



SCIENCES

MATHÉMATIQUES

Théorie des files d'attente et applications

Théorie des files d'attente 1

tendances avancées

sous la direction de

Vladimir Anisimov

Nikolaos Limnios

Théorie des files d'attente 1

First published 2021 in Great Britain by ISTE Editions Ltd.

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms and licenses issued by the CLA. Enquiries concerning reproduction outside these terms should be sent to the publishers at the undermentioned address:

ISTE Editions Ltd
27-37 St George's Road
London SW19 4EU
UK

© ISTE Editions Ltd 2021

The rights of the authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

British Library Cataloguing-in-Publication Data

A CIP record for this book is available from the British Library

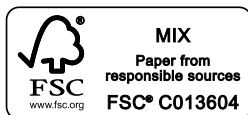
ISBN: 978-1-78948-001-6 (print)

ISBN: 978-1-78949-001-5 (e-book)

ERC code:

PE1 Mathematics

PE1_21 Application of mathematics in industry and society



Printed and bound in Great Britain by CPI Group (UK) Ltd., Croydon, Surrey CR0 4YY, January 2021

Exemplaire réservé à Janos Sztrik

Encyclopédie SCIENCES

Mathématiques, domaine dirigé par Nikolaos Limnios

Théorie des files d'attente et applications, thème dirigé
par Vladimir Anisimov

Théorie des files d'attente 1

tendances avancées

sous la direction de
Vladimir Anisimov
Nikolaos Limnios

ISTE
editions

Résultats récents en files d'attente avec rappels à sources finies avec collisions

Anatoly NAZAROV¹, János SZTRIK² et Anna KVACH¹

¹Université d'État de Tomsk, Tomsk, Russie

²Université de Debrecen, Debrecen, Hongrie

L'objectif de ce chapitre est de donner un aperçu des résultats récents sur les systèmes de files d'attente avec rappel à source finie à serveur unique avec collision de clients. Il s'agit notamment de déterminer lorsque le serveur est fiable et de modéliser les pannes et les réparations aléatoires du serveur selon qu'il est inactif ou occupé. Les méthodes numériques, de simulation et asymptotiques, assistées par des outils, sont considérées sous la condition d'un nombre croissant et illimité de sources. Plusieurs cas et exemples sont traités et les résultats des différentes approches sont comparés les uns aux autres pour montrer les avantages et les inconvénients de la méthode donnée. En général, nous prouvons que la distribution à l'état d'équilibre du nombre de clients dans le système de service peut être approchée par une distribution normale avec une moyenne et une variance données. En utilisant des méthodes asymptotiques dans certaines conditions en régime permanent, la distribution du temps de séjour en orbite et dans le système peut être approchée par une distribution exponentielle généralisée. En outre, il est prouvé que la distribution du nombre de nouveaux rappels jusqu'au succès du service dans la limite est distribuée géométriquement. À l'aide d'une simulation stochastique, plusieurs systèmes sont analysés, montrant la direction à suivre pour d'autres investigations analytiques. Des tableaux et des figures sont présentés pour illustrer certaines caractéristiques particulières de ces systèmes.

8.1. Introduction

Les files d'attente à rappel à source finie sont des systèmes stochastiques très utiles et efficaces pour modéliser plusieurs problèmes survenant dans les systèmes de commutation téléphonique et les réseaux de télécommunication, y compris les réseaux cellulaires, les réseaux informatiques locaux CSMA (écoute d'un support à accès multiple), les centres d'appels et les réseaux maillés sans fil CSMA à l'échelle de la trame. Pour en mesurer l'importance, le lecteur intéressé est invité à consulter les ouvrages et les références cités ci-après : (Artalejo et Corral 2008 ; Falin et Artalejo 1998 ; Fiems et Phung-Duc 2017 ; Gómez-Corral et Phung-Duc 2016 ; Kim et Kim 2016).

Après avoir recherché dans les bases de données scientifiques, nous avons remarqué qu'un nombre relativement restreint d'articles seulement ont été consacrés aux systèmes de files d'attente lorsque les appels entrants (primaires ou secondaires) causent des collisions à la demande en service et que les deux vont en orbite (voir par exemple, (Ali et Wei 2015 ; Choi *et al.* 1992 ; Kim 2010 ; Kumar *et al.* 2010 ; Lakaour *et al.* 2018 ; Peng *et al.* 2014 ; Takeda et Yoshihiro 2017)).

Dans les systèmes CSMA réels, les collisions sont inévitables et elles diminuent l'efficacité des performances du système, c'est pourquoi de nouveaux protocoles devraient être développés pour les éviter (voir (Cao *et al.* 2018 ; Jinsoo *et al.* 2018 ; Kwak *et al.* 2018 ; Wentink 2017 ; Yeo *et al.* 2017)). Reith (2017) et Takeda et Yoshihiro (2017) ont modélisé des situations réelles où des collisions peuvent se produire.

La modélisation stochastique des systèmes avec collisions est nécessaire non seulement d'un point de vue technique, mais c'est un défi mathématique puisqu'elle nécessite des approches plus sophistiquées.

Nazarov et son groupe de recherche ont mis au point une méthode asymptotique très efficace (Nazarov et Moiseeva 2006) à l'aide de laquelle divers systèmes ont été examinés. En ce qui concerne les systèmes à rappel à source finie avec collisions, il convient de mentionner les documents suivants : (Kvach et Nazarov 2015b ; Kvach 2014 ; Kvach et Nazarov 2015a, 2015c ; Nazarov *et al.* 2014).

Sztrik et son groupe de recherche se sont dédiés aux systèmes dont les serveurs ne sont pas fiables, comme on peut le voir dans (Almási *et al.* 2005 ; Sztrik 2005 ; Sztrik *et al.* 2006 ; Wüchner *et al.* 2010) ; c'est pourquoi on comprend que les deux groupes de recherche aient commencé leur coopération en 2017.

Nos recherches ont été basées sur des approches analytiques, numériques, de simulation et asymptotiques telles que traitées dans (Anisimov et Sztrik 1989 ;

Anisimov et Artalejo 2001 ; Anisimov 1999 ; Artalejo et Corral 2008 ; Bhat 2015 ; Bossel 2013 ; Falin et Templeton 1997 ; Harchol-Balter 2013 ; Kobayashi et Mark 2009 ; Kulkarni 2016 ; Lakatos *et al.* 2013 ; Law et Kelton 1991 ; Nazarov et Terpugov 2004 ; Nazarov et Moiseeva 2006 ; Rubinstein et Kroese 2016 ; Stewart 2009 ; Yao 2016).

L'objectif principal de ce chapitre est de donner un aperçu des résultats récemment obtenus dans ce champ au moyen de différentes méthodes. Ce faisant, nous avons essayé d'unifier la notation qui est apparue dans différentes publications et d'utiliser la notation standard des articles de style occidental, qui diffère de celle des articles de style russe. Nous sommes confiants dans le fait que nos modèles avec collision peuvent être utilisés dans des situations réelles pour décrire, par exemple, le fonctionnement de systèmes à accès aléatoire traités dans (Fiems et Phung-Duc 2017) sans collision ; les réseaux maillés sans fil basés qui utilisent CSMA au niveau des trames et des paquets examinés dans (Takeda et Yoshihiro 2017) ; les systèmes avec erreurs de transmission analysés dans (Lakaour *et al.* 2018). Lorsqu'une collision se produit, les signaux sont superposés, les paquets sont déformés et doivent donc être retransmis : tout doit être renvoyé en orbite.

Nous reconnaissons que l'approche du trafic très peu dense appliquée dans (Fiems et Phung-Duc 2017) pour obtenir la distribution du nombre de clients dans la structure de service pour les systèmes sans collision pourrait être intégrée pour étudier les systèmes avec collision. Toutefois, nous devons souligner qu'outre ces derniers, nous sommes en mesure d'obtenir une approximation pour la distribution d'un certain nombre de nouveaux essais et pour la distribution du temps de réponse/attente d'un client. Ces mesures n'ont pas été traitées dans le document susmentionné. De plus, dans nos différents cas, le temps de service est généralement réparti et le serveur n'est pas fiable.

Du point de vue théorique, l'une des principales difficultés par rapport aux systèmes sans collision est que le processus de service est interrompu plusieurs fois par les collisions jusqu'à son achèvement. Les pannes de serveur non fiables interrompent également le service, ce qui entraîne une structure de temps de service total assez compliquée. En outre, en cas de collision, le nombre de clients en orbite augmente de deux. L'utilisation d'une approche algorithmique signifie que la $n^{\text{ième}}$ itération ne dépend pas seulement de l'itération précédente mais des deux itérations précédentes. Notre contribution la plus importante est qu'en plus d'obtenir la distribution asymptotique du nombre de clients en orbite, la distribution asymptotique du nombre de rappels et la distribution du temps de réponse/attente d'un appel sont déterminées dans certaines conditions.

Les nouveaux aperçus numériques pour les modèles avec collisions et serveur peu fiable sont obtenus à l'aide d'approches algorithmiques, numériques (MOSEL),

de simulation. La combinaison de ces méthodes nous permet de valider l'exactitude de la méthode asymptotique. Nous montrons qu'avec un certain paramétrage, la propriété caractéristique des files d'attente avec rappels à source finie ayant un maximum de temps d'attente moyen pour un taux d'arrivée croissant reste valable même en cas de collision. Cela signifie que le maximum est dû à la source finie et non à la collision. Cette caractéristique surprenante a été remarquée par plusieurs auteurs et expliquée de différentes façons. Cependant, dans les systèmes avec collisions, nous devons faire face à la durée totale du service, qui est souvent interrompu par des collisions et des pannes. Ceci est dû à la collision et le meilleur temps de service est la dernière phase de ces mesures. D'après nos expériences, la moyenne du temps de service total et du meilleur temps dépend fortement de la distribution du temps de service et le minimum ne peut être obtenu qu'avec un paramétrage très spécial.

Ce chapitre est une version agrandie et modifiée de notre récent article (Nazarov *et al.* 2018).

Le reste du document est organisé comme suit. La section 8.2 donne une description du modèle et définit le processus non markovien multidimensionnel correspondant.

Dans les sections 8.3 et 8.4, les systèmes ayant un serveur fiable et un serveur non fiable sont traités respectivement. Dans ces sections, les modèles ayant des temps de service exponentiellement et généralement distribués sont étudiés puis analysés au moyen de méthodes basées sur des outils, algorithmiques, de simulation et asymptotiques, respectivement. Les principaux résultats des articles sont rassemblés et plusieurs figures illustrent les caractéristiques les plus intéressantes du système donné.

Enfin, le présent chapitre se termine par une conclusion et met en lumière certains projets futurs.

8.2. Description du modèle et notations

Dans ce qui suit, nous présentons le modèle dans sa forme la plus générale telle qu'elle a été traitée à l'aide de méthodes numériques et asymptotiques.

Considérons un système de file d'attente de type $M/GI/1//N$ avec collision de clients et un serveur peu fiable (figure 8.1). Le nombre de sources est N et chacune d'entre elles peut générer une requête primaire pendant un temps exponentiellement distribué avec le taux λ/N . Une source ne peut pas générer un nouvel appel avant la fin du service complet de ce client.

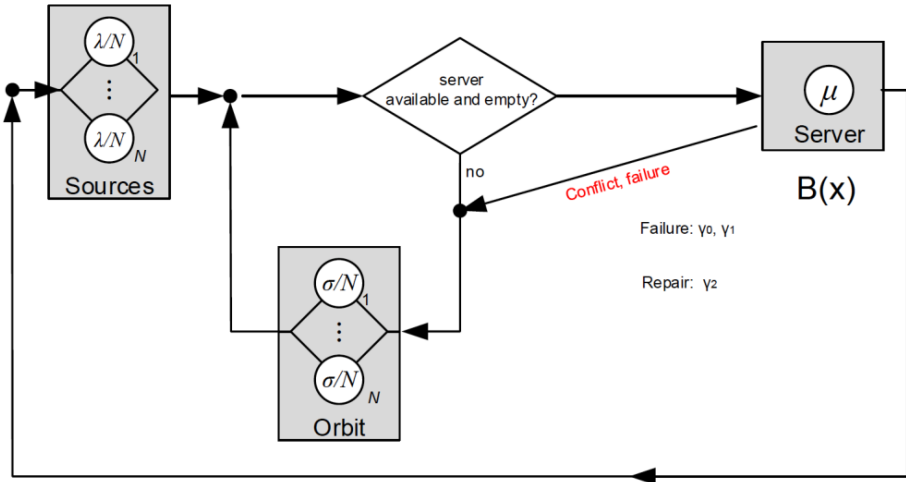


Figure 8.1. Système de file d'attente de type $M/GI/1//N$ avec collisions de clients et serveur peu fiable

Si lorsqu'une requête primaire atteint le serveur, celui-ci est inactif, alors le serveur entre en service immédiatement ; dans ce cas, le temps de service requis a une fonction de distribution de probabilité $B(x)$. Désignons sa fonction de taux de service par $\mu(y) = B'(y) (1 - B(y))^{-1}$ et sa transformée de Laplace-Stieltjes par $B^*(y)$, respectivement. Si le serveur est occupé, un client arrivant (primaire ou répété) entre en collision avec un client en service et les deux sont placés en orbite. Les temps d'interception des clients sont censés être répartis de façon exponentielle avec le taux σ/N , c'est-à-dire que cela ne dépend pas du nombre de fois qu'il a été bloqué. Nous supposons que le serveur n'est pas fiable, c'est-à-dire que sa durée de vie est censée être distribuée de façon exponentielle avec un taux de défaillance γ_0 si le serveur est inactif et un taux de défaillance γ_1 s'il est occupé. Lorsque le serveur tombe en panne, il est immédiatement envoyé pour se faire réparer et le temps de réparation est supposé être réparti de façon exponentielle avec le taux γ_2 . Dans le cas où le serveur est en panne, toutes les sources continuent de générer des clients et les envoient en orbite ; de même, les clients peuvent réessayer de l'orbite au serveur mais tous les clients qui arrivent sont immédiatement placés en orbite. De plus, dans ce modèle peu fiable, nous supposons que la requête interrompue va immédiatement en orbite et que son service suivant est indépendant de celui qui est interrompu. Ceci bien sûr, dans le cas d'un serveur fiable $\gamma_0 = \gamma_1 = 0$. Toutes les variables aléatoires

impliquées dans la construction du modèle sont supposées être indépendantes les unes des autres.

Soit $J(t)$ le nombre de clients dans le système au temps t , c'est-à-dire le nombre total de clients en orbite et en service. De même, soit $K(t)$ l'état du serveur au temps t , c'est-à-dire :

$$K(t) = \begin{cases} 0, & \text{si le serveur est inoccupé} \\ 1, & \text{si le serveur est occupé} \\ 2, & \text{si le serveur est en panne (en réparation)} \end{cases}$$

Ainsi, nous allons étudier le processus $\{K(t), J(t)\}$, qui n'est pas un processus de Markov à moins que le temps de service ne soit distribué de façon exponentielle. Pour être un processus de Markov, nous devons utiliser la méthode des variables supplémentaires, à savoir, nous allons considérer deux variantes : la méthode du temps de service résiduel et la méthode du temps de service écoulé en fonction du but de l'étude.

