni rappel, avec une probabilité égale à $1 - (\lambda + j\mu) \Delta t + \theta(\Delta t)$

D'où les équations des états sont données par :

$$P_{0j}(t+\Delta t) = \nu P_{1j}(t) \Delta t + (1 - (\lambda + j\mu) \Delta t) P_{0j}(t) + \theta(\Delta t) \quad (2.17)$$

$$P_{1j}(t+\Delta t) = \lambda P_{0j}(t) \Delta t + (j+1)\mu P_{0j+1}(t) \Delta t + \lambda P_{1j-1}(t) \Delta t + (1 - (\lambda + \nu) \Delta t) P_{1j}(t) + \theta(\Delta t) \quad (2.18)$$

On divise les deux équations par Δt et on prends la limite lorsque $\Delta t \rightarrow 0$ pour obtenir :

$$dP_{0j}(t)/dt = (\lambda + j\mu) P_{0j} - \nu P_{1j} ;$$

$$dP_{1j}(t)/dt = (\lambda + \nu) P_{1j} - \lambda P_{0j} - (j+1)\mu P_{0j+1} - \lambda P_{1j-1} ;$$

A l'état stationnaire, on a que P_{ij} existe et est indépendante de t et d'où, $dP_{ij}(t)/dt = 0$.

D'où les équations suivantes :

$$(\lambda + j\mu) P_{0j} = \nu P_{1j} ; \quad (2.19)$$

$$(\lambda + \nu) P_{1j} = \lambda P_{0j} + (j+1)\mu P_{0j+1} + \lambda P_{1j-1} ; \quad (2.20)$$

En combinant les équations (2.19) et (2.20) on a l'équation :

$$(j+1)\mu \nu P_{0j+1} - \lambda (\lambda + j\mu) P_{0j} = j\mu \nu P_{0j} - \lambda (\lambda + (j-1)\mu) P_{0j-1} . \quad (2.21)$$

D'après l'équation (1.8), on obtient l'équation, $j\mu \nu P_{0j} - \lambda (\lambda + (j-1)\mu) P_{0j-1} = 0$.

Et par conséquent:

$$P_{0j} = \frac{\lambda (\lambda + (j-1)\mu)}{j\mu \nu} P_{0j-1} = \prod_{i=0}^{j-1} \frac{\lambda (\lambda + i\mu)}{(i+1)\mu \nu} P_{00}$$

d'où

$$P_{0j} = \frac{\rho^j}{\mu^j j!} \prod_{i=0}^{j-1} (\lambda + i\mu) P_{00} .$$

A partir de l'équation (2.19) on conclut :

$$P_{1j} = \frac{\rho^j}{\mu^j j!} \prod_{i=0}^j (\lambda + i\mu) P_{00}$$

$$P_{1j} = \frac{\rho^{j+1}}{\mu^j j!} \prod_{i=1}^j (\lambda + i\mu) P_{00} .$$

A cet ensemble d'équations on ajoute la condition de normalisation ;

$$\sum_{j=0}^{\infty} P_{0j} + \sum_{j=0}^{\infty} P_{1j} = 1 , \text{ pour obtenir :}$$

$$P_{00} = \left[\sum_{j=0}^{\infty} \frac{\rho^j}{j! \mu^j} \prod_{i=0}^{j-1} (\lambda + i\mu) + \sum_{j=0}^{\infty} \frac{\rho^{j+1}}{j! \mu^j} \prod_{i=1}^j (\lambda + i\mu) \right]^{-1}$$

mais,

$$\sum_{j=0}^{\infty} \frac{\rho^j}{j! \mu^j} \prod_{i=0}^{j-1} (\lambda + i\mu) = \sum_{j=0}^{\infty} \frac{\rho^j}{j!} \prod_{i=0}^{j-1} \frac{\lambda + i\mu}{\mu} = \sum_{j=0}^{\infty} \frac{(-\rho)^j}{j!} \prod_{i=0}^{j-1} \left(-\frac{\lambda}{\mu} - i\right) \quad (2.22)$$

Rappelons la formule du Binôme classique :

$$(1 + x)^m = \sum_{j=0}^{\infty} \frac{x^j}{j!} \prod_{i=0}^{j-1} (m - i) \quad (2.23)$$

En appliquant la formule (2.23) sur la somme (2.22), on peut déterminer l'expression ;

$$\sum_{j=0}^{\infty} \frac{\rho^j}{j! \mu^j} \prod_{i=0}^{j-1} (\lambda + i\mu) = (1 - \rho)^{-\frac{\lambda}{\mu}},$$

et

$$\sum_{j=0}^{\infty} \frac{\rho^{j+1}}{j! \mu^j} \prod_{i=1}^j (\lambda + i\mu) = \rho (1 - \rho)^{-\frac{\lambda}{\mu} - 1}$$

ce qui mène à ;

$$P_{00} = \left[(1 - \rho)^{-\frac{\lambda}{\mu}} + \rho (1 - \rho)^{-\frac{\lambda}{\mu} - 1} \right]^{-1} = (1 - \rho)^{\frac{\lambda}{\mu} + 1} .$$

Finalement, après un calcul simple, on obtient le résultat du théorème ;

$$P_{0j} = \frac{\rho^j}{j! \mu^j} \prod_{i=0}^{j-1} (\lambda + i\mu) (1 - \rho)^{\frac{\lambda}{\mu} + 1} ,$$

et

$$P_{1j} = \frac{\rho^{j+1}}{j! \mu^j} \prod_{i=1}^j (\lambda + i\mu) (1 - \rho)^{\frac{\lambda}{\mu} + 1} .$$

$P_0(z)$ et $P_1(z)$ sont obtenues par la formule du Binôme :

$$\begin{aligned}
P_0(z) &= \sum_{j=0}^{\infty} z^j \frac{\rho^j}{j! \mu^j} \prod_{i=0}^{j-1} (\lambda + i\mu) (1-\rho)^{\frac{\lambda}{\mu}+1} \\
&= \sum_{j=0}^{\infty} \frac{(-z\rho)^j}{j!} \prod_{i=0}^{j-1} \left(-\frac{\lambda}{\mu} - i\right) (1-\rho)^{\frac{\lambda}{\mu}+1} \\
&= (1-\rho)^{\frac{\lambda}{\mu}+1} (1-z\rho)^{-\frac{\lambda}{\mu}} \\
&= (1-\rho) \left(\frac{1-\rho}{1-\rho z} \right)^{\frac{\lambda}{\mu}}, \text{ qui est exactement le résultat du théorème.}
\end{aligned}$$

Et de même pour $P_1(z)$:

$$P_1(z) = \sum_{j=0}^{\infty} z^j P_{1j} = \rho \left(\frac{1-\rho}{1-\rho z} \right)^{\frac{\lambda}{\mu}+1}$$

2.3.1. Nombre moyen de clients dans l'orbite

La distribution stationnaire du nombre de clients dans l'orbite $q_j = \Pr\{N = j\}$ a pour fonction génératrice $P(z)$ suivante :

$$P(z) = P_0(z) + P_1(z) = (1 + \rho - \rho z) \left(\frac{1-\rho}{1-\rho z} \right)^{\frac{\lambda}{\mu}+1} . \quad (2.24)$$

A partir de l'équation (2.7) , on obtient:

$$E(N) = \frac{\rho(\lambda + \rho\mu)}{(1-\rho)\mu} . \quad (2.25)$$

2.3.2. Probabilité de blocage

La probabilité de blocage B est la probabilité que le serveur soit occupé ;

$$B = \sum_{j=0}^{\infty} P_{1j} = P_1(1) = \rho = \frac{\lambda}{\nu} \quad (2.26)$$

2.4. Le système M/M/2

Pour le cas $c=2$ et rappels constant; Artalejo[2] donne les formules suivantes :

$$P_{0j} = \frac{\lambda^2}{(\lambda + \mu)^2 + v\mu} \left(\frac{\lambda((\lambda + \mu)^2 + v\mu)}{v\mu(3\lambda + 2(v + \mu))} \right)^j P_{00} \quad j \geq 1 \quad (2.27)$$

$$P_{1j} = \frac{\lambda + \mu(1 - \delta_{0j})}{v} P_{0j} \quad j \geq 0 \quad (2.28)$$

$$P_{2j} = \frac{\lambda^2(\lambda + v + \mu)}{v^2(3\lambda + 2(v + \mu))} \left(\frac{\lambda((\lambda + \mu)^2 + v\mu)}{v\mu(3\lambda + 2(v + \mu))} \right)^j P_{00} \quad j \geq 0 \quad (2.29)$$

$$P_{00} = \frac{2v\mu(\lambda + v + \mu) - \lambda(\lambda + \mu)^2}{\mu(2v(\lambda + v) + (\lambda + \mu)(\lambda + 2v))} \quad (2.30)$$

La condition nécessaire et suffisante pour la stabilité est donnée par $\rho_2 < 1$

$$\text{avec } \rho_2 = \frac{\lambda(\lambda + \mu)^2}{2v\mu(\lambda + v + \mu)} \quad (2.31)$$

preuve

De la même manière que dans le cas $c = 1$, on obtient les équations des probabilités $\{P_{ij} / (i \in \{0,1,2\}, j \geq 0)\}$ suivantes :

$$(\lambda + \mu(1 - \delta_{0j})) P_{0j} = v P_{1j} \quad j \geq 0 \quad ; \quad (2.32)$$

$$(\lambda + v + \mu(1 - \delta_{0j})) P_{1j} = \lambda P_{0j} + 2v P_{2j} + \mu P_{0j+1} \quad j \geq 0 \quad ; \quad (2.33)$$

$$(\lambda + 2v) P_{2j} = \lambda P_{1j} + \mu P_{1j+1} + \lambda(1 - \delta_{0j}) P_{2j-1} \quad j \geq 0 \quad ; \quad (2.34)$$

a partir de l'équation (2.32), on a:

$$P_{1j} = \frac{\lambda + \mu(1 - \delta_{0j})}{v} P_{0j} \quad j \geq 0 \quad ; \quad (2.35)$$

des équations (2.33) et (2.35), on peut exprimer les probabilités $\{P_{2j}\}$ en fonction de $\{P_{0j}\}$;

$$P_{2j} = \frac{(\lambda + \mu(1 - \delta_{0j}))^2 + v\mu(1 - \delta_{0j})}{2v^2} P_{0j} - \frac{\mu}{2v} P_{0j+1} \quad j \geq 0 \quad ; \quad (2.36)$$

En combinant (2.36), (2.35) et (2.34), on obtient;

$$(\alpha_j + \beta_j) P_{0j} = \alpha_{j-1} (1 - \delta_{0j}) P_{0j-1} + \beta_{j+1} P_{0j+1} \quad j \geq 0 \quad ; \quad (2.37)$$

où,

$$\alpha_j = \frac{\lambda((\lambda + \mu(1 - \delta_{0j}))^2 + \nu\mu(1 - \delta_{0j}))}{2\nu^2} \quad j \geq 0 \quad ;$$

$$\beta_j = \frac{\mu}{\nu} \left(\frac{3}{2} \lambda + \nu + \mu \right) (1 - \delta_{0j}) \quad j \geq 0 \quad ;$$

On note bien que les probabilités $\{P_{0j}, j \geq 0\}$ satisfont un système d'équation de naissance et de mort donné par (1.8) avec un taux de naissance égal à α_j et de taux de mort β_j .

Ce qui permet d'écrire :
$$P_{0j} = P_{00} \prod_{k=0}^{j-1} \frac{\alpha_k}{\beta_{k+1}}$$

Et d'après le paragraphe 1.7.3 on constate que l'existence de la solution (la condition d'ergodicité) est liée à la convergence de la série S tel que,

$$S = \sum_{j=0}^{\infty} \prod_{k=0}^{j-1} \frac{\alpha_k}{\beta_{k+1}}, \text{ on définit :}$$

$$\alpha = \frac{\lambda((\lambda + \mu)^2 + \nu\mu)}{2\nu^2}, \quad \beta = \frac{\mu}{\nu} \left(\frac{3}{2} \lambda + \nu + \mu \right), \text{ et } \gamma = \frac{\lambda^3}{2\nu^2}$$

d'où

$$S = \frac{\gamma}{\beta} \sum_{j=0}^{\infty} \left(\frac{\alpha}{\beta} \right)^j.$$

Et par conséquent, S converge si et seulement si $\alpha < \beta$, après un calcul simple on déduit

l'équation (2.31) du théorème ;
$$\frac{\lambda(\lambda + \mu)^2}{2\nu\mu(\lambda + \nu + \mu)} < 1.$$

Et

$$P_{0j} = P_{00} \prod_{k=0}^{j-1} \frac{\alpha_k}{\beta_{k+1}} = \frac{\gamma}{\alpha} \left(\frac{\alpha}{\beta} \right)^j P_{00}.$$

Des d'équations (2.36) et la condition de normalisation, on déduit également les équations (2.29) et (2.30).

2.5. Système M/M/c

Mêmes hypothèses et notations que ci-dessus, avec un nombre de c serveurs identiques en service.

L'état du système à l'instant t peut être décrit au moyen d'un processus bidimensionnel $X(t) = (C(t), N(t))$, où N(t) est le nombre de clients dans l'orbite et C(t) est le nombre de serveurs occupés à l'instant t.

On note que le processus X est un processus de Markov qui prend ses valeurs dans l'ensemble des états $S = \{0, 1, \dots, c\} \times \mathbb{N}$.

