

## THESE

Présentée en vue de l'obtention du diplôme de

**Doctorat en Sciences**

**Spécialité : Recherche Opérationnelle**

**Sur Les Modèles avec Rappels et Recherche des Clients en Orbite.**

Par:

*Mme BOUSSAHA Zina*

Sous la direction de **Prof. OUKID Nadia,**

U.S.D.Blida 1

Et **Prof. ZEGHDOUDI Halim,**

U.B.M. Annaba

Devant le jury

<b>Président</b>	<b>BENMOSTEFA Fatima Zohra</b>	<b>Prof.</b>	<b>U.B.M. ANNABA</b>
<b>Examineur</b>	<b>KERBOUA Mourad</b>	<b>MCA</b>	<b>Univ. 8 Mai 45 Guelma</b>
<b>Examineur</b>	<b>EZZEBSA Abdelali</b>	<b>MCA</b>	<b>Univ. 8 Mai 45 Guelma</b>

**Année : 2023.**

# Dédicace

*Cette thèse est dédiée à toute ma famille...  
et mes meilleures amies...*

# Remerciements

Tout d'abord, je remercie Allah le tout puissant de m'avoir donné le courage et la patience nécessaires à mener ce travail à son terme.

Je tiens d'exprimer toute ma reconnaissance à ma directrice de thèse **Prof. Oukid Nadia**, de m'avoir encadré, orienté, aidé et conseillé, et ainsi pour la confiance et la liberté qu'elle m'a accordé, sans oublier ses efforts fournis pendant ces longues années.

Je tiens à remercier également mon co-directeur de thèse **Prof. Zeghdoudi Halim** pour son aide immense, la qualité de son suivi ainsi que pour tous les conseils et les informations qu'il m'a prodigués avec un degré de patience et de professionnalisme sans égal.

Je tiens aussi à adresser mes plus sincères remerciements à mon enseignante **Prof. Djellab Nathalia**, qui ma initié aux valeurs authentiques, en signe d'un profond respect et d'un profond amour.

Que les membres de jury trouvent, ici, l'expression de mes sincères remerciements pour l'honneur qu'ils me font en prenant le temps de lire et d'évaluer ce travail.

Ce travail n'aurait pu arriver à terme sans le soutien, la prière et l'encouragement familial sur tout mon mari et mes chères parents que dieu les garde pour moi en bonne santé.

Pour finir, je souhaite remercier toute personne ayant contribué de près ou de loin à la réalisation de ce travail.

# Résumé

Dans cette thèse, nous nous sommes intéressés à l'étude des modèles de files d'attente avec rappels et recherche des clients en orbite. Dans un premier temps, nous avons effectué une synthèse sur les systèmes de files d'attente avec rappels de type  $M/M/1$ ,  $M/G/1$  et  $M/G/1$  avec rappels et feedback.

Dans un deuxième temps, nous avons présenté quelques systèmes de files d'attentes avec rappels et recherche des clients en orbite (temps de recherche négligeable) et nous nous sommes intéressés à cet effet aux mesures de performances de ces modèles. L'étude a deux objectifs principaux : Le premier consiste à introduire la recherche des clients en orbite dans un modèle de file d'attente avec rappels, ce qui permet de réduire le temps d'inactivité du serveur. Le deuxième objectif est de donner un aperçu du lien entre la file d'attente avec rappels correspondante et la file d'attente classique (sans rappels).

Nous nous sommes intéressés par la suite à l'étude du modèle de file d'attente  $M/M/1$  avec rappels et recherche des clients en orbite (temps de recherche non négligeable), ou l'analyse de sensibilité des mesures de performance par rapport à leurs paramètres critiques est discutée.

Enfin, nous avons généralisé l'étude en considérant le modèle  $M/G/1$  avec rappels, Bernoulli feedback et recherche des clients en orbite (le temps de recherche suit une loi exponentielle). Les fonctions génératrices partielles du nombre de clients dans le système lorsque le serveur est libre, occupé et à la recherche des clients en orbite sont trouvées. Plusieurs résultats numériques sont présentés afin de montrer l'impact de certains paramètres du système sur les probabilités de l'états du système.

Ce travail ouvre quelques perspectives de recherche telles que :

- Analyse des systèmes de files d'attente avec rappels et recherche des clients en orbite plus complexe. Modéliser des systèmes concrets avec ce type de modèle.
- Extension du modèle étudié avec arrivée des clients floue.

---

**Mots clés :** File d'attente avec rappels, Orbite, Chaine de Markov induite, Condition d'érgodicité, Feedback, Distribution stationnaire, Mesures de performance.

---

# Abstract

In this thesis, we are interested in the study of queueing models with retrial and search for customers in orbit. First, we made a synthesis on queueing systems of type  $M/M/1$ ,  $M/G/1$  and  $M/G/1$  with retrial and feedback.

Secondly, we presented some queueing systems with retrial and search for customers in orbit (the search time is negligible) and we are interested for this purpose in the performance measures of these models. The study has two main objectives : The first is to introduce the search for customers in a queueing model with retrials, which helps reduce server idle time. The second objective is to give an overview of the link between the retrial queue and the classic queue (without retrial).

We then focused on the study of the  $M/M/1$  queueing model with retrial and search for customers in orbit (significant search time), or sensitivity analysis of performance measures in relation to their critical parameters is discuss.

Finally, we generalized the study by considering the  $M/G/1$  model with retrial, Bernoulli feedback and search for customers in orbit (the search time follows an exponential law). The partial generating functions of the number of customers in the system when the sever is free, busy and looking for customers in orbit are found. Several numerical results are presented in order to show the impact of certain parameters on the probabilities of the system states.

This work opens some research perspectives such as :

- Analysis of queueing systems with retrial and orbital search of customers in more complex. Model concrete systems with this type of model.

- Extension of the model studied with fuzzy customers arrivals.

---

**Keywords :**Retrial queue, Orbit, Embedded Markov chain , Ergodicity condition, Feedback, Steady-state distribution, Performance measure. .

---

## ملخص

في هذه الرسالة ، نهتم بدراسة نماذج الإنتظار مع عمليات الإسترجاعات والبحث عن عملاء في المدار. أولاً ، قمنا بعمل توليف حول أنظمة الطابور باستخدام عمليات الاسترجاع  $M/M/1$  ،  $M/G/1$  و  $M/G/1$  مع عمليات الاسترجاع و التغذيةيات الرجعية.

في الخطوة الثانية ، قدمنا بعض أنظمة الطابور مع عمليات الاسترجاع والبحث عن عملاء في المدار (وقت بحث منعدم) وكنا مهتمين لهذا الغرض بقياسات أداء هذه النماذج. الدراسة لها هدفان رئيسيان : الأ هو تقديم البحث عن العملاء في المدار في نموذج الطابور مع عمليات الاسترجاعات ، مما يسمح بتقليل وقت الحمول للخادم. الهدف الثاني هو إعطاء نظرة عامة على الارتباط بين قائمة الانتظار مع عمليات الاسترجاع المقابلة وقائمة الانتظار الكلاسيكية (بدون عمليات الاسترجاع).

ركزنا بعد ذلك على دراسة نموذج قائمة الانتظار  $M/M/1$  مع عمليات الاسترجاعات والبحث عن العملاء في المدار (وقت بحث معتبر) ، حيث ناقشنا و حللنا حساسية الأداء مقابل معاييرهم الحرجة.

أخيراً ، قمنا بتعميم الدراسة على دراية بنموذج  $M/G/1$  مع تذكيرات ، و تغذية راجعة من نوع Bernoulli والبحث عن العملاء في المدار (يتبع وقت البحث قانوناً أسياً). يتم العثور على وظائف التوليد الجزئي لعدد العملاء في النظام عندما يكون شاغرا ، ومشغولا ويبحث عن عملاء في المدار. يتم عرض العديد من النتائج الرقمية من أجل إظهار تأثير بعض المعاملات الخاصة على احتمالات حالات النظام.

يفتح هذا العمل بعض وجهات النظر البحثية مثل :

- تحليل أنظمة الطابور مع عمليات الاسترجاع وإيجاد عملاء في مدار أكثر تعقيداً. تجسيد نموذج أنظمة ملموسة مع هذا النوع من النماذج.
- تمديد النموذج المدروس مع وصول العملاء الغامضين.

---

**كلمات مفتاحية :** نظم طوابير الإنتظار التي لها خاصية إعادة التنبية بطلب الخدمة، المدار ، سلسلة ماركوف المضمنة، شرط احتمال حدوث الخدمة مرة أخرى، التغذيةية الرجعية، لتوزيع الثابت لحالة النظام، مقاييس الأداء.

# Table des matières

<b>Dédicace</b>	I
<b>Remerciements</b>	II
<b>Résumé</b>	III
<b>Abstract</b>	IV
<b>Introduction Générale</b>	1
<b>1 Généralités sur les modèles de files d'attente avec rappels</b>	4
1.1 Introduction	4
1.2 Modèle de file d'attente avec rappels	5
1.2.1 Description d'un modèle d'attente avec rappels	5
1.2.2 Les objectifs de la théorie des files d'attente	5
1.2.3 Domaines d'application	6
1.2.4 Méthodes d'analyse des modèles d'attente	7
1.3 Analyse de quelques modèles de file d'attente avec rappels	7
1.3.1 Modèle de file d'attente $M/M/1$ avec rappels	7
1.3.2 Modèle de file d'attente $M/G/1$ avec rappels	9
1.3.3 Modèle de file d'attente $M/G/1$ avec rappels et feedback	15
1.4 Conclusion	21
<b>2 Modèles de files d'attente avec rappels et recherche des clients en orbite.</b>	22
2.1 Introduction	22
2.2 Modèles d'attente avec rappels linéaire et recherche des clients en orbite	23
2.2.1 Modèle d'attente $M/M/1$ avec rappels linéaire et recherche des clients	23
2.2.2 Modèle d'attente $M/G/1$ avec rappels et recherche des clients	28
2.3 Conclusion	34
<b>3 Modèle d'attente <math>M/M/1</math> avec rappels et temps de recherche des clients exponentiel</b>	35
3.1 Introduction	35
3.2 Description du modèle	35
3.3 Etude analytique du système	36

3.4	Formules récursives	39
3.5	Cas limite	40
3.6	Analyse de sensibilité des mesures de performance du modèle	41
3.6.1	L'effet du taux d'arrivée $\lambda$ sur le modèle	41
3.6.2	L'effet du taux de recherche $\nu$ sur le modèle	43
3.6.3	L'effet du taux d'inactivité $\alpha$ sur le modèle	44
3.6.4	Effet de variations de $\nu$ et $\alpha$ sur le nombre de clients en orbite	46
3.6.5	Effet de variations de $\nu$ et $\alpha$ sur le nombre de clients dans le système	46
3.6.6	Effet de variations de $\nu$ et $\alpha$ sur le temps moyen de séjour des clients dans le système	47
3.7	Conclusion	48
<b>4</b>	<b>Modèle d'attente <math>M/G/1</math> avec rappels, feedback et recherche des clients en orbite</b>	<b>49</b>
4.1	Introduction	49
4.2	Description mathématique du modèle	49
4.3	La condition de stabilité	50
4.4	Etude de l'état stationnaire	51
4.5	Résultats numériques	56
4.5.1	Effet de $c$ sur les probabilités de l'état du serveur	56
4.5.2	Effet de $\mu$ sur les probabilités de l'état du serveur	57
4.5.3	Effet de $\alpha$ sur les probabilités de l'état du serveur	58
4.5.4	Effet de $\lambda$ sur les probabilités de l'état du serveur	59
4.6	Conclusion	60
	<b>Conclusion Générale</b>	<b>61</b>
	<b>Bibliographie</b>	<b>62</b>



# Table des figures

1.1	Schéma illustratif d'un système de file d'attente avec rappels.	5
1.2	Schéma illustratif d'un système de file d'attente avec rappels et feedback.	16
2.1	Représentation schématique d'un système de files d'attente avec rappels et recherche de clients en orbite.	23
2.2	Graphe des transitions du modèle $M/M/1$ avec rappels et recherche de clients en orbite.	24
2.3	Graphe des transitions du modèle $M/G/1$ avec rappels et recherche de clients en orbite.	30
3.1	Graphe des transitions du modèle.	36
3.2	Effet de $\lambda$ sur les probabilités de l'état du serveur.	42
3.3	Effet de $\lambda$ sur le nombre moyen de clients dans le système et dans l'orbite.	42
3.4	Effet de $\lambda$ sur le temps moyen de séjour des clients dans le système et dans l'orbite.	43
3.5	Effet de $\nu$ sur les probabilités de l'état du serveur.	44
3.6	Effet de $\nu$ sur le nombre moyen de clients dans le système et dans l'orbite.	44
3.7	Effet de $\nu$ sur le temps moyen de séjour des clients dans le système et dans l'orbite.	44
3.8	Effet de $\alpha$ sur les probabilités de l'état du serveur.	45
3.9	Effet de $\alpha$ sur le nombre moyen de clients dans le système et dans l'orbite.	45
3.10	Effet de $\alpha$ sur le temps moyen de séjour des clients dans le système et dans l'orbite.	46
3.11	Variation de nombre moyen de clients dans l'orbite en fonction de $\alpha$ et $\nu$ .	46
3.12	Variation de nombre moyen de clients dans le système en fonction de $\alpha$ et $\nu$ .	47
3.13	Variation de temps moyen de séjour des clients dans le système en fonction de $\alpha$ et $\nu$ .	47
4.1	Graphe des transitions du modèle $M/G/1$ avec rappels, feedback et recherche des clients en orbite.	51
4.2	Effet de $c$ sur les probabilités de l'état du serveur.	57
4.3	Effet de $\mu$ sur les probabilités de l'état du serveur.	58
4.4	Effet de $\alpha$ sur les probabilités de l'état du serveur.	59
4.5	Effet de $\lambda$ sur les probabilités de l'état du serveur.	59

# Liste des tableaux

3.1	Variations des mesure de performances dans le système en fonction de $\lambda$ .	. . . . .	42
3.2	Variations des mesures de performances dans le système en fonction de $\nu$ .	. . . . .	43
3.3	Variations des mesures de performances dans le système en fonction de $\alpha$ .	. . . . .	45
4.1	Effet de $c$ sur les probabilités de l'état du serveur.	. . . . .	57
4.2	Effet de $\mu$ sur les probabilités de l'état du serveur.	. . . . .	57
4.3	Effet de $\alpha$ sur les probabilités de l'état du serveur.	. . . . .	58
4.4	Effet de $\lambda$ sur les probabilités de l'état du serveur.	. . . . .	59

# Introduction Générale

LA théorie des files d'attente est une théorie mathématique relevant du domaine des probabilités, né en 1917, des travaux de l'ingénieur danois Erlang sur la gestion des réseaux téléphoniques de Copenhague. Entre 1909 et 1920, elle a étudié notamment les systèmes d'arrivée dans une queue, les différentes priorités de chaque nouvel arrivant, ainsi que la modélisation statistique des temps d'exécution. L'application de cette théorie n'a été généralisée à divers types de problèmes qu'après la seconde guerre mondiale. C'est grâce aux apports des mathématiciens Khintchine, Palm, Kendall, Pollaczek et Kolmogorov que la théorie s'est vraiment développée. La théorie de files d'attente est aujourd'hui largement utilisée et les applications de ces travaux sont alors très pratiques et concernent les disciplines de recherche opérationnelle et génie industriel.

Cette théorie classique s'est très vite montrée inefficace face à des systèmes réels de plus en plus complexes. Dès la fin des années 1940, des chercheurs tels que Kosten (1947) et Wilkinson (1956) ont mis en évidence les limites de la théorie classique des files d'attente qui ne permettait pas d'expliquer le comportement stochastique des systèmes téléphoniques où les abonnés répétaient leurs appels en recomposant le numéro plusieurs fois jusqu'à l'obtention de la communication. Ce phénomène de répétition de demandes du service a poussé certains chercheurs à étendre le modèle d'attente classique à celui dit avec *rappels* (Cohen, 1957). Cependant, l'influence de ce phénomène a été longtemps négligée durant les décennies suivantes. Ce n'est que vers les années 1970-1980 qu'on a vu un net regain d'intérêt pour cette catégorie de modèles, avec l'avènement de nouvelles technologies, notamment dans les systèmes de télécommunication : réseaux ATM. Les systèmes de files d'attente avec rappels peuvent être appliqués pour résoudre les problèmes pratiques, tels que l'analyse du comportement des abonnés dans les réseaux téléphoniques, l'analyse du temps d'attente pour accéder à la mémoire sur les disques magnétiques,... (TranGia, Mandjes, 1997). Ce type de modèles se rencontre également dans la modélisation de protocoles spécifiques de communication, tels que CSMA (Carrier Sense Multiple Access) ou encore les disciplines Auto-Repeat, Ring-Back-When-Free, Repeat-Last-Number,... (Khomichkov, 1995). Ainsi l'analyse mathématique de modèles de files d'attente pour la modélisation et l'évaluation des performances dans le Cloud Data Centers qui permettent d'offrir des services de qualité plus accessibles pour les fournisseurs de service Cloud et qui répondent aux caractéristiques dynamiques de l'environnement du Cloud Computing [25], [3], [65], ... .

La majorité des études sur les systèmes d'attente avec rappels considère le modèle sans feedback. Néanmoins, plusieurs situations réelles peuvent être modélisées comme des systèmes de files d'attente avec rappels et feedback. Dans les télécommunications, les transmissions des don-

nées de protocole sont parfois répétées. Ceci arrive fréquemment à cause de la médiocrité du service. En industrie (reproduction du produit mal composé) est un exemple de files d'attente avec feedback. Les modèles de files d'attente avec feedback ont été largement étudiés par un grand nombre de chercheurs. La notion de feedback a été initialement introduite par Takas (1963) [76] pour l'étude de certains systèmes d'attente classiques, et depuis plusieurs papiers sont apparus sur ce sujet en considérant d'autres types de systèmes avec différentes variantes (rappels, vacances, pannes,...) (Krishna Kumar et al., 2010). Le phénomène de feedback dans les systèmes d'attente avec rappels peut apparaître dans plusieurs situations pratiques (Artalejo, Gómez-Corral, 2008) [8], par exemple, dans les systèmes MATS (Multiple Access Telecommunication Systems) où des messages s'avérant comme erreurs à la destination (messages perdus ou corrompus) sont renvoyés. En particulier, les systèmes d'attente avec rappels et feedback sont utilisés pour modéliser le protocole ARQ (Automatic Repeat Request) dans un réseau de communication à haute fréquence. Le modèle  $M/G/1$  avec rappels et feedback a été étudié dans Djellab (2005) [32], l'auteur a obtenu la distribution du nombre de clients dans le système à l'état stationnaire, en utilisant la méthode de la chaîne de Markov auxiliaire. Pour le même modèle, Boualem et al. (2012) [17] ont utilisé la méthode de comparaison stochastique des chaînes de Markov pour obtenir quelques approximations qualitatives. De plus, les auteurs ont obtenu la propriété de la décomposition stochastique du nombre de clients dans le système au régime stationnaire. Dans [16], les auteurs ont dérivé certaines mesures de performance, importantes et intéressantes, pour le même système considéré.

La plupart des modèles avec rappels étudié dans la littérature supposent qu'après chaque service, le serveur restera inactif dans le système jusqu'à l'arrivée du prochain client primaire ou un client de l'orbite. Même s'il y a des clients dans le système qui veulent obtenir un service, ils ne peuvent pas occuper le serveur immédiatement, en raison de leur ignorance de l'état du serveur. Par conséquent, après l'achèvement de chaque service, le prochain client entre en service uniquement après un certain intervalle de temps pendant lequel le serveur est libre alors qu'il peut y avoir des clients en attente dans l'orbite. Mais dans la vraie vie, nous voulons toujours minimiser le temps d'inactivité du serveur et minimiser les coûts de conservation. Donc, il est nécessaire d'étudier les systèmes de file d'attente avec recherche des clients. La recherche de clients immédiatement à la fin d'un service était d'abord discuté dans le cas des files d'attente classiques par Neuts et Ramalhoto (1984) [62]. Un système de file d'attente à serveur unique, deux flux indépendants de clients Poissonniens, et une fonction de distribution générale pour le temps de service est étudié dans [13], où la recherche de différents types de clients est réalisée avec une priorité non préemptive. Un modèle d'attente avec rappels et clients qui cherchent un serveur, avec arrivées Poissonniennes et capacité finie de la file a été analysé dans [15], Les résultats obtenus sont généralisés au cas des arrivées qui suivent une loi  $k$  Poisson dans [14].

Artalejo et al. [4] ont envisagé un système de files d'attente avec rappels dans lequel, immédiatement après l'achèvement d'un service, le serveur recherche un client en orbite selon une certaine probabilité ou reste inactif. Si une recherche est faite, un service est suivi d'un autre service, sinon, il sera suivi d'un temps d'inactivité. Cependant, dans Artalejo et al. [4] il n'y a pas de temps libre ni de temps de recherche (le temps de recherche est égal à zéro). Dudin et al. [35] considèrent le même modèle que dans [4] avec une entrée  $BMAP$  et rechercher des clients en orbite. Cependant, le mécanisme de recherche est lancé juste après l'achèvement du service. Le mo-

dèle  $MAP/M/c$  avec rappels et recherche des clients en orbite à été étudié dans [23]. L'analyse en régime permanent du modèle à été effectuée à l'aide de la troncature directe, approximation géométrique matricielle et algorithme efficace pour le calcul des performances ont été présentés. Quelques autres extensions se trouvent dans ([77], [49], [78], [79], [52], [24], [73], [75], [1], [40], [70], [26], [74], [9], [68], [71], [50], [11], [51], [12], [58], [64], [61], [41], [21], [69], [28], [53], ...). Sous divers types de modèles étudiés dans ces travaux, ils se sont unis en considérant le temps de la recherche des clients en orbite négligeable.

La forme géométrique matricielle des probabilités d'état stationnaire à été dérivée dans [37] pour le système de file d'attente  $PH/PH/1/r$ , ils ont supposé que le processus de recherche et celui du service sont de type phase d'ordre  $m$ . Dans D'Apice et al.(2004) [36] les auteurs ont considéré un système de file d'attente  $M/G/1/r$  avec rappels, serveur qui recherche des clients en orbite selon une loi exponentiel, et service prioritaire pour les clients primaires. Dudin et al. [35] considèrent le modèle  $BMAP/G/1$  avec rappels et durée de recherche des clients aléatoire et aussi probablement en fonction du nombre de clients dans l'orbite, ils ont effectué l'analyse de l'état stationnaire du modèle. Jusqu'à présent, peu de travaux sur les modèles de files d'attente avec rappels et recherche orbitale de clients ( temps de recherche non négligeable) ont été effectués, les lecteurs sont référés dans ([31], [30], [66], [40], [34], [59], [48], [60], [46], ...).

Cette thèse est organisée en quatre chapitres. Dans le premier chapitre, nous présentons quelques rappels des résultats sur les systèmes de files d'attente avec rappels de type  $M/M/1$ ,  $M/G/1$ , et  $M/G/1$  avec rappels et feedback qui permettra d'introduire les chapitres suivants où seront abordés de nouveaux résultats.

La description du phénomène de recherche des clients en orbite de service dans les systèmes d'attente avec rappel fait l'objet du second chapitre. Nous considérons les modèles avec le temps de recherche des clients en orbite négligeable.

Le troisième chapitre contient l'analyse du système de file d'attente  $M/M/1$  avec rappels et recherche des clients en orbite, le temps de recherche est non négligeable et on a étudié la sensibilité des mesures de performance par rapport à leurs paramètres critique.

Le chapitre quatre est concerné par nos investigations sur le modèle d'attente  $M/G/1$  avec rappels, feedback et recherche des clients en orbite.

Les résultats issues des deux derniers chapitres ont fait l'objet d'une publication [22], et de plusieurs communications et publications dans des conférences et proceedings internationaux ([21], [20], [19], et [18]).

# Chapitre 1

## Généralités sur les modèles de files d'attente avec rappels

### 1.1 Introduction

Les systèmes de files d'attente avec rappels sont des systèmes utilisés dans la modélisation des réseaux de télécommunication et dans les systèmes informatiques. Après son arrivée à une station donnée, un client qui trouve tous les serveurs occupés et ne peut pas recevoir le service immédiatement, quitte le système pour être rappelé ultérieurement à des instants aléatoires jusqu'à satisfaction de sa demande. C'est le cas pour les appels téléphoniques par exemple, entre deux rappels successifs, le client en question se trouve en orbite. Un tel système est appelé système de files d'attente avec rappel (retour). Ce phénomène de files d'attente avec rappel, a fait inspirer plusieurs chercheurs. Des monographies et des études bien connus sur les files d'attentes avec rappels incluent Falin et Templeton [39] et Artalejo et Gomez [8]. Une classification bibliographique est donnée dans les articles de Aissani [2], Artalejo [5], Kim et al. [47], Ghimire et al. [44], Phung-Duc [67], Sasikala et Indhira [72].

Les systèmes de files d'attente avec rappels sont identifiés par la nature stochastique du processus des arrivées, la distribution du temps de service, le nombre de serveurs qui composent l'espace de service, la capacité et la discipline d'attente ainsi que la spécification concernant le processus de répétition d'appels.

Une file d'attente est notée  $A/B/c/m/N/H$  selon Kendall, où  $A$  et  $B$  décrivent respectivement, la distribution du temps inter-arrivées et la distribution du temps de service,  $c$  est le nombre de serveurs identiques et indépendants,  $(m - c)$  est la capacité de la file d'attente,  $m$  est la capacité de l'orbite et  $H$  représente une fonction de persistance  $H = \{H_n, n \geq 1\}$ . Si  $m, N$  et  $H$  sont absents dans la notation de Kendall, alors  $c = m, N = \infty$  et  $H_n = 1$  pour tout  $n \geq 1$ . La distribution du temps inter-rappels n'est pas indiquée.

Dans ce chapitre nous présentons les éléments essentiels de quelques systèmes de files d'attente avec rappels dont l'étude nous sera nécessaire pour la compréhension des prochains chapitres.

## 1.2 Modèle de file d'attente avec rappels

### 1.2.1 Description d'un modèle d'attente avec rappels

En général, un système de files d'attente avec rappels contient un espace de service composé de  $c \geq 1$  dispositifs de service et un espace d'attente de  $m - c$  ( $m \geq c$ ) positions d'attente. Les clients arrivent dans le système selon un processus aléatoire avec une loi de probabilité donnée, et forment un flux d'appels primaires. A l'arrivée d'un client, s'il y a un ou plusieurs serveurs libres, le client sera immédiatement pris en charge. Sinon, s'il y a des positions d'attente libres, le client va rejoindre la file d'attente selon la discipline de service adaptée dans le système. D'autre part, si un client arrive et trouve tous les serveurs et positions d'attentes occupés, il quitte le système définitivement avec une probabilité  $1 - H_0$  ou fait des tentatives de service après une durée de temps aléatoire avec une probabilité  $H_0$ . Les clients qui reviendront et rappelleront pour le service sont dits en "orbite". Cette dernière peut être finie ou infinie. Dans le cas d'une orbite à capacité finie, si elle est pleine, un client qui trouve tous les serveurs et les positions d'attente du buffer occupés, sera obligé de quitter le système définitivement sans être servi. Chaque client en orbite appelé aussi client secondaire (*rappels*), est supposé rappeler pour le service à des intervalles de temps suivant une loi de probabilité et une intensité de rappels bien définie [1.1].

Il existe dans la littérature des modèles avec rappels essentiellement trois politiques de rappels : La politique classique, la politique constante et la politique versatile (une combinaison des deux précédentes). Chacun de ces clients secondaires est traité comme un client primaire, c'est-à-dire un nouveau client qui arrive de l'extérieur du système. S'il trouve un serveur libre, il sera servi immédiatement puis quittera le système. Sinon, s'il y a des positions d'attente disponibles dans le buffer, il le rejoindra. Par contre, si tous les serveurs et les positions d'attente sont encore occupés, le client quittera le système pour toujours avec la probabilité  $1 - H_n$  (si c'est le  $n^{\text{ième}}$  rappel sans succès) ou bien entre en orbite avec la probabilité  $H_n$  si l'orbite n'est pas pleine.

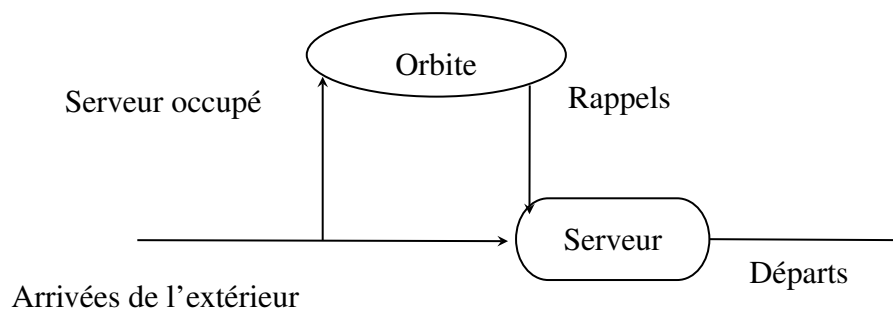


FIGURE 1.1 – Schéma illustratif d'un système de file d'attente avec rappels.

### 1.2.2 Les objectifs de la théorie des files d'attente

L'objectif est de diminuer le temps d'attente du client tout en minimisant les couts de fonctionnement. Il s'agit de réduire l'impact négatif du temps d'attente des clients et de lui attribuer

une perception plus agréable, tout en apportant une valeur ajoutée à l'organisation. La file d'attente avec rappels aide les managers pour comprendre et prendre la meilleure décision concernant les opérations du système impliquant. Se dévoile comme un renfort permettant de gérer la file d'attente avec rappels dans un établissement et garantir la satisfaction de tous les clients pour prendre les meilleures décisions. cette théorie a pour objet l'étude de système où des entités, appelées clients, cherchent à accéder à des ressources, généralement limitées, afin d'en obtenir un service. Cette théorie est mise en oeuvre afin d'optimiser les activités de l'entreprise améliorer la gestion des clients. Elle constitue en fait, un des outils clés dans sa gestion.

Dans le domaine économique, L'objectif de la théorie des files d'attente est de satisfaire aux mieux les clients tout en minimisant ses couts qui se subdivisent en deux types :

- Les couts de service : c'est les couts résultants du fonctionnement des serveurs des matériels, appareils, des employés ... etc, et c'est évident que ces couts augmentent avec l'augmentation du niveau de service.

- Les couts d'attente : ce sont les couts résultants de la diminution de la qualité du service dans le système pour le servir, et elle contient la perte des clients qui quittent la file d'attente pour sa longueur et la perte d'autres clients qui les influencent et leur reportent la mauvaise réputation du système.

### 1.2.3 Domaines d'application

La théorie des files d'attente a fait l'objet de nombreuses recherches depuis l'apparition des premiers réseaux téléphoniques, est resté la seule jusqu'en 1950 ([27], [38]). Mais la tendance a commencé à changer immédiatement après la seconde guerre mondiale et de nombreuses autres applications ont été trouvées. Pendant ce temps, des enquêtes dans une autre branche de la probabilité appliquée, à savoir la "*fiabilité*" a également commencé. L'apparition de la machine qui est un cas particulier de la théorie des files d'attente et de fiabilité ont été développés au même moment. On a également découvert que les modèles de fiabilité du complexe systèmes pourraient être formulés en termes de files d'attente (pannes et réparations). Ces trois domaines de la probabilité appliquée ont beaucoup en commun et peuvent être traités par les mêmes techniques et procédures mathématiques. Dans les années 60, la modélisation des systèmes informatiques et des systèmes de transmission de données s'est ouverte la voie à des études de files d'attente caractérisées par des disciplines de service complexes et ont créé le besoin d'analyser des systèmes interconnectés. Des progrès dans ce domaine ont été rapides et de nombreuses applications industrielles ont été largement acceptées depuis les années 70 ([33], [54]). Les méthodes de file d'attente des réseaux ont toujours été une composante de base de l'étude des systèmes de communication [56]. L'introduction généralisée de l'ordinateur dans ces systèmes a introduit l'utilisation de nouveaux résultats sur la théorie des réseaux dans les études de performance [43]. Autres applications importantes de la théorie des files d'attente ([80], [44], [42], [45], [57], [63]) : les réseaux optiques, le cloud computing, optimisation des stocks (gestion à flux tendu), l'évitement de collision dans les réseaux locaux sous des protocoles CSMA/CD (CarrierSense Multiple Access with Collision Detection), l'analyse du temps d'attente pour accéder à la mémoire sur les disques magnétiques, gestion des avions au décollage ou à l'atterrissage, attente des clients à un guichet, ou bien encore traitement informatique de données par un serveur.



## 1.2.4 Méthodes d'analyse des modèles d'attente

L'une des méthodes utilisées pour résoudre les modèles de file d'attente markoviens consiste à construire une équation différentielle. Habituellement, la meilleure méthode pour faire face à de tels équations est de les convertir, si possible en une seule équation pour une fonction génératrice.

Lorsque l'on s'écarte de l'hypothèse d'exponentialité de l'une des deux quantités stochastiques : le temps des inter-arrivées et la durée de service, ou en prenant en compte certaines spécificités des problèmes par introduction de paramètres supplémentaires, on aboutit à un modèle non markovien. La combinaison de tous ces facteurs rend l'étude mathématique du modèle très délicate, voire impossible on essaye alors de se ramener à un processus de Markov judicieusement choisi à l'aide de l'une des méthodes d'analyse suivantes :

1- *Méthode des étapes d'Erlang* : Son principe est d'approximer toute loi de probabilité ayant une transformation de Laplace rationnelle par une loi de Cox (mélange de lois exponentielles), cette dernière possède la propriété d'absence de mémoire par étape.

2- *Méthode de la chaîne de Markov induite* : Elaborée par Kendall, et souvent utilisée, elle consiste à choisir une suite d'instants  $1, 2, 3, \dots, n$  (déterministes ou aléatoires) tels que la chaîne induite  $\{N_n, n \geq 0\}$ , où  $N_n = N(n)$ , soit markovienne et homogène.

3- *Méthode des variables supplémentaires* : Elle consiste à compléter l'information sur le processus  $\{N(t), t \geq 0\}$  de telle manière à lui donner le caractère markovien. Ainsi, on se ramène à l'étude du processus  $\{N(t), T(t_1), T(t_2), \dots, T(t_n), t \geq 0\}$ , Les variables  $T(t_k), k \in \{0, 1, 2, \dots, n\}$  sont dites supplémentaires.

4- *Méthode des événements fictifs* : Le principe est d'introduire des événements fictifs qui permettent de donner une interprétation probabiliste aux transformées de Laplace et aux variables aléatoires décrivant le système étudié.

5- *Simulation* : C'est un procédé d'imitation artificielle d'un processus réel effectué sur ordinateur. Elle nous permet d'étudier les systèmes les plus complexes, de prévoir leurs comportements et de calculer leurs caractéristiques. Les résultats obtenus ne sont qu'approximatifs, mais peuvent être utilisés avec une bonne précision. Cette technique se base sur la génération de variables aléatoires suivant les lois gouvernant le système.

## 1.3 Analyse de quelques modèles de file d'attente avec rappels

### 1.3.1 Modèle de file d'attente $M/M/1$ avec rappels

#### Description du modèle

On considère un système de files d'attente sans positions d'attente. Le service est assuré par un seul serveur.

- **Processus des arrivées** : Les clients primaires arrivent selon un processus de Poisson de taux  $\lambda > 0$ .
- **Processus de service** : Les durées de service suivent une loi exponentielle de fonction de répartition ;  $B(x) = 1 - e^{-\mu x}, x \geq 0$ , et de moyenne finie  $1/\mu$ .

— **Processus des rappels** : Les temps entre deux rappels consécutifs sont également exponentiels de paramètre  $\theta > 0$ , la fonction de répartition,  $T(x) = 1 - e^{-\theta x}$ ,  $x \geq 0$ .

Nous admettons que les durées de service, les durées entre deux rappels consécutifs ainsi que entre deux arrivées primaires successives sont mutuellement indépendantes. L'état du système peut être décrit par le processus  $\{C(t), N_o(t), t \geq 0\}$ , où  $C(t)$  est égale à 0 ou 1 selon le fait que le serveur est libre ou non,  $N_o(t)$  est le nombre de clients en orbite à l'instant  $t$ .

### Distribution stationnaire de l'état du système

Supposons que le régime stationnaire existe ( $\rho = \frac{\lambda}{\mu} < 1$ ) Le processus est de Markov d'espace d'états  $S = \{0, 1\} \times \mathbb{N}$ . Les équations d'équilibre statistique sont :

$$\begin{aligned}(\lambda + j\theta)p_{0j} &= \mu p_{1j}, \\ (\lambda + \mu)p_{1j} &= \lambda p_{0j} + (j+1)\theta p_{0,j+1} + \lambda p_{1,j-1},\end{aligned}$$

ici :

$$p_{ij} = \lim_{t \rightarrow \infty} P(C(t) = i, N_o(t) = j), \quad i = 0, 1 \text{ et } j \geq 0,$$

les probabilités de transition en régime stationnaire sont définies par :

$$q_{i,j}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (i+1, j), \\ \mu & \text{si } (k, l) = (i-1, j), \\ j\theta & \text{si } (k, l) = (i+1, j-1), \\ -(\lambda + \mu + j\theta) & \text{si } (k, l) = (i, j) \\ 0 & \text{sinon} \end{cases}$$

représentent la distribution stationnaire conjointe de l'état du serveur et du nombre de clients en orbite. Introduisons les fonctions génératrices suivantes :

$$P_0(z) = \sum_{n=0}^{\infty} z^n p_{0n}, \quad P_1(z) = \sum_{n=0}^{\infty} z^n p_{1n}.$$

A l'aide de ses fonctions et à partir des équations, on obtient les fonctions génératrices partielles :

$$\begin{aligned}P_0(z) &= (1 - \rho) \left( \frac{1 - \rho}{1 - z\rho} \right)^{\frac{\lambda}{\theta}}, \\ P_1(z) &= \rho \left( \frac{1 - \rho}{1 - z\rho} \right)^{\frac{\lambda}{\theta} + 1}.\end{aligned}$$

Les transformées inverses nous donnent les formules analytiques explicites :

$$\begin{aligned}p_{0j} &= \frac{\rho^j}{j! \theta^j} \prod_{k=0}^{j-1} (1 - k\theta) (\lambda - \rho)^{\frac{\lambda}{\theta} + 1}, \\ p_{1j} &= \frac{\rho^{j+1}}{j! \theta^j} \prod_{k=1}^j (\lambda + k\theta) (1 - \rho)^{\frac{\lambda}{\theta} + 1}.\end{aligned}$$

Par conséquent, la fonction génératrice de la distribution de la taille de l'orbite sera définie par :

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

Et celle de la distribution stationnaire du nombre de client dans le système s'obtiendra de la manière suivante :

$$Q(z) = P_0(z) + zP_1(z) = \left( \frac{1 - \rho}{1 - z\rho} \right)^{\frac{\lambda}{\theta} + 1}.$$

On peut également trouver la distribution du nombre de serveur occupé :

$$\begin{aligned} p_0 &= \lim_{t \rightarrow \infty} P(C(t) = 0) = P_0(1) = 1 - \rho, \\ p_1 &= \lim_{t \rightarrow \infty} P(C(t) = 1) = P_1(1) = \rho. \end{aligned}$$

### Mesures de performance

Les caractéristiques du modèle sont :

- Nombre moyen de clients dans le système

$$\bar{n} = Q'(1) = \frac{\rho(\lambda + \theta)}{(1 - \frac{\rho}{\theta})}.$$

- Nombre moyen de clients en orbite

$$\bar{n}_0 = P'(1) = \frac{\rho(\lambda + \rho\theta)}{\theta(1 - \rho)}.$$

### 1.3.2 Modèle de file d'attente $M/G/1$ avec rappels

#### Description du modèle

- **Processus des arrivés** : Les clients arrivent dans le système selon un processus de Poisson de taux  $\lambda > 0$  :  $P(\tau_n^e \leq x) = 1 - e^{-\lambda x}$ .
- **Processus de service** : Le service des clients est assuré par un seul serveur. La durée de service  $\tau$  est de loi générale  $P(\tau_n^s \leq x) = B(x)$  et de transformée de Laplace-Stieltjes  $\tilde{B}(s)$ ,  $Re(s) > 0$ . Soient les moments  $\beta_k = (-1)^k \tilde{B}(k)(0)$ , l'intensité du trafic  $\rho = \lambda\beta_1$  et  $\gamma = \frac{1}{\beta_1}$ .
- **Processus des rappels** : La durée entre deux rappels successifs d'une même source secondaire est exponentiellement distribuée de paramètre  $\theta > 0$  :  $T(x) = P(\tau_n^r \leq x) = 1 - e^{-\theta x}$ .

Le système évolue de la manière suivante : On suppose que le  $(n - 1)^{i\text{ème}}$  client termine son service à l'instant  $\xi_{n-1}$  ( les clients sont numérotés dans l'ordre de service) et le serveur devient libre, même s'il y a des clients dans le système, ils ne peuvent pas occuper le serveur immédiatement à cause de leur ignorance de l'état de ce dernier. Donc il existe un intervalle de temps  $R_n$  durant

lequel le serveur reste libre avant que le  $n^{i\grave{e}me}$  client n'entre en service. A l'instant  $\xi_n = \eta_n + R_n$  le  $n^{i\grave{e}me}$  client débute son service durant un temps  $\tau_n^s$ . Les rappels qui arrivent durant ce temps de service n'influent pas sur ce processus. A l'instant  $\xi_n = \eta_n + \tau_n^s$  le  $n^{i\grave{e}me}$  client achève son service, le serveur devient libre et ainsi de suite.

### Chaîne de Markov induite

Considérons le processus  $\{C(t), N_0(t), t \geq 0\}$ , où  $C(t)$  représente l'état du serveur

$$C(t) = \begin{cases} 0 & \text{si le serveur est libre,} \\ 1 & \text{si le serveur est occupé.} \end{cases}$$

et  $N_0(t)$  est le nombre de clients en orbite à la date  $t$ . En général, ce processus n'est pas un processus de Markov, mais il possède une chaîne de Markov induite.

Soit  $(q_n)$  la chaîne de Markov induite aux instants de départs, où  $q_n = N_0(\xi_n)$  représente le nombre de clients en orbite après le  $n^{i\grave{e}me}$  départ, dont l'équation fondamentale est :

$$q_{n+1} = q_n - \delta_{q_n} + \nu_{n+1},$$

où  $\nu_{n+1}$  est le nombre des clients primaires arrivant dans le système durant le service du  $(n+1)^{i\grave{e}me}$  client. Elle ne dépend pas des événements qui se sont produits avant l'instant  $\eta_{n+1}$  (où l'instant 0 en faisant une translation) du début de service du  $(n+1)^{i\grave{e}me}$  client. La distribution de  $\nu_{n+1}$  est la suivante :

$$P(\nu_n = i) = a_i = \int_0^\infty \frac{(\lambda x)^i}{i!} \exp(-\lambda x) dB(x),$$

où  $a_i > 0, i \geq 0$ . On a les résultats suivants si

$$\nu = \lim_{n \rightarrow \infty} \nu_n, \quad E[\nu] = \rho; \quad \text{alors } A(z) = \sum_0^\infty a_i z^i = \tilde{B}(\lambda - \lambda z).$$

La variable aléatoire  $\delta_{q_n}$  est une variable de Bernoulli définie par :

$$\delta_{q_n} = \begin{cases} 1 & \text{si le } (n+1)^{i\grave{e}me} \text{ client servi provient de l'orbite,} \\ 0 & \text{si le } (n+1)^{i\grave{e}me} \text{ client servi est primaire.} \end{cases}$$

Elle dépend de  $q_n$  et sa distribution est :

$$P(\delta_{q_n} = 1/q_n = i) = \frac{i\theta}{\lambda + i\theta},$$

$$P(\delta_{q_n} = 0/q_n = i) = \frac{i}{\lambda + i\theta}.$$

Les probabilités de transition de l'état  $i$  à l'état  $j$ , ( $\forall j \geq 0$  et  $0 \leq i \leq j$ ) sont :

$$r_{ij} = P(q_{n+1} = j/q_n = i) = a_{j-i} \frac{\lambda}{\lambda + i\theta} + a_{j-i+1} \frac{i\theta}{\lambda + i\theta}.$$

La condition d'existence du régime stationnaire peut être obtenue comme suit :  
L'accroissement moyen de la chaîne vaut

$$\begin{aligned} E[q_{n+1} - q_n / q_n = i] &= E[\nu_{n+1}] - E[\delta q_n = 1 / q_n = i] \\ &= \rho - \frac{i\theta}{\lambda + i\theta}. \end{aligned}$$

Si  $\rho < 1$ , alors :

$$\lim_{i \rightarrow \infty} E[q_{n+1} - q_n / q_n = i] = \rho - 1 < 0$$

et la chaîne est donc ergodique. Par contre, si  $\rho \geq 1$ , alors :

$$\lim_{i \rightarrow \infty} E[q_{n+1} - q_n / q_n = i] = \rho - \frac{i\theta}{\lambda + i\theta} \geq 1 - \frac{i\theta}{\lambda + i\theta} = \frac{\lambda}{\lambda + i\theta} > 0.$$

Puisque la chaîne est bornée inférieurement par la chaîne induite du système  $M/G/1$  classique, donc la chaîne n'est pas ergodique (elle est transitoire).

Soit  $P_n = \lim P(N_0(\xi_i) = n)$ . Les équations de Kolmogorov se présentent de la manière suivante :

$$P_n = \sum_{m=0}^n p_m \frac{\lambda}{\lambda + m\theta} a_{n-m} + \sum_{m=1}^{n+1} p_m \frac{m\theta}{\lambda + m\theta} a_{n-m+1}. \text{ et } n = 0, 1, 2, \dots$$

Vu la présence de convolution, cette équation peut être transformée, à l'aide des fonctions génératrices  $\phi(z) = \sum_{n=0}^{\infty} z^n p_n$  et  $\psi(z) = \sum_{n=0}^{\infty} z^n \frac{p_n}{\lambda + n\theta}$ ,

$$\phi(z) = A(z)(\lambda\phi(z) + \theta\psi'(z)).$$

D'un autre côté ;

$$\begin{aligned} \phi(z) &= \sum_{n=0}^{\infty} z^n p_n = \sum_{n=0}^{\infty} z^n p_n \frac{\lambda + n\theta}{\lambda + n\theta}, \\ &= \lambda \sum_{n=0}^{\infty} z^n p_n \frac{p_n}{\lambda + n\theta} + \theta \sum_{n=0}^{\infty} n z^n \frac{p_n}{\lambda + n\theta}, \\ &= \lambda\psi(z) + \theta\psi'(z). \end{aligned}$$

Par conséquent ;

$$\begin{aligned} \lambda\psi(z) + \theta\psi'(z) &= A(z)(\lambda\psi(z) + \theta\psi'(z)), \\ \theta\psi'(z)[A(z) - z] &= \lambda\psi(z)[1 - A(z)]. \end{aligned}$$

**Lemme 1.3.2.1.** *La fonction analytique  $f(z) = A(z) - z$  est positive, croissante et pour  $z \in [0, 1]$ ,  $\rho < 1 : z < A(z) < 1$ .*

*Démonstration.* Soit

$$f(z) = \tilde{B}(\lambda - \lambda z) - z, \quad f(1) = \tilde{B}(0) - 1 = 0$$

en plus :

$$f'(z) = -\lambda \tilde{B}'(\lambda - \lambda z) - 1 \text{ et } f'(1) = \rho - 1 < 0,$$

alors 1 est le seul zéro de  $f$ . En outre ;

$$f''(z) = -\lambda \tilde{B}'(\lambda - \lambda z) + \lambda^2 \tilde{B}''(\lambda - \lambda z) \geq 0.$$

Alors  $f(z)$  est décroissante sur  $[0, 1]$ , positive pour  $\rho = \frac{\lambda}{\gamma} < 1$  et pour  $z \in [0, 1]$  :

$$z < f(z) < 1.$$

Notons aussi que :

$$\lim_{z \rightarrow 1^-} \frac{1 - \tilde{B}(\lambda - \lambda z)}{\tilde{B}(\lambda - \lambda z) - z} = \frac{\rho}{1 - \rho} < \infty.$$

□

**Théorème 1.3.2.1.** Soit  $\rho < 1$ , la distribution stationnaire de la chaîne de Markov induite possède la fonction génératrice suivante :

$$\phi(z) = \sum_{n=0}^{\infty} z^n P_n = \frac{(1 - \rho)(1 - z)A(z)}{A(z) - z} \exp \left\{ \frac{\lambda}{\theta} \int_1^z \frac{1 - A(u)}{A(u) - u} du \right\},$$

où  $A(z) = \tilde{B}(\lambda - \lambda z)$ .

### Distribution stationnaire de l'état du système

Le premier résultat sur le système  $M/G/1$  avec rappels est basé sur la méthode des variables supplémentaires. Une des approches permettant de trouver la distribution stationnaire jointe de l'état du serveur et de la taille de l'orbite a été introduite par De Kok (1984) [29]. Elle consiste à décrire le processus des arrivées comme processus de Markov avec dépendance de l'état de paramètre  $\lambda_{in}$  quand  $\{C(t), N_0(t)\}$  est dans l'état  $(i, n)$  et à appliquer les schémas récursifs. L'état du système peut être décrit par le processus

$$X(t) = \begin{cases} N_0(t) & \text{si } C(t) = 0 \\ \{C(t), N_0(t), \xi(t)\} & \text{si } C(t) = 1 \end{cases},$$

où  $\xi(t)$  est une variable aléatoire supplémentaire à valeurs dans  $\mathbb{R}^+$ , et désignant la durée de service écoulé à la date  $t$ . Notons par :

$$P_{0n} = \lim_{t \rightarrow \infty} P(C(t) = 0, N_0(t) = n),$$

$$P_{1n}(x) = \lim_{t \rightarrow \infty} \frac{d}{dx} P(C(t) = 1, \xi(t) \leq x, N_0(t) = n).$$

Les probabilités  $P_{0n}$  et  $P_{1n}(x)$  vérifient le système d'équations de balance :

$$\begin{aligned}(\lambda + n\theta)p_{0n} &= \int_0^\infty p_{1n}(x)b(x)dx, \\ p'_{1n}(x) &= -(\lambda + b(x))p_{1n}(x) + \lambda p_{1n-1}(x), \\ p_{1n}(0) &= -\lambda p_{0n} + (n + 1)\theta p_{0n+1}.\end{aligned}$$

où  $b(x) = \frac{B'(x)}{1-B(x)}$  est l'intensité instantanée du service étant donné que la durée écoulée est égale à  $x$ . Soient les fonctions génératrices, telles que  $P_0(z) = \sum_{n=0}^\infty z^n p_{0n}$  et  $P_1(z, x) = \sum_{n=0}^\infty z^n p_{1n}(x)$ . Le système d'équations de balance devient :

$$\begin{aligned}\lambda \sum_{n=0}^\infty z^n p_{0n} + \theta \sum_{n=0}^\infty z^n p_{0n} &= \int_0^\infty \sum_{n=0}^\infty z^n p_{1n}(x)dx, \\ \sum_{n=0}^\infty z^n p'_{1n}(x) &= -(\lambda + b(x)) \sum_{n=0}^\infty z^n p_{1n}(x) + \lambda \sum_{n=0}^\infty z^n p_{1n-1}(x), \\ \sum_{n=0}^\infty z^n p_{1n}(0) &= \lambda \sum_{n=0}^\infty z^n p_{0n} + \theta \sum_{n=0}^\infty z^n (n + 1)p_{0n+1}.\end{aligned}$$

D'où ;

$$\begin{cases} \lambda P_0(z) + \theta z \lambda P'_0(z) = \int_0^\infty P_1(z, x)b(x)dx, \\ P'_1(z, x) = (\lambda z - \lambda - b(x))P_1(z, x), \\ P_1(z, 0) = \lambda P_0(z). \end{cases} \quad (1.1)$$

De la deuxième équation de (1.1), on a :

$$P_1(z, x) = P_1(z, 0) [1 - B(x)] \exp(-(\lambda - \lambda z)x).$$

Donc, la première équation de (1.1), devient ;

$$\begin{aligned}\lambda P_0(z) + \theta z P'_0(z) &= \int_0^\infty P_1(z, 0) [1 - B(x)] \exp(-(\lambda - \lambda z)x)b(x)dx, \\ &= P_1(z, 0) \tilde{B}(\lambda - \lambda z) = P_1(z, 0)A(z).\end{aligned} \quad (1.2)$$

A partir des équations (1.1) et (1.2), on a :

$$P_1(z, 0)f(z) = \lambda P_0(z) + \theta z \left( \frac{P_1(z, 0)}{\theta} - \frac{\lambda}{\theta} P_0(z) \right), \quad (1.3)$$

$$P_1(z, 0) = \frac{\lambda - \lambda z}{A(z) - z} P_0(z) [1 - B(x)] \exp(-\lambda(\lambda - \lambda z)x). \quad (1.4)$$

En intégrant cette équation, et en utilisant la formule

$$\int_0^\infty \exp(-sx) [1 - B(x)] dx = \frac{(1 - \tilde{B}(s))}{s},$$

On obtient ;

$$P_1(z) = \int_0^\infty P_1(z, x) dx = P_0(z) \frac{1 - A(z)}{A(z) - z}.$$

De (1.2) et (1.4), on peut obtenir  $P_0(z)$

$$\lambda P_0(z) + \theta z P_0'(z) = A(z) [\lambda P_0(z) + \theta P_0'(z)],$$

$$\theta [A(z) - z] P_0'(z) = \lambda [1 - A(z)] P_0(z).$$

Considérons  $f(z) = A(z) - z$ . Du lemme 1.3.2.1,  $f(z)$  est une fonction décroissante sur  $[0, 1]$ , positive et pour  $\rho < 1$  et  $z \in [0, 1] : z < A(z) < 1$ . En plus,

$$\lim_{z \rightarrow 1^-} \frac{1 - A(z)}{A(z) - z} = \frac{A'(1)}{A(1) - 1} = \frac{\rho}{1 - \rho} < \infty.$$

De ce fait, pour  $z = 1$ , la fonction  $\frac{1 - A(z)}{A(z) - z} = \frac{\rho}{1 - \rho}$ .

**Théorème 1.3.2.2.** Si  $\rho = \lambda\beta_1 < 1$ , le système est en régime stationnaire et les fonctions génératrices de la distribution conjointe de l'état du serveur et de la taille de l'orbite sont données par :

$$P_0(z) = \sum_{n=0}^{\infty} z^n p_{0n} = (1 - \rho) \exp \left[ \frac{\lambda}{\theta} \int_1^z \frac{1 - A(u)}{A(u) - u} du \right].$$

$$P_1(z) = \sum_{n=0}^{\infty} z^n p_{1n} = \frac{1 - A(z)}{A(z) - z} P_0(z).$$

Par conséquent, la distribution marginale du nombre de serveurs occupés s'exprime de la manière suivante :

$$P_0 = \lim_{t \rightarrow \infty} P(C(t) = 0) = 1 - \rho; \quad P_1 = \lim_{t \rightarrow \infty} P(C(t) = 1) = \rho.$$

La fonction génératrice de la distribution marginale de la taille de l'orbite est définie par :

$$P(z) = P_0(z) + P_1(z) = \frac{1 - z}{A(z) - z} P_0(z)$$

$$= \frac{(1 - \rho)(1 - z)}{A(z) - z} \exp \left[ \frac{\lambda}{\theta} \int_1^z \frac{1 - A(u)}{A(u) - u} du \right],$$

et la fonction génératrice de la distribution de l'état stationnaire du nombre de clients dans le système est :

$$Q(z) = P_0(z) + zP_1(z) = \frac{(1 - \rho)(1 - z)A(z)}{A(z) - z} \exp \left[ \frac{\lambda}{\theta} \int_1^z \frac{1 - A(u)}{A(u) - u} du \right].$$



## Mesures de performance

Les caractéristiques du modèle sont :

- Nombre moyen de clients dans le système

$$\bar{n} = Q'(1) = \rho + \frac{\lambda^2 \beta_2}{2(1-\rho)} + \frac{\lambda \rho}{\theta(1-\rho)},$$

- Nombre moyen de clients en orbite

$$\bar{n}_0 = P'(1) = \bar{n} - \rho = \frac{\lambda^2 \beta_2}{2(1-\rho)} + \frac{\lambda \rho}{\theta(1-\rho)},$$

- Temps moyen d'attente d'un client

$$\bar{W} = \frac{\bar{n}_0}{\lambda} = \frac{\lambda \beta_2}{2(1-\rho)} + \frac{\rho}{\theta(1-\rho)},$$

- Nombre moyen de rappels par client (d'après la formule de Little)

$$\bar{R} = \bar{W} \theta = \frac{\lambda \theta \beta_2}{2(1-\rho)} + \frac{\rho}{1-\rho}.$$

### 1.3.3 Modèle de file d'attente $M/G/1$ avec rappels et feedback

La majorité des études sur les systèmes d'attente avec rappels considère le modèle sans feedback. Néanmoins, plusieurs situations réelles peuvent être modélisées comme des systèmes de files d'attente avec rappels et feedback. La notion de feedback a été initialement introduite par Takas (1963) [76] pour l'étude de certains systèmes d'attente classiques, et depuis plusieurs papiers sont apparus sur ce sujet en considérant d'autres types de systèmes avec différentes variantes (rappels, vacances, pannes,...) ([55], [17], [16], [9], [10], ...). Le phénomène de feedback dans ce type de systèmes peut apparaître dans plusieurs situations pratiques, par exemple, dans les systèmes MATS (Multiple Access Telecommunication Systems) où des messages s'avérant comme erreurs à la destination (messages perdus ou corrompus) sont renvoyés. Encore, ils sont utilisés pour modéliser le protocole ARQ (Automatic Repeat Request) dans un réseau de communication à haute fréquence. Dans ces systèmes, à chaque fois qu'un client arrive et trouve le serveur occupé, quitte le service pour rejoindre l'orbite, sinon il reçoit immédiatement son service. Une fois servi, il décide avec une probabilité  $c$  de retourner en orbite pour un autre service ou de quitter le système définitivement avec une probabilité complémentaire  $1 - c$ .

#### Description du modèle

Nous considérons un système de file d'attente à un seul serveur et espace d'attente inexistant, avec :

- **Processus des arrivés** : Les appels arrivent selon un processus homogène de Poisson de taux  $\lambda > 0$ .
- **Processus de service** : La durée de service suit une loi générale  $B(x)$ , éventuellement exponentielle, de TL-S  $\tilde{B}(s)$ ,  $Re(s) > 0$  et de moyenne finie  $\frac{1}{\gamma}$ .
- **Processus de rappels** : Chaque client primaire reçoit immédiatement son service si le serveur est libre, sinon il rejoint l'orbite. Chaque client en orbite produit un flux poissonnien de paramètre  $\theta > 0$ .
- **Protocole de feedback** : le client servi doit décider, soit il rejoint l'orbite pour un autre service avec une probabilité  $c$  où il quitte le système définitivement avec une probabilité  $\bar{c} = 1 - c$ .

Toutes les variables aléatoires introduites sont mutuellement indépendantes.  
Le modèle est représenté par la figure 1.2.

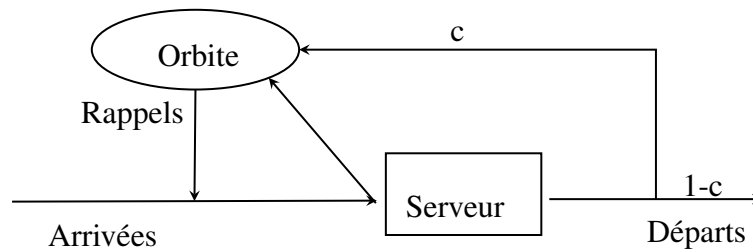


FIGURE 1.2 – Schéma illustratif d'un système de file d'attente avec rappels et feedback.

L'état du système à la date  $t$  peut être décrit par le processus stochastique  $\{C(t), N_o(t), \zeta(t), t \geq 0\}$ , où  $N_o(t)$  est le nombre de clients en orbite,  $C(t)$  est l'état du serveur à la date  $t$  ( $C(t) = 0$  : serveur libre ;  $C(t) = 1$  : serveur occupé). Si  $C(t) = 1$ ,  $\zeta(t)$  représente le temps de service écoulé du client en service.

### La chaîne de Markov induite et la condition d'ergodicité

Soit  $\xi_n$  l'instant où le serveur devient libre pour la  $n^{\text{ième}}$  fois. La suite de variables aléatoire  $\{q_n = N_o(\xi_n), n \geq 1\}$  forme une chaîne de Markov induite, dont l'équation fondamentale est :

$$q_{n+1} = q_n - \delta_{q_n} + v_{n+1} + u. \quad (1.5)$$

La variable aléatoire  $v_{n+1}$  représente le nombre de clients primaires arrivant au système durant le  $(n+1)^{\text{ième}}$  temps de service. Sa distribution est donnée par :

$$k_i = P(v_{n+1} = i) = \int_0^{\infty} \frac{(\lambda x)^i}{i!} e^{-\lambda x} dB(x)$$

avec la fonction génératrice  $K(z) = \tilde{B}(\lambda - \lambda z)$ .

La variable aléatoire  $\delta_{q_n}$  est définie par

$$\delta_{q_n} = \begin{cases} 1 & \text{si le } (n+1)^{\text{ième}} \text{ client servi provient de l'orbite} \\ 0 & \text{si le } (n+1)^{\text{ième}} \text{ client servi est primaire,} \end{cases}$$

elle dépend de  $q_n$ , et sa distribution est donnée par

$$P[\delta_{q_n} = 1/q_n = k] = \frac{k\theta}{\lambda + k\theta}; P[\delta_{q_n} = 0/q_n = k] = \frac{\lambda}{\lambda + k\theta}.$$

La variable aléatoire  $u$  est définie par

$$u = \begin{cases} 1 & \text{le client servi décide de rejoindre l'orbite} \\ 0 & \text{le client servi décide de quitter le système.} \end{cases}$$

Encore,  $P[u = 0] = \bar{c}$  et  $P[u = 1] = c$ . Pour le calcul des probabilités de transition à une étape de la chaîne, on a :

$$\begin{aligned} r_{nm} &= P(q_{i+1} = m/q_i = n) \\ &= P(q_i - \delta_{q_i} + v_{n+1} + u = m/q_i = n) \\ &= P(v_{i+1} = m - k + \delta_{q_i} - u/q_i = n) \\ &= P(v_{i+1} = m - n/q_i = n, u = 0, \delta_{q_i} = 0) P(u = 0) P(\delta_{q_i} = 0/q_i = n) \\ &\quad + P(v_{i+1} = m - n + 1/q_i = n, u = 0, \delta_{q_i} = 1) P(u = 0) P(\delta_{q_i} = 1/q_i = n) \\ &\quad + P(v_{i+1} = m - n - 1/q_i = n, u = 1, \delta_{q_i} = 0) P(u = 1) P(\delta_{q_i} = 0/q_i = n) \\ &\quad + P(v_{i+1} = m - n/q_i = n, u = 1, \delta_{q_i} = 1) P(u = 1) P(\delta_{q_i} = 1/q_i = n) \\ &= P(v_{i+1} = m - n) P(u = 0) P(\delta_{q_i} = 0/q_i = n) \\ &\quad + P(v_{i+1} = m - n + 1) P(u = 0) P(\delta_{q_i} = 1/q_i = n) \\ &\quad + P(v_{i+1} = m - n - 1) P(u = 1) P(\delta_{q_i} = 0/q_i = n) \\ &\quad + P(Pv_{i+1} = m - n) P(u = 1) P(\delta_{q_i} = 1/q_i = n) \\ &= k_{m-n} \frac{\lambda}{\lambda + n\theta} \bar{c} + k_{m-n+1} \frac{n\theta}{\lambda + n\theta} \bar{c} + k_{m-n-1} \frac{\lambda}{\lambda + n\theta} c + k_{m-n} \frac{n\theta}{\lambda + n\theta} c. \end{aligned} \quad (1.6)$$

Notons que  $r_{km} \neq 0$  uniquement pour  $k = 0, 1, \dots, m + 1$ .

**Théorème 1.3.3.1.** *La chaîne de Markov induite  $\{q_n, n \geq 1\}$  est ergodique ssi*

$$\rho = \frac{\lambda}{\gamma} + c < 1.$$

*Démonstration.* De (1.5), il est claire que  $\{q_n, n \geq 1\}$  est une chaîne de Markov irréductible et périodique. Pour montrer la condition suffisante, nous utilisons le critère de Foster. On établit l'existence d'une fonction non-négative  $f(s)$ ,  $s \in S$ , et  $\varepsilon > 0$  telle que l'accroissement moyen de la chaîne  $x_s = E[f(q_{n+1}) - f(q_n) | q_n = s]$  est finie pour tout  $s \in S$  et  $x_s \leq -\varepsilon$  pour tout  $s \in S$  sauf peut être pour un nombre fini. Soit  $f(q_n) = k$ . Alors

$$\begin{aligned} x_k &= E[q_{n+1} - q_n / q_n = k] \\ &= E[v_{n+1} - \delta_{q_n} + u / q_n = k] \\ &= E[v_{n+1} / q_n = k] - E[\delta_{q_n} / q_n = k] + E[u / q_n = k] \\ &= \frac{\lambda}{\gamma} - \frac{k\theta}{\lambda + k\theta} + c. \end{aligned}$$

Soit  $x = \lim_{k \rightarrow \infty} x_k$ . Donc  $x = \frac{\lambda}{\gamma} - 1 + c < 0$ . Par conséquent, la condition suffisante est  $\frac{\lambda}{\gamma} + c < 1$ . Pour prouver que la condition ci-dessus est aussi la condition nécessaire, nous appliquons la condition de Kaplan : une chaîne de Markov irréductible et périodique n'est pas ergodique si  $x_k < \infty$  pour tout  $k \geq 0$  et il existe  $k_0 \in \mathbb{N}$ , tel que  $x_k \geq 0$  pour tout  $k \geq k_0$ . Dans notre cas, cette condition est vérifiée parce que  $r_{km} = 0$  pour  $m < k - 1$  et  $k > 0$ .

Cependant,  $\frac{\lambda}{\gamma} + c \geq 1$  est la négation de l'ergodicité pour la chaîne de Markov induite  $\{q_n, n \geq 1\}$ .

Finalement,  $\{q_n, n \geq 1\}$  est ergodique ssi  $\rho = \frac{\lambda}{\gamma} + c < 1$  [17].  $\square$

### Distribution stationnaire de l'état du système

Sous la condition que  $\rho < 1$ , nous cherchons la distribution stationnaire  $\pi_m = \lim_{n \rightarrow \infty} P(q_n = m)$ . De l'équation (1.6), nous obtenons les équations de Kolmogorov pour  $d_m$  :

$$\pi_m = \sum_{n=0}^m k_{m-n} \frac{\lambda}{\lambda + n\theta} \bar{c} \pi_n + \sum_{n=1}^{m+1} k_{m-n+1} \frac{n\theta}{\lambda + n\theta} \bar{c} \pi_n + \sum_{n=0}^{m-1} k_{m-n-1} \frac{\lambda}{\lambda + n\theta} c \pi_n + \sum_{n=0}^m k_{m-n} \frac{n\theta}{\lambda + n\theta} c \pi_n.$$

A l'aide des fonctions génératrices  $\Pi(z) = \sum_{m=0}^{\infty} \pi_m z^m$ ,  $L(z) = \sum_{m=0}^{\infty} \frac{\pi_m}{\lambda + m\theta} z^m$  et en utilisant l'équation (1.5), on obtient

$$\begin{aligned} \Pi(z) &= E[z^{q_{n+1}}] = E[z^{q_n - \delta_{q_n} + v_{n+1} + u}] \\ &= \sum_{m=0}^{\infty} E[z^{m - \delta_{q_n} + v_{n+1} + u} / q_n = m] P[q_n = m] \\ &= \sum_{m=0}^{\infty} z^m E[z^{-\delta_{q_n}} / q_n = m] E[z^{v_{n+1}}] E[z^u] P[q_n = m] \\ &= \sum_{m=0}^{\infty} z^m \left( \frac{\lambda}{\lambda + m\theta} \cdot 1 + \frac{m\theta}{\lambda + m\theta} \cdot z^{-1} \right) K(z) (\bar{c} + cz) \pi_m \\ &= K(z) (\bar{c} + cz) \left( \lambda \sum_{m=0}^{\infty} \frac{\pi_m}{\lambda + m\theta} z^m + \theta \sum_{m=0}^{\infty} m \frac{\pi_m}{\lambda + m\theta} z^{m-1} \right) \\ &= K(z) (\bar{c} + cz) [\lambda L(z) + \theta L'(z)]. \end{aligned} \tag{1.7}$$

D'autre part,

$$\begin{aligned} \Pi(z) &= \sum_{m=0}^{\infty} \pi_m z^m \frac{\lambda + m\theta}{\lambda + m\theta} \\ &= \lambda L(z) + \theta z L'(z). \end{aligned} \tag{1.8}$$

De (1.7), nous avons :

$$\lambda L(z) + \theta z L'(z) = K(z) (\bar{c} + cz) \left[ \lambda L(z) + \theta L'(z) \right]$$

ou

$$\theta L'(z) \left[ (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z \right] = \lambda L(z) \left[ 1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) \right].$$

Considérons la fonction  $f(z) = (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z$ . On a :

$$\begin{aligned} f(1) &= \tilde{B}(0) - 1 = 1 - 1 = 0; \\ f'(z) &= -\lambda (\bar{c} + cz) \tilde{B}'(\lambda - \lambda z) - c \tilde{B}(\lambda - \lambda z) - 1 \text{ et} \\ f'(1) &= -\lambda \tilde{B}'(0) + c \tilde{B}(0) - 1 = \rho - 1 < 0; \\ f''(z) &= \lambda^2 (\bar{c} + cz) \tilde{B}''(\lambda - \lambda z) \geq 0. \end{aligned}$$

Par conséquent, la fonction  $f(z)$  est décroissante dans  $[0, 1]$ ,  $z = 1$  est le seule zéro, et pour  $z \in [0, 1]$  la fonction est positive : comme  $\rho < 1$ , nous avons

$$z < (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) \leq 1.$$

D'où

$$\lim_{z \rightarrow 1} \frac{1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} = \frac{\rho}{1 - \rho} < \infty.$$

Donc pour  $z \in [0, 1]$ , on a

$$L'(z) = \frac{\lambda 1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{\theta (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} L(z).$$

La solution de cet équation est :

$$L(z) = L(1) \exp \left\{ \frac{\lambda}{\theta} \int_1^z \frac{1 - (\bar{c} + cu) \tilde{B}(\lambda - \lambda u)}{(\bar{c} + cu) \tilde{B}(\lambda - \lambda u) - u} du \right\}.$$

De (1.8),

$$\begin{aligned} \Pi(z) &= \lambda L(z) + \theta z L'(z) \\ &= \lambda L(z) + \lambda z \frac{1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} L(z) \\ &= \lambda \frac{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) (1 - z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} L(1) \exp \left\{ \frac{\lambda}{\theta} \int_1^z \frac{1 - (\bar{c} + cu) \tilde{B}(\lambda - \lambda u)}{(\bar{c} + cu) \tilde{B}(\lambda - \lambda u) - u} du \right\}. \end{aligned}$$

Puisque  $\Pi(1) = 1$ , nous avons  $L(1) = \sum_{m=0}^{\infty} \frac{\pi_m}{\lambda + m\theta} = \frac{1 - \rho}{\lambda}$ .

Finalement, on obtient l'expression de la fonction génératrice de la distribution stationnaire du nombre de clients dans le système aux instants de départs est donnée par :

$$\Pi(z) = \left[ \frac{(1 - \rho) \tilde{B}(\lambda - \lambda z)(1 - z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} \right] \left[ (\bar{c} + cz) \exp \left\{ \frac{\lambda}{\theta} \int_1^z \frac{1 - (\bar{c} + cu) \tilde{B}(\lambda - \lambda u)}{(\bar{c} + cu) \tilde{B}(\lambda - \lambda u) - u} du \right\} \right]. \quad (1.9)$$

**Remarque 1.3.3.1.** Le facteur  $\frac{(1-\rho)\tilde{B}(\lambda-\lambda z)(1-z)}{(\bar{c}+cz)\tilde{B}(\lambda-\lambda z)-z}$  est la fonction génératrice du nombre de clients dans le système  $M/G/1$  avec feedback [76];

le second facteur  $(\bar{c} + cz) \exp \left\{ \frac{\lambda}{\theta} \int_1^z \frac{1 - (\bar{c} + cu) \tilde{B}(\lambda - \lambda u)}{(\bar{c} + cu) \tilde{B}(\lambda - \lambda u) - u} du \right\}$  est la fonction génératrice du nombre de clients dans le système de file d'attente  $M/G/1$  avec rappels et feedback étant donné que le serveur est libre.

## Mesures de performance

Les caractéristiques du modèle sont :

- Nombre moyen de clients dans le système :

$$\bar{n} = \Pi'(1) = \rho + c(1 - \lambda\beta_1) + \frac{\bar{c}\lambda\rho}{\mu(1 - \rho)} + \frac{c\lambda\beta_1}{(1 - \rho)} + \frac{\lambda^2\beta_2}{2(1 - \rho)}.$$

- Nombre moyen de clients en orbite :

$$\bar{n}_0 = \bar{n} - \rho = c(1 - \lambda\beta_1) + \frac{\bar{c}\lambda\rho}{\mu(1 - \rho)} + \frac{c\lambda\beta_1}{(1 - \rho)} + \frac{\lambda^2\beta_2}{2(1 - \rho)}.$$

- Temps moyen d'attente :

$$\bar{W} = \frac{c(1 - \lambda\beta_1)}{\lambda} + \frac{\bar{c}\rho}{\mu(1 - \rho)} + \frac{c\beta_1}{(1 - \rho)} + \frac{\lambda\beta_2}{2(1 - \rho)}.$$

- Nombre moyen de rappels par client :

$$\bar{R} = \mu\bar{W} = \frac{c\mu(1 - \lambda\beta_1)}{\lambda} + \frac{\bar{c}\rho}{(1 - \rho)} + \frac{c\mu\beta_1}{(1 - \rho)} + \frac{\lambda\mu\beta_2}{2(1 - \rho)}.$$

- Nombre moyen de clients qui demandent un service supplémentaire :

$$\bar{F} = c\Pi'(1) = c\rho + c^2(1 - \lambda\beta_1) + \frac{c\bar{c}\lambda\rho}{\mu(1 - \rho)} + \frac{c^2\lambda\beta_1}{(1 - \rho)} + \frac{\lambda^2c\beta_2}{2(1 - \rho)}.$$

## 1.4 Conclusion

Dans ce chapitre, nous avons rappelé et présenté les notions et techniques de base sur les systèmes de files d'attente avec rappels, nous nous sommes intéressés aux modèles d'attente de type  $M/M/1$ ,  $M/G/1$  avec rappels, et le  $M/G/1$  avec rappels et feedback. Une étude plus poussée de ce genre de systèmes est nécessaire pour améliorer et mieux évaluer les performances des systèmes informatiques, des réseaux de communications, systèmes industriels et systèmes complexes dans nombreux domaines. Cette technique est devenu inconcevable pour construire un système quelconque sans avoir fait une analyse des performances au préalable. Les modèles d'attente développés ces dernières décennies tentent de prendre en considération la recherche des clients en orbite par le serveur. Ces phénomènes affectent les caractéristiques de performance des systèmes réels.

## Chapitre 2

# Modèles de files d'attente avec rappels et recherche des clients en orbite.

### 2.1 Introduction

Nous considérons des modèles de files d'attente avec rappels sans salle d'attente dans lesquels, chaque époque d'achèvement du service ne doit pas nécessairement être suivi d'un temps d'inactivité. le temps d'inactivité du serveur est réduit par l'introduction de la recherche des clients en orbite immédiatement après l'achèvement du service. Ceci est réalisé comme suit : immédiatement à l'achèvement de chaque service, le serveur sélectionne un client de l'orbite avec une probabilité  $p_j$ , ou il y a  $j$  clients dans l'orbite ( il est supposé que le serveur est au courant de l'état de l'orbite, par exemple, il a un registre des clients en orbite par contre ces derniers ignorent son état). Avec une probabilité  $1 - p_j$  le serveur n'effectue aucune recherche des clients à la fin de son service et dans ce cas, comme dans la file d'attente avec rappel classique, une concurrence prend place entre les clients primaire et les clients secondaire pour atteindre le service. Ainsi, si la recherche est faite, le service sera suivi d'un autre service si non, il sera suivi d'un temps d'inactivité. L'étude à deux objectifs principaux :

1. Le premier consiste à introduire la recherche des clients en orbite dans un modèle de file d'attente avec rappels, ce qui permet de réduire le temps d'inactivité du serveur. Si le tenu des coûts et de mise en œuvre de la recherche de clients sont introduits. Les résultats obtenus peuvent être utilisés pour le réglage optimal des paramètres du mécanisme de la recherche.
2. Le deuxième objectif est de donner un aperçu du lien entre la file d'attente avec rappels correspondante et la file d'attente classique (sans rappels).

Cependant, on observe que si  $p_j = 1$ , le modèle sera réduit à celui d'une file d'attente classique (sans rappels) et quand  $p_j = 0$ , il devient le modèle de file d'attente avec rappels.



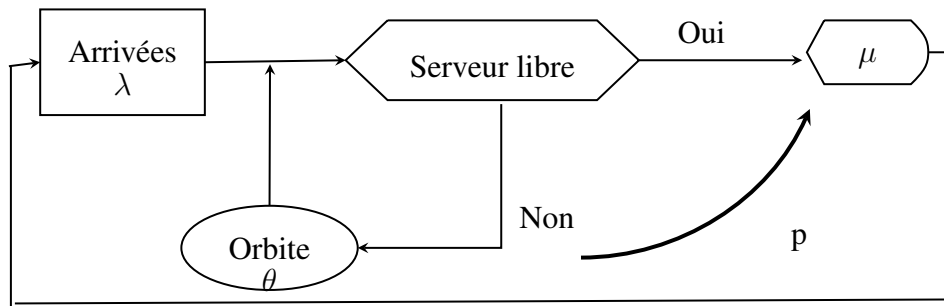


FIGURE 2.1 – Représentation schématique d'un système de files d'attente avec rappels et recherche de clients en orbite.

## 2.2 Modèles d'attente avec rappels linéaire et recherche des clients en orbite

Les temps d'inter-rappels peuvent être modélisés selon différentes disciplines en fonction de chaque application particulière. Dans les systèmes téléphoniques, les tentatives répétées sont effectués individuellement par chaque client bloqué suivant une loi exponentielle de taux  $\theta$ . C'est la discipline de rappels classique dont le taux est  $j\theta$ , quand la taille de l'orbite est  $j \geq 1$ . En revanche, il existe d'autres types de situations de file d'attente dont les temps d'inter-rappels sont indépendants du nombre de clients en orbite. Cette deuxième possibilité est la discipline de rappels constante, i.e. le taux d'inter-rappels est  $\alpha(1 - \delta_{j0})$  où  $\delta_{j0}$  est la fonction de Kronecker, Artalejo et Gomez-Corral [7] unissent les deux disciplines en définissant la politique d'inter-rappels linéaire avec le taux  $\alpha(1 - \delta_{j0}) + j\theta$ . L'application suivante motive l'analyse du modèle considéré ci-dessous.

Réparation par le service avec la recherche des clients : le serveur tient un registre des clients qui sont obligés de quitter le système car ils ont rencontré un serveur occupé au moment de leurs arrivées. À la fin d'un service, le serveur décide de démarrer immédiatement le prochain service en récupérant un client non satisfait (*de l'orbite*) avec une probabilité  $p_j$ .

Le temps de recherche est supposé *négligeable*. La probabilité de ne pas effectuée une recherche des clients est  $q_j = 1 - p_j$ . Si le serveur ne récupère pas le prochain client à servir de l'orbite, alors il va y avoir une concurrence entre les clients primaires et ceux de l'orbite pour attendre le prochain service. Ainsi, le travail actuel inclut comme cas particulier, la file d'attente classique lorsque  $p_j = 1$ ,  $j \neq 0$  et la file d'attente avec rappels lorsque  $p_j = 0$ .

### 2.2.1 Modèle d'attente $M/M/1$ avec rappels linéaire et recherche des clients

#### Description du modèle :

Nous considérons un système de file d'attente à serveur unique auquel les clients primaire arrivent selon un flux Poissonienne de taux  $\lambda$ . Tout client qui trouve, le serveur occupé quitte immédiatement la zone de service et rejoint l'orbite. L'intervalle entre deux tentatives répétées successives est distribué de manière exponentielle avec un taux  $\alpha(1 - \delta_{j0}) + j\theta$ , étant donné que

le nombre de clients en orbite est  $j$ ,  $j \geq 1$ . La distribution du temps de service est exponentiel de taux  $\mu$ . Soit  $\xi_j$  l'instant de l'achèvement de service du  $j^{\text{ème}}$  client. Immédiatement après ce moment, le serveur passe à la recherche d'un client en orbite avec une probabilité  $p_j$ ,  $p_0 = 0$  qui dépend du nombre de clients  $j$  en orbite. Avec la probabilité  $q_j = 1 - p_j$  le serveur reste libre. Dans ce dernier cas, l'événement à suivre dépend d'une concurrence entre les arrivées primaires du taux  $\lambda$  et le flux d'inter-rappels de taux  $\alpha(1 - \delta_{j0}) + j\theta$ . Le temps de recherche est supposé être *négligeable*.

Les flux d'arrivées primaires, d'inter-rappels et du temps de service sont supposés être mutuellement indépendants.

### Distribution stationnaire de l'état du système

Soit  $N(t)$  le nombre de clients en orbite et  $C(t)$  est l'état du serveur à l'instant  $t$ . Nous avons :

$$C(t) = \begin{cases} 0 & \text{si le serveur est libre,} \\ 1 & \text{si le serveur est occupé.} \end{cases}$$

L'état du système à la date  $t$  peut être décrit par le processus stochastique suivant :

$$\{C(t), N(t), \xi(t); t \geq 0\},$$

l'espace d'état du processus est  $S = \{0, 1\} \times \mathbb{N}$ .  $\rho = \frac{\lambda}{\mu} < 1$ , est la condition de stabilité du système proposé. Soit les probabilités de l'état du système données par  $P_{ij} = P[C(t) = i, N(t) = j]$ , les transitions entre les états sont illustrées dans la figure 2.2.

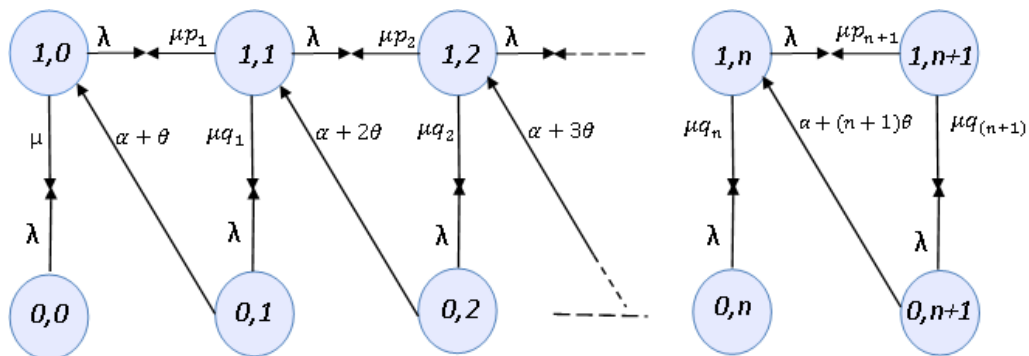


FIGURE 2.2 – Graphe des transitions du modèle  $M/M/1$  avec rappels et recherche de clients en orbite.

**Théorème 2.2.1.1.** La distribution de l'état du serveur  $C(t)$  et du nombre de clients dans l'orbite  $N(t)$  du système  $M/M/1$  au régime stationnaires est donnée par :

$$P_{0j} = P_{00}q_j\rho^j \prod_{k=0}^{j-1} \frac{\lambda + \alpha(1 - \delta_{k0}) + k\theta}{p_{k+1}\lambda + \alpha + (k+1)\theta}, \quad j \geq 1, \quad (2.1)$$

$$P_{1j} = P_{00}\rho^{j+1} \prod_{k=1}^j \frac{\lambda + \alpha + k\theta}{p_k\lambda + \alpha + k\theta}, j \geq 0, \quad (2.2)$$

$$P_{00}^{-1} = \sum_{n=0}^{\infty} \rho^{j+1} \left( 1 + \frac{q_j\mu}{\lambda + \alpha(1 - \delta_{j0}) + j\theta} \right) \prod_{k=1}^j \frac{\lambda + \alpha + k\theta}{p_k\lambda + \alpha + k\theta}. \quad (2.3)$$

*Démonstration.* L'ensemble des équations d'équilibre statistique pour les probabilités  $P_{0j}$  et  $P_{1j}$  est :

$$(\lambda + \alpha(1 - \delta_{j0}) + j\theta)P_{0j} = q_j\mu P_{1j}, j \geq 0, \quad (2.4)$$

$$(\lambda + \mu)P_{1j} = \lambda P_{1,j-1} + \lambda P_{0j} + [\alpha + (j + 1)\theta] P_{0,j+1} + \mu p_{j+1} P_{1,j+1}, j \geq 0, \quad (2.5)$$

En utilisant l'équation (2.4), éliminer les probabilités  $P_{1j}$  de l'équation (2.5). Après quelques calculs sur l'équation résultante, nous obtenons :

$$\begin{aligned} & \mu q_{j-1} q_j (\alpha + (j + 1)\theta + \lambda p_{j+1}) P_{0,j+1} - \lambda q_{j-1} q_{j+1} (\lambda + \alpha + j\theta) P_{0j} \\ &= \mu q_{j-1} q_{j+1} (\alpha + j\theta + \lambda p_j) P_{0j} - \lambda q_j q_{j+1} (\lambda + \alpha(1 - \delta_{j-1,0}) + (j - 1)\theta) P_{0,j-1}, \end{aligned}$$

cela implique que :

$$\mu q_{j-1} q_{j+1} (\alpha + j\theta + \lambda p_j) P_{0j} - \lambda q_j q_{j+1} (\lambda + \alpha(1 - \delta_{j-1,0}) + (j - 1)\theta) P_{0,j-1} = 0,$$

donc :

$$P_{0j} = \frac{\lambda q_j (\lambda + \alpha(1 - \delta_{j-1,0}) + (j - 1)\theta)}{\mu q_{j-1} (\alpha + j\theta + \lambda p_j)} P_{0,j-1}.$$

En résolvant récursivement, on trouve les équations (2.1), (2.2) et (2.3).  $\square$

Il semble impossible d'exprimer les formules (2.1), (2.3) en terme d'une fonction connue, même dans le cas de *la recherche géométrique* des clients (c-à-d  $p_j = 1 - p^j$ ,  $p \in [0, 1]$ ,  $j \geq 1$ ). Cependant, nous supposons le cas de *la recherche constante*  $p_j = p$ ,  $p \in [0, 1]$ ,  $j \geq 1$ , pour obtenir des équations fermées. Premièrement, nous introduisons quelques notations élémentaire. Soit  $F$  la série hypergéométrique donnée par :

$$F(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!},$$

où  $(x)_k$  est le symbole de Pochhammer défini par :

$$(x)_k = \begin{cases} 1, & \text{si } k = 0, \\ x(x + 1)\dots(x + k - 1), & \text{si } k \geq 1. \end{cases}$$

Nous introduisons également les fonctions génératrice partielles :

$$P_i(z) = \sum_{j=0}^{\infty} z^j P_{ij}, i \in \{0, 1\}, |z| \leq 1,$$

et les moments factoriels partiels  $M_k^i$  définis par

$$M_0^i = \sum_{j=0}^{\infty} P_{ij}, i \in \{0, 1\},$$

$$M_k^i = \sum_{j=k}^{\infty} j(j-1)\dots(j-k+1)P_{ij}, i \in \{0, 1\}, k \geq 1.$$

Supposons que  $\{N(t); t \geq 0\}$  est positif récurrent, alors :

**Théorème 2.2.1.2.** -Les probabilités limites  $\{P_{ij}\}_{(i,j) \in s}$  sont données par :

$$P_{0j} = P_{00} \frac{(1-p)\lambda}{\lambda+\alpha} \rho^j \frac{\left(\frac{\lambda+\alpha}{\theta}\right)_j}{\left(\frac{p\lambda+\alpha}{\theta}+1\right)_j}, j \geq 1, \quad (2.6)$$

$$P_{1j} = P_{00} \rho^{j+1} \frac{\left(\frac{\lambda+\alpha}{\theta}+1\right)_j}{\left(\frac{p\lambda+\alpha}{\theta}+1\right)_j}, j \geq 0, \quad (2.7)$$

$$P_{00}^{-1} = F\left(1, \frac{\lambda+\alpha}{\theta}+1; \frac{p\lambda+\alpha}{\theta}+1; \rho\right). \quad (2.8)$$

-Les fonctions génératrice partielles  $P_i(z), 0 \leq i \leq 1$ , sont données par :

$$P_0(z) = P_{00}(1-\rho z)F\left(1, \frac{\lambda+\alpha}{\theta}+1; \frac{p\lambda+\alpha}{\theta}+1; \rho z\right). \quad (2.9)$$

$$P_1(z) = P_{00}\rho F\left(1, \frac{\lambda+\alpha}{\theta}+1; \frac{p\lambda+\alpha}{\theta}+1; \rho z\right) \quad (2.10)$$

-Les moments factoriels partiels  $M_k^i, i \in \{0, 1\}, k \geq 0$ , sont donnés par :

$$M_0^0 = 1 - \rho,$$

$$M_k^0 = P_{00} k! \frac{(1-p)\lambda}{\lambda+\alpha} \rho^k \frac{\left(\frac{\lambda+\alpha}{\theta}\right)_k}{\left(\frac{p\lambda+\alpha}{\theta}+1\right)_k} F\left(k+1, \frac{\lambda+\alpha}{\theta}+k; \frac{p\lambda+\alpha}{\theta}+k+1; \rho\right), k \geq 1,$$

$$M_0^1 = \rho,$$

$$M_k^1 = P_{00} k! \rho^{k+1} \frac{\left(\frac{\lambda+\alpha}{\theta}+1\right)_k}{\left(\frac{p\lambda+\alpha}{\theta}+1\right)_k} F\left(k+1, \frac{\lambda+\alpha}{\theta}+k+1; \frac{p\lambda+\alpha}{\theta}+k+1; \rho\right), k \geq 1,$$

*Démonstration.* Pour le cas  $p_j = p, j \geq 1$ , les formules (2.1) et (2.3) sont réduites à (2.6) et (2.8). On utilise les fonctions génératrice, nous obtenons (2.10) et

$$P_0(z) = \frac{P_{00}}{\lambda+\alpha} \left( p\lambda + \alpha + (1-p)\lambda F\left(1, \frac{\lambda+\alpha}{\theta}; \frac{p\lambda+\alpha}{\theta}+1; \rho z\right) \right). \quad (2.11)$$

Après un certain réarrangement (2.11) donne l'expression alternative (2.9). La clé pour calculer les moments factoriels partiels est l'identité suivante :

$$P_i(1+z) = \sum_{k=0}^{\infty} M_k^i \frac{z^k}{k!}, i \in \{0, 1\}.$$

Après avoir remplacé  $(1+z)^j$  par  $\sum_{k=0}^j \binom{j}{k} z^k$ , nous pouvons obtenir  $M_k^i$  par une identification directe des coefficients de la série  $P_i(1+z)$ .  $\square$

Nous donnons ensuite les expressions correspondant à la discipline de rappels classique et constante.

**Corollaire 2.2.1.1.** (Discipline de rappels classique) : Considérons  $\alpha = 0$  et  $\theta > 0$ , alors -Les probabilités limites sont données par :

$$\begin{aligned} P_{0j} &= P_{00}(1-p)\rho^j \frac{\left(\frac{\lambda}{\theta}\right)_j}{\left(\frac{\rho\lambda}{\theta} + 1\right)_j}, \quad j \geq 1, \\ P_{1j} &= P_{00}\rho^{j+1} \frac{\left(\frac{\lambda}{\theta} + 1\right)_j}{\left(\frac{\rho\lambda}{\theta} + 1\right)_j}, \quad j \geq 0, \\ P_{00}^{-1} &= F\left(1, \frac{\lambda}{\theta} + 1; \frac{\rho\lambda}{\theta} + 1; \rho\right). \end{aligned}$$

-Les fonctions génératrices partielles sont données par :

$$\begin{aligned} P_0(z) &= P_{00}(1-\rho z)F\left(1, \frac{\lambda}{\theta} + 1; \frac{\rho\lambda}{\theta} + 1; \rho z\right), \\ P_1(z) &= P_{00}\rho F\left(1, \frac{\lambda}{\theta} + 1; \frac{\rho\lambda}{\theta} + 1; \rho z\right). \end{aligned}$$

-Les moments factoriels partiels sont donnés par :

$$\begin{aligned} M_0^0 &= 1 - \rho, \\ M_k^0 &= P_{00}k!(1-p)\rho^k \frac{\left(\frac{\lambda}{\theta}\right)_k}{\left(\frac{\rho\lambda}{\theta} + 1\right)_k} F\left(k+1, \frac{\lambda}{\theta} + k; \frac{\rho\lambda}{\theta} + k+1; \rho\right), \quad k \geq 1, \\ M_0^1 &= \rho, \\ M_k^1 &= P_{00}k!\rho^{k+1} \frac{\left(\frac{\lambda}{\theta} + 1\right)_k}{\left(\frac{\rho\lambda}{\theta} + 1\right)_k} F\left(k+1, \frac{\lambda}{\theta} + k+1; \frac{\rho\lambda}{\theta} + k+1; \rho\right), \quad k \geq 1. \end{aligned}$$

**Corollaire 2.2.1.2.** (Discipline de rappels constante) : Considérons  $\alpha > 0$ ,  $\theta = 0$ , et soit  $\beta = \frac{\rho(\lambda+\alpha)}{\tau\lambda+\alpha} < 1$  (voir la proposition 2.2.2.3) alors

-Les probabilités limites sont données par :

$$\begin{aligned} P_{0j} &= P_{00} \frac{(1-p)\lambda}{\lambda + \alpha} \beta^j, \quad j \geq 1, \\ P_{1j} &= P_{00} \rho \beta^j, \quad j \geq 0, \\ P_{00} &= 1 - \beta. \end{aligned}$$

-Les fonctions génératrices partielles sont données par :

$$\begin{aligned} P_0(z) &= \frac{(1-pz)(1-\beta)}{1-\beta z}, \\ P_1(z) &= \frac{\rho(1-\beta)}{1-\beta z}. \end{aligned}$$

-Les moments factoriels partiels sont donnés par :

$$\begin{aligned} M_0^0 &= 1 - \rho, \\ M_k^0 &= k! \frac{(1-p)\lambda}{\lambda + \alpha} \left( \frac{\beta}{1-\beta} \right)^k, \quad k \geq 1, \\ M_0^1 &= \rho, \\ M_k^1 &= k! \rho \left( \frac{\beta}{1-\beta} \right)^k, \quad k \geq 1. \end{aligned}$$

Pour le choix du  $p_j = 1, j \geq 1$ , et  $p_j = 0, j \geq 1$ , on peut déduire les mesures de performances de la file d'attente  $M/M/1$  classique et ceux du  $M/M/1$  avec rappels.

## 2.2.2 Modèle d'attente $M/G/1$ avec rappels et recherche des clients

Cette partie est consacrée à l'étude du système de files d'attente  $M/G/1$  avec rappel et recherche de clients en orbite, ce système est composé d'un seul serveur avec des arrivées de clients Poissonniennes, et les durées de service suivent une loi générale. La première étude de ce système était faite par Artalejo et al. [4], puis d'autres extensions ont été réalisées par Joushua dans [77]. Nous décrivons d'abord le modèle associé à ce type de files d'attente puis nous déduisons les équations de Chapman Kolmogorov et nous calculons les mesures de performance de la file  $M/G/1$  avec rappel exponentiel.

### Description du modèle

Pour ce modèle, qui est composé d'un seul serveur. Le processus d'arrivée des clients dans le système est de Poisson de taux  $\lambda > 0$  et la durée de service  $\mu$  est de loi générale  $B(x)$  et de transformée de Laplace-Stieltjes  $\tilde{B}(s), Re(s) > 0$ . Soient les moments  $\beta_k = (-1)^k \tilde{B}^{(k)}(0)$ , l'intensité du trafic est  $\rho = \lambda \beta_1, b(x) = \frac{B'(x)}{1-B(x)}$  est l'intensité instantané du service donné quand le temps de service écoulé est  $x$ . Soit  $\eta_j$  l'instant de l'achèvement du service du  $j^{ième}$  client,

immédiatement après cet instant le serveur part pour une recherche de clients en orbite avec une probabilité  $p_j$ , avec  $q_j = 1 - p_j$  le serveur reste libre. La durée entre deux rappels successifs d'une même source secondaire est linéaire de taux  $\alpha(1 - \delta_{j0}) + j\theta$ . Le temps de recherche est supposé être *négligeable*.

Les flux d'arrivées primaires, d'inter-rappels et du temps de service sont supposés être mutuellement indépendants.

### Chaîne de Markov induite

Considérons le processus  $\{C(t), N(t), \xi(t); t \geq 0\}$ ,  $\xi(t)$  représente le temps de service écoulé du client servi ( $\xi(t) = 0$  si  $C(t) = 0$ ). Soit  $N(t)$  le nombre de clients en orbite et  $C(t)$  représente l'état du serveur. On a  $C(t)$  égal à 1 ou 0 selon que le serveur est occupé ou libre. On remarque que la séquence  $N_n = N(\eta_n^+)$  forme une chaîne de Markov induite, dont l'équation fondamentale est :

$$N_{n+1} = N_n - \delta_{N_n} + v_{n+1}.$$

Où  $v_{n+1}$  est le nombre de clients primaires arrivant dans le système durant le service du  $(n+1)^{\text{ème}}$  client. La distribution de  $v_{n+1}$  est la suivante :

$$P(v_n = i) = k_i = \int_0^{\infty} \frac{(\lambda x)^i}{i!} e^{-\lambda x} dB(x),$$

où  $k_i > 0, i \geq 0$ . On a les résultats suivants

$$\text{si } v = \lim_{n \rightarrow \infty} v_n, E[v] = \rho, k(z) = \sum_{i=0}^{\infty} k_i z^i = \tilde{B}(\lambda - \lambda z).$$

La variable aléatoire  $\delta_{N_n}$  est une variable de Bernoulli définie par

$$\delta_{N_n} = \begin{cases} 1 & \text{si le } (n+1)^{\text{ème}} \text{ client servi provient de l'orbite,} \\ 0 & \text{si le } (n+1)^{\text{ème}} \text{ client servi est primaire.} \end{cases}$$

**Proposition 2.2.2.1.** *Supposons que  $\lim_{j \rightarrow \infty} p_j$  existe. Si  $\theta > 0$ , alors  $\{N_n\}_{n=1}^{\infty}$  est positif récurrent si et seulement si  $\rho < 1$ .*

*Démonstration.* Noter que  $\{N_n\}_{n=1}^{\infty}$  est irréductible et apériodique donc pour étudier la récurrence positive nous utiliserons le critère de Foster qui dit qu'une chaîne de Markov irréductible et apériodique est récurrente positivement s'il existe une fonction non négative  $f(j), j \in \mathbb{N}$ , et  $\varepsilon > 0$  telle que l'accroissement moyen de la chaîne vaut  $\varphi_j = E[f(N_{n+1}) - f(N_n) / N_n = j]$  est fini pour tout  $j \in \mathbb{N}$ , et  $\varphi_j \leq -\varepsilon$  sauf peut-être pour un nombre fini. En choisissant la fonction  $f(j) = j$ , on obtient :

$$\begin{aligned} \varphi_j &= E[N_{n+1} - N_n / N_n = j] = E[-\delta_{N_n} + v_{n+1} / N_n = j] \\ &= \begin{cases} \rho & \text{si } j = 0, \\ \rho - \left( p_j + q_j \frac{\alpha + j\theta}{\lambda + \alpha + j\theta} \right) & \text{si } j \geq 1. \end{cases} \end{aligned}$$

On a  $\lim_{j \rightarrow \infty} \varphi_j = \rho - 1$ . Ainsi,  $\rho < 1$ , est suffisant pour la récurrence positive. Pour étudier la non-ergodicité, il suffit de montrer que  $\{N_n\}_{n=1}^{\infty}$  est non-ergodique si la condition de Kaplan est réalisée,  $\varphi_j < \infty$ , pour  $j \geq j_0$ . Si  $\rho \geq 1$ ; il est évident que  $\varphi_j \geq 0$ , pour  $j \geq 1$  car il existe un indice  $k$  tel que  $P_{ij} = 0$ , pour  $j < i - k, i > 0$ .  $\square$

**Proposition 2.2.2.2.** Si  $\lim_{j \rightarrow \infty} p_j = 1$  alors  $\{N_n\}_{n=1}^{\infty}$  est récurrent positif si et seulement si  $\rho < 1$ .

*Démonstration.* On trouve facilement que  $\lim_{j \rightarrow \infty} \varphi_j = \rho - 1$ , donc le critère de Foster vérifie encore que  $\rho < 1$  est suffisante pour la récurrence positive. La nécessité suit de l'argument donné dans la proposition [2.2.2.1](#).  $\square$

Nous analysons également le cas  $\theta = 0$  et  $p_j = r$ , pour  $j \geq 1$ .

**Proposition 2.2.2.3.** Si  $\theta = 0$ ,  $\alpha > 0$  et  $p_j = r$ , pour  $j \geq 1$  alors  $\{N_n\}_{n=1}^{\infty}$  est récurrent positif si et seulement si  $\beta = \frac{\rho(\lambda + \alpha)}{r\lambda + \alpha} < 1$ .

*Démonstration.* L'accroissement moyen donné par

$$\varphi_j = \rho - \frac{r\lambda + \alpha}{\lambda + \alpha}, j \geq 1$$

Alors comme dans des propositions précédentes, on obtiens la condition.  $\square$

### Distribution stationnaire de l'état du système

Les transitions entre les états sont illustrées dans la figure [2.3](#).

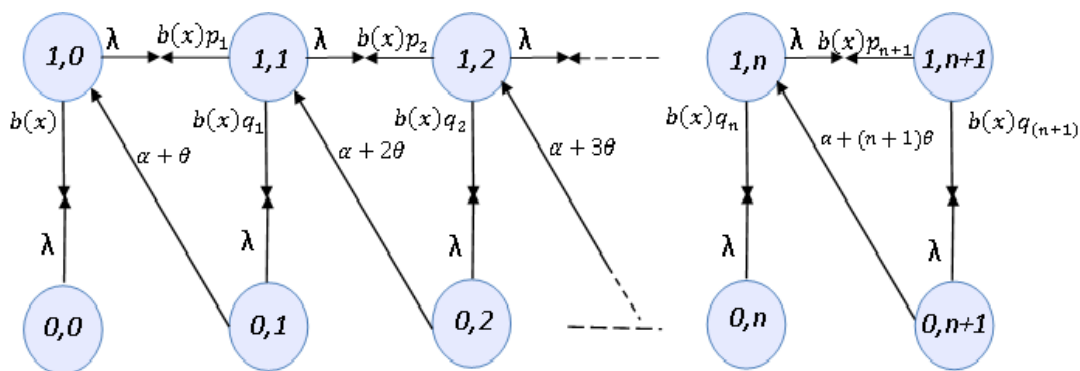


FIGURE 2.3 – Graphe des transitions du modèle  $M/G/1$  avec rappels et recherche de clients en orbite.

Pour simplifier, nous considérons  $\alpha = 0$  (rappels classique) et  $p_j = p, p \in [0, 1], j \geq 1$ .



**Théorème 2.2.2.1.** Si  $\rho < 1$ , et le système est à l'état stationnaire ; alors la distribution jointe de l'état du serveur et de la taille de la file d'attente (orbite)

$$\begin{aligned} P_{0j} &= P[C(t) = 0, N(t) = j], \\ P_{1j}(x) &= \frac{\partial}{\partial x} P[C(t) = 1, \xi(t) \leq x, N(t) = j], \end{aligned}$$

ont les fonctions génératrices partielles suivantes :

$$\begin{aligned} P_0(z) &= \sum_{j=0}^{\infty} P_{0j} z^j \\ &= \frac{\lambda p P_{00}}{\theta} (z)^{-\frac{\lambda p}{\theta}} \int_0^z (t)^{\frac{\lambda p}{\theta}-1} s(z, t) dt. \end{aligned} \quad (2.12)$$

$$\begin{aligned} P_1(z, x) &= \sum_{j=0}^{\infty} P_{1j}(x) z^j \\ &= \frac{\lambda(1-z)}{\tilde{B}(\lambda - \lambda z) - z} P_0(z) [1 - B(x)] e^{-\lambda(1-z)x}. \end{aligned} \quad (2.13)$$

où

$$s(t_1, t_2) = \exp \left\{ \frac{\lambda(1-p)}{\theta} \int_{t_1}^{t_2} \frac{1 - \tilde{B}(\lambda - \lambda u)}{u - \tilde{B}(\lambda - \lambda u)} du \right\}. \quad (2.14)$$

$$P_{00} = \frac{\theta(1 - \lambda\beta_1) s(0, 1)}{\lambda p \int_0^1 (t)^{\frac{\lambda p}{\theta}-1} dt}. \quad (2.15)$$

Dans le cas  $C(t) = 1$ , on néglige le temps de service écoulé  $\xi(t)$ , alors pour les probabilités  $P_{1j} = P[C(t) = 1, N(t) = j]$ ,

$$\begin{aligned} P_1(z) &= \sum_{j=0}^{\infty} P_{1j} z^j, \\ &= \frac{1 - \tilde{B}(\lambda - \lambda z)}{\tilde{B}(\lambda - \lambda z) - z} P_0(z). \end{aligned} \quad (2.16)$$

*Démonstration.* L'ensemble des équations d'équilibre statistique est obtenu comme suit :

$$\begin{aligned} (\lambda + j\theta) P_{0j} &= [1 - (1 - \delta_{j0})p] \int_0^{\infty} P_{1j}(x) b(x) dx, \\ P'_{1j}(x) &= -(\lambda + b(x)) P_{1j}(x) + \lambda P_{1,j-1}(x) \\ P_{1j}(0) &= \lambda P_{0j} + (j+1)\theta P_{0,j+1} + p \int_0^{\infty} P_{1,j+1}(x) b(x) dx, \end{aligned}$$

A l'aide des fonctions génératrices, ces équations sont transformées en :

$$\lambda P_0(z) + \theta z P_0'(z) = (1-p) \int_0^\infty P_{1j}(z, x) b(x) dx + \lambda p P_{00}, \quad (2.17)$$

$$\frac{\partial P_1(z, x)}{\partial x} = -(\lambda - \lambda z + b(x)) P_1(z, x), \quad (2.18)$$

$$(z + p\tilde{B}(\lambda - \lambda z)) P_1(z, 0) + \lambda p P_{00} = \theta z P_0'(z) + \lambda z P_0(z). \quad (2.19)$$

La résolution de (2.18) donne

$$P_1(z, x) = P_1(z, 0) (1 - B(x)) e^{-\lambda(1-z)x} \quad (2.20)$$

En combinant (2.17), (2.19) et (2.20), on obtient :

$$\begin{aligned} & \theta z (z - \tilde{B}(\lambda - \lambda z)) P_0'(z) + (\lambda z - ((1-p)\lambda z + \lambda p)\tilde{B}(\lambda - \lambda z)) P_0(z) \\ &= \lambda p P_{00} (z - \tilde{B}(\lambda - \lambda z)) \end{aligned} \quad (2.21)$$

Le coefficient de  $P_0'(z)$  à deux zéros  $z = 0$  et  $z = 1$ , il suffit de choisir un point quelconque  $a \in ]0, 1[$  puis résoudre (2.21) pour  $z \in ]0, a]$ , on obtient :

$$P_0(z) = \left[ \left( \frac{z}{a} \right)^{\frac{\lambda p}{\theta}} s(a, z) \right]^{-1} \left\{ P_0(a) + \frac{\lambda p P_{00}}{\theta(a)^{\frac{\lambda p}{\theta}}} \int_a^z (t)^{\frac{\lambda p}{\theta}-1} s(a, t) dt \right\}$$

Comme  $z \rightarrow 0^+$ ,  $P_0(0) < \infty$  et  $\left(\frac{z}{a}\right)^{-\frac{\lambda p}{\theta}}$  diverge. Ainsi,

$$P_0(a) = \frac{\lambda p P_{00}}{\theta(a)^{\frac{\lambda p}{\theta}}} \int_0^a (t)^{\frac{\lambda p}{\theta}-1} s(a, t) dt \quad (2.22)$$

D'autre part la résolution de (2.21) pour  $z \in [a, 1[$ , et en prenant comme limite  $z \rightarrow 1^-$ , on obtient :

$$P_0(a) = \frac{P_0(1) s(a, 1)}{(a)^{\frac{\lambda p}{\theta}}} - \frac{\lambda p P_{00}}{\theta(a)^{\frac{\lambda p}{\theta}}} \int_a^1 (t)^{\frac{\lambda p}{\theta}-1} s(a, t) dt. \quad (2.23)$$

Pour obtenir la relation (2.23), il faut prendre en considération que :

$$\lim_{z \rightarrow 1^+} \frac{1 - \tilde{B}(\lambda - \lambda z)}{\tilde{B}(\lambda - \lambda z) - z} = \frac{\lambda \beta_1}{1 - \lambda \beta_1} < \infty.$$

En égalant (2.22) et (2.23) on obtient :

$$P_0(1) = \frac{\lambda p P_{00}}{\theta s(0, 1)} \int_0^1 (t)^{\frac{\lambda p}{\theta}-1} s(0, t) dt. \quad (2.24)$$

On peut alors réécrire la solution de (2.21) sous la forme (2.12).

En combinant les équations (2.19), (2.20) et (2.21), nous obtenons (2.13), puisque  $P_1(z) = \int_0^\infty P_1(z, x) dx$ , on obtient (2.16).

Appliquons maintenant la condition de normalisation  $P_0(1) + P_1(1) = 1$ , on aura :

$$P_0(1) = 1 - \lambda\beta_1 = 1 - \rho. \quad (2.25)$$

En utilisant (2.24) et (2.25), on obtient l'expression de  $P_{00}$  comme dans (2.15).  $\square$

**Corollaire 2.2.2.1.** Les moments factoriels partiels  $M_k^i, i \in \{0, 1\}, k \in \{0, 1\}$  sont donnés par :

$$\begin{aligned} M_0^0 &= 1 - \rho, \\ M_0^1 &= \rho, \\ M_1^0 &= \frac{\lambda p P_{00}}{\theta} - \frac{\lambda(\rho - p)}{\theta}, \\ M_1^1 &= \frac{\lambda^2}{(1 - \rho)\theta} \{ \theta\beta_2 + \lambda\beta_1^2 - p\beta_1(1 - P_{00}) \}. \end{aligned}$$

### Mesures de performance du système

- La fonction de probabilité de masse de l'état du serveur

$$\begin{aligned} \Pr[\text{le serveur est libre}] &= M_0^0 = 1 - \rho, \\ \Pr[\text{le serveur est occupé}] &= M_0^1 = \rho. \end{aligned}$$

En particulier, la probabilité de blocage (ie, la probabilité qu'un client primaire soit bloqué)  $= M_0^1 = \rho$ .

- La fonction de probabilité de masse du nombre de clients en orbite

$$\Pr[\text{il y a } i \text{ clients dans l'orbite}] = P_{0i} + P_{1i}.$$

- Le nombre moyen de clients en orbite

$$\bar{n} = M_1^0 + M_1^1.$$

- La période d'occupation prévue  $E(L)$  (ie, la période qui commence à l'instant où un client qui arrive trouve le système vide et se termine à l'instant de son départ c-à-d le système est à nouveau vide).

$$E(L) = \frac{1}{\lambda} \frac{(1 - P_{00})}{P_{00}}.$$

- Le taux global des rappels  $\tilde{\theta}_1$  aux quels les clients en orbite demandent un service est donné par

$$\tilde{\theta}_1 = \theta \sum_{i=0}^{\infty} i (P_{0i} + P_{1i}) = \theta \bar{n}.$$

- La vitesse à laquelle les clients en orbite atteignent le serveur avec succès est donnée par

$$\tilde{\theta}_2 = \theta \sum_{i=0}^{\infty} iP_{0i} = \theta M_1^0.$$

- La fraction du taux de réussite des rappels est donnée par

$$\frac{\tilde{\theta}_2}{\tilde{\theta}_1} = \frac{M_1^0}{\bar{n}}.$$

## 2.3 Conclusion

Nous avons établi l'analyse stochastique complète des modèles  $M/M/1$  et  $M/G/1$  avec la discipline de rappels linéaire et recherche des clients en orbite par le serveur ( le temps de recherche est considéré négligeable), nous avons mis en évidence la chaîne de Markov induite associée au modèles. Par la méthode des variables supplémentaires et à l'aide du graphe des transitions nous avons obtenu les fonctions génératrices de la distribution conjointe de l'état du serveur et de la taille de l'orbite des deux modèles, ainsi que les fonctions génératrice de la distribution stationnaire du nombre de clients dans le système et la distribution marginale du nombre de serveurs occupés. Quelques résultats comparatifs avec le système standard et celui avec rappels classique et constant ont été calculer. Enfin, nous avons établi aussi les mesures de performances de ces modèles.

Dans le prochain chapitre nous étudions un modèle d'attente  $M/M/1$  avec rappels classique et temps de recherche des clients en orbite non négligeable, de plus plusieurs résultats numérique serrent présenté.

# Chapitre 3

## Modèle d'attente $M/M/1$ avec rappels et temps de recherche des clients exponentiel

### 3.1 Introduction

Dans ce chapitre, un modèle de files d'attente avec rappels et recherche des clients en orbite est proposé par Tuan Phung [66] est motivé par les systèmes d'informatique en nuage ( Cloud Computing) où l'unité de traitement et l'unité de stockage sont séparés. L'unité de traitement à la capacité de servir une seule tâche à chaque fois. Nous allons voir comment les fournisseurs de services Cloud peuvent garantir aux utilisateurs de service Cloud la qualité de service appropriée, en considérant le temps de recherche des clients en orbite *non négligeable*.

Nous proposons de notre part, plusieurs résultats numériques pour étudier l'effet de certains paramètres clés sur les caractéristiques du modèle, ce travail a été présenté en 2020 [20].

### 3.2 Description du modèle

Les clients arrivent au serveur selon un processus Poissonnien de taux  $\lambda$ . Le temps de service des clients entrants suit la loi de distribution exponentielle avec une moyenne  $1/\mu$ . Après l'achèvement du service, le serveur reste inactif pendant une durée de temps qui suit la loi exponentielle avec une moyenne  $1/\alpha$ . Pendant cette période d'inactivité, un client arrivant (soit primaire ou secondaire ) est immédiatement servi. Après le temps de repos, le serveur commence la chercher d'un client en orbite. Le temps de recherche suit la distribution exponentielle avec la moyenne  $1/\nu$ . Les clients qui arrivent et trouvent le serveur occupé (par un service d'un client ou part la recherche) rejoignent l'orbite. Ainsi, les clients de l'orbite tentent d'attendre le serveur après un certain temps exponentiellement distribué avec moyenne  $1/\theta$ . Ce modèle été analysé pour la premières fois dans la littérature dans [66].

### 3.3 Etude analytique du système

Soit  $C(t)$  l'état du serveur à l'instant  $t$ ,  $t \geq 0$ .

$$C(t) = \begin{cases} 0 & \text{le serveur est libre,} \\ 1 & \text{le serveur est occupé,} \\ 2 & \text{le serveur recherche un client.} \end{cases}$$

$N(t)$  est le nombre de clients dans l'orbite à l'instant  $t$ ,  $t \geq 0$ . Nous avons  $\{X(t) = (C(t), N(t)), t \geq 0\}$  qui forme une chaîne de Markov sur l'espace d'état  $S = \{0, 1, 2\} \times \{0, 1, 2, \dots\}$ . Voir la figure 3.1, pour les transitions entre les états. Nous supposons que le système est stable, c'est-à-dire que la distribution stationnaire existe [19].

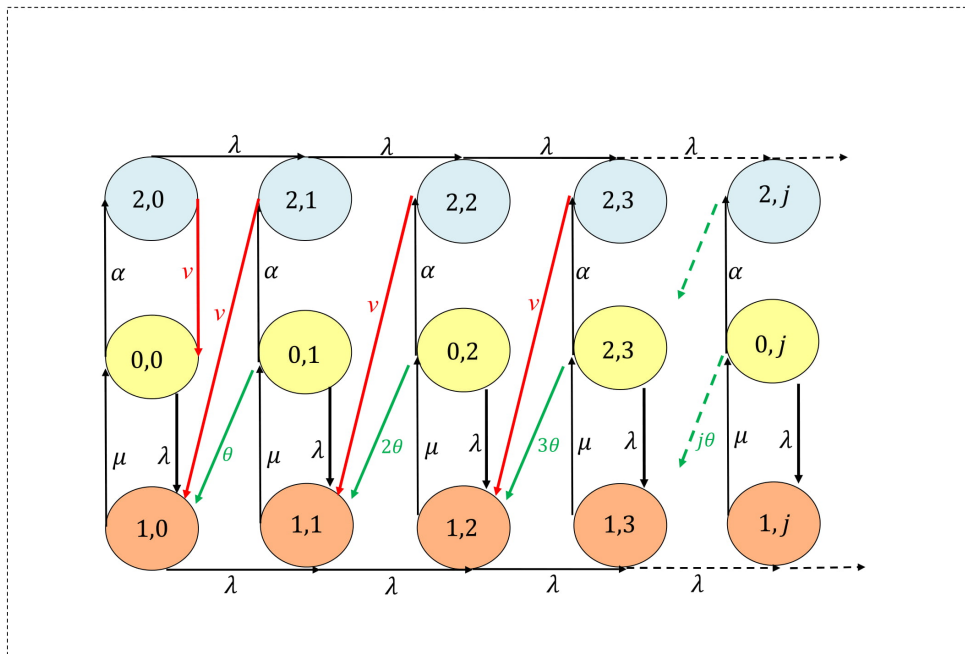


FIGURE 3.1 – Graphe des transitions du modèle.

La condition nécessaire et suffisante pour que le système soit stable est  $\lambda < \mu$  qui sera obtenu plus tard dans l'analyse.

$$p_{ij} = \lim_{t \rightarrow \infty} p(C(t) = i, N(t) = j),$$

Les équations d'équilibre de balance pour les états  $(i, j)$  sont indiqués comme suit :

$$(\lambda + \alpha)p_{0,0} = \mu p_{1,0} + \nu p_{2,0}, \quad (3.1)$$

$$(\lambda + \alpha + j\theta)p_{0,j} = \mu p_{1,j} \quad j \geq 1, \quad (3.2)$$

$$\begin{aligned} (\lambda + \mu)p_{1,j} &= (j + 1)\theta p_{0,j+1} + \nu p_{2,j+1} + \lambda p_{1,j-1} + \lambda p_{0,j} \quad j \geq 0, \\ (\lambda + \nu)p_{2,j} &= \alpha p_{0,j} + \lambda p_{2,j-1} \quad j \geq 0, \end{aligned} \quad (3.3)$$

où  $p_{i,-1} = 0 (i = 1, 2)$ . Soient  $P_i(z)$  les fonctions génératrice de  $p_{i,j}$ , ie.  $P_i(z) = \sum_{j=0}^{\infty} p_{i,j} z^j$  ( $i = 0, 1, 2$ ). On Transforment les équations d'équilibre ci-dessus en fonctions génératrice nous obtenons :

$$(\lambda + \alpha)P_0(z) + \theta z P_0'(z) = \mu P_1(z) + \nu p_{2,0}, \quad (3.4)$$

$$(\lambda + \mu)P_1(z) = \theta P_0'(z) + \frac{\nu}{z}(P_2(z) - p_{2,0}) + \lambda z P_1(z) + \lambda P_0(z), \quad (3.5)$$

$$(\lambda + \nu)P_2(z) = \alpha P_0(z) + \lambda z P_2(z). \quad (3.6)$$

En résumant les équations ci-dessus et organisent les rendements des résultats en aura :

$$\lambda(P_1(z) + P_2(z)) = \mu P_0'(z) + \frac{\nu(P_2(z) - p_{2,0})}{z}. \quad (3.7)$$

Cette équation représente l'équilibre entre les entrées et les sorties de l'orbite. A partir de (3.4) et (3.6), nous obtenons :

$$P_1(z) = \frac{(\lambda + \alpha)P_0(z) + \mu z P_0'(z) - \nu p_{2,0}}{\mu}, \quad (3.8)$$

$$P_2(z) = \frac{\alpha P_0(z)}{\lambda + \nu - \lambda z}. \quad (3.9)$$

Substituons ces deux expressions dans l'équation d'équilibre de l'orbite (3.7) et organisons les rendements du résultat en aura :

$$P_0'(z) = A(z)P_0(z) + B(z). \quad (3.10)$$

où

$$A(z) = \frac{\frac{\lambda(\lambda + \alpha)}{\mu} + \frac{\alpha(\lambda - \frac{\nu}{z})}{\lambda + \nu - \lambda z}}{\theta \left(1 - \frac{\lambda z}{\theta}\right)}, \quad B(z) = \frac{p_{2,0}\nu}{\theta z}.$$

Nous décomposons  $A(z)$  comme suit :

$$A(z) = \frac{a}{z} + \frac{b}{1 - \frac{\lambda z}{\mu}} + \frac{c}{1 - \frac{\lambda z}{\lambda + \nu}},$$

où  $a, b$  et  $c$  sont donnés par :

$$a = -\frac{\alpha\nu}{\theta(\lambda + \nu)}, \quad b = \frac{\lambda^2(\lambda + \alpha + \nu - \mu)}{\theta\mu(\lambda + \nu - \mu)}, \quad c = \frac{\lambda^2\alpha\mu}{(\lambda + \nu)^2\theta(\mu - \lambda - \nu)}.$$

nous résolvons d'abord l'équation différentielle non homogène ;

$$P_0'(z) = A(z)P_0(z),$$

qui se transforme en ;

$$\frac{P_0'(z)}{P_0(z)} = \frac{a}{z} + \frac{b}{1 - \frac{\lambda z}{\mu}} + \frac{c}{1 - \frac{\lambda z}{\lambda + \nu}}.$$

La solution de cette équation différentielle est donnée par :

$$P_0(z) = Cz^a \left( \frac{\mu - \lambda}{\mu - \lambda z} \right)^{\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda z} \right)^{\frac{c(\lambda + \nu)}{\lambda}},$$

où  $C$  est un nombre constant. Comme d'habitude, nous avons la solution pour notre équation différentielle original (3.10) sous la forme suivante :

$$P_0(z) = C(z)z^a \left( \frac{\mu - \lambda}{\mu - \lambda z} \right)^{\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda z} \right)^{\frac{c(\lambda + \nu)}{\lambda}},$$

où  $C(z)$  est une fonction inconnue. Substituons dans l'équation différentielle (3.10) les rendements on aura :

$$C'(z)z^a \left( \frac{\mu - \lambda}{\mu - \lambda z} \right)^{\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda z} \right)^{\frac{c(\lambda + \nu)}{\lambda}} = \frac{p_{2,0}\nu}{\theta z},$$

ou d'une manière équivalente :

$$C'(z) = \frac{p_{2,0}\nu}{\theta} z^{-(a+1)} \left( \frac{\mu - \lambda}{\mu - \lambda z} \right)^{-\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda z} \right)^{-\frac{c(\lambda + \nu)}{\lambda}}.$$

Donc, nous avons

$$C(z) = C_0 - \frac{p_{2,0}\nu}{\theta} \int_z^1 u^{-(a+1)} \left( \frac{\mu - \lambda}{\mu - \lambda u} \right)^{-\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda u} \right)^{-\frac{c(\lambda + \nu)}{\lambda}} du,$$

où  $C_0$  est un nombre constant. Puisque  $P_0(z)$  est analytique en  $z = 0$  et  $a < 0$ , nous devons avoir  $C(0) = 0$  d'où :

$$C_0 = \frac{p_{2,0}\nu}{\theta} \int_0^1 u^{-(a+1)} \left( \frac{\mu - \lambda}{\mu - \lambda u} \right)^{-\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda u} \right)^{-\frac{c(\lambda + \nu)}{\lambda}} du.$$

La solution finale de  $P_0(z)$  est donnée par :

$$P_0(z) = \frac{p_{2,0}\nu}{\theta} z^a \left( \frac{\mu - \lambda}{\mu - \lambda z} \right)^{\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda z} \right)^{\frac{c(\lambda + \nu)}{\lambda}} \times \int_0^z u^{-(a+1)} \left( \frac{\mu - \lambda}{\mu - \lambda u} \right)^{-\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda u} \right)^{-\frac{c(\lambda + \nu)}{\lambda}} du. \quad (3.11)$$



De (3.7), (3.9) et (3.10), nous obtenons :

$$P_1(1) + P_2(1) = \left( \frac{\theta}{\lambda} A(1) + \frac{\alpha}{\lambda} \right) P_0(z). \quad (3.12)$$

Nous avons aussi la condition de normalisation :

$$P_0(1) + P_1(1) + P_2(1) = 1. \quad (3.13)$$

De (3.12) et (3.13), nous obtenons :

- La probabilité que le serveur est libre mais le système n'est pas vide  $P_0(1)$  :

$$P_0(1) = \frac{\nu(1 - \frac{\lambda}{\mu})}{\alpha + \nu},$$

où l'expression de  $A(1)$  en termes de paramètres donnés est utilisée dans (3.8) et (3.9) :

- La probabilité que le serveur est occupé  $P_1(1)$  :

$$P_1(1) = \frac{\lambda}{\nu},$$

- La probabilité que le serveur est en recherche  $P_2(1)$  :

$$P_2(1) = \frac{\alpha(1 - \frac{\lambda}{\mu})}{\alpha + \nu}.$$

Par conséquent, de l'expression  $P_0(z)$ , nous obtenons l'expression de  $p_{2,0}$  comme suit :

$$p_{2,0} = \frac{\theta(1 - \frac{\lambda}{\mu})}{(\alpha + \nu) \int_0^1 u^{-(a+1)} \left( \frac{\mu - \lambda}{\mu - \lambda u} \right)^{-\frac{b\mu}{\lambda}} \left( \frac{\nu}{\lambda + \nu - \lambda u} \right)^{-\frac{c(\lambda + \nu)}{\lambda}} du}. \quad (3.14)$$

De cette expression, nous obtenons le fait que la condition de stabilité pour le modèle est  $\lambda < \mu$ .

## 3.4 Formules récursives

Nous allons dériver un schéma récursif de la distribution stationnaire. A partir de l'équation d'équilibre orbital, on obtient :

$$\lambda(p_{1,j} + p_{2,j}) = (j + 1)\theta p_{0,j+1} + \nu p_{2,j+1}.$$

De cette équation et de (3.3) avec  $j := j + 1$ , on obtient :

$$\lambda p_{1,j} + \frac{\lambda^2}{\lambda + \nu} p_{2,j} = \left( (j + 1)\theta + \frac{\alpha\nu}{\nu + \lambda} \right) p_{0,j+1}$$

Par conséquent, nous avons le schéma récursif suivant pour la distribution stationnaire.

$$p_{0,j} = \frac{\lambda [(\lambda + \nu) p_{1,j-1} + \lambda p_{2,j-1}]}{j\theta(\lambda + \nu) + \alpha\nu}, j \geq 1.$$

$$p_{1,j} = \frac{(\lambda + \alpha + j\theta) p_{0,j}}{\mu}, j \geq 1.$$

$$p_{2,j} = \frac{\alpha p_{0,j} + \lambda p_{2,j-1}}{\lambda + \nu}, j \geq 1.$$

Où  $p_{0,0}$ ,  $p_{1,0}$  et  $p_{2,0}$  sont donnés à l'avance. En particulier,  $p_{2,0}$  est obtenu par (3.14) et  $p_{0,0}$  s'obtient de (3.3) avec  $j = 0$  alors que  $p_{1,0}$  s'obtient en sommant (3.1) et (3.3) avec  $j = 0$ , i.e.,  $p_{1,0} = \frac{\lambda(p_{0,0} + p_{2,0})}{\mu}$ . Il est à noter que la deuxième et la troisième équations découlent de (3.2) et (3.3), respectivement.

**Remarque 3.4.0.1.** Cette formule récursive permet de calculer toute probabilité  $p_{i,j}$ . De plus, le schéma récursif peut être implémenté de manière numérique et symbolique.

**Remarque 3.4.0.2.** En prenant les dérivées en  $z = 1$  pour l'équation différentielle (3.10), nous pouvons obtenir  $P_0^{(n)}(1)$  pour tout  $n$ . Puisque  $P_1(1)$  et  $P_2(1)$  sont exprimées en termes de  $P_0(1)$ , on peut aussi calculer  $P_1^{(n)}(1)$  et  $P_2^{(n)}(1)$  pour tout  $n$ .

### 3.5 Cas limite

Nous étudions le cas où  $\nu \rightarrow \infty$  ce qui signifie qu'un client dans l'orbite est choisi part le serveur après un temps d'inactivité distribué de manière exponentielle avec une moyenne  $\frac{1}{\alpha}$ . Cela équivaut à la politique de rappels linéaire présentée dans [6].

Notamment, on observe que lorsque  $\nu \rightarrow \infty$ ,

$$a = \frac{\alpha}{\theta}, \quad b = \frac{\lambda^2}{\theta\mu}, \quad c = 0.$$

De plus,

$$\lim_{\nu \rightarrow \infty} P_2(z) = 0,$$

ce qui signifie que les états de recherche n'existent pas. Nous avons

$$\begin{aligned} \lim_{\nu \rightarrow \infty} \frac{\nu p_{2,0}}{\theta} &= \lim_{\nu \rightarrow \infty} \frac{\nu \left(1 - \frac{\lambda}{\mu}\right)}{(\lambda + \nu) \int_0^1 u^{-(a-1)} \left(\frac{\mu-\lambda}{\mu-\lambda u}\right)^{-\frac{b\mu}{\lambda}} \left(\frac{\nu}{\lambda+\nu-\lambda u}\right)^{-\frac{c(\lambda+\nu)}{\lambda}} du} \\ &= \frac{1 - \frac{\lambda}{\mu}}{\int_0^1 u^{-(a-1)} \left(\frac{\mu-\lambda}{\mu-\lambda u}\right)^{-\frac{b\mu}{\lambda}} du}. \end{aligned}$$

Ainsi, il découle de (3.11) que

$$P_0(z) = \left(1 - \frac{\lambda}{\mu}\right) z^{-\frac{\alpha}{\theta}} \left(\frac{\mu - \lambda}{\mu - \lambda z}\right)^{\frac{\lambda}{\theta}} \frac{\int_0^z u^{\frac{\alpha}{\theta}-1} \left(\frac{\mu-\lambda}{\mu-\lambda u}\right) du}{\int_0^1 u^{\frac{\alpha}{\theta}-1} \left(\frac{\mu-\lambda}{\mu-\lambda u}\right) du}.$$

Substitution (3.10) dans (3.8), on obtient :

$$P_1(z) = \frac{(\lambda + \alpha + \theta z A(z)) P_0(z)}{\mu}.$$

## 3.6 Analyse de sensibilité des mesures de performance du modèle

Notre objectif est d'étudier le comportement des mesures de performance de ce système par rapport à quelques paramètres. Le calcul numérique est réalisé à l'aide du logiciel Matlab. Les différents résultats obtenus sont présentés dans des tableaux et des figures [18], [20].

Dans tout ce qui suit, on va présenter l'influence des paramètres sur :

$P_0$  : la probabilité que le serveur est libre mais le système n'est pas vide.

$P_1$  : la probabilité que le serveur est occupé.

$P_2$  : la probabilité que le serveur est en recherche.

$\bar{n}_0$  : le nombre moyen de clients dans l'orbite.

$\bar{n}_s$  : le nombre moyen de clients dans le système.

$W$  : le temps moyen de séjour des clients dans le système.

$W_q$  : le temps moyen d'attente des clients dans l'orbite.

### 3.6.1 L'effet du taux d'arrivée $\lambda$ sur le modèle

On fait varier le taux d'arrivée  $\lambda$  sur les mesures de performance pour  $\mu = 10$ ,  $\nu = 2.5$ ,  $\alpha = 3$ , et  $\theta = 1.34$ .

**Commentaire :**

À partir du tableau 3.1 et de la figure 3.2, nous pouvons voir que le taux des arrivées influence sur les probabilités de l'état du serveur  $P_0$ ,  $P_1$  ou elle décroissent de façon monotone, par contre la probabilité  $P_2$  est croissante. Pour le nombre moyen de clients dans le système et dans l'orbite, ont tendance à augmenter de façon monotone. Ce qui est conforme à notre prévision.

$\lambda$	$P_0$	$P_1$	$P_2$	$\bar{n}_0$	$\bar{n}_s$	$W$	$W_q$
1	0.4091	0.1	0.4909	2.9917	3.0917	3.0917	2.9917
1.2	0.4	0.12	0.48	4.5178	4.6178	3.8482	3.7648
1.4	0.3909	0.14	0.4691	6.1354	6.2354	4.4539	4.3824
1.6	0.3818	0.16	0.4582	7.7792	7.8792	4.9245	4.862
1.8	0.3727	0.18	0.4473	9.3653	9.4653	5.2585	5.2029
2	0.3636	0.2	0.4364	10.7964	10.8964	5.4482	5.3982
2.2	0.3545	0.22	0.4255	11.9708	12.0708	5.4867	5.4413
2.4	0.3455	0.24	0.4145	12.9758	12.8958	5.3732	5.3316
2.6	0.3364	0.26	0.4036	13.2037	13.3037	5.1168	5.0783
2.8	0.3273	0.28	0.3927	13.1686	13.2686	4.7388	4.7031
3	0.3182	0.3	0.3818	12.7189	12.8189	4.2730	4.2396

TABLE 3.1 – Variations des mesure de performances dans le système en fonction de  $\lambda$ .

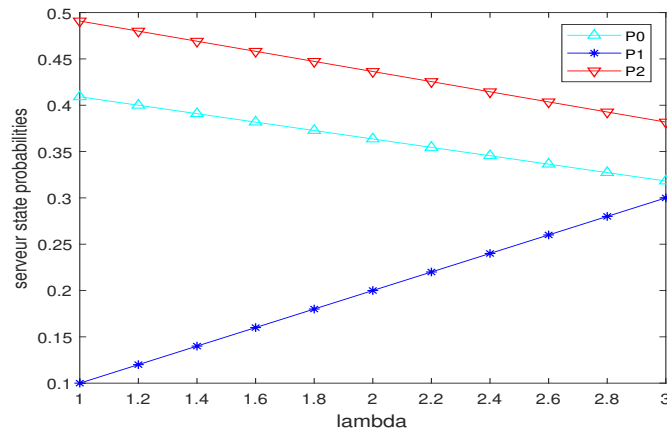


FIGURE 3.2 – Effet de  $\lambda$  sur les probabilités de l'état du serveur.

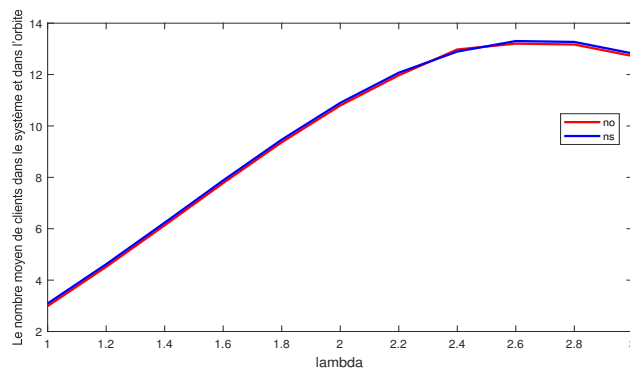


FIGURE 3.3 – Effet de  $\lambda$  sur le nombre moyen de clients dans le système et dans l'orbite.

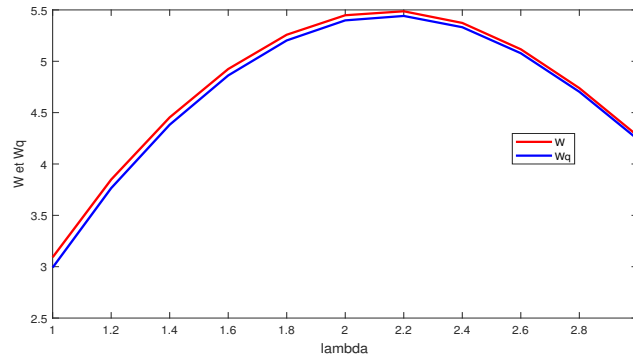


FIGURE 3.4 – Effet de  $\lambda$  sur le temps moyen de séjour des clients dans le système et dans l'orbite.

### 3.6.2 L'effet du taux de recherche $\nu$ sur le modèle

On fait varier le taux de recherche  $\nu$  sur les mesures de performance pour  $\lambda = 2.2$ ,  $\mu = 4$ ,  $\theta = 1.44$ , et  $\alpha = 1.2$ .

$\nu$	$P_0$	$P_1$	$P_2$	$\bar{n}_0$	$\bar{n}_s$	$W$	$W_q$
1	0.2045	0.55	0.2455	2.5629	3.1129	1.415	1.165
2	0.2812	0.55	0.1687	17.3015	17.8515	8.1143	7.8643
3	0.3214	0.55	0.1286	4.0757	4.6257	2.1026	1.8526
4	0.3462	0.55	0.1038	3.0197	3.5697	1.6226	1.3726
5	0.3629	0.55	0.0871	2.5876	3.1376	1.4262	1.1762
6	0.375	0.55	0.075	2.3456	2.9856	1.3162	1.0662
7	0.3841	0.55	0.0659	2.1893	2.7393	1.2451	0.9951
8	0.3913	0.55	0.0587	2.0795	2.6295	1.1952	0.9452
9	0.3971	0.55	0.0529	1.9978	2.5478	1.1581	0.9081
10	0.4018	0.55	0.0442	1.9347	2.4847	1.1294	0.8791
11	0.4057	0.55	0.0443	1.8843	2.4343	1.1065	0.8565

TABLE 3.2 – Variations des mesures de performances dans le système en fonction de  $\nu$ .

#### Commentaire :

Dans le tableau 3.2, on voit clairement la stabilité de la probabilité  $P_1$ ,  $P_2$  qui décroît contrairement à  $P_0$  ce qui est montré par la figure 3.5. On remarque dans 3.6 que l'augmentation de la valeur du taux  $\nu$  induit une augmentation du nombre moyen et du temps moyen de séjour des clients dans le système et dans l'orbite, puis elle baisse juste au moment où la valeur de  $\nu$  se rapproche du taux des arrivées  $\lambda$ .

De plus en plus le serveur est en recherche, le nombre moyen de clients dans le système est plus petit, voir la figure 3.7.

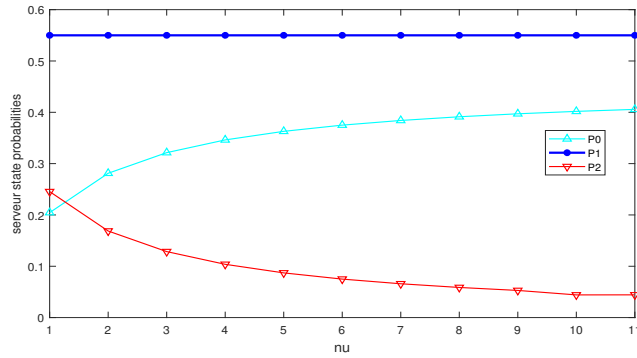


FIGURE 3.5 – Effet de  $\nu$  sur les probabilités de l'état du serveur.

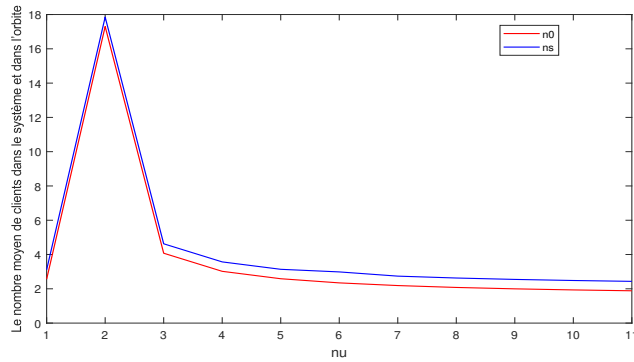


FIGURE 3.6 – Effet de  $\nu$  sur le nombre moyen de clients dans le système et dans l'orbite.

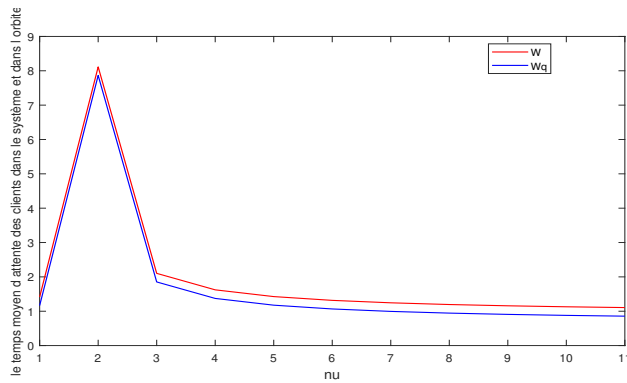


FIGURE 3.7 – Effet de  $\nu$  sur le temps moyen de séjour des clients dans le système et dans l'orbite.

### 3.6.3 L'effet du taux d'inactivité $\alpha$ sur le modèle

On fait varier le taux d'inactivité  $\alpha$  sur les mesures de performance pour  $\lambda = 5$ ,  $\mu = 10$ ,  $\theta = 1.33$ , et  $\nu = 2.5$ .

**Commentaire :** Dans le tableau [3.3](#) et d'après les figures [3.9](#), [3.10](#) en remarque :

$\alpha$	$P_0$	$P_1$	$P_2$	$\bar{n}_0$	$\bar{n}_s$	$W$	$W_q$
1	0.3571	0.5	0.1492	3.838	4.338	0.8676	0.7676
1.2	0.3378	0.5	0.1622	4.4081	4.581	0.9162	0.8162
1.4	0.3205	0.5	0.1795	4.391	4.891	0.9782	0.8782
1.6	0.3049	0.5	0.1951	4.758	5.258	1.0516	0.9516
1.8	0.2907	0.5	0.2093	5.1741	5.6741	1.1348	1.0348
2	0.2778	0.5	0.2222	5.6329	6.1329	1.2266	1.1266
2.2	0.266	0.5	0.234	6.1287	6.6287	1.3257	1.2257
2.4	0.2551	0.5	0.2449	6.6572	7.1572	1.4314	1.3314
2.6	0.2451	0.5	0.2549	7.2144	7.7144	1.5429	1.4429
2.8	0.2358	0.5	0.2642	7.7972	8.2972	1.6594	1.5594
3	0.2273	0.5	0.2727	8.4026	8.9026	1.7805	1.6805

TABLE 3.3 – Variations des mesures de performances dans le système en fonction de  $\alpha$ .

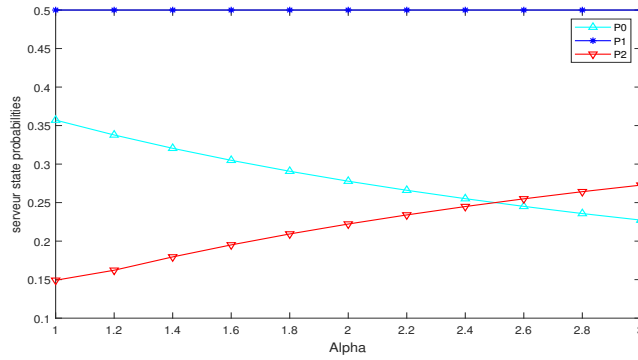


FIGURE 3.8 – Effet de  $\alpha$  sur les probabilités de l'état du serveur.

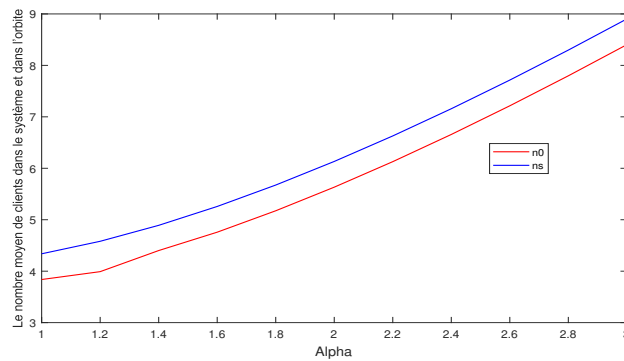


FIGURE 3.9 – Effet de  $\alpha$  sur le nombre moyen de clients dans le système et dans l'orbite.

L'augmentation de  $\alpha$  induit l'augmentation du nombre moyen de clients dans le système et dans l'orbite, ainsi que le temps de séjours des clients dans le système et dans l'orbite.

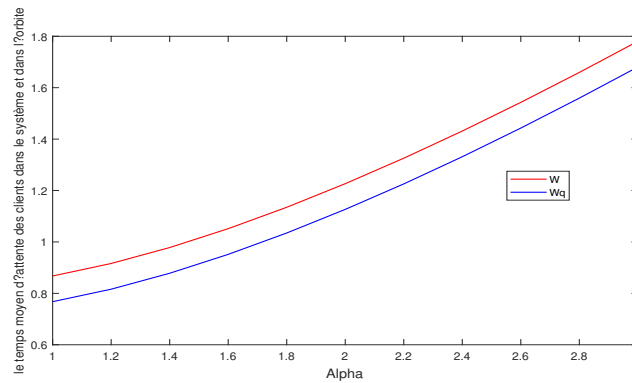


FIGURE 3.10 – Effet de  $\alpha$  sur le temps moyen de séjour des clients dans le système et dans l'orbite.

### 3.6.4 Effet de variations de $\nu$ et $\alpha$ sur le nombre de clients en orbite

Dans cette partie, on fait varier les deux paramètres  $\nu$  et  $\alpha$  au même temps, afin de voir la variation de nombre moyen de clients dans l'orbite avec  $\lambda = 7$ ,  $\mu = 8$  et  $\theta = 10$ .

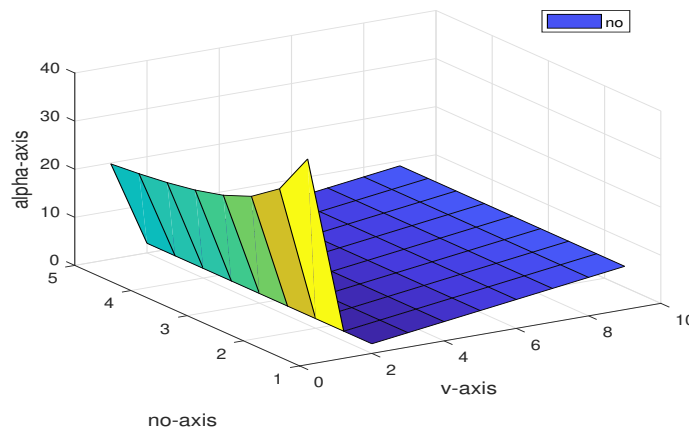


FIGURE 3.11 – Variation de nombre moyen de clients dans l'orbite en fonction de  $\alpha$  et  $\nu$ .

**Commentaire :** Le nombre moyen de clients dans l'orbite diminue avec l'augmentation du taux de recherche, et augmente avec l'augmentation du taux d'inactivité.

### 3.6.5 Effet de variations de $\nu$ et $\alpha$ sur le nombre de clients dans le système

On fait varier les deux paramètres  $\nu$  et  $\alpha$  au même temps, afin de voir la variation du nombre moyen de clients dans le système avec  $\lambda = 7$ ,  $\mu = 8$  et  $\theta = 10$ .

**Commentaire :** Le nombre moyen de clients dans le système augmentent avec l'augmentation du taux d'inactivité et diminue avec l'augmentation du taux de recherche, voir la figure [3.12](#).



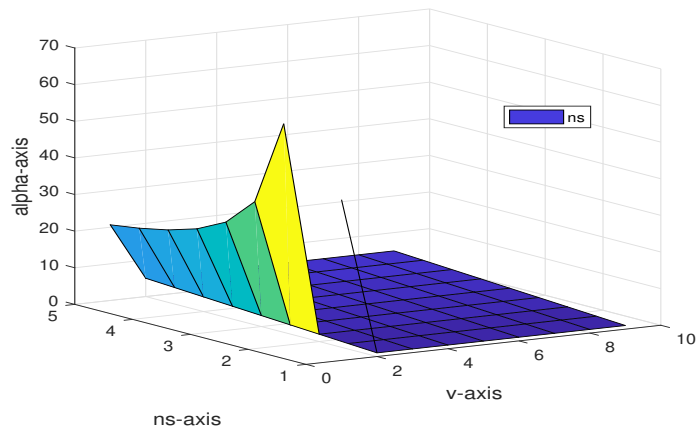


FIGURE 3.12 – Variation de nombre moyen de clients dans le système en fonction de  $\alpha$  et  $\nu$ .

### 3.6.6 Effet de variations de $\nu$ et $\alpha$ sur le temps moyen de séjour des clients dans le système

On fait varier les deux paramètres  $\nu$  et  $\alpha$  au même temps, afin de voir la variation du temps moyen de séjour des clients dans le système avec  $\lambda = 5$ ,  $\mu = 8$  et  $\theta = 10$ .

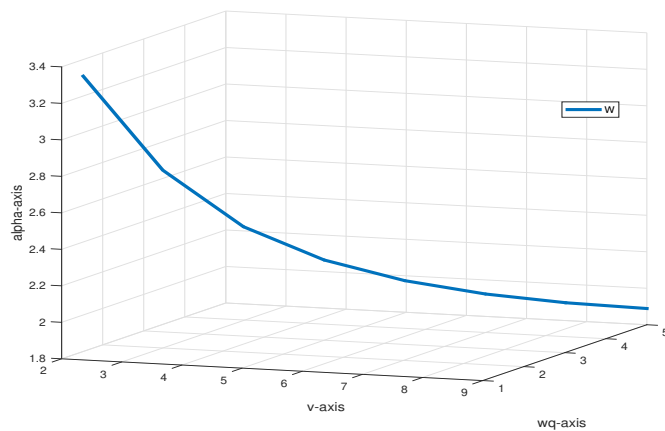


FIGURE 3.13 – Variation de temps moyen de séjour des clients dans le système en fonction de  $\alpha$  et  $\nu$ .

**Commentaire :** À partir de la figure [3.13](#), on remarque que le temps moyen de séjour des clients dans le système diminue avec l'augmentation du taux de recherche  $\nu$ , et augmente avec l'augmentation du taux d'inactivité.

## 3.7 Conclusion

Dans ce chapitre, nous avons considéré l'analyse de sensibilité des mesures de performance par rapport à leurs paramètres critique, dans le système de file d'attente  $M/M/1$  avec rappels et recherche des clients en orbite, le temps de recherche est non négligable. Ce modèle proposé par Tuan Phung-Duc en 2015 [66], est motivé par les systèmes du Cloud Computing. Nous avons effectuée une analyse numérique montrant l'effet de la variation de quelques paramètres sur les mesures de performance, (le nombre moyen de clients dans le système, dans l'orbite, et les probabilités  $P_0$ ,  $P_1$  et  $P_2$  ainsi que le temps moyen de séjour des clients dans le système et dans l'orbite) du système proposée, Ces résultats sont présentés dans [20] et [18].

## Chapitre 4

# Modèle d'attente $M/G/1$ avec rappels, feedback et recherche des clients en orbite

### 4.1 Introduction

Dans ce chapitre, nous étudions un système d'attente avec rappels, Bernoulli feedback et recherche des clients en orbite, La notion de feedback introduite en général pour exploiter des situations d'attente où tous les clients demandent le principal service et seulement quelques uns parmi eux ont besoin de demander un autre service. Nous avons obtenu quelques mesures de performance, importantes et intéressantes du système, en utilisant l'analyse stationnaire. Des exemples numériques sont donnés pour illustrer les résultats théoriques obtenus. Ces résultats font l'objet de l'étude réalisée dans [22]

### 4.2 Description mathématique du modèle

On considère un système de files d'attente à un seul serveur où les clients primaires arrivent suivant un flux poissonnien de taux  $\lambda$ . Un client qui arrive et trouve le serveur occupé, quitte l'aire du service pour rejoindre un groupe de clients bloqués appelé orbite. Après un certain temps aléatoire, il renouvelle sa tentative d'entrer en service jusqu'à ce qu'il le trouve disponible. Les intervalles de temps inter-rappels suivent une distribution exponentielle de taux  $\theta$ . La durée de service est de loi générale, de fonction de distribution  $B(x)$  de transformée de Laplace-Stieltjes  $\tilde{B}(s) = \int_0^\infty e^{-sx} dB(x)$ , et des deux premiers moments finis  $\beta_1$  et  $\beta_2$ , respectivement. Après l'achèvement du service, le serveur reste inactif pendant une durée de temps qui suit la loi exponentielle avec une moyenne  $\frac{1}{\alpha}$ . Pendant cette période d'inactivité, un client arrivant (soit primaire ou secondaire) est immédiatement servi. Après la période d'inactivité, le serveur commence la recherche des clients en orbite, selon une distribution exponentielle de taux  $\mu$ . Les clients qui arrivent et trouvent le serveur occupé (par un service d'un client ou par la recherche) rejoignent l'orbite. Après le temps de recherche, le serveur récupère un client de l'orbite s'il y en a un, sinon il reste libre. Une fois servi, le client doit décider, soit de rejoindre l'orbite pour un autre service avec une probabilité  $c$  ( $0 \leq c < 1$ ) où de quitter le système définitivement avec

une probabilité complémentaire  $\bar{c} = 1 - c$ .

Toutes les variables aléatoires introduites sont mutuellement indépendantes.

### 4.3 La condition de stabilité

Soit  $\{\xi_n, n \in \mathbb{N}\}$  une suite d'instants de la fin d'un service, La suite de variables aléatoires  $\{q_n = N_o(\xi_n), n \geq 1\}$  forme une chaîne de Markov induite, dont l'équation fondamentale est :

$$q_{n+1} = q_n - \delta_{q_n} + v_{n+1} + u. \quad (4.1)$$

où,  $v_{n+1}$  représente le nombre de clients qui arrivent pendant un temps de service qui se termine à l'instant  $t_{n+1}$ . Sa distribution est donnée par :

$$P(v_n = i) = k_i = \int_0^\infty \frac{(\lambda x)^i}{i!} e^{-\lambda x} dB(x),$$

avec la fonction génératrice  $K(z) = \sum_{n=0}^\infty k_n z^n = \tilde{B}(\lambda - \lambda z)$ . La variable de Bernoulli  $\delta_{q_n}$  est définie par :

$$\delta_{q_n} = \begin{cases} 1 & \text{si le } (n+1)^{\text{ème}} \text{ client provient de l'orbite,} \\ 0 & \text{si non.} \end{cases}$$

Sa distribution conditionnelle est donnée par :

$$P(\delta_{q_n} = 1/q_n = i) = \frac{i\theta}{\lambda + i\theta}, i \neq 0.$$

$$P(\delta_{q_n} = 0/q_n = i) = \frac{\lambda}{\lambda + i\theta}.$$

La variable aléatoire  $u$  est définie par :

$$u = \begin{cases} 1, & \text{si le client servi décide de rejoindre l'orbite;} \\ 0, & \text{si le client servi décide de quitter le système.} \end{cases}$$

En outre,  $P[u = 1] = c$  et  $P[u = 0] = \bar{c} = 1 - c$ .

Nous avons les probabilités de transition en une étape suivante

$$r_{ij} = k_{j-i} \frac{\lambda}{\lambda + i\theta} \bar{c} + k_{j-i+1} \frac{i\theta}{\lambda + i\theta} \bar{c} + k_{j-i-1} \frac{\lambda}{\lambda + i\theta} c + k_{j-i} \frac{i\theta}{\lambda + i\theta} c.$$

Tel que  $r_{ij} \neq 0$  pour  $i = 0, 1, 2, \dots, j + 1$ .

**Théorème 4.3.0.1.** *La chaîne de Markov induite  $\{q_n, n \geq 1\}$  est ergodique si et seulement si  $\rho = \lambda\beta_1 + c < 1$ .*

*Démonstration.* De (4.1), on peut voir que  $\{q_n, n \geq 1\}$  est une chaîne de Markov irréductible et apériodique. Pour trouver la condition suffisante, nous utilisons le critère de Foster. D'après cette dernière, on montre l'existence d'une fonction non négative  $f(k), k \in S$ , (dans notre cas  $f(q_n) = q_n$ ) et  $\varepsilon > 0$  telle que l'accroissement moyen  $\chi_k = E[f(q_{n+1}) - f(q_n) / q_n = k]$  est fini pour tout  $k \in S$  et  $\chi_k \leq -\varepsilon$  pour tous  $k \in S$  sauf peut-être un nombre fini, c'est-à-dire  $\chi_k = E[q_{n+1} - q_n / q_n = k] = \lambda\beta_1 - \frac{k\theta}{\lambda+k\theta} + c$ . Soit  $\chi = \lim_{k \rightarrow \infty} \chi_k$ . Alors  $\chi = \lambda\beta_1 - 1 + c < 0$ . Donc la condition suffisante est  $\lambda\beta_1 + c < 1$ . Pour prouver qu'il y a aussi une condition nécessaire pour notre chaîne de Markov induite, on applique la condition de Kaplan : une chaîne de Markov irréductible et apériodique n'est pas ergodique si  $\chi_k < \infty$  pour tout  $k \geq 0, k_0 \in \mathbb{N}$  et  $k \geq k_0$  dans notre cas, cette condition est vérifiée car  $r_{ij} = 0$  pour  $j < i - 1$  et  $i > 0$ . Par conséquent,  $\lambda\beta_1 + c \geq 1$  donne la non ergodicité de la chaîne de Markov induite  $\{q_n, n \geq 1\}$ . Finalement,  $\{q_n, n \geq 1\}$  est ergodique si et seulement si  $\rho = \lambda\beta_1 + c < 1$ .  $\square$

## 4.4 Etude de l'état stationnaire

Soit  $C(t)$  l'état du serveur à l'instant  $t \geq 0$ .

$$C(t) = \begin{cases} 0 & \text{le serveur est libre,} \\ 1 & \text{le serveur est occupé,} \\ 2 & \text{le serveur recherche un client.} \end{cases}$$

$N(t)$  est le nombre de clients dans l'orbite à l'instant  $t, t \geq 0$ . Nous avons  $\{X(t) = (C(t), N(t)), t \geq 0\}$  qui forme une chaîne de Markov sur l'espace d'état  $S = \{0, 1, 2\} \times \{0, 1, 2, \dots\}$ . Les transitions entre les états sont illustrées dans la figure 4.1.

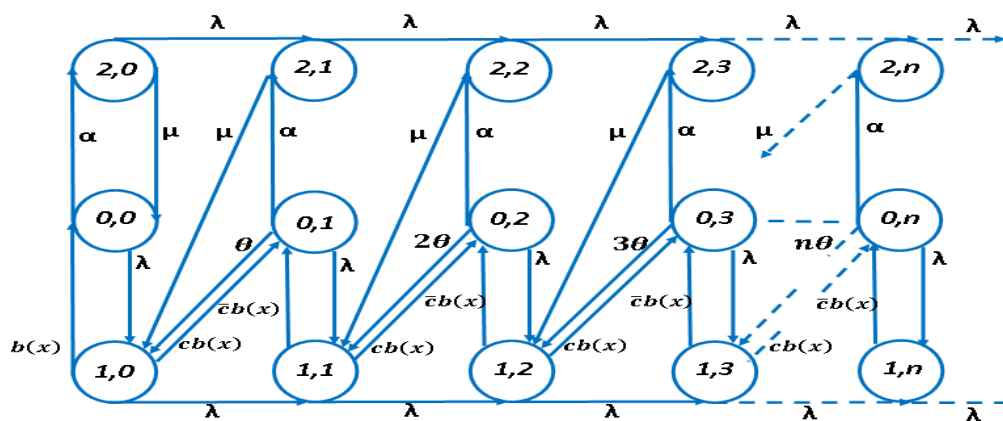


FIGURE 4.1 – Graphe des transitions du modèle  $M/G/1$  avec rappels, feedback et recherche des clients en orbite.

Si le système est en régime stationnaire  $\rho < 1$ , alors la distribution conjointe de l'état du

serveur et de la longueur de la file d'attente (orbite).

$$\begin{aligned} p_{0,n} &= P \{C(t) = 0, N(t) = n\}, \\ p_{1,n}(x) &= P \{C(t) = 1, \xi(t) < x, N(t) = n\}, \\ p_{2,n} &= P \{C(t) = 2, N(t) = n\}. \end{aligned}$$

La technique des PGF est utilisée ici pour obtenir la solution à l'état d'équilibre du modèle de file d'attente avec rappels feedback et recherche des clients en orbite. Pour résoudre les équations ci-dessus, on définit les fonctions génératrices pour  $|z| \leq 1$ , comme suit :

$$P_0(z) = \sum_{n=0}^{\infty} z^n p_{0,n}, P_1(z) = \sum_{n=0}^{\infty} z^n p_{1,n}, P_1(z, x) = \sum_{n=0}^{\infty} z^n p_{1,n}(x), P_2(z) = \sum_{n=0}^{\infty} z^n p_{2,n}.$$

L'ensemble d'équations d'équilibre statistique est obtenu comme suit :

$$(\lambda + \alpha)p_{0,0} = \bar{c} \int_0^{\infty} p_{1,0}(x) b(x) dx + \mu p_{2,0}, \quad (4.3)$$

$$(\lambda + \alpha + n\theta)p_{0,n} = (1 - (1 - \delta_{n,0})c) \int_0^{\infty} p_{1,n}(x) b(x) dx + c \int_0^{\infty} p_{1,n-1}(x) b(x) dx, n \geq 1,$$

$$p'_{1,n}(x) = -(\lambda + b(x))p_{1,n}(x) + \lambda p_{1,n-1}(x), n \geq 0, \quad (4.4)$$

$$p_{1,n}(0) = \lambda p_{0,n} + (n+1)\theta p_{0,n+1} + \mu p_{2,n+1}, n \geq 0, \quad (4.5)$$

$$(\lambda + \mu)p_{2,n} = \lambda p_{2,n-1} + \alpha p_{0,n}, n \geq 0. \quad (4.6)$$

Où  $p_{i,-1} = 0$  ( $i = 1, 2$ ). En transformant les équations d'équilibre ci-dessus en fonctions génératrices, nous obtenons,

$$(\lambda + \alpha) P_0(z) + \theta z P'_0(z) = (1 - c(1 - z)) \int_0^{\infty} P_1(z, x) b(x) dx + \mu p_{2,0}, \quad (4.7)$$

$$\frac{\partial P_1(z, x)}{\partial x} = -(\lambda - \lambda z + b(x)) P_1(z, x), \quad (4.8)$$

$$P_1(z, 0) = \lambda P_0(z) + \theta \frac{dP_0(z)}{dz} + \frac{\mu}{z} (P_2(z) - p_{2,0}), \quad (4.9)$$

$$(\lambda - \lambda z + \mu) P_2(z) = \alpha P_0(z). \quad (4.10)$$

La résolution de (4.8) donne,

$$P_1(z, x) = P_1(z, 0) (1 - B(x)) e^{-\lambda(1-z)x}. \quad (4.11)$$

Combiner (4.11), (4.9), (4.10) et (4.7) et après quelques manipulations algébriques on obtient :

$$\begin{aligned} & \theta z \left[ (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z \right] P'_0(z) \\ &= z \left[ \lambda \left( 1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) \right) + \alpha \left( 1 - \frac{\mu (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{z(\lambda - \lambda z + \mu)} \right) \right] P_0(z) \\ &+ \mu p_{2,0} \left( (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z \right). \end{aligned} \quad (4.12)$$

Coefficient de  $P'_0(z)$  a deux zéros  $z_1 = 0$  et  $z_2 = 1$ . Nous choisissons un point arbitraire  $a \in (0, 1)$  et résoudre (4.12) pour  $z \in (0, a]$  puis pour  $z \in [a, 1)$ .

On considère la fonction  $f(z) = (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z$ . Nous avons :

$$\begin{aligned} f(1) &= \tilde{B}(0) - 1 = 1 - 1 = 0, \\ f'(z) &= -\lambda(\bar{c} + cz) \tilde{B}'(\lambda - \lambda z) + c\tilde{B}(\lambda - \lambda z) - 1, \\ f'(1) &= -\lambda\tilde{B}'(0) + c\tilde{B}(0) - 1 = \rho - 1 < 0, \\ f''(z) &= \lambda^2(\bar{c} + cz) \tilde{B}''(\lambda - \lambda z) - 2\lambda c\tilde{B}'(\lambda - \lambda z), \\ f''(1) &= \lambda^2\beta_2 + 2\lambda c\beta_1 \geq 0. \end{aligned}$$

Donc la fonction  $f(z)$  est décroissante sur l'intervalle  $[0, 1]$ ,  $z_2 = 1$  est le seul zéro ici et pour  $z \in [a, 1)$ , nous avons :  $z < (\bar{c} + cz) \tilde{B}(\lambda - \lambda z) \leq 1$ .

A côté

$$\frac{1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} = \frac{\rho}{1 - \rho} < \infty, \quad (4.13)$$

c'est-à-dire que la fonction  $\frac{1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z}$  peut être défini au point  $z_2 = 1$  comme  $\frac{\rho}{1 - \rho}$ . Cela signifie que pour  $z \in (0, a]$  et pour  $z \in [a, 1)$  nous pouvons réécrire l'équation (4.12) comme suit :

$$P'_0(z) = \left[ \frac{\lambda}{\theta} \left( \frac{1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} \right) + \frac{\alpha}{\theta} \left( \frac{1 - \frac{\mu(\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{z(\lambda - \lambda z + \mu)}}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} \right) \right] P_0(z) + \mu p_{2,0} \frac{1}{z}. \quad (4.14)$$

D'abord, nous trouvons la solution générale de l'équation homogène, qui est transformée en

$$\frac{P'_0(z)}{P_0(z)} = \frac{\lambda}{\theta} \left( \frac{1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} \right) + \frac{\alpha}{\theta} \left( \frac{\frac{\lambda^2}{\lambda + \mu}}{\lambda + \mu - \lambda z} - \frac{\frac{\mu}{\lambda + \mu}}{z} + \frac{1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} \frac{\lambda}{\lambda + \mu - \lambda z} \right) \quad (4.15)$$

nous posons

$$s(t_1, t_2) = \exp \left\{ \frac{\lambda}{\theta} \int_{t_1}^{t_2} \frac{1 - (\bar{c} + cu) \tilde{B}(\lambda - \lambda u)}{(\bar{c} + cu) \tilde{B}(\lambda - \lambda u) - u} du \right\}. \quad (4.16)$$

En résolvant (4.15) pour  $z \in (0, a]$  on obtient :

$$P_0(z) = c(z) s(a, z)^{\left(1 + \frac{\alpha}{\lambda + \mu - \lambda z}\right)} \left[ \left( \frac{a}{z} \right)^\mu \left( \frac{\lambda + \mu - \lambda a}{\lambda + \mu - \lambda z} \right)^\lambda \right]^{\frac{\alpha}{\theta(\lambda + \mu)}} \exp \left\{ -\alpha \lambda \int_a^z \frac{\ln s(a, u)}{(\lambda + \mu - \lambda u)^2} du \right\}. \quad (4.17)$$

En substituant cette solution dans l'équation différentielle nonhomogène (4.14), on peut déter-

miner la fonction  $c(z)$  comme :

$$c(z) = \frac{\mu p_{2,0}}{\left[ a^\mu (\lambda + \mu - \lambda a)^\lambda \right]^{\frac{\alpha}{\theta(\lambda+\mu)}}} \int_a^z s(a, t)^{-\left(1 + \frac{\alpha}{\lambda + \mu - \lambda t}\right)} t^{\left(\frac{\alpha\mu}{\theta(\lambda+\mu)} - 1\right)} (\lambda + \mu - \lambda t)^{\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \exp \left\{ \alpha\lambda \int_a^t \frac{\ln s(a, u)}{(\lambda + \mu - \lambda u)^2} du \right\} dt + P_0(a).$$

Comme  $z \rightarrow 0^+$ ,  $P_0(0) < \infty$  et  $\left(\frac{a}{z}\right)^{\frac{\mu\alpha}{\theta(\lambda+\mu)}}$  est divergente. Ainsi

$$P_0(a) = \frac{\mu p_{2,0}}{\left[ a^\mu (\lambda + \mu - \lambda a)^\lambda \right]^{\frac{\alpha}{\theta(\lambda+\mu)}}} \int_0^a s(a, t)^{-\left(1 + \frac{\alpha}{\lambda + \mu - \lambda z}\right)} t^{\left(\frac{\alpha\mu}{\theta(\lambda+\mu)} - 1\right)} (\lambda + \mu - \lambda t)^{\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \exp \left\{ \alpha\lambda \int_a^t \frac{\ln s(a, u)}{(\lambda + \mu - \lambda u)^2} du \right\} dt. \quad (4.18)$$

Par contre, en résolvant (4.15) pour  $z \in [a, 1)$ , et en prenant comme limite  $z \rightarrow 1$ , avec (4.14) on obtient :

$$P_0(a) = \left[ a^\mu (\lambda + \mu - \lambda a)^\lambda \right]^{-\frac{\alpha}{\theta(\lambda+\mu)}} \left[ \begin{aligned} & P_0(1) s(a, 1)^{-\left(1 + \frac{\alpha}{\mu}\right)} \mu^{\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \exp \left\{ \alpha\lambda \int_a^1 \frac{\ln s(a, u)}{(\lambda + \mu - \lambda u)^2} du \right\} \\ & - \mu p_{2,0} \int_a^1 s(a, t)^{-\left(1 + \frac{\alpha}{\lambda + \mu - \lambda t}\right)} t^{\left(\frac{\alpha\mu}{\theta(\lambda+\mu)} - 1\right)} (\lambda + \mu - \lambda t)^{\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \exp \left\{ \alpha\lambda \int_a^t \frac{\ln s(a, u)}{(\lambda + \mu - \lambda u)^2} du \right\} dt \end{aligned} \right]. \quad (4.19)$$

En mettant en équation (4.18) et (4.19) on obtient :

$$P_0(1) = p_{2,0} \mu^{1 - \frac{\lambda\alpha}{\theta(\mu+\lambda)}} \int_0^1 s(1, t)^{-\left(1 + \frac{\alpha}{\lambda + \mu - \lambda t}\right)} t^{\left(\frac{\alpha\mu}{\theta(\lambda+\mu)} - 1\right)} (\lambda + \mu - \lambda t)^{\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \exp \left\{ \alpha\lambda \int_1^t \frac{\ln s(1, u)}{(\lambda + \mu - \lambda u)^2} du \right\} dt, \quad (4.20)$$

alors on peut réécrire la solution de (4.14) comme :

$$P_0(z) = \mu p_{2,0} \left( \frac{1}{z^\mu (\lambda + \mu - \lambda z)^\lambda} \right)^{\frac{\alpha}{\theta(\lambda+\mu)}} \int_0^z s(z, t)^{-\left(1 + \frac{\alpha}{\lambda + \mu - \lambda t}\right)} t^{\left(\frac{\alpha\mu}{\theta(\lambda+\mu)} - 1\right)} (\lambda + \mu - \lambda t)^{\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \exp \left\{ \alpha\lambda \int_z^t \frac{\ln s(1, u)}{(\lambda + \mu - \lambda u)^2} du \right\} dt. \quad (4.21)$$

En combinant (4.10), (4.9) et (4.11), on obtient

$$P_1(z, x) = \left[ \lambda \left( 1 + \frac{\alpha}{\lambda + \mu - \lambda z} \right) \left( 1 + \frac{1 - (\bar{c} + cz) \tilde{B}(\lambda - \lambda z)}{(\bar{c} + cz) \tilde{B}(\lambda - \lambda z) - z} \right) P_0(z) + \frac{\mu p_{2,0}}{z} (\theta - 1) \right] (1 - B(x)) e^{-\lambda(1-z)x}.$$



Puisque  $P_1(z) = \int_0^\infty P_1(z, x) dx$ , on obtient

$$P_1(z) = \frac{1-\tilde{B}(\lambda-\lambda z)}{1-z} \left[ \left(1 + \frac{\alpha}{\lambda+\mu-\lambda z}\right) \left(1 + \frac{1-(\bar{c}+cz)\tilde{B}(\lambda-\lambda z)}{(\bar{c}+cz)\tilde{B}(\lambda-\lambda z)-z}\right) P_0(z) + \frac{\mu p_{2,0}}{\lambda z} (\theta - 1) \right]. \quad (4.22)$$

Appliquer maintenant la condition de normalisation  $P_0(1) + P_1(1) + P_2(1) = 1$ , et nous avons pris (4.22) avec  $\lim_{z \rightarrow 1} \frac{1-\tilde{B}(\lambda-\lambda z)}{1-z} = \lambda\beta_1$  on a :

$$P_0(1) = \frac{1 - \beta_1 (\theta - 1) \mu p_{2,0}}{\left(1 + \frac{\alpha}{\mu}\right) \left(1 + \frac{\lambda\beta_1}{1-\rho}\right)}. \quad (4.23)$$

En utilisant (4.23) et (4.20), nous obtenons l'expression de  $p_{2,0}$  comme

$$p_{2,0} = \left[ \left(1 + \frac{\alpha}{\mu}\right) \left(1 + \frac{\lambda\beta_1}{1-\rho}\right) \mu^{1-\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \int_0^1 s(1, t)^{-(1+\frac{\alpha}{\lambda+\mu-\lambda t})} t^{\left(\frac{\alpha\mu}{\theta(\lambda+\mu)}-1\right)} (\lambda + \mu - \lambda t)^{\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \exp \left\{ \alpha\lambda \int_1^t \frac{\ln s(1, u)}{(\lambda + \mu - \lambda u)^2} du \right\} dt. + \beta_1 (\theta - 1) \mu \right]^{-1}. \quad (4.24)$$

Maintenant, nous résumons les résultats ci-dessus dans le théorème suivant.

**Théorème 4.4.0.1.** *Sous la condition de stabilité  $\rho = \lambda\beta_1 + c < 1$ , les fonctions génératrices de la distribution conjointe stationnaire de la taille de l'orbite et de l'état du serveur sont données par :*

$$P_0(z) = \mu p_{2,0} \left( \frac{1}{z^\mu (\lambda + \mu - \lambda z)^\lambda} \right)^{\frac{\alpha}{\theta(\lambda+\mu)}} \int_0^z s(z, t)^{-(1+\frac{\alpha}{\lambda+\mu-\lambda t})} t^{\left(\frac{\alpha\mu}{\theta(\lambda+\mu)}-1\right)} (\lambda + \mu - \lambda t)^{\frac{\lambda\alpha}{\theta(\mu+\lambda)}} \exp \left\{ \alpha\lambda \int_z^t \frac{\ln s(1, u)}{(\lambda + \mu - \lambda u)^2} du \right\} dt.$$

où  $p_{2,0}$  est donnée dans l'équation (4.24).

$$P_1(z) = \frac{1-\tilde{B}(\lambda-\lambda z)}{1-z} \left[ \left(1 + \frac{\alpha}{\lambda+\mu-\lambda z}\right) \left(1 + \frac{1-(\bar{c}+cz)\tilde{B}(\lambda-\lambda z)}{(\bar{c}+cz)\tilde{B}(\lambda-\lambda z)-z}\right) P_0(z) + \frac{\mu p_{2,0}}{\lambda z} (\theta - 1) \right]$$

$$P_2(z) = \frac{\alpha}{(\lambda - \lambda z + \mu)} P_0(z).$$

**Corollaire 4.4.0.1.** *- Sous la condition de stabilité  $\rho < 1$ ,*

- La fonction génératrice de la taille de l'orbite  $P(z)$ , est donnée par

$$P(z) = P_0(z) + P_1(z) + P_2(z).$$

- La fonction génératrice de la taille du système  $\phi(z)$ , est donnée par

$$\phi(z) = P_0(z) + zP_1(z) + P_2(z) = P(z) + (z - 1) P_1(z).$$

A partir des résultats ci-dessus, nous pouvons obtenir certaines mesures de performance du système en régime permanent.

**Corollaire 4.4.0.2.** • *La probabilité que le serveur est libre mais le système n'est pas vide, notée  $P_0$ , est donnée par*

$$P_0 = P_0(1) = \frac{1 - \beta_1(\theta - 1)\mu p_{2,0}}{\left(1 + \frac{\alpha}{\nu}\right)\left(1 + \frac{\lambda\beta_1}{1-\rho}\right)}.$$

• *La probabilité que le serveur est occupé, notée  $P_1$ , est donnée par*

$$P_1 = P_1(1) = \lambda\beta_1 \left[ \left(1 + \frac{\alpha}{\mu}\right) \left(1 + \frac{\rho}{1-\rho}\right) P_0 + \frac{\mu p_{2,0}}{\lambda} (\theta - 1) \right].$$

• *La probabilité que le serveur recherche un client en orbite, notée  $P_2$ , est donnée par*

$$P_2 = P_2(1) = \frac{\alpha}{\mu} P_0.$$

• *Nombre moyen de clients en orbite  $L_q$ , est donnée par*

$$L_q = P'(1).$$

• *Nombre moyen de clients dans le système  $L_s$ , est donnée par*

$$L_s = \phi'(1) = L_q + P_1(1).$$

• *Le temps d'attente d'un client dans le système  $W_s$  et le temps d'attente d'un client dans la file  $W_q$  sont déterminés à l'aide des formules de Little*

$$W_s = \frac{L_s}{\lambda} \text{ et } W_q = \frac{L_q}{\lambda}.$$

## 4.5 Résultats numériques

Basé sur les résultats obtenus dans les sections précédentes, nous présenterons quelques exemples numériques utilisant MATLAB afin d'illustrer l'effet de divers paramètres dans le système. Les valeurs arbitraires des paramètres sont choisies de manière à satisfaire la condition de stabilité. Nous étudions l'effet de certains paramètres sur la sensibilité du système et sur la disponibilité du serveur. Nous supposons que le temps de service suit la distribution exponentielle.

### 4.5.1 Effet de $c$ sur les probabilités de l'état du serveur

Nous supposons que  $\lambda = 0.06$ ,  $\theta = 1.25$ ,  $\alpha = 2.5$ ,  $\mu = 1.5$ , et  $\beta_1 = 1$ .

**Remarque 4.5.1.1.** *Pour le tableau 4.1 et la Fig. 4.2 on peut voir que les probabilités telle que le serveur est libre  $P_0$ , est à la recherche de clients  $P_2$  diminuent de manière monotone. La probabilité que le serveur est occupé  $P_1$  augmente, ce qui est conforme à nos attentes.*

$c$	$P_0$	$P_1$	$P_2$
0.5	0.2141	0.3120	0.3568
0.6	0.1846	0.3608	0.3077
0.7	0.1811	0.3666	0.3019
0.8	0.1796	0.3691	0.2994
0.9	0.1788	0.3704	0.2981

TABLE 4.1 – Effet de  $c$  sur les probabilités de l'état du serveur.

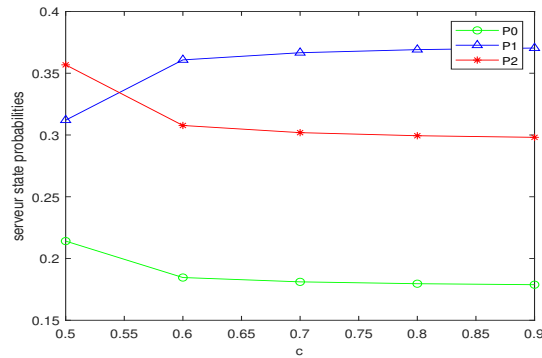


FIGURE 4.2 – Effet de  $c$  sur les probabilités de l'état du serveur.

#### 4.5.2 Effet de $\mu$ sur les probabilités de l'état du serveur

Nous supposons que  $\lambda = 6$ ,  $\theta = 1$ ,  $\alpha = 2$ ,  $\beta_1 = \frac{1}{30}$ , et  $c = 0.5$ .

$\mu$	$P_0$	$P_1$	$P_2$
0.1	0.0285	0.4179	0.5704
1.1	0.1838	0.4695	0.3342
2.1	0.2076	0.4247	0.1977
3.1	0.2086	0.3669	0.1346
4.1	0.2130	0.3263	0.1039
5.1	0.2154	0.2980	0.0845
6.1	0.2136	0.2755	0.0700
7.1	0.2078	0.2559	0.0585
8.1	0.1989	0.2377	0.0491
9.1	0.1874	0.2204	0.0412

TABLE 4.2 – Effet de  $\mu$  sur les probabilités de l'état du serveur.

**Remarque 4.5.2.1.** A partir du tableau 4.2 et de la Fig. 4.3, nous pouvons observer que lorsque le taux de recherche en orbite  $\mu$  augmente, la probabilité que le serveur soit inactif augmente de manière monotone  $P_0$ , lorsque les probabilités que le serveur soit occupé et en recherche d'un client diminuent.

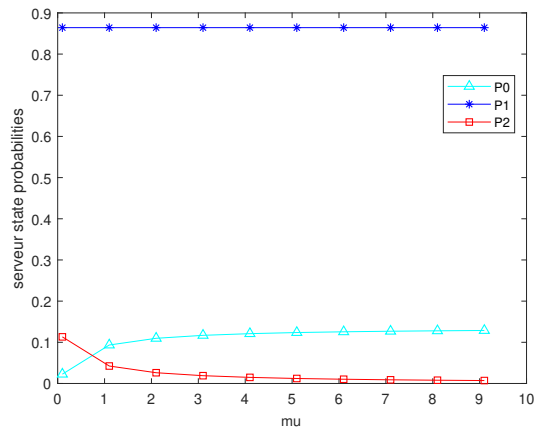


FIGURE 4.3 – Effet de  $\mu$  sur les probabilités de l'état du serveur.

### 4.5.3 Effet de $\alpha$ sur les probabilités de l'état du serveur

Nous supposons que  $\lambda = 5$ ,  $\theta = 2$ ,  $\mu = 2$ ,  $c = 0.2$ , et  $\beta_1 = \frac{1}{7}$ .

$\alpha$	$P_0$	$P_1$	$P_2$
1	0.0300	0.6651	0.0150
1.5	0.0295	0.6892	0.0221
2	0.0285	0.7087	0.0285
2.5	0.0272	0.7248	0.0341
3	0.0260	0.7383	0.0390
3.5	0.0248	0.7498	0.0434
4	0.0236	0.7597	0.0472
4.5	0.0225	0.7683	0.0507
5	0.0215	0.7759	0.0537
5.5	0.0206	0.7826	0.0565
6	0.0197	0.7885	0.0590

TABLE 4.3 – Effet de  $\alpha$  sur les probabilités de l'état du serveur.

**Remarque 4.5.3.1.** D'après le tableau 4.3 et la Fig.4.4, nous pouvons observer que, les probabilités que le serveur soit occupé  $P_1$ , et à la recherche de clients  $P_2$  augmentent de façon monotone avec l'augmentation du temps d'inactivité après l'achèvement d'un service  $\alpha$ .

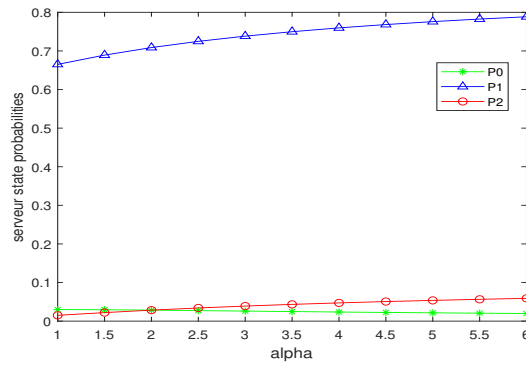


FIGURE 4.4 – Effet de  $\alpha$  sur les probabilités de l'état du serveur.

#### 4.5.4 Effet de $\lambda$ sur les probabilités de l'état du serveur

Nous supposons que :  $\theta = 3, \alpha = 2, \mu = 1.5, \beta_1 = \frac{1}{25}$ , et  $c = 0.5$ .

$\lambda$	$P_0$	$P_1$	$P_2$
1	0.3022	0.2171	0.4029
2	0.2351	0.3267	0.3134
3	0.2044	0.3727	0.2726
4	0.1842	0.3948	0.2456
5	0.1672	0.4027	0.2230
6	0.1510	0.3987	0.2014
7	0.1342	0.3828	0.1789
8	0.1159	0.3540	0.1546
9	0.0956	0.3103	0.1274

TABLE 4.4 – Effet de  $\lambda$  sur les probabilités de l'état du serveur.

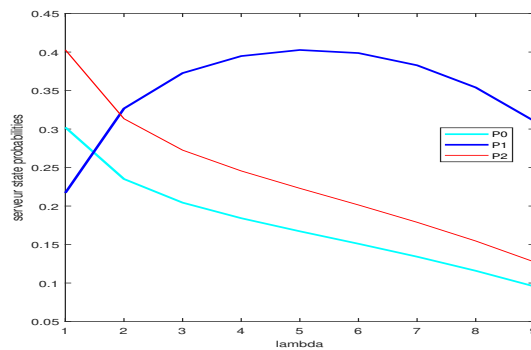


FIGURE 4.5 – Effet de  $\lambda$  sur les probabilités de l'état du serveur.

**Remarque 4.5.4.1.** *D'après le tableau 4.4 et la figure 4.5 nous pouvons voir que la probabilité que le serveur est occupé  $P_1$  augmente avec l'augmentation de  $\lambda$ , lorsque les probabilités que le serveur est inactif  $P_0$ , et celle que le serveur est en recherche des clients  $P_2$  diminuent de manière monotone.*

## 4.6 Conclusion

Dans ce chapitre, nous avons considéré un système d'attente avec rappels, feedback et recherche des clients en orbite. La méthode de la chaîne de Markov induite est proposée pour déterminer la condition d'ergodicité. Les fonctions génératrices partielles du nombre de clients dans le système lorsque le serveur est libre, occupé et à la recherche des clients en orbite sont trouvées. Ces études théoriques sont soutenues par des illustrations numériques. Le modèle présenté dans ce travail est nouveau et il est réaliste car il peut s'appliquer à diverses situations de congestion rencontrées dans les systèmes de télécommunication, les hôpitaux, les banques, les systèmes de fabrication qui impliquent la recherche des clients pour réduire le temps d'attente et rendre le système plus fonctionnel.

# Conclusion Générale

Dans ce travail, nous nous sommes intéressés aux systèmes de files d'attente avec rappels et recherche des clients en orbite :

Dans un premier temps, nous avons passé en revue certains résultats consacrés aux systèmes de files d'attente de type  $M/M/1$ ,  $M/G/1$  avec rappels et  $M/G/1$  avec rappels et feedback ainsi que sur les approches utilisées par les chercheurs dans leurs investigations.

Dans un deuxième temps, nous avons établi l'analyse stochastique complète des modèles  $M/M/1$  et  $M/G/1$  avec la discipline de rappels linéaire et recherche des clients en orbite par le serveur ; le temps de recherche est considéré négligeable.

Dans les deux derniers chapitres, nous avons réalisé l'analyse mathématique des systèmes de files d'attente avec rappels et recherche des clients en orbite le temps de recherche est considéré non négligeable. A cet effet, la chaîne de Markov induite est décrite, la condition d'ergodicité est établie à l'aide du critère de Foster et condition de Kaplan, la distribution stationnaire de l'état du système est trouvée en appliquant la méthode des fonctions génératrices. De plus Les fonctions génératrices partielles du nombre de clients dans le système lorsque le serveur est libre, occupé et à la recherche des clients en orbite sont trouvées. Notons que ces fonctions génératrices pourront être utilisées pour obtenir des mesures de performance importantes telles que la taille moyenne du système et le temps d'attente moyen. Nos investigations théoriques sont appuyées par des illustrations numériques.

Les travaux réalisés durant cette thèse et les résultats obtenus ouvrent un ensemble de perspectives :

- Analyser des systèmes de files d'attente avec rappels et recherche des clients en orbite plus complexe. Modéliser des systèmes concrets avec ce type de modèle.
- Extension du modèle étudié avec arrivée des clients floue.

# Bibliographie

- [1] S. N. Ahammed, P. Rajadurai, K. Indhira, and V. Chandrasekaran. A repairable retrial queueing system with two types of service, modified vacations and orbit search. *International Journal of Applied Engineering Research*, 9(18) :4337–4357, 2014.
- [2] A. Aissani. A survey on retrial queueing models. *Actes des Journées Statistiques Appliquées, USTHB, Alger*, pages 1–11, 1994.
- [3] A. Anupama and G. S. Keerthi. Using queueing theory the performance measures of cloud with infinite servers. *International Journal of Computer Science and Engineering Technology (IJCSET)*, 5(1), 2014.
- [4] J. Artalejo, V. Joshua, and A. Krishnamoorthy. An  $M/G/I$  retrial queue with orbital search by the server. *Advances in stochastic modelling*, pages 41–54, 2002.
- [5] J. R. Artalejo. Accessible bibliography on retrial queues : progress in 2000–2009. *Mathematical and computer modelling*, 51(9-10) :1071–1081, 2010.
- [6] J. R. Artalejo and A. Gómez-Corral. Steady state solution of a single-server queue with linear repeated requests. *Journal of Applied Probability*, 34(1) :223–233, 1997.
- [7] J. R. Artalejo and A. Gómez-Corral. Advances in retrial queues. *European journal of operational research*, 189(3) :1041–1041, 2008.
- [8] J. R. Artalejo and A. Gómez-Corral. *Retrial Queueing Systems : A Computational Approach*. Springer-Verlag Berlin Heidelberg, 2008.
- [9] G. Ayyappan and P. Thamizhselvi. Transient analysis of bulk arrival general service retrial queueing system with priority, bernoulli feedback, collisions, orbital search, modified bernoulli vacation, random breakdown and delayed repair. *International Journal of Statistics and Systems (IJSS)*, 12(1) :57–70, 2017.
- [10] G. Ayyappan and J. Udayageetha. Transient analysis of  $M[X_1], M[X_2]/G_1, G_2/1$  retrial queueing system with priority services, collisions, orbital search, working breakdown, start up/close down time, feedback, modified bernoulli vacation and balking. *International Journal of Applied Engineering Research*, 13(11) :8783–8795, 2018.
- [11] D. Babu, A. Krishnamoorthy, and V. Joshua.  $MAP/PH/I$  retrial queue with abandonment, flush out and search of customers. In *International Conference on Distributed Computer and Communication Networks*, pages 144–156. Springer, 2018.
- [12] D. Babu, A. Krishnamoorthy, and V. C. Joshua. *Retrial queue with search of interrupted customers from the finite orbit*, pages 360–371. Springer, 2018.



- [13] P. Bocharov, C. D'Apice, B. D'Auria, and S. Salerno. A queueing system of finite capacity with the server requiring a priority search for customers. *Vestnik RUDN, Seria Prikladnaia Matematika I Informatika*, (12) :50–61, 2000.
- [14] P. Bocharov, C. D'Apice, N. Phong, and G. Rizelian. Retrial servicing of multivariate poisson flow with customer-searching server with finite buffer. *Vestnik RUDN, Seria Prikladnaia Matematika I Informatika*, (1) :98–106, 2002.
- [15] P. Bocharov, C. D'Apice, N. Phong, and G. Rizelian. Retrial servicing of poisson flow in a system of finite capacity with customer-searching server. *Vestnik RUDN, Seria Prikladnaia Matematika I Informatika*, (1) :87–97, 2002.
- [16] M. Boualem, M. Cherfaoui, N. Djellab, and D. Aïssani. Analyse des performances du systeme  $M/G/1$  avec rappels et bernoulli feedback. *Journal Européen des Systèmes Automatisés*, 47(1-3) :181–193, 2013.
- [17] M. Boualem, N. Djellab, and D. Aïssani. Stochastic approximations and monotonicity of a single server feedback retrial queue. *Mathematical Problems in Engineering*, 2012, 2012.
- [18] Z. Boussaha, N. Oukid, and Zeghdoudi. On  $M/M/1$  retrial queue with orbital search. *Colloque International MOAD'22, Méthodes et Outils d'Aide à la Décision, Béjaia, ISBN : 978-9931-884-15-6*, pages 702–707, 15-17 Novembre 2022.
- [19] Z. Boussaha, N. Oukid, and Zeghdoudi. Stability analysis of a retrial queue with orbital search of customer. *Workshop IDEPT2020*, October 12, 2021.
- [20] Z. Boussaha, N. Oukid, and H. Zeghdoudi. Performance and analysis of an  $M/M/1$  retrial queue with orbital search. 1<sup>st</sup> International Symposium on Materials, Energy and Environment, MEE'2020; El Oued, ALGERIA, 2020.
- [21] Z. Boussaha, N. Oukid, and H. Zeghdoudi. On the orbital search of customers. *MSS'2019 (Colloque International Modélisation Stochastique et Statistique), USTHB Alger, ISBN 978-9931-9222-2-3. hal-02593238.*, pages 122–125, 24-26 Novembre 2019.
- [22] Z. Boussaha, N. Oukid, H. Zeghdoudi, S. Soualhi, and N. Djellab. On the  $M/G/1$  feedback retrial queueing with orbital search of customers. *Advances in Mathematics : Scientific Journal*, 11(8) :723–739, 2022.
- [23] S. R. Chakravarthy, A. Krishnamoorthy, and V. C. Joshua. Analysis of a multi-server retrial queue with search of customers from the orbit. *Performance evaluation*, 63(8) :776–798, 2006.
- [24] K. U. Chandrika and D. Sumitha. Performance measures of feedback batch arrival retrial queue with optional second service and orbital search. *Advances in Applied Research*, 3(2) :179–186, 2011.
- [25] X. Chang, B. Wang, J. K. Muppala, and J. Liu. Modeling active virtual machines on iaas clouds using an  $M/G/m/m+K$  queue. *IEEE Transactions on services computing*, 9(3) :408–420, 2014.
- [26] P. Chen, Y. Zhou, and C. Li. Batch arrival retrial g-queue with orbital search and non-persistent customers. *Journal of Interdisciplinary Mathematics*, 19(1) :95–109, 2016.

- [27] J. Cohen. Basic problems of telephone traffic theory and the influence of repeated calls. *Philips Telecommunication Review*, 18(2) :49–100, 1957.
- [28] S. Damodaran, A. Subramanian, and G. Sekar. Time dependent retrial queueing model with orbital search under non-preemptive priority services. *J. Math. Comput. Sci.*, 11(3) :3276–3299, 2021.
- [29] A. De Kok. Algorithmic methods for single server systems with repeated attempts. *Statistica Neerlandica*, 38(1) :23–32, 1984.
- [30] T. Deepak. On a retrial queueing model with single/batch service and search of customers from the orbit. *Top*, 23(2) :493–520, 2015.
- [31] T. Deepak, A. Dudin, V. Joshua, and A. Krishnamoorthy. On an  $M(X)/G/1$  retrial system with two types of search of customers from the orbit. *Stochastic Analysis and Applications*, 31(1) :92–107, 2013.
- [32] N. DJELLAB. On the  $M/G/1$  retrial queue with feedback. 2005.
- [33] J. H. Dshalalow. *Frontiers in queueing : models and applications in science and engineering*, volume 7. CRC press, 1997.
- [34] A. Dudin, T. Deepak, V. C. Joshua, A. Krishnamoorthy, and V. Vishnevsky. On a  $BMAP/G/1$  retrial system with two types of search of customers from the orbit. In *International Conference on Information Technologies and Mathematical Modelling*, pages 1–12. Springer, 2017.
- [35] A. N. Dudin, A. Krishnamoorthy, V. Joshua, and G. V. Tsarenkov. Analysis of the  $BMAP/G/1$  retrial system with search of customers from the orbit. *European Journal of Operational Research*, 157(1) :169–179, 2004.
- [36] C. D’Apice, T. De Simone, R. Manzo, and G. Rizelian. Priority service of primary customers in the  $M/G/1/r$  retrial queueing system with server searching for customers. *J. Information Theory and information processing*, 4(1) :13–23, 2004.
- [37] C. D’Apice and R. Manzo. Search for customers in a finite capacity queueing system with phase-type distributions. *Information Processes, Electronic Scientific Journal*, 3(1) :61–69, 2003.
- [38] A. Eldin. Approach of theoretical description of repeated call attempts. *Ericsson Thecnics*, 23(3) :346–407, 1967.
- [39] G. Falin and J. G. Templeton. *Retrial queues*, volume 75. CRC Press, 1997.
- [40] S. Gao and J. Wang. Performance and reliability analysis of an  $M/G/1-G$  retrial queue with orbital search and non-persistent customers. *European Journal of Operational Research*, 236(2) :561–572, 2014.
- [41] S. Gao and J. Wang. Stochastic analysis of a preemptive retrial queue with orbital search and multiple vacations. *RAIRO-Operations Research*, 54(1) :231–249, 2020.
- [42] S. Gao and J. Zhang. Strategic joining and pricing policies in a retrial queue with orbital search and its application to call centers. *IEEE Access*, 7 :129317–129326, 2019.

- [43] E. Gelenbe, G. Pujolle, E. Gelenbe, and G. Pujolle. *Introduction to queueing networks*, volume 2. Wiley New York, 1998.
- [44] S. Ghimire, G. B. Thapa, R. Ghinire, and S. Silvestrov. A survey on queueing systems with mathematical models and applications. *American Journal of Operation Research*, 7(1) :1–14, 2017.
- [45] A. Gómez-Corral and T. Phung-Duc. Retrial queues and related models. *Annals of Operations Research*, 247(1) :1–2, 2016.
- [46] V. C. Joshua, A. P. Mathew, and A. Krishnamoorthy. A retrial queueing system in which server searches to accumulate customers for optimal bulk serving. In *International Conference on Distributed Computer and Communication Networks*, pages 311–321. Springer, 2020.
- [47] J. Kim and B. Kim. A survey of retrial queueing systems. *Annals of operations research*, 247(1) :3–36, 2016.
- [48] K. Kirupa and K. U. Chandrika. Unreliable batch arrival retrial g-queue with fluctuating modes of service, preemptive priority and orbital search. *Int J Math Trends Technol (Special Issue)*, pages 45–53, 2018.
- [49] A. Krishnamoorthy, T. Deepak, and V. Joshua. An  $M/G/1$  retrial queue with nonpersistent customers and orbital search. *Stochastic Analysis and Applications*, 23(5) :975–997, 2005.
- [50] A. Krishnamoorthy, V. Joshua, and A. P. Mathew. A retrial queueing system with abandonment and search for priority customers. In *International Conference on Distributed Computer and Communication Networks*, pages 98–107. Springer, 2017.
- [51] A. Krishnamoorthy, V. Joshua, and A. P. Mathew. A retrial queueing system with multiple hierarchical orbits and orbital search. In *International Conference on Distributed Computer and Communication Networks*, pages 224–233. Springer, 2018.
- [52] A. Krishnamoorthy, V. Narayanan, V. Joshua, and T. Deepak. An  $MAP|PH|1$  retrial queue with service interruption and orbital search. 2009.
- [53] A. Krishnamoorthy, V. Vishnevsky, T. Deepak, and V. Joshua. On a retrial queueing model with orbital search of customers—application to telecommunication on highways1.
- [54] V. Kulkarni and H. Liang. Retrial queues revisited. *Frontiers in queueing : Models and applications in science and engineering*, 7 :19, 1997.
- [55] B. K. Kumar, G. Vijayalakshmi, A. Krishnamoorthy, and S. S. Basha. A single server feedback retrial queue with collisions. *Computers & Operations Research*, 37(7) :1247–1255, 2010.
- [56] L. Lakatos, L. Szeidl, and M. Telek. *Introduction to queueing systems with telecommunication applications*, volume 388. Springer, 2013.
- [57] C. Lakshmi and S. A. Iyer. Application of queueing theory in health care : A literature review. *Operations research for health care*, 2(1-2) :25–39, 2013.
- [58] J. Li, T. Li, and M. An. An  $M/M/1$  retrial queue with working vacation orbit search and balking. *Engineering Letters*, 27(1) :97–102, 2019.

- [59] A. P. Mathew, A. Krishnamoorthy, and V. Joshua. *M/MAP/M/* retrial queue with search for customers completed service in the offer zone. In *12th International Workshop on Retrial Queues and Related Topics*, pages 20–21, 2018.
- [60] A. P. Mathew, A. Krishnamoorthy, and V. C. Joshua. *A retrial queueing system with orbital search of customers lost from an offer zone*, pages 39–54. Springer, 2018.
- [61] S. P. B. Murugan and R. Vijaykrishnaraj. A bulk arrival retrial queue with orbital search and exponentially distributed multiple working vacation. In *AIP Conference Proceedings*, volume 2177, page 020065. AIP Publishing LLC, 2019.
- [62] M. F. Neuts and M. Ramalhoto. A service model in which the server is required to search for customers. *Journal of Applied Probability*, 21(1) :157–166, 1984.
- [63] D. S. Nguyen. Application of queuing theory in service design. In *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 837–840. IEEE, 2017.
- [64] M. Nila and D. Sumitha. Batch arrival retrial queueing model with starting failures, customer impatience, multi optional second phase and orbital search. *International Journal of Engineering Research Technology (IJERT)*, 2019.
- [65] A. Outamazirt, M. Escheikh, D. Aïssani, K. Barkaoui, and O. Lekadir. Performance analysis of the *M/G/c/c+r* queuing system for cloud computing data centres. *International Journal of Critical Computer-Based Systems*, 8(3-4) :234–257, 2018.
- [66] T. Phung-Duc. Retrial queue for cloud systems with separated processing and storage units. In *International Conference on Queueing Theory and Network Applications*, pages 143–151. Springer, 2015.
- [67] T. Phung-Duc. Retrial queueing models : A survey on theory and applications. *arXiv preprint arXiv :1906.09560*, 2019.
- [68] J. Radha, K. Indhira, and V. Chandrasekaran. A group arrival retrial g-queue with multi optional stages of service, orbital search and server breakdown. In *IOP Conference Series : Materials Science and Engineering*, volume 263, page 042150. IOP Publishing, 2017.
- [69] J. Radha, K. Indhira, and V. Chandrasekaran. Analysis of *M [X]/Gk/1* retrial queue with bernoulli-2 vacation, search for customers and hot stand-by. *International Journal of Process Management and Benchmarking*, 10(2) :196–223, 2020.
- [70] P. Rajadurai, V. Chandrasekaran, and M. Saravananarajan. Steady state analysis of batch arrival feedback retrial queue with two phases of service, negative customers, bernoulli vacation and server breakdown. *International Journal of Mathematics in Operational Research*, 7(5) :519–546, 2015.
- [71] K. P. Rani, K. U. Chandrika, and J. E. A. Bagyam. Bulk arrival retrial queue with non persistent customers, active breakdown, delayed repair and orbital search. *International Journal of Scientific Engineering Research*, 8(12), 2017.
- [72] S. Sasikala and K. Indhira. Bulk service queueing models—a survey. *International Journal of Pure and Applied Mathematics*, 106(6) :43–56, 2016.

- [73] G. Sekar, G. Ayyappan, and A. Subdramanian. Single server retrial queueing system with orbital search under erlang k service. *Int. J. Appl. Math*, 3 :454–463, 2012.
- [74] M. Sharma and R. Singh. Study of retrial chain with non continued clients along with orbital search. *International Journal of Mathematics And its Applications*, 55 :879–891, 2017.
- [75] D. Sumitha and K. U. Chandrika. Retrial queueing system with starting failure, single vacation and orbital search. *International journal of computer applications*, 40(13) :29–33, 2012.
- [76] L. Takács. A single-server queue with feedback. *Bell system Technical journal*, 42(2) :505–519, 1963.
- [77] J. C. Varghese and A. Krishnamoorthy. *Retrial Queues with Orbital Search*. PhD thesis, Department of Mathematics, 2003.
- [78] P. Wüchner, J. Sztrik, and H. de Meer. *Structured Markov Chains Arising from Homogeneous Finite Source Retrial Queues with Orbital Search*. Citeseer, 2008.
- [79] P. Wüchner, J. Sztrik, and H. de Meer. Finite-source *M/M/S* retrial queue with search for balking and impatient customers from the orbit. *Computer Networks*, 53(8) :1264–1273, 2009.
- [80] P. Zeephongsekul, A. Bedford, J. Broberg, P. Dimopoulos, and Z. Tari. *Queuing theory applications to communication systems : Control of traffic flows and load balancing*, pages 991–1022. RMIT University, 2006.