Désignons par $Y(t)$ le processus aléatoire supplémentaire égal au temps de service écoulé du client jusqu'au moment t et par $Z(t)$ le temps de service résiduel, c'est-à-dire l'intervalle de temps à partir du moment jusqu'à la fin du service réussi du client, respectivement. Il est évident que $\{K(t), J(t), Y(t)\}$ et $\{K(t), J(t), Z(t)\}$ sont des processus de Markov. Notons que $Y(t)$ et $Z(t)$ ne sont définis que dans les moments où le serveur est occupé, c'est-à-dire lorsque $K(t) = 1$.

Définissons les probabilités stationnaires comme suit :

$$\begin{aligned} P_0(j) &= P\{K = 0, J = j\} \\ P_1(j, y) &= P\{K = 1, J = j, Y < y\} \\ P_1(j, z) &= P\{K = 1, J = j, Z < z\} \\ P_2(j) &= P\{K = 2, J = j\} \end{aligned}$$

Bien entendu, dans le cas d'un temps de service exponentiellement distribué, les probabilités d'équilibre sont notées comme suit :

$$P_k(j) = P\{K = k, J = j\}, \quad k = 0, 1, 2, \quad j = 0, \dots, N$$

La distribution en régime permanent de l'état du serveur est indiquée par :

$$R_k = P(K = k), k = 0, 1, 2$$

et la répartition du nombre de clients dans le système est désignée par :

$$P(j) = P(J = j), j = 0, \dots, N$$

Il est clair que dans le cas d'un serveur fiable, toutes les probabilités où $K = 2$ sont 0.

L'objectif principal des études est d'obtenir ces distributions et d'autres mesures de performance des systèmes, telles que la distribution du temps de séjour dans le système, la distribution du temps total de service et la distribution du nombre de nouveaux essais. Il s'agit de problèmes très complexes et, à la connaissance des auteurs, il n'existe pas de formules analytiques précises pour trouver des solutions. C'est la raison pour laquelle nous avons essayé d'obtenir les caractéristiques des différents systèmes à l'aide de méthodes de simulation stochastique, algorithmique et asymptotique.

Puisque, dans ce qui suit, nous allons utiliser la distribution Gamma pour montrer les caractéristiques spéciales de certains systèmes, nous l'introduisons ici pour être sûrs de la forme car elle peut être définie de plusieurs façons.

On dit d'une variable aléatoire X qu'elle a une [distribution Gamma](#) avec des paramètres (α, β) si sa fonction de densité est donnée par :

$$f(x) = \begin{cases} 0 & , \text{si } x < 0 \\ \frac{\beta(\beta x)^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} & , \text{si } x \geq 0 \end{cases}$$

où $\alpha > 0$, $\beta > 0$, et :

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$$

est la [fonction Gamma](#).

Sa fonction de distribution ne peut être obtenue sous une forme explicite, à l'exception de $\alpha = n$. Dans ce cas, elle se réduit à la distribution d'Erlang.

On note que, selon les paramètres, elle peut prendre de petites valeurs avec une probabilité élevée, c'est-à-dire lorsque le paramètre de forme est faible, dans ce cas le coefficient de variation au carré C_X^2 est élevé. Sa fonction de densité a la forme présentée en figure 8.2.

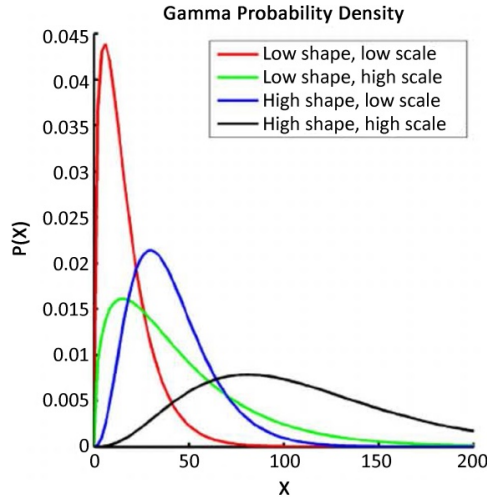


Figure 8.2. Fonction de densité de la distribution Gamma

On peut démontrer que :

$$E(X) = \frac{\alpha}{\beta}, \quad Var(X) = \frac{\alpha}{\beta^2}, \quad C_X^2 = \frac{1}{\alpha}$$

α et β sont appelés respectivement *paramètre de forme* et *paramètre d'échelle*, et peuvent être exprimés par la moyenne et la variance de la manière suivante si l'on souhaite estimer les paramètres de l'échantillon :

$$\alpha = \frac{(E(X))^2}{Var(X)}, \quad \beta = \frac{E(X)}{Var(X)}$$

On peut montrer que sa transformée de Laplace a la forme :

$$L_X(s) = \left(\frac{\beta}{\beta + s} \right)^\alpha = \left(1 + \frac{s}{\beta} \right)^{-\alpha}$$

8.3. Systèmes avec serveur fiable

8.3.1. Systèmes M/M/1

8.3.1.1. Approche algorithmique

Dans des articles précédents (Kvach 2014 ; Nazarov *et al.* 2014), on a dérivé les équations de Kolmogorov en régime permanent et la distribution de l'état du système a été obtenue par une approche algorithmique. Puis, la distribution du nombre de clients dans le système a été calculée et utilisée pour valider les résultats asymptotiques.

8.3.1.2. Approche asymptotique

La principale contribution de Nazarov *et al.* (2014) est que la fonction caractéristique à l'état d'équilibre de la distribution prélimite du nombre de clients dans le système peut être approchée par la fonction caractéristique asymptotique du troisième ordre suivante, à savoir :

$$\text{Eexp}\{iuJ\} \approx \exp \left\{ iu\kappa_1 N + \frac{(iu)^2}{2} \kappa_2 N + \frac{(iu)^3}{3!} \kappa_3 N \right\}$$

où la distribution de l'état du serveur est explicitement donnée sous la forme suivante :

$$R_1 = \frac{\sigma(2\lambda + \mu) - \sqrt{\sigma^2(2\lambda - \mu)^2 + 8\sigma\mu\lambda^2}}{4\mu(\sigma - \lambda)}$$

$$R_0 = 1 - \frac{\sigma(2\lambda + \mu) - \sqrt{\sigma^2(2\lambda - \mu)^2 + 8\sigma\mu\lambda^2}}{4\mu(\sigma - \lambda)}$$

et les cumulants (semi-invariants) (Nazarov et Sudyko 2010) sont :

$$\kappa_1 = \frac{2\mu R_1^2}{\sigma(1 - 2R_1)}, \quad \kappa_2 = \mu R_1 \frac{1 + (R_1 - R_0)R_0}{\lambda - (\lambda - \sigma)(R_1 - R_0)^2}$$

$$a = R_0 R_1 (R_1 - R_0)$$

$$\kappa_3 = 2 \cdot \left\{ \frac{\left[(\lambda - \sigma)(R_0 - R_1)^2 \kappa_2 + \mu a \right] \cdot \left[\frac{1}{2} - ((R_0 - R_1)(\sigma - \lambda) + \lambda) \frac{\kappa_2}{\mu R_1} \right]}{\lambda + (\sigma - \lambda)(R_0 - R_1)^2} \right. \\ \left. + \frac{\lambda \kappa_2 \cdot \left[\frac{a}{R_0} + \frac{1}{2} \right] + \mu \left[\frac{1}{R_0} - \frac{a}{2} \right]}{\lambda + (\sigma - \lambda)(R_0 - R_1)^2} \right\}$$

En fait, κ_1 est la solution positive de l'équation :

$$\lambda(1 - \kappa_1) = \mu R_1(\kappa_1)$$

et numériquement, il est facile de vérifier que :

$$R_0(\kappa_1) = \frac{\delta(\kappa_1) + \mu}{2\delta(\kappa_1) + \mu}$$

$$R_1(\kappa_1) = \frac{\delta(\kappa_1)}{2\delta(\kappa_1) + \mu}$$

où $\delta(\kappa_1) = \lambda(1 - \kappa_1) + \sigma\kappa_1$.

On peut facilement voir que l'approximation du second ordre aboutit à une distribution normale. Dans ce cas, dénotons par $G(x)$ la fonction de distribution de la distribution gaussienne avec la moyenne $N\kappa_1$ et variance $N\kappa_2$.

De plus, dénotons par $P_{as}(j)$ la distribution asymptotique discrète obtenue à l'aide de l'approximation gaussienne, à savoir :

$$P_{as}(j) = \{G(j+0,5) - G(j-0,5)\} [G(N+0,5) - G(-0,5)]^{-1}$$

$$j = 0, \dots, N \quad [8.1]$$

que nous appellerons approximation gaussienne de la distribution discrète prélimite $P(j)$.

Dans ce qui suit, nous montrons comment nous pouvons obtenir la distribution à l'aide de la fonction caractéristique inverse (voir (Nazarov et Sudyko 2010)).

Soit $D_m(j)$ une distribution de probabilité avec $j = 0, 1, 2, \dots, N$ et $h_m(u)$ sa fonction caractéristique, c'est-à-dire :

$$h_m(u) = \sum_{j=0}^N e^{iuj} D_m(j)$$

où $i = \sqrt{-1}$ est l'unité imaginaire et $h_m(u)$ la fonction caractéristique asymptotique du $m^{\text{ième}}$ ordre. Ensuite, en utilisant la transformation inverse $D_m(j)$ est donnée par la formule :

$$D_m(j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iuj} h_m(u) du$$

$$\text{Il est évident que : } \sum_{j=0}^N D_m(j) = h_m(0) = 1.$$

Si les valeurs de $D_m(j)$ sont réelles et non négatives pour tous $j = 0, \dots, N$, alors :

$$D_m(j) = P_m(j)$$

où $P_m(j)$ s'appelle l'approximation du $m^{\text{ième}}$ ordre de la distribution prélimite $P(j)$.

Comme $D_m(j)$ est calculé par intégration numérique, parmi les nombres $D_m(j)$ il peut apparaître des nombres complexes avec des parties imaginaires assez petites ou des nombres négatifs réels avec des valeurs absolues assez petites.

Dans ce cas, nous définissons les nombres réels non négatifs $g_m(j)$ par la formule suivante :

$$g_m(j) = \frac{1}{4} \{D_m(j) + \overline{D_m(j)} + |D_m(j) + \overline{D_m(j)}|\}$$

où la barre au-dessus indique une conjugaison complexe. Définissons ensuite l'approximation de la distribution prélimite de la distribution limite par l'égalité :

$$P_m(j) = g_m(j) / \sum_{l=0}^N g_m(l)$$

Pour comparer les distributions de probabilités $P(j)$ et $P_m(j)$, nous utiliserons la distance de Kolmogorov définie comme suit :

$$\Delta_m = \max_{0 \leq n \leq N} \left| \sum_{j=0}^n P(j) - \sum_{j=0}^n P_m(j) \right|$$

Par cette distance, nous pouvons estimer l'erreur d'approximation du $m^{\text{ième}}$ ordre de la distribution prélimite.

Dans un article précédent (Nazarov *et al.* 2014), les approximations de deuxième et troisième ordre de la distribution prélimite ont été comparées à la distribution exacte obtenue par la méthode algorithmique. L'applicabilité de la méthode asymptotique a été validée à l'aide d'une configuration différente de paramètres et pour différents paramètres N , et certaines conclusions ont été tirées.

Un problème plus compliqué, à savoir la répartition du temps de séjour dans la structure de service, a été étudié par Kvach et Nazarov (2015b) à l'aide de méthodes asymptotiques lorsque N tend vers l'infini. Il a été prouvé que la fonction caractéristique du temps de séjour T d'un client dans le système de service peut être approchée de la manière suivante :

$$\text{Eexp}\{iuT\} \approx q + (1-q) \frac{\sigma q / N}{\sigma q / N - iu}$$

$$\text{où } q = \frac{\mu R_0}{\delta + \mu}.$$

Il est alors facile de montrer que l'approximation de la fonction de distribution prélimite du temps de séjour peut s'écrire :

$$A(x) \approx 1 - (1-q)e^{-\frac{\sigma q}{N}x}$$

ce qui signifie que le temps de séjour asymptotique peut être nul avec probabilité q et qu'il est distribué de manière exponentielle avec une probabilité $1 - q$. En conséquence, le temps de séjour moyen :

$$E(T) \approx (1-q)N / (\sigma q)$$

À partir du raisonnement probabiliste :

$$\lambda(1 - \kappa_1) = \mu R_1$$

ce qui signifie que le taux d'arrivée moyen est égal au taux de départ moyen et de la formule de Little :

$$\mu R_1(1-q) = \sigma q \kappa_1$$

devrait être valide. Ces formules peuvent être utilisées pour la validation des résultats numériques et asymptotiques obtenus. Les exemples d'échantillons asymptotiques concernant la distribution du temps de séjour n'ont pas été validés par simulation dans cet article.

8.3.2. Système M/GI/1

Cette section traite des résultats lorsque les temps de service requis sont généralement distribués, mais dans les exemples, la distribution Gamma est utilisée en raison de ses propriétés utiles. En effet, on voit facilement que son coefficient de variation au carré peut être inférieur, égal ou supérieur à 1 dépendant des valeurs des paramètres de forme et d'échelle.

8.3.2.1. Approche algorithmique

Une étude précédente (Kvach et Nazarov 2015a) traite de l'approche algorithmique de la façon d'obtenir la distribution en régime permanent du système. La méthode de la technique des variables supplémentaires avec temps de service résiduels a été appliquée et plusieurs exemples numériques ont été traités avec Gamma temps de service distribués. Les résultats ont aidé à valider les résultats asymptotiques pour le même modèle.

8.3.2.2. Simulation stochastique

Les articles de Nazarov *et al.* (2017b, 2017c) sont consacrés à l'analyse asymptotique du temps de service total moyen, de la distribution du temps de séjour dans le système et de la distribution du nombre de nouveaux rappels. Il est à noter que les résultats n'ont pas été validés par simulation. Parallèlement, des simulations ont été effectuées, les estimations de la moyenne et de la variance de la durée du séjour ont été obtenues et la distribution du nombre de nouveaux rappels a également été déterminée. L'analyse de simulation sera publiée prochainement.

8.3.2.3. Approche asymptotique

Les résultats asymptotiques publiés dans (Nazarov *et al.* 2017b, 2017c) sont résumés dans cette section. Auparavant, nous avons besoin de quelques notations, à savoir :

$$B^*(\alpha) = \int_0^\infty e^{-\alpha x} dB(x), \quad \delta(\kappa_1) = \lambda + (\sigma - \lambda)\kappa_1$$

Alors, κ_1 peut s'obtenir de :

$$\kappa_1 = 1 - \frac{\delta(\kappa_1)}{\lambda} \cdot \frac{B^*(\delta(\kappa_1))}{2 - B^*(\delta(\kappa_1))} \quad [8.2]$$

et la distribution de l'état du serveur peut être déterminée par :

$$R_0 = \frac{1}{2 - B^*(\delta)}, \quad R_1 = \frac{1 - B^*(\delta)}{2 - B^*(\delta)}$$

Dans le cas du temps de service résiduel jusqu'à l'approche d'achèvement de service avec succès pour l'état du serveur occupé, nous avons :

$$R_1(z) = e^{\delta z} \int_0^z e^{-\delta x} \cdot \{\lambda(1 - \kappa_1) - \delta R_0 B(x)\} dx$$

et dans le cas de l'approche du temps de service écoulé, nous obtenons :

$$\int_0^\infty R_1(y) \mu(y) dy = \lambda(1 - \kappa)$$

En introduisant les notations :

$$A_1 = \lambda(1 - \kappa_1), \quad R_1^*(\alpha) = -\delta R_0 [B^*(\alpha)]$$

nous obtenons :

$$\kappa_2 = \frac{A_1(R_0 \cdot B^*(\delta) [\delta + A_1] - (\delta + A_1 R_0))}{A_1(\sigma - \lambda)(R_1^*(\delta) - R_1 - R_0(B^*(\delta) - 1)) + \delta((\sigma - \lambda)(R_1^*(\delta) - R_0 B^*(\delta)) - \lambda)}$$

Par conséquent, la distribution prélimite en régime permanent du nombre de clients dans le système peut être approchée par une distribution normale avec moyenne $N\kappa_1$ et variance $N\kappa_2$.

Il est intéressant de noter que l'équation [8.2] peut avoir une, deux ou trois racines $0 < \kappa_1 < 1$. Par exemple, pour la fonction de distribution Gamma $B(x)$ avec un paramètre de forme α et un paramètre d'échelle β avec des valeurs $\alpha = \beta = 2$, $\lambda = 0,29$, $\sigma = 20$ l'équation [8.2] a trois racines, à savoir $\kappa_1^{(1)} = 0,031$, $\kappa_1^{(2)} = 0,188$ et $\kappa_1^{(3)} = 0,549$.

Considérons ces valeurs des paramètres de service α , β et le système λ , σ pour lequel l'équation [8.2] a trois racines ; la distribution de probabilité $P(i) = P\{i(t) = i\}$ peut alors être d'un type bimodal. En particulier, pour $\alpha = \beta = 2$, $\lambda = 0,29$, $\sigma = 19,7$ par simulation et méthodes numériques à $N = 200$ pour $P(i)$ nous obtenons le graphique suivant présenté à la figure 8.3.

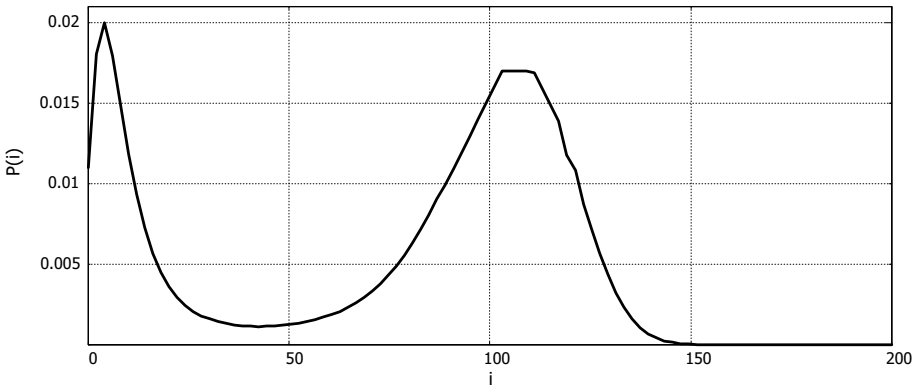


Figure 8.3. Distribution de probabilité bimodale du nombre de clients dans le système

La raison de cette caractéristique est expliquée par Nazarov *et al.* (2017c).

En règle générale, l'équation [8.2] a trois racines dans des cas exceptionnels à des valeurs spéciales de paramètres et une telle situation est extrêmement rare. Considérons donc dans ce qui suit quelques propriétés du système lorsque l'équation principale [8.2] a une racine unique $0 < \kappa_1 < 1$.

Dans l'article de Nazarov *et al.* (2017b), pour la durée moyenne de séjour du client en service, on obtient l'expression suivante :

$$\lim_{N \rightarrow \infty} E(T_S) = \frac{1 - B^*(\delta)}{\delta B^*(\delta)}$$

Considérons ensuite l'influence des paramètres du système sur le temps de séjour moyen $E(T_S)$ du client dans le serveur. Nous choisirons $\sigma = 20$ et les valeurs des paramètres λ et $\alpha = \beta$ sont spécifiées dans le tableau 8.1.

Comme nous pouvons le voir, pour $\alpha < 1$ les valeurs du temps de séjour moyen total $E(T_S)$ du client sous service prend des valeurs inférieures à l'unité. Pour $\alpha < 1$, il y a une forte probabilité d'émergence de petites valeurs de temps de service et ce fait influence sans doute le temps de séjour moyen total $E(T_S)$ du client sur le serveur et il prend plutôt de faibles valeurs. De plus, notons qu'avec l'augmentation des valeurs du paramètre λ les valeurs d' $E(T_S)$ diminuent.

		$\alpha = \beta$							
		0,1	0,3	0,5	0,8	1	2	3	5
λ	0,5	0,324	0,546	0,682	0,857	1	5,038	20,816	153,239
	1	0,204	0,367	0,488	0,727	1	5,576	21,695	154,766
	5	0,067	0,165	0,301	0,640	1	5,937	22,360	155,975
	10	0,046	0,140	0,280	0,632	1	5,979	22,441	156,125
	15	0,039	0,131	0,273	0,629	1	5,993	22,468	156,175

Tableau 8.1. Temps de séjour moyen $E(T_S)$ du client en service à différentes valeurs de λ et $\alpha = \beta$

Dans le cas où $\alpha > 1$, les valeurs d' $E(T_S)$ deviennent plus grandes que l'unité. Le tableau 8.1 montre qu'avec une augmentation du paramètre de service α , les valeurs du temps de séjour moyen total $E(T_S)$ du client sous service augmentent considérablement et atteignent des valeurs très élevées. Notons que dans cette

situation, le paramètre λ n'influence pratiquement pas $E(T_S)$ et avec un paramètre de service croissant α cette influence devient de moins en moins importante.

Le tableau 8.1 montre également qu'en cas de temps de service exponentiel, c'est-à-dire que $\alpha = \beta = 1$, les valeurs du temps de séjour moyen total $E(T_S)$ du client sur le serveur sont égales à l'unité.

Pour la répartition du nombre de nouveaux rappels/transitions du client placé en orbite, nous avons les résultats suivants.

Soit ν le nombre de transitions du client en orbite, puis :

$$\lim_{N \rightarrow \infty} E z^\nu = \frac{q}{1 - (1-q)z}$$

où la valeur du paramètre q a une forme : $q = R_0 B^*(\delta)$.

Du théorème prouvé, il s'ensuit évidemment que la distribution de probabilité $P\{\nu = n\}$, $n = \overline{0, \infty}$ du nombre de transitions du client placé en orbite est géométrique et :

$$P\{\nu = n\} = q(1-q)^n, \quad n = \overline{0, \infty}$$

Par conséquent, en utilisant la loi de probabilité totale pour la fonction caractéristique du temps de séjour/attente W du client en orbite, on obtient :

$$E e^{iuW} \approx q + (1-q) \frac{\sigma q}{\sigma q - iuN}$$

Dans le cas où $N \rightarrow \infty$, les distributions de probabilités limites du temps de séjour du client dans le système T et du temps de séjour du client en orbite W coïncident, à savoir :

$$\lim_{N \rightarrow \infty} E \exp \left\{ iu \frac{T}{N} \right\} = \lim_{N \rightarrow \infty} E \exp \left\{ iu \frac{W}{N} \right\} = q + (1-q) \frac{\sigma q}{\sigma q - iu}$$

Considérons l'influence des paramètres du système sur les valeurs du nombre moyen de rappels ν . Nous choisirons les mêmes paramètres du système qui ont été pris en compte dans les exemples précédents, à savoir $\sigma = 20$, et les valeurs des paramètres λ et $\alpha = \beta$, et qui sont spécifiés dans le tableau 8.2.

		$\alpha = \beta$							
		0,1	0,3	0,5	0,8	1	2	3	5
λ	0,5	0,470	1,131	1,86	3,65	6,32	162,7	793	6 091
	1	0,656	2,002	4,19	11,57	21,83	204,1	848	6 172
	5	1,114	4,192	9,31	22,73	37,07	234,4	891	6 236
	10	1,278	4,755	10,28	24,30	39,02	238,1	896	6 244
	15	1,355	4,969	10,63	24,83	39,67	239,3	898	6 247

Tableau 8.2. Nombre moyen de nouveaux rappels pour diverses valeurs de λ et $\alpha = \beta$

Le tableau 8.2 montre que le nombre moyen de nouveaux rappels dépend évidemment du paramètre de service α qui est le résultat attendu et logique. Il est à noter que pour les petites valeurs du paramètre α , l'influence du paramètre λ sur le nombre moyen de rappels est significative et évidente. Mais avec l'augmentation du paramètre α , cette influence diminue et disparaît pratiquement déjà pour des valeurs importantes de α .

On peut tirer les conclusions suivantes : le nombre moyen de rappels et la durée moyenne du séjour du client sous service sont une conséquence de la collision de clients et de l'admissibilité des tentatives répétées de service par le même client. La durée du service à la clientèle pour les tentatives répétées a la même distribution de probabilité $B(x)$, mais sa réalisation répétée, naturellement, accepte des valeurs différentes. Si, pour la distribution $B(x)$, il y a une forte probabilité d'émergence de faibles valeurs de temps de service comme dans la distribution Gamma avec le paramètre de forme $\alpha < 1$, alors un petit nombre de rappels est suffisant pour réaliser une faible valeur du temps de service qui sera complété avec succès. Si les petites valeurs du temps de service sont peu probables pour la distribution de probabilité $B(x)$, comme dans la distribution Gamma avec le paramètre de forme $\alpha > 1$, alors le nombre de tentatives de service infructueuses devient important, comme on peut le voir au tableau 8.2, le serveur fonctionne sans résultat, la durée moyenne du séjour du client sous service augmente (tableau 8.1).

8.4. Systèmes avec serveur peu fiable

Dans de nombreuses situations pratiques, le serveur n'est pas fiable et après un temps aléatoire, il peut tomber en panne et nécessite une réparation qui prend également une durée aléatoire (Almási *et al.* 2005 ; Dragieva 2014 ; Gharbi et

Dutheillet 2011 ; Gharbi et Ioualalalen 2006 ; Gharbi *et al.* 2014 ; Krishnamoorthy *et al.* 2014 ; Roszik 2004 ; Sztrik *et al.* 2006 ; Wang *et al.* 2010, 2011 ; Zhang et Wang 2013). Dans les sections qui suivent, nous résumons nos résultats obtenus par différentes méthodes.

8.4.1. Système M/M/1

8.4.1.1. Approche assistée par outil par MOSEL

Puisque, dans de nombreuses situations pratiques, l'espace d'état de la chaîne de Markov descriptive est très large, il est assez difficile de calculer les mesures du système au moyen de la manière traditionnelle d'écrire et de résoudre les équations de régime permanent sous-jacentes.

Pour simplifier cette procédure, plusieurs progiciels ont été développés et utilisés efficacement pour l'évaluation de la performance de systèmes complexes (voir par exemple, (Gharbi et Dutheillet 2011 ; Gharbi et Ioualalalen 2006 ; Gharbi *et al.* 2015 ; Gharbi *et al.* 2014 ; Ikhlef *et al.* 2016)). Dans le cadre de nos recherches, un outil logiciel similaire appelé MOSEL (*Modeling, Specification and Evaluation Language*) a été utilisé pour formuler le modèle et obtenir les mesures de performance. Une étude précédente de Bérczes *et al.* (2017) adresse la formulation du modèle, la dérivation de plusieurs mesures de performance et de la génération d'exemples illustratifs montrant un phénomène intéressant de files d'attente pour un nouveau rappel à source finie, c'est-à-dire que le temps de séjour moyen a un maximum lorsque l'intensité de l'arrivée augmente, dans une configuration paramétrique spécifique. Nous avons montré que cela subsiste encore dans le cas des collisions, comme le montre l'exemple suivant.

Cas	N	λ/N	γ_0	γ_1	γ_2	σ/N	μ
Figure 8.4 cas 1	100	0,01	0,01	0,01	1	0,1	1
Figure 8.4 cas 2	100	0,01	0,1	0,1	1	0,1	1
Figure 8.4 cas 3	100	0,01	1	1	1	0,1	1
Figure 8.5	100	0,03 – 8,1	0,01	0,01	1	0,1	1
Figure 8.6	100	0,03 – 8,1	0,1	0,1	1	0,1	1

Tableau 8.3. Valeurs numériques des paramètres du modèle

La figure 8.4 montre la répartition à l'état d'équilibre des trois cas étudiés. Dans cette figure, on peut également voir l'effet de la panne du serveur. Nous constatons que le nombre moyen de clients augmente à mesure que l'intensité de la panne augmente. D'après la forme des courbes, il est clairement visible que la distribution à l'état d'équilibre des cas est normalement distribuée.

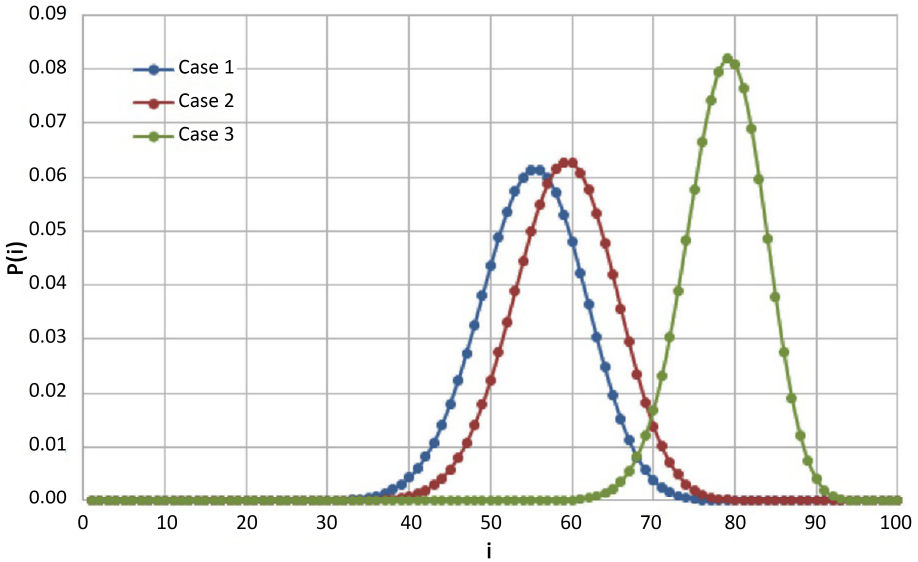


Figure 8.4. Distributions à l'état d'équilibre

Les figures 8.5 et 8.6 montrent le temps de réponse moyen en fonction du taux de génération des clients. Comme nous pouvons le voir, le temps de réponse moyen sera plus élevé à mesure que nous augmentons le taux de génération, mais après que λ / N soit supérieur à 1.5 le temps de réponse moyen commence à diminuer.

8.4.1.2. Approche algorithmique

L'avantage d'une approche algorithmique de la méthode générale supportée par un outil réside dans le fait que dans ce cas nous n'avons pas d'explosion d'espace d'état et que N peut être beaucoup plus élevé que dans le cas des calculs effectués par MOSEL comme nous le verrons plus tard. La publication des résultats a été soumise (voir (Kuki *et al.* 2019)). Le calcul a été effectué par un tableur, MS Excel, et N peut être de taille arbitraire. Lorsque les calculs sont effectués par l'outil MOSEL-2 (voir (Bérczes *et al.* 2017)), on se heurte à des limites strictes, à savoir que l'espace d'état croît extrêmement vite ; par conséquent, le nombre de sources ne peut dépasser 200. Dans Excel, nous pouvons aller bien au-delà de 200. Deux autres

avantages de l'utilisation de cette feuille de calcul : l'effet des modifications des paramètres est immédiatement visible et l'ensemble des probabilités d'équilibre, tant bidimensionnelles qu'unidimensionnelles, sont présentes dans des colonnes séparées et peuvent être utilisées directement pour des études plus approfondies.

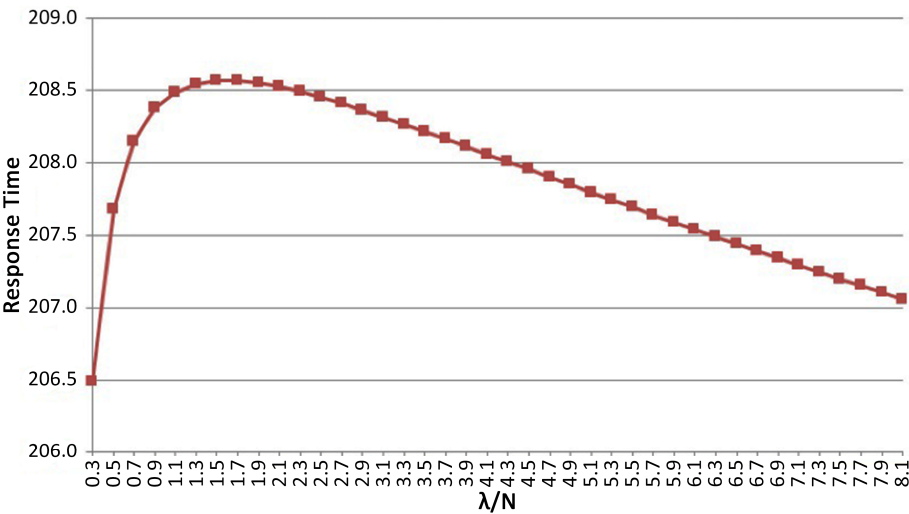


Figure 8.5. Temps de réponse moyen versus λ / N , $\gamma_0 = \gamma_1 = 0,01$

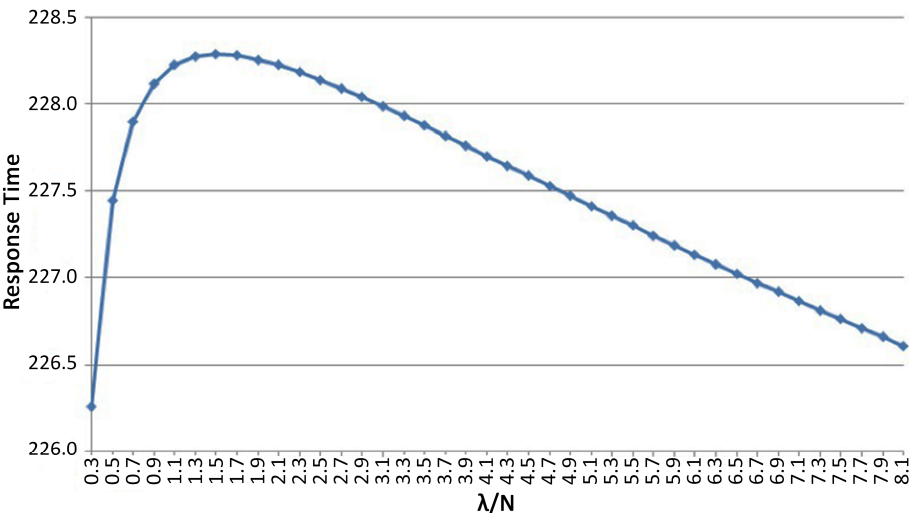


Figure 8.6. Temps de réponse moyen versus λ / N , $\gamma_0 = \gamma_1 = 0,1$

À titre d'exemple, nous considérons trois systèmes avec les paramètres d'entrée suivants :

$$N = 100, \quad \lambda = 1, \quad \mu = 1, \quad \sigma = 5, \quad \gamma_0 = 0,1, \quad \gamma_1 = 0,1, \quad \gamma_2 = 1$$

La figure 8.7 présente les probabilités d'équilibre de trois modèles : un système fiable de base sans conflit, un système fiable avec conflit et un système peu fiable avec conflit. En l'absence de conflit, les attentes des états sont moins élevées que dans les autres cas. Les probabilités des états ont la moyenne la plus élevée dans le système peu fiable comme on s'y attendait.

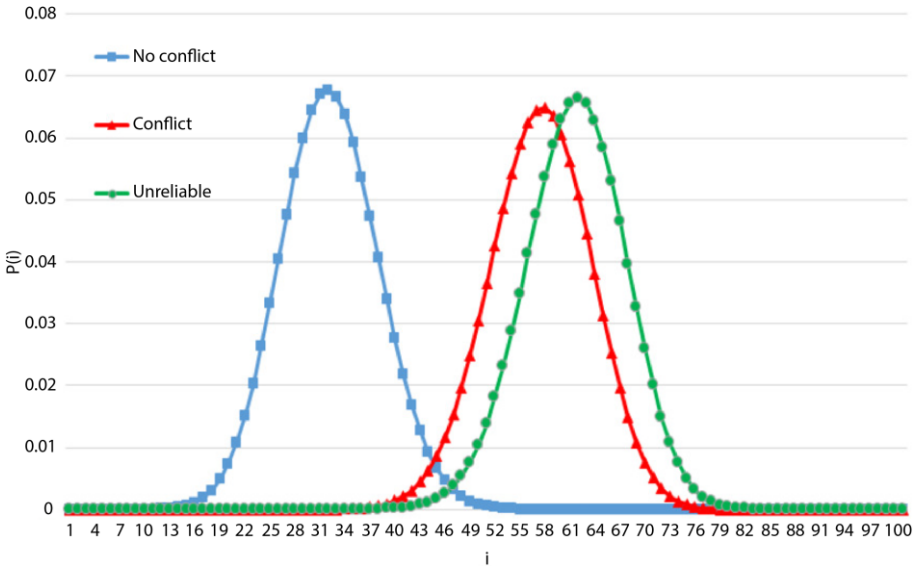


Figure 8.7. *Fiable sans conflit, fiable avec conflit et peu fiable avec conflit*

8.4.1.3. Simulation stochastique

Afin de valider l'applicabilité de l'approche asymptotique, nous avons besoin de résultats numériques ou de simulation. Le bon fonctionnement du logiciel de simulation a été testé à l'aide des exemples d'échantillons numériques. Les recherches réalisées par les méthodes de simulation et asymptotiques ont été soumises pour être publiées (voir (Nazarov, Sztrik, Kvach et Bérczes 2018 ; Nazarov, Sztrik, Kvach et Tóth 2018)).

8.4.1.4. Approche asymptotique

Nous traitons premièrement de la répartition du nombre de clients dans le système telle qu'elle a été publiée par Nazarov *et al.* (2018). Les résultats asymptotiques de premier ordre sont les suivants :

$$\lim_{N \rightarrow \infty} E \exp \left\{ i w \frac{J}{N} \right\} = \exp \{ i w \kappa_1 \}$$

où κ_1 est la solution positive de l'équation :

$$(1 - \kappa_1) \lambda - \mu R_1(\kappa_1) = 0$$

où les distributions stationnaires des probabilités $R_k(\kappa_1)$ de l'état du serveur $k = 0, 1, 2$ sont obtenues comme suit :

$$R_0(\kappa_1) = \left\{ \frac{\gamma_0 + \gamma_2}{\gamma_2} + \frac{\gamma_1 + \gamma_2}{\gamma_2} \cdot \frac{a(\kappa_1)}{a(\kappa_1) + \gamma_1 + \mu} \right\}^{-1}$$

$$R_1(\kappa_1) = \frac{a(\kappa_1)}{a(\kappa_1) + \gamma_1 + \mu} \cdot R_0(\kappa_1)$$

$$R_2(\kappa_1) = \frac{1}{\gamma_2} [\gamma_0 R_0(\kappa_1) + \gamma_1 R_1(\kappa_1)]$$

ici $a(\kappa_1)$ est :

$$a(\kappa_1) = (1 - \kappa_1) \lambda + \sigma \kappa_1$$

Les résultats asymptotiques du deuxième ordre sont :

$$\lim_{N \rightarrow \infty} E \exp \left\{ i w \frac{J - \kappa_1 N}{\sqrt{N}} \right\} = \exp \left\{ \frac{(i w)^2}{2} \kappa_2 \right\}$$

où κ_2 est :

$$\kappa_2 = \frac{\gamma_2 \mu (R_1 - b_1) + (1 - \kappa_1) \lambda \{ (\gamma_1 + \gamma_2) b_1 + (1 - \kappa_1) \lambda R_2 \}}{(\lambda + \mu b_2) \gamma_2 - (1 - \kappa_1) \lambda (\gamma_1 + \gamma_2) b_2}$$

et :

$$b_1 = \frac{(1 - \kappa_1) \lambda}{a + \gamma_1 + \mu} R_0, \quad b_2 = \frac{(\sigma - \lambda)(R_0 - R_1)}{a + \gamma_1 + \mu}$$

Par conséquent, la distribution prélimite du nombre de clients dans le système peut être approchée par une distribution normale avec moyenne $N \kappa_1$ et écart $N \kappa_2$.

Déterminons l'exactitude et le champ d'application de cette approximation.

Nous avons les paramètres d'entrée suivants :

$$\mu = 1, \quad \sigma = 5, \quad \gamma_0 = 0,1, \quad \gamma_1 = 0,2, \quad \gamma_2 = 1$$

et nous fournirons les valeurs de la distance de Kolmogorov Δ dans le tableau 8.4.

	<i>N</i> = 5	<i>N</i> = 10	<i>N</i> = 20	<i>N</i> = 30	<i>N</i> = 50
$\lambda = 0,5$	0,095	0,059	0,032	0,023	0,017
$\lambda = 1,0$	0,037	0,023	0,017	0,014	0,011
$\lambda = 2,0$	0,078	0,046	0,022	0,014	0,013

Tableau 8.4. Distance de Kolmogorov entre la distribution prélimite $P(i)$ et la distribution asymptotique $P_{as}(i)$ pour des valeurs différentes des paramètres N et λ

Le tableau 8.5 montre la distance de Kolmogorov pour les paramètres suivants :

$$\lambda = 1, \quad \mu = 1, \quad \gamma_0 = 0,1, \quad \gamma_1 = 0,2, \quad \gamma_2 = 1$$

et pour diverses valeurs des paramètres N et σ .

	$N = 5$	$N = 10$	$N = 20$	$N = 30$	$N = 50$
$\sigma = 0,2$	0,066	0,035	0,018	0,013	0,010
$\sigma = 1,0$	0,033	0,018	0,014	0,014	0,008
$\sigma = 5,0$	0,037	0,023	0,017	0,014	0,011

Tableau 8.5. Distance de Kolmogorov entre la distribution prélimite $P(i)$ et la distribution asymptotique $P_{as}(i)$ pour différentes valeurs des paramètres N et σ

En supposant une erreur acceptable $\Delta \leq 0,05$ des valeurs données dans les tableaux 8.4 et 8.5, nous pouvons conclure que l'approximation a une erreur importante à $N \leq 10$, à $10 < N < 20$ l'acceptabilité de l'approximation est douteuse, et à $N \geq 20$ l'approximation a une précision assez acceptable.

Passons maintenant à l'analyse de la répartition du temps d'attente et du temps de séjour.

En utilisant la formule de Little, nous avons :

$$(1 - \kappa_1) \lambda E\left(\frac{1}{N} T\right) = \kappa_1$$

ainsi :

$$E\left(\frac{1}{N} T\right) = \frac{\kappa_1}{(1 - \kappa_1) \lambda} \quad [8.3]$$

où κ_1 est la valeur moyenne asymptotique du nombre de clients dans le système et $(1 - \kappa_1) \lambda$ l'intensité moyenne d'arrivée du flux entrant.

Toutefois, il convient de souligner que l'égalité ci-avant [8.3] ne définit que le temps de séjour moyen du client dans le système. Nous aimerions effectuer une étude plus détaillée pour le client T marqué.

L'une des principales contributions de Nazarov *et al.* (2018) est que pour la limite de la fonction caractéristique du temps de séjour normalisé nous avons :

$$\lim_{N \rightarrow \infty} E \exp \left\{ i w \frac{T}{N} \right\} = q + (1-q) \frac{\sigma q}{\sigma q - i w}$$

où q est :

$$q = \frac{(1 - \kappa_1) \lambda}{(1 - \kappa_1) \lambda + \sigma \kappa_1}$$

Par conséquent, la fonction caractéristique du temps de séjour du client dans le système dans la situation prélimite avec N fini peut être approximée par :

$$E e^{i w T} \approx q + (1-q) \frac{\sigma q}{\sigma q - i w N} \quad [8.4]$$

Dans ce qui suit, trouvons la valeur moyenne du temps de séjour normalisé du client dans le système qui est :

$$E \left(\frac{1}{N} T \right) = (1-q) \frac{1}{\sigma q} = \frac{1}{\sigma} \cdot \frac{1-q}{q} = \frac{1}{\sigma} \frac{\sigma \kappa_1}{(1 - \kappa_1) \lambda} = \frac{\kappa_1}{(1 - \kappa_1) \lambda}$$

et cela coïncide avec le résultat obtenu plus tôt par la formule de Little.

Pour la distribution du nombre de transitions/rappels du client placé en orbite, nous obtenons les résultats suivants. Soit ν le nombre de transitions du client en orbite, alors :

$$\lim_{N \rightarrow \infty} E z^\nu = \frac{q}{1 - (1-q)z}$$

résultant en la distribution de probabilité $P\{\nu = n\}$, $n = \overline{0, \infty}$ du nombre de transitions du client en orbite qui est géométrique et ayant la forme :

$$P\{\nu = n\} = q(1-q)^n, \quad n = \overline{0, \infty}$$

Par conséquent, la fonction caractéristique prélimite du temps de séjour ou d'attente W du client sélectionné dans une orbite peut être approchée comme :

$$E e^{i w W} \approx q + (1-q) \frac{\sigma q}{\sigma q - i w N}$$

Dans le cas où $N \rightarrow \infty$, les distributions des probabilités limites du temps de séjour du client dans le système T et le temps de séjour du client en orbite W coïncident, à savoir :

$$\lim_{N \rightarrow \infty} \text{Eexp} \left\{ iu \frac{T}{N} \right\} = \lim_{N \rightarrow \infty} \text{Eexp} \left\{ iu \frac{W}{N} \right\} = q + (1-q) \frac{\sigma q}{\sigma q - iu}$$

Auparavant, nous avons obtenu que la distribution de probabilité du nombre de transitions du client en orbite est géométrique avec le paramètre q . Voyons à quel point les résultats de limitation sont proches des résultats de simulation et à quelles valeurs N cette approximation est admissible.

Dans ce but, dénotons par $P_{as}(\nu = n)$ la distribution géométrique asymptotique des probabilités avec paramètre q et par $P_s(\nu = n)$ la distribution de probabilité du nombre de transitions du client sélectionné dans l'orbite obtenue à l'aide du programme de simulation. En outre, déterminons la précision (erreur) de l'approximation de la distribution au moyen de la distance de Kolmogorov Δ , qui, pour les distributions de probabilité, $P_{as}(\nu = n)$ et $P_s(\nu = n)$ est définie comme :

$$\Delta = \max_{0 \leq i < \infty} \left| \sum_{n=0}^i (P_{as}(\nu = n) - P_s(\nu = n)) \right|$$

En exécutant le programme de simulation pour :

$$\lambda = 1, \quad \mu = 1, \quad \sigma = 4, \quad \gamma_2 = 1$$

et en appliquant l'approximation, nous fournirons la distance de Kolmogorov Δ pour diverses valeurs N et $\gamma = \gamma_0 = \gamma_1$ dans le tableau 8.6.

Nous pouvons voir, comme prévu, qu'en augmentant N la distance de Kolmogorov devrait diminuer, mais avec ce paramétrage, il n'y a pas de réduction essentielle si $N > 50$.

Voyons le nombre moyen de nouveaux rappels dans les mêmes conditions qu'auparavant.

Encore une fois, nous pouvons observer ce qui était attendu : à mesure que N augmente, le nombre moyen de rappels augmente puisqu'il y a de plus en plus de clients dans le système, ce qui entraîne de plus en plus de collisions. En même temps, on peut également voir comment le nombre moyen de nouveaux rappels

augmente à mesure que le taux de défaillance du serveur augmente. On souligne qu'à chaque fois les valeurs limites donnent de très bonnes approximations montrant l'efficacité de la méthode asymptotique.

	<i>N</i> = 20	<i>N</i> = 30	<i>N</i> = 50	<i>N</i> = 100	<i>N</i> = 200
$\gamma = 0,05$	0,026	0,016	0,009	0,005	0,003
$\gamma = 0,1$	0,024	0,015	0,009	0,004	0,002
$\gamma = 0,5$	0,017	0,011	0,006	0,004	0,001

Tableau 8.6. Distance de Kolmogorov entre la distribution $P_s(i)$ et l'approximation de la distribution géométrique $P_{as}(i)$ pour différentes valeurs des paramètres *N* et γ

	<i>N</i> = 20	<i>N</i> = 30	<i>N</i> = 50	<i>N</i> = 100	<i>N</i> = 200	Limite
$\gamma = 0,05$	5,512	5,727	5,842	5,900	5,930	5,977
$\gamma = 0,1$	6,090	6,233	6,334	6,415	6,442	6,494
$\gamma = 0,5$	10,336	10,501	10,640	10,715	10,777	10,821

Tableau 8.7. Nombre moyen de nouveaux rappels dans des situations de prélimination et de limitation pour diverses valeurs des paramètres *N* et γ

8.4.2. Système *M/GI/1*

8.4.2.1. Simulation stochastique

Dans (Tóth *et al.* 2017), le temps de service requis est supposé être distribué en Gamma et les paramètres d'entrée du système sont recueillis dans le tableau 8.8.

Cas	<i>N</i>	λ/N	γ_0	γ_1	γ_2	σ/N	α	β
1	100	0,01	0,1	0,1	1	0,01	0,5	0,5
2	100	0,01	0,1	0,1	1	0,01	1	1
3	100	0,01	0,1	0,1	1	0,01	2	2

Tableau 8.8. Valeurs numériques des paramètres du modèle

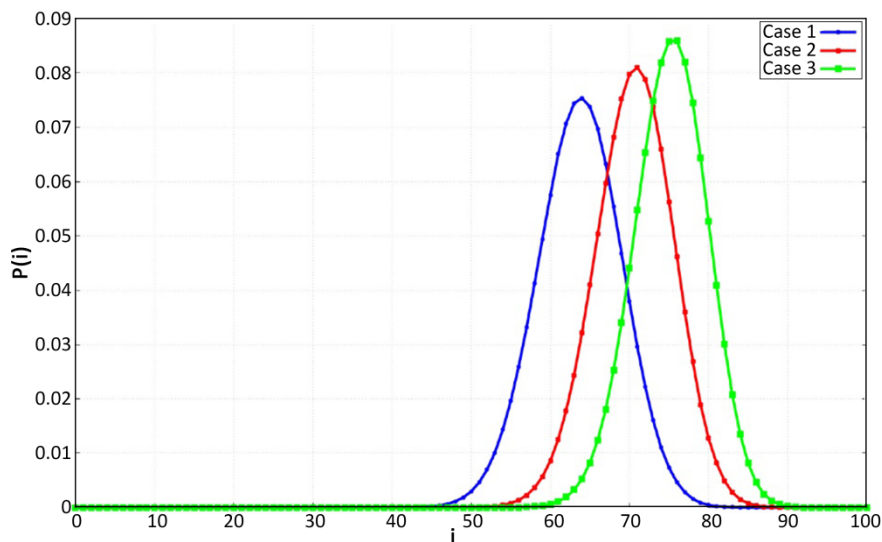


Figure 8.8. Comparaison des distributions en régime permanent

La figure 8.8 montre la distribution à l'état d'équilibre des trois cas à l'étude. On observe que le nombre moyen de clients augmente au fur et à mesure que α et β sont élevés. Le cas 2 est un cas particulier car lorsque $\alpha=1$, il représente la distribution exponentielle. D'après la forme des courbes, il est clairement visible que la distribution à l'état d'équilibre des cas est normalement distribuée. Le tableau suivant présente les mesures de rendement envisagées en fonction des différents cas.

Dans le tableau 8.9, les notes signifient ce qui suit : $E(J)$ et $Var(J)$ sont le nombre moyen et la variance des clients dans le système, $E(T)$ et $Var(T)$ sont la moyenne et la variance du temps de réponse, $E(W)$ et $Var(W)$ sont la moyenne et la variance du temps d'attente, $E(S)$ et $Var(S)$ sont la moyenne et la variance du temps de service complété avec succès, et $E(IS)$ est le temps de service moyen interrompu.

La figure 8.9 représente la confirmation du temps d'attente moyen. Les mêmes paramètres sont utilisés (voir tableau 8.9) comme indiqué dans la figure 8.8 mais ici le paramètre courant est λ/N . Comme on peut s'y attendre avec l'augmentation de λ/N , le temps d'attente moyen augmente également, mais un phénomène intéressant est à noter, à savoir que lorsque λ/N est plus grand que 0,1, le temps d'attente moyen commence à diminuer.

Cas	$E(J)$	$Var(J)$	$E(T)$	$Var(T)$	$E(W)$	$Var(W)$	$E(S)$	$Var(S)$	$E(IS)$
1	63,6	27,9	175,3	65 657,3	174,5	65 434,6	0,3	0,1	0,4
2	70,5	24,3	239,9	105 273,4	238,9	104 918,6	0,4	0,2	0,5
3	75,1	21,2	302,8	151 781,1	301,5	151 277,6	0,6	0,2	0,6

Tableau 8.9. Résultats de la simulation

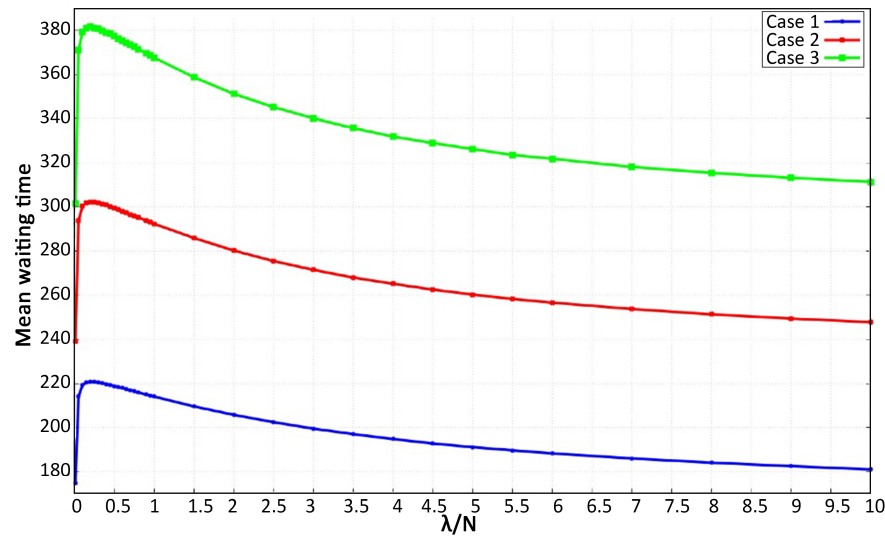


Figure 8.9. Temps d'attente moyen par rapport à l'intensité des clients entrants

8.4.2.2. Approche asymptotique

Ces résultats ont été publiés par Nazarov *et al.* (2017a) à l'aide d'une technique de variables supplémentaires.

La limite de la fonction caractéristique du nombre de clients échelonné dans les systèmes peut être écrite sous la forme suivante :

$$\lim_{N \rightarrow \infty} E \exp \left\{ i w \frac{J}{N} \right\} = \exp \{ i w \kappa_1 \}$$

où κ_1 est la solution positive de l'équation :

$$(1 - \kappa_1) \lambda - \delta(\kappa_1) [R_0(\kappa_1) - R_1(\kappa_1)] + \gamma_1 R_1(\kappa_1) = 0$$

Ici $\delta(\kappa_1)$ est :

$$\delta(\kappa_1) = (1 - \kappa_1) \lambda + \sigma \kappa_1$$

et les distributions stationnaires des probabilités $R_k(\kappa_1)$ de l'état du serveur $k = 0, 1, 2$ sont déterminées comme suit :

$$R_0(\kappa_1) = \left\{ \frac{\gamma_0 + \gamma_2}{\gamma_2} + \frac{\gamma_1 + \gamma_2}{\gamma_2} \cdot \frac{\delta(\kappa_1)}{\delta(\kappa_1) + \gamma_1} [1 - B^*(\delta(\kappa_1) + \gamma_1)] \right\}^{-1}$$

$$R_1(\kappa_1) = R_0(\kappa_1) \frac{\delta(\kappa_1)}{\delta(\kappa_1) + \gamma_1} \cdot [1 - B^*(\delta(\kappa_1) + \gamma_1)]$$

$$R_2(\kappa_1) = \frac{1}{\gamma_2} [\gamma_0 R_0(\kappa_1) + \gamma_1 R_1(\kappa_1)]$$

8.4.3. Simulation stochastique de systèmes spéciaux

Dans (Tóth *et al.* 2017), on a étudié des systèmes ayant non seulement des temps de service distribués selon une loi Gamma, mais aussi des temps d'interarrivées et de rappels distribués selon une loi Gamma.

Le tableau 8.10 présente les paramètres d'entrée de ce cas.

Cas	N	α	β	γ_0	γ_1	γ_2	σ/N	α_1	β_1/N
1	100	1	1	0,1	0,1	1	0,01	0,5	0,01
2	100	1	1	0,1	0,1	1	0,01	1	0,01
3	100	1	1	0,1	0,1	1	0,01	2	0,01

Tableau 8.10. Valeurs numériques des paramètres des temps d'interarrivées distribués selon une loi Gamma

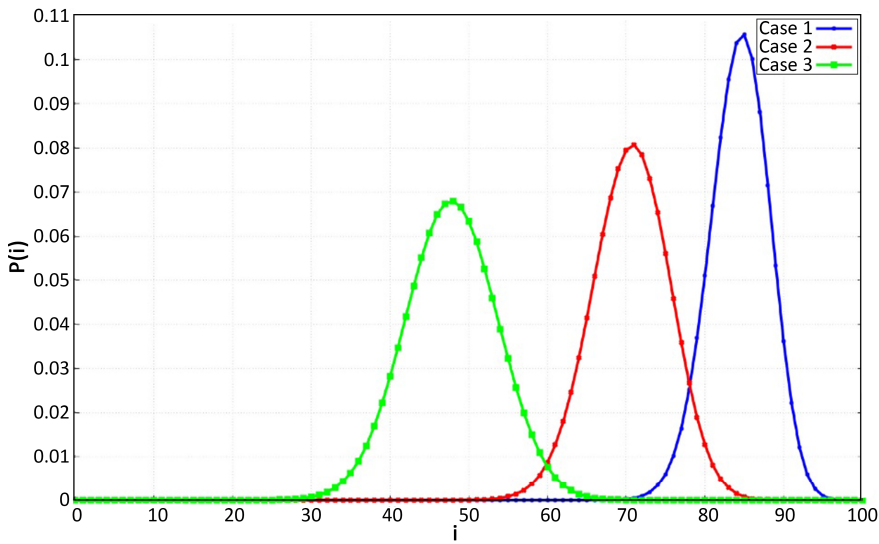


Figure 8.10. Distributions à l'état d'équilibre des temps d'interarrivées distribués selon une loi Gamma

La figure 8.10 illustre la distribution en régime permanent de ce cas. Le temps de service est maintenant distribué de façon exponentielle ($\alpha = 1$) et le temps d'arrivée est distribué selon une loi Gamma. Il est intéressant de noter que, dans ces cas, les distributions à l'état d'équilibre sont toujours distribuées normalement et lorsque α_1 est supérieur à 1, le nombre moyen de demandes dans le système est nettement inférieur à celui des autres cas. Le tableau 8.11 présente les estimations des principales mesures du rendement.

Cas	$E(J)$	$Var(J)$	$E(T)$	$Var(T)$	$E(W)$	$Var(W)$	$E(S)$	$Var(S)$	$E(IS)$
1	84,3	14,1	270	128 831,5	269	128 420,4	0,4	0,2	0,5
2	70,5	24,3	239,9	105 273,4	238,9	104 918,6	0,4	0,2	0,5
3	47,7	34,2	183	69 830,9	182	69 573,9	0,5	0,2	0,4

Tableau 8.11. Résultats de la simulation

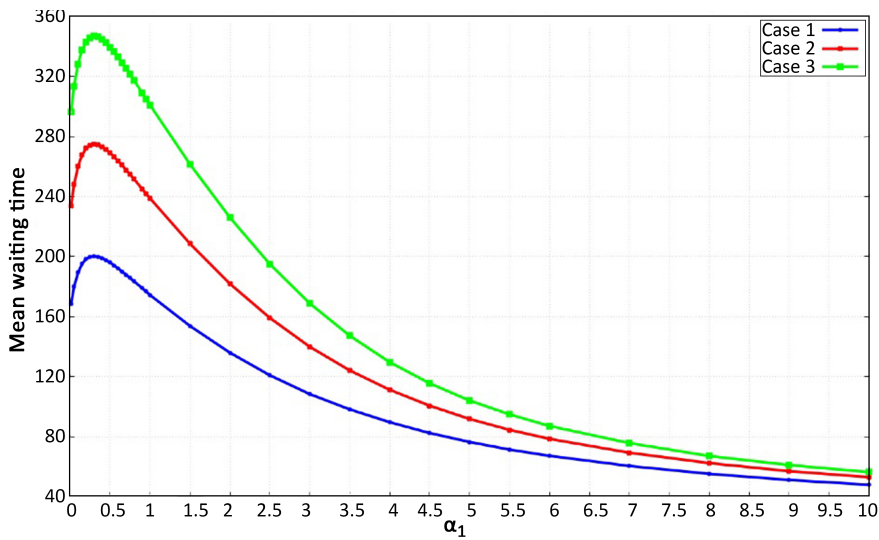


Figure 8.11. Temps d'attente moyen par rapport au paramètre de forme

Dans les figures 8.11 et 8.13, la distribution du temps de service des *cas* possède les paramètres suivants : $\alpha = \beta = 0,5, 1, 2$ et tous les autres paramètres sont les mêmes que ceux du tableau 8.11. Le paramètre en cours d'exécution est α_1 , de sorte que l'impact des différentes distributions sur les différentes mesures de rendement peut être découvert de cette façon. Tout d'abord, le temps d'attente moyen (figure 8.11), après une augmentation initiale du temps d'attente moyen, commence à diminuer de façon

monotone, ce qui fait que lorsque α_1 augmente les clients passent moins de temps dans le système. En fin de compte, les valeurs des différents cas sont presque égales. Il s'ensuit (figure 8.11) qu'il n'est donc pas surprenant que le temps moyen d'interruption de service varie en fonction d' α_1 . Il est intéressant de noter que vers $\alpha_1 \approx 4,5$, on constate que les courbes se croisent.

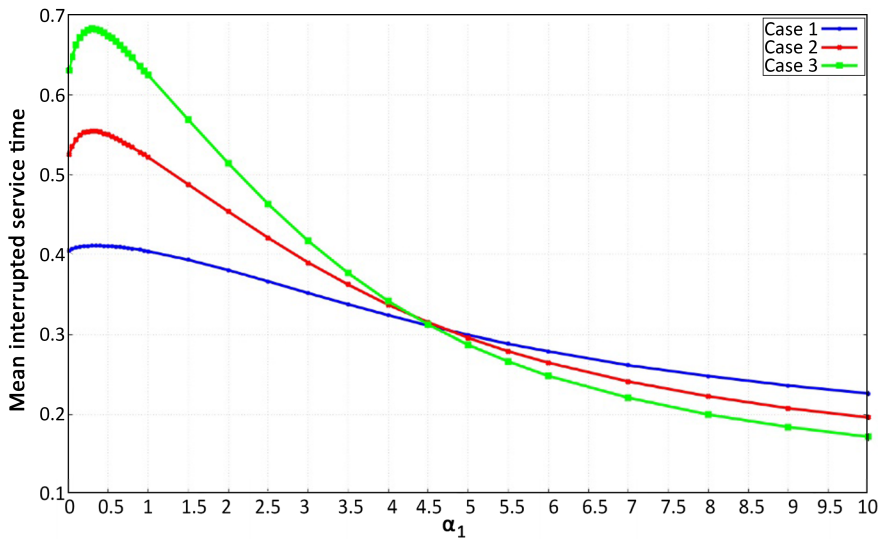


Figure 8.12. Temps moyen d'interruption de service en fonction du paramètre de forme

8.4.4. Temps de rappel distribué selon une loi Gamma

Dans ce cas, non seulement le temps d'interarrivée est distribué selon une loi Gamma avec les mêmes paramètres, mais aussi la durée du nouveau rappel. Le tableau 8.12 présente les paramètres d'entrée de ce système.

Cas	N	α	β	γ_0	γ_1	γ_2	α_1	β_1/N
1	100	1	1	0,1	0,1	1	0,5	0,01
2	100	1	1	0,1	0,1	1	1	0,01
3	100	1	1	0,1	0,1	1	2	0,01

Tableau 8.12. Valeurs numériques des paramètres des temps de rappel distribués selon une loi Gamma

Le tableau 8.13 présente les principales mesures du rendement relatives à ces cas.

Cas	$E(J)$	$Var(J)$	$E(T)$	$Var(T)$	$E(W)$	$Var(W)$	$E(S)$	$Var(S)$	$E(IS)$
1	81,5	16,9	220,8	82 735,7	219,8	82 377,8	0,3	0,1	0,6
2	70,5	24,3	239,9	105 273,4	238,9	104 918,6	0,4	0,2	0,5
3	56,6	30,1	261,4	146 264,4	260,4	145 919	0,6	0,3	0,3

Tableau 8.13. Résultats de la simulation

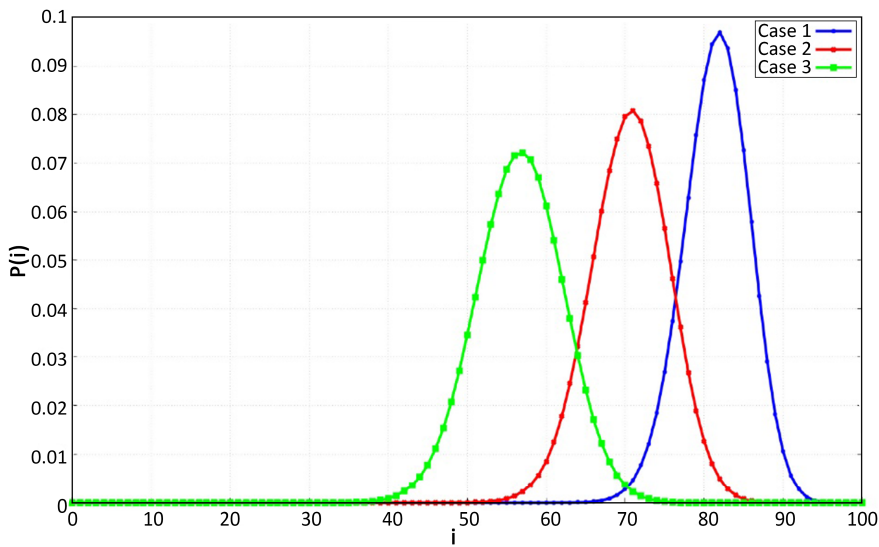


Figure 8.13. Répartition à l'état d'équilibre des temps de rappel distribués selon une loi Gamma

Notons que dans le cadre de ce paramétrage, cette modification n'a pas d'effet significatif sur la distribution en régime permanent (voir figure 8.13). Bien sûr, le nombre moyen de clients dans le système est assez disparate, mais la distribution reste normale. De plus, lorsque α_1 est inférieur à 1, le nombre moyen de clients dans le système est plus élevé que lorsqu'il est supérieur à 1.

8.4.5. L'effet des pannes

Dans (Tóth *et al.* 2019), les systèmes $M/G/1//N$ et $G/M/1//N$ ont été étudiés avec des temps de fonctionnement et de réparation exponentiellement distribués. En cas de panne du serveur, deux modes de fonctionnement sont envisagés :

- la requête interrompue est instantanément placée en orbite (mode 1) ;
- le service de la requête interrompue est suspendu et se poursuit après la réparation du serveur (mode 2).

8.4.5.1. Scénario A, temps de service distribué selon une loi Gamma

Le tableau 8.14 présente les paramètres d'entrée du scénario A.

Cas	N	λ/N	γ_0	γ_1	γ_2	σ/N	α	β
1	100	0,01	0,1	0,1	1	0,01	0,5	0,5
2	100	0,01	0,1	0,1	1	0,01	1	1
3	100	0,01	0,1	0,1	1	0,01	2	2

Tableau 8.14. Valeurs numériques des paramètres du modèle

La figure 8.14 montre la distribution en régime permanent des cas étudiés pour le mode de fonctionnement 2. Le cas 2 est un cas particulier car lorsque le paramètre α est égal à 1, il aboutit à la distribution exponentielle. Le nombre moyen de clients augmente avec l'augmentation des paramètres α et β , si l'on examine de plus près la forme des courbes, la distribution en régime permanent des cas suit la distribution normale. Le tableau 8.15 présente les mesures de rendement envisagées par rapport aux différents cas.

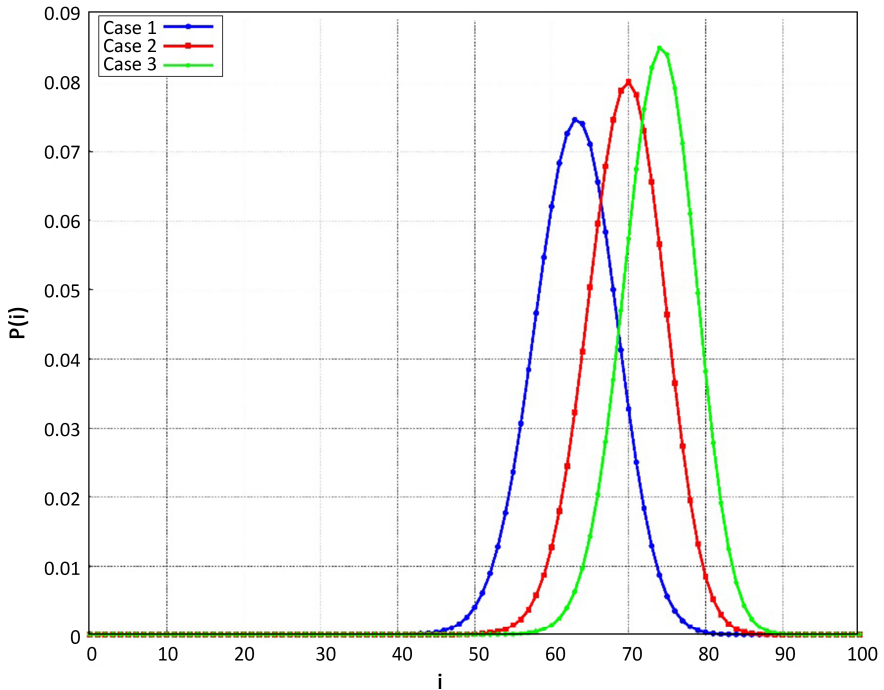


Figure 8.14. Comparaison des distributions en régime permanent pour le mode 2

Les figures 8.15 à 8.17 comparent le temps d'attente moyen des deux différents modes de fonctionnement des cas étudiés. Le mode de fonctionnement 1 reflète le mode lorsque les requêtes interrompues entrent instantanément en orbite ; en mode de fonctionnement 2, nous considérons le mode lorsque le service de la requête interrompue est suspendu et il continue après la réparation du serveur.

Dans tous les cas, les résultats confirment que l'application du mode opératoire 2 entraîne une réduction du temps d'attente moyen. Lorsque les valeurs des paramètres α et β sont plus élevées, la différence entre les modes est également plus grande. Avec l'augmentation de l'intensité des arrivées, il faut s'attendre à des temps d'attente plus longs, mais une fois que λ/N a atteint 0,15, ils commencent à diminuer de façon monotone, ce qui est un phénomène intéressant.

Cas	$E(J)$	$Var(J)$	$E(T)$	$Var(T)$	$E(W)$	$Var(W)$	$E(S)$	$Var(S)$
1	63	28,3	170,5	63 092,8	169,7	62 855,9	0,3	0,2
2	69,6	24,6	228,9	97 974,6	227,8	97 613,9	0,5	0,2
3	73,9	21,9	283,1	136 396,9	281,7	135 909,2	0,6	0,2

Tableau 8.15. Résultats numériques du scénario A

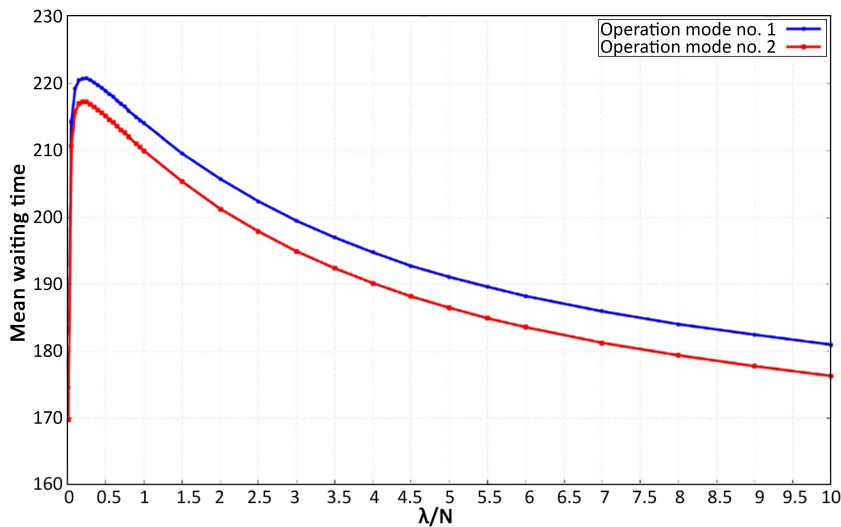


Figure 8.15. Temps d'attente moyen par rapport à l'intensité des nouveaux clients du cas 1

La figure 8.15 montre le temps d'attente moyen en fonction du taux d'échec. Comme on peut s'y attendre, le temps d'attente moyen augmente avec un taux d'échec plus élevé dans tous les cas. En comparant les modes de fonctionnement, la différence est assez évidente. En utilisant le mode de fonctionnement numéro 2, la croissance semble linéaire, et dans le cas du mode de fonctionnement numéro 1, l'augmentation est très significative, surtout dans le cas 3.

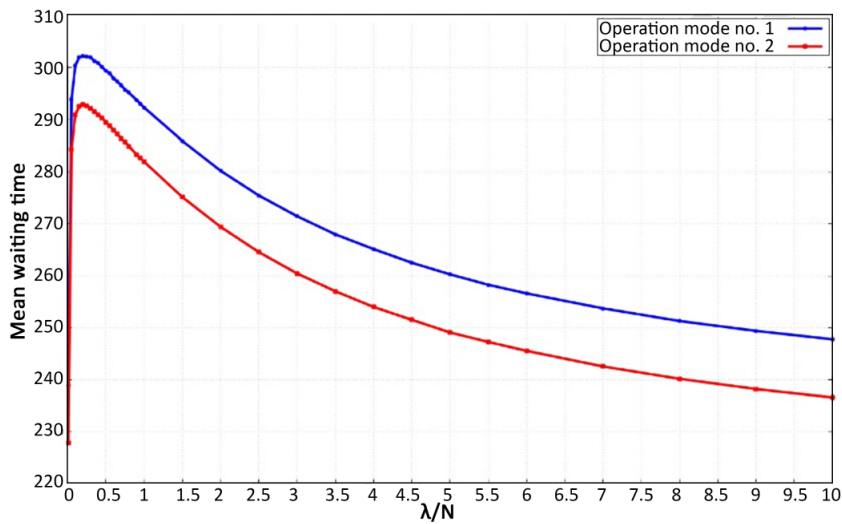


Figure 8.16. Temps d'attente moyen par rapport à l'intensité des nouveaux clients du cas 2

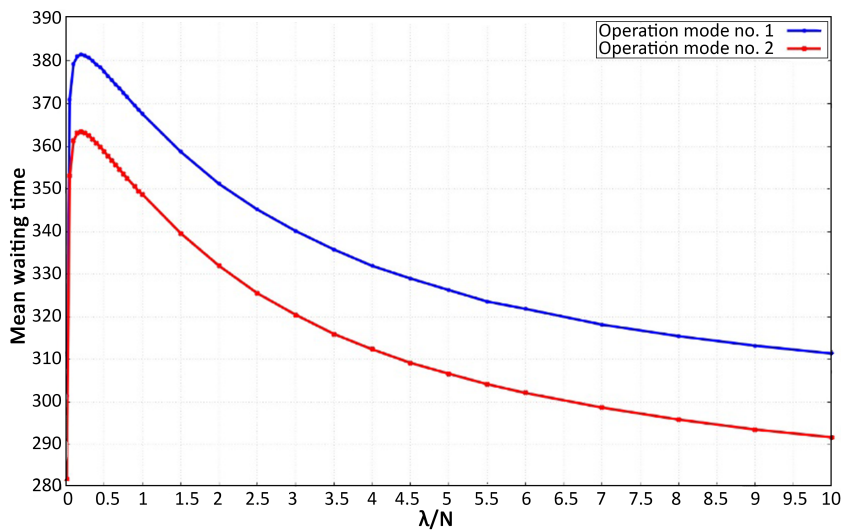


Figure 8.17. Temps d'attente moyen par rapport à l'intensité des nouveaux clients du cas 3

8.4.5.2. Scénario B, temps d'interarrivées distribués selon une loi Gamma

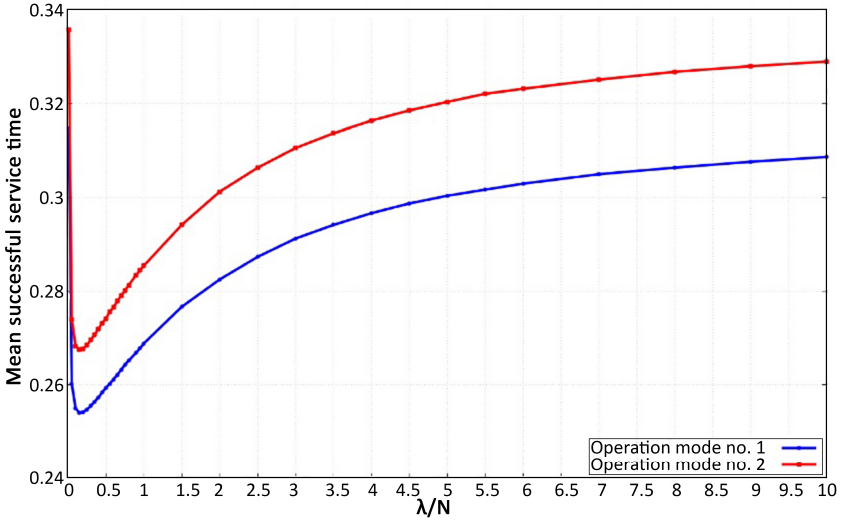


Figure 8.18. Temps moyen de service réussi par rapport à l'intensité des clients entrants du cas 1

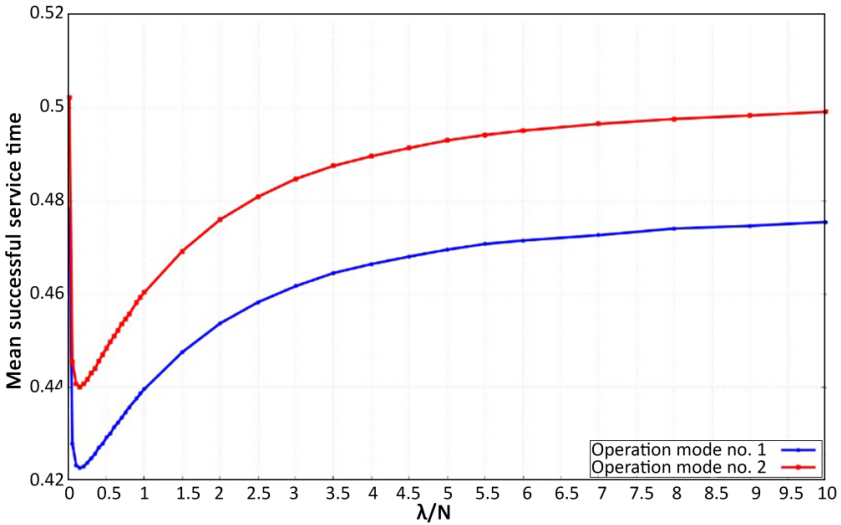


Figure 8.19. Temps moyen de service réussi par rapport à l'intensité des clients entrants du cas 2

Dans le scénario B, la distribution des temps d'interarrivées des clients est distribuée selon une loi Gamma avec le paramètre λ_1 et β_1 .

Le tableau 8.16 présente les valeurs numériques des paramètres du scénario B.

Cas	N	α	β	γ_0	γ_1	γ_2	σ/N	α_1	β_1/N
1	100	1	1	0,1	0,1	1	0,01	0,5	0,01
2	100	1	1	0,1	0,1	1	0,01	1	0,01
3	100	1	1	0,1	0,1	1	0,01	2	0,01

Tableau 8.16. Valeurs numériques des paramètres du scénario B

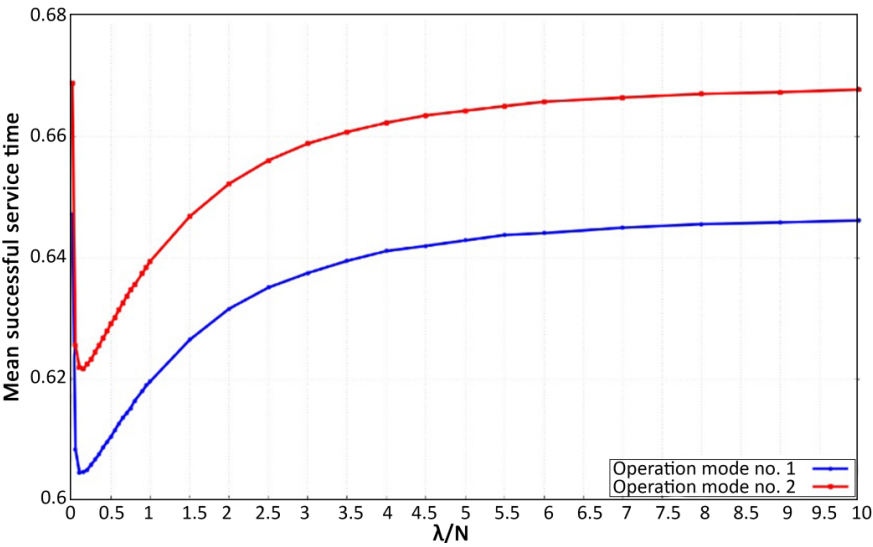


Figure 8.20. Temps moyen de service réussi par rapport à l'intensité des clients entrants du cas 3

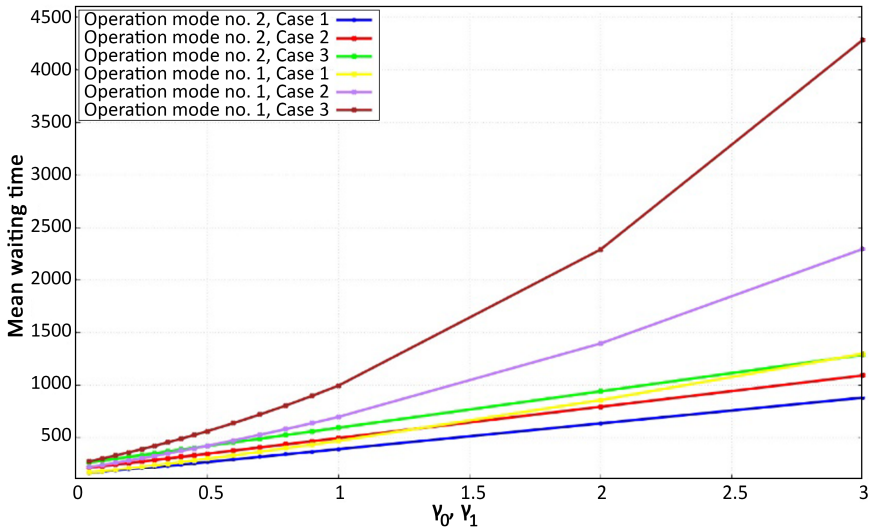


Figure 8.21. Temps d'attente moyen par rapport à l'intensité du taux d'échec

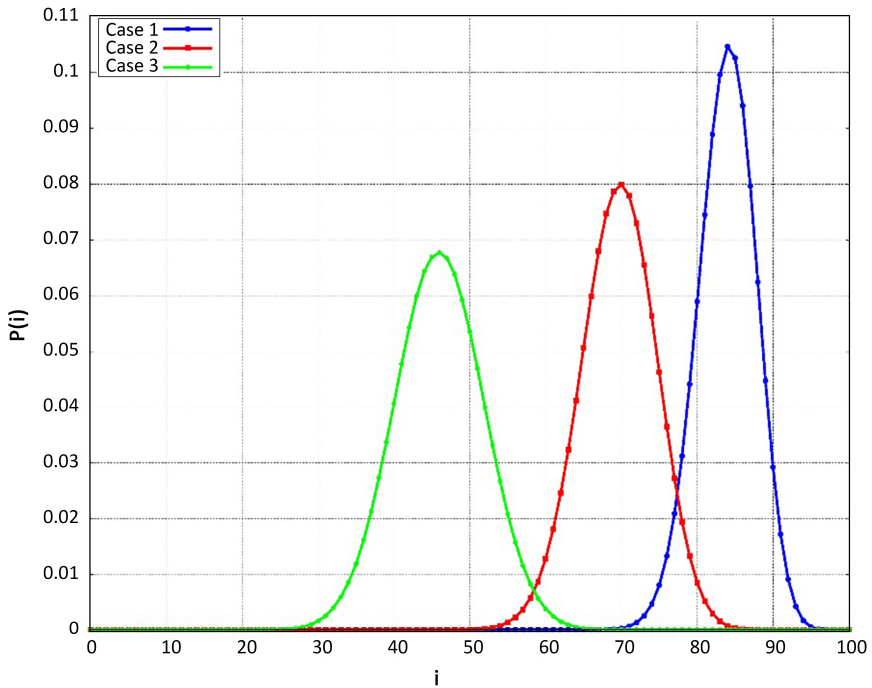


Figure 8.22. Distributions à l'état d'équilibre du scénario B

Cas	E(J)	Var(J)	E(T)	Var(T)	E(W)	Var(W)	E(S)	Var(S)
1	83,8	14,4	259,7	121 328,2	258,6	120 904,8	0,4	0,2
2	69,5	24,7	228,9	97 997,6	227,8	97 637,2	0,5	0,2
3	45,9	34,3	170	62 628,4	168,9	62 379	0,5	0,3

Tableau 8.17. Résultats numériques du scénario B

La figure 8.22 montre les distributions à l'état d'équilibre du scénario B. Maintenant, le temps de service est censé être distribué de façon exponentielle ($\alpha = 1$). Avec cette modification par rapport au scénario A, les distributions à l'état d'équilibre suivent toujours une distribution normale et comme la valeur de α_1 augmente, le nombre moyen de demandes dans le système diminue.

Dans le *scénario 3*, le nombre moyen de clients dans le système est beaucoup plus faible dans les deux cas. Dans le tableau 8.17 se trouve une estimation des principales mesures du rendement avec les cas en question.

Le paramètre courant est α_1 et aide à découvrir l'impact des différentes distributions sur les diverses mesures de performance examinées. Les figures 8.23 à 8.25 montrent la comparaison du temps d'attente moyen entre les deux modes de fonctionnement des cas étudiés. Comme dans le scénario A, l'utilisation du mode de fonctionnement numéro 2 assure des temps d'attente moyens plus courts, lorsque les requêtes interrompues restent sur le serveur en cas de panne de ce dernier et que leurs services continuent après que le serveur soit à nouveau en mesure d'effectuer les tâches.

Dans tous les cas, le temps d'attente moyen commence à augmenter jusqu'à ce qu' α_1 atteigne 0,3, puis diminue de façon monotone. Avec des valeurs plus élevées de α et β , la différence entre les modes de fonctionnement est également plus grande.

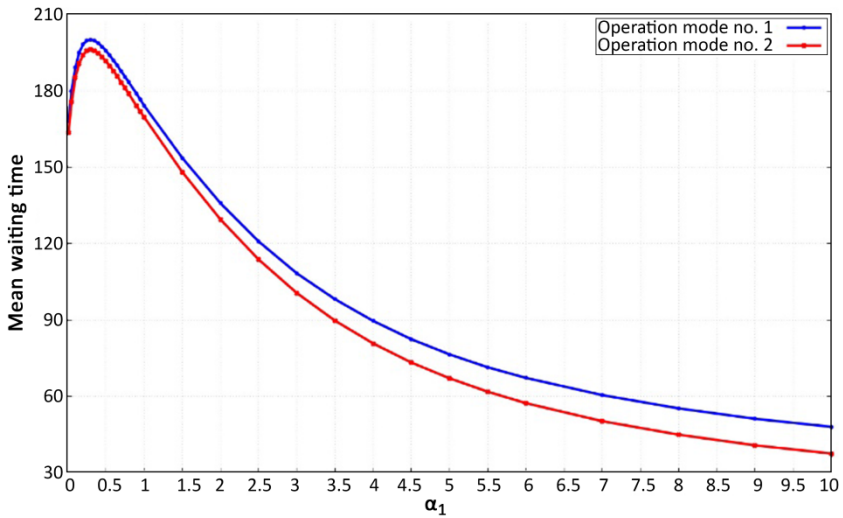


Figure 8.23. Temps d'attente moyen par rapport au paramètre de forme, $\alpha = \beta = 0,5$

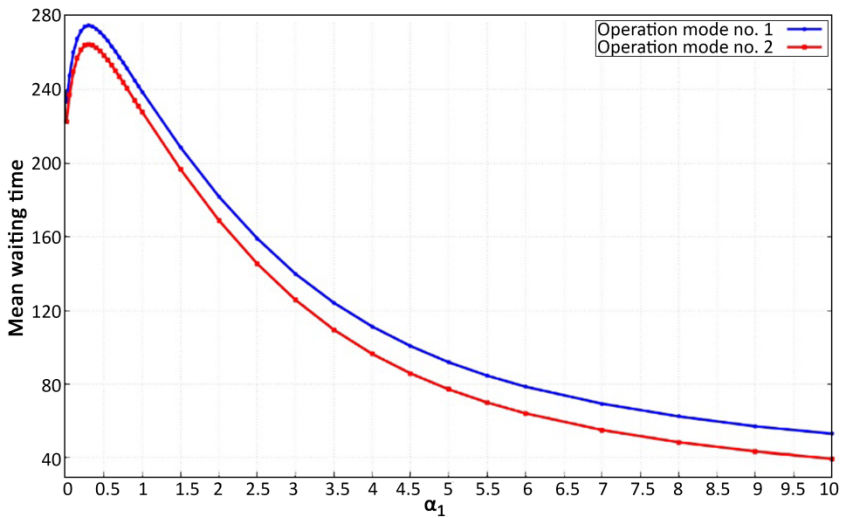


Figure 8.24. Temps d'attente moyen par rapport au paramètre de forme, $\alpha = \beta = 1$

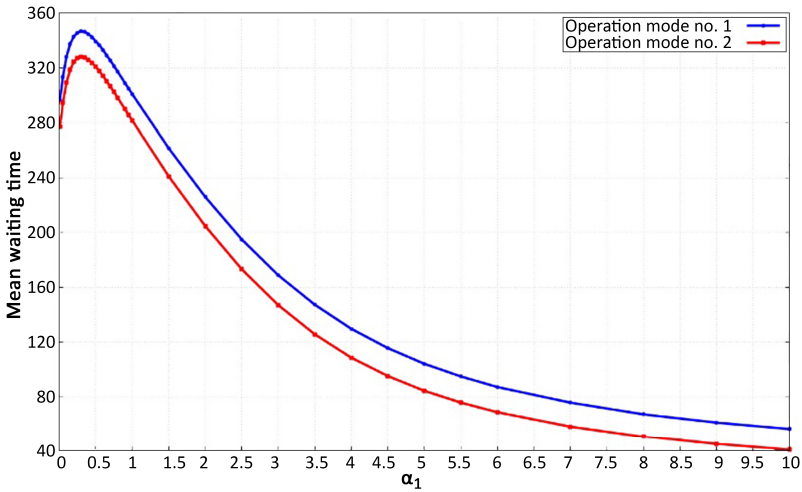


Figure 8.25. Temps d'attente moyen par rapport au paramètre de forme, $\alpha = \beta = 2$

Les figures 8.26 à 8.28 montrent le temps moyen de service réussi par rapport au paramètre de forme du temps d'interarrivée à l'aide des deux modes de fonctionnement. Comme nous pouvons le voir dans le scénario A, nous obtenons ce à quoi nous nous attendions puisque l'utilisation du mode de fonctionnement numéro 2 fournit des valeurs plus élevées de la durée moyenne de service réussie. La différence entre les modes de fonctionnement appliqués est assez élevée dans tous les cas, en particulier lorsque α et β est égale à 0,5. Le temps moyen de service réussi se comporte à l'inverse du temps d'attente moyen, car lorsque le temps d'attente moyen augmente, le temps moyen de service réussi diminue et *vice versa*.

8.5. Conclusion

Dans ce chapitre, on a considéré les méthodes numériques, de simulation et asymptotiques, supportées par des outils, sous la condition d'un nombre croissant et illimité de sources dans une file d'attente avec rappels à source finie avec des collisions de clients et un serveur peu fiable. Au cours de l'étude, plusieurs cas et exemples ont été traités et les résultats des différentes approches ont été comparés entre eux, montrant les avantages et les inconvénients de la méthode donnée. Les tableaux et figures utilisés dans ce chapitre illustrent certaines caractéristiques particulières de ces systèmes. Dans un proche avenir, les deux groupes de recherche aimeraient poursuivre leurs travaux dans ce sens, notamment sur les systèmes avec des clients impatients, les systèmes intégrés dans un environnement aléatoire et

les systèmes à communication bidirectionnelle, pour ne mentionner que quelques généralisations alternatives.

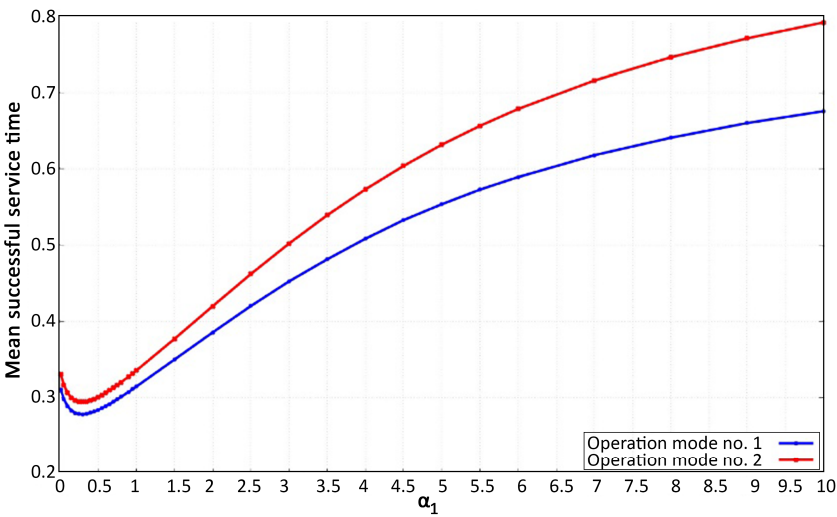


Figure 8.26. Temps de service moyen par rapport au paramètre de forme, $\alpha = \beta = 0,5$

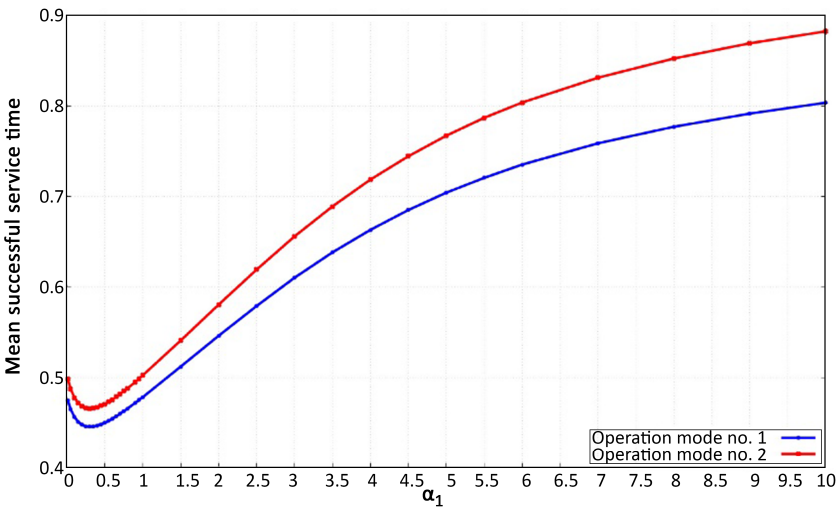


Figure 8.27. Temps de service moyen par rapport au paramètre de forme, $\alpha = \beta = 1$

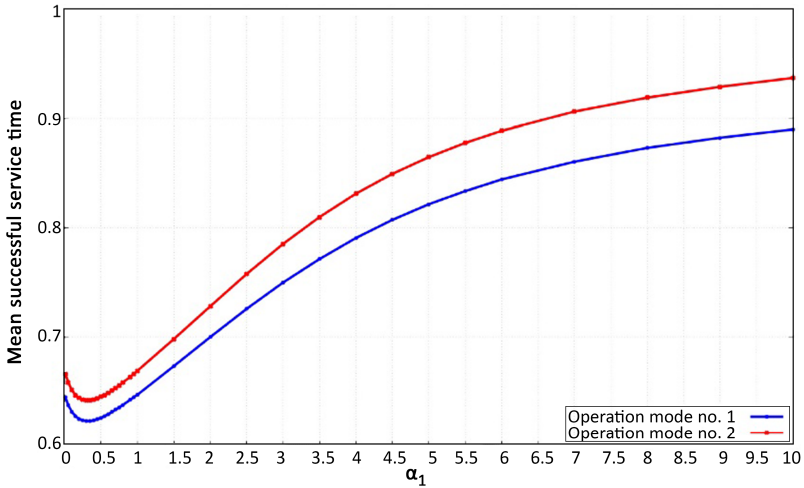


Figure 8.28. Temps de service moyen par rapport au paramètre de forme, $\alpha = \beta = 2$

8.6. Remerciements

Les auteurs sont très reconnaissants envers l'examineur pour ses commentaires et ses recommandations, qui ont amélioré la présentation du chapitre.

Le travail/publication de J. Sztrik est soutenu par le projet EFOP-3.6.1-16-2016-00022. Le projet est cofinancé par l'Union européenne et le Fonds social européen.

8.7. Bibliographie

- Ali, A.-A., Wei, S. (2015). Modeling of coupled collision and congestion in finite source wireless access systems. *Wireless Communications and Networking Conference (WCNC)*, IEEE, New Orleans, 1113–1118.
- Almási, B., Roszik, J., Sztrik, J. (2005). Homogeneous finite-source retrial queues with server subject to breakdowns and repairs. *Math. Comput. Model.*, 42(5–6), 673–682.
- Anisimov, V. (1999). Averaging methods for transient regimes in overloading retrial queueing systems. *Math. Comput. Model.*, 30(3–4), 65–78.
- Anisimov, V., Sztrik, J. (1989). Asymptotic analysis of some complex renewable systems operating in random environments. *European Journal of Operational Research*, 41(2), 162–168.

Cette bibliographie est identique à celle de l'ouvrage correspondant en anglais publié par ISTE.

- Anisimov, V., Artalejo, J. R. (2001). Analysis of Markov multiserver retrial queues with negative arrivals. *Queueing Systems*, 39(2–3), 157–182.
- Artalejo, J., Corral, A.G. (2008). *Retrial Queueing Systems: A Computational Approach*. Springer, Berlin.
- Bérczes, T., Sztrik, J., Tóth, Á., Nazarov, A. (2017). Performance modeling of finite-source retrial queueing systems with collisions and non-reliable server using MOSEL. *International Conference on Distributed Computer and Communication Networks*, Springer, Berlin, 248–258.
- Bhat, U. N. (2015). *An Introduction to Queueing Theory. Modeling and Analysis in Applications*. Birkhäuser, Boston.
- Bossel, H. (2013). *Modeling and Simulation*. Springer-Verlag, Berlin.
- Cao, Y., Khosla, D., Chen, Y., Huber, D. J. (2018). System and method for real-time collision detection. US Patent 9,934,437.
- Choi, B.D., Shin, Y.W., Ahn, W.C. (1992). Retrial queues with collision arising from unslotted CSMA/CD protocol. *Queueing Syst.*, 11(4), 335–356.
- Dragieva, V. I. (2014). Number of retrials in a finite source retrial queue with unreliable server. *Asia-Pac. J. Oper. Res.*, 31(2), 23.
- Falin, G., Artalejo, J. (1998). A finite source retrial queue. *European Journal of Operational Research*, 108, 409–424.
- Falin, G., Templeton, J.G.C. (1997), *Retrial Queues*. Chapman and Hall, London.
- Fiems, D., Phung-Duc, T. (2017), Light-traffic analysis of random access systems without collisions. *Annals of Operations Research*, 277(2), 311–327
- Gharbi, N., Dutheillet, C. (2011). An algorithmic approach for analysis of finite-source retrial systems with unreliable servers. *Computers & Mathematics with Applications*, 62(6), 2535–2546.
- Gharbi, N., Ioualalen, M. (2006). GSPN analysis of retrial systems with servers breakdowns and repairs. *Applied Mathematics and Computation*, 174(2), 1151–1168.
- Gharbi, N., Mokdad, L., Ben-Othman, J. (2015). A performance study of next generation cellular networks with base stations channels vacations. *Global Communications Conference (GLOBECOM)*, IEEE, 1–6.
- Gharbi, N., Nemmouchi, B., Mokdad, L., Ben-Othman, J. (2014). The impact of breakdowns disciplines and repeated attempts on performances of small cell networks. *Journal of Computational Science*, 5(4), 633–644.
- Gómez-Corral, A., Phung-Duc, T. (2016). Retrial queues and related models. *Annals of Operations Research*, 247(1), 1–2.
- Harchol-Balter, M. (2013). *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*. Cambridge University Press, Cambridge.

- Ikhlef, L., Lekadir, O., Aïssani, D. (2016). MRSPN analysis of semi-Markovian finite source retrial queues. *Annals of Operations Research*, 247(1), 141–167.
- Jinsoo, A., Kim, Y., Kwak, J., Son, J. (2018). Wireless communication method for multi-user transmission scheduling, and wireless communication terminal using same. US Patent App. 15/736,968.
- Kim, J., Kim, B. (2016). A survey of retrial queueing systems. *Annals of Operations Research*, 247(1), 3–36.
- Kim, J.S. (2010). Retrial queueing system with collision and impatience. *Communications of the Korean Mathematical Society*, 25(4), 647–653.
- Kobayashi, H., Mark, B. L. (2009). *System Modeling and Analysis: Foundations of System Performance Evaluation*. Pearson Education India, Chennai.
- Krishnamoorthy, A., Pramod, P.K., Chakravarthy, S.R. (2014). Queues with interruptions: a survey. *TOP*, 22(1), 290–320.
- Kuki, A., Bérczes, T., Sztrik, J., Kvach, A. (2019). Numerical analysis of retrial queueing systems with conflict of customers. *Journal of Mathematical Sciences*, 237, 637–683.
- Kulkarni, V.G. (2016). *Modeling and Analysis of Stochastic Systems*. CRC Press, Boca Raton.
- Kumar, B.K., Vijayalakshmi, G., Krishnamoorthy, A., Basha, S.S. (2010). A single server feedback retrial queue with collisions. *Computers & Operations Research*, 37(7), 1247–1255.
- Kvach, A. (2014). Numerical research of a Markov closed retrial queueing system without collisions and with the collision of the customers. *Proceedings of Tomsk State University*. TSU Publishing House, Tomsk, 105–112 (in Russian).
- Kvach, A., Nazarov, A. (2015a). Numerical research of a closed retrial queueing system M/GI/1//N with collision of the customers. *Proceedings of Tomsk State University*. TSU Publishing House, Tomsk, 65–70 (in Russian).
- Kvach, A., Nazarov, A. (2015b). *Sojourn Time Analysis of Finite Source Markov Retrial Queueing System with Collision*. Springer International Publishing, Cham.
- Kvach, A., Nazarov, A. (2015c). The research of a closed RQ-system M/GI/1//N with collision of the customers in the condition of an unlimited increasing number of sources. *Probability Theory, Random Processes, Mathematical Statistics and Applications*, 65–70 (in Russian).
- Kwak, B.-J., Rhee, J.-K., Kim, J., Kyounghee, K. (2018). Random access method and terminal supporting the same. US Patent 9,954,754.
- Lakaour, L., Aïssani, D., Adel-Aissanou, K., Barkaoui, K. (2018). M/M/1 retrial queue with collisions and transmission errors. *Methodology and Computing in Applied Probability*, 1–12.
- Lakatos, L., Szeidl, L., Telek, M. (2013). *Introduction to Queueing Systems with Telecommunication Applications*. Springer, New York.

- Law, A.M., Kelton, W.D. (1991). *Simulation Modeling and Analysis*. McGraw-Hill, New York.
- Nazarov, A., Kvach, A., Yampolsky, V. (2014). *Asymptotic Analysis of Closed Markov Retrial Queuing System with Collision*. Springer International Publishing, Cham.
- Nazarov, A., Moiseeva, S.P. (2006). *Methods of Asymptotic Analysis in Queueing Theory*. NTL Publishing House of Tomsk University, Tomsk (in Russian).
- Nazarov, A., Sudyko, E. (2010). Method of asymptotic semi-invariants for studying a mathematical model of a random access network. *Probl. Inf. Transm.*, 46(1), 86–102.
- Nazarov, A., Sztrik, J., Kvach, A. (2017a). Comparative analysis of methods of residual and elapsed service time in the study of the closed retrial queuing system M/GI/1/N with collision of the customers and unreliable server. *International Conference on Information Technologies and Mathematical Modelling*, 97–110.
- Nazarov, A., Sztrik, J., Kvach, A. (2017b). Some features of a finite-source M/GI/1 retrial queuing system with collisions of customers. *International Conference on Distributed Computer and Communication Networks*, 186–200.
- Nazarov, A., Sztrik, J., Kvach, A. (2017c). Some features of a finite-source M/GI/1 retrial queuing system with collisions of customers. *International Conference on Distributed Computer and Communication Networks*, 79–86.
- Nazarov, A., Sztrik, J., Kvach, A. (2018). A survey of recent results in finite-source retrial queues with collisions. In *Information Technologies and Mathematical Modelling. Queueing Theory and Applications*, Dudin, A., Nazarov, A., Moiseev, A. (eds). Springer, Berlin.
- Nazarov, A., Sztrik, J., Kvach, A., Bérczes, T. (2018). Asymptotic analysis of finite-source M/M/1 retrial queueing system with collisions and server subject to breakdowns and repairs. *Annals of Operations Research*, 277(2), 213–229.
- Nazarov, A., Sztrik, J., Kvach, A., Tóth, A. (2018). Asymptotic sojourn time analysis of Markov finite-source M/M/1 retrial queueing system with collisions and server subject to breakdowns and repairs. *Probability in the Engineering and Informational Sciences*, 33(3), 387–403.
- Nazarov, A., Terpugov, A. (2004). *Theory of Mass Service*. NTL Publishing House of Tomsk University, Tomsk (in Russian).
- Peng, Y., Liu, Z., Wu, J. (2014). An M/G/1 retrial G-queue with preemptive resume priority and collisions subject to the server breakdowns and delayed repairs. *J. Appl. