

N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE
SCIENTIFIQUE



UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL ABBES

THESE DE DOCTORAT

Présentée par

Aicha MESSABIHI

Spécialité : MATHEMATIQUES

Option : PROBABILITES

Intitulée

**Etude des systèmes de files d'attente avec
impatience et feedback**

Soutenu le 14/01/2021

Devant le jury composé de :

Président :

Mr. Boubaker MECHAB (Professeur à l'Université Djillali Liabès)

Examineurs :

Mr. Hakim OUADJED (MCA à l'Université de Mascara)

Mr. Tayeb BLOUHI (MCA à l'Université USTO Oran)

Directeur de thèse :

Madame Amina Angelika BOUCHENTOUF (Professeur à l'Université
Djillali Liabès)

Année universitaire : 2020/2021

Dédicace

Je tiens à remercier les membres de ma famille pour tout ce qu'ils ont fait pour moi. Mon frère et mes soeurs m'ont souvent soutenu, encouragé et offert leur amitié. En ce qui concerne mes parents, ils ont joué un rôle très important durant toute ma vie surtout ma mère qui a suivi mon éducation avec le plus grand intérêt. Son encouragement, sa motivation débordante et souvent contagieuse ainsi que sa confiance. Elle m'a soutenue tout au long de mes études. J'ai partagé avec elle toutes mes réussites, et je n'ai pas pu surmonter certaines difficultés que grâce à elle. J'ai de la chance d'avoir une mère exceptionnelle et j'en remercie Dieu.

Remerciements

Au terme de ce travail de recherche, de cette expérience professionnelle et humaine fortement marquante, j'ai l'honneur et l'immense plaisir de remercier toutes les personnes ayant contribué d'une manière ou d'une autre à l'accomplissement de cette thèse de doctorat. Tout d'abord, mes remerciements les plus cordiaux sont adressés à ma directrice de thèse Professeur Amina Angelika Bouchentouf. L'empreinte de ses conseils avisés et de ses idées pertinentes est incontestablement visible dans ce manuscrit. Par ailleurs, je ne peux penser à son encadrement sans me souvenir de sa disponibilité à chaque fois que j'ai eu besoin d'elle, de son écoute ainsi que de la liberté qu'elle m'a accordée sur l'organisation du travail. Tous ces points positifs de son encadrement m'ont permis de mener ce travail de thèse à bon port dans un cadre particulièrement agréable et d'être sincèrement satisfaite de ce que j'ai appris durant mes années de thèse. J'adresse également mes remerciements les plus respectueux aux membres du jury : Professeur Boubaker Mechab, Docteur Hakim Ouadjed et Docteur Tayeb Blouhi de l'honneur qu'ils m'ont fait en acceptant de rapporter et d'examiner cette thèse de doctorat.

Abstract

The main objective of this thesis is to analyze different queueing systems with customers impatience and feedback, in particular we focus on heterogenous servers queues.

At first, we consider a heterogenous two-server queueing model with Bernoulli feedback, reverse balking, renegeing and retention of renegeed customers. Secondary, we deal with a queueing model with two heterogenous servers, no waiting line, Bernoulli feedback, renegeing, and retention of renegeed customers. Using the recursive method, we obtain the steady state probabilities for the proposed queueing models. Then, we derive useful performance measures of the considered systems. Numerical analyzes are carried out to show the impact of systems parameters on the behaviors of the systems. A cost model as well as a cost profit analysis are presented for the second queueing system considered in this thesis.

Keywords :

Queueing models, feedback, balking, renegeing.

Résumé

L'objectif de cette thèse est d'analyser différents systèmes de files d'attente avec clients impatients et feedback, nous nous intéressons plus précisément aux files d'attentes avec serveurs hétérogènes.

En premier lieu, nous traitons un modèle de files d'attente avec serveurs hétérogènes, Bernoulli feedback, dérobade inversée (reverse balking), abandon (reneging) et retention de clients abandonnés. En second lieu, nous considérons un système d'attente sans buffer avec serveurs hétérogènes, Bernoulli feedback, dérobade (balking), abandon et retention de clients abandonnés. En utilisant la récursivité, nous obtenons l'état stationnaire des probabilités associée aux modèles de files d'attente proposés. Les mesures de performances des systèmes considérées sont obtenues. Des résultats numériques sont donnés pour montrer l'influence des différents paramètres sur le comportement du système. Une analyse de coût bénéfique est présentée pour le second système de file d'attente considéré dans cette thèse.

Mots clés :

Système de files d'attente, feedback, dérobade, abandon.

مختص

الهدف الرئيسي من هذه الرسالة هو تحليل مختلف أنظمة قوائم انتظار مع نفاذ صبر الزبائن وإمكانية العودة وفق نظام برتولي. على وجه الخصوص نركز دراستنا على قوائم انتظار ذات الخوادم المتعددة غير متجانسة.

أولاً، ندرس نموذج لنظام قائمة انتظار ذات خادمين ذوي رتبة مختلفة في الخدمة، بتغذية راجعة بتوزيع برتولي مع العزوف والعزوف العكسي مع إمكانية التصل الزبائن و الاحتفاظ بالزبائن المتصلين باستعمال طريقة تراجعية.

ثانياً، ندرس نموذج لنظام بخادمين غير متجانسين، ذات سعة قائمة انتظار معدومة. بتغذية راجعة تتبع التوزيع البرتولي، مع إمكانية التصل الزبائن و الاحتفاظ بالزبائن المتصلين عن الخدمة.

حصل على التوزيع الاحتمالي لحالة استقرار أنظمة قوائم الانتظار المقترحة ثم نستطيع مقاييس أداء مفيدة للأنظمة المدروسة. نقدم التحليلات العددية لتوضيح تأثير معاملات الأنظمة على أداء النظام واستقراره. نقوم أيضاً بتحليل تكلفة وتحليل تكلفة ربح نظام قائمة الانتظار الأول المدروس في الأطروحة.

الكلمات المفتاحية

أنظمة قوائم انتظار، نظام تغذية راجعة بتوزيع برتولي، العزوف العكسي، التصل

Table des matières

1	Introduction et présentation	1
1.1	Un peu d'histoire sur le développement de la théorie des files d'attente	2
1.2	Description d'un système de files d'attente	4
1.2.1	Définition d'une file d'attente	4
1.2.2	Les caractéristiques d'une file d'attente	4
1.2.3	Notations	6
1.2.4	Les mesures de performances	7
1.2.5	Préliminaires mathématiques	8
1.2.6	Les modèles de files d'attente	10
1.2.6.1	Modèle de files d'attente Markovien	10
1.2.6.2	Modèle de files d'attente non Markovien	13
1.3	Files d'attente avec feedback	17
1.4	Files d'attente avec serveurs hétérogènes	18
1.5	Files d'attente avec clients impatientes	19
1.5.1	Notation de files d'attente avec clients impatientes	20
1.5.2	Disciplines de service	21
1.6	Comportement des clients impatientes dans un système de files d'attente	22
1.6.1	Dérobade (balking)	22
1.6.2	Abandon (reneging)	26
1.7	Contribution de la thèse	30
1.8	Présentation du manuscrit	34
	Bibliographie	36
2	Heterogenous two server queueing system with reverse balking and reneging	43
2.1	Introduction	44
2.2	Model description and main results	46
2.2.1	Steady-State Solution	47
2.3	Performance measures and numerical analysis	48
2.3.1	Impact of arrival and service rates λ , μ_1 and μ_2 respectively on the characteristics of the system	49
2.3.2	Impact of impatience rate ξ on the characteristics of the system	51

2.3.3	Impact of balking probability θ' on the characteristics of the system	52
2.3.4	Impact of feedback probability δ' on the characteristics of the system	53
2.3.5	Impact of retention of renegeed customers probability γ' on the characteristics of the system	54
2.4	Economic analysis	55
2.4.1	Effect of the probability of retaining the renegeed customers γ' on the costs of the system	56
2.4.2	Effect of probability of retaining the renegeed customers γ' on total expected profit	57
2.4.3	Effect of the probability of balking θ' on the costs of the system	58
2.5	Conclusion and future work	58
Bibliography		59
3	Queueing system with renegeing and no waiting line	61
3.1	Introduction	62
3.2	Model description	64
3.3	Steady state solution	65
3.4	Performance measures	66
3.5	Numerical solution and graphical representation	67
3.5.1	Impact of service rates on the characteristics of the system	67
3.5.2	Impact of renegeing (impatient) rates on the characteristics of the system	68
3.5.3	Impact of probability of retention renegeed customers on the characteristics of the system	70
3.5.4	Impact of probability of feedback customers on the characteristics of the system	72
3.6	Conclusion	73
Bibliography		74
Conclusion et perspectives		76
Bibliographie générale		77

Chapitre 1

Introduction et présentation

La théorie des files d'attente est l'un des outils analytiques les plus puissants pour la modélisation de la dynamique des systèmes logistiques et de communications. Cette théorie a pour but d'étudier différents systèmes d'attente et évaluer les résultats mathématiques en termes de différentes mesures de performance. Les résultats de la théorie des files d'attente sont nécessaires pour obtenir les caractéristiques des modèles d'attente. Ils fournissent des informations nécessaires aux décideurs en prédisant diverses caractéristiques des lignes d'attente telles que le temps d'attente moyen dans la file d'attente, le nombre moyen de clients dans la file d'attente, le temps de service,... etc.

Les files d'attente avec clients impatientes (balking et/ou reneging) sont devenues un domaine vital dans la théorie des files d'attente. Ces systèmes ont une large application dans l'ingénierie des télécommunications, les réseaux informatiques, les systèmes de production et d'autres systèmes stochastiques. Une autre caractéristique très importante des systèmes de files d'attente est la présence du mécanisme de Bernoulli feedback. Dans de nombreuses situations de file d'attente, les clients peuvent être servis plusieurs fois pour certaines raisons. Si le service d'un client n'est pas satisfaisant, il peut être réessayé encore et encore jusqu'à la fin du service. Ces modèles d'attente apparaissent dans la modélisation stochastique de nombreuses situations réelles. Par exemple, dans la transmission de données, un paquet transmis de la source à la destination peut être renvoyé et il peut continuer ce processus jusqu'à ce que le paquet soit finalement transmis.

Les files d'attente sont des phénomènes que l'on rencontre dans de très nombreux domaines. Par exemple dans les banques, dans les aéroports, dans les bureaux de poste,... etc. Il s'agit de systèmes de files d'attente classiques. D'autres systèmes sont également considérés tels que les centres d'appels, les services d'urgence des hôpitaux, les bureaux de poste et les ordinateurs. Pour ces systèmes, les serveurs sont généralement différents les uns des autres. Ils peuvent posséder différents types de compétences. Même lorsqu'ils ont la même compétence, leur capacité peut être différente. De tels systèmes hétérogènes ne sont pas encore suffisamment étudiés dans la littérature. Dans cette thèse nous nous intéressons à l'étude de différents modèles de files d'attente avec serveurs hétérogènes, Bernoulli feedback et clients impatientes qui s'avèrent d'une por-

tée limitée en raison de la complexité des résultats. Au cours des dernières décennies, beaucoup d'efforts ont été consacrés à l'étude de ce type de systèmes d'attente, en raison de leurs vastes applications dans de nombreux domaines tels que les centres d'appels, les réseaux informatiques, les systèmes de communication et de télécommunication, les systèmes de fabrication flexibles, la modélisation des performances,... etc. Dans ce chapitre, nous fournissons dans la section 1.1 un peu d'historique sur l'étude de la théorie de files d'attente ainsi qu'une revue de la littérature. Dans la section 1.2 nous rappelons un vocabulaire fondamental et les résultats de différents systèmes de files d'attente. Ensuite, dans les sections 1.3 et 1.4 nous donnons successivement des définitions sur les systèmes de files d'attente avec feedback et clients impatientes. Après cela, dans les sections 1.5 et 1.6, nous présentons le concept de systèmes de files d'attente avec des clients impatientes et leur comportement dans le système "dérobade et abandon". La section 1.7 est consacré à la contribution de le thèse et nous terminons ce chapitre par une présentation du manuscrit dans la section 1.8.

1.1 Un peu d'historique sur le développement de la théorie des files d'attente

La théorie des files d'attente est une théorie de la recherche opérationnelle relevant du domaine de probabilité. Elle a été développé pour fournir des modèles mathématiques pour prédire l'évolution des files d'attente, l'élaboration et l'optimisation des indicateurs de mesures des performances du service. Le caractère massif des demandes de service, ainsi que la diversité des facteurs d'influence externe conduisent naturellement à une formulation dans le langage de la théorie des probabilités.

La première publication concernant ce domaine était sur le sujet d'attente et le nombre d'appels de Johansen publié en 1910. L'influence de la mécanique statistique de l'époque l'a vraisemblablement conduit à introduire la notion d'équilibre statistique qui lui a permis d'obtenir la formule pour la probabilité de rejet qui porte aujourd'hui son nom. Ses travaux ont constitué le point de départ pour des recherches s'articulant autour de la confirmation ou de l'infirmité de ses résultats en télétrafic et du point de vue théorique sur la généralisation de sa fameuse formule. La première généralisation a été effectuée par Engset (1918) dans le cas d'une source finie. D'autres modèles ont été étudiés par O'Dell et Gibson (1926). En 1928, Fry publia un ouvrage concernant la probabilité et ses applications en ingénierie comprenant un chapitre sur les problèmes de télétrafic. Parmi les précurseurs de cette période, nous citons Vulot (1925), Cromelin (1932,1934), Wilkinson (1931), et Khintchine (1960).

Dans les années 30, l'introduction des processus de naissance et de mort marque un tournant nouveau dans la description des phénomènes d'attente. Kolmogorov s'est intéressé à cette question en 1931. En 1933, il publie son ouvrage où il propose son axiomatique de la théorie des probabilités qui est adopté jusqu'à aujourd'hui. L'idée d'Erlang sur l'équilibre statistique est déjà assimilée à la notion de mesures stationnaire de processus markovien. Cependant, Pollaczek remarque en 1934 que l'équilibre

statistique ne permet pas toujours de décrire la micro-dynamique de tous les processus aléatoires en théorie de file d'attente. A cette époque (1934), Pollaczek et Khintchine s'intéressent au cas de durées de service arbitraires (la formule d'Erlang est obtenue dans le cas d'une distribution exponentielle). Les formules obtenues par Pollaczek vers 1950 sont déduites de manière plus simple par Kendall (1951) par la méthode de la chaîne de Markov induite (ou incluse).

Les travaux du suédois Palm (1936, 1938) et surtout en 1943 ont été à la base de nombreuses études sur la variabilité des charges téléphoniques. Le flux de Palm appelé également processus de renouvellement a trouvé des applications intéressantes en théorie de fiabilité et de maintenance. Gnedenko a recueilli une série d'articles de Khintchine proposant diverses extensions parmi lesquelles les formules connues aujourd'hui sous le nom de formules de Pollaczek-Khintchine.

L'introduction de nouvelles classes de processus aléatoires a permis d'élaborer de nouvelles approches intéressantes qui sont les processus semi-markoviens (Smith (1953, 1956), Cox (1955)), les processus markoviens à trajectoire linéaires par morceaux (Belyaev (1962)) qui est une méthode similaire à celle utilisée par Takacs (1955) pour obtenir l'équation intégral-différentielle de la loi de la durée d'attente. Les processus de marquage (Franken et al. (1981)) constituent un cadre de formalisation plus élaboré. Les processus de diffusion conduisent à des méthodes approximatives très précisées par les praticiens (Courtois (1977) et Disney et al. (1972)).

Du point de vue mathématiques, il est intéressant de citer la formule de conservation de Little (1961). L'analyse opérationnelle bien que souvent controversée constitue une approche complémentaire intéressante à la théorie classique des files d'attente. Elle contribue à prouver la validité des résultats classiques pour les systèmes complexes lorsque les hypothèses conventionnelles ne peuvent être justifiées. Elle permet en outre d'expliquer la robustesse de ce type de résultats.

La méthode des événements fictifs dont l'idée remonte à Van Dantzig (1960) constitue une autre approche non conventionnelle, appelée également méthode des marques collectives ou méthodes des catastrophes (Kleinrock (1975), Klimov (1966)). Elle permet de donner une interprétation probabiliste aux résultats complexes souvent obtenus en termes de transformées de Laplace ou de fonctions génératrices. Cette méthode a été utilisée avec succès pour l'étude des systèmes avec priorités (Bronstein et Dukhovny (1976)) et plus récemment, des systèmes avec vacances (Gong (1992)).

Les travaux de Lindley, Kingman et Marshall sont à la base de nombreuses méthodes mathématiques plus complexes axées sur les questions d'approximation et de stabilité. Nous pourrions se référer à ce sujet aux ouvrages de Borokov (1976), Stoyan (1983). L'élaboration de l'approche des opérateurs de la théorie de stabilité (connue sous le nom de méthode de stabilité forte) permet d'obtenir des estimations qualitatives de stabilité avec un calcul exact des constantes Aissani (1983) et Kartashov (1996).

La théorie de files d'attente est un pôle de recherche important dans un monde plus soucieux et exigeant comme l'industrie, le commerce, marketing, télécommunications, qui demande de plus en plus d'exploitation et développement. Cette théorie a contribué à une concurrence internationale qui a poussé les chercheurs à étudier des systèmes

complexes en ajoutant aux systèmes classiques quelques propriétés qui caractérisent les serveurs (homogénéité, hétérogénéité, vacances,...) ou/et les arrivées (rappel, dérobade, feedback, abandon, classe des clients) qui ont conçu une large attention et qui font l'objet de la recherche jusqu'à ce jour (Choi et Chang (1999), Feng et al. (2000), Wang et Chang (2002), Legros et Jouini (2017), Bouchentouf et Guendouzi (2019), Kim (2020), Gao et Wang(2020)),... etc.

1.2 Description d'un système de files d'attente

1.2.1 Définition d'une file d'attente

La théorie de file d'attente consiste à étudier des systèmes où les clients se présentent à un dispositif de service appelé serveur. Le client occupe le serveur pendant un certain temps, les autres clients doivent attendre avant d'être servi formant ainsi une file d'attente.

Exemple :

- Réseau informatiques : serveur = routeur, client = paquet
- Ateliers : serveur = machines, client = tâche.

Les mesures de performances les plus utiles sont :

- Taille moyenne de la file d'attente.
- Taux d'utilisation du serveur.
- Temps moyen d'attente d'un client.

Pour évaluer ces mesures de performances il faut bien décrire le modèle de notre file d'attente .

1.2.2 Les caractéristiques d'une file d'attente

Dans le cadre de la théorie des files d'attentes, l'amélioration de l'analyse des files d'attentes repose surtout sur le choix du modèle approprié en tenant compte des caractéristiques suivantes :

- La population (clients),
- Le nombre de serveurs,
- La nature du processus d'arrivées et du service,
- La discipline de service.



FIGURE 1.1 – Exemple d’un système de file d’attente

1. La population

Dans la théorie de files d’attente nous appelons population la source de clients potentiels qui se forment par deux cas possibles. Deux cas se présentent : le nombre de clients est infiniment grand en tout temps. C’est le cas des clients des supermarchés, des banques, des centres d’appels...etc. Le nombre de clients est fini ou limité par exemple dans une salle d’attentes il y’a une capacité qui limite le nombre de clients.

2. Le nombre de serveurs

La capacité de service dépend de la rapidité de ce dernier et du nombre de serveurs disponibles. Les systèmes de files d’attente fonctionnent soit avec un seul serveur soit avec des serveurs multiples. Les exemples de systèmes de files d’attentes avec serveur unique sont nombreux : les magasins avec une seule caisse, les véhicules qui attendent aux carrefours, les patients qui attendent un traitement...etc., et pour les systèmes avec plus d’un serveur : les banques, les appels en attente dans un centre d’appel, les machines des usines industrielles...etc. Nous constatons que le taux de service peut être équidistribué pour tout le(s) serveur(s) c’est ce qu’on appelle un système de file d’attente avec serveurs homogènes et dans le cas contraire lorsque le taux de service se diffère nous disons que c’est un systèmes de file d’attente avec serveur(s) hétérogène(s).

3. La nature du processus des arrivées et des services

Les files d’attente résultent de la variabilité des tendances d’arrivée et de service. Elles se forment à cause des congestions temporaires causées par le degré élevé de variation des arrivées et leur service. D’habitude nous considérons ses variations par des distributions théoriques de probabilité. Dans les principaux modèles utilisés, nous supposons que le nombre d’arrivées dans un intervalle donné suit la loi de poisson de taux λ alors que le temps de service suit la loi exponentielle de taux μ . En général, la distribution de Poisson donne une bonne

estimation du nombre de clients qui arrivent par unité de temps (par exemple le nombre de clients par minute, le nombre d'accident par an, ...etc.) et aussi donne une bonne approximation des temps de service. Généralement, les temps de services sont soit assez courts, certains sont proche de zéro ou assez long. C'est la caractéristique de la loi exponentielle .

4. La discipline de service

La discipline du service concerne l'ordre de traitement des clients. Souvent, la discipline la plus fréquente est premier arrivé premier servi (FIFO). Elle est appliquée dans les banques, les entreprises de service, ...etc. Donc la discipline de file d'attente présente la manière dont un serveur choisit le prochain client à partir d'une file d'attente. Les disciplines fréquentes sont :

- (First In First Out) : premier arrivé premier servi.

- LIFO (Last In First Out) : dernier arrivé premier servi.

- Priorité relative : les clients sont servis suivant leur importance ou sur base des exigences de leur service.

- Priorité absolu : le service d'un client est interrompu lorsqu'un client de priorité supérieure se présente dans la file d'attente. Le client dont ce service est interrompu est remis en tête de la file.

- Random : le client accède au serveur de manière aléatoire, indépendamment de l'ordre des arrivées.

- PS (Processor sharing) : tous les clients sont servis en même temps avec le même taux de service.

1.2.3 Notations

La notation introduite par David George Kendall en 1953 de forme $A/S/C$ permet de décrire les éléments qui constituent une file d'attente simple tels que :

- A : la distribution des interarrivées.
- S : la distribution de temps de service.
- C : le nombre de serveurs dans le système.

Généralement nous supposons que :

- S serveurs existent en parallèle et une discipline FIFO qui alimente tous les serveurs.
- A_1, A_2, \dots sont des variables aléatoires indépendantes et identiquement distribuées.
- S_1, S_2, \dots sont des variables aléatoires indépendantes et identiquement distribuées.
- Les A_i et les S_i sont indépendants.

Les paramètres A et S sont représentés par les distributions suivantes :

- M : distribution exponentielle (Markovienne).

- E_k : distribution Erlang d'ordre k .
- G : distribution générale (aléatoire).
- D : distribution déterministe (fixée).
- GI : distribution générale indépendante.

La notation de Kendall a été développée par Lee (1966) en ajoutant un autre symbole $A/S/C/D/E$ pour représenter plus de caractéristiques au système de file d'attente. Dans cette notation D représente la taille maximale de la file et E désigne la discipline du serveur qui est soit FIFO, LIFO, ...etc. Par exemple :

- $M/G/1$ indique que c'est une file d'attente à serveur unique dans le système dont la distribution des temps des interarrivées est exponentielle, la distribution du temps de service est arbitraire (générale) .

- $M/D/1/\infty$ est une file d'attente a capacité illimitée admettant une distribution exponentielle pour le temps des interarrivées, le temps de service est déterministe, un seul serveur et la discipline du service est *FCFS*.

Il existe une autre notation définie par $M^x/G/1$ qui considère une file d'attente dont les arrivées sont poissoniennes, le temps de service suit une distribution générale, serveur unique, la capacité de la file est illimitée (infinie) et discipline de service FIFO. L'exposant x se réfère aux clients arrivants au système sous forme d'un groupe (batch).

1.2.4 Les mesures de performances

L'évolution permanente des systèmes de files d'attente souligne le besoin croissant d'outils facilitant l'étude de leur comportement. En effet, le développement d'un système complexe demande non seulement une modélisation qualitative mais aussi une validation a priori des performances des systèmes lors de la phase de la conception.

Les mesures de performances les plus recommandées dans une analyse de files d'attente sont :

- Le nombre moyen de clients qui attendent dans la file ou dans le système.
- le temps moyen d'attente dans la file et dans le système.
- Le taux d'utilisation du système, c'est à dire le pourcentage de la capacité utilisée.
- Le cout associé au niveau de service (capacité) mis en place.
- La probabilité qu'un client potentiel attende pour être servi.
- L'intensité de trafic.

Parmi ces mesures de performances, le taux d'utilisation du système nécessite quelques éclaircissement. Il reflète le pourcentage de l'occupation des serveurs plutôt que leur inactivité. Il est logique de penser qu'une bonne gestion des ressources implique un

taux d'utilisation de cent pour cent. Cependant, le fait d'augmenter le taux d'utilisation revient à augmenter à la fois le nombre de clients qui attendent et le temps moyen d'attente, ces deux mesures croient indéfiniment lorsque le taux d'utilisation s'approche de cent pour cent. Si tous les serveurs sont occupés nous sommes sûres que les clients potentiels qui arrivent vont attendre et cela contredit notre approche de cent pour cent pour notre taux d'utilisation surtout dans les opérations normales. Ceci nous ramène à essayer d'équilibrer le système de telle sorte que la somme des coûts de service et d'attente soit minimale.

1.2.5 Préliminaires mathématiques

Nous décrivons très brièvement les outils mathématiques utilisés pour la théorie de file d'attente.

1. La loi de Little

La loi de Little est l'une des relations fondamentales dans la théorie de files d'attente développée par Little en 1960. Généralement cette loi ne concerne que le régime permanent du système.

Formule de Little : Le nombre moyen de clients \bar{N} , le temps moyen passé dans le système \bar{T} , et est le taux d'entrée dans le système λ , se relie de la façon suivante :

$$\bar{N} = \lambda \bar{T}.$$

La loi de Little peut s'appliquer aussi en considérant uniquement l'attente dans la file (sans le service). Elle permet alors de relier le nombre moyen de client en attente \bar{N}_Q au temps moyen d'attente d'un client \bar{T}_Q avant le service :

$$\bar{N}_Q = \lambda \bar{T}_Q.$$

Enfin, on peut appliquer la loi de Little en ne considérant que les serveurs. De la même manière cette loi peut relier le nombre moyen de clients en service \bar{N}_s au temps moyen de séjour d'un client \bar{T}_s dans le service qui n'est rien d'autre que le temps moyen de service par la relation :

$$\bar{N}_s = \lambda \bar{T}_s.$$

Ces trois relations ne sont pas indépendantes, nous pouvons déduire l'une d'entre elles à partir des deux autres relations en remarquant que $\bar{N} = \bar{N}_Q + \bar{N}_s$ et $\bar{T} = \bar{T}_Q + \bar{T}_s$.

2. Intensité de trafic :

Dans les modèles de files d'attente, il existe certaines relations de base (entre certains paramètres et les mesures de performance) qui permettent de déterminer les mesures de performances désirées .

Par exemple pour une file d'attente $M/M/c$ nous assumons que λ est le taux d'arrivée et μ est le taux de service. Une importante mesure est donnée à partir de ces deux paramètres qui est le taux d'utilisation du système (intensité de trafic) représenté par le rapport entre la demande et le taux de service par :

$$\rho = \frac{\lambda}{c\mu}, \quad \text{où } c \text{ est le nombre de serveurs dans le système.}$$

La condition nécessaire pour dire que le système $M/M/c$ est stable est l'existence d'une intensité de trafic inférieure strictement à 1 ($\lambda < c\mu$). Il n'y a aucune utilité à analyser les systèmes dans lesquels $\lambda > c\mu$ car il est évident que dans tels cas le système diverge.

Pour les systèmes de files d'attente $M/M/c$ dont les serveurs sont hétérogènes l'intensité de trafic est égale à $\rho = \frac{\lambda}{\sum_{i=1}^c \mu_i}$.

$$\sum_{i=1}^c \mu_i$$

3. Description d'une file d'attente à l'état d'équilibre

Lorsque nous commençons à analyser un système de files d'attente, l'état de ce dernier dépend beaucoup de l'état initial et du temps écoulé. On dit alors que le système est en situation transiente et son étude est alors très complexe.

C'est pourquoi dans la théorie des files d'attente, nous préférons faire l'étude une fois que le système a atteint sa situation d'équilibre où les états du système sont essentiellement indépendantes de l'état initiale.

Nous supposons en quelque sorte que le système est en opération depuis un très long moment et satisfait la condition d'ergodicité (stabilité) ainsi le processus d'arrivé spécifie les instants auxquels les clients arrivent dans le système qui est défini par la distribution des intervalles séparant deux arrivées consécutives, les clients arrivent au système en décrivant un processus de naissance. Ils peuvent être réguliers et leurs arrivées sont espacées par un temps égale à Δt . Le modèle le plus simple est le plus courant est celui des arrivées complètement aléatoires qui est caractérisé par le processus de poisson ainsi le nombre de clients dans le système à l'instant t est un processus stochastique à espace d'état discret à temps continu et sans mémoire donc c'est une chaine de Markov à temps continu.

Les clients demandent le service qui est définit par une distribution aléatoire. Dans la majorité des cas, nous supposons que la population (clients) est homogène; les services demandés sont identiquement distribués. Dans la pratique nous rencontrons souvent la distribution exponentielle de taux $\mu(n)$ qui est caractérisé par la propriété d'absence de mémoire ainsi nous pouvons définir le taux de transition d'un état n quelconque vers l'état $n - 1$ est égal à $\mu(n)$ et le taux de transition d'un état n vers l'état $n + 1$ est égale à $\lambda(n)$. Ce type de chaine de Markov est connu sous le nom de processus de naissance et de mort.

C'est à partir de l'état stationnaire que nous pouvons analyser les performances

du système, les résultats analytiques du comportement transient du système de files d'attente nous aide à étudier le comportement dynamique du système.

1.2.6 Les modèles de files d'attente

Les modèles de files d'attente sont modélisés par des processus stochastiques markovien et non markovien adaptés au contexte du modèle qui repose sur certaines hypothèses.

1.2.6.1 Modèle de files d'attente Markovien

Les systèmes markoviens sont des systèmes où les temps des inter-arrivées et le temps de service sont des variables aléatoires indépendantes exponentiellement distribuées. Ces systèmes sont faciles à étudier vu la propriété d'absence de mémoire de la loi exponentielle.

Exemple d'une file d'attente $M/M/1$.

Nous considérons un modèle de base le plus utilisé dans la réalité vu son importance dans plusieurs domaines. Le but n'est pas d'étudier de façon exhaustive ce modèle mais plutôt d'analyser ce type de système en régime permanent sous les hypothèses suivantes :

- Le taux d'arrivée est distribué selon la loi de Poisson de paramètre λ .
- Le taux de service est distribué selon la loi exponentielle de paramètre μ .
- La discipline de service est FIFO.
- Le système comprend un seul serveur.

La file est suggérée comme étant un processus de naissance et de mort pour lequel :

$$\lambda_n = \lambda, \quad \forall n \geq 0.$$
$$\mu_n = \begin{cases} \mu, & \forall n \geq 1, \\ 0, & \text{si } n = 0. \end{cases}$$

A-Régime transient

Soit $N(t)$ le nombre de clients présents dans le système à l'instant t grâce aux propriétés fondamentales du processus de Poisson et de la loi exponentielle, $N(t)$ est un processus markovien homogène.

Les probabilités d'état $P_n(t)$ seront calculées par les équations différentielles ci-dessous :

$$\begin{cases} P'_n(t) = -(\lambda + \mu)P_n(t) + \lambda P_{n-1}(t) + \mu P_{n+1}(t), \\ P'_0(t) = \lambda P_0(t) + \mu P_1(t). \end{cases}$$

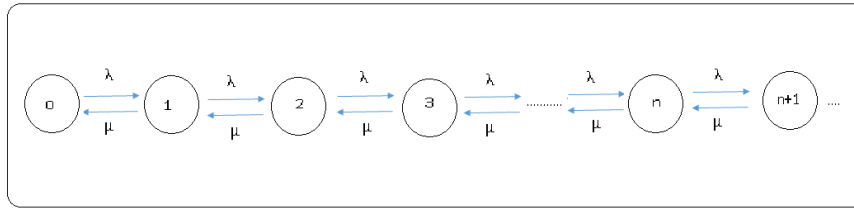


FIGURE 1.2 – Diagramme de transition d'un système de files d'attente $M/M/1$

B-Régime stationnaire

Dans notre cas nous disposons de la file $M/M/1$ dont la description d'états est présentée par $\{N(t)\}_{t \geq 0}$ où $N(t)$ est le nombre de clients présents dans le système à l'instant t . C'est la plus simple description d'états que l'on peut avoir. Notre processus $\{X(t)\}_{t \geq 0}$ à espace d'état discret à temps continu vérifie la propriété d'un processus sans mémoire c'est donc bien une chaîne de Markov à temps continu (CMTC).

Dans ce qui suit nous nous intéressons à l'analyse stationnaire d'une file d'attente stable ($\lambda < \mu$). Soit P_n la probabilité stationnaire à l'état n . Ces probabilités peuvent être calculées de plusieurs façons.

1^{re} méthode : En résolvant le système d'équations linéaire $\mathbf{P}Q = 0$ et $\sum_{n=0}^{\infty} P_n = 1$, où

$\mathbf{P} = [P_0, P_1, \dots]$ est le vecteur des probabilités stationnaires et Q est le générateur infinitésimal de la CMTC :

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & & \\ \mu & -(\lambda + \mu) & \lambda & 0 & \dots & \\ 0 & \mu & -(\lambda + \mu) & \lambda & 0 & \dots \\ \vdots & 0 & \mu & -(\lambda + \mu) & \lambda & 0 \\ & \vdots & 0 & \mu & -(\lambda + \mu) & \lambda \\ & & \vdots & 0 & \mu & \ddots \end{pmatrix}$$

2^{me} méthode : Soient les équations d'états à l'équilibre :

$$\begin{aligned} \lambda P_0 &= \mu P_1, \\ (\lambda + \mu) P_1 &= \lambda P_0 + \mu P_2, \\ &\vdots \\ P_n(\lambda + \mu) &= \lambda P_{n-1} + \mu P_{n+1}, \text{ pour tout } n \geq 1. \end{aligned}$$

En utilisant la méthode récursive, nous trouvons les équations suivantes :

$$\begin{aligned}\lambda P_0 &= \mu P_1, \\ \lambda P_1 &= \mu P_2, \\ &\vdots \\ \lambda P_n &= \mu P_{n-1}, \text{ pour tout } n \geq 0.\end{aligned}$$

3^{me} méthode : Le système étant décomposé en deux systèmes E_1 et E_2 . Nous égalisons les flux entre ces deux systèmes où $E_1 = \{0, 1, \dots, n-1\}$ et $E_2 = \{n, n+1, \dots\}$, nous pouvons écrire l'équation :

$$\lambda P_{n-1} = \mu P_n, \text{ pour tout } n \geq 1.$$

Nous retrouvons le même système d'équations obtenu par les méthodes 1 et 2.

En notant $\rho = \frac{\lambda}{\mu}$, nous avons alors quelle que soit la méthode utilisée :

$$P_n = \rho P_{n-1}, \text{ pour tout } n \geq 1.$$

Ces équations permettent d'exprimer P_n en fonction de P_0 :

$$P_n = \rho^n P_0, \text{ pour tout } n \geq 0.$$

Pour calculer P_0 , nous utilisons la condition de normalisation $\sum_{n=0}^{\infty} P_n = 1$ pour achever les calculs qui nous permet d'obtenir :

$$P_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n} = 1 - \rho.$$

Notons que $\rho < 1$ nous garantit la convergence de la série et qui est cohérent avec la condition de la stabilité. Ainsi il existe certaines relations de base qui permettent de déterminer les mesures de performances désirées grâce à quelques valeurs clés qui sont présentées ci-dessous :

- Nombre moyen de client dans la file :

$$L_q = \sum_{n \geq 1} (n-1) P_n = \frac{\rho^2}{1-\rho}.$$

- Nombre moyen de clients dans le système :

$$L_s = \sum_{n \geq 0} n P_n = \frac{\rho}{1-\rho}.$$

- Temps moyen d'attente dans la file :

$$T_a = \frac{\lambda}{\mu(\mu - \lambda)}.$$

- Temps moyen passé dans le système :

$$T_s = \frac{1}{\mu - \lambda}.$$

- Probabilité qu'il ait zéro client dans le système :

$$P_0 = 1 - \frac{\lambda}{\mu}.$$

- Probabilité qu'il ait n clients dans le système :

$$P_n = P_0 \left(\frac{\lambda}{\mu} \right)^n.$$

- Probabilité qu'il ait moins de n unités dans le système :

$$P_{<n} = 1 - \left(\frac{\lambda}{\mu} \right)^n.$$

1.2.6.2 Modèle de files d'attente non Markovien

Les files d'attente sont supposées non markoviennes si le temps des interarrivées ou la durée de services ne suit pas la loi exponentielle. Ce facteur rend l'étude mathématique du modèle très délicate. Dans ce qui suit nous présentons des méthodes analytiques (voir Abbas (2003)) pour éliminer l'information non markovienne :

1. Méthode des étapes d'Erlang :

Cette méthode est proposée par Erlang et généralisée par Cox. Son principe est d'approximer toute loi de probabilité ayant une transformée de Laplace rationnelle par une loi de Cox qui possède la propriété d'absence de mémoire.

2. Méthode de la chaîne de Markov induite :

Cette méthode élaborée par Kendall a pour principe de choisir une suite d'instantants t_1, t_2, \dots, t_k d'une façon déterministe ou aléatoire telle que le processus induit $\{X_n, n \geq 0\}$ avec $X_n = X_{t_n}$ est une chaîne de Markov homogène qui prédit l'évolution ultérieure du système.

3. Méthode d'approximation :

Nous caractérisons l'état du système étudié par :

- Des méthodes asymptotiques décrivant l'état du système.
- L'estimation par bornes de certaines de ces caractéristiques.

4. Simulation :

La simulation est une technique de modélisation. Elle permet de présenter le

fonctionnement d'un système composé de différents centres d'activité, de mettre en évidence les caractéristiques et de décrire la circulation des différents objets traités par ces processus et enfin observer le comportement du système.

5. Méthode des évènements fictifs :

Le principe de cette méthode est d'introduire des évènements fictifs qui permettent de donner une interprétation probabiliste aux transformées de Laplace et aux variables aléatoires décrivant le système étudié.

La file d'attente $M/G/1$ a été extensivement étudiée par plusieurs chercheurs pour ses applications dans plusieurs domaines (voir par exemple, Levy et Yechiali (1975), Scholl et Kleinrock (1983), Madan et Seneta (1990), Li et Zhu (1996), Choudhury (2005, 2006) et Taha (2007)).

Exemple d'une file d'attente $M/G/1$

Nous considérons un système de files d'attente à capacité illimitée avec un seul serveur étudié dans la thèse de Baruah (2015). Le processus d'arrivée est poissonien dont la loi est exponentielle de taux λ , le temps de service d'un client est distribué selon une variable aléatoire générale X qui n'est plus exponentielle, la distribution de service X est caractérisée par sa fonction de densité de probabilité $f_X(t)$, X est également caractérisée par la donnée de tous ses moments par :

$$E[X^k] = \int_0^{\infty} t^k f_X(t) dt \quad \text{pour tout } k \geq 1.$$

Une autre façon de caractériser parfaitement X est de donner la fonction de densité B^* du temps de service définie par :

$$B^*(s) = \int_0^{\infty} e^{-st} f_X(t) dt.$$

Pour l'étude d'état du système, nous considérons les probabilités stationnaires suivantes :

$$P_n(x) = \lim_{t \rightarrow \infty} P_n(x, t),$$

$$P_n = \lim_{t \rightarrow \infty} P_n(t) = \lim_{t \rightarrow \infty} \int_0^{\infty} P_n(x, t) dx,$$

Les équations d'équilibre sont :

$$P_n(x + \Delta x) = P_n(x)[1 - (\lambda + \mu(x))\Delta x] + P_{n-1}(x)\lambda\Delta x, \quad n \geq 1, \quad (1.1)$$

$$P_0(x + \Delta x) = P_0(x)[1 - (\lambda + \mu(x))\Delta x], \quad (1.2)$$

$$Q = \int_0^{\infty} \mu(x)\Delta x P_0(x) dx + Q(1 - \lambda\Delta x). \quad (1.3)$$

Ainsi nous obtenons les équations qui engendrent le système $M/G/1$:

$$\frac{d}{dx}P_n(x) + (\lambda + \mu(x))P_n(x) = \lambda P_{n-1}(x), \quad n \geq 1, \quad (1.4)$$

$$\frac{d}{dx}P_0(x) + (\lambda + \mu(x))P_0(x) = 0, \quad (1.5)$$

$$\lambda Q = \int_0^{\infty} P_0(x)\mu(x)dx. \quad (1.6)$$

Les équations définies si-dessus sont résolues par l'utilisation des conditions d'équilibre et la condition de normalisation suivantes :

$$P_n(0) = \int_0^{\infty} P_n(x)\mu(x)dx, \quad n \geq 1, \quad (1.7)$$

$$P_0(0) = \int_0^{\infty} P_1(x)\mu(x)dx + \lambda Q, \quad (1.8)$$

$$Q + \sum_{n=0}^{\infty} \int_0^{\infty} P_n(x)dx = 1. \quad (1.9)$$

La fonction génératrice est définie par

$$P_q(x, z) = \sum_{n=0}^{\infty} z^n P_n(x), \quad (1.10)$$

$$P_q(z) = \sum_{n=0}^{\infty} z^n P_n.$$

En multipliant les équations (1.4) par z , sommant sur $n = 1, 2, \dots$, additionnant avec l'équation (1.5) et en utilisant la fonction génératrice définie par l'équation (1.10), nous obtenons

$$\frac{d}{dx}P_q(x, z) + (\lambda - \lambda z + \mu(x))P_q(x, z) = 0. \quad (1.11)$$

Similairement, pour les équations (1.7) et (1.8) nous obtenons

$$zP_q(0, z) = \int_0^{\infty} P_q(x, z)\mu(x)dx + \lambda(z-1)Q. \quad (1.12)$$

En résolvant les équations (1.11) et (1.12), nous obtenons

$$P_q(x, z) = P_q(0, z)e^{-(\lambda - \lambda z)x - \int_0^x \mu(t)dt}. \quad (1.13)$$

En intégrant l'équation (1.13) par partie, nous obtenons

$$P_q(z) = P_q(0, z) \left[\frac{1 - B^*(\lambda - \lambda z)}{\lambda - \lambda z} \right], \quad (1.14)$$

où $B^*(\lambda - \lambda z)$ est la transformé de Laplace-Stieltjes de la distribution du temps de service.

Maintenant, en multipliant l'équation (1.13) par $\mu(x)dx$ et intégrant sur x , nous obtenons

$$\int_0^{\infty} P_q(x, z) \mu(x) dx = P_q(0, z) B^*(\lambda - \lambda z). \quad (1.15)$$

En utilisons les équations (1.15) et (1.12), nous obtenons :

$$P_q(0, z) = \frac{\lambda(z-1)Q}{z - B^*(\lambda - \lambda z)}. \quad (1.16)$$

En substituant $P_q(0, z)$ dans l'équation (1.14) et en utilisant la condition de normalisation pour obtenir Q , nous obtenons :

$$P_q(z) = \frac{[1 - B^*(\lambda - \lambda z)][1 - \lambda E(s)]}{B^*(\lambda - \lambda z - z)}, \quad (1.17)$$

où $E(s)$ est la moyenne de temps de service.

L'équation (1.17) donne la fonction génératrice du nombre de clients dans la file. En utilisant l'équation (1.17) et la relation $L_q = \frac{d}{dz} P_q(z)|_{z=1}$, nous parvenons à dériver les mesures de performances suivantes :

- Probabilité pour que le système soit vide :

$$P_0 = 1 - \rho = 1 - \lambda E(s),$$

- Taux d'utilisation du système :

$$\rho = 1 - P_0 = \lambda E(s).$$

- Nombre moyen de clients dans la file :

$$L_a = \frac{\lambda^2 E(s^2)}{2(1 - \rho)}.$$

- Nombre moyen de clients dans le système :

$$L_s = \rho + L_q.$$

- Temps moyen d'attente dans la file :

$$T_a = \frac{\lambda E(s^2)}{2(1 - \rho)}.$$

- Temps moyen de service dans le système :

$$T = \frac{\rho}{\lambda} + \frac{\lambda E(s^2)}{2(1 - \rho)}.$$

1.3 Files d'attente avec feedback

Plusieurs situations réelles peuvent être modélisées comme des systèmes de files d'attente avec feedback où le client peut être renvoyé au système pour bénéficier d'un autre service. Dans les télécommunications, les transmissions des données 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. Takacs (1963) a étudié la file $M/M/1$ avec feedback, il a déterminé le processus stationnaire de la longueur de la file et la distribution d'attente des clients dans le système. D'Avignon et Disney (1976) ont analysé une file d'attente à un serveur avec feedback dépendant. Bengtsson (1984) a donné une approximation de la modélisation et du contrôle pour les systèmes de files d'attente avec feedback. Santhakumaran et Thangaraj (2000) ont étudié une file d'attente à un serveur avec feedback et clients impatientes. Choudhury et Paul (2005) ont analysé la file $M/G/1$ avec deux phases de serveurs hétérogène et bernoulli feedback. Thangaraj et Vanitha (2009) ont obtenu la solution transiente de la file $M/M/1$ avec feedback et catastrophes. Salehirad et Badamchizadeh (2009) ont étudié la file $M/G/1$ avec k phases de services hétérogènes et feedback. Ayyappan et al. (2010) ont présenté la file $M/M/1$ avec feedback et clients impatientes sous la discipline de service avec priorité par la méthode de la matrice géométrique. Kumar et sharma (2012) ont donné la solution stationnaire de la file $M/M/K$ avec feedback, abandon et retention. Melikov et al. (2015) ont présenté l'analyse numérique d'un système de file d'attente avec feedback. Kumar et Taneja (2017) ont discuté un modèle de files d'attente avec feedback, et trois serveurs dont l'un est relié de manière centralisée aux deux autres serveurs, les clients peuvent revenir pour le service au plus d'une fois. Des équations ont été dérivées pour trouver la longueur moyenne de la file d'attente en utilisant la technique de la fonction génératrice des probabilités.

Exemple de files d'attente avec feedback

Nous considérons un système à un seul serveur, les arrivées sont poissonniennes et les temps de service exponentielles. Nous supposons qu'un client quittant le serveur s'écarte du système avec la probabilité δ et retourne à la file d'attente pour un service supplémentaire avec la probabilité $\delta' = 1 - \delta$. Nous considérons aussi que le système ne distingue pas entre le retour du client et un client arrivant de l'extérieur demandant un service alors il s'en suit que le processus $N(t)$ est bien un processus de naissance et de mort tel que :

$$\lambda_n = \lambda, \quad n = 0, 1, \dots$$

$$\mu_n = \delta\mu, \quad n = 0, 1, \dots$$

$$P_0 = 1 - \frac{\lambda}{\delta\mu}.$$

Donc, quand $\lambda < \mu$:

$$P_k = \left(1 - \frac{\lambda}{\delta\mu}\right) \left(\frac{\lambda}{\delta\mu}\right)^k .$$

1.4 Files d'attente avec serveurs hétérogènes

Dans la littérature, la plupart des travaux sont réalisés pour les files d'attente multi-serveur avec homogénéité de service. Ceci n'est valable que lorsque le processus est à contrôle mécanique ou électrique. Mais en réalité ce mode n'est pas toujours valide, le service peut être long, court, ou normal. C'est à dire que le serveur ne peut pas fournir un service de même taux pour tous les clients, l'exemple le plus concret si le serveur est humain, donc ce dernier ne peut pas garder le même taux de service tout au long de son service vue la nature humaine, il perd son efficacité de service instantanément si la file est trop longue ainsi nous sommes entrain de parler de files d'attente avec serveurs hétérogènes. Dans la vie réelle, les modèles des systèmes de files d'attente avec différentes intensités de service sont utilisés pour l'étude des processus de télécommunication, en informatique, en industrie,...,etc.

L'hétérogénéité est un mécanisme de serveurs qui fournit aux clients une qualité de service différente. Le service hétérogène est clairement une caractéristique principale de fonctionnement de presque tous les systèmes de fabrication. Le rôle de la qualité et la performance de service sont des aspects cruciaux pour la clientèle, par exemple dans les entreprises, une attention particulière doit être accordée aux clients lors de la conception et de la mise en œuvre de leurs opérations. Pour ces raisons, les files d'attente avec serveurs hétérogènes ont reçu une attention considérable. Morse (1958) est le premier qui a introduit la notion d'hétérogénéité dans le service et a obtenu des résultats en régime permanent pour son modèle de files d'attente. Saaty (1961) a examiné le problème de Morse et a obtenu les expressions explicites des probabilités en régime stationnaire et du nombre moyen dans le système. Ancker et Cafarian (1963) ont étudié l'état stationnaire de la file $M/M/S/N$ avec abandon et serveur hétérogènes. Ils ont dérivé les mesures de performance de ce système et une comparaison était faite entre système avec serveurs hétérogènes et système avec serveurs homogènes équivalent. Heffer (1969) a analysé la distribution du temps d'attente des files d'attente avec serveurs hétérogènes. Singh (1970) a étudié un système de files d'attente Markovien avec balking et deux serveurs hétérogènes. Dans Neuts et Takahashi (1981), le comportement asymptotique de la distribution stationnaire de la file d'attente $GI/PH/c$ a été examiné. Sharma et Dass (1988) ont analysé les distributions stationnaire pour le système de files d'attente $M/M/2/N$ à serveurs hétérogènes.

Ensuite, Rykov (2001) a étudié le contrôle monotone d'un système de files d'attente avec serveurs hétérogènes . Dans Al Seedy (2004), l'auteur proposait la solution transiente pour la file $M/M/2$ avec dérobage. Dans Kumar et Madheswari (2005), les auteurs ont présenté la solution transiente de la file $M/M/2$ avec catastrophe (retard ou interruption) de service. Kumar et Madheswari (2005) ont analysé un système de

files d'attente $M/M/2$ avec vacances multiple en utilisant la méthode de la matrice géométrique. Dans Kumar et al. (2007), les auteurs ont dérivé la solution transiente pour un système de files d'attente avec deux serveurs hétérogènes avec catastrophe on définissant une suite de fonction génératrice. Artalejo et Gomez-corrall (2008) ont analysé le système de files d'attente $M/G/1$ avec rappel et serveurs en vacance.

Rykov et Efrosinin (2009) ont étudié le contrôle optimal de système de files d'attente avec serveurs hétérogènes. Yue (2009) ont examiné plus en détail le modèle de Kumar et Madheswari (2005).

Krishnamoorthy et Sreenivasan (2012) ont présenté un modèle de files d'attente $M/M/2$ à deux serveurs hétérogènes dont l'un des serveurs reste inactif et le second serveur est en vacance lorsque le système est vide. Sridhar et Pitchai (2015) ont analysé le système de files d'attente $M/M/2$ avec serveurs hétérogènes et vacances de serveurs, l'état stationnaire de ce système est étudié par la méthode de la matrice géométrique. Rajan (2017) a considéré un système de files d'attente $M/M/2$ avec catastrophe dont le premier serveur est tout le temps disponible et le deuxième serveur disponible par intermittence. Ce modèle est résolu par la technique de la matrice géométrique. Kalyanaraman et Senthilkumar (2018) ont analysé une file d'attente avec deux serveurs hétérogènes et clients impatientes (reneging). Kalyanaraman et Kalaiselvi (2019) ont présenté l'analyse d'un système de files d'attente à deux serveur hétérogènes dont un des serveurs admet un seuil de service.

1.5 Files d'attente avec clients impatientes

Les systèmes de files d'attente avec clients impatientes (dérobade 'balking' et abandon 'reneging') apparaissent dans de nombreuses situations de la vie réelle, leur application est potentielle dans différents domaines tels que les systèmes de communication, les centres d'appels, les systèmes d'inventaire de production,... etc. De tels phénomènes sont souvent rencontrés dans les systèmes de télécommunications, pour plus de détails, nous donnons une petite description du phénomène inspiré de Bacceli (1984) :

- Dans un réseau de télécommunication, un abonné impatient peut abandonner avant que la connexion qu'il demande soit complètement établie, entraînant une utilisation inefficace des ressources.

- Dans un réseau à commutation de paquets, les nœuds de commutation ont des capacités de buffer limitées. Par conséquent, un client arrivant n'est accepté que si sa taille ajoutée aux tailles des paquets déjà présents dans le nœud est inférieure à la capacité totale. Le débit de sortie étant constant, cela équivaut à une limitation de son temps d'attente.

Les systèmes avec des temps d'attente limités peuvent être classés comme suit :

- La limitation n'agit que sur le temps d'attente ou sur le temps de séjour (attente + service).
- Le client peut calculer son temps d'attente potentiel à partir de son arrivée et dérobe

si cela dépasse sa patience ou il rejoint la file d'attente, mais quitte le système si sa patience expire.

La combinaison de ces deux distinctions donne quatre systèmes de file d'attente avec des clients impatientes :

1. limitation du temps de séjour, clients avertis. Le client entrant quitte immédiatement s'il sait que son temps de séjour total dépasse sa patience.
2. limitation du temps de séjour, clients non avertis. C'est le cas de clients qui n'ont aucune information sur le système ni sur le début du service (par exemple, un abonné en ligne attendant une tonalité).
3. limitation du temps d'attente, clients avertis. Identique au cas 1. L'impatience n'agit que sur le temps d'attente.
4. limitation du temps d'attente, clients non avertis. Identique au deuxième cas. L'impatience agit uniquement sur le temps d'attente.



FIGURE 1.3 – Système de files d'attente avec impatience

1.5.1 Notation de files d'attente avec clients impatientes

Nous pouvons décrire la file d'attente avec clients impatientes avec la nomenclature introduite par Barrer (1957) sous la forme $A/S/s/s + D - X$ où :

- A, S et s représentent respectivement la loi du processus des interarrivés, des temps de services et le nombre de serveurs comme la nomenclature de Kendall.
- D représente la taille de la file.
- X représente la discipline de service.

En comparant les deux nomenclatures de Kendall et Barrer nous constatons que le paramètre D remplace alors le paramètre N (la taille de la file) comme facteur limitant la

capacité du système. La file d'attente avec clients impatientes est donc un modèle avec perte.

1.5.2 Disciplines de service

La dynamique de service d'une file d'attente avec clients impatientes est la même pour une file d'attente classique. Les disciplines de service mentionnées dans la partie 1.2.2 sont bien sur applicables pour une file d'attente avec clients impatientes en prenant en considération les clients qui ne sont plus à servir. Les disciplines fréquentes dans les systèmes de files d'attente avec clients impatientes sont :

- EDF (Earliest Deadline First) non préemptive : à la fin de service le serveur sert un client dont le délai résiduel à cet instant est le plus court.
- EDF préemptive (EDFp) : le serveur sert à chaque instant un client dont le délai résiduel à cet instant est le plus court.
- LDF (Largest Deadline First) non préemptive : le serveur sert à chaque fois qu'il complète un service du client dont le délai résiduel à cet instant est le plus long.
- LDF préemptive notée (LDFp) : le serveur sert à chaque instant le client dont le délai résiduel à cet instant est le plus long.

Les différentes variables aléatoires et processus stochastiques caractérisant une file d'attente avec clients impatientes sont définies par :

- W_n : le temps d'attente associé au client C_n avant d'être servi. Lorsque ce temps est supérieur à D_n (délai d'attente) alors le client C_n est perdu.
- T_N^a : le temps d'attente d'un client C_n dans la file :

$$T_n^a = W_n \mathbb{1}_{(W_n < D_n)} + D_n \mathbb{1}_{(W_n \geq D_n)}.$$

- T_n^s : le temps de séjour du client C_n , c'est le temps que passe effectivement C_n dans le système :

$$T_n^s = (W_n + \sigma_n) \mathbb{1}_{(W_n < D_n)} + D_n \mathbb{1}_{(W_n \geq D_n)}.$$

où σ_n est la durée de service demandée par le client C_n .

- P_k : la probabilité de perte du client C_k :

$$P_k = P[C_k \text{ est perdu}] = P[W_k > D_k].$$

- $(N_t)_{t \geq 0}$: le processus de comptage qui compte le nombre de clients dans le système à chaque instant t

- $(Q_t)_{t \geq 0}$: le processus qui compte le nombre de clients dans la file d'attente à l'instant t :

$$Q_t = (N_t - s)^+.$$

- $(S_t)_{t \geq 0}$: le processus des fins de services comptant le nombre de services terminés à l'instant t :

$$S_t = \sum_{n \in \mathbb{N}^*} 1_{(T_n^- \leq t)},$$

où T_n^- est le $n^{ième}$ instant de fin service .

- $(Y_t)_{t \geq 0}$: le processus qui compte le nombre de pertes des clients jusqu'à l'instant t :

$$Y_t = \sum_{n \in \mathbb{N}^*} 1_{(T_n^{\cdot} \leq t)},$$

où T_n^{\cdot} est le $n^{ième}$ instant où un client est perdu.

- $(D_t)_{t \geq 0}$: le processus de départ du système comptant les instants de sortie du système (service ou perte) jusqu'à l'instant t :

$$D_t = S_t + Y_t.$$

1.6 Comportement des clients impatientes dans un système de files d'attente

L'études de files d'attente avec clients impatientes est similaire à celle de files d'attente ordinaire. La différence entre les deux systèmes se concentre sur le mécanisme du client impatient qui est présenté dans cette section par :

1.6.1 Dérobade (balking)

Le mécanisme balking se produit dans des systèmes comme les centres d'appel où l'on indique généralement au client qui appelle comme bien de temps il doit attendre avant qu'un opérateur ne soit disponible pour répondre à l'appel. Ensuite le client a le choix d'attendre (joindre la file) ou de raccrocher (quitter la file).

Le client impatient est confronté à la décision d'adhérer ou non lorsque le(s) serveur(s) est (sont) actif(s) une fois arrivé au système. Le principal facteur qui résulte de la décision du client de joindre ou non la file d'attente est le temps d'attente dans la file avant d'acquérir le service (Liu et Kulkarni (2006, 2008)). Cependant, le client prend toujours sa décision conformément à la longueur de la file d'attente.

Haight (1957) est le premier qui a travaillé sur les files d'attente avec balking. Il a supposé que le client admet une valeur de seuil N qui est la longueur de la file avant qu'il n'arrive au service, si le client observe que la longueur de la file est inférieure à N alors il rejoint la file d'attente dans le cas contraire il quitte la file. Pour les travaux sur le balking avec un seuil, nous nous référons aux (Yechiali (1971), Krishna Kumar

et al (1993), Ke et Chang (2009) et Bae et Kim (2010)). Beaucoup de recherches ont traité des files d'attente avec balking en assumant que la probabilité de balking dépend du nombre de clients n qui se trouvent dans le système. Ainsi, la probabilité de balking est croissante lorsque le nombre de clients n augmente dans le système (voir par exemple Ancker et Grafian (1963(a)), Rao (1965), Subba Rao (1967), Singh (1970), VanTits et Van der Veeken (1980), Serhan (2006), Yechiali (2007), Lozano et Moreno (2008), et Perel et Yachiali (2010)). Les études ci-dessus supposent que la probabilité de balking ne diminue pas le nombre de clients dans la file d'attente. Cette hypothèse est presque raisonnable en réalité. Cependant, Haviv et Kerner (2007) ont abordé un phénomène inverse dans lequel les clients se dérobent d'une file d'attente vide. Ce type de mécanisme est appliqué aux différents systèmes d'industrie.

Exemple d'un modèle de files d'attente M/M/1 avec dérobage

Pour donner une bonne modélisation mathématique pour ce système de files d'attente avec dérobage (balking) il faut décrire ce modèle comme une file d'attente classique en ajoutant le caractère balking que subit le client impatient et le modéliser mathématiquement. Nous traitons une file d'attente M/M/1 avec dérobage étudié par (Lozano 2008) sous les hypothèses suivantes :

- Les clients arrivent selon un processus de Bernoulli avec une probabilité p , où p est la probabilité de rejoindre la file d'attente et $q = 1 - p$ est la probabilité de ne pas rejoindre la file d'attente.
- Le système possède un seul serveur dont le taux de service est distribué selon une loi exponentielle μ .
- Si le système est vide à l'instant d'arrivée d'un client, le service commence immédiatement et le client quitte le système une fois le service est terminée.
- Si le client arrivant trouve le système occupé, alors il a le choix de rejoindre la file pour être servi avec une probabilité r_k si la capacité du système est égale à k (client persistant), soit il abandonne le système sans service avec une probabilité complémentaire (client non persistant).
- la probabilité de balking est égale à $1 - r_k$ si la file est de longueur $k + 1$.
- Les durées de service sont indépendants et distribuées selon la loi géométrique avec une probabilité $\bar{s} = 1 - s$ où s est la probabilité qu'un client ne termine pas son service .
- Il est supposé que la capacité du système est égale à N .

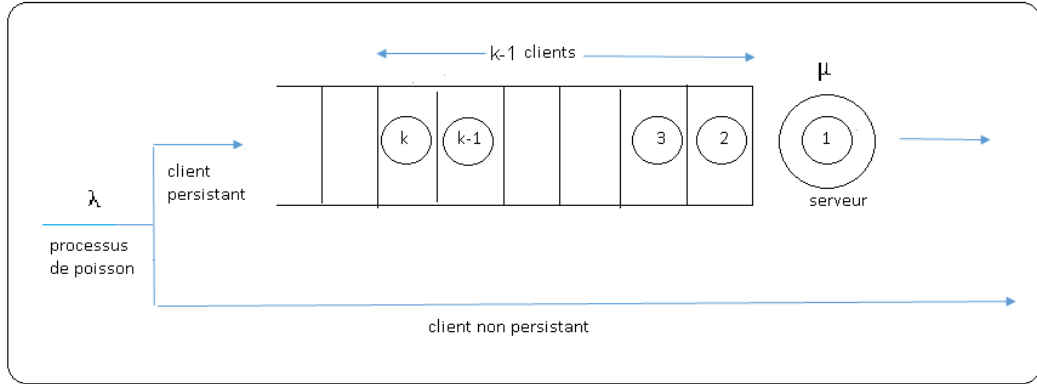


FIGURE 1.4 – Le modèle mathématique d’un système M/M/1 avec dérobade (Lozano et Moreno (2008))

Soit $r_k = r$ la probabilité de balking (une constante). De tels modèles existent lorsque le client n’a pas d’information sur la capacité du système. Il est clair que si la capacité du système augmente, le temps d’attente est plus long et les arrivées sont impatientes. Nous supposons $0 < p < 1$ et $0 < s < 1$. l’intensité du trafic est donnée par $\rho = \frac{p}{s}$. Le système peut être décrit par la variable aléatoire X_m , qui désigne le nombre de clients dans le système.

Si $r_k = r = 1$, les clients rejoignent le système pour recevoir un service. Dans le cas où $r_k = 0$, les clients arrivants ne rejoignent le système que lorsque leur temps d’attente dans la file est nul, lorsque le système est totalement vide (c’est un cas d’impatience extrême).

D’après Lozano et Moreno (2008), la chaîne de Markov $(X_m)_{m \in \mathbb{N}}$ est ergodique si et seulement si $\rho \delta_{r,1} \delta_{N,\infty} < 1$ et sa distribution stationnaire est donnée par la formule :

$$P_0 = \left(1 + \sum_{k=1}^{N+1} \frac{\rho^k s^{k-1} r^{k(k-1)/2}}{\prod_{j=0}^{k-1} (1 - pr^j)} \right)^{-1},$$

et

$$P_k = \frac{\rho^k s^{k-1} r^{k(k-1)/2}}{\prod_{j=0}^{k-1} (1 - pr^j)} \left(1 + \sum_{l=1}^{N+1} \frac{\rho^l s^{l-1} r^{(l-1)l/2}}{\prod_{i=0}^{l-1} (1 - pr^i)} \right)^{-1}, \quad k = 1, \dots, N+1.$$

Il faut souligner l'importance de trouver l'hypothèse d'ergodicité qui est liée au terme "propagation du chaos" (Brokov (1998)). Bien que les systèmes à caractère chaotique obéissent à certaines règles qui peuvent être décrites par des équations mathématiques, la théorie du chaos montre la difficulté de prédire leur comportement à long terme. Ce système de files d'attente pourrait être abordé d'une autre manière en utilisant la théorie du chaos (Brokov (1998), Friedman et al. (1996), Haxholdt (2003)); plus précisément, nous devrions suivre les deux étapes cruciales données par Borovkov (1998). Bien sûr, lorsque les valeurs paramétriques tendent vers la condition d'ergodicité, le système devient instable et tend vers la divergence. De plus, bien que le système soit toujours ergodique pour un système à capacité fini, il est sérieusement perturbé et déséquilibré lorsque les valeurs paramétriques s'approchent de la condition d'ergodicité d'un modèle de files d'attente à capacité infini.

Cet exemple étudie la distribution stationnaire du temps d'attente dans le système sous la discipline de service *FIFO*.

Soit W le temps d'attente dans le système sous la discipline de service *FIFO*. D'après Lozano et Moreno (2008) :

– La distribution stationnaire du temps d'attente T^a dans le système *FCFS* est donnée par :

$$P[T^a = 0] = P_0 + \bar{s}P_1,$$

$$P[T^a = j] = \sum_{k=1}^{\min(j, N+1)} \binom{j-1}{k-1} \bar{s}^k s^{j-k+1} P_k + \sum_{k=2}^{\min(j+1, N+1)} \binom{j-1}{k-2} \bar{s}^k s^{j-k+1} P_k, \quad j \geq 1.$$

– La distribution stationnaire du temps séjour T^s dans le système *FCFS* est donnée par :

$$P[T^s = j] = \bar{s}s^{j-1}(P_0 + \bar{s}P_1) + (1 - \delta_{j,1}) \sum_{k=1}^{\min(j-1, N+1)} \binom{j-1}{k} \bar{s}^{k+1} s^{j-k} P_k$$

$$+ (1 - \delta_{j,1}) \sum_{k=2}^{\min(j, N+1)} \binom{j-1}{k-1} \bar{s}^{k+1} s^{j-k} P_k, \quad j \geq 1.$$

1.6.2 Abandon (reneging)

Dans cette partie, nous nous intéressons à un type particulier d'impatience, à savoir, le reneging. Dans de nombreux cas pratiques les systèmes tels les standard téléphonique, le traitement des patients dans les hôpitaux, les systèmes d'inventaire de stockage de biens ...etc, les clients peuvent devenir impatients et quittent le système sans obtenir le service lorsque le temps d'attente est intolérable (long). En effet, lorsqu'un client considère qu'il a passé trop de temps à attendre pour le service demandé nous disons qu'il s'impatiente.

Deux catégories d'impatience se présentent (voir Salch (2013)) :

◇ **Impatience sans abandon** : Dans ce cas, le client trouve qu'il a patienté trop longtemps par rapport au service qu'il a demandé mais il reste toutefois dans la file d'attente. Cette information peut être obtenue par le biais d'un questionnaire de satisfaction. Ce cas se produit surtout pour des services ou des biens qui font face à une pénurie ou encore à un monopole. Par exemple lorsqu'un client achète une voiture neuve, il n'est pas inhabituel qu'il ait à attendre plusieurs mois avant d'obtenir le modèle qu'il avait demandé. Cette situation décrit le phénomène de pénurie. En effet, les véhicules étant pour la plupart fabriqués à la demande du client pour respecter exactement le panel d'options qu'il a choisi, cela entraîne un long délai de production. Dans ce cas, le client va éventuellement négocier un rabais si les délais de livraison ne sont pas respectés. Il s'impatiente mais va rarement annuler sa commande pour se tourner vers un concurrent, il ne va donc pas abandonner. La situation de monopole se produit par exemple à un péage autoroutier lors des heures de pointe. Les véhicules forment des files d'attente à chaque portique de péage, les automobilistes peuvent s'impatienter mais ils n'ont pas d'autre choix que de passer par ce péage. Là encore, ils peuvent s'impatienter mais n'abandonneront pas.

◇ **Impatience avec abandon** : Dans ce cas, les clients qui s'impatientent quittent immédiatement le système. Cela se produit lorsque la ressource demandée est accessible à un faible coût pour le client ou lorsque le nombre de demandes pour cette ressource varie fortement. Dans ce cas, le client peut avoir intérêt à abandonner sa demande pour la reformuler plus tard. C'est ce qui se produit souvent dans les centres d'appels téléphoniques. Les clients qui s'impatientent raccrochent, car ils n'ont pas été servis assez vite.

Les principales règles de reneging dans beaucoup de recherches sont : Le cas le plus simple lorsque T est une constante fixe (Boots et Tijms (1999), Xiong et Altiok (2009)). Dans d'autres travaux le temps d'attente maximal T est supposé distribuer selon la loi exponentielle de paramètre i où i est le nombre de clients dans le système (Rao (1965), Wang et Chang (2002)). Le cas où le temps d'attente est arbitrairement distribuée se trouve dans (Andreas et Manfred (1999), Ward et Glynn (2005), Zeltyn et Mandelbaum (2005)). De plus, la lenteur du service (Omarah (2006)) et la panne de service (Blackburn (1972), Nasrallah (2009)) peuvent provoquer l'impatience (Perel et Yachiali (2010)). Il y a également d'autres hypothèses particulières concernant le comportement du reneging. Adan et al. (2009) ont examiné le cas où les clients aban-

donnent la file simultanément, par exemple le cas des systèmes à distance où les clients abandonnent le système une fois l'installation de transport est disponible. Il existe encore beaucoup de travaux qui examinent le comportement des clients impatientes dans la file d'attente (Jones (1999), Zohar et al.(2002), Zeltyn et Mendelbaum(2005)).

Les modèles de files d'attente avec clients impatientes ont un grand impact dans de nombreuses situations de la vie réelle. Ces modèles sont naturellement présents dans des domaines tels que l'informatique, les systèmes de communication, systèmes de production, services téléphoniques, systèmes de lignes de production, systèmes d'exploitation de machines, postes,... etc. Ces dernières années, l'étude des systèmes de files d'attente avec des clients impatientes 'balking and renegeing' a suscité un intérêt croissant. Nous pouvons se référer à Shin et Choo (2009), El-Paoumy et Nabwey (2011), Kumar et al. (2014), Kumar et Sharma (2014), Bouchentouf et al. (2014), Baek (2017), et Bouchentouf et Guendouzi (2020).

Dans la littérature sur les files d'attente, les modèles où les clients peuvent être impatientes en raison des vacances des serveurs ont été analysés extensivement. Yue et al. (2006) a présenté l'analyse des performances optimales d'un système de file d'attente $M/M/1/N$ avec balking, renegeing et vacance du serveur. Altman et Yechiali (2006) ont analysé certains modèles de files d'attente tels que $M/M/1$, $M/G/1$ et $M/M/c$ avec vacances des serveurs et des clients impatientes, des cas de vacances uniques et multiples ont été étudiés. Altman et Yechiali (2008) ont étudié la file d'attente avec une infinité de serveurs avec vacances et clients impatientes. Ils ont obtenu la fonction génératrice de probabilité du nombre de clients dans le modèle et les mesures de performance qui en résultent du système. Les systèmes de files d'attente avec vacances et abandons synchronisés ont été réalisés par Adan et al. (2009). Wu et Ke (2010) ont présenté un système avec multiserveurs, peu de fiabilité et clients impatientes. Dans Padmavathy et al. (2011), les auteurs ont étudié le comportement à l'état d'équilibre des files d'attente avec vacance, clients impatientes et waiting server. Nous pouvons élargir notre littérature, on se référant aux recherches plus récentes telles que Yue et al. (2014), Yue et al. (2016) et Ammar (2015, 2017).

Dans la réalité c'est bien rare de trouver des systèmes avec balking uniquement, nous trouvons toujours ce mécanisme accompagné avec le renegeing (voir Haghghi et al.(1986), Rényi (1973), Kumar et Arivudainambi (2002), , El-Paoumy (2008), Ma et al. (2013)).

Exemple d'un système de files d'attente avec abandon

Nous considérons un modèle de file d'attente $M/M/m/m + c$ avec abandon étudié par Takagi (2014) dont l'objectif était de présenter l'analyse stochastique de ce système et de dériver quelques mesures de performances qui seront cités dans la suite.

Ce système doit être modélisé en décrivant ses caractéristiques suivantes :

- Le processus d'arrivée est poissonien de taux λ .
- Le taux de service suit une distribution exponentielle de moyenne $\frac{1}{\mu}$.
- Les clients sont servis selon la discipline *FIFO*
- La capacité du système est limité à c
- Le temps d'impatience du client suit une distribution exponentielle de moyenne $\frac{1}{\xi}$.

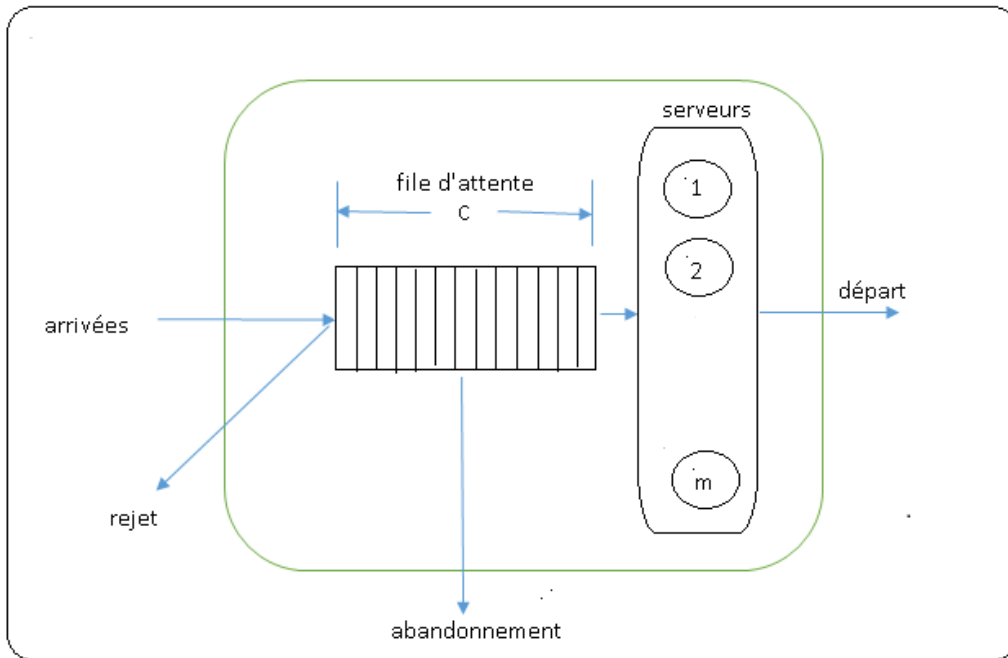


FIGURE 1.5 – Système de file d’attente M/M/m/m+c avec clients impatientes

Soit $N(t)$ le nombre de clients présents dans le système à l’instant t à espace d’état $S = \{0, 1, 2, \dots, m + c\}$.

Le processus est considéré comme étant un processus de naissance et de mort dans le système pour le quel :

$$\lambda_k = \lambda, \quad 0 \leq k \leq m + c - 1,$$

$$\mu_k = \begin{cases} k\mu, & 1 \leq k \leq m - 1, \\ m\mu + (k - m)\xi, & m \leq k \leq m + c. \end{cases}$$

La distribution stationnaire est définie par :

$$P_k = \lim_{t \rightarrow +\infty} P\{N(t) = k\}, \quad 0 \leq k \leq m + c,$$

dont les équations d’équilibre à l’état stationnaire sont :

$$-\lambda P_0 + \mu P_1 = 0,$$

$$\begin{aligned}
& -(\lambda + k\mu)P_k + (k+1)\mu P_{k+1} + \lambda P_{k-1} = 0, \quad 1 \leq k \leq m-1, \\
& -[\lambda + m\mu + (k-m)\xi]P_k + \lambda P_{k-1} + [m\mu + (k-m+1)\xi]P_{k+1} = 0, \quad m \leq k \leq m+c-1, \\
& -(m\mu + c\xi)P_{m+c} + \lambda P_{m+c-1} = 0.
\end{aligned}$$

La condition de normalisation est comme suit $\sum_{k=0}^{m+c} P_k = 1$. En résolvant les équations écrites ci-dessus, nous obtenons pour $1 \leq k \leq m$:

$$P_k = P_0 \frac{\rho^k}{k!}, \quad \text{où } \rho = \frac{\lambda}{\mu} \text{ est l'intensité de trafic.}$$

Pour $1 \leq k \leq c$, nous obtenons :

$$P_{m+k} = P_0 \frac{\rho^m}{m!} \frac{(\rho/m)^k}{\prod_{j=1}^k (1 + j\theta/m)}, \quad \text{où } \theta = \frac{\xi}{\mu}.$$

En utilisant la condition de normalisation, nous obtenons

$$\frac{1}{P_0} = \sum_{k=0}^{m-1} \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{k=0}^c \frac{(\rho/m)^k}{\prod_{j=0}^k (1 + j\theta/m)}.$$

Les mesures de performances du système sont :

- La probabilité de rejet $P\{B\}$.

$$P\{B\} = P_{m+c} = P_0 \frac{\rho^m}{m!} \frac{(\rho/m)^c}{\prod_{j=1}^c (1 + j\theta/m)}.$$

- La probabilité d'abandon P_{ren} .

$$P_{ren} = \frac{1}{\lambda} \sum_{k=1}^c (k\theta) P_{m+k} = \frac{\theta}{\lambda} P_0 \frac{\rho^m}{m!} \sum_{k=1}^c \frac{k(\rho/m)^k}{\prod_{j=1}^c (1 + j\theta/m)}.$$

- Nombre moyen de clients dans la file L_q .

$$P\{L_q = 0\} = P_0 \sum_{k=0}^m \frac{\rho^k}{k!},$$

$$P\{L_q = k\} = P_{m+k} = P_0 \frac{\rho^m}{m!} \frac{(\rho/m)^k}{\prod_{j=1}^k (1 + j\theta/m)}, \quad 1 \leq k \leq c,$$

où L_q est le nombre de clients dans la file à l'instant t ainsi le nombre moyen des clients dans la file est donné par

$$L_q = P_0 \frac{\rho^m}{m!} \sum_{k=1}^c \frac{k(\rho/m)^k}{\prod_{j=1}^k (1 + j\theta/m)}.$$

- Nombre moyen de clients servis $\mathbb{E}(C.S)$.

$$\mathbb{E}(C.s) = \rho \sum_{k=0}^{m-1} P_k + m \sum_{k=1}^c P_{m+k}.$$

- Taux moyen de service λ' .

$$\lambda' = \mu \mathbb{E}(C.S.)$$

1.7 Contribution de la thèse

La contribution de cette thèse consiste principalement à étudier certains systèmes de files d'attente avec Bernoulli feedback, serveurs hétérogènes, et clients impatientes.

- **Le premier résultat : Heterogeneous two server queueing system with reverse balking and reneging.**

Dans ce travail, nous avons établi l'analyse d'un système de files d'attente $M/M/2$ à capacité finie N à deux serveurs hétérogènes avec feedback, muni du mécanisme de clients impatientes (reverse-balking, balking, reneging). Les clients arrivent dans le système un par un sous forme d'un processus de poisson de paramètre λ . Ils sont servis auprès des serveurs 1 et 2 selon le principe premier arrivée premier servi (FIFO) avec un taux de service μ_1 et μ_2 respectivement où les temps de services sont considérés indépendants. Lorsque le système est vide ou l'un des serveurs est libre, le client peut se dérober et ne rejoint pas la file avec une probabilité θ' et peut rejoindre la file d'attente avec une probabilité $\theta = 1 - \theta'$. Lorsque le nombre de clients dans le système est $n \geq 2$ le client entre dans le système avec une probabilité $\frac{1}{[(n-2)+1]}$. Les clients s'impatientent en raison de la longueur de la file ou bien la lenteur du service, ainsi chaque client active un délai d'impatience T de distribution exponentielle de paramètre ξ , dès son arrivé au système. Si le serveur se libère avant l'expiration du temps d'impatience T , le client reste dans la file d'attente et si le temps d'impatience T s'écoule et le serveur est en occupation alors le client abandonne le système sans recevoir un service avec une probabilité γ . Via un certain mécanisme, il peut être maintenu dans le système avec

une probabilité $\gamma' = 1 - \gamma$. Si le client n'est pas satisfait de son service, il peut rejoindre la file d'attente avec une probabilité $\delta' = 1 - \delta$ ou quitter définitivement le système avec une probabilité δ .

Théorème 1.7.1. *Les équations d'équilibre d'un système de file d'attente à serveurs hétérogènes, Bernoulli feedback, reverse balking, abandon et retention des clients impatientes sont :*

$$\begin{aligned} \lambda\theta P_{00} &= \delta\mu_1 P_{10} + \delta\mu_2 P_{01}, \quad n = 0, \\ \begin{cases} (\lambda\theta + \delta\mu_1)P_{10} = \delta\mu_2 P_{11} + \lambda\theta\pi_1 P_{00}, \\ (\lambda\theta + \delta\mu_2)P_{01} = \delta\mu_1 P_{11} + \lambda\theta\pi_2 P_{00}, \end{cases} & \quad n=1, \\ (\lambda + \delta(\mu_1 + \mu_2))P_2 &= (\delta(\mu_1 + \mu_2) + \xi\gamma)P_3 + \lambda P_1, \quad n = 2, \\ \left(\frac{\lambda}{(n-2)+1} \right) + (\delta(\mu_1 + \mu_2) + (n-2)\xi\gamma) & P_n = (\delta(\mu_1 + \mu_2) + ((n+1) - 2)\xi\gamma) P_{n+1} \\ & + \left(\frac{\lambda}{(n-3)+1} \right) \frac{dP_{n-1}(t)}{dt}, \quad 3 \leq n \leq N-1, \\ \left(\frac{\lambda}{(N-2)+1} \right) P_{N-1} &= (\delta(\mu_1 + \mu_2) + (N-2)\xi\gamma) P_N. \end{aligned}$$

Les probabilités d'état stationnaire sont données par :

$$\begin{aligned} P_{10} &= \frac{\lambda\theta}{\delta\mu_1} \left(\frac{\lambda\theta + \delta(\mu_1 + \mu_2)\pi_1}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) P_{00}, \\ P_{01} &= \frac{\lambda\theta}{\delta\mu_2} \left(\frac{\lambda\theta + \delta(\mu_1 + \mu_2)\pi_2}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) P_{00}, \\ P_1 &= \left(\frac{\lambda(\mu_1 + \mu_2)}{\delta\mu_1\mu_2} \left(\frac{\lambda\theta + \delta\mu_1\pi_2 + \delta\mu_2\pi_1}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) \right) \theta P_{00}, \end{aligned}$$

et

$$P_n = \left(\frac{1}{((n-2)+1)!} \prod_{k=2}^n \frac{\lambda}{\delta(\mu_1 + \mu_2) + (k-2)\xi\gamma} \left(\frac{\lambda(\mu_1 + \mu_2)}{\delta\mu_1\mu_2} \left(\frac{\lambda\theta + \delta\mu_1\pi_2 + \delta\mu_2\pi_1}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) \right) \right) \theta P_{00}, \quad 2 \leq n \leq N,$$

avec

$$P_{00} = \left\{ 1 + \left\{ \left(\frac{\lambda(\mu_1 + \mu_2)}{\delta\mu_1\mu_2} \left(\frac{\lambda\theta + \delta\mu_1\pi_2 + \delta\mu_2\pi_1}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) \right) \theta \left(1 + \sum_{n=2}^N \left(\frac{1}{((n-2)+1)!} \prod_{k=2}^n \frac{\lambda}{\delta(\mu_1 + \mu_2) + (k-2)\xi\gamma} \right) \right) \right\} \right\}^{-1}.$$

Les mesures de performances associées à notre modèle de file d'attente sont :

– Le nombre moyen de clients dans le système, L_s :

$$L_s = \sum_{n=1}^N nP_n = P_1 + \sum_{n=2}^N nP_n.$$

– Le nombre moyen de clients servis, $\mathbb{E}(C, S)$:

$$\mathbb{E}(C, S) = \sum_{n=1}^N \delta(\mu_1 + \mu_2) nP_n.$$

– Le taux moyen de reverse-balking, $R_{rev-balk}$:

$$R_{rev-balk} = \theta' \lambda P_{00} + \theta' \lambda P_1 + \sum_{n=2}^N \lambda \left(1 - \frac{1}{(n-2)+1} \right) P_n.$$

– Le taux moyen d'abandon, R_{ren} :

$$R_{ren} = \sum_{n=2}^N (n-2) \xi \gamma P_n.$$

– Le taux moyen de retention, R_{ret} :

$$R_{ret} = \sum_{n=2}^N (n-2) \xi (1 - \gamma) P_n.$$

Afin d'analyser l'impact de la probabilité de retention et la probabilité de reverse-balking sur le système, une analyse économique a été effectuée.

- Le coût total prévu par unité de temps Γ par unité du temps est donné par :

$$\Gamma = (\mu_1 + \mu_2)(C_{service} + \delta' C_{service-feedback}) + C_{holding} L_s + C_{lost} \lambda P_N + C_{ren} R_{ren} + C_{ret} R_{ret} + C_{rev-balk} R_{rev-balk} + 2C_{fixed}.$$

- Le total des revenus attendus dans le système Δ :

$$\Delta = R(\mu_1 + \mu_2)(1 - P_{00}).$$

où $R(\mu_1 + \mu_2)(1 - P_{00})$ est le taux de revenu pour la contribution au service des clients dans le système.

-Le profit total Θ :

$$\Theta = \Delta - \Gamma.$$

-Le deuxième résultat : Queueing system with renegeing and no waiting line

Ce travail traite un système de file d'attente à deux serveurs hétérogènes sans ligne d'attente (sans buffer), avec Bernoulli feedback, abandon, et retention des clients abandonnés sans file d'attente.

Les clients arrivent à la station de service un par un suivant un processus de poisson de paramètre λ , il y'a deux serveurs hétérogènes qui fournissent un service à tous les clients adhérant le système selon leur ordre d'arrivée (FIFO) avec un taux de service indépendant de loi exponentielle de paramère μ_i avec $i = 1, 2$. Après avoir reçu son service, le client peut rejoindre la file d'attente avec une probabilité δ' ou quitter le système avec une probabilité $\delta = 1 - \beta'$, il n'y a pas de distinction entre les arrivées et le retour des clients déjà servis. Les clients dans le système peuvent devenir impatients lorsque le service est si long et peuvent quitter la file d'attente avant la fin de service. Dans ce cas, au début de service chaque client s'attribue un délai d'attente. Le

temps d'impatience suit une distribution exponentielle de paramètre ν_1, ν_2 pour les serveurs 1 et 2 respectivement. C'est suivant ce temps d'impatience que le client décide soit d'abandonner la file d'attente avec une probabilité γ soit de rester avec une probabilité $\gamma' = 1 - \gamma$.

Théorème 1.7.2. *Si nous avons un système d'attente à deux serveurs hétérogènes, sans buffer avec Bernoulli feedback, abandon, et retention des clients abandonnés, alors*

1. les équations d'équilibre sont :

$$\begin{aligned}\lambda P_{0,0} &= (\delta\mu_1 + \gamma\nu_1)P_{1,0} + (\delta\mu_2 + \gamma\nu_2)P_{0,1}, \\ (\lambda + \delta\mu_1 + \gamma\nu_1)P_{1,0} &= (\delta\mu_2 + \gamma\nu_2)P_{1,1} + \lambda\pi_1 P_{0,0}, \\ (\lambda + \delta\mu_2 + \gamma\nu_2)P_{0,1} &= (\delta\mu_1 + \gamma\nu_1)P_{1,1} + \lambda\pi_2 P_{0,0}, \\ \lambda P_1 &= (\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))P_2.\end{aligned}$$

2. Les probabilités d'état stationnaire P_n sont données par :

$$P_{1,0} = \left\{ \frac{\lambda + (\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))\pi_1}{2\lambda + \beta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)} \frac{\lambda}{\delta\mu_1 + \gamma\nu_1} \right\} P_{0,0},$$

$$P_{0,1} = \left\{ \frac{\lambda + (\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))\pi_2}{2\lambda + \beta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)} \frac{\lambda}{\delta\mu_2 + \gamma\nu_2} \right\} P_{0,0},$$

$$P_1 = \frac{\lambda(\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))}{(\delta\mu_1 + \gamma\nu_1)(\delta\mu_2 + \gamma\nu_2)} \left\{ \frac{\lambda + (\delta\mu_1 + \gamma\nu_1)\pi_2 + (\delta\mu_2 + \gamma\nu_2)\pi_1}{2\lambda + \delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)} \right\} P_{0,0},$$

et

$$P_2 = \frac{\lambda^2}{(\delta\mu_1 + \gamma\nu_1)(\delta\mu_2 + \gamma\nu_2)} \left\{ \frac{\lambda + (\delta\mu_1 + \gamma\nu_1)\pi_2 + (\delta\mu_2 + \gamma\nu_2)\pi_1}{2\lambda + \delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)} \right\} P_{0,0},$$

avec

$$P_{0,0} = \left(1 + \left\{ \frac{(\lambda + (\delta\mu_1 + \gamma\nu_1)\pi_2 + (\delta\mu_2 + \gamma\nu_2)\pi_1)(\lambda^2 + \lambda(\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)))}{(2\lambda + \delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))((\delta\mu_1 + \gamma\nu_1)(\delta\mu_2 + \gamma\nu_2))} \right\} \right)^{-1}.$$

Les expressions des différentes mesures de performance sont dérivées.

– Le nombre moyen de clients dans le système L_s .

$$L_s = P_1 + 2P_2.$$

– Le taux d'arrivée effectif dans le système λ' :

$$\lambda' = \lambda(1 - P_2).$$

– Le nombre de clients servis $E(C.S)$:

$$\mathbb{E}(C.S) = \delta\mu_1 P_{1,0} + \delta\mu_2 P_{0,1} + 2\delta(\mu_1 + \mu_2)P_2.$$

– Le taux moyen d’abandon R_{ren} :

$$R_{ren} = \gamma \nu_1 P_{1,0} + \gamma \nu_2 P_{0,1} + 2\gamma(\nu_1 + \nu_2)P_2.$$

– Le taux moyen de retention R_{ret} :

$$R_{ret} = (1 - \gamma)\nu_1 P_{1,0} + (1 - \gamma)\nu_2 P_{0,1} + 2(1 - \gamma)(\nu_1 + \nu_2)P_2.$$

– Le taux d’abandon R_a :

$$R_a = \lambda - \delta\mu_1 P_{1,0} - \delta\mu_2 P_{0,1} - 2\delta(\mu_1 + \mu_2)P_2.$$

– La proportion de clients perdus P_{CLost} :

$$P_{CLost} = \text{Le taux moyen d’abandon}/\lambda'.$$

1.8 Présentation du manuscrit

Notre thèse est focalisée sur l’analyse des systèmes de files d’attente avec serveurs hétérogènes, Bernoulli feedback, et clients impatientes. Les chapitres suivants représentent la manière dont la recherche a abouti. Notre manuscrit est composé de trois chapitres :

Dans le premier chapitre (chapitre introductif), nous avons présenté un travail de synthèse sur les files d’attente classiques, en donnant les notions essentielles liées aux systèmes d’attente, dans ce chapitre nous avons abordé les types et les méthode analytiques pour les files d’attente classique, nous nous sommes intéressés aussi aux différents mécanismes des clients impatientes (dérobade et abandon), l’hétérogénéité de service, et le feedback. Dans ce chapitre nous avons présenté une bonne partie de littérature.

Dans le deuxième chapitre, nous établissons l’analyse d’un système de files d’attente à deux serveurs hétérogènes à capacité finie avec Bernoulli feedback, reverse-balking, et abandon. A la fin du service, le client peut quitter le système comme il peut réintégrer la file d’attente pour recevoir un autre service avec une certaine probabilité. Les clients impatientes peuvent être retenus dans le système avec une certaine probabilité en utilisant certain mécanisme. L’analyse stationnaire du système est établie. Les mesures de performances du système sont dérivées et une analyse économique était établie pour ce système. Ce travail a été publié dans **OPSEARCH 55, volume 251-267, 2018.**

Le troisième chapitre est consacré à étude d’un système de files d’attente $M/M/2$ sans buffer (no waiting line), deux serveurs hétérogènes dont la distribution de service est exponentielle de taux μ_i , $i = 1, 2$, Benoulli feedback, abandon, et retention des clients impatientes. Le flux d’arrivé est poissonien, les clients se dirigent vers les serveurs avec une certaine probabilité. Le client dans le système active un délai d’attente sur lequel il décide de recevoir un service ou de quitter le système. Une fois le client est servi,

il peut revenir au système pour un autre service (Bernoulli feedback). Les probabilités d'état stationnaire sont obtenues. Différents mesures de performances du système sont dérivées ce résultat a fait l'objet d'une publication internationale dans **Probst Forum, 11, volume 67-76, 2018.**

Bibliographie

- Adan .I., Economou .A., Kapodistria .S., Synchronized reneging in queueing systems with vacations, *Queueing Systems*, 62, 1-33, (2009).
- Aissani .A., Kartashov .N.V., Ergodicity and stability of markov chains with respect to operator topology in the space of transition kernels, *Sciences U.S.S.R, serie A* 11, 3-5, (1983).
- Al-Seedy .R.O., A transient solution of the non-truncated queue M/M/2 with balking and an additional server for longer queues (Krishnamoorthi discipline), *Applied Mathematics and Computation* , 3, 763-769, (2004).
- Altman .E., Yechiali .U., Analysis of customers impatience in queues with server vacation, *Queueing Systems*, 52(4), 261-279, (2006).
- Altman .E., Yechiali .U., Infinite server queues with systems additional tasks and impatient customers, *Probability in the Engineering and Informational Sciences*, 22(4), 477-493, (2008).
- Ammar .S.I., Transient analysis of an M/M/1 queue with impatient behavior and multiple vacations, *Applied Mathematics and Computation*, 260, 97-105, (2015).
- Ammar .S.I., Transient solution of an M/M/1 vacation queue with a waiting server and impatient customers, *Journal of the Egyptian Mathematical Society*, 25, 337-342, (2017).
- Ancker.C.J. and Gafarian.A.V., Queueing with reneging and multiple heterogeneous servers, *Nav. Res. Logist. Q.*, 10, 125-145, (1963).
- Ancker Jr.C. J. and Gafarian. A. V., Some queueing problems with balking and reneging. *I, Operations Research* , 11, 88- 100, (1963 (a)).
- Andreas, Manfred, On the M(n)/M(n)/s queue with impatient calls, *Performance Evaluation*, 35, 1-18, (1999).
- Artalejo.J. R., Gomez-Corral.A., Advances in retrial queues, *European Journal of Operation Research*, 189(3), 1042-1233, (2008).
- Ayyapan .G., Muthu Ganapathi Subramanian .A., Sekar .G., M/M/1 Retrial queueing system with loss and feedback under non-pre-emptive priority service by matrix geometric method, *Applied Mathematical Sciences*, 48, 2379-2389, (2010).
- Baccelli .F., Boyer .P., Hebutte .G., Single-server queues with impatient customers, *Advances in Applied Probability*, 16, 887-905, (1984).
- Bae .J., Kim .S., Lee .E.Y., The virtual waiting time of the M/G/1 queue with impatient customers, *Queueing Systems*, 38, 485-494, (2001).
- Baek .J., Dudina .O., Kim .C., A queueing system with heterogeneous impatient customers and consumable additional items, *Int. J. Appl. Math.Comput. Sci.*, 27 (2), 367-384, (2017).
- Barrer .D.Y., Queueing with impatient customers and ordered service, *Operations Res*, 5, 650-656, (1957).
- Baruah .M., Analysis of some batch arrival queueing systems with balking, reneging, random breakdowns fluctuating modes of service and bernoulli scheduled server vacations thesis,

(2015).

Belyaev. Y.K., Linear markov process and their application to reliability problems, Probability Theory And Mathematical Statistics, 309-323, (1962).

Bengtsson .B., On some control problems for queues, Linkoping Studies In Science And Technology, Dissertation Department of Electrical Engineering, Linkoping University, Sweden, 225, (1984).

Blackburn .J.D., Optimal control of a single-server queue with balking and renegeing, Management Science, 19, 297-313, (1972).

Boots .N., Tijms .H., An M/M/c queue with impatient customers, TOP, 7, 213-220, (1999).

Borovkov .A., A Stochastic Processes in queueing theory, Springer-Verlag, Berlin, (1976).

Borovkov .K.A., Propagation of chaos for queueing networks, Theory of Probability and its Applications, 42, 385-394, (1998) .

Bouchentouf .A.A, Guendouzi .A.H., Cost Optimization dor an $M^x/M/c$ vacation queueing system with waiting servers and impatient customers, SeMA Journal, 76 (2), 309-341, (2019).

Bouchentouf .A.A, Guendouzi .A.H., The $M^x/M/c$ bernoulli feedback queue with variant multiple working vacations and impatient customers performance and economic analysis, Arabian Journal Of Mathematics, 9 (2), 309-327, (2020).

Bouchentouf .A.A., Kadi .K., Rabhi .A., Analysis of two heterogeneous server queueing model with balking, renegeing and feedback, Math.Sci. Appl. E Notes, 2 (2), 10-21, (2014).

Bronstein .O.I., Dukhovny .I., Priority queueing models in computer systems, Nauka, Moscow, (1976) .

Brown .L, Gans .N., Mandelbaum .A., Sakov .A., Shen.H., Zeltyn.S. and Zhao.L., Statistical analysis of a telephone call center : A queueing-science perspective, J. Amer. Statist. Assoc., 100, 36-50, (2005).

Choi .B.D., Chang .Y., Single server retrial queues with priority calls, Mathematical and computer modelling, 30, 7-32, (1999).

Choudhury .G., Paul .M. , A two phase queueing system with Bernoulli feedback, International Journal of Information and Management Sciences, 16(1), 35-52,(2005).

Courtois .P.J., Decomposability queueing and computer system applications, Academic Press, (1977).

Cox .D.R., Some statistical methods connected with series of events, Journal of The Royal Statistical Society : series B (methodological), 2517-6161, (1955).

Crommelin .C.D., Delay probability formulae when the holding times are constant, PostOffice Electrical Engineers Journal, 25, 41-50, (1932).

Crommelin .C.D., Delay probability formulae, Post Office Electrical Engineers Journal, 26, 266-274, (1934).

D'Avignon .G.R., Disney .R.L., Single Server Queue with State dependent feedback, INFOR, 14,71-85, (1973).

Disney .R.L., A matrix solution for the two server queue with overflow, Mngement science, 254-265, (1972).

Disney .R.L., Farrell .R.L., De Morais .P.R., A Characterization of $M/G/1$ queues with renewal departure processes, Mananagement science, 11, 1213-1338, (1973) .

El-Paoumy .M.S., On poisson arrival queue : $M^x/ M /2/N$ with balking, renegeing and heterogeneous servers, Applied Mathematical Sciences, 2(24), 1169-1175, (2008).

El-Paoumy .M. S., Nabwey .H. A., The poissonian queue with balking function, renegeing and two heterogeneous servers, International Journal of Basic and Applied Sciences., 11(6),

149-152, (2011).

Engset. T., Emploi du calcul des probabilités pour la détermination du nombre des selecteurs dans les bureaux téléphoniques centraux, Rev. Gen. D'élect, 9, 40-138, (1918).

Erlang .A. K., The theory of probabilities and telephone conversations. Nyt Tidsskrift for Matematik B, 20, 33, (1909).

Erlang .A. K., Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges, Elektrotekniker, 13, 5, (1917).

Feng .W., Kowada .M., Adachi .K., Analysis of multi-server queue with two priority classes and (M,N) threshold service schedule : non-preemptive priority, 7, 653-671, (2000).

Franken .D., Konig .D., Arndt .U., Schmidt .V., Queues and point processes, Akademie- verlag, Berlin, (1981).

Friedman .E.J., Landsberg .A.S., Long-run dynamics of queues : stability and chaos, Operations Research Letters , 18, 185-191, (1996).

Fry .T.C., Probability and its engineering uses, Van nostrand reinhold, 15, 175-177, (1928).

Gao.J.Z., Wang.X., Analysis of a retrial queue with two type breakdowns and delayed repairs, Ieee Access, 428-442, (2020).

Gnedenko .B.V., Kovalenko .I.N., Introduction to Queueing Theory, Birkhuser, (1989).

Gong .W.B., Hu .J.K. , The Maclaurin series of the $GI/G/1$ queues, Journal of Applied Probability, 29 , 184-276, (1992).

Haghighi .A., Medhi .J., Mohanty .S.G., On a multiserver markovian queueing system with balking and reneging, Com. Oper. Res, 13, 421-425, (1986).

Haight .F.A., Queueing with balking, I, Biometrika, 44, 360-369, (1957).

Haight .F.A., Queueing with reneging, Metrika, 2, 186-197, (1959).

Haviv .M., Kerner .Y., On balking from an empty queue, Queueing Systems, 55, 239-249, (2007).

Haxholdt .C., Larsen .E.R., Van Ackere .A., Mode locking and chaos in a deterministic queueing model with feedback, Management Science, 49, 816-830 , (2003).

Heffer .J.C., Steady state solution of the $M/E_k/c$ (FIFO) queueing system, CORSJ, 7, 16-30, (1969).

Johansen .S.G., Stidham .S.Jr., Control of arrivals to a stochastic input-output system, Adv. Appl. Probab, 12, 972-999, (1980) .

Jones .L.K., Inferring balking behavior from transactional data, Operations Research, 47, 778-784, (1999).

Kalyanaraman .R., Kalaiselvi .S., Heterogeneous server queue with a threshold on slow server, International Journal of Recent Technology and Engineering (IJRTE), 7, 2277-3878, (2019).

Kalyanaraman .R., Senthilkumar .R., Heterogeneous server Markovian queue with restricted admissibility and with reneging, Mathematical Sciences International Research Journal, 7(1), 309-315, (2018).

Kartachov .N.V., Strongly stable markov chains, Journal Of Soviet Mathematics, 34(2), 1493-1498, (1996).

Ke .J.C., Chang .F. M., Modified vacation policy for $M/G/1$ retrial queue with balking and feedback, Computers and Industrial Engineering, 57, 433-443, (2009).

Kendall .D.G., Some problems in the theory of queues, Journal of the Royal Statistical Society-Series B, 13(2), 151-185, (1951).

Khintchine .A.Y, Mathematical Methods in the theory of Queueing, Chapter I, Hafner, New York, (1960).

Kim .C.S., Finite-Source queueing models for Analysis of complex communication Systems, Journal of the Society of Korea Industrial and System Engineering , 26, 62-67, (2003) .

Kingman .J.F.C., The effect of queue discipline on waiting time variance, Proceeding of the Cambridge Philosophical Society, 63, 163-164, (1961).

Kleinrock .L., Queueing Systems, John Wiley, New York, 417, (1975).

Klimov .V.P., Stochastic service systems, Nauka, Moscow, (1966).

Krishna Kumar.B., Parthasarathy. P.R., Sharafali.M. ,Transient solution of an $M/M/1$ queue with balking, Queueing Systems ,13, 441-448, (1993).

Krishnamoorthy .A. , Sreenivasan .C., An $M/M/2$ queueing system with heterogeneous servers including one with working vacation, International Journal of Stochastic Analysis 1-16, (2012).

Kumar .B.K., Arivudainambi .D., Vijaykumar .A., An $M/G/1/1$ queue with unreliable server and no waiting capacity, International Journal of Information and Management Science, 13(2), 35-50, (2001).

Kumar .B.K., Madheswari .S.P., An $M/M/2$ queueing system with heterogeneous server and multiple vacations, Mathematical and Computer Modelling, 41, 1415-1429, (2005).

Kumar.B.K., Madheswari.S.P., Venkatakrishnan.K.S., Transient solution of an $M/M/2$ queue with heterogeneous servers subject to catastrophes, Information and Management Sciences, 63-80, (2007).

Kumar .R., Sharma .S.K., A Markovian feedback queue with retention of renege customers and balking, AMO-Advanced Modeling and Optimization, 14(3), (2012).

Kumar.R., Sharma.S.K., A Single-server markovian queueing system with discouraged arrivals and retention of renege customers, Yugoslav Journal of Operations Research, 24,119-126, (2014).

Kumar .S., Taneja .G., A feedback queueing model with chances of providing service at most twice by one or more out of three Servers, Aryabhata Journal of Mathematics and Informatics, 9 (1), 171-184, (2017).

Legros .B., Jouini .O., Routing in a queueing system with two heterogeneous servers in speed and in quality of resolution, Stochastic models, 1532-4214 , (2017).

Li .H.,Zhu .Y., Analysis of $M/G/1$ queue with delayed vacation and exhaustive service discipline, European Journal of Operation Research, 92(1), 125-134, (1996).

Lindley .D.V., The theory of queues with a single server, Proc. Camb.Phil, Soc, 48, 277-289, (1952).

Liu .L., Kulkarni .V., Explicit solutions for the steady state distributions in $M/PH/1$ queues with workload dependent balking, Queueing Systems, 52, 251-260, (2006).

Liu .L., Kulkarni.V., Busy period analysis for $M/PH/1$ queues with workload dependent balking, Queueing Systems, 59, 37-51, (2008).

Little .J.D., A proof of the queueing formula $L = \lambda W$, Oper. Res, 9, 383-387, (1961).

Lozano .M., Moreno .P., A discrete time single-server queue with balking : economic applications, Applied Economics, 40, 735-748, (2008).

Ma .Y., Liu .W.Q., , Li .J.H., Equilibrium balking behavior in the $Geo/Geo/1$ queueing system with multiple vacations, Applied Mathematical Modelling, 37, 3861-3878, (2013).

Marshall. K.T., Wolff. R.W., Customer average and time average queue lengths and waiting times, Journal Of Applied Probability, 8(3), 535-542, (1971) .

Melikov .A.Z., Ponomarenko .L.A., Kuliyeva .Kh.N., Numerical analysis of a queueing system with feedback, Cybernetics and Systems Analysis, 51, 4, (2015).

Morse .P.M., Queues inventories and maintenance, Operations Research Society Of America, 359-520, (1958).

Nasrallah .W.F., How pre-emptive priority affects completion rate in an M/M/1 queue with poisson renegeing, European Journal of Operational Research, 193, 317-320, (2009).

Neuts .M.F., Takahashi .Y., Asymptotic behavior of the stationary distribution in the GI/PH/c queue with heterogeneous servers, 441-452, 57 (1981).

O'Dell.G.F., Gibson.W.W., Automatic trunking in theory and practice. Inst.P.O.Elect.Engrs,41, 107, (1926).

Omarah .A.S.R., A transient solution of queues with variable channel considering balking concept when $S = 2$ and $[\sigma] = 1$, Applied Mathematics and Computation, 174, 337- 344, (2006).

Padmavathy .R., Kalidass .K., Ramanath .K., Vacation queues with impatient customers and a waiting server, International Journal of Latest Trends in Engineering and Technology, 1(1), 10-19, (2011).

Palm.C., Etude des delais d'attente, Ericsson Technics, 5 (2), 39-56, (1936).

Palm.C., Analysis of the Erlang traffic formula for busy signal arrangements, Ericsson Technics, 6, 39-58, (1938) .

Perel .N. , Yechiali .U., Queues with slow servers and impatient customers, European Journal of Operational Research, 201, 247-258, (2010).

Pollaczek.F., Theorie des wartens, 19, 71-78, (1930).

Pollaczek.F., Sur quelques lois asymptotiques de la théorie de l'encombrement des réseaux téléphoniques, Ann.Univ.Lyon, 5, 21-35, (1942).

Pollaczek .F., Sur une theorie unifiée des problemes stochastiques soulevés par l'encombrement d'un faisceau parfait de lignes telphoniques, 254, 3965-3967, (1962).

Rajan .V., Queuing analysis of markovian queue Having two heterogeneous servers with catastrophes using matrix geometric technique, International Journal of Statistics and Systems, 12, 205-212, (2017).

Renyi .A. , Probability theory (in Hungarian), Tankonyvkiado, Budapest, (1973).

Rykov .V., Monotone control of queueing systems with heterogeneous servers, Queueing Sys, 37(4), 391-403, (2001).

Rykov .V.V., Efrosinin .D.V., On the slow server problem, Automa. Remote Cont, 70(12), 2013-2023, (2009).

Saaty .T.L., Elements of queueing theory with applications, McGraw-Hill, (1961).

Salehirad .M.R., Badamchizadeh .A., On the multi-phase M/G/1 queueing system with random feedback, Springer-Verlag, 17, 131-139, (2009).

Salch .A., Thèse Ordonnancement stochastique avec impatience , (2013).

Santhakumaran .A., Thangaraj .V., A Single server queue with impatient and feedback customers, Information and Management Science, 11, 71-79, (2000).

Serhan Ziya .H.A., Foley .R.D, Pekoz .E., A monotonicity result for a G/GI/c queue with balking or renegeing, J. Appl. Probab., 43, 1201-1205, (2006).

Sharma .D. P., Dass .S., Multiserver markovian queues with finite waiting space, Sankhya, B-50, 328-331, (1998).

Shin .Y.W., Choo .T.S., M/M/s queue with impatient customers and retrials, Applied Mathematical Modelling, 33, 2596-2606, (2009).

Singh .V.P., Two-server markovian queues with balking : heterogeneous vs. homogeneous servers, Operations Research, 18, 145-159, (1970).

Smith .W.L., On the distribution of queueing times, *proc.Soc*, 49, 449-461, (1953).

Smith .W.E., Various optimizers for single-stage production, *Naval Research Logistics Quarterly*, 3(1-2), 59-66, (1956).

Sridhar .A., Allah Pitcha .R., Analyses of a Markovian queue with two heterogeneous servers and working vacation, *International Journal of Applied Operational Research*, 5, 1-15, (2015).

Stoyan .D., Comparison methods of queue and other stochastic models, Wileys, new york, (1983).

Subba Rao .S., Queueing with balking and reneing in $M/G/1$ systems, *Metrika*, 12, 173-188, (1967) .

Taha .H.A., *Operation research : an Introduction*, 8th edn, New Jersey, Pearson Prentice Hall, (2007).

Takac's .I., Investigation of waiting time problems by reduction to markov process, *Acta Math. Acd.Scie.Hungar.*, 101-129, (1955).

Takacs .L., A Single server queue with feedback, *The Bell System Tech. Journal*, 42, 134-149, (1963).

Takagi .H., Waiting time in the $M/M/m/(m+c)$ queue with impatient customers, *International Journal of Pure and Applied Mathematics*, 90, 519-559, (2014).

Thangaraj .V., Vanitha .S., On the Analysis of $M/M/1$ feedback queue with catastrophes using Continued fractions, *International Journal of Pure and Applied Mathematics*, 53, 131-151, (2009).

Van Dantzig .G.B., Wolfe .P., Decomposition principle for linear programs, *Operations Research*, 8(1), 101-111, (1960).

Van Tits .M. H. L., Van der Veecken .H. J. M., Simulation of a queueing problem with balking, *Simulation*, 35, 88-93, (1980).

Vaulot .E., Application du calcul des probabilités a l'exploitation telephonique, *Annales des Postes, Télégraphes et Téléphones* , 14, 136-156,(1925).

Wang .K.H., Chang .Y.C., Cost analysis of a finite $M/M/R$ queueing system with balking, reneing and server breakdowns, *Mathematical Methods of Operations Research*, 56, 169,(2002).

Ward .A.R., Glynn .P.W., A diffusion approximation for a $GI/GI/1$ queue with balking or reneing, *Queueing Systems*, 50, 371-400, (2005).

Wilkinson .R.I., The interconnection of telephone systems, *The Bell System Technical Journal*, 10, 531-564, (1931).

Wu .C.H, Ke .J.C., Computational algorithm and parameter optimization for a multi server system with unreliable servers and impatient customers, *Journal of Computational and Applied Mathematics*, 235, 547-562, (2010).

Xiong .W., Altiok .T., An approximation for multi-server queues with deterministic reneing times, *Annals of Operations Research*, 172, 143-151, (2009).

Yechiali .U., On optimal balking rules and toll charges in the $GI/M/1$ queueing process, *Operations Research*, 19, 349-370, (1971).

Yechiali .U., Queues with system disasters and impatient customers when system is down, *Queueing Systems*, 56, 195-202, (2007).

Yue .D., Yue .W., A heterogeneous two-server network system with balking and a Bernoulli vacation schedule, in : *Proceedings of the 4th International Conference on Queueing Theory and Network Applications*, Article No. 20, (2009).

Yue .D., Zhang .Y., Yue .W., Optimal performance analysis of an $M/M/1/N$ queue system with

balking, reneging and server vacation, *International Journal of Pure and Applied Mathematics*, 28 (1), 1001-115, (2006).

Yue .D., Yue .W., Saffer .Z., Chen .X., Analysis of an M/M/1 queueing system with impatient customers and a variant of multiple vacation policy, *Journal of Industrial and Management Optimization*, 10 (1), 89-112, (2014).

Yue .D., Yue .W., Zhao .G., Analysis of an M/M/1 queue with vacations and impatience timers which depend on the server's states, *Journal of Industrial and Management Optimization*, 12 (2), 653-666, (2016).

Zeltyn .S., Mandelbaum .A., Call Centers with Impatient Customers : Many-Server Asymptotics of the $M/M/n + G$ Queue, *Queueing Systems*, 51, 361-402, (2005).

Zohar .E., Mandelbaum .A., Shimkin .N., Adaptive behavior of impatient customers in tele-queues : theory and empirical support, *Management Science*, 48, 566-583, (2002).

Chapitre 2

Heterogenous two server queueing system with reverse balking and renegeing

Ce chapitre fut l'objet d'une publication internationale dans le journal OPSEARCH
2018, volume 55(2), pages 251-267

Heterogenous two server queueing system with reverse balking and renegeing

Amina Angelika BOUCHENTOUF , Aicha MESSABIHI

Depertement of Mathematics, University of Sidi Bel Abbes, B. P.89, Sidi Bel Abbes, 22000, Algerie.

E-mail : bouchentouf-amina@yahoo.fr

Departement of Mathematics, University Center of Ain Temouchent, Ain Temouchent, Algerie

E-mail : amessabihi@yahoo.fr

Abstract. This paper presents an analysis of heterogeneous two-server queueing system with feedback, reverse balking, renegeing and retention of renegeed customers. The steady-state solution of the model is obtained and different measures of effectiveness are derived. Economic analysis of the model is performed studying the impact of balking probability and retention probability on the total expected profit of the system.

Keywords. Queueing systems, Heterogeneity, Markovian model ,Reneging ,Bernoulli feedback and Reverse balking

Mathematics Subject Classification (2000). 60K25. 68M20. 90B22

2.1 Introduction

Queueing with customer impatience has special significance for the business world as it has negative impact on the performance of any business. A customer is said to be impatient if he tends to join the queue only when a short wait is expected and tends to remain in the line if his wait has been sufficiently small. Impatience generally takes three forms. The first is balking, deciding not to join the queue at all up on arrival; the second is renegeing, the reluctance to remain in the waiting line after joining and waiting, and the third is jockeying between lines when each of a number of parallel service channels has its own queue Gross and Harris (1985).

In several realistic systems, like telephone switchboard customers and hospital emergency rooms' handling of critical patients, customers may become impatient and leave (balk; customer might decide to not enter the system when the queue length is too long or renege; a waiting customer in the queue might hang up and leave the queue if the waiting time is too long) the system without getting services.

Queueing systems with balking, renegeing, or both have attracted much attention from numerous researchers. For related literature, interested readers may refer to Haight (1957,1959), Ancker and Gafarian (1963), Abou-EI-Ata and Hariri (1992), Wang and Chang (2002), and references therein.

Nevertheless, in the case of the business connected with investment, the number of customers with a peculiar firm becomes interesting and an attracting factor for investing customers. Therefore, the probability to join such a firm is high. In this case more

the system size is large more customers enter the system. Thus the probability of balking will be small when the system size is large and vice-versa, which is balking in reverse sense "called as reverse balking". Till now, the study of queueing models with reverse balking is limited, Jain and al. (2014) studied an $M/M/1/N$ queueing system with reverse balking, Kumar and Som (2014) developed and incorporated the concept of reverse balking and reverse reneging into the $M/M/1/N$ queueing system. Then, recently Som (2016) considered a multi-server Markovian queueing system with reverse balking and impatient customers.

Heterogeneity of service is a common feature of many real multi-server queueing situations. The heterogeneous service mechanisms are invaluable scheduling methods that allow customers to receive different quality of service. Heterogeneous service is clearly a main feature of the operation of almost any manufacturing system. The role of quality and service performance is crucial aspects in customer perceptions and firms must dedicate special attention to them when designing and implementing their operations. For this reason, queues with heterogeneous servers have received considerable attention in the literature Kumar et al.(2007).

Analytical results for queueing systems with more than two heterogeneous servers are difficult, that's was pointed out in Neuts and Takahashi (1981) Therefore, many researchers have focused on queueing systems with two heterogeneous servers. A control model for a machine center with two heterogeneous servers has been introduced by Liu and Kumar (1984). Kumar and Madheswari (2002) presented the transient solution of the $M/M/2$ queue with catastrophe at the service station. Madan and al.(2003) investigated an $M/M/2$ queue with Bernoulli schedules and a single vacation policy where the two servers provide heterogeneous exponential service to customers. They obtained the steady-state probability generating functions of the system size for various states of the servers. In Kumar and Madheswari (2005), authors analyzed an $M/M/2$ queueing system with two heterogeneous servers and multiple vacations by using the matrix-geometric solution method. Kumar et al. (2007) derived the transient solution for the probabilities in the two-server queueing system subject to catastrophes, where one server is faster than the other. Further, Yue at al. (2009) considered the model in Kumar and Madheswari (2005). They obtained the explicit expression of the rate matrix and presented the conditional stochastic decomposition results for the queue length and the waiting time. In Yue and Yue (2010) , authors discussed an $M/M/2$ queueing system with two heterogeneous servers under a variant vacation policy, where the two servers may take together at most J vacations when the system is empty. Krishnamoorthy and Sreenivasan (2012) displayed an $M/M/2$ queueing model with heterogeneous servers where one server remains idle but the other goes on vacation in the absence of waiting customers.

Queueing systems with two heterogeneous servers and impatience have attracted many researchers. In Abou-El-Ata (1983), authors discussed the two-channels queue $M/M/2$ with both balking, heterogeneity and considering different probabilities in choosing the server. Yue and Yue (2009) introduced an $M/M/2$ queueing system with balking and two heterogeneous servers under Bernoulli schedules and a single vacation policy.

They presented a generalization of Model B in Madan and al.(2003)obtained the explicit expressions of the steady-state condition, the stationary distribution of the system size, and the mean system size. Mahdy El-Paoumy and Hossam Nabwey (2011) studied the $M/M/2/N$ queue with general balk function, renegeing and two heterogeneous servers. And recently, in Ammar (2014) transient analysis of a two-heterogeneous servers queue with impatient behavior is given.

Feedback represents the case where after getting partial or incomplete service, customer retries for service. In computer communication, the transmission of protocol data unit is sometimes repeated due to occurrence of an error. This usually happens because of non-satisfactory quality of service. Rework in industrial operations is also an example of a queue with feedback Kumar and Sharma (2012) Ayyapan et al.(2010) studied $M/M/1$ retrial queueing system with loss and feedback under non preemptive priority service by matrix geometric method. Sharma and Kumar (2012) considered a single server, finite capacity Markovian feedback queue with renegeing, balking and retention of renegeed customers in which the inter-arrival and service times follow exponential distribution. Then, recently Bouchentouf et al.(2014)established an analysis of two heterogeneous server queueing model with balking, renegeing and feedback.

The two heterogeneous servers are extensively studied as mentioned above, nevertheless, in the literature there is no work on an $M/M/2$ queue with heterogeneous servers subject to reverse balking, renegeing and retention of renegeed customers. Based on this observation, we have investigated the analysis of heterogeneous two-server queueing system with reverse balking, renegeing and retention of renegeed customers, and we have presented the impact of retention probability of renegeed customers has a positive economic impact.

The rest of the paper is arranged as follows : In section 2, we describe the model and give the main result. Then in section 3, various performance measures are derived, after that in section 4 an economic analysis of the system is studied. Finally, we finish the paper by a small conclusion.

2.2 Model description and main results

Consider $M/M/2$ feedback queue with heterogeneous servers subject to reverse balking, renegeing and retention of renegeed customers, the service times at each server are independently, identically and exponentially distributed with parameters μ_1 and μ_2 .

The customers are served according to the following discipline :

- If the two servers are busy the customers wait in the queue.
- If one server is free , the head customer in the queue goes to it.
- If both servers are free, the head customer of the queue chooses server 1 with probability π_1 and server 2 with probability, π_2 .

After completion of each service, the customer can either leave the system definitively with probability δ or come back to the system and join the end of the queue with probability δ' , where $\delta + \delta' = 1$. The customers both newly arrived and those that are fed

back are served in order in which they join the tail of original queue. The queue discipline is First-Come-First-Served. We do not discriminate between the regular arrival and feedback arrival. The capacity of the system is taken as finite (say, N).

When the system is empty or one of servers is free, customers can balk and do not enter the system with probability θ' , and may enter the system with probability θ ($\theta = 1 - \theta'$). Such kind of balking is referred to as reverse balking.

When $n \geq 2$, the arriving customers may enter the system with probability $1/[(n-2)+1]$, which depends on the number of customers present in the system at that time.

Each customer upon joining the queue will wait a certain length of time for his service to begin. If it does not begin by then, he will lose his impatience and may leave the queue (renege) without getting service with probability γ and may remain in the queue with probability $\gamma' (= 1 - \gamma)$. So, a reneged customer can be retained by employing certain convincing mechanisms to stay in the queue for completion of his service. Thus, this latter can be retained in the queueing system with some probability γ' and may leave the queue without receiving service with probability γ .

The renegeing times T are independently, identically and exponentially distributed with parameter ξ . This effort to retain the renegeing customer in the queue for his service has positive effect on business of any firm. Now, let define the equilibrium probabilities : $P_{0,0} : \mathbb{P}(\text{there is no customer in the system})$. $P_{1,0} : \mathbb{P}(\text{there is one customer being served by server 1})$. $P_{0,1} : \mathbb{P}(\text{there is one customer being served by server 2})$. $P_n : \mathbb{P}(\text{there are } n \text{ customers in the system})$, $n = 2, 3, \dots, N$. Also, $P_0 = P_{0,0}$, $P_1 = P_{1,0} + P_{0,1}$, and $P_2 = P_{1,1}$.

2.2.1 Steady-State Solution

In this section, the steady state probabilities are derived by using the Markov process method.

Let $P(n)$ denote the probability that are n customers in the system in the steady state. So, by applying the Markov process theory, the differential-difference equations of the model are :

$$\lambda\theta \frac{dP_{00}(t)}{dt} = \delta\mu_1 \frac{dP_{10}(t)}{dt} + \delta\mu_2 \frac{dP_{01}(t)}{dt}, \quad n = 0 \quad (2.1)$$

$$\begin{cases} (\lambda\theta + \delta\mu_1) \frac{dP_{10}(t)}{dt} = \delta\mu_2 P_{11} + \lambda\theta\pi_1 \frac{dP_{00}(t)}{dt}, & n=1 \\ (\lambda\theta + \delta\mu_2) \frac{dP_{01}(t)}{dt} = \delta\mu_1 P_{11} + \lambda\theta\pi_2 \frac{dP_{00}(t)}{dt}, \end{cases} \quad (2.2)$$

$$(\lambda + \delta(\mu_1 + \mu_2)) \frac{dP_2(t)}{dt} = (\delta(\mu_1 + \mu_2) + \xi\gamma) \frac{dP_3(t)}{dt} + \lambda \frac{dP_1(t)}{dt}, \quad n = 2 \quad (2.3)$$

$$\left(\left(\frac{\lambda}{(n-2)+1} \right) + (\delta(\mu_1 + \mu_2) + (n-2)\xi\gamma) \right) \frac{dP_n(t)}{dt} = (\delta(\mu_1 + \mu_2) + ((n+1)-2)\xi\gamma) \frac{dP_{n+1}(t)}{dt} + \left(\frac{\lambda}{(n-3)+1} \right) \frac{dP_{n-1}(t)}{dt}, \quad 3 \leq n \leq N-1 \quad (2.4)$$

$$\left(\frac{\lambda}{(N-2)+1} \right) \frac{dP_{N-1}(t)}{dt} = (\delta(\mu_1 + \mu_2) + (N-2)\xi\gamma) \frac{dP_N(t)}{dt}. \quad (2.5)$$

Theorem 2.2.1. *If the steady-state equations of heterogeneous two-server queueing system with feedback, reverse balking, renegeing and retention of renegeed customers are*

$$\lambda\theta P_{00} = \delta\mu_1 P_{10} + \delta\mu_2 P_{01}, \quad n = 0 \quad (2.6)$$

$$\begin{cases} (\lambda\theta + \delta\mu_1)P_{10} = \delta\mu_2 P_{11} + \lambda\theta\pi_1 P_{00}, \\ (\lambda\theta + \delta\mu_2)P_{01} = \delta\mu_1 P_{11} + \lambda\theta\pi_2 P_{00}, \end{cases} \quad n=1 \quad (2.7)$$

$$(\lambda + \delta(\mu_1 + \mu_2))P_2 = (\delta(\mu_1 + \mu_2) + \xi\gamma)P_3 + \lambda P_1, \quad n = 2 \quad (2.8)$$

$$\left(\left(\frac{\lambda}{(n-2)+1} \right) + \delta(\mu_1 + \mu_2) + (n-2)\xi\gamma \right) P_n = (\delta(\mu_1 + \mu_2) + ((n+1) - 2)\xi\gamma) P_{n+1} + \left(\frac{\lambda}{(n-3)+1} \right) P_{n-1}, \quad 3 \leq n \leq N-1 \quad (2.9)$$

$$\left(\frac{\lambda}{(N-2)+1} \right) P_{N-1} = (\delta(\mu_1 + \mu_2) + (N-2)\xi\gamma) P_N. \quad (2.10)$$

Then

$$P_{10} = \frac{\lambda\theta}{\delta\mu_1} \left(\frac{\lambda\theta + \delta(\mu_1 + \mu_2)\pi_1}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) P_{00}, \quad (2.11)$$

$$P_{01} = \frac{\lambda\theta}{\delta\mu_2} \left(\frac{\lambda\theta + \delta(\mu_1 + \mu_2)\pi_2}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) P_{00}, \quad (2.12)$$

$$P_1 = \left(\frac{\lambda(\mu_1 + \mu_2)}{\delta\mu_1\mu_2} \left(\frac{\lambda\theta + \delta\mu_1\pi_2 + \delta\mu_2\pi_1}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) \right) \theta P_{00}, \quad (2.13)$$

$$P_n = \left(\frac{1}{((n-2)+1)!} \prod_{k=2}^n \frac{\lambda}{\delta(\mu_1 + \mu_2) + (k-2)\xi\gamma} \left(\frac{\lambda(\mu_1 + \mu_2)}{\delta\mu_1\mu_2} \left(\frac{\lambda\theta + \delta\mu_1\pi_2 + \delta\mu_2\pi_1}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) \right) \right) \theta P_{00}, \quad 2 \leq n \leq N, \quad (2.14)$$

with

$$P_{00} = \left\{ 1 + \left\{ \left(\frac{\lambda(\mu_1 + \mu_2)}{\delta\mu_1\mu_2} \left(\frac{\lambda\theta + \delta\mu_1\pi_2 + \delta\mu_2\pi_1}{2\lambda\theta + \delta(\mu_1 + \mu_2)} \right) \right) \theta \left(1 + \sum_{n=2}^N \left(\frac{1}{((n-2)+1)!} \prod_{k=2}^n \frac{\lambda}{\delta(\mu_1 + \mu_2) + (k-2)\xi\gamma} \right) \right) \right\} \right\}^{-1}. \quad (2.15)$$

Démonstration. We obtain the steady-state-probabilities by using iterative method. Solving (2.6)-(2.7) we get (2.11), (2.12) and (2.13). Then solving (2.8), (2.9) and (2.10) we get (2.14). Finally, by using the normalization condition we find (2.15). \square

2.3 Performance measures and numerical analysis

This section presents different measures of effectiveness of the queueing model under investigation, moreover a numerical analysis will be given.

Expected system size, L_s

$$L_s = \sum_{n=1}^N nP_n = P_1 + \sum_{n=2}^N nP_n.$$

Expected number of customers served, $\mathbb{E}(C, S)$

$$\mathbb{E}(C, S) = \sum_{n=1}^N \delta(\mu_1 + \mu_2)nP_n.$$

Average reverse balking rate, $R_{rev-balk}$

$$R_{rev-balk} = \theta' \lambda P_{00} + \theta' \lambda P_1 + \sum_{n=2}^N \lambda \left(1 - \frac{1}{(n-2)+1}\right) P_n.$$

Average renegeing rate, R_{ren}

$$R_{ren} = \sum_{n=2}^N (n-2)\xi \gamma P_n.$$

Average retention rate, R_{ret}

$$R_{ret} = \sum_{n=2}^N (n-2)\xi(1-\gamma)P_n.$$

Next, we present some numerical examples to show the impact of different parameters on the characteristics of the system.

2.3.1 Impact of arrival and service rates λ , μ_1 and μ_2 respectively on the characteristics of the system

Let us present the evolution of the system when λ varies from 0.5 to 2.5 with a pitch 0.5, considering two cases, $\mu_1 = 10$, $\mu_2 = 12$ and $\mu_1 = 4$, $\mu_2 = 2$. Put $\xi = 18$, $N = 5$, $\pi_1 = 0.7$, $\pi_2 = 0.3$, $\theta = 0.4$, $\gamma = 0.1$, and $\delta = 0.6$.

The numerical results obtained for these situations are stored in Tables 2.1 and 2.2.

λ	L_s	$\mathbb{E}(C, S)$	$R_{rev-balk}$	R_{ren}	R_{ret}
0.5	0.126083	0.454038	0.294291	0.001179	0.010609
1	0.278354	1.002075	0.562193	0.008628	0.077653
1.5	0.443810	1.597717	0.792572	0.026214	0.235922
2	0.615174	2.214627	0.990189	0.055588	0.500296
2.5	0.788179	2.837446	1.163936	0.097002	0.873018

TABLE 2.1: System characteristics vs. λ when $\mu_1 = 4$ and $\mu_2 = 2$

λ	L_s	$\mathbb{E}(C, S)$	$R_{rev-balk}$	R_{ren}	R_{ret}
0.5	0.32983	0.435388	0.299438	0.000035	0.000319
1	0.068435	0.903335	0.595861	0.000279	0.002513
1.5	0,106057	1.399955	0.886434	0.000925	0.008322
2	0,145625	1.922253	1.169418	0.002147	0.019324
2.5	0,186893	2.466991	1.443306	0.004102	0.036914

TABLE 2.2: System characteristics vs. λ when $\mu_1 = 10$ and $\mu_2 = 12$

Graphical representations (see Figure 2.1) illustrate the details of the results given in Tables 2.1-2.2.

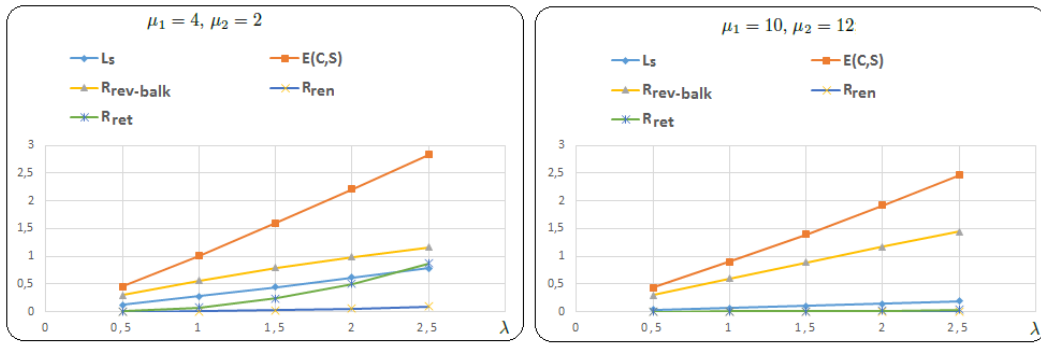


FIGURE 2.1 – System performance measures vs. λ when $\mu_1 = 10, \mu_2 = 12$ and $\mu_1 = 4, \mu_2 = 2$

According to Tables 3.1–3.2 and Figure ?? we constat that :

- Along the increase of λ , the mean number of customers in the system L_s , the mean number of customers served $\mathbb{E}(C, S)$, the average reneing rate R_{ren} , the average retention rate R_{ret} and the average reverse balking rate $R_{rev-balk}$ increase.
- From Table 2.1, it is clearly visible that, with increase in average arrival rate λ , the size of the system increases. An increasing expected system size leads to high confidence of customers with the system, therefore the rate of reverse balking increases. Due to this, more and more arriving customers join the particular system.
- An increasing service rate generates a large number of customers served.
- By comparing the results when $\mu_1 = 4, \mu_2 = 2$ and $\mu_1 = 10, \mu_2 = 12$, we constat that when $\mu_1 = 10, \mu_2 = 12$ (the mean service rates small), the average reneing rate and the average retention rate R_{ret} and R_{ren} respectively are smaller than the case where $\mu_1 = 4, \mu_2 = 2$. On the other side the average reverse balking rate $R_{rev-balk}$ is bigger in the case where $\mu_1 = 10, \mu_2 = 12$. This implies a decreasing in the size of the system L_s , and the number of customers served in the case $\mu_1 = 10, \mu_2 = 12$ is smaller than the first one. This is absolutely agree with our intuition.

2.3.2 Impact of impatience rate ξ on the characteristics of the system

We examine the evolution of the system when ξ varies from 1 to 9 with a pitch equals to 2, considering two cases $\gamma = 0.3$, and $\gamma = 0.9$. Let $\lambda = 0.5$, $N = 5$, $\pi_1 = 0.4$, $\pi_2 = 0.6$, $\delta = 0.6$, $\mu_1 = 8$, $\mu_2 = 5$, and $\theta = 0.6$.

The numerical results obtained for these situations are given in Tables 2.3 and 2.4.

ξ	L_s	$\mathbb{E}(C, S)$	$R_{rever-balk}$	R_{ren}	R_{ret}
1	0.075017	0.682682	0.198788	0.000031	0.000071
3	0.075026	0.682737	0.198862	0.000085	0.000199
5	0.075007	0.682561	0.198851	0.000135	0.000314
7	0.075002	0.682522	0.198878	0.000178	0.000415
9	0.074998	0.682485	0.198901	0.000217	0.000505

TABLE 2.3: System characteristics vs. λ when $\gamma = 0.3$

ξ	L_s	$\mathbb{E}(C, S)$	$R_{rever-balk}$	R_{ren}	R_{ret}
1	0.075014	0.682649	0.198828	0.000086	0.000010
3	0.075005	0.682541	0.198917	0.000217	0.000024
5	0.074988	0.682387	0.198955	0.000312	0.000035
7	0.074979	0.682305	0.198994	0.000385	0.000043
9	0.074971	0.682239	0.199024	0.000442	0.000049

TABLE 2.4: System characteristics vs. λ when $\gamma = 0.9$

Graphical representations (see Figure 2.2) illustrate the details of some results given in Tables 2.3-2.4.

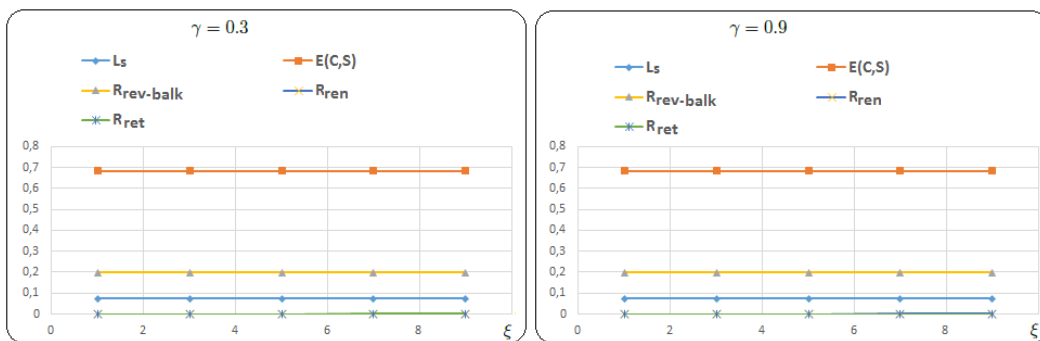


FIGURE 2.2 – System performance measures vs. ξ when $\gamma = 0.3$ and $\gamma = 0.9$

- From Tables 2.3-2.4 and Figure 2.2 we constat that along the increase of the impatience rate ξ , the reneing rate R_{ren} increases, that is, customers become more impatient and consequently they leave the system, this implies that retention rate R_{ret} increases, more customers abandon the system, more the probability of retention becomes high. Moreover, the increase of ξ leads to a decrease in the size of the system L_s and in the mean number of customers served $\mathbb{E}(C, S)$ which is absolutely explicable; more the average rate of impatience becomes small, more customers leave the system without getting their complete service. On the other hand the increase of the impatient rate ξ implies the increase of the rate of reverse balking $R_{rever-balk}$, the rate that the customers do not enter the system increases because the size of the system becomes empty.

- By comparing the results when $\gamma = 0.3$ and $\gamma = 0.9$, we constat that when the probability of reneing is big $\gamma = 0.9$, the size of the system L_s and the mean number of customers served $\mathbb{E}(C, S)$ are smaller than the first case, contrariwise the rate of reneing R_{ren} , and the rate of retention R_{ret} are bigger when $\gamma = 0.9$. Also the rate that the customers do not enter the system $R_{rev-balk}$ is bigger in the second case because more customers are reneged more the system becomes empty and consequently the customers balk.

2.3.3 Impact of balking probability θ' on the characteristics of the system

Let us present the evolution of the system when θ' varies from 0 to 1 with pitch 0.25. Let $\lambda = 4$, $N = 5$, $\pi_1 = 0.6$, $\pi_2 = 0.4$, $\delta = 0.6$, $\mu_1 = 8$, $\mu_2 = 10$, $\gamma = 0.3$, and $\xi = 12$. The numerical results obtained for these situations are arranged in Table 2.5.

θ'	L_s	$\mathbb{E}(C, S)$	$R_{rever-balk}$	R_{ren}	R_{ret}
0	0.665793	7.228796	0.039564	0.074733	0.174378
0.25	0.577965	6.242019	0.866805	0.064875	0.151375
0.5	0.457426	4.940206	1.761986	0.051345	0.119805
0.75	0.281504	3.040247	2.771922	0.031598	0.073729
1	0.000000	0.000000	4.000000	0.000000	0.000000

TABLE 2.5: System performance measures vs. θ'

Graphical representations are given in Figure 2.3.

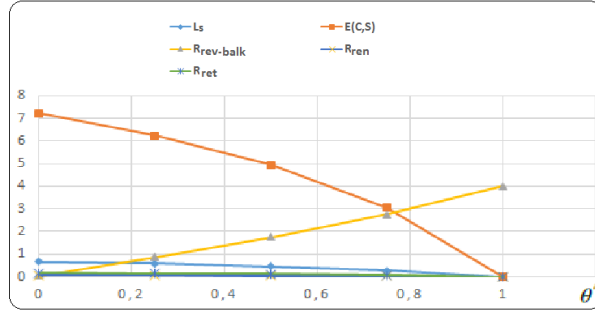


FIGURE 2.3 – System performance measures vs. θ'

- It is visible from Table 2.5 and Figure 2.3 that higher is the probability of balking θ' (when there are zero or one customers in the system) lower are the expected system size L_s , the mean number of customers served $\mathbb{E}(C,S)$, the renegeing rate R_{ren} and the retention rate of renegeed customers R_{ret} . While reverse balking rate $R_{rev-balk}$ increases with increase in probability of balking.

2.3.4 Impact of feedback probability δ' on the characteristics of the system

To study the impact of feedback probability δ' on the evolution of the system, we vary δ' from 0 to 0.8 with pitch 0.2, and take $\lambda = 6$, $N = 5$, $\pi_1 = 0.8$, $\pi_2 = 0.2$, $\theta = 0.3$, $\mu_1 = 12$, $\mu_2 = 9$, $\gamma = 0.4$, and $\xi = 15$.

The numerical results obtained for these situations are stored in Table 2.6.

δ'	L_s	$\mathbb{E}(C,S)$	$R_{rever-balk}$	R_{ren}	R_{ret}
0	0.220373	4.627841	3.973570	0.028298	0.042447
0.2	0.290127	4.874138	3.882742	0.050240	0.075359
0.4	0.412201	5.193734	3.710110	0.100988	0.151482
0.6	0.665690	5.591793	3.381425	0.245392	0.368089
0.8	1.394208	5.576832	2.635374	0.858339	1.287508

TABLE 2.6: System characteristics vs. δ'

Graphical representations are presented in Figure 2.4.

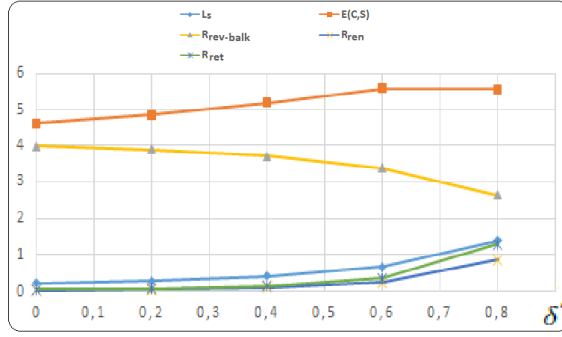


FIGURE 2.4 – System performance measures vs. δ'

- According to Table 2.6 and Figure 2.4, we constat that along the increase of the feedback probability δ' , the size of the system L_s and the number of customers served $\mathbb{E}(C, S)$ increase. Furthermore, the high probability of feedback leads to high rate of renegeing R_{ren} and consequently to a high rate of retaining renegeed customers R_{ret} . Contrariwise with the increase in the feedback probability the rate that the customers do not enter the system $R_{rev-balk}$ decreases, because more the system is full more customers have the desire to enter the system.

2.3.5 Impact of retention of renegeed customers probability γ' on the characteristics of the system

To examine the effect of retention of renegeed customers probability on the behavior of the system, we vary γ' from 0 to 1 with pitch 0.25, and fixe the other parameters. Let $\lambda = 4$, $N = 5$, $\pi_1 = 0.2$, $\pi_2 = 0.8$, $\delta = 0.5$, $\mu_1 = 7$, $\mu_2 = 8$, $\theta = 0.3$, and $\xi = 10$.

γ'	L_s	$\mathbb{E}(C, S)$	$R_{rever-balk}$	R_{ren}	R_{ret}
0	0.462946	3.490239	2.438400	0.137458	0.000000
0.25	0.468447	3.513351	2.433867	0.122273	0.040758
0.5	0.475043	3.562822	2.420011	0.100263	0.100263
0.75	0.486599	3.649495	2.402920	0.065769	0.197306
1	0.510080	3.825598	2.372389	0.000000	0.396339

TABLE 2.7: System characteristics vs. γ'

The results are given in the following Figure.

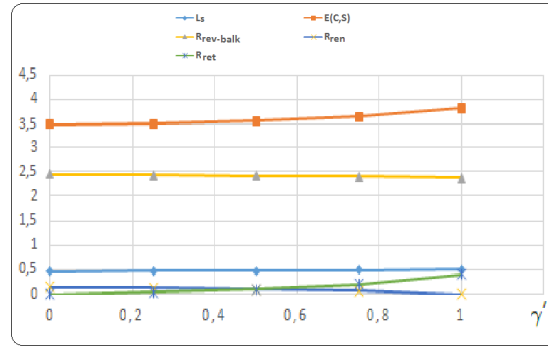


FIGURE 2.5 – System performance measures vs. γ'

- From Table 2.7 and Figure 2.5, we remark that with the increase in the retention probability, the size of the system L_s , the mean number of customers served $\mathbb{E}(C,S)$, the reneing rate R_{ren} and the retention rate R_{ret} increase, on the other hand the rate of reverse balking $R_{rev-balk}$ decreases, more customers are kept in the system, more the size of the system is big, more customers are entering.

2.4 Economic analysis

In this section we perform an economic analysis of two heterogenous server feed-back queueing system with reverse balking, reneing and retention of reneged customers. We analyze the impact of probability of retaining reneged customers and of the reverse balking probability, to this end let's use the following symbols :

$1/\lambda$: mean inter-arrival time

$1/\mu_1$: mean service time of the first server.

$1/\mu_2$: mean service time of the second server.

L_s : expected number of customers in the system.

R_{ren} : average rate of reneing.

R_{ret} : average rate of retention.

$R_{rev-balk}$: average rate of reverse balking.

$C_{service}$: cost of service per unit time.

$C_{holding}$: unit holding cost per unit time.

C_{lost} : cost associated with each lost customer per unit time.

$C_{rever-balk}$: cost associated with each reverse balked customer per unit time.

C_{ren} : cost associated with each reneged customer per unit time.

C_{ret} : cost of retaining a reneged customer per unit time.

$C_{service-feedback}$: cost of serving a feedback customer per unit time.

C_{fixed} : fixed server purchase cost per unit.

R : revenue earned by providing service to a customer.

Γ : total expected cost per unit time of the system.

Δ : total expected revenue per unit time.

* The expected cost Γ per unit time is given by :

$$\Gamma = (\mu_1 + \mu_2)(C_{service} + \delta' C_{service-feedback}) + C_{holding}L_s + C_{lost}\lambda P_N \\ + C_{ren}R_{ren} + C_{ret}R_{ret} + C_{reve-balk}R_{rev-balk} + 2C_{fixed}$$

Let $R(\mu_1 + \mu_2)(1 - P_{00})$ be the earning revenue rate for contributing service to customers in the system, thus

* The total expected revenue Δ of the system is given by :

$$\Delta = R(\mu_1 + \mu_2)(1 - P_{00})$$

* The total expected profit Θ is presented as

$$\Theta = \Delta - \Gamma$$

Next, we present the variation in the different types of costs involved with the change in probability of retaining the renege customers γ' and in the probability of balking θ' . For the whole numerical study we fix the costs at $R = 30$, $C_{holding} = 5$, $C_{lost} = 6$, $C_{ren} = 4$, $C_{ret} = 7$, $C_{rev-balk} = 3$, $C_{fixed} = 2$, $C_{service} = 6$, and $C_{service-feedback} = 8$.

2.4.1 Effect of the probability of retaining the renege customers γ' on the costs of the system

We study the variation in the different types of costs involved with the change in probability of retaining the renege customers γ' . We take $\lambda = 6$, $N = 5$, $\theta = 0.4$, $\xi = 17$, $\delta' = 0.3$, $\mu_1 = 7$, $\mu_2 = 8$, $\pi_1 = 0.4$, $\pi_2 = 0.6$, Let $\Sigma = C_{service} + \delta' C_{service-feedback}$. We study numerically the cost-profit aspects. The following Table presents the numerical results associated with the economic analysis of the model.

γ'	$(\mu_1 + \mu_2) \times \Sigma$	$C_{holding}L_s$	$C_{lost}\lambda P_N$	$C_{ren}R_{ren}$	$C_{ret}R_{ret}$	$C_{rev-balk}R_{rev-balk}$	$2C_{fixed}$	Γ
0	109.2	3.611663	0.001131	1.595059	0.000000	8.559071	4	126.9669241
0.25	109.2	3.744245	0.020884	2.027240	1.1825564	8.490101	4	128.6650264
0.5	109.2	3.720956	0.003715	1.228695	2.1502162	8.496311	4	128.7998937
0.75	109.2	3.838505	0.011263	0.852672	4.4765305	8.436002	4	130.8149728
1	109.2	4.137463	0.073279	0.000000	10.3519386	8.336844	4	136.0995234

TABLE 2.8: Costs involved vs. probability of retaining the renege customers γ'

- From the analysis of Table 2.8, we constat that the service cost per customer and the fixed server purchase cost remain the same, because the number of servers and the service rate as well as the server purchase cost remain the same all over. The holding costs,

the costs due to lost customers increase because of the increase in expected system size due to increase in probability of retaining the renege customers. On the other hand the costs due to reverse balking decreases because of the increase in expected system size. The renege cost per customer decreases as retention leads to the decrease in abandonments due to renege. Along the increase of the probability of retaining the renege customers the cost of retention per customer also increases. Moreover, the total expected cost of the system increases with the increase in probability of retaining the renege customers.

2.4.2 Effect of probability of retaining the renege customers γ' on total expected profit

Let's present the effect of probability of retaining the renege customers γ' on total expected profit. The variations in total expected costs, total expected revenue and total expected profit due to change in probability of retaining the renege customers are presented in Tables 2.9 and 2.10.

γ'	$R(\mu_1 + \mu_2)$	$1 - P_{00}$	Δ
0	390	0.508538	198.329911
0.25	390	0.519894	202.758666
0.5	390	0.518397	202.174898
0.75	390	0.528433	206.088685
1	390	0.549649	214.363121

TABLE 2.9: Total expected revenue Δ vs. γ'

γ'	Δ	Γ	Θ
0	198.329911	126.9669241	71.362987
0.25	202.7586655	128.6650264	74.093639
0.5	202.1748983	128.7998937	73.375005
0.75	206.0886846	130.8149728	75.273712
1	214.3631213	136.0995234	78.263598

TABLE 2.10: Total expected profit Θ vs. γ'

According to Tables 2.9 and 2.10 we can analyze the effect of probability of retaining the renege customers on the total expected profit. It is quite evident from the results given below that the total expected profit increases with the increase in probability of retaining the renege customers. Thus, one can have a snapshot of the costs and profits associated with the system in long period of time. Furthermore the retention of renege customers has a positive impact in the economy.

2.4.3 Effect of the probability of balking θ' on the costs of the system

We analyze the variation in the different types of cost-profit involved with the change in balking probability θ' . We take $\lambda = 3$, $N = 5$, $\delta' = 0.3$, $\mu_1 = 12$, $\mu_2 = 9$, $\pi_1 = 0.3$, $\pi_2 = 0.7$, $\xi = 7$, $\gamma = 0.7$, and we vary θ' . The following Table presents the numerical results associated with the economic analysis of the model.

θ'	Γ	Δ	Θ
0	0.622391	392.106447	332.5715109
0.25	0.552889	348.319902	287.577846
0.5	0.454053	286.053690	224.2103493
0.75	0.295189	185.969307	123.1667248
1	0.000000	0.000000	-63.4

TABLE 2.11: Total costs of the system vs. θ'

- From table 2.11, we constat that total expected cost Γ , total expected revenue Δ , total expected profit Θ of the system decrease with the increase in probability of balking. Thus, it is quite evident that a balked customers have a negative impact in the economy.

2.5 Conclusion and future work

In this paper the concept of feedback, reverse balking, renegeing and retention of renegeed customers is incorporated into two heterogeneous servers queueing system with finite capacity. The steady-state analysis of the model is performed and some important measures of performance are derived. Economic analysis of the model is also analyzed. This model finds its application in investment business facing customer impatience. For further works, the model can be studied in non-Markovian environment and the time-dependent studies of the model can also be performed.

Bibliography

- Abou-El-Ata .M.O., On poisson queues with both balking and heterogeneous servers, Delta Journal of Sciences, 2, 292-303, (1983).
- Abou El-Ata .M.O., Hariri .A.M.A., The $M/M/c/N$ queue with balking and reneging, Computers and Operations Research, 19(13), 713-716, (1992).
- Ancker .Jr., C.J., Gafarian .A.V., Some queuing problems with balking and reneging : I, Operations Research, 11(1), 88-100, (1963).
- Ancker .Jr., C. J., Gafarian .A.V., Some queuing problems with balking and reneging : II, Operations Research, 11(6), 928-937, (1963).
- Ayyapan .G., Muthu Ganapathi Subramanian .A., Sekar .G., $M/M/1$ Retrial queueing system with loss and feedback under non-pre-emptive priority service by matrix geometric method, Applied Mathematical Sciences, 4(48), 2379-2389, (2010).
- Bouchentouf .A.A., Kadi .M., Rabhi .A., Analysis of two heterogeneous server queueing model with balking, reneging and feedback, Mathematical Sciences And Applications E-Notes, 2(2), 10-21, (2014).
- Bouchentouf .A.A., Messabihi .A., Heterogenous two server queueing system with reverse balking and reneging, Operational Research, 55(2), 251-267, (2018).
- D'Avignon .G.R., Disney .R.L., Single server queue with State dependent feedback, INFOR, 14, 71-85, (1976).
- El-Paoumy Mahdy .S., Nabwey Hossam .A., The Poissonian queue with balking function, reneging and two heterogeneous servers, International Journal of Basic & Applied Sciences IJBAS-IJENS, 11(6), 149-152, (2011).
- Gross .D., Harris .C.M., Fundamentals of queuing theory, (Second Edition), Wiley, Hoboken, NJ, (1985).
- Haight .F.A., Queueing with balking, Biometrika, 44, 360-369, (1957).
- Haight .F.A., Queueing with reneging, Metrika, 2, 186-197, (1959).
- Jain .N.K., Kumar .R., Som .B.K., An $M/M/1/N$ queueing system with reverse balking, American Journal of Operational Research, 4(2), 17-20, (2014).
- Kopytko.B., Zhernovyi.K., Stationary characteristics of $M^\theta/M/1$ queue with switching of service mode, Sientific Research Of The Institute of Mathematics and Computer Science, 2 (10), 117-128, (2011).
- Krishnamoorthy .A., Sreenivasan .C., An $M/M/2$ queueing system with heterogeneous servers including one with working vacation, International Journal of Stochastic Analysis, 1-16, (2012).
- Kumar .B.K., Madheswari .S.P., Transient behavior of the $M/M/2$ queue with catastrophes, Statistica, 62, 129-136, (2002).

Kumar .B.K., Madheswari .S.P., An $M/M/2$ queueing system with heterogeneous servers and multiple vacations, *Mathematical and Computer Modelling*, 41, 1415-1429, (2005).

Kumar .B.K., Madheswari .S.P., Venkatakrisnan .K.S., Transient solution of an $M/M/2$ queue with heterogeneous servers subject to catastrophes, *Information and Management Sciences*, 18, 63-80, (2007).

Kumar .R., Sharma .S.K., $M/M/1/N$ Queueing system with retention of reneged customers, *Pakistan Journal of Statistics and Operation Research*, 8, 859-866, (2012).

Kumar .R., Sharma .S.K., An $M/M/1/N$ Queueing model with retention of reneged customers and balking, *American Journal of Operational Research*, 2(1), 1-5, (2012).

Kumar .R., Sharma .S.K., A Markovian feedback queue with retention of reneged customers and balking, *AMO-Advanced Modeling and Optimization*, 14(3), 681-688, (2012).

Kumar .R., Som .B.K., An $M/M/1/N$ queueing system with reverse balking and reverse renegeing, *AMO-Advanced Modeling and Optimization*, 16(2), (2014).

Liu .W., Kumar .P., Optimal control of a queueing system with two heterogeneous servers, *IEEE Transactions on Automatic Control*, 29, 696-703, (1984).

Madan .K.C., Abu-Dayyeh .W., Taiyyan .F., A two server queue with Bernoulli schedules and a single vacation policy, *Applied Mathematics and Computation*, 145, 59-71, (2003).

Neuts .M.F., Takahashi .Y., Asymptotic behavior of the stationary distribution in the $GI/PH/c$ queue with heterogeneous servers, *Zeitschrift fur Wahrscheinlichkeitsh*, 57, 441-452, (1981).

Sharma .S.K., Kumar .R., A Markovian feedback queue with retention of reneged customers and balking, *AMO-Advanced Modeling and Optimization*, 14(3), 681-688, (2012).

Sherif .I.A., Transient analysis of a two-heterogeneous servers queue with impatient behavior, *Journal of the Egyptian Mathematical Society*, 22, 90-95, (2014).

Som .B.K., Decision making uncertain environment-A queueing theory approach international journal of advanced engineering, Management and Science (IJAEMS), 2(6), (2016).

Wang .K-H., Chang .Y-C., Cost analysis of a finite $M/M/R$ queueing system with balking, renegeing and server breakdowns, *Mathematical Methods of Operations Research*, 56(2), 169-180, (2002).

Yue .D., Yu .J., Yue .W., A Markovian queue with two heterogeneous servers and multiple vacations, *Journal of Industrial and Management Optimization*, 5, 453-465, (2009).

Yue .D., Yue .W., A heterogeneous two-server network system with balking and a Bernoulli vacation schedule, *Proceedings of the 4th International Conference on Queueing Theory and Network Applications*, Article No. 20, (2009).

Yue .D., Yue .W., Analysis of two server queues with a variant vacation policy, *The Ninth International Symposium on Operations Research and Its Applications (ISORA'10)*, 483-491, (2010).

Chapitre 3

Queueing system with renegeing and no waiting line

Ce chapitre fut l'objet d'une publication dans un journal international PROBSTAT FORUM, Volume 11, April 2018, Pages 67 – 76

Queueing system with renegeing and no waiting line

Amina Angelika BOUCHENTOUF , Aicha MESSABIHI

Departement of Mathematics, University of Sidi Bel Abbas, B. P.89, Sidi Bel Abbas, 22000, Algerie.

E-mail : bouchentouf-amina@yahoo.fr

Departement of Mathematics, University Center of Ain Temouchent, Ain Temouchent, Algerie

E-mail : amessabihi@yahoo.fr

Abstract. This paper deals with a heterogeneous two-server queueing system with renegeing and no waiting line. In this work, renegeing involves the situation where customers begin receiving service may disengage before service completion, and via certain mechanism, these latter are retained in the system. After the completion of service, each customer may rejoin the system as a feedback customer for receiving another regular service with some probability or can leave the system. We obtain the stationary state probabilities and deduce the explicit expressions of different performance measures of the system. Finally, we present some numerical examples to demonstrate how the various parameters of the model influence the behavior of the system.

Mathematics Subject Classification. 60K25, 68M20, 90B22.

Keywords. Queueing models, no waiting line, renegeing, Bernoulli feedback.

3.1 Introduction

Motivated by the application to telephone call centers and more general customer contact centers, this work focuses on the queueing systems with impatient customers. Recently there has been a great interest in multiserver queueing systems with impatient customer. The majority of call centers can be classified into two categories : revenue-generating and service-oriented. An important aspect for modeling of service-oriented call centers is the impatience behavior of the customers. Two usual modes in which customers advertise their impatience are balking and renegeing ; a customer refuses to enter the queue if the wait is too long or the queue is too big, this is the balking behavior. On the other hand, a customer who is waiting to be served might hang up (renege) before getting service if the wait in line becomes too long ; this is the renegeing behavior. Of course, there can be a combination of the two. The first who considered this type of queues was in Haight (1957), where he studied an $M/M/1$ queue with balking. An $M/M/1$ queue with renegeed customers was also proposed by Haight (1959) The combined effects of balking and renegeing in an $M/M/1/N$ queue have been investigated by Ancker and Gafarian (1963) After that multiple works were given, let's cite for instance Abou elata and Hariri(1992) considered the multiple servers queueing system $M/M/c/N$ with balking and renegeing, then Wang and Chang

(2002) extended this work to study an $M/M/c/N$ queue with balking, reneging and server breakdowns. Zeltyn and Mandelbaum (2005) studied call centers with impatient customers : many server asymptotic of the $M/M/n+G$ queue, where this queue is characterized by Poisson arrivals, exponential service times, n service agents and generally distributed patience times of customers. Dai and He (2013) studied Many-server queues with customer abandonment, numerical analysis of their diffusion models was given. Tkachenko(2013) focused on the multi-channel queueing system with heterogeneous servers, regenerative input flow and balking, where servers times are random variables but not necessary exponential. Jose and Manoharan (2011) dealt with Markovian queueing system with random balking, after that Jose and Manoharan (2011) presented the optimum system capacity for Markovian queueing system with an adaptive balking. Feedback represents the case where customer after getting incomplete or dissatisfying service comes back to the system for another regular service. In computer communication, the transmission of protocol data unit is sometimes repeated because of occurrence of an error. This generally occurs due to dissatisfying quality of service. Rework in industrial operations is also an example of a queue with feedback. Takacs(1963), Choi and Kim (1998) and Santhakumaran and Thangaraj(2000) considered a single server feedback queue with impatient and feedback customers, they studied $M/M/1$ queueing model for queue length at arrival epochs and obtained result for stationary distribution, mean and variance of queue length. Thangaraj and Vanitha (2009) obtained transient solution of $M/M/1$ feedback queue with catastrophes using continued fractions, the steady-state solution, moments under steady state and busy period analysis were calculated. Ayyapan and al.(2010) studied $M/M/1$ retrial queueing system with loss and feedback under non preemptive priority service by matrix geometric method. Kumar and Sharma (2012)($M/M/1/N$ Queueing systeme with retention of reneging customers) studied a single server queueing system with retention of renege customers. Kumar and Sharma (2012) ($M/M/1/N$ Queueing model with retention of reneging customers and balking) studied a single server queueing system with retention of renege customers and balking. Sharma and Kumar (2012) considered a single server, finite capacity Markovian feedback queue with reneging, balking and retention of renege customers in which the inter-arrival and service times follow exponential distribution. El-Paoumy and Nabwey(2011) studied the $M/M/2/N$ queue with general balk function, reneging and two heterogeneous servers. Most of the earlier work on multiserver queueing models deal with homogeneous servers; that is, the individual service rates are same for all servers in the system. But this assumption may be valid only when the service process is mechanically or electronically controlled. In a queueing system with human servers the above assumption is highly unrealistic. Often servers providing identical service, serves at different rates. This motivated the researchers to study multiserver queueing system with heterogeneous servers by Krishnamoorthy and Sreenivasan (2012) Heterogeneity of service is a common feature of many real multi-server queueing situations. The heterogeneous service mechanisms are invaluable scheduling methods that allow customers to receive different quality of service. Heterogeneous service is clearly a main feature of the operation of almost

any manufacturing system. The role of quality and service performance are crucial aspects in customer perceptions and firms must dedicate special attention to them when designing and implementing their operations Dharmaraja and Kumar (2015). The queueing systems with heterogeneous servers have received a considerable attention in the literature, let's cite for instance Cooper (1972), Gross and Harris (1998), Saglam and Torun (2005) and Saglam and Shahbazov (2007).

In this work we consider an heterogeneous two-server queueing model with Bernoulli feedback, renegeed customers, retention of renegeed customers and no waiting line. In our model customers being receiving service may disengage before service completion and using certain mechanism they can be retained with some probability. We obtain the stationary state probabilities and deduce the explicit expressions of the mean system size, and the mean waiting time of customers in the system. Finally, we present some numerical examples to demonstrate how the various parameters of the model influence the behavior of the system.

3.2 Model description

Consider an $M/M/2$ queue with Bernoulli feedback, renegeed customers, retention of renegeed customers and no waiting space. Customers arrive at the service station one by one according to Poisson stream with arrival rate λ . There is a two heterogeneous servers which provide service to all arriving customers. Service times are independently and identically distributed exponential random variables with rates $\mu_i; i = 1, 2$. After completion of each service, the customer can either rejoin with probability δ' or he can leave the system with probability δ where $\delta + \delta' = 1$. There is no distinguishing between the regular arrival and feedback arrival. The regular arrivals or feedback arrivals may become impatient when the service is so long and can disengage before service completion. In fact, each customer, at the beginning of services activates an individual timer, which follows an exponential distribution with parameters ν_1, ν_2 for services 1 and 2 respectively. This time is renegeing time of an individual customer after which customer either decide to abandon the queue with probability γ or retained with complimentary probability. The customers are served according to the following discipline :

- If the two servers are busy the customers leave the system.
- If one server is free, the first customer who come to the system goes to it.
- If both servers are free, the head customer chooses server 1 with probability π_1 and server 2 with probability π_2 , where $\pi_1 + \pi_2 = 1$.

Let P_n be the probability that there are n customers in the system. Applying the Markov process theory, we obtain the following set of steady-state equations :

$$\begin{aligned}
 P_{0,0} &= \mathbb{P}(\text{there is no customer in the system}); \\
 P_{1,0} &= \mathbb{P}(\text{there is one customer being served by server 1}); \\
 P_{0,1} &= \mathbb{P}(\text{there is one customer being served by server 2}); \\
 P_2 &= \mathbb{P}(\text{there are 2 customers in the system}).
 \end{aligned}$$

Also, $P_0 = P_{0,0}$; $P_1 = P_{1,0} + P_{0,1}$ and $P_2 = P_{1,1}$.

3.3 Steady state solution

In this section, the steady state probabilities are derived by the Markov process method.

Let $P(n)$ denotes the probability that are n customers in the system in steady state.

Applying the Markov process theory, the differential-difference equations of the model are :

$$\frac{dP_{0,0}(t)}{dt} = -\lambda P_{0,0}(t) + (\delta\mu_1 + \gamma\nu_1)P_{1,0} + (\delta\mu_2 + \gamma\nu_2)P_{0,1}(t), \quad (3.1)$$

$$\frac{dP_{1,0}(t)}{dt} = -(\lambda + \delta\mu_1 + \gamma\nu_1)P_{1,0}(t) + (\delta\mu_2 + \gamma\nu_2)P_{1,1}(t) + \lambda\pi_1 P_{0,0}(t), \quad (3.2)$$

$$\frac{dP_{0,1}(t)}{dt} = -(\lambda + \delta\mu_2 + \gamma\nu_2)P_{0,1}(t) + (\delta\mu_1 + \gamma\nu_1)P_{1,1}(t) + \lambda\pi_2 P_{0,0}(t), \quad (3.3)$$

$$\frac{dP_1(t)}{dt} = -\lambda P_1(t) + (\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))P_2(t) \quad (3.4)$$

Theorem 3.3.1. *If we have heterogeneous two-server queueing system with Bernoulli feedback, reneging, retention of reneged customers and no waiting line, then 1. The steady-state equations are*

$$\lambda P_{0,0} = (\delta\mu_1 + \gamma\nu_1)P_{1,0} + (\delta\mu_2 + \gamma\nu_2)P_{0,1} \quad (3.5)$$

$$(\lambda + \delta\mu_1 + \gamma\nu_1)P_{1,0} = (\delta\mu_2 + \gamma\nu_2)P_{1,1} + \lambda\pi_1 P_{0,0} \quad (3.6)$$

$$(\lambda + \delta\mu_2 + \gamma\nu_2)P_{0,1} = (\delta\mu_1 + \gamma\nu_1)P_{1,1} + \lambda\pi_2 P_{0,0} \quad (3.7)$$

$$\lambda P_1 = (\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))P_2 \quad (3.8)$$

2. The steady-state-probabilities P_n of system size are given by

$$P_{1,0} = \left\{ \frac{\lambda + (\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))\pi_1}{2\lambda + \delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)} \frac{\lambda}{\delta\mu_1 + \gamma\nu_1} \right\} P_{0,0}. \quad (3.9)$$

$$P_{0,1} = \left\{ \frac{\lambda + (\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))\pi_2}{2\lambda + \delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)} \frac{\lambda}{\delta\mu_2 + \gamma\nu_2} \right\} P_{0,0}. \quad (3.10)$$

$$P_1 = \frac{\lambda(\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))}{(\delta\mu_1 + \gamma\nu_1)(\delta\mu_2 + \gamma\nu_2)} \left\{ \frac{\lambda + (\delta\mu_1 + \gamma\nu_1)\pi_2 + (\delta\mu_2 + \gamma\nu_2)\pi_1}{2\lambda + \delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)} \right\} P_{0,0}, \quad (3.11)$$

and

$$P_2 = \frac{\lambda^2}{(\delta\mu_1 + \gamma\nu_1)(\delta\mu_2 + \gamma\nu_2)} \left\{ \frac{\lambda + (\delta\mu_1 + \gamma\nu_1)\pi_2 + (\delta\mu_2 + \gamma\nu_2)\pi_1}{2\lambda + \delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)} \right\} P_{0,0}, \quad (3.12)$$

with

$$P_{0,0} = \left(1 + \left\{ \frac{(\lambda + (\delta\mu_1 + \gamma\nu_1)\pi_2 + (\delta\mu_2 + \gamma\nu_2)\pi_1)(\lambda^2 + \lambda(\delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2)))}{(2\lambda + \delta(\mu_1 + \mu_2) + \gamma(\nu_1 + \nu_2))((\delta\mu_1 + \gamma\nu_1)(\delta\mu_2 + \gamma\nu_2))} \right\} \right)^{-1}. \quad (3.13)$$

Proof We obtain the steady-state-probabilities by using iterative method. Solving (3.6)-(3.7) we get (3.9) and (3.10). Then adding (3.9) and (3.10) we get (3.11). After that it suffices to replace (3.11) in (3.8) to obtain (3.13). Finally, by using the normalization condition we find (3.13).

3.4 Performance measures

This part of paper is devoted to present some of performance measures that are of general interest for the evaluation of the characteristic of the existing queueing system.

- The mean number of customers in the systems L_s .

$$L_s = P_1 + 2P_2. \quad (3.14)$$

Customers arrive into the system at the rate of λ . But not all the customers who arrive can join the system because of finite buffer restriction. The effective arrival rate into the system is thus different from the overall arrival rate and is given by

$$\lambda' = \lambda(1 - P_2). \quad (3.15)$$

- The expected number of customers served

$$\mathbb{E}(C.S) = \delta\mu_1 P_{1,0} + \delta\mu_2 P_{0,1} + 2\delta(\mu_1 + \mu_2)P_2. \quad (3.16)$$

We assumed that each customer has a random patience time following $exp(\gamma)$. Thus, the renegeing rate of the system would depend on the state of the system. Thus

- The average renegeing rate

$$R_{ren} = \gamma\nu_1 P_{1,0} + \gamma\nu_2 P_{0,1} + 2\gamma(\nu_1 + \nu_2)P_2. \quad (3.17)$$

- Average retention rate

$$R_{ret} = (1 - \gamma)\nu_1 P_{10} + (1 - \gamma)\nu_2 P_{01} + 2(1 - \gamma)(\nu_1 + \nu_2)P_2. \quad (3.18)$$

- Rate of abandonment

$$R_a = \lambda - \delta\mu_1 P_{1,0} - \delta\mu_2 P_{0,1} - 2\delta(\mu_1 + \mu_2)P_2. \quad (3.19)$$

In system management, customers who leave the system (renege) represent business lost. Consequently, It is interesting to present the proportion of customers lost. So, using (3.15) and (3.17), we get proportion of customers lost due to renegeing out of

those arriving and joining the system

$$P_{C_{Lost}} = \text{Average reneing rate}/\lambda'. \quad (3.20)$$

3.5 Numerical solution and graphical representation

In this section some numerical examples are carried out in order to show the impact of different parameters .

3.5.1 Impact of service rates on the characteristics of the system

Let us vary λ from 2 to 7 with a pitch 1, and take $\gamma = 0.6$, $\delta = 0.2$, $\nu_1 = 5$, $\nu_2 = 4$, $\pi_1 = 0.3$, $\pi_2 = 0.7$, and two cases of service rates, $\mu_1 = 2$, $\mu_2 = 6$ then $\mu_1 = 14$, $\mu_2 = 20$. The numerical results obtained for these situations are stored in Tables 3.1 and 3.2.

λ	L_s	$\mathbb{E}(C.S)$	R_{ren}	R_{ret}	R_a	$P_{C_{Lost}}$	λ'
2	0.476439791	0.843754891	2.702721806	1.801814538	1.156245109	1.601536556	1.687580465
3	0.743414884	1.277981417	4.172534311	2.781689541	1.722018583	1.919303611	2.173983463
4	0.967139887	1.634283246	8.149404788	3.594413177	2.365716754	3.302165405	2.467897209
5	1.144495874	1.912847892	9.755495839	4.233589431	3.087152108	3.70179907	2.635339102
6	1.282730368	2.127761543	11.02277795	4.728473629	3.872238457	4.045139641	2.724943742
7	1.390566941	2.311588583	12.15289377	5.19987643	4.688411417	4.39075051	2.767839744

TABLE 3.1: Variation of different characteristics of the system vs. λ when $\mu_1 = 2$, and $\mu_2 = 6$

λ	L_s	$\mathbb{E}(C.S)$	R_{ret}	R_a	R_{ren}	$P_{C_{Lost}}$	λ'
2	0.17842094	1.632472836	0.827021316	0.367527164	1.240531974	0.650572996	1.906829796
3	0.306987035	2.651645267	1.359762368	0.348354733	2.039643553	0.751530268	2.713987235
4	0.444346861	3.690555575	1.907466425	0.309444425	3.958728761	1.167676489	3.390261598
5	0.581065815	4.691155416	2.437710915	0.308844584	5.187925768	1.317884415	3.936555976
6	0.711087561	5.619142347	2.931154309	0.380857653	6.35835232	1.456561768	4.365315952
7	0.831120551	6.545269871	3.440492887	0.454730129	7.532680202	1.604764861	4.693946375

TABLE 3.2: Variation of different characteristics of the system vs. λ when $\mu_1 = 14$, $\mu_2 = 20$

Graphical representations (see Figure 3.1) illustrate the details of some results given in Tables 3.1-3.2.

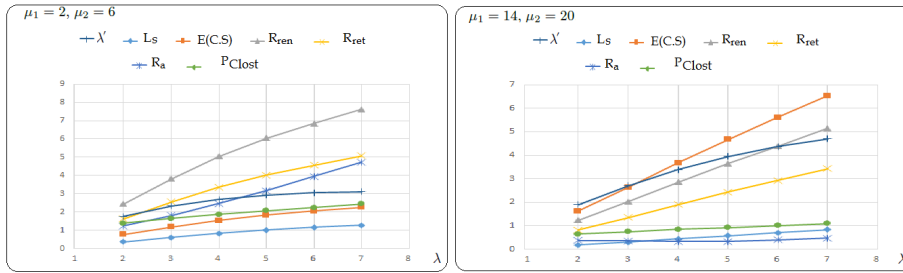


FIGURE 3.1 – Variation of different characteristics of the system vs. λ

According to Tables 3.1–3.2 and Figure 3.1 we observe that :

- Along the increase of λ , the mean number of customer in the system L_s , the effective arrival rate λ' , the mean number of customers served $\mathbb{E}(C.S)$, the average reneing rate R_{ren} , the average retention rate R_{ret} and the average abandonment rate R_a increase.
- Comparing the results when $\mu_1 = 2, \mu_2 = 6$ and $\mu_1 = 14, \mu_2 = 20$, we remark that when $\mu_1 = 14, \mu_2 = 20$ (the mean service rates small), the size of the system L_s , the average reneing rate R_{ren} , the average retention rate R_{ret} and the average abandonment rate R_a are smaller than the case where $\mu_1 = 2, \mu_2 = 6$. On the other side the effective arrival rate λ' , and the mean number of customers served $\mathbb{E}(C.S)$ are bigger in the first case. This is absolutely agree with our intuition.

3.5.2 Impact of reneing (impatient) rates on the characteristics of the system

Let's vary ν_1 from 2 to 12 with a pitch 2, and take $\lambda = 10, \delta = 0.6, \gamma = 0.7, \mu = \mu_1 = 2, \mu_2 = 10, \nu_2 = 3, \pi_1 = 0.6$, and $\pi_2 = 0.4$.

Graphical representation is presented in Figure 3.2.

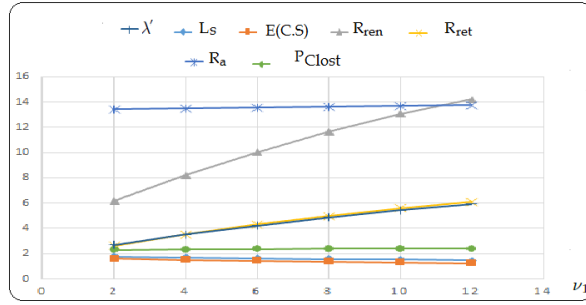


FIGURE 3.2 – Variation of different characteristics of the system vs. ν_1

From Figure 3.2 we observe that along the increase of ν_1 , the renegeing rate R_{ren} increases, that is customers become more impatient and consequently they leave the system, this implies that R_a and R_{ret} increase, more customers abandon the system, more the probability of retention becomes high. On the other side the increase of ν_1 leads to a decrease in L_s and of the mean number of customers served $E(C.S)$ which is absolutely explicable; more the average rate of impatience becomes small, more customers leave the system without getting their complete service.

- Now, let's vary μ_1 from 2 to 6 and put $\lambda = 8$, $\delta = 0.6$, $\gamma = 0.8$, $\mu_2 = 6$, $\pi_1 = 0.3$, $\pi_2 = 0.7$ then take $\nu_1 = 1$, $\nu_2 = 8$, and $\nu_1 = 10$, $\nu_2 = 12$.

The numerical results obtained for these situations are stored in Tables 3.3 and 3.4.

μ_1	L_s	$E(C.S)$	R_{ren}	R_{ret}	R_a	P_{Clost}	λ'
2	0.618092354	3.135806207	11.41074943	2.852687358	4.864193793	1.823125302	6.258894777
2.8	0.586723744	3.329861899	11.01283117	2.753207793	4.670138101	1.732212858	6.357666217
3.6	0.558990039	3.508801206	10.65486087	2.663715217	4.491198794	1.653174801	6.445090294
4.4	0.5342895	3.674267347	10.33104325	2.582760812	4.325732653	1.583776864	6.523042156
5.2	0.512146455	3.827695969	10.03666234	2.509165586	4.172304031	1.522320019	6.593004244
6	0.492179919	3.958998281	9.74263519	2.435658797	4.041001719	1.463701184	6.656164044

TABLE 3.3: Variation of different characteristics of the system vs. μ_1 when $\nu_1 = 10$, $\nu_2 = 12$

μ_1	L_s	$E(C.S)$	R_{ren}	R_{ret}	R_a	P_{Clost}	λ'
2	0.917824773	5.275301511	7.37001571	1.842503927	2.724698489	1.473113043	5.003021148
2.8	0.882597554	5.64189026	7.084369574	1.771092394	2.35810974	1.378532005	5.139067897
3.6	0.850905468	5.95482707	6.848031147	1.712007787	2.04517293	1.30146951	5.261768405
4.4	0.822227934	6.229937518	6.64563771	1.661409428	1.770062482	1.236840976	5.373073692
5.2	0.796142615	6.476308669	6.468488286	1.617122072	1.523691331	1.181552398	5.474567439
6	0.772303227	6.667360009	6.303881429	1.575970357	1.332639991	1.132254346	5.567548891

TABLE 3.4: Variation of different characteristics of the system vs. μ_1 when $\nu_1 = 1$, $\nu_2 = 8$

Graphical representations (see Figure 3.3) illustrate the details of some results given in Tables 3.3-3.4.

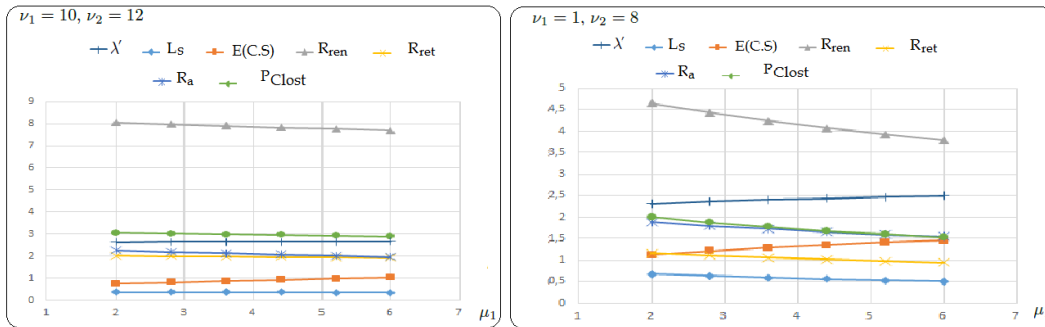


FIGURE 3.3 – Variation of different characteristics of the system vs. μ_1

- According to Tables 3.3-3.4 and Figure 3.3 we observe that the increase of μ_1 implies that customers are served faster, on the other hand when the mean service rate is small, R_a , R_{ren} , R_{ret} and the mean number of customers served $\mathbb{E}(C.S)$ are high, thus the mean number of customers in the system decreases which agree absolutely with our intuition.

- Comparing results given in Figure 3.3, we remark that when the reneing rates are $\nu_1 = 1, \nu_2 = 8$ (mean reneing rates large) the mean number of customers in the systems L_s , the average reneing rate R_{ren} , the average of retention reneeged customers R_{ret} and rate of abandonment R_a , are higher than the case where $\nu_1 = 10, \nu_2 = 12$ (average reneing rates small), contrariwise the mean number of customers served in the first case is smaller than the second one.

3.5.3 Impact of probability of retention reneeged customers on the characteristics of the system

Let's vary λ from 1 to 11, and take $(\gamma = 0.3, 0.8)$, $\delta = 0.6$, $\nu_1 = 5$, $\nu_2 = 7$, $\mu_1 = 6$, $\mu_2 = 4$, $\pi_1 = 0.6$ and $\pi_2 = 0.4$.

The numerical results obtained for these situations are presented in Tables 3.5 and 3.6.

λ	L_s	$E(C.S)$	R_{ren}	R_{ret}	R_a	$P_{C_{Lost}}$	λ'
1	0.153932923	0.78692642	0.459842337	1.072965454	0.21307358	0.469148707	0.980163283
3	0.505336371	2.725929897	1.619914928	3.7798015	0.274070103	0.619882094	2.613262981
5	0.810148495	4.487651909	2.68165434	6.25719346	0.512348091	0.721871089	3.714865966
7	1.042030913	5.864515497	3.512749359	8.196415172	1.135484503	0.795456582	4.41601646
9	1.213423609	6.901824353	4.138933535	9.657511582	2.098175647	0.851049532	4.863328605
11	1.340917486	7.914702367	4.707209673	10.98348924	3.085297633	0.913095298	5.155222774

TABLE 3.5: Variation of different characteristics of the system vs. λ when $\gamma = 0.3$

λ	L_s	$\mathbb{E}(C.S)$	R_{ren}	R_{ret}	R_a	$P_{C_{Lost}}$	λ'
1	0.090851848	0.464880042	0.701788314	0.175447078	0.535119958	0.70744613	0.992002478
3	0.303447552	1.616796358	2.499795805	0.624948951	1.383203642	0.885761802	2.8221987
5	0.518642404	2.830052809	4.42561302	1.106403255	2.169947191	1.022018865	4.33026549
7	0.712786561	3.951319698	6.218383555	1.554595889	3.048680302	1.129617646	5.504856956
9	0.878906357	4.927308153	7.784914751	1.946228688	4.072691847	1.216778002	6.3979746
11	1.017835087	5.954850279	9.33829074	2.334572685	5.045149721	1.320316579	7.072766404

TABLE 3.6: Variation of different characteristics of the system vs. λ when $\gamma = 0.8$

Graphical representations (see Figure 3.4) illustrate the details of some results given in Tables 3.5-3.6.

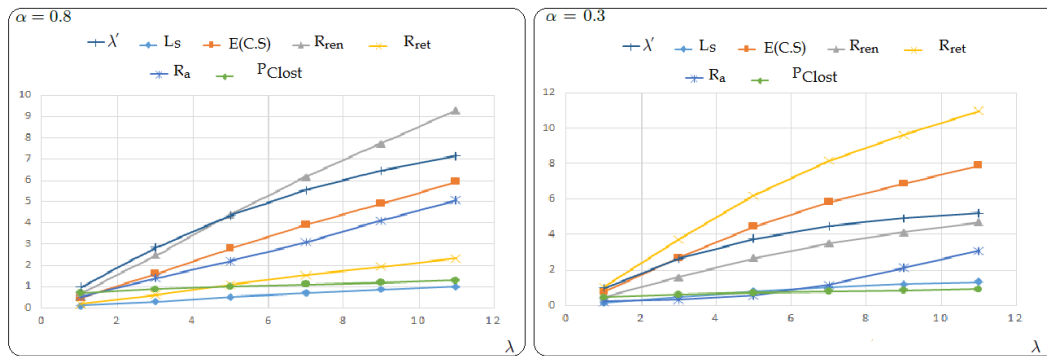


FIGURE 3.4 – Variation of different characteristics of the system vs. λ

Comparing the evolution of different performance measures given in Figure 3.4 and Tables 3.5-3.6, we remark that - With the increase in the probability of non-retention γ , the size of the system decreases and consequently we get a decrease in the number of customers served, this leads to a lose in potential customers which has a negative impact.

- When the probability of retaining customers is big $\gamma' = 0.7$ ($\gamma = 0.3$), the size of the system L_s , the mean number of customers served $\mathbb{E}(C.S)$, the average retention rate are bigger compared to the case where $\gamma' = 0.2$ ($\gamma = 0.8$). Then when the probability of retaining customers is big $\gamma' = 0.7$ ($\gamma = 0.3$) the average reneing rate, the rate of abandonment, the proportion of lost customers and the effective arrival rate are smaller compared to the case where $\gamma' = 0.2$ ($\gamma = 0.8$).

3.5.4 Impact of probability of feedback customers on the characteristics of the system

Let's vary λ from 1 to 2, and take ($\delta = 0.2, 0.8$), $\gamma = 0.4$, $\nu_1 = 8$, $\nu_2 = 7$, $\mu_1 = 6$, $\mu_2 = 4$, $\pi_1 = 0.3$ and $\pi_2 = 0.7$.

The numerical results obtained for these situations are stored in Tables 3.7 and 3.8.

λ	L_s	$\mathbb{E}(C.S)$	R_{ren}	R_{ret}	R_a	$P_{C_{Lost}}$	λ'
1	0.165067179	0.358919856	1.113772518	1.670658778	0.641080144	1.157184609	0.962484732
1.2	0.20866857	0.451297297	1.39557527	2.093362905	0.748702703	1.226987841	1.137399429
1.4	0.254284479	0.547399675	1.687827038	2.531740556	0.852600325	1.294192737	1.304154312
1.6	0.301345291	0.646066042	1.987117815	2.980676722	0.953933958	1.358952627	1.462242152
1.8	0.349337015	0.746256427	2.290396607	3.43559491	1.053743573	1.421407651	1.611358012
2	0.397804327	0.852847204	2.610426828	3.915640241	1.147152796	1.490503666	1.751372296

TABLE 3.7: Variation of different characteristics of the system in the case $\delta = 0.2$

λ	L_s	$\mathbb{E}(C.S)$	R_{ren}	R_{ret}	R_a	$P_{C_{Lost}}$	λ'
1	0.060082509	0.72037064	0.506094201	0.759141301	0.27962936	0.510815908	0.990756537
1.2	0.07551928	0.893405645	0.625745401	0.938618101	0.306594355	0.528394319	1.184239455
1.4	0.091929453	1.074431595	0.750458779	1.125688168	0.325568405	0.545662176	1.375317572
1.6	0.109240323	1.262740181	0.879772802	1.319659203	0.337259819	0.562629285	1.563681139
1.8	0.127379465	1.457632941	1.01323224	1.51984836	0.342367059	0.579304922	1.749048214
2	0.146275168	1.663917485	1.153645485	1.730468228	0.336082515	0.597383294	1.931164627

TABLE 3.8: Variation of different characteristics of the system in the case $\delta = 0.8$

Graphical representations (see Figure 3.5) illustrate the details of some results given in Tables 3.7-3.8.

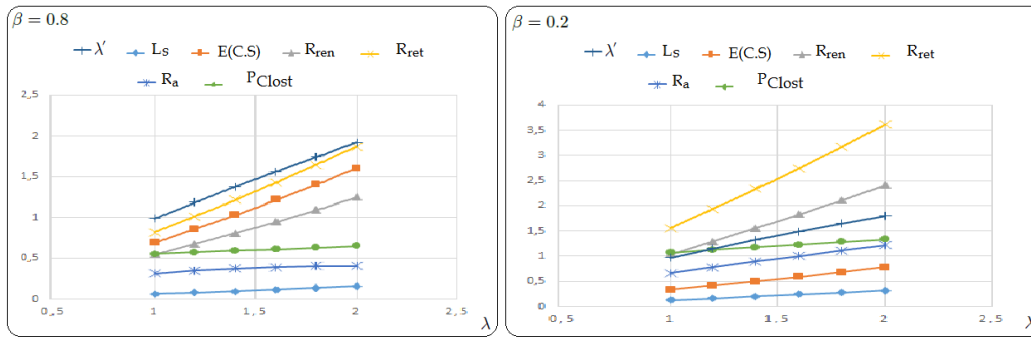


FIGURE 3.5 – Variation of different characteristics of the system vs. λ

Comparing the results in Figure 3.5 and Tables 3.7-3.8 we observe that - With the increase of λ , the mean number of customers in the system L_s , the average reneing rate, average retention rate, proportion of customer lost due to reneing out of those arriving and joining the system, $P_{C_{Lost}}$, the expected of customers served $\mathbb{E}(C.S)$ increase, and when $\delta = 0.2$ the rate of abandonment increases, while when $\delta = 0.8$ it decreases.

- When the probability of feedback $\delta' = 0.8$, ($\delta = 0.2$) is big, the average reneing rate, average retention rate, proportion of customer lost due to reneing out of those arriving and joining the system, $P_{C_{Lost}}$, rate of abandonment are bigger then the case where $\delta' = 0.2$, except $\mathbb{E}(C.S)$ and L_s which are smaller in the first case.

3.6 Conclusion

In this paper we analyzed a heterogeneous two-server queueing system with Bernoulli feedback, reneing, retention of reneged customers and no waiting line. The stationary state probabilities were obtained and the explicit expressions of different characteristics of the system were deduced. Finally, some numerical examples have been presented to demonstrate how the various parameters of the model influence the behavior of the system.

For further work, this model can be studied under the provision of time dependent arrival and service rate. The cost-profit analysis of the model can also be carried out to study its economic analysis.

Bibliography

- Abou El-Ata .M.O., Hariri. A. M. A. , The $M/M/c/N$ queue with balking and reneging, Computers and Operations Research, 19(13), 713-716, (1992).
- Ancker .Jr, C.J., Gafarian .A.V., Some queuing problems with balking and reneging I, Operations Research, 11(1), 88-100,(1963).
- Ancker .Jr, C. J., Gafarian .A.V., Some queuing problems with balking and reneging II, Operations Research, 11(6), 928-937, (1963).
- Ayyapan .G., Muthu Ganapathi Subramanian .A., Sekar .G. , $M/M/1$ Retrial queueing system with loss and feedback under non-pre-emptive priority service by matrix geometric method, Applied Mathematical Sciences, 4(48), 2379-2389.(2010)
- Bouchentouf .A.A., Messabihi .A., Queueing system with reneging and no waiting line, Probst Forum, 11, 67-76, (2018).
- Choi .B.D., Kim .Y.C., The $M/M/c$ Retrial queue with geometric loss and feedback, Computers Math. Applic, 36(6), 41-52.(1998).
- Choudhury .G., Paul .M. , A Two phase queueing system with Bernoulli feedback, Information and Management Sciences, 16(1), 35-52.(2005).
- Cooper .R.B., Introduction to queuing theory, The Macmillan Company, New York, NY, USA.(1972).
- Dai .J.G., He .S., Many-server queues with customer abandonment : numerical analysis of their diffusion models, Stochastic Systems, 3(1), 96-146, (2013).
- D'Avignon .G.R., Disney .R.L., Single-server queues with state-dependent Feedback, INFOR , 14, 71-85,(1976).
- Dharmaraja .S., Kumar .R., Transient solution of a markovian queueing model with heterogeneous servers and catastrophes, OPSEARCH, 52(4), 810-826.(2015).
- El-Paoumy .M.S., Nabwey .H.A., The poissonian queue with balking function, reneging and two heterogeneous servers, International Journal of Basic and Applied Sciences IJBAS-IJENS., 11(6), 149-152,(2011).
- Gross .D., Harris .C.M., Fundamentals of queueing theory, John Wiley and Sons, New York, NY,USA, 3rd edition, (1998).
- Haight .F.A., Queueing With reneging, Metrika, 2, 186-197,(1959).
- Jose, Joby .K., Manoharan .M., Markovian queueing system with random balking, OPSEARCH., 48(3), 236-246, (2011).
- Jose, Joby .K., Manoharan .M., Optimum system capacity for markovian queueing system with an adaptive balking, Journal of Statistics & Management Systems, 17(2), 155-164, (2014).
- Krishnamoorthy .A., Sreenivasan .C., An $M/M/2$ queueing system with heterogeneous servers including one with working vacation, International Journal of Stochastic Analysis , Article ID 145867, 16 pages,(2012).

Kumar .R., Sharma .S.K., *M/M/1/N* queuing system with retention of renege customers, Pakistan Journal of Statistics and Operation Research, 8, 859-866,(2012).

Kumar .R., Sharma .S.K., An *M/M/1/N* Queuing model with retention of renege customers and balking, American Journal of Operational Research, 2(1), 1-5,(2012).

Saglam .V.,Torun .H., On optimization of stochastic service system with two heterogeneous channels, International Journal of Applied Mathematics, 17(1), 1-6,(2005).

Saglam .V., Shahbazov .A., Minimizing loss probability in queuing systems with heterogeneous servers, Iranian Journal of Science and Technology, 31(2), 199-206,(2007).

Santhakumaran .A., Thangara .V., A Single server queue with impatient and feedback customers, Information and Management Science, 11(3), 71-79, (2000).

Sharma .S.K., Kumar .R.A., Markovian feedback queue with retention of renege customers and balking, AMO-Advanced Modeling and Optimization, 14(3), 681-688, (2012).

Takacs .L., A single-server queue with Feedback, The Bell System Tech. Journal, 42, 134-149.(1963).

Thangaraj .V., Vanitha .S., On the analysis of *M/M/1* feedback queue with catastrophes using continued fractions, International Journal of Pure and Applied Mathematics, 53(1), 131-151,(2009).

Tkachenko .A., Multichannel queuing systems with balking and regenerative input flow, Higher School of Economics Research, WP BRP 14/STI/(2013).

Wang .K.H., Chang .Y.C., Cost analysis of a finite *M/M/R* queueing system with balking, renege and server breakdowns, Mathematical Methods of Operations Research, 56(2), 169-180,(2002).

Zeltyn .S., Mandelbaum .A., Call centers with impatient customers : many-server asymptotics of the *M/M/n + G* queue, Queueing Systems, 51(3-4), 361-402.(2005).

Conclusion et perspectives

Les systèmes de files d'attente à serveurs hétérogènes et clients impatientes sont rencontrés dans plusieurs domaines. L'étude de tels systèmes est certainement très importante pour les applications pratique, car l'hétérogénéité du service et le mécanisme des clients impatientes dans le système a une influence importante sur l'ergodicité et les principaux mesures de performances du système.

Dans ce mémoire, nous avons étudié un modèle de files d'attente à serveurs hétérogènes avec feedback, dérobade et abandon des clients. En utilisant la méthode des équations de balance en état stationnaire, nous avons pu obtenir des résultats analytiques et certaines mesures de performance du modèle en question, telles que le nombre moyen de clients dans le système, ...etc. De plus, nous avons effectué une analyse de sensibilité du modèle d'attente étudié, tout en montrant l'effet de l'influence de changement des valeurs des paramètres du modèle sur ses caractéristiques.

Dans ce travail, nous nous sommes intéressés aux modèles $M/M/2$ et $M/M/2/N$ à serveurs hétérogènes muni du feedback, balking et reneging. Notre étude a pour objectifs d'analyser l'impacte de reneging et du balking sur le système de files d'attente et de fournir l'état stationnaire de chaque système. La contribution principale de ce travail est de donner une bonne modélisation pour les systèmes de files d'attente grâce à la théorie opérationnelle ainsi évaluer les mesures de performances. Nous avons d'abord rappelé quelques concepts et techniques de base de la théorie de files d'attente et nous avons introduit quelques systèmes de files d'attente classiques. Ensuite, nous avons présenté les systèmes de files d'attente avec dérobade et abandon. Puis une étude approfondis sur les modèles $M/M/2$ et $M/M/2/N$ avec feedback, dérobade et abandon a été présenté. Enfin, nous avons effectué une analyse numérique montrant l'effet de la variation de quelques paramètres sur les mesures de performances approprié aux systèmes. En terme de continuité ce travail, plusieurs perspectives de recherche peuvent être envisagées, on peut citer :

- Entreprendre la même démarche pour généraliser cette étude pour un modèle de files d'attentes à plusieurs serveurs hétérogènes.
- Etude de ces modèles de files d'attente avec rappels.

Bibliographie générale

- Abou-El-Ata .M.O., On poisson queues with both balking and heterogeneous servers, Delta Journal of Sciences, 2, 292-303, (1983).
- Abou El-Ata .M.O., Hariri .A.M.A., The $M/M/c/N$ queue with balking and reneging, Computers and Operations Research, 19(13), 713-716, (1992).
- Adan .I., Economou .A., Kapodistria .S., Synchronized reneging in queueing systems with vacations, Queueing Systems, 62, 1-33, (2009).
- Aissani .A., Kartashov .N.V., Ergodicity and stability of markov chains with respect to operator topology in the space of transition kernels, Sciences U.S.S.R, serie A 11, 3-5, (1983).
- Al-Seedy .R.O., A transient solution of the non-truncated queue $M/M/2$ with balking and an additional server for longer queues (Krishnamoorthi discipline), Applied Mathematics and Computation , 3, 763-769, (2004).
- Altman .E., Yechiali .U., Analysis of customers impatience in queues with server vacation, Queueing Systems, 52(4), 261-279, (2006).
- Altman .E., Yechiali .U., Infinite server queues with system's additional tasks and impatient customers, Probability in the Engineering and Informational Sciences, 22(4), 477-493, (2008).
- Ammar .S.I., Transient analysis of an $M/M/1$ queue with impatient behavior and multiple vacations, Applied Mathematics and Computation, 260, 97-105, (2015).
- Ammar .S.I., Transient solution of an $M/M/1$ vacation queue with a waiting server and impatient customers, Journal of the Egyptian Mathematical Society, 25, 337-342, (2017).
- Ancker.C.J. and Gafarian.A.V., Queueing with reneging and multiple heterogeneous servers, Nav. Res. Logist. Q., 10, 125-145, (1963).
- Ancker Jr.C. J. and Gafarian. A. V., Some queueing problems with balking and reneging. II, Operations Research , 11, 88- 100, (1963(a)).
- Ancker Jr.C. J. and Gafarian. A. V., Some queueing problems with balking and reneging. I, Operations Research , 11, 88- 100, (1963(a)).
- Andreas, Manfred, On the $M(n)/M(n)/s$ queue with impatient calls, Performance Evaluation, 35, 1-18, (1999).
- Artalejo.J. R., Gomez-Corral.A., Advances in retrial queues, European Journal of Operation Research, 189(3), 1042-1233, (2008).
- Ayyapan .G., Muthu Ganapathi Subramanian .A., Sekar .G., $M/M/1$ Retrial queueing system with loss and feedback under non-pre-emptive priority service by matrix geometric method, Applied Mathematical Sciences, 48, 2379-2389, (2010).
- Baccelli .F., Boyer .P., Hebutte .G., Single-server queues with impatient customers, Advances in Applied Probability, 16, 887-905, (1984).
- Bae .J., Kim .S., Lee .E.Y., The virtual waiting time of the $M/G/1$ queue with impatient customers, Queueing Systems, 38, 485-494, (2001).

Baek .J., Dudina .O., Kim .C., A queueing system with heterogeneous impatient customers and consumable additional items, *Int. J. Appl. Math.Comput. Sci.*, 27 (2), 367-384, (2017).

Baruah .M., Analysis of some batch arrival queueing systems with balking, reneging, random breakdowns fluctuating modes of service and bernoulli scheduled server vacations thesis, (2015).

Barrer .D.Y., Queueing with impatient customers and ordered service, *Operations Res*, 5, 650-656, (1957).

Belyaev. Y.K., Linear markov process and their application to reliability problems, *Probability Theory And Mathematical Statistics*, 309-323, (1962).

Bengtsson .B., On some control problems for queues, *Linkoping Studies In Science And Technology*, Dissertation Department of Electrical Engineering, Linkoping University, Sweden, 225, (1982).

Blackburn J.D., Optimal control of a single-server queue with balking and reneging, *Management Science*, 19, 297-313, (1972).

Boots .N., Tijms .H., An $M/M/c$ queue with impatient customers, *TOP*, 7, 213-220, (1999).

Borovkov .A., *A Stochastic Processes in queueing theory*, Springer-Verlag, Berlin, (1976).

Borovkov .K.A., Propagation of chaos for queueing networks, *Theory of Probability and its Applications*, 42, 385-394, (1998) .

Bouchentouf .A.A, Guendouzi .A.H., Cost Optimization dor an $M^x/M/c$ vacation queueing system with waiting servers and impatient customers, *Sema Journal*, 76 (2), 309-341, (2019).

Bouchentouf .A.A, Guendouzi .A.H., The $M^x/M/c$ bernoulli feedback queue with variant multiple working vacations and impatient customers performance and economic analysis, *Arabian Journal Of Mathematics*, 9 (2), 309-327, (2020).

Bouchentouf .A.A., Kadi .K., Rabhi .A., Analysis of two heterogeneous server queueing model with balking, reneging and feedback, *Math.Sci. Appl.E Notes*, 2 (2), 10-21, (2014).

Bouchentouf .A.A., Messabihi .A., Heterogenous two server queueing system with reverse balking and reneging, *Operational Research*, 55(2), 251-267, (2018)..

Bouchentouf .A.A., Messabihi .A., Queueing system with reneging and no waiting line, *Probst Forum*, 11, 67-76, (2018).

Bronstein .O.I., Dukhovny .I., *Priority queueing models in computer systems*, Nauka, Moscow, (1976) .

Brown .L, Gans .N., Mandelbaum .A., Sakov .A., Shen.H., Zeltyn.S. and Zhao.L., Statistical analysis of a telephone call center : A queueing-science perspective, *J. Amer. Statist. Assoc.*, 100, 36-50, (2005).

Choi .B.D., Chang .Y., Single server retrial queues with priority calls, *Mathematical and computer modelling*, 30, 7-32, (1999).

Choi .B.D., Kim .Y.C., The $M/M/c$ Retrial queue with geometric loss and feedback, *Computers Math. Applic*, 36(6), 41-52.(1998).

Choudhury .G., Paul .M. , A two phase queueing system with Bernoulli feedback, *International Journal of Information and Management Sciences*, 16(1), 35-52,(2005).

Cooper .R.B., *Introduction to queueing theory*, The Macmillan Company, New York, NY, USA.(1972).

Courtois .P.J., *Decomposability queueing and computer system applications*, Academic Press, (1977).

Cox .D.R., Some statistical methods connected with series of events, *Journal of The Royal Statistical Society : series B (methodological)*, 2517-6161, (1955).

Crommelin .C.D., Delay probability formulae when the holding times are constant, PostOffice Electrical Engineers Journal, 25, 41-50, (1932).

Crommelin .C.D., Delay probability formulae, Post Office Electrical Engineers Journal, 26, 266-274, (1934).

D'Avignon .G.R., Disney .R.L., Single Server Queue with State dependent feedback, INFOR, 14,71-85, (1976).

Dharmaraja .S., Kumar .R., Transient solution of a markovian queuing model with heterogeneous servers and catastrophes, OPSEARCH, 52(4), 810-826.(2015).

Disney .R.L., A matrix solution for the two server queue with overflow, Mangement science, 254-265, (1972).

Disney .R.L., Farrell .R.L., De Morais .P.R., A Characterization of $M/G/1$ queues with renewal departure processes, Manangement science, 11, 1213-1338, (1973) .

El-Paoumy .M.S., On poisson arrival queue : $M^x/M/2/N$ with balking, reneging and heterogeneous servers, Applied Mathematical Sciences, 2(24), 1169-1175, (2008).

El-Paoumy .M. S., Nabwey .H. A., The poissonian queue with balking function, reneging and two heterogeneous servers, International Journal of Basic and Applied Sciences., 11(6), 149-152, (2011).

Erlang .A. K., The theory of probabilities and telephone conversations. Nyt Tidsskrift for Matematik B, 20, 33, (1909).

Erlang .A. K., Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges, Elektroteknikeren, 13, 5, (1917).

Feng .W., Kowada .M., Adachi .K., Analysis of multi-server queue with two priority classes and (M,N) threshold service schedule : non-preemptive priority, 7, 653-671, (2000).

Franken .D., Konig .D., Arndt .U., Schmidt .V., Queues and point processes, Academie- verlag, Berlein, (1981).

Friedman .E.J., Landsberg .A.S., Long-run dynamics of queues : stability and chaos, Operations Research Letters , 18, 185-191, (1996)

Fry .T.C., Probability and its engineering uses, Van nostrand reinhold,15, 175-177, (1928).

Gao.J.Z., Wang.X., Analysis of a retrial queue with two type breakdowns and delayed rapairs, Ieee Access, 428-442, (2020).

Gnedenko .B.V., Kovalenko .I.N., Introduction to Queueing Theory, Birkhuser, (1989).

Gnedenko .B.V., Kovalenko .I.N., Introduction to Queueing Theory, Birkhuser, (1989).

Gong .W.B., Hu .J.K. , The Maclaurin series of the $GI/G/1$ queues, Journal of Applied Probability, 29 , 184-276, (1992).

Gross .D., Harris .C.M., Fundamentals of queuing theory, (Second Edition), Wiley, Hoboken, NJ, (1985).

Gross .D., Harris .C.M., Fundamentals of queueing theory, John Wiley and Sons, New York, NY,USA, 3rd edition, (1998).

Haight .F.A., Queuing with balking, I, Biometrika, 44, 360-369, (1957).

Haghighi .A., Medhi .J., Mohanty .S.G., On a multiserver markovian queuing system with balking and reneging, Com.Oper.Res, 13, 421-425, (1986).

Haight .F.A., Queueing with reneging, Metrika, 2, 186-197, (1959).

Haviv .M., Kerner .Y., On balking from an empty queue, Queueing Systems, 55, 239-249, (2007).

Haxholdt .C., Larsen .E.R., Van Ackere .A., Mode locking and chaos in a deterministic queueing model with feedback, Management Science, 49, 816-830 , (2003).

Heffer .J.C., Steady state solution of the $M/E_k/c$ (FIFO) queuing system, *CORSJ*, 7, 16-30, (1969).

Jain .N.K., Kumar .R., Som .B.K., An $M/M/1/N$ queuing system with reverse balking, *American Journal of Operational Research*, 4(2), 17-20, (2014).

Johansen .S.G., Stidham .S.Jr., Control of arrivals to a stochastic input-output system, *Adv. Appl. Probab*, 12, 972-999, (1980) .

Jones .L.K., Inferring balking behavior from transactional data, *Operations Research*, 47, 778-784, (1999).

Jose, Joby .K., Manoharan .M., Optimum system capacity for markovian queuing system with an adaptive balking, *Journal of Statistics and Management Systems*, 17(2), 155-164, (2014).

Kalyanaraman .R., Kalaiselvi .S., Heterogeneous server queue with A threshold on slow server, *International Journal of Recent Technology and Engineering (IJRTE)*, 7, 2277-3878, (2019).

Kalyanaraman .R., Senthilkumar .R., Heterogeneous server Markovian queue with restricted admissibility and with reneging, *Mathematical Sciences International Research Journal*, 7(1), 309-315, (2018).

Kartachov .N.V., Strongly stable markov chains, *Stability problems for stochastic models*, VSP utreicht, TbiMC Scientific Publishers, 54-91, (1981).

Kartachov .N.V., Strongly stable markov chains, *Journal Of Soviet Mathematics*, 34(2), 1493-1498, (1986).

Ke .J.C., Chang .F. M., Modified vacation policy for $M/G/1$ retrial queue with balking and feedback, *Computers and Industrial Engineering*, 57,

Kendall .D.G., Some problems in the theory of queues, *Journal of the Royal Statistical Society-Series B*, 13(2), 151-185, (1951).

Khintchine .A.Y, *Mathematical Methods in the theory of Queueing*, Chapter I, Hafner, New York, (1960).

Kim .C.S., Finite-Source queueing models for Analysis of complex communication Systems, *Journal of the Society of Korea Industrial and System Engineering* , 26, 62-67, (2003) .

Kingman .J.F.C., The effect of queue discipline on waiting time variance, *Proceeding of the Cambridge Philosophical Society*, 63, 163-164, (1961).

Kleinrock .L., *Queueing Systems*, John Wiley, New York, 417p, (1975).

Klimov .V.P., *Stochastic service systems*, Nauka, Moscow, (1966).

Kopytko. B., Zhernovyi.K., Stationary characteristics of $M^{\theta}/M/1$ queue with swtching of service modes, *Scientific Research Of The Institute Of Mathematics And Computer Science*, 2 (10), 117-128, (2011).

Krishna Kumar.B., Parthasarathy. P.R., Sharafali.M. ,Transient solution of an $M/M/1$ queue with balking, *Queueing Systems* ,13, 441-448, (1993).

Krishnamoorthy .A. , Sreenivasan .C., An $M/M/2$ queueing system with heterogeneous servers including one with working vacation, *International Journal of Stochastic Analysis* 1-16, (2012).

Kumar .B.K., Arivudainambi .D., Vijaykumar .A., An $M/G/1/1$ queue with unreliable server and no waiting capacity, *International Journal of Information and Management Science*, 13(2), 35-50, (2002).

Kumar .B.K., Madheswari .S.P., An $M/M/2$ queuing system with heterogeneous server and multiple vacations, *Mathematical and Computer Modelling*, 41, 1415-1429, (2005).

Kumar.B.K., Madheswari.S.P., Venkatakrishnan.K.S., Transient solution of an $M/M/2$ queue with heterogeneous servers subject to catastrophes, *Information and Management Sciences*,

63-80, (2007).

Kumar .B.K., Madheswari .S.P., Transient behavior of the M/M/2 queue with catastrophes, *Statistica*, 62, 129-136, (2002).

Kumar .R., Sharma .S.K., A Markovian feedback queue with retention of reneged customers and balking, *AMO-Advanced Modeling and Optimization*, 14(3), (2012).

Kumar.R., Sharma.S.K., A Single-server markovian queuing system with discouraged arrivals and retention of reneged customers, *Yugoslav Journal of Operations Research*, 24,119-126, (2014).

Kumar .S., Taneja .G., A feedback queuing model with chances of providing service at most twice by one or more out of three Servers, *Aryabhatta Journal of Mathematics and Informatics*, 9 (1), 171-184, (2017).

Legros .B., Jouini .O., Routing in a queueing system with two heterogeneous servers in speed and in quality of resolution, *Stochastic models*, 1532-4214 , (2017).

Li .H.,Zhu .Y., Analysis of M/G/1 queue with delayed vacation and exhaustive service discipline, *European Journal of Operation Research*, 92(1), 125-134, (1996).

Lindley .D.V., The theory of queues with a single server, *Proc. Camb.Phil, Soc*, 48, 277-289, (1952).

Liu .L., Kulkarni.V., Busy period analysis for M/PH/1 queues with workload dependent balking, *Queueing Systems*, 59, 37-51, (2008).

Liu .L., Kulkarni .V., Explicit solutions for the steady state distributions in M/PH/1 queues with workload dependent balking, *Queueing Systems*, 52, 251-260, (2006).

Liu .W., Kumar .P., Optimal control of a queueing system with two heterogeneous servers, *IEEE Transactions on Automatic Control*, 29, 696-703, (1984).

Little .J.D., A proof of the queueing formula $L = \lambda W$, *Oper. Res*, 9, 383-387, (1961).

Lozano .M., Moreno .P., A discrete time single-server queue with balking : economic applications, *Applied Economics*, 40, 735-748, (2008).

Ma .Y., Liu .W.Q., , Li .J.H., Equilibrium balking behavior in the *Geo/Geo/1* queueing system with multiple vacations, *Applied Mathematical Modelling*, 37, 3861-3878, (2013).

Madan .K.C., Abu-Dayyeh .W., Taiyyan .F., A two server queue with Bernoulli schedules and a single vacation policy, *Applied Mathematics and Computation*, 145 , 59-71, (2003).

Marshall. K.T., Wolff. R.W., Customer average and time average queue lengths and waiting times, *Journal Of Applied Probability*, 8(3), 535-542, (1971) .

Melikov .A.Z., Ponomarenko .L.A., Kuliyeva .Kh.N., Numerical analysis of a queueing system with feedback, *Cybernetics and Systems Analysis*, 51, 4, (2015).

Morse .P.M., Queues inventories and maintenance, *Operations Research Society Of America*, 359-520, (1958).

Nasrallah .W.F., How pre-emptive priority affects completion rate in an M/M/1 queue with poisson reneging, *European Journal of Operational Research*, 193, 317-320, (2009).

Neuts .M.F., Takahashi .Y., Asymptotic behavior of the stationary distribution in the GI/PH/c queue with heterogeneous servers, 57, 441-452, (1981).

O'Dell .G.F., Theoretical principles of the traffic capacity of automatic switches, *P.O. Elect. Engrs'J*, 13, 23-209, (1920).

O'Dell.G.F., Gibson.W.W., Automatic trunking in theory and practice. *Inst.P.O.Elect.Engrs*,41, 107, (1926).

Omarah .A.S.R., A transient solution of queues with variable channel considering balking concept when $S = 2$ and $[\sigma] = 1$, *Applied Mathematics and Computation*, 174, 337- 344,

(2006).

Padmavathy .R., Kalidass .K., Ramanath .K., Vacation queues with impatient customers and a waiting server, *International Journal of Latest Trends in Engineering and Technology*, 1(1), 10-19, (2011).

Palm.C., Etude des delais d'attente, *Ericsson Technics*, 5 (2), 39-56, (1936).

Palm.C., Analysis of the Erlang traffic formula for busy signal arrangements, *Ericsson Technics*, 6, 39-58, (1938) .

Perel .N. , Yechiali .U., Queues with slow servers and impatient customers, *European Journal of Operational Research*, 201, 247-258, (2010).

Pollaczek.F., *Theorie des wartens*, 19, 71-78, (1930).

Pollaczek.F., Sur quelques lois asymptotiques de la theorie de l'encombrement des reseaux telephoniques, *Ann.Univ.Lyon*, 5, 21-35, (1942).

Rajan .V., Queuing analysis of markovian queue Having two heterogeneous servers with catastrophes using matrix geometric technique, *International Journal of Statistics and Systems*, 12, 205-212, (2017).

Rao .S.S., Queueing models with balking, renegeing and interruptions, *Operations Research*, 13, 596, (1965).

Renyi .A. , *Probability theory (in Hungarian)*, Tankonyvkiado, Budapest, (1973).

Rykov .V., Monotone control of queueing systems with heterogeneous servers, *Queueing Sys*, 37(4), 391-403, (2001).

Rykov .V.V., Efrosinin .D.V., On the slow server problem, *Automa. Remote Cont*, 70(12), 2013-2023, (2009).

Saaty .T.L., *Elements of queueing theory with applications*, McGraw-Hill, (1961).

Saglam .V.,Torun .H., On optimization of stochastic service system with two heterogeneous channels, *International Journal of Applied Mathematics*, 17(1), 1-6,(2005).

Saglam .V., Shahbazov .A., Minimizing loss probability in queueing systems with heterogeneous servers, *Iranian Journal of Science and Technology*, 31(2), 199-206,(2007).

Salehirad .M.R., Badamchizadeh .A., On the multi-phase $M/G/1$ queueing system with random feedback, *Springer-Verlag*, 17, 131-139, (2009).

Santhakumaran .A., Thangaraj .V., A Single server queue with impatient and feedback customers, *Information and Management Science*, 11, 71-79, (2000).

Serhan Ziya .H.A., Foley .R.D, Pekoz .E., A monotonicity result for a $G/GI/c$ queue with balking or renegeing, *J. Appl. Probab.*, 43, 1201-1205, (2006).

Sharma .D. P., Dass .S., Multiserver markovian queues with finite waiting space, *Sankhya*, B-50, 328-331, (1998).

Sharma .S.K., Kumar .R., A Markovian feedback queue with retention of renegeed customers and balking, *AMO-Advanced Modeling and Optimization*, 14(3), 681-688, (2012).

Sherif .I.A., Transient analysis of a two-heterogeneous servers queue with impatient behavior, *Journal of the Egyptian Mathematical Society*, 22, 90-95, (2014).

Shin .Y.W., Choo .T.S., $M/M/s$ queue with impatient customers and retrials, *Applied Mathematical Modelling*, 33, 2596-2606, (2009).

Singh .V.P., Two servers Markovian queues with balking, Heterogeneous Vs Homogeneous servers, *Oper. Res*, 18(1), 145 -159, (1970).

Singh .V.P., Two-server markovian queues with balking : heterogeneous vs. homogeneous servers, *Operations Research*, 18, 145-159, (1970).

Smith .W.E., Various optimizers for single-stage production, *Naval Research Logistics*

Quarterly, 3(1-2), 59-66, (1956).

Smith .W.L., On the distribution of queueing times, *proc.Soc*, 49, 449-461, (1953).

Som .B.K., Decision making uncertain environment-A queueing theory approach international journal of advanced engineering, Management and Science (IJAEMS), 2(6), (2016).

Sridhar .A., Allah Pitcha .R., Analyses of a Markovian queue with two heterogeneous servers and working vacation, *International Journal of Applied Operational Research*, 5, 1-15, (2015).

Stoyan .D., Comparison methods of queue and other stochastic models, Wileys, new york, (1983).

Subba Rao .S., Queueing with balking and reneging in M/G/1 systems, *Metrika*, 12, 173-188, (1967) .

Taha .H.A., Operation research : an Introduction, 8th edn, New Jersey, Pearson Prentice Hall, (2007).

Takac's .I., Investigation of waiting time problems by reduction to markov process, *Acta Math. Acad.Scie.Hungar.*, 101-129, (1955).

Takac's .L., A Single server queue with feedback, *The Bell System Tech. Journal*, 42, 134-149, (1963).

Takagi .H., Waiting time in the $M/M/m/(m+c)$ queue with impatient customers, *International Journal of Pure and Applied Mathematics*, 90, 519-559, (2014).

Thangaraj .V., Vanitha .S., On the Analysis of M/M/1 feedback queue with catastrophes using Continued fractions, *International Journal of Pure and Applied Mathematics*, 53, 131-151, (2009).

VanDantzig .G.B., Wolfe .P., Decomposition principle for linear programs, *Operations Research*, 8(1), 101-111, (1960).

Van Tits .M. H. L., van der Veecken .H. J. M., Simulation of a queueing problem with balking, *Simulation*, 35, 88-93, (1980).

Vaulot .E., Application du calcul des probabilités a l'exploitation telephonique, *Annales des Postes, Telegraphes et Telephones* , 14, 136-156,(1925).

Ward .A.R., Glynn .P.W., A diffusion approximation for a GI/GI/1 queue with balking or reneging, *Queueing Systems*, 50, 371-400, (2005).

Wilkinson .R.I., The interconnection of telephone systems, *The Bell System Technical Journal*, 10, 531-564, (1931).

Wu .C.H, Ke .J.C., Computational algorithm and parameter optimization for a multi server system with unreliable servers and impatient customers, *Journal of Computational and Applied Mathematics*, 235, 547-562, (2010).

Xiong .W., Altiok .T., An approximation for multi-server queues with deterministic reneging times, *Annals of Operations Research*, 172, 143-151, (2009).

Yechiali .U., On optimal balking rules and toll charges in the GI/M/1 queueing process, *Operations Research*, 19, 349-370, (1971).

Yechiali .U., Queues with system disasters and impatient customers when system is down, *Queueing Systems*, 56, 195-202, (2007).

Yue .D., Yue .W., A heterogeneous two-server network system with balking and a Bernoulli vacation schedule, in : *Proceedings of the 4th International Conference on Queueing Theory and Network Applications*, Article No. 20, (2009).

Yue .D., Zhang .Y., Yue .W., Optimal performance analysis of an M/M/1/N queue system with balking, reneging and server vacation, *International Journal of Pure and Applied Mathematics*, 28 (1), 1001-115, (2006).

Yue .D., Yue .W., Saffer .Z., Chen .X., Analysis of an $M/M/1$ queueing system with impatient customers and a variant of multiple vacation policy, *Journal of Industrial and Management Optimization*, 10 (1), 89-112, (2014).

Yue .D., Yue .W., Zhao .G., Analysis of an $M/M/1$ queue with vacations and impatience timers which depend on the server's states, *Journal of Industrial and Management Optimization*, 12 (2), 653-666, (2016).

Zeltyn .S., Mandelbaum .A., Call Centers with Impatient Customers : Many-Server Asymptotics of the $M/M/n + G$ Queue, *Queueing Systems*, 51, 361-402, (2005).

Zohar .E., Mandelbaum .A., Shimkin .N., Adaptive behavior of impatient customers in tele-queues : theory and empirical support, *Management Science*, 48, 566-583, (2002).