Au temps (t+Δt) l'état (i, j) ∈ S est atteint de l'un des états suivants :

- (i-1, j) : arrivée de client primaire avec une probabilité égale à $\lambda \Delta t + o(\Delta t)$

- $(i-1, j+1)$: rappel d'un client de l'orbite pour service avec une probabilité égale à $(j+1)\mu \Delta t + o(\Delta t)$
- $(i+1, j)$: complétion de service avec une probabilité égale à $(i+1)v \Delta t + o(\Delta t)$
- (i, j) : ni arrivée, ni service complété, ni rappel avec une probabilité égale à $1 - (\lambda + iv + j\mu) \Delta t + o(\Delta t)$

On définit les taux de transition d'un état $(i, j) \in S$ à un état $(n, m) \in S$ par $q_{(i,j)(n,m)}$;

- pour $0 \leq i \leq c-1$

$$q_{(i,j)(n,m)} = \begin{cases} \lambda & \text{si } (n, m) = (i+1, j) \\ i v & \text{si } (n, m) = (i-1, j) \\ j\mu & \text{si } (n, m) = (i+1, j-1) \\ -(\lambda + iv + j\mu) & \text{si } (n, m) = (i, j) \\ 0 & \text{si non} \end{cases}$$

- pour $i=c$

$$q_{(i,j)(n,m)} = \begin{cases} cv & \text{si } (n, m) = (c-1, j) \\ \lambda & \text{si } (n, m) = (c, j+1) \\ -(\lambda + cv) & \text{si } (n, m) = (c, j) \\ 0 & \text{si non} \end{cases}$$

2.5.1. Formules explicites des caractéristiques de performance

On considérera que le processus X est ergodique ce qui a lieu si et seulement si la condition suivante est vérifiée [27]: $\rho < 1$ où $\rho = \lambda / cv$

On définit aussi les probabilités limites; $P_{ij} = \lim_{t \rightarrow \infty} \Pr(C(t)=i, N(t)=j)$, pour $(i, j) \in S$, qui sont positives si et seulement si la condition $\rho < 1$ est satisfaite.

D'après les probabilités de transition, on peut écrire l'équation suivante :

$$P_{ij}(t+\Delta t) = \lambda \Delta t P_{i-1,j}(t) + (j+1)\mu \Delta t P_{i-1,j+1}(t) + (i+1)v\Delta t P_{i+1,j}(t) + (1 - (\lambda + iv + j\mu) \Delta t) P_{ij}(t) + o(\Delta t)$$

On divise l'équation par Δt et on prend la limite lorsque $\Delta t \rightarrow 0$ pour obtenir :

$$dP_{ij}(t)/dt = \lambda P_{i-1,j}(t) + (j+1)\mu P_{i-1,j+1}(t) + (i+1)v P_{i+1,j}(t) - (\lambda + iv + j\mu) P_{ij}(t); \begin{matrix} 0 \leq i \leq c-1 \\ j \geq 0 \end{matrix}$$

A l'état stationnaire, on a que P_{ij} existe et est indépendante de t d'où $dP_{ij}/dt = 0$.

Ce qui nous permet d'écrire :

$$(\lambda + i v + j \mu) P_{ij} = \lambda P_{i-1j} + (j+1)\mu P_{i-1j+1} + (i+1)v P_{i+1j} \quad \text{si } 0 \leq i \leq c-1 \quad (2.38)$$

$$(\lambda + c v) P_{cj} = \lambda P_{c-1j} + (j+1)\mu P_{c-1j+1} + \lambda P_{cj-1} \quad \text{si } i = c \quad (2.39)$$

On pose les fonctions génératrices partielles $P_i(z) = \sum_{j=0}^{\infty} z^j P_{ij}$; pour $0 \leq i \leq c$

En multipliant les termes des équations (2.38) et (2.39) par z^j et en sommant sur les $j \geq 0$, on obtient ;

$$(\lambda + i v) \sum_{j=0}^{\infty} z^j P_{ij} + \mu \sum_{j=0}^{\infty} j z^j P_{ij} = \lambda \sum_{j=0}^{\infty} z^j P_{i-1j} + \mu \sum_{j=0}^{\infty} (j+1) z^j P_{i-1j+1} + (i+1)v \sum_{j=0}^{\infty} z^j P_{i+1j} \quad 0 \leq i \leq c-1$$

$$(\lambda + c v) \sum_{j=0}^{\infty} z^j P_{cj} = \mu \sum_{j=0}^{\infty} (j+1) z^j P_{c-1j+1} + \lambda \sum_{j=0}^{\infty} z^j P_{c-1j} + \lambda \sum_{j=0}^{\infty} z^j P_{cj-1} \quad i=c$$

D'où les équations ;

$$(\lambda + i v)P_i(z) + \mu z P'_i(z) = \lambda P_{i-1}(z) + \mu P'_{i-1}(z) + (i+1)v P_{i+1}(z) ; 0 \leq i \leq c-1 \quad (2.40)$$

$$(\lambda + c v) P_c(z) = \lambda P_{c-1}(z) + \mu P'_{c-1}(z) + \lambda z P_c(z); \quad i=c \quad (2.41)$$

En introduisant la fonction génératrice bivariée $P(x, z)$ tel que,

$$P(x, z) = \sum_{i=0}^c \sum_{j=0}^{\infty} x^i z^j P_{ij} = \sum_{i=0}^c x^i P_i(z)$$

En multipliant les termes de l'équation (2.40) et (2.41) par x^i et en sommant les deux équations résultantes terme à terme sur les i $0 \leq i \leq c$, on obtient l'équation suivante :

$$\lambda (1-x) P(x, z) + \mu (z-x) P'_z(x, z) + v (x-1) P'_x(x, z) + \lambda x^c (x-z) P_c(z) + \mu x^c (x-z) P'_c(z) = 0 \quad (2.42)$$

En dérivant l'équation (2.42) par rapport à z, x, xx, xz, zz respectivement aux point $x=z=1$, on obtient :

$$\mu P'_{z}(1,1) - \lambda P_c(1) - \mu P'_c(1) = 0 \quad (2.43)$$

$$\lambda + \mu P'_{z}(1,1) - v P'_{x}(1,1) - \lambda P_c(1) - \mu P'_c(1) = 0 \quad (2.44)$$

$$\lambda P'_{x}(1,1) + \mu P''_{xz} - v P''_{xx} - \lambda c P_c(1) - \mu c P'_c(1) = 0 \quad (2.45)$$

$$-\lambda p'_{z,z}(1,1) - \mu p''_{zz} + (v + \mu) p''_{xz} + \lambda p'_{c,c}(1) - \mu c p'_{c,c}(1) - \lambda c p_c(1) + \mu p''_{zz} = 0 \quad (2.46)$$

$$\mu p''_{z,z} - \lambda p'_{c,c}(1) - \mu p''_{zz} = 0 \quad (2.47)$$

On utilise le fait que :

$$p'_{z,z}(x, z) = \sum_i \sum_j j x^i z^{j-1} P_{ij}$$

$$\begin{aligned} \bullet p'_{z,z}(1,1) &= \sum_i \sum_j j P_{ij} \\ &= \lim_{t \rightarrow \infty} E(N(t)) = N \end{aligned}$$

$$\bullet p_c(1) = \sum_{j=0}^{\infty} P_{c,j} = P\{C(t) = c\} = B$$

$$\text{et que } p'_{x,x}(x, z) = \sum_i \sum_j i x^{i-1} z^j P_{ij}$$

On a aussi

$$\begin{aligned} \bullet p'_{x,x}(1,1) &= \sum_i \sum_j i P_{ij} \\ &= \lim_{t \rightarrow \infty} E(C(t)) = Y \end{aligned}$$

$$\bullet p'_{c,c}(1) = \lim_{t \rightarrow \infty} E(N(t), C(t)=c) = N_c$$

La résolution de 2.43 et 2.44 donne :

$$\lambda = v p'_{x,x}(1,1)$$

D'où

$$Y = \lambda / v \quad (2.48)$$

Après un calcul long, La résolution 2.44, 2.45, 2.46 et 2.47 donne :

$$N = \frac{v + \mu}{\mu} \frac{\lambda - v \cdot \text{var}(c(t))}{c - \lambda} \quad (2.49)$$

2.5.2. Remarque sur la nature numérique de la solution stationnaire

Remarquons que pour $c \leq 2$, la distribution stationnaire $\{P_{ij}, (i, j) \in S\}$ est telle que les suites partielles $\{P_{ij}, i \geq 0\}$ satisfont un système d'équations de type « naissance et de mort », mais quand $c > 2$, cette structure de naissance et de mort n'est pas préservée ; d'où vient la difficulté analytique de tels systèmes.

2.6. Conclusion

La seule manière d'obtenir la distribution stationnaire $P_{i,j}$ consiste en la résolution numérique des équations de Kolmogorov (2.38) et (2.39).

Comme ce système d'équations est très compliqué, il ne peut pas être résolu directement, et on a besoin donc de méthodes d'approximation numérique de la solution de notre système. Cependant une implémentation pratique de la file M/M/c avec rappel exige une variété d'approximations et de méthodes de troncature afin de calculer les principales caractéristiques de performance.

CHAPITRE TROISIEME

*APPROXIMATION DU SYSTEME M /M/ c
AVEC RAPPELS*

3.0.Introduction

Un examen détaillé de la littérature montre que des efforts ont été fait pour l'investigation de systèmes markoviens à plusieurs serveurs avec rappel exponentiel [5, 17, 42, 43,...]. Cependant, il convient de noter qu'aucune solution analytique satisfaisante des équations de Kolmogorov pour la distribution stationnaire du nombre de clients occupés et la longueur de file d'attente n'a été obtenue , à l'exception des cas $c = 1$ et $c = 2$.

L'utilisation d'un modèle tronqué pour la solution numérique des équations de Kolmogorov pour le modèle original avec un nombre illimité de sources a été suggéré par Wilkinson [46]. Il a noté que des probabilités stationnaires pour le modèle tronqué peuvent être calculées par les équations de Kolmogorov, mais n'a pas donné un algorithme détaillé.

Ce chapitre est consacré à l'étude de ce système en introduisant une solution algorithmique basée sur l'approximation numérique du système dynamique. Dans la première partie de ce chapitre, on présente une analyse du système basée sur la méthode de troncature directe.

Le reste de cette première partie est organisée de la manière suivante : Dans la section .1., on décrit le modèle mathématique. Dans la section .2., on présente l'analyse algorithmique basée sur un algorithme numérique pour calculer la distribution stationnaire du modèle tronqué et différentes mesures de performance.

Dans la section .3.on présente des formules explicites de différentes caractéristiques de performance du modèle tronqué. Dans la section .4.on présente la relation existant entre le système tronqué et le système initial. Dans la section .5. On étudie la convergence de la solution stationnaire du modèle tronqué vers celle du modèle original. En fin un choix de M est discuté.

La deuxième partie est consacrée à l'approximation du système $M/ M/c$ avec des modèles tronqués généralisés . Dans ce système d'attente on constate l'absence de résultats explicites ; l'analyse est réduite à une solution par la méthode matricielle géométrique.

Dans la section .7. de cette partie des modèles mathématiques sont décrits. Dans la section .8. on étudie la méthode matricielle géométrique, dans la section .9.on étudie l'analyse de quelque modèles tronqués généralisés .

La troisième partie est consacrée à la validation des approximations et leurs comparaisons.

3.1. Modèle tronqué

3.1.1. Description du modèle

Dans ce modèle, la capacité de l'orbite est finie, elle est limitée par une constante M ; si le nombre de clients dans l'orbite est égal à M , dans ce cas, certains clients qui arrivent vers le système n'ont pas la possibilité d'y entrer .

Ce système peut être décrit par un processus bivarié $(C^{(M)}(t), N^{(M)}(t))$ où $C^{(M)}(t)$ est le nombre de serveurs occupés et $N^{(M)}(t)$ est le nombre de clients dans l'orbite à l'instant t . Sous ces hypothèses, le processus $(C^{(M)}(t), N^{(M)}(t))$ est markovien avec espace des états fini , $S^{(M)} = \{0, 1, \dots, c\} \times \{0, 1, \dots, M\}$.

Les taux de transition $q_{(ij)(nm)}^{(M)}$ sont donnés par :

1. pour $0 \leq i \leq c-1, 0 \leq j \leq M$

$$q_{(ij)(nm)}^{(M)} = \begin{cases} \lambda & \text{si } (n, m) = (i+1, j) \\ i v & \text{si } (n, m) = (i-1, j) \\ j \mu & \text{si } (n, m) = (i+1, j-1) \\ -(\lambda + i v + j \mu) & \text{si } (n, m) = (i, j) \\ 0 & \text{si non} \end{cases}$$

• pour $i=c, 0 \leq j \leq M-1$

$$q_{(cj)(nm)}^{(M)} = \begin{cases} c v & \text{si } (n, m) = (c-1, j) \\ \lambda & \text{si } (n, m) = (c, j+1) \\ -(\lambda + c v) & \text{si } (n, m) = (c, j) \\ 0 & \text{si non} \end{cases}$$

3. pour $i=c, j=M$

$$q_{(cM)(nm)}^{(M)} = \begin{cases} c v & \text{si } (n, m) = (c-1, M) \\ -c v & \text{si } (n, m) = (c, M) \\ 0 & \text{si non,} \end{cases}$$

D'où $q_{(ij)(nm)}^{(M)}$ sont les mêmes que ceux du modèle original sauf pour l'état ($i = c, j = M$).
Et par suite on obtient les équations d'équilibre ;

• Pour $0 \leq i \leq c-1, 0 \leq j \leq M-1$

$$P_{ij}^{(M)}(t + \Delta t) = \lambda \Delta t P_{i-1j}^{(M)}(t) + (j+1) \mu \Delta t P_{i-1j+1}^{(M)}(t) + (i+1) \nu \Delta t P_{i+1j}^{(M)}(t) + 1 - (\lambda + i \nu + j \mu) \Delta t P_{ij}^{(M)}(t) + \theta(\Delta t) \quad (3.1)$$

On divise l'équation par Δt et on prend la limite lorsque $\Delta t \rightarrow 0$ pour obtenir :

$$\frac{dP_{ij}^{(M)}(t)}{dt} = \lambda P_{i-1j}^{(M)}(t) + (j+1)\mu P_{i-1j+1}^{(M)}(t) + (i+1)\nu P_{i+1j}^{(M)}(t) - (\lambda + i \nu + j \mu) P_{ij}^{(M)}(t) \quad ; 0 \leq i \leq c-1, j \geq 0 \quad (3.2)$$

A l'état stationnaire, on a que P_{ij} existe et est indépendante de t et d'où $P_{ij}^{(M)}(t)/dt = 0$
D'où les équations suivantes :

$$(\lambda + i \nu + j \mu) P_{ij}^{(M)} = \lambda P_{i-1j}^{(M)} + (j+1)\mu P_{i-1j+1}^{(M)} + (i+1) \nu P_{i+1j}^{(M)}, \quad 0 \leq i \leq c-1 \quad (3.3)$$

$$(\lambda + c \nu) P_{cj}^{(M)} = \lambda P_{c-1j}^{(M)} + (j+1) \mu P_{c-1j+1}^{(M)} + \lambda P_{cj-1}^{(M)} \quad (3.4)$$

3.1.2. Calcul de la distribution stationnaire

Comme l'espace des états $S^{(M)}$ est fini et la chaîne est irréductible ; du moment que tout état (i, j) peut être atteint par n'importe quel autre état ; Le processus $X^{(M)}$ est donc ergodique d'où l'existence d'une distribution stationnaire $P_{ij}^{(M)}$;

$P_{ij}^{(M)} = \lim_{t \rightarrow \infty} \Pr\{C^{(M)}(t) = i, N^{(M)}(t) = j\}$ est une solution des équations linéaires suivantes :

$$(\lambda + i \nu + j \mu) P_{ij}^{(M)} = \lambda P_{i-1j}^{(M)} + (j+1)\mu P_{i-1j+1}^{(M)} + (i+1) \nu P_{i+1j}^{(M)} \quad \text{si } 0 \leq i \leq c-1, 0 \leq j \leq M-1, \quad (3.5)$$

$$(\lambda + i \nu + M \mu) P_{iM}^{(M)} = \lambda P_{i-1M}^{(M)} + (i+1) \nu P_{i+1M}^{(M)} \quad \text{si } 0 \leq i \leq c-1, \quad (3.6)$$

$$(\lambda + c \nu) P_{cj}^{(M)} = \lambda P_{c-1j}^{(M)} + (j+1)\mu P_{c-1j+1}^{(M)} + \lambda P_{cj-1}^{(M)} \quad \text{si } 0 \leq j \leq M-1 \quad (3.7)$$

$$c \nu P_{cM}^{(M)} = \lambda P_{c-1M}^{(M)} + \lambda P_{cM-1}^{(M)} \quad (3.8)$$

et $P_{ij}^{(M)}$ satisfont la condition de normalisation ;

$$\sum_{i=0}^c \sum_{j=0}^{\infty} P_{ij}^{(M)} = 1 \quad (3.9)$$

Ces probabilités vont être calculées à l'aide d'un algorithme récursif suivant :

•**Algorithme 1**

On introduit des variables $R_{ij}^{(M)}$, $0 \leq i \leq c$, $0 \leq j \leq M$ tel que,

$$R_{ij}^{(M)} = \frac{P_{ij}^{(M)}}{P_{0M}^{(M)}} .$$

D'après la condition de normalisation, on a :

$$\sum_{i=0}^c \sum_{j=0}^{\infty} P_{ij}^{(M)} = \sum_{i=0}^c \sum_{j=0}^{\infty} R_{ij}^{(M)} P_{0M}^{(M)} = 1 .$$

D'où ;

$$P_{0M}^{(M)} = \frac{1}{\sum_{i=0}^c \sum_{j=0}^{\infty} R_{ij}^{(M)}} .$$

ET PAR CONSEQUENT , LES VARIABLES $R_{ij}^{(M)}$ SATISFONT L'ENSEMBLE DES EQUATIONS SUIVANTES :

$$R_{0M}^{(M)} = 1$$

$$(\lambda + i v + j \mu) R_{ij}^{(M)} = \lambda R_{i-1j}^{(M)} + (j+1)\mu R_{i-1j+1}^{(M)} + (i+1)v R_{i+1j}^{(M)} \quad \text{si } i \leq c-1 \quad (3.10)$$

$$(\lambda + i v + M \mu) R_{iM}^{(M)} = \lambda R_{i-1M}^{(M)} + (i+1)v R_{i+1M}^{(M)} \quad \text{si } 0 \leq i \leq c-1 \quad (3.11)$$

$$(\lambda + c v) R_{cj}^{(M)} = \lambda R_{c-1j}^{(M)} + (j+1)\mu R_{c-1j+1}^{(M)} + \lambda v R_{c,j-1}^{(M)} \quad \text{si } 0 \leq j \leq M-1 \quad (3.12)$$

$$c v R_{cM}^{(M)} = \lambda R_{c-1M}^{(M)} + \lambda R_{cM-1}^{(M)} \quad \text{si } 0 \leq j \leq M-1 \quad (3.13)$$

1. On pose $j = M$ pour calculer les variables $R_{0M}^{(M)}$, $R_{cM}^{(M)}$

1.1 On réécrit l'équation (3.11) par :

$$R_{i+1M}^{(M)} = \frac{(\lambda + i v + M \mu) R_{iM}^{(M)} - \lambda R_{i-1M}^{(M)}}{(i+1)v} \quad 0 \leq i \leq c-1$$

Où bien ;

$$R_{iM}^{(M)} = \frac{(\lambda + (i-1)v + M\mu)R_{i-1M}^{(M)} - \lambda R_{i-2M}^{(M)}}{iv} ; \quad 1 \leq i \leq c$$

A partir de la relation $R_{0M}^{(M)} = 1$, on peut calculer les variables $R_{1M}^{(M)}$ $R_{cM}^{(M)}$ d'une manière récursive.

2. On pose $j = M-1$, pour pouvoir calculer les variables : $R_{0M-1}^{(M)}$ $R_{cM-1}^{(M)}$.

2.1. La variable $R_{cM-1}^{(M)}$ peut être calculée directement à partir de la relation (3.13) ;

$$R_{cM-1}^{(M)} = \frac{cvR_{cM}^{(M)} - \lambda R_{c-1M}}{\lambda} .$$

2.2. Pour trouver les variables $R_{0j}^{(M)}$ $R_{c-1j}^{(M)}$, on utilise l'équation (3.10) pour $i=0, \dots, c-1$; ces équations s'écrivent sous la forme suivante :

$$\alpha_i X_{i-1} + b_i X_i + \gamma_i X_{i+1} = \delta_i, \quad 0 \leq i \leq c-1 \quad (3.14)$$

Où,

$$X_i = R_{ij}^{(M)}, \quad \alpha_i = -\lambda, \quad b_i = \lambda + i v + j \mu, \quad \gamma_i = -(i+1)v, \quad \delta_i = (j+1)\mu R_{i-1j+1}^{(M)}$$

Avec $X_{-1} = 0, X_c = R_{cj}^{(M)}$.

L'équation (3.14) est une équation aux différences du second ordre dont la solution est fournie par l'algorithme de Gauss qui est destiné à la résolution numérique d'équations linéaires algébriques et considéré comme une méthode efficace .

L'algorithme consiste à transformer le système tri-diagonal (3.14) à un système triangulaire par la procédure suivante :

On considère la matrice augmentée équivalente au système d'équations 3.14 ;

$$\left(\begin{array}{cccccccc|c} b_0 & \gamma_0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \delta_0 \\ \alpha_1 & b_1 & \gamma_1 & 0 & 0 & \dots & 0 & 0 & 0 & \delta_1 \\ 0 & \alpha_2 & b_2 & \gamma_2 & 0 & \dots & 0 & 0 & 0 & \delta_2 \\ \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \alpha_{c-2} & b_{c-2} & \gamma_{c-2} & 0 & \delta_{c-2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha_{c-1} & b_{c-1} & \gamma_{c-1} & \delta_{c-1} \end{array} \right)$$

On commence par éliminer les éléments sous la diagonale, et on obtient :

$$\left(\begin{array}{cccccccc|c} B_0 & \gamma_0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & D_0 \\ 0 & B_1 & \gamma_1 & 0 & 0 & \dots & 0 & 0 & 0 & D_1 \\ 0 & 0 & B_2 & \gamma_2 & 0 & \dots & 0 & 0 & 0 & D_2 \\ \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & B_{c-2} & \gamma_{c-2} & 0 & D_{c-2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_{c-1} & \gamma_{c-1} & D_{c-1} \end{array} \right)$$

Cette dernière est équivalente au système d'équations :

$$B_i X_i + \gamma_i X_{i+1} = D_i \quad 0 \leq i \leq c-1 \quad (3.15)$$

avec: $\beta_0 = b_0$, $D_0 = \delta_0$;

$$\text{d'où l'équation : } B_{i-1} X_{i-1} + \gamma_{i-1} X_i = D_{i-1} \quad (3.16)$$

On multiplie l'équation (3.16) par $\frac{\alpha_i}{B_{i-1}}$, et on obtient ;

$$\alpha_i X_{i-1} + \frac{\alpha_i}{B_{i-1}} \gamma_{i-1} X_i = \frac{\alpha_i}{B_{i-1}} D_{i-1} \quad (3.17)$$

Si on retranche l'équation (3.17) de l'équation (3.14), on a :

$$\left(b_i - \frac{\alpha_i \gamma_{i-1}}{B_{i-1}} \right) X_i + \gamma_i X_{i+1} = \delta_i - \frac{\alpha_i D_{i-1}}{B_{i-1}}, \quad (3.18)$$

qui est exactement l'équation (3.15) avec:

$$B_i = b_i - \frac{\alpha_i \gamma_{i-1}}{B_{i-1}} ; \quad D_i = \delta_i - \frac{\alpha_i D_{i-1}}{B_{i-1}} \quad 1 \leq i \leq c-1 \quad (3.19)$$

•• On peut donc calculer les variables B_i , D_i , pour $1 \leq i \leq c-1$ par la formule récursive (3.19) en considérant les valeurs initiales $B_0 = b_0$ et $D_0 = \delta$, on aura :

$$B_0 = \lambda + j\mu ; \quad B_i = \lambda + i\nu + j\mu - \frac{\lambda i\nu}{B_{i-1}}, \quad 0 \leq i \leq c-1 \quad (3.20)$$

$$D_0 = 0 ; \quad D_i = (j+1)\mu R_{i-1}^{(M)} + \frac{\lambda D_{i-1}}{B_{i-1}}$$

•• On calcule maintenant, les variables X_{c-1}, \dots, X_0 dans l'ordre inverse à partir de l'équation (3.15) dont on peut tirer la formule suivante :

$$X_i = \frac{D_{ij} - \gamma_i X_{i+1}}{B_i} \quad (3.21)$$

Qui donne :

$$R_{ij}^{(M)} = \frac{D_{ij} + (i+1)vR_{i+1j}^{(M)}}{B_{ij}}, \quad i=c-1, c-2, \dots, 1, 0 \quad (3.22)$$

Afin d'éviter une soustraction dans l'équation (3.20), on introduit les variables ;
 $\beta_i = B_i - (\lambda + j\mu)$ et donc on aura :

$$\beta_{0j} = 0 \quad ; \quad \beta_i = iv \frac{j\mu + \beta_{i-1j}}{\lambda + j\mu + \beta_{i-1j}}, \quad 1 \leq i \leq c-1 \quad (3.23)$$

$D_{0j} = 0$;

$$D_{ij} = (j+1)\mu R_{i-1j+1} + \frac{\lambda D_{i-1j}}{\lambda + j\mu + \beta_{i-1j}}, \quad 1 \leq i \leq c-1 \quad (3.24)$$

Par suite, les variables $R_{ij}^{(M)}$ sont calculées par ;

$$R_{ij}^{(M)} = \frac{D_{ij} + (i+1)vR_{i+1j}^{(M)}}{\lambda + j\mu + \beta_{ij}} \quad . \quad i = c-1, c-2, \dots, 1, 0 \quad (3.25)$$

3.Répéter l'étape **2** pour $j \geq 0$, c'est à dire pour $j = M-2, M-3, \dots, 0$; on trouve toutes les variables $R_{ij}^{(M)}$.

4.On peut maintenant calculer les probabilités $P_{ij}^{(M)}$; $P_{ij}^{(M)} = R_{ij}^{(M)} \cdot P_{0M}^{(M)}$.
D'où ;

$$P_{ij}^{(M)} = \frac{R_{ij}^{(M)}}{\sum_{i=0}^C \sum_{j=0}^M R_{ij}^{(M)}} \quad . \quad (3.26)$$

3.1.3 Caractéristiques de performance du système tronqué

APRES AVOIR CALCULE LES PROBABILITES $P_{ij}^{(M)}$, LES DIFFERENTES MESURES DE PERFORMANCE PEUVENT ETRE EXPLICITEMENT CALCULEES :

a) La probabilité de blocage $B^{(M)}$;

$$B^{(M)} = \Pr \{ C^{(M)} = c \} = \sum_{j=0}^M P_{cj}^{(M)} = \sum_{j=0}^M R_{cj}^{(M)} P_{0M}^{(M)} \quad . \quad (3.27)$$

b)Le nombre moyen de serveurs occupés $Y^{(M)}$;

$$\begin{aligned}
Y^{(M)} = E(C^{(M)}) &= \sum_{i=0}^C i \Pr\{C^{(M)} = i\} \\
&= \sum_{i=0}^C \sum_{j=0}^M i \Pr\{C^{(M)} = i, N^{(M)} = j\} ;
\end{aligned}$$

$$Y^{(M)} = \sum_{i=0}^C \sum_{j=0}^M i R_{ij}^{(M)} P_{0M}^{(M)} . \quad (3.28)$$

C) Le nombre moyen de clients dans l'orbite $N^{(M)}$;

$$\begin{aligned}
N^{(M)} = E(N^{(M)}) &= \sum_{j=0}^M j \Pr\{N^{(M)} = j\} \\
&= \sum_{i=0}^C \sum_{j=0}^M j \Pr\{C^{(M)} = i, N^{(M)} = j\} ;
\end{aligned}$$

$$N^{(M)} = \sum_{i=0}^C \sum_{j=0}^M j R_{ij}^{(M)} P_{0M}^{(M)} . \quad (3.29)$$

D) La variance du nombre de serveurs occupés $V^{(M)}$;

$$V^{(M)} = E(C^{(M)})^2 - (E(C^{(M)}))^2 = \sum_{i=0}^C \sum_{j=0}^M i^2 R_{ij}^{(M)} P_{0M}^{(M)} - (Y^{(M)})^2 . \quad (3.30)$$

3.1.4. Relation existant entre le système tronqué et le système initial

Pour étudier l'existence de cette relation, on a besoin de quelques notions et résultats concernant la théorie des processus stochastique.

a) Comparaison stochastique de variables aléatoires :

Soit $(\Omega, \mathfrak{F}, \wp)$ un espace de probabilité et (E, \mathcal{M}) un espace mesurable, où E est partiellement ordonné par la relation $<$. Soient $\xi^{(1)} : \Omega \rightarrow E$ et $\xi^{(2)} : \Omega \rightarrow E$ des variables aléatoires sur $(\Omega, \mathfrak{F}, \wp)$.

ON DIT QUE $\xi^{(1)}$ EST STOCHASTIQUEMENT INFÉRIEURE À $\xi^{(2)}$ ET ON NOTE

$$\xi^{(1)} \leq_{st} \xi^{(2)} \text{ SI}$$

$$\wp\{\xi^{(1)} \in A\} \leq \wp\{\xi^{(2)} \in A\} \quad \forall A.$$

Avec A est un sous ensemble monotone d'espace d'états E . (on dit que A est monotone si $x \in A$ et $x < y$ alors $y \in A$).

Théorème3. 1. [27]

SI E EST UN ESPACE METRIQUE SEPARABLE ET COMPLET, M EST UNE σ -ALGEBRE ALORS $\xi^{(1)} \leq_{st} \xi^{(2)}$ SI ET SEULEMENT SI IL EXISTE UN ESPACE DE PROBABILITE $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ ET DES VARIABLES ALEATOIRES ;

$\eta^{(1)} : \Omega^* \rightarrow E$ ET $\eta^{(2)} : \Omega^* \rightarrow E$ TEL QUE :

• $\xi^{(1)}$ A LA MEME DISTRIBUTION QUE $\eta^{(1)}$ ET $\xi^{(2)}$ A LA MEME DISTRIBUTION QUE $\eta^{(2)}$

• $\eta^{(1)}(\omega) < \eta^{(2)}(\omega) \forall \omega \in \Omega^*$.

DEFINITION 3.1.

ON DIT QU'UN PROCESSUS DE MARKOV ξ_t DEFINIT SUR $A \subset \mathbb{Z}_+^N$ EST UN PROCESSUS DE MIGRATION SI ET SEULEMENT SI LES TAUX DE TRANSITIONS SUIVANTES, A PARTIR DE L'ETAT $N = (N_1, N_2, \dots, N_N)$ SONT POSSIBLES VERS LES ETATS :

- 1. $n + e_i$ de taux $\lambda_{n,i}^1$, $1 \leq i \leq N$**
- 2. $n - e_i$ de taux $\mu_{n,i}^1$, $1 \leq i \leq N$**
- 3. $n - e_i + e_j$ de taux $a_{n,i,j}^{ij}$, $j \leq N$, $i \geq 1$.**

Où $e_i = (0, \dots, 1, 0, \dots, 0)$ vecteur de dimension N dont la i-ième coordonnée égale à 1 et les autres sont nulles.

On remarque que les processus de migration de dimension un sont les processus de naissance et de mort classiques. Et par suite les processus de migration de dimension deux peuvent être utilisés pour décrire notre système avec rappel.

Par exemple, le processus $(C(t), N(t))$ est un processus de migration défini par les caractéristiques suivants :

$$\lambda_{(n,m)}^1 = \begin{cases} \lambda & \text{si } 0 \leq n \leq c-1 \\ 0 & \text{si } n = c \end{cases}$$

$$\lambda_{(n,m)}^2 = \begin{cases} 0 & \text{si } 0 \leq n \leq c-1 \\ \lambda & \text{si } n = c \end{cases}$$

$$\mu_{(n,m)}^1 = n\nu ;$$

$$\mu_{(n,m)}^2 = 0 ;$$

$$a_{(n,m)}^{1,2} = 0 ;$$

$$\left\{ \begin{array}{l} m\mu \text{ si } 0 \leq n \leq c-1 \end{array} \right.$$

$$a^{2,1}_{(n,m)} = 0 \quad \text{si } n=c.$$

Le processus $(C^{(M)}(t), N^{(M)}(t))$ du modèle tronqué est un processus de migration défini par les caractéristiques suivants :

$$\lambda^{1(M)}_{(n,m)} = \begin{cases} \lambda & \text{si } 0 \leq n \leq c-1 \\ 0 & \text{si } n=c \end{cases}$$

$$\lambda^{2(M)}_{(n,m)} = \begin{cases} 0 & \text{si } 0 \leq n \leq c-1 \\ \lambda & \text{si } n=c, 0 \leq m \leq M-1 \\ 0 & \text{si } n=c, m=M \end{cases}$$

$$\begin{aligned} \mu^{1(M)}_{(n,m)} &= n\nu; \\ \mu^{2(M)}_{(n,m)} &= 0; \\ a^{1,2(M)}_{(n,m)} &= 0; \end{aligned}$$

$$a^{2,1(M)}_{(n,m)} = \begin{cases} m\mu & \text{si } 0 \leq n \leq c-1 \\ 0 & \text{si } n=c \end{cases}$$

Théorème 3.2. [24]

Soient $\xi_t^{(1)}, \xi_t^{(2)}$ deux processus de migration avec espace des états $A_1, A_2 \subset N^2$ respectivement et de caractéristiques infinitésimales :

$\lambda_n^{i(1)}, \mu_n^{i(1)}, a_n^{i,j(1)} \quad (n \in A_1)$ et $\lambda_n^{i(2)}, \mu_n^{i(2)}, a_n^{i,j(2)} \quad (n \in A_2)$ respectivement.

Si ;

$$\bullet \xi_0^{(1)} \leq_{st} \xi_0^{(2)};$$

$$\bullet \lambda_n^{i(1)} \leq \lambda_n^{i(2)}; i=1,2, n \in A_1 \cap A_2;$$

$$\bullet \mu_n^{i(1)} \geq \mu_n^{i(2)}; i=1,2, n \in A_1 \cap A_2;$$

• $\forall n \in A_1, m \in A_2$ tel que $n_i = m_i$ et $n_j \leq m_j \quad (i, j=1, 2, i \neq j)$ On a :

$$\lambda_n^{i(1)} + a_n^{i,j(1)} \leq \lambda_m^{i(2)} + a_m^{i,j(2)};$$

$$\mu_n^{i(1)} + a_n^{i,j(1)} \geq \mu_m^{i(2)} + a_m^{i,j(2)};$$

Alors il existe un espace de probabilité $(\Omega, \mathfrak{F}, \wp)$ et un processus de Markov de dimension deux $(\eta_t^{(1)}, \eta_t^{(2)})$ tels que ;

• $\eta_t^{(1)}, \eta_t^{(2)}$ ont même distribution que $\xi_t^{(1)}$ et $\xi_t^{(2)}$ respectivement.

• $\eta_t^{(1)}(\omega) \leq \eta_t^{(2)}(\omega) \forall t \geq 0, \forall \omega \in \Omega$.

En appliquant le théorème 3.2 aux deux modèles $\{C(t), N(t)\}$ et $\{C^{(M)}(t), N^{(M)}(t)\}$ qui sont des processus de migration, et en utilisant le théorème 3.1 on a les deux théorèmes suivants :

Théorème 3.3. [27]

Si $(C^{(M)}(0), N^{(M)}(0)) \leq_{st} (C(0), N(0))$;

ALORS $\forall T \geq 0$, ON A : $(C^{(M)}(T), N^{(M)}(T)) \leq_{st} (C(T), N(T))$, ET EN PARTICULIER POUR LA DISTRIBUTION STATIONNAIRE ON A :

$$\left\{ P_{ij}^{(M)} \right\} \leq_{st} \left\{ P_{ij}^{(M+1)} \right\}$$

Théorème 3.4. [27]

Si $(C^{(M)}(0), N^{(M)}(0)) \leq_{st} (C^{(M+1)}(0), N^{(M+1)}(0))$;

ALORS $\forall T \geq 0$, ON A :

$(C^{(M)}(t), N^{(M)}(t)) \leq_{st} (C^{(M+1)}(t), N^{(M+1)}(t))$;

et en particulier, pour la distribution stationnaire on a :

$$\left\{ P_{ij}^{(M)} \right\} \leq_{st} \left\{ P_{ij}^{(M+1)} \right\}$$

3.1.5. Convergence de solutions du modèle tronqué vers celles du modèle original

La question qui doit se poser maintenant, est-ce que ces caractéristiques stationnaires du modèle tronqué convergent vers celles du modèle initial lorsque le niveau de troncature M tend vers l'infini ?

On pose ;

$$\overline{P}_{ij} = \sum_{n=i}^{\infty} \sum_{m=j}^{\infty} P_{nm} \quad . \quad (3.31)$$

En utilisant l'inégalité $\{P_{ij}^{(M)}\} \leq_{st} \{P_{ij}^{(M+1)}\} \leq_{st} \{P_{ij}\}$.

ON A EN PARTICULIER ;

$$\overline{P}_{ij}^{(M)} \leq \overline{P}_{ij}^{(M+1)} \leq \overline{P}_{ij} \quad . \quad (3.32)$$

La suite $\{\overline{P}_{ij}^{(M)}\}$ est donc monotone et bornée, et par conséquent, elle est convergente. On en déduit l'existence d'une limite $a_{ij} = \lim_{M \rightarrow \infty} \overline{P}_{ij}^{(M)}$ pour $M \rightarrow \infty$ et par suite $a_{ij} \leq \overline{P}_{ij}$.

Mais ;

$$P_{ij}^{(M)} = \overline{P}_{ij}^{(M)} + \overline{P}_{i+1j+1}^{(M)} - \overline{P}_{i+1j+1}^{(M)} - \overline{P}_{ij+1}^{(M)} \quad (3.33)$$

D'où l'existence de la limite $b_{ij} = \lim_{M \rightarrow \infty} P_{ij}^{(M)}$ pour $M \rightarrow \infty$.

Par définition, $b_{ij} = a_{ij} + a_{i+1j+1} - a_{i+1j} - a_{ij+1}$, et $\overline{b}_{ij} = a_{ij}$.

Ce qui mène à :

$$\begin{aligned} \overline{b}_{00} &= \sum_{i=0}^C \sum_{j=0}^{\infty} b_{ij} = a_{00} \\ &= \lim_{M \rightarrow \infty} \overline{P}_{00}^{(M)} = \lim_{M \rightarrow \infty} 1 = 1, \end{aligned} \quad (3.34)$$

c'est à dire que les variables b_{ij} satisfont la condition de normalisation et donc on peut conclure que $b_{ij} = P_{ij}$.

Finalement on a :

$$\lim_{M \rightarrow \infty} P_{ij}^{(M)} = P_{ij} \quad (3.35)$$

On a confirmé donc la convergence de la distribution stationnaire du modèle tronqué vers celle du modèle original, ainsi que celles des différentes mesures de performance qui s'écrivent en fonction de cette dernière :

3.1.6. Choix du niveau de troncature M

Aucune base analytique pour le choix de M n'est disponible, une approche d'épreuves et d'erreurs doit être adoptée, de telle sorte à minimiser l'effet de cette approximation sur la solution stationnaire P_{ij} .

On peut commencer par une valeur initiale raisonnable pour M, et on l'augmente progressivement jusqu'à ce qu'un critère convenablement choisi soit rencontré.

On choisit d'augmenter M jusqu'à ce que les différents éléments du vecteur de probabilité ne changent pas de manière significative. Chaque fois que M est augmenté, une valeur δ , définie comme le plus grand changement individuel des éléments de $P_{ij}^{(M)}$ est déterminé. M est choisi comme plus petite valeur de j pour laquelle $\delta < \varepsilon$, où ε est une valeur infinitésimale prédéterminée.

M sera le plus petit entier qui vérifie l'inégalité suivante :

$$\text{Max}_{(i,j)} \left| P_{ij}^{(M)} - P_{ij}^{(M+1)} \right| \leq \varepsilon$$

3.2. Modèle tronqué généralisé

On décrit maintenant d'autres méthodes d'approximation numérique de file d'attente M/M/c avec rappel en les comparant avec la méthode de troncature directe.

L'idée de base est d'imposer une hypothèse simplifiée qui donne un modèle de file d'attente auxiliaire avec un état du système infini et un générateur infinitésimal plus approprié.

3.2.1. Description de quelques modèles tronqués généralisés

G.I.Falin introduit un modèle généralisé simple en considérant que le taux de rappel sera infini lorsque le nombre de clients dans l'orbite dépasse une certaine valeur M.

Soit $X^F = \{(C^{(F)}(t), N^{(F)}(t)), t \geq 0\}$ le processus qui caractérise l'état du système dans ce modèle.

μ_j est le taux de rappel donné par ;

$$\mu_j = \begin{cases} j\mu, & 0 \leq j \leq M \\ \infty, & j > M \end{cases}$$

Le processus $X^F = \{(C^{(F)}(t), N^{(F)}(t)), t \geq 0\}$, où $C^{(F)}(t)$ est le nombre de serveurs occupés et $N^{(F)}(t)$ est le nombre de clients dans l'orbite dans ce système, est un processus de Markov avec espace des états $S^F = \{0, 1, \dots, c-1\} \times \{0, 1, \dots, M\} \cup \{c\} \times \mathbb{N}$.

Les taux de transition d'un point $(i, j) \in S^F$ à un point $(n, m) \in S^F$ sont donnés par :

1. pour $0 \leq i \leq c-1, 0 \leq j \leq M$;

$$q_{(i,j)(n,m)}^{(F)} = \begin{cases} \lambda & \text{si } (n, m) = (i+1, j) \\ i v & \text{si } (n, m) = (i-1, j) \\ j\mu & \text{si } (n, m) = (i+1, j-1) \\ -(\lambda + i v + j\mu) & \text{si } (n, m) = (i, j) \\ 0 & \text{si non} \end{cases}$$

2. pour $i=c, 0 \leq j \leq M-1$;

$$q_{(c,j)(n,m)}^{(F)} = \begin{cases} c v & \text{si } (n, m) = (c-1, j) \\ \lambda & \text{si } (n, m) = (c, j+1) \\ -(\lambda + c v) & \text{si } (n, m) = (c, j) \\ 0 & \text{si non} \end{cases}$$

3. pour $i = c, j \geq M+1$;

$$q_{(c,j)(n,m)}^{(F)} = \begin{cases} cv & \text{si } (n, m) = (c-1, j) \\ \lambda & \text{si } (n, m) = (c, j+1) \\ -(\lambda + cv) & \text{si } (n, m) = (c, j) \\ 0 & \text{si non} \end{cases}$$

Une autre méthode pour ramener le modèle de rappel initial à un modèle numérique traitable, consiste en un modèle proposé par Neuts et Rao [42]. Ils proposent que quand le nombre de clients dans l'orbite excède un niveau N , seulement N d'entre eux sont autorisés à effectuer des rappels tandis que les autres attendent jusqu'à ce que cette période de surcharge expire.

Ce modèle a un taux de rappel ;

$$\mu_j = \begin{cases} j\mu, & 0 \leq j \leq N \\ N\mu, & j > N \end{cases}$$

Cette approximation mène à un générateur infinitésimal qui est homogène à partir du niveau N ci-dessus (où le taux de rappel devient constant).

Le processus associé $X^{NR} = \{(C^{(NR)}(t), N^{(NR)}(t)), t \geq 0\}$, où $C^{(NR)}(t)$ est le nombre de serveurs occupés et $N^{(NR)}(t)$ est le nombre de clients dans l'orbite dans ce système, est un processus de Markov avec espace des états $S^{NR} = \{0, 1, \dots, c\} \times \mathbb{N}$.

Ses taux de transition $q_{(i,j)(n,m)}^{(NR)}$ d'un point $(i,j) \in S^{NR}$ à un point $(n,m) \in S^{NR}$ sont donnés par :

1. Pour $0 \leq i \leq c-1, 0 \leq j \leq N$;

$$q_{(i,j)(n,m)}^{(NR)} = \begin{cases} \lambda & \text{si } (n,m) = (i+1,j) \\ iv & \text{si } (n,m) = (i-1,j) \\ j\mu & \text{si } (n,m) = (i+1,j-1) \\ -(\lambda + iv + j\mu) & \text{si } (n,m) = (i,j) \\ 0 & \text{si non} \end{cases}$$

2. Pour $i = c, 0 \leq j \leq N-1$;

$$q_{(c,j)(n,m)}^{(NR)} = \begin{cases} cv & \text{si } (n,m) = (c-1,j) \\ \lambda & \text{si } (n,m) = (c,j+1) \\ -(\lambda + cv) & \text{si } (n,m) = (c,j) \\ 0 & \text{si non} \end{cases}$$

Par analogie au cas scalaire, puisque les transitions d'état sont entre les blocs voisins, on peut constater que la valeur de Π_j est une fonction seulement des taux de transition entre les états de niveau j et les états de niveau $j-1$ (d'après le système 3.38). Puisque ces taux de transition ne dépendent pas de j , ceci suggère qu'il existe une matrice constante R telle que :

$$\Pi_j = \Pi_{j-1} R, \quad j = 1, 2, 3, \dots \quad (3.39)$$

et que les valeurs de Π_j $j = 2, 3, \dots$, ont une forme matricielle géométrique, donnée par :

$$\Pi_j = \Pi_0 R^j, \quad j = 1, 2, 3, \dots \quad (3.40)$$

En remplaçant la relation (3.40) dans (3.38), on obtient :

$$\Pi_0 R^{j-1} A_0 + \Pi_0 R^j A_1 + \Pi_0 R^{j+1} A_2 = 0, \quad j = 1, 2, 3, \dots \quad (3.41)$$

pour $j=1$ on a:

$$R^2 A_2 + R A_1 + A_0 = 0 \quad (3.42)$$

la condition de normalisation est satisfaite si,

$$\sum_{j=1}^{\infty} R^j < \infty \quad (3.43)$$

La somme (3.43) converge si le rayon spectral de R est inférieur à l'unité. Ceci est dû au fait que toutes les valeurs propres de R doivent être de modules inférieures à 1 pour que la somme converge.

Posons,

$$S = \sum_{j=0}^{\infty} R^j = I + R + R^2 + \dots$$

D'où,

$$S(I - R) = (I + R + R^2 + \dots) - (R + R^2 + \dots) = I.$$

Avec I est la matrice unité. Puis, à condition que la somme converge, nous pouvons écrire :

$$S = (I - R)^{-1}.$$

Théorème 3.5.[41]

Un processus de générateur Q^{\sim} est récurrent positif si et seulement si la matrice R qui satisfait l'équation ; $R^2 A_2 + R A_1 + A_0 = 0$, a toutes ses valeurs propres de modules inférieures à un. et le système d'équations :

$$\begin{cases} \Pi_0 (B_0 + R B_1) = 0 \\ \Pi_0 (I - R)^{-1} E = 1 \end{cases} \quad (3.44)$$

Où $A_1 = A_{1N}$ et $A_2 = A_{2N}$

On étudiera par la suite le choix de N par des méthodes algorithmiques de telle sorte qu'on minimise l'effet de cette approximation sur les probabilités d'équilibre .

Sous les conditions du théorème 3.5., on peut constater que Π a la forme d'une matrice géométrique modifiée qui peut être efficacement calculée par la relation récurrente ;

$$\Pi_j = \Pi_N R^{j-N} \quad , j \geq N \quad (3.47)$$

Où la matrice R est l'unique solution non- négative de rayon spectral inférieure à un qui satisfait l'équation ;

$$R^2 A_2 + R A_1 + A_0 = 0 \quad (3.48)$$

Théoriquement , R est donnée par $\lim R_n$ lorsque $n \rightarrow \infty$, où la suite $\{ R_n \}$ est donnée par la formule récursive suivante :

$$R_0 = 0 \quad , R_{n+1} = - A_0 A_1^{-1} - R_n^2 A_2 A_1^{-1} \quad , \quad n \geq 1. \quad (3.49)$$

et le test d'arrêt est donné par : $\text{Max} [R_{i,j(n+1)} - R_{i,j(n)}] < \varepsilon$.

Théorème.3.6.

Sous la condition de stabilité du système et pour $0 \leq j \leq N-1$, le vecteur de probabilité stationnaire Π satisfaisant ; $\Pi Q^* = 0$ et $\Pi E = 1$ est donné par :

$$\Pi_j = \Pi_N \prod_{m=N}^{j+1} H_m \quad . \quad \text{si } 0 \leq j \leq N-1. \quad (3.50)$$

OU ;

$$H_j = - A_{2j} (H_{j-1} A_0 + A_{1j-1})^{-1} \quad j \geq 1. \quad (3.51)$$

LE VECTEUR Π_N EST OBTENU PAR LA SOLUTION DE L'EQUATION ;

$$\Pi_N (H_N A_0 + A_{N-1} + R A_2) = 0 ; \quad (3.52)$$

et la condition de normalisation ;

$$\Pi_N \left[\sum_{j=0}^{N-1} \prod_{m=j+1}^N H_m + (I - R)^{-1} \right] E = 1 ; \quad (3.53)$$

PREUVE.

L'EQUATION MATRICIELLE $\Pi Q^* = 0$ S'ECRIT DE LA FAÇON SUIVANTE :

$$\left\{ \begin{array}{l} \Pi_0 A_{10} + \Pi_1 A_{21} = 0 \\ \Pi_0 A_0 + \Pi_1 A_{11} + \Pi_2 A_{22} = 0 \\ \cdot \\ \cdot \\ \Pi_{j-1} A_0 + \Pi_j A_{1j} + \Pi_{j+1} A_{2j+1} = 0 ; \quad 0 \leq j \leq N-1 \\ \Pi_{N-1} A_0 + \Pi_N A_N + \Pi_{N+1} A_2 = 0 \\ \Pi_j A_0 + \Pi_{j+1} A_1 + \Pi_{j+2} A_2 = 0 ; \quad j \geq N \end{array} \right.$$

A partir de ce système d'équations, on obtient ;

$$\Pi_0 = - \Pi_1 A_{21} A_{10}^{-1} = \Pi_1 H_1$$

$$\Pi_1 = - \Pi_2 A_{22} [(E - A_{21} A_{10}^{-1} A_0 A_{10}^{-1}) A_{10}]^{-1}$$

$$= - \Pi_2 A_{22} [A_{10} - A_{21} A_{10}^{-1} A_0]^{-1}$$

$$= - \Pi_2 A_{22} [H_1 A_0 + A_{10}]^{-1} = \Pi_2 H_2.$$

Par récurrence on a $\Pi_{j-1} = \Pi_j H_j ; \quad 0 \leq j \leq N-1$

Avec :

$$H_j = - A_{2j-1} [H_{j-1} A_0 + A_{1j-1}]^{-1} . \quad 0 \leq j \leq N-1$$

Et par suite $\Pi_j = \Pi_N \prod_{m=j+1}^N H_m \quad \text{pour } 0 \leq j \leq N-1$

Il reste à trouver le vecteur Π_N en combinant les deux équations :

$$\Pi_{N-1} A_0 + \Pi_N A_1 + \Pi_{N+1} A_2 = 0 \quad \text{et} \quad \Pi_{j+1} = \Pi_j R , j \geq N .$$

Ce qui mène à : $\Pi_N [H_N A_0 + A_1 + R A_2] = 0 .$

D'où , Π_N va être déterminé par l'équation (3.52) et la condition de normalisation ;

$$\sum_{j=0}^{\infty} \Pi_j E = \sum_{j=0}^{N-1} \Pi_j E + \sum_{j=N}^{\infty} \Pi_j E = \sum_{j=0}^{N-1} \Pi_N \prod_{m=i+1}^N H_m E + \sum_{j=N}^{\infty} \Pi_m R^{j-N} E = 1$$

qui est exactement l'équation (3.53) du théorème.

3.2.3. MESURES DE PERFORMANCE

**LORSQUE LE VECTEUR DE PROBABILITE STATIONNAIRE PIEST CALCULE ,
UNE VARIETE DE CARACTERISTIQUES DE PERFORMANCE PEUVENT ETRE
EXPLICITEMENT CALCULEES :**

1. la probabilité de blocage B;

$$\begin{aligned} B = \sum_{j=0}^{\infty} \Pi_{c,j} E &= \sum_{j=0}^{N-1} \Pi_{c,j} + \left(\sum_{j=N}^{\infty} \Pi_N R^{j-N} \right) c \\ &= \sum_{j=0}^{N-1} \Pi_{c,j} + (\Pi_N (I-R)^{-1}) c \end{aligned} \quad (3.54)$$

2. Le nombre moyen de serveurs occupés Y;

$$\begin{aligned} Y &= \sum_{i=1}^c \sum_{j=0}^{\infty} i \Pi_{i,j} \\ &= \sum_{i=1}^c \sum_{j=0}^{N-1} i \Pi_{i,j} + \Pi_N (I-R)^{-1} a \end{aligned} \quad (3.55)$$

où $a = (0, 1, \dots, c)$.

3. Le nombre moyen de clients dans l'orbite N ;

$$N = \sum_{j=0}^{\infty} \sum_{i=0}^c j \Pi_{ij} = \left[\sum_{j=0}^{N-1} \sum_{i=0}^c j \Pi_{ij} \right] + \Pi_N \sum_{j=N}^{\infty} \sum_{i=0}^c j R^{j-N}$$

$$N = \left[\sum_{j=0}^{N-1} \sum_{i=0}^c j \Pi_{ij} \right] + \Pi_N \sum_{j=N}^{\infty} \sum_{i=0}^c j R^{j-N}$$

$$N = \left[\sum_{j=0}^{N-1} \sum_{i=0}^c j \Pi_{ij} \right] + \Pi_N R (I-R)^{-2} E \quad (3.56)$$

4. Le temps moyen d'attente d'un client dans l'orbite W_o ;

$$W_o = \frac{N}{\lambda} \quad (3.57)$$

3.2.4. La solution algorithmique du modèle X^F

On considère les états (i, j) , $j \geq 0$ et $0 \leq i \leq c$ arrangés dans l'ordre lexicographique suivant:

$(0, 0), (1, 0), \dots, (c, 0), (0, 1), \dots, (c, 1), (0, 2), (1, 2), \dots, (c, M), (c, M+1), \dots$

L'ensemble des états $S_j = \{(0, j), (1, j), \dots, (c, j)\}$ est appelé les états de niveau j .

Q^* peut être écrit par :

$$Q^* = \left(\begin{array}{cccccc|ccc} A_{10} & A_0 & & & & & 0 & \dots & 0 \\ A_{21} & A_{11} & A_0 & & & & & & \\ & A_{22} & A_{12} & A_0 & & & & & \\ & & \cdot & \cdot & \cdot & & & & \\ & & & & & A_{2M-2} & A_{1M-2} & A_0 & \\ & & & & & & A_{2M-1} & A_{1M-1} & \\ & & & & & & & \lambda & 0 \\ \hline 0 & \dots & & & & & cv & & \\ & \cdot & & & & & 0 & & \\ & & & & & & 0 & & \end{array} \right)$$

$$\left(\begin{array}{ccc} -(cv + \lambda) & \lambda & 0 \\ cv & -(cv + \lambda) & \lambda \\ 0 & cv & -(cv + \lambda) \end{array} \right)$$

Où $A_0 = \lambda$, A_{1j} et A_{2j} , $j \geq 0$ sont des matrices carrées d'ordre $c+1$, données par :

$$A_{1j} = \left(\begin{array}{cccc|ccc} q_{0j} & \lambda & 0 & & & 0 & 0 \\ v & q_{1j} & \lambda & 0 & & & \\ 0 & 2v & q_{2j} & \lambda & & & \\ \cdot & \cdot & \cdot & \cdot & & & 0 \\ 0 & & 0 & & & & \\ & & & & & & \\ & & & & & (c-1)v & q_{c-1j} & \lambda \\ 0 & 0 & & & & 0 & cv & q_{cj} \end{array} \right)$$

Où $q_{ij} = -(\lambda + iv + j\mu)$, $0 \leq i \leq c-1$ et $q_{cj} = -(\lambda + cv)$.

$$A_{2j} = \left(\begin{array}{cccccc|ccc} 0 & j\mu & 0 & & & & 0 & 0 \\ 0 & 0 & j\mu & 0 & & & & 0 \\ 0 & 0 & 0 & 0 & j\mu & 0 & & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & & 0 \end{array} \right)$$

0

L'application de la méthode matricielle géométrique se base sur la procédure suivante :

1. La chaîne de Markov induite de générateur infinitésimal A, a un seul état c et donc, admet une distribution stationnaire $\Pi c = 1$.

2. Les matrices A_0, A_1, A_2 sont des matrices 1×1 ; c'est à dire sont des nombres réels :
 $A_0 = \lambda, A_1 = -(\lambda + c v), A_2 = c v$.

Ainsi la condition de stabilité a lieu pour $\lambda < c v$.

3. La matrice R satisfait l'équation :

$R^2 c v - R(\lambda + c v) + \lambda = 0$; cette équation admet les solutions ; $R=1$ où $R = \frac{\lambda}{c v}$.

Comme la deuxième solution est inférieure à 1 ,

$$R = \frac{\lambda}{c v} \quad (3.58)$$

D'après l'équation (3.47) , on a pour $j > M$:

$$P_{cj}^{(F)} = P_{cM}^{(F)} \left(\frac{\lambda}{c} \right)^{j-M} \quad (3.59)$$

Les probabilités $P_{ij}^{(F)}$, $0 \leq i \leq c, 0 \leq j \leq M$, satisfont les équations de Kolmogorov données par (3.1), (3.2) ,(3.3) et (3.4). On écrit :

$$(\lambda + i v + j \mu) P_{ij}^{(F)} = \lambda P_{i-1j}^{(F)} + (j+1) \mu P_{i-1j+1}^{(F)} + (i+1) v P_{i+1j}^{(F)} \quad \text{si } 0 \leq i \leq c-1, \quad 0 \leq j \leq M-1 \quad (3.60)$$

$$(\lambda + i v + M \mu) P_{iM}^{(F)} = \lambda P_{i-1M}^{(F)} + (i+1) v P_{i+1M}^{(F)} \quad \text{si } 0 \leq i \leq c-1 \quad (3.61)$$

$$(\lambda + c v) P_{cj}^{(F)} = \lambda P_{c-1j}^{(F)} + (j+1) \mu P_{c-1j+1}^{(F)} + \lambda v P_{c,j-1}^{(F)} \quad \text{si } 0 \leq j \leq M-1$$

$$c v P_{cM}^{(F)} = \lambda P_{c-1M}^{(F)} + \lambda P_{cM-1}^{(F)} \quad (3.62)$$

D'où, les probabilités $P_{ij}^{(F)}$ coïncident à une constante près avec $P_{ij}^{(M)}$, c'est à dire ;

$$P_{ij}^{(F)} = \text{const} \cdot P_{ij}^{(M)} ; 0 \leq i \leq c, 0 \leq j \leq M \quad (3.63)$$

D'après la condition de normalisation ; $\sum_{i=0}^c \sum_{j=0}^{\infty} P_{ij}^{(F)} = 1$.

On a :

$$\sum_{i=0}^c \sum_{j=0}^M P_{ij}^{(F)} + \sum_{j=M+1}^{\infty} P_{cj}^{(F)} = 1 \quad (3.64)$$

En combinant les équations (3.64) et (3.59), on obtient ,

$$\sum_{i=0}^C \sum_{j=0}^M P_{ij}^{(F)} + P_{cM}^{(F)} \sum_{j=M+1}^{\infty} \left(\frac{\lambda}{cv} \right)^{j-M} = 1 \quad (3.65)$$

Ainsi l'équation (3.65) donne :

$$\sum_{i=0}^C \sum_{j=0}^M \text{const } P_{ij}^{(M)} + P_{cM}^{(F)} \cdot \frac{\lambda}{cv-\lambda} = 1 \quad ,$$

et par suite,

$$\text{Const} = 1 - \frac{\lambda}{cv-\lambda} P_{cM}^{(F)} \quad (3.66)$$

EN REMPLAÇANT LA VALEUR DE LA CONST DANS (3.64) ; ON A :

$$P_{cM}^{(F)} = \frac{P_{cM}^{(M)}}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} \quad .$$

D'où:

$$\text{const} = \frac{1}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} \quad (3.67)$$

On peut calculer maintenant, les différents paramètres de performance associés au modèle $X^{(F)}$, en particulier la probabilité de blocage $B^{(F)}$, le nombre moyen de clients dans l'orbite $N^{(F)}$ et le nombre moyen de serveurs occupés $Y^{(F)}$.

Théorème

$B^{(F)}$, $N^{(F)}$ et $Y^{(F)}$ sont données par les formules suivantes :

$$B^{(F)} = \frac{B^{(M)} + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} \quad (3.68)$$

Où

$$N^{(F)} = \frac{N^{(M)} + \frac{(M+1)cv - M\lambda}{(cv - \lambda)^2} \lambda P_{cM}^{(M)}}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} \quad (3.69)$$

$$Y^{(F)} = \frac{Y^{(M)} + \frac{c\lambda}{(cv - \lambda)} P_{cM}^{(M)}}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} \quad (3.70)$$

Preuve

Par définition, de

$$\begin{aligned} B^{(F)} = P(c^{(F)}(t) = j) &= \sum_{j=0}^{\infty} P_{cj}^{(F)} = \sum_{j=0}^M P_{cJ}^{(F)} + \sum_{j=M+1}^{\infty} P_{cM}^{(F)} \left(\frac{\lambda}{cv}\right)^{j-M} \\ &= \text{const} \sum_{j=0}^M P_{cj}^{(M)} + \text{const} \cdot P_{cM}^{(M)} \cdot \frac{\lambda}{cv - \lambda} \\ &= \frac{B^{(M)}}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} + \frac{P_{cM}^{(M)} \lambda}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)} cv - \lambda} \\ &= \frac{B^{(M)} + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} \quad ; \text{ d'où le résultat.} \end{aligned}$$

De la même manière :

$$\begin{aligned} N^{(F)} &= \sum_{i=0}^C \sum_{j=0}^{\infty} j P_{ij}^{(F)} = \text{const} \sum_{i=0}^C \sum_{j=0}^M j P_{ij}^{(M)} + \sum_{j=M+1}^{\infty} j P_{cM}^{(F)} = \text{const} N^{(M)} + \text{const} P_{cM}^{(M)} \sum_{j=M+1}^{\infty} j \left(\frac{\lambda}{cv}\right)^{j-M} \\ &= \text{const} \left(N^{(M)} + P_{cM}^{(M)} \left((M+1) \frac{\lambda}{cv - \lambda} + \frac{\lambda^2}{(cv - \lambda)^2} \right) \right) \\ &= \frac{N^{(M)} + \frac{(M+1)(cv - \lambda) + \lambda}{(cv - \lambda)^2} \lambda P_{cM}^{(M)}}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} \\ N^{(F)} &= \frac{N^{(M)} + \frac{(M+1)cv - M\lambda}{(cv - \lambda)^2} \lambda P_{cM}^{(M)}}{1 + \frac{\lambda}{cv - \lambda} P_{cM}^{(M)}} \quad ; \text{ d'où le résultat.} \end{aligned}$$

Et finalement,

$$Y^{(F)} = E(C^{(F)}(t)) = \sum_{i=0}^c \sum_{j=0}^{\infty} i P_{ij}^{(F)} = \sum_{i=0}^c \sum_{j=0}^M i P_{ij}^{(F)} + c \cdot \text{const.} P_{cM}^{(M)} \cdot \frac{\lambda}{c\nu - \lambda}$$

$$Y^{(F)} = \text{const} \left(Y^{(M)} + \frac{c \lambda P_{cM}^{(M)}}{c\nu - \lambda} \right)$$

$$= \frac{Y^{(M)} + \frac{c \lambda}{(c\nu - \lambda)} P_{cM}^{(M)}}{1 + \frac{\lambda}{c\nu - \lambda} P_{cM}^{(M)}} ; \text{ d'où le résultat du théorème.}$$

3.2.5. Solutions algorithmique du modèle $X^{(NR)}$

On partitionne l'espace des états $S^{NR} = \{0, 1, \dots, c\} \times \mathbb{N}$ du processus $\{(C^{(NR)}(t), N^{(NR)}(t)), t \geq 0\}$ par $S_0 \cup S_1 \cup \dots$, où les $S_j = \{(0, j), (1, j), \dots, (c, j), j \geq N\}$.

On considère les états initiaux (i, j) pour $0 \leq i \leq c, 0 \leq j \leq N-1$ et à partir des taux de transitions $q_{(i,j)(n,m)}^{(NR)}$, on peut écrire le générateur infinitésimal du processus $X^{(NR)}$ sous la forme (3.37) avec les matrices carrées suivantes :

$$A_0 = \begin{pmatrix} 0 & 0 & & & & & & & 0 \\ 0 & 0 & & & & & & & 0 \\ \cdot & \cdot & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ 0 & 0 & & & & & & & \\ 0 & 0 & & & & & & & \lambda \end{pmatrix}$$

$$A_1 = A_{1N} \text{ et } A_2 = A_{2N}$$

L'application de la méthode matricielle géométrique implique la procédure suivante :

1. $A = A_0 + A_1 + A_2$; est donnée par:

$$\Pi A_0 E = (\Pi(0), \Pi(1), \dots, \Pi(c)) \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \lambda \end{pmatrix} = \lambda \Pi(c)$$

$$\Pi A_2 E = (\Pi(0), \Pi(1), \dots, \Pi(c)) \begin{pmatrix} 0 & 0 & N\mu & 0 \dots & 0 \\ 0 & 0 & 0 & N\mu \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & \vdots & N\mu \\ & & & & 0 \end{pmatrix} \cdot E$$

$$= (\Pi(0), \Pi(1), \dots, \Pi(c)) \cdot \begin{pmatrix} N\mu \\ N\mu \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

$$= (\Pi(0) + \Pi(1) + \dots + \Pi(c-1)) N\mu$$

$$= (1 - \Pi(c)) N\mu$$

Ainsi $\lambda < \frac{N\mu(1-\Pi_c)}{\Pi_c}$ est la condition nécessaire et suffisante de l'ergodicité du processus $X^{(NR)}$.

$$\text{Soit } \rho(N) = \lambda \left[\frac{N\mu(1-\Pi_c)}{\Pi_c} \right]^{-1} \quad (3.73)$$

Autrement dit, la condition d'équilibre est vérifiée lorsque $\rho(N) < 1$.

En appliquant la règle de l'Hôpital de manières répétées on obtient, pour $N \rightarrow \infty$, que la

$$\text{quantité } \left[\frac{N\mu(1-\Pi_c)}{\Pi_c} \right] \text{ tend vers } c.$$

Et par suite $\rho(N)$ tend vers $\frac{\lambda}{c\nu} = \rho$; qui est la condition de stabilité du système initial.

•Evaluation de la matrice R :

Pour évaluer la matrice R, on prend l'avantage de la structure matricielle spéciale associée à la matrice A_0 . On remarque que cette dernière peut être écrite comme suit :

$$A_0 = \lambda e_c e_c^t$$

où e_c est un vecteur colonne de dimension $(c + 1)$ tel que tous ses éléments sont égaux à zéro excepté le dernier qui est égal à un.

Puisque A_0 est de cette forme, la matrice R peut être écrite explicitement et déterminé en prenant en ligne de compte que R est donnée par $\text{Lim}_{n \rightarrow \infty} R_n$, où $\{R_n\}$ est une suite de matrices définies par la formule récursive:

$$R_0 = 0, R_{n+1} = -A_0 A_1^{-1} - R_n^2 A_2 A_1^{-1}, \quad n \geq 1.$$

Ainsi, on remarque qu'une ligne zéro dans A_0 produit une ligne zéro dans R_1 , et par conséquent, dans tous les itérés successifs.

Pour être précis, R est de la forme suivante :

$$R = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ u \end{pmatrix} \quad (3.74)$$

Donc le problème se réduit à la détermination du vecteur ligne u :

De (3.69) on trouve :

$$\begin{aligned} u R &= u_c u \\ \text{et } \eta &= u_c \end{aligned} \quad (3.75)$$

où u_c désigne le dernier élément de u et η est le rayon spectral de la matrice R

En multipliant l'équation fondamentale $R^2 A_2 + R A_1 + A_0 = 0$ par le vecteur u et en utilisant (3.75), on obtient :

$$u (\eta^2 A_2 + \eta A_1 + A_0) = 0 \quad (3.76)$$

En fait, la théorie générale de Neuts établit une procédure pour calculer le rayon spectral η donnée par le lemme (3.1).

En appliquant une méthode élémentaire comme celle de bisection, η peut être évalué comme une solution de l'équation $\eta = x(\eta)$ dans l'intervalle $[z_1, z_2]$ dans $[0, 1]$. (l'objectif de cette méthode est la résolution d'équations non linéaires du type $f(x) = 0$).

Cette méthode repose sur le constat que si le produit $f(z_1).f(z_2)$ est négatif, alors la fonction f s'annule au moins une fois sur l'intervalle $[z_1, z_2]$. Les différentes étapes de la méthode peuvent être résumées comme suit :

1. Choisir un l'intervalle $[\eta_1 = z_1, \eta_2 = z_2]$ tel que $f(z_1).f(z_2) < 0$
2. Calculer la valeur de la fonction en $\eta_3 = (z_1 + z_2) / 2$

3. On choisit un nouvel intervalle $[\eta_1, \eta_3]$ ou $[\eta_3, \eta_2]$ en respectant la condition du 1. On est alors assuré de toujours encadrer la racine.
4. Répéter les étapes 2 et 3 jusqu'à l'obtention de la précision désirée, c'est-à-dire jusqu'à ce que $f(\eta_3) < \varepsilon$, ε étant la précision désirée.

On pose $f(\eta) = x(\eta) - \eta$.

On utilise le fait que $x(1) = 1$, on prend $z_2 = 1$. Il suffit donc de choisir z_1 comme un petit entier positif inférieur à un pour vérifier la condition 1, et d'où la solution η est fournie.

Enfin, le vecteur u est explicitement déterminé par les équations (3.76) et (3.36) dont on a le système d'équations :

$$\begin{cases} u (\eta^2 A_2 + \eta A_1 + A_0) = 0 \\ u e = (\lambda / (N+1) \mu) + \eta \end{cases} \quad (3.77)$$

•Le choix du niveau de troncature N

On utilise le fait que R est en fonction de N , on écrit $R(N)$ au lieu de N .

Dans la méthode matricielle géométrique, la solution stationnaire des états (i, j) dépend de $\eta(N)$; le rayon spectral de $R(N)$. Pour minimiser l'effet de cette approximation sur cette solution, N doit être choisi de telle sorte que $\eta(N)$ converge vers $\eta(\infty)$, autrement dit $|\eta(N) - \eta(\infty)| < \varepsilon$.

une implémentation de cet algorithme est donnée dans la section 4 de ce chapitre.

.3.3.Conclusion

A travers ce chapitre, on a étudié les deux méthodes d'approximation de système M/M/c avec rappels; un algorithme numérique est proposé pour le calcul de la distribution stationnaire qui est au centre de l'analyse du système tronqué, puisque les différents paramètres de performance sont complètement déterminés par cette distribution.

La deuxième approximation est basée sur la théorie de Marcel Neuts qui est appliquée aux processus quasi-de naissance et de mort.

3.4.Résultats numériques

Dans cette section, on présente quelques résultats numériques de l'implémentation de chacune des méthodes de troncature directe et la méthode matricielle géométrique.

Notre premier objectif est de montrer l'efficacité de la méthode matricielle géométrique appliquée à ce genre de système.

Notre second objectif est d'étudier l'effet du taux de rappel sur les différents résultats obtenus.

Les résultats sont obtenus pour $v = 1$.

3.4.1. Différentes mesures de performance obtenues par la méthode de troncature directe (algorithme 1)

M	$B^{(M)}$	$N^{(M)}$	$N_{APP}^{(M)}$	$Y^{(M)}$
1	0.2906	0.1705	2.2449	3.4379
5	0.4539	0.9797	2.2870	3.8206
10	0.5102	1.7057	2.3583	3.9463
15	0.5268	2.0893	2.3844	3.9828
20	0.5321	2.2687	2.3931	3.9944
25	0.5338	2.3456	2.3960	3.9981
30	0.5344	2.3773	2.3970	3.9994
40	0.5347	2.3946	2.3974	3.9999
45	0.5347	2.3964	2.3974	4

Tableau 4.1 :Dépendance des mesures de performance avec le niveau de troncature M pour $c= 5$, $\lambda= 4$ et $\mu= 20$.

NOTONS COMME PREMIER RESULTAT IMPORTANT, QUE LORSQUE LE NIVEAU DE TRONCATURE M AUGMENTE, LES PARAMETRES DE PERFORMANCE DU SYSTEME TRONQUE S'APPROCHENT DE CEUX DU SYSTEME INITIAL (VOIR TABLEAU 4.1.).

Remarque

On utilise la formule de l'équation (2.49) pour approcher le nombre moyen de clients dans l'orbite, on obtient :

$$N_{app}^{(M)} = \frac{v + \mu}{\mu} \cdot \frac{\lambda - v \cdot \text{var}(C^{(M)})}{cv - \lambda}$$

3.4.2. Résultats numériques pour le modèle $X^{(F)}$

On considère maintenant l'approximation des mesures de performance du système initial à l'aide du modèle $X^{(F)}$.

Dans le tableau 4.2. on donne le nombre moyen de serveurs occupés et la probabilité de blocage dans les deux modèles pour plusieurs valeurs de M.

M	$Y^{(M)}$	$Y^{(F)}$	$B^{(M)}$	$B^{(F)}$
0	3.4379	4	0.2906	0.5458
5	3.8206	4	0.4539	0.5370
10	3.9463	4	0.5102	0.5352
15	3.9828	4	0.5268	0.5348
20	3.9944	4	0.5321	0.5347
25	3.9981	4	0.5338	0.5347
30	3.9994	4	0.5344	0.5347
40	3.999	4	0.5347	0.5347
45	4.000	4	0.5347	0.5347

Tableau 4.2 : Comparaison de la probabilité de blocage et le nombre moyen de serveurs occupés dans les deux systèmes $X^{(M)}$ et $X^{(F)}$ pour $c = 5$, $v = 1$, $\lambda = 4$, $\mu = 20$.

Un exemple comparatif de quelques mesures de performance obtenues par les deux modèles $X^{(M)}$ et $X^{(F)}$ est donné dans le tableau 4.2. Au delà de cette comparaison, on notera que le nombre moyen de serveurs occupés et la probabilité de blocage du modèle $X^{(F)}$ s'approchent plus de ceux du système initial que celles données par le modèle $X^{(M)}$. Nous en déduisons donc que le modèle $X^{(F)}$ est plus consistant. Ce résultat est testé par le cas $c = 1$ où des fonctions explicites sont déjà obtenues par les équations (2.25) et (2.26). (tableau 4.3).

M	$Y^{(M)}$	$Y^{(F)}$	$B^{(M)}$	$B^{(F)}$	$N^{(M)}$	$N^{(F)}$
5	0.4957	0.5000	0.4957	0.5000	0.4334	0.5040
10	0.4999	0.5000	0.4999	0.5000	0.5468	0.5475
20	0.5000	0.5000	0.5000	0.5000	0.5500	0.5500
25	0.5000	0.5000	0.5000	0.5000	0.5500	0.5500
30	0.5000	0.5000	0.5000	0.5000	0.5500	0.5500

Tableau 4.3 : Comparaison de la probabilité de blocage et le nombre moyen de serveurs occupés dans les deux systèmes $X^{(M)}$ et $X^{(F)}$ pour $c = 1$, $\lambda = 0.5$ et $\mu = 10$.

3.4.3. Choix du niveau de troncature M

On peut commencer par une valeur M, et on l'augmente progressivement jusqu'à ce que le critère donné par le paragraphe 3.1.6 soit vérifié.

Les résultats sont obtenus pour $\varepsilon = 10^{-6}$.

On présente dans le tableau suivant les valeurs appropriés de M obtenues par ce critère ainsi que les mesures de performance associées pour différentes valeurs de ρ et μ .

c	ρ	μ	M	$B^{(M)}$	$Y^{(M)}$	$N^{(M)}$
7	0.5	50	8	0.0737	3.5	0.0614
		5	10	0.0640	3.5	0.1183
		0.5	11	0.0514	3.5	0.4395
5	0.8	50	36	0.5453	3.9998	1.3688
		5	40	0.5035	3.9998	2.0951
		0.5	48	0.4395	3.9998	7.9690
1	0.95	50	95	0.95	0.95	16.4800
		5	120	0.95	0.95	17.4291
		0.5	200	0.95	0.95	53.8625

Tableau 4.4 : les valeurs critiques M et les mesures de performance associées.

3.4.4. Résultats numériques obtenus pour le modèle $X^{(NR)}$

3.4.4.1. Choix du niveau de troncature N

Pour comprendre comment varie l'intensité de trafic $\rho(N)$ et le rayon spectral $\eta(N)$ en fonction de N, des expériences numériques sont graphiquement présentées par les figures 1, 2 et 3 ci-dessous :

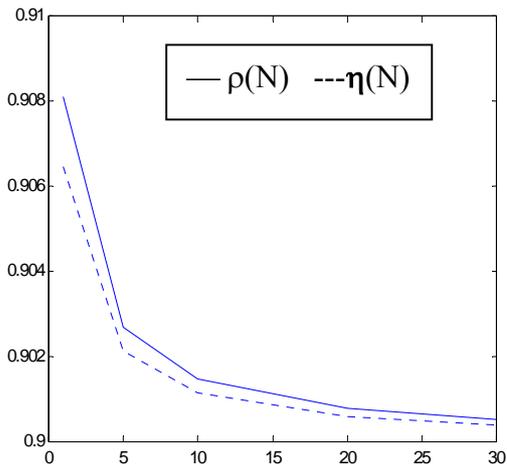


Fig.1 :rayon spectral ($\rho = 0.9, \mu = 50$)

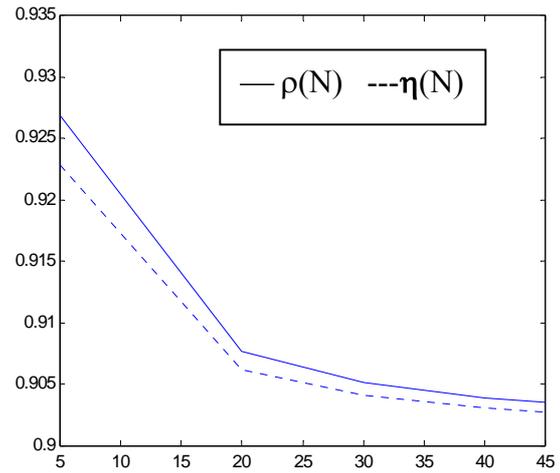


Fig.2 :rayon spectral ($\rho = 0.9, \mu = 5$)

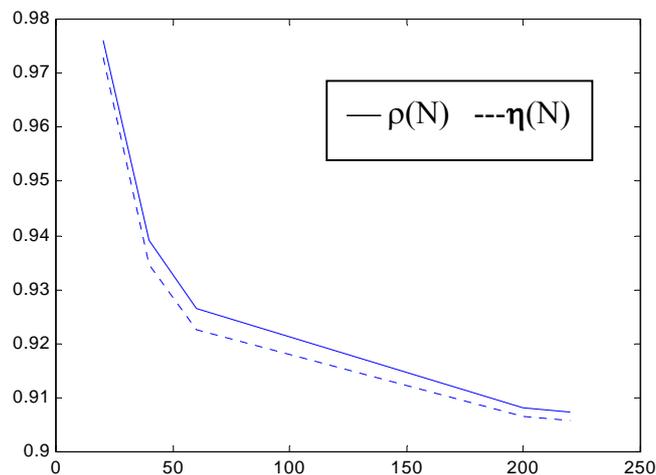


Fig.3 :rayon spectral ($\rho = 0.9, \mu = 0.5$)

A partir des graphes, on constate que plus la valeur de N augmente, plus les valeurs $\rho(N)$ et du rayon spectral $\eta(N)$ s'approchent de ρ , qui est la condition de stabilité du système initial. Cette approche s'améliore avec l'augmentation du taux de rappel.

D'où l'importance des conditions suivantes pour le choix de N :

C₁. La condition de stabilité donnée par la relation (3.72) doit être vérifiée.

C₂. Le rayon spectral $\eta(N)$ doit converger vers $\eta(\infty) = \rho$, autrement dit $|\eta(N) - \rho| < \varepsilon$.

C₃. Choisir N tel que $|\eta(N) - \eta(N+1)| < \varepsilon$.

A partir de ce critère de sélection, on donne les valeurs appropriées de N et les mesures de performance associées, pour différentes valeurs de c et ρ , pour $\varepsilon = 0.001$ (table 4.5).

C	ρ	μ	N	$B^{(NR)}$	$Y^{(NR)}$	$N^{(NR)}$
7	0.5	50	5	0.0738	3.5	0.0803
		5	8	0.0640	3.5	0.1188
		0.5	10	0.0514	3.5	0.4554
5	0.8	50	5	0.5453	4	2.3688
		5	10	0.5035	4	2.8850
		0.5	35	0.4395	4	8.1178
1	0.95	50	5	0.95	0.95	17.5491
		5	18	0.95	0.95	21.3009
		0.5	45	0.95	0.95	74.5410

Tableau 4.5 : Valeurs critiques N et les mesures de performance associées.

On constate que les valeurs du nombre moyen de clients dans l'orbite $N^{(NR)}$, obtenues par le modèle $X^{(NR)}$ sont plus grandes. Etant donné le raisonnement de l'approximation, la limitation des rappels à un nombre fixé de clients fait déplacer la masse de probabilité vers de plus grandes valeurs de j ayant pour résultat la surestimation du nombre moyen de clients dans l'orbite.

On peut également constater que le nombre moyen de serveurs occupés est indépendant de la valeur de μ et est toujours égal à λ / v .

3.4.5. Comparaison numérique des résultats

Pour simplifier la notation, nous pouvons appeler les paramètres M , K et N comme niveaux de troncature correspondant aux modèles $X^{(M)}$, $X^{(F)}$ et $X^{(NR)}$ respectivement. Selon ce qui précède, les caractéristiques de performance des modèles tronqués convergent vers celles du modèle initial lorsque le niveau de troncature tend vers l'infini. Ainsi, nous mesurerons l'exactitude d'approximation par la vitesse de la convergence. Dans ce sens, le plus petit niveau de troncature pour atteindre la valeur vraie de n'importe quelle mesure de performance signifie une convergence plus rapide et par conséquent, une meilleure approximation.

On va considérer par la suite la probabilité de blocage comme l'une des mesures de performance afin de déterminer les valeurs appropriées de M , K et N .

• La première étape consiste à considérer que la vraie valeur de B peut être obtenue en choisissant un niveau de troncature suffisamment grand dans l'un des deux modèles. Choisir $X^{(M)}$ par exemple. On fait augmenter la valeur de troncature de telle sorte à avoir un niveau T qui vérifie la condition :

$$|B^{(M)}(T) - B^{(M)}(T+1)| \leq 10^{-4}.$$

On considère par suite que $B^{(M)}(T)$ est choisie comme la vraie valeur de B .

• La deuxième étape consiste à déterminer la valeur critique M comme le premier entier qui vérifie l'inégalité : $|B^{(M)}(M) - B| \leq 10^{-2}$.

Le même critère est appliqué aux processus $X^{(F)}$ et $X^{(NR)}$ afin d'avoir les valeurs critique, K et N .

Les valeurs de M , K et N sont présentées dans les tableaux ci-dessous.

Pour une comparaison des résultats nous balayons tous les domaines possibles des différentes paramètres du système ainsi que leur influences mutuelles :

Différentes valeurs de ρ pour étudier l'évolution des systèmes dans le régime chargé et peu chargé.

Différentes valeurs de μ afin de faire ressortir une possible influence du taux de rappel.

A partir du tableau 4.4 on constate que :

• **Au niveau élevé de la congestion du système ; qui est indiqué par les grandes valeurs de ρ et petites valeurs de μ , les niveaux de troncature M et K sont suffisamment grands, autrement dit, ces valeurs augmentent avec l'augmentation de ρ et la diminution de μ**

• Pour des valeurs assez élevées de μ et petites valeurs de ρ , on constate qu'on a des valeurs raisonnables de M . et dans ce cas; la première approche devient pratique pour calculer le vecteur de probabilité stationnaire. En outre, avec l'augmentation du nombre des équations du système, le nombre d'itérations exigées pour obtenir la convergence désirée augmente très rapidement .

• Les valeurs de K sont sensiblement plus petites que M , ce qui nous permet de conclure que le modèle $X^{(F)}$ est le plus consistant.

• L'influence notable de l'intensité de trafic et du taux de rappel μ sur le comportement du système et le choix des approximations les plus performantes n'altère en rien le fait que le modèle $X^{(NR)}$ reste le plus consistant avec plus de précision dans le régime chargé (ρ proche de 1).

	$\mu = 0.1$	$\mu = 0.5$	$\mu = 20$
$\rho = 0.3$	10	10	10
	10	10	10
	10	10	10
$\rho = 0.5$	18	10	10
	14	10	10
	12	10	10
$\rho = 0.8$	100	40	20
	66	25	15
	45	22	12

Tableau 4.6 : Valeurs critiques M , K et N

3.4.6. Conclusion

Aux niveaux élevés de la congestion du système, une approche directe demande trop de calculs, car on a $M.(c+1)$ équations à résoudre pour calculer le vecteur de la distribution stationnaire. Par conséquent, la première approximation est considérée critique pour des systèmes de congestion extrêmes, où la convergence des méthodes itératives est extrêmement lente. Ce qui nous exige un temps d'exécution très long.

Pour surmonter cet inconvénient, nous développons une solution algorithmique basée sur la méthode matricielle géométrique en utilisant le modèle $X^{(F)}$ et $X^{(NR)}$.

Il est clair que l'intérêt de la dernière règle de rappel est lié à une meilleure utilisation des ressources du système.

CONCLUSION GENERALE

Dans cette thèse, nous avons présenté l'investigation numérique du modèle d'attente M/M/c avec rappel opérant sous la politique de rappel classique. Notre étude s'est focalisée sur le calcul de la distribution stationnaire du système et des différentes mesures de performance associées.

Soulignons que pour ce type de système de file d'attente il est difficile d'utiliser une analyse directe conduisant à la solution exacte. Pour arriver à une solution algorithmique, nous avons exploité plusieurs stratégies d'approximation des différents caractéristiques du système.

En premier lieu, il est montré que le modèle de file d'attente considéré peut être approché par un système fini en faisant une troncature de l'orbite à un niveau M. Cependant cette troncature implique le choix d'une grande valeur de M, quand le niveau de congestion du système est élevé. Cet inconvénient peut être évité en utilisant des méthodes de troncature plus sophistiquées où en imposant une hypothèse simplifiée qui donne un modèle de file d'attente auxiliaire avec un état du système infini et un générateur infinitésimal plus approprié.

Il est montré que les modèles présentés peuvent être décrits par des processus quasi de naissance et de mort à plusieurs états limites (QMN). L'avantage de travailler avec les QMN est que l'on peut présenter les traits de base de la file M/M/c en pensant en termes d'états blocs et sous matrices de transition.

Ce travail constitue un pas important dans l'évaluation des systèmes à plusieurs serveurs présentant des arrivées, des rappels et des services exponentiels, en montrant l'efficacité de la méthode matricielle géométrique appliquée à ce genre de système. L'étude est significative due à l'intérêt du système lui même et son analyse spécifique.

Annexe : Rappels de probabilités

1. Formule de Little

Considérons un système général, caractérisé par :

- le nombre moyen d'arrivée par unité de temps : λ
- le nombre moyen de clients présents dans le système : \overline{N}_S
- le temps moyen d'attente d'un clients dans le système : \overline{W}_S

Le résultat général établi par Little stipule que les variables λ , \overline{N}_S et \overline{W}_S sont liées par la relation :

$$\overline{N}_S = \lambda \overline{W}_S$$

cette formule est intéressante pour différentes raisons :

- La théorie des file d'attente donne assez facilement des résultats sur les probabilités d'états, desquelles on déduit le nombre moyen de clients dans un système. La formule de Little permet de calculer le temps de réponse (ou le temps d'attente) d'un système à partir de l'évaluation du nombre moyen.
- Il est extrêmement difficile de mesurer sur un système réel le temps d'attente. par contre les données sur le nombre moyen d'arrivées par unité de temps d'une part et sur le nombre moyen de clients présents s'obtiennent facilement (par des mesures de comptage ou par échantillonnage). On peut donc évaluer le temps d'attente en faisant des mesures sur des variables beaucoup plus faciles à évaluer.
- Le résultat de Little ne dépend pas de la variance du processus d'arrivée mais seulement de sa moyenne. Il ne dépend pas non plus du temps de service d'un client dans le système.
- Cette formule est valable sous les seules conditions de stationnarité et d'existence des moyennes stationnaires λ , \overline{W}_S , \overline{N}_S . En particulier elle ne dépend pas de la discipline de service, pourvu que celle-ci ne modifie pas la durée de service des clients [36].

Référence

- [1]. Alan Ruegg, *Processus stochastiques avec applications aux phénomènes d'attente et de fiabilité*, Ecole polytechnique Fédérale de Lausanne, (1989).
- [2]. J.R. Artalejo, "Stationary analysis of the characteristics of the M/M/2 queue with constant repeated attempts". *Opsearch* 33, 83-95 (1996).
- [3]. J.R. Artalejo, A. Gomez-corrall, "Stochastic analysis of the departure and quasi-input processes in a versatile single server queue". *J.of Appl Mathematics and Stochastic Analysis* 9, 171-183 (1996).
- [4]. J.R. Artalejo, A. Gomez-Corrall, "Steady state solution of a single-server queue with linear repeated requests". *j . of Applied Probability* 34, 223-233 (1997).
- [5]. J.R. Artalejo, A. Gomez-corrall, M.F. Neuts, " Numerical Analysis of multiserver retrial queues operating under a full access policy". *Operations research*, 1-17 (1998)
- [6]. J.R. Artalejo, A. Gomez-corrall, " Analysis of a stochastic clearing system with repeated attempts". *Commun. statist- stochastic models* 14, 623-645 (1998)
- [7]. J.R. Artalejo, A. Gomez-corrall, M.F. Neuts, "Analysis of multiserver queues with constant retrial rate". *European J. of Operational research* 135, 569- 581 (2001).
- [8]. J.R. Artalejo, M.Pozo, "Numerical calculation of the stationary distribution of the main multiserver retrial queue". *Annals of Operations Research* 111, 1-18 (2002).
- [9]. B. Baynat, " Théorie des files d'attente". Hermes, Paris, Collection réseaux et télécommunication, (2000).
- [10]. S.R. Cakravarthy, A.Krishnamoorthy, V.C.Joshua, "Analysis of multiserver retrial queue with search of customers from the orbit". *Research Award* 30-55 (1998).
- [11]. S.Chakravarthy, Alexander. Dudin, " A multiserver retrial queue with bmap arrivals and group services". *Queueing systems* 42, 5-31 (2002).
- [12]. B.D. Choi, K.K. park, C.E.M. pearce, "An M/M/1 retrial queue with control policy and general retrial times", *Queueing Systems* 14, 275- 292 (1993).
- [13]. B.D. Choi, Y.Chang, B.Kim, "MAP/M/c retrial queue with gard channels and its application to cellular networks". *Sociedad de Estadistica* 7, 231-248 (1999).
- [14]. B.D. Choi, Y.Chang, "MAP/M/c retrial queue with the retrial group of finite capacity and geometric loss". *Mathematical and computer modelling* 30, 99-113 (1999).
- [15]. B.D. Choi, Y.Chang, "Single server retrial queue with priority calls". *Mathematical and computer modelling* 30, 7-32 (1999).

- [16]. B.D. Choi, K.B. Choi, Y.W.Lee “M/G/1 retrial queueing system with two types of calls and finite capacity”. Queueing system 19, 215-229 (1995).
- [17]. J.W.Cohen, “Basic problems of telepgone traffic theory and the influence of repeated calls”. Philips Telecom Rev.18, 49- 100 (1957).
- [18]. T. Engset, “The probability theory of computing the nombre of switching equipments in automatic exchanges”. E.T.Z.31, 304 –305 (1918)
- [19]. A. K. Erlang, « Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges ».Post Office Electrical Engineer’s Journal 10, 189-197 (1917).
- [20]. J.I. Falin, “On the accuracy of a numerical method of calculation of characteristics with repeated calls”. Elektrosvyas, No.8, 35-36(1983).
- [21]. J.I. Falin, “On sufficient condition for ergodicity of multichannel queueing system with repeated calls”. Advanced in App.Prob 16, 474-448 (1984).
- [22]. J.I. Falin, “Double –channel queueing system with repeated calls”, All-union institute for scientific and technical information, (1984).
- [23]. J.I. Falin, “On ergodicity of multilinear queueing systems with repeated calls”. Soviet j of computer and system sciences, 25,No.1, 60-65 (1987)
- [24]. J.I. Falin, “Comparability of migration processes”. Theory of prob and its app, 33,No2,, 370-372 (1988).
- [25]. J.I. Falin, “A survey of retrial queue”. Queueing systems 7,No2, 127-167 (1990).
- [26]. J.I. Falin, j.R. Artalejo, “Approximations for multiserver queues with balking/retrial discipline”, OR Spektrum 17, 239- 244 (1995).
- [27]. J.I. Falin, J.G.C.Templeton, Retrial queues, Chapman & Hall, London, 1997.
- [28]. J.I. Falin, “ A multiserver retrial queue with a finite number of sources of primary calls”, Mathematical and computer modelling 30, 33-49 (1999).
- [29]. G.Fayolle, A simple telephone exchange with delayed feedbacks, in: Teletraffic Analysis and Computer performance evaluation, Elsevier, 1986.
- [30]. W. Feller, “ An introduction to probability theory and its applications”. John Wiley and Sons, 1971.
- [31]. D.Gross et C.Harris, Fundamentals of Queueing theory, Jon Wiley et Sons, New york, 1974.

- [32]. E.Gelenbe et G.Pujolle, **Introduction aux réseaux de files d'attente**, Eyrolles, Paris,1985.
- [33]. A.Gomez-Corral, « **queueing system with repeated attempts and negative arrivals**». Thesis, departement of statistics and operations research, computense univ of Madrid (1996).
- [34]. T. Hanschke, “**Explicit formulas for the characteristics of the M/M/2/2 queue with repeated attempts**”, *J.Appl. prob* 24, 486- 494 (1987).
- [35]. J. Kelson, J. Cozzolino, H. Young, **A service system with unfilled requests repeated**” 28 (1968).
- [36]. L. Kleinrock, “**Queueing systems**”. Computer application 2, Wiley , New york (1976).
- [37]. A.G. Kok, “**computational methods for single server systems with repeated attempts**”. Interfaculteit der actuariale wetenschappen en econometrie, Report 89 (1982).
- [38]. A.G. Kok, “**algorithmic methods for single server systems with repeated attempts**”. *Statistica neerlandica*, 38, No.1, 23-32 (1984).
- [39]. L. Kosten, “ **On the influence of repeated calls in the theory of probabilities of blocking**” . *De Ingenieur*. 59. P01(1974)
- [40]. E. D. Lazowska, J. Zahaorjan, G.S.Graham et K.C.Sevcik, “**Quantitative system performance: computer system analysis using queueing network models**”. Prentice Hall, Englewood cliffs, N.J. (1984).
- [41]. M.F. Neuts, “**Matrix-geometric Solutions in stochastic models**” – An Algorithmic Approach, Johns Hopkins University press, Baltimore, MD, 1981.
- [42]. M.F. Neuts, B. M.Rao “**Numerical investigation of a multiserver retrial model**”. *Queueing systems* 7, 169-190 (1990).
- [43]. C.E.M. Pearce, “**Extended continued fraction, recurrence relations and two-dimensional Markov processes**”. *Advances in applied probability* 21, 357-375 (1989).
- [44]. J.Pellaumail, P.Boyer et Leguesdron, « **Reseaux ATMet psimulation** ». Hermes paris(1994).
- [45]. C.H. Saver et K.M.Chandy, “**Computer systems performance modeling**. Prentice Hall, Englewood cliffs, N.J. (1981).
- [46]. R. I. Wilkinson, “**Theories for toll traffic engineering in the U.S.A**”.*Bell system technical. J.*35, 421 – 514 (1956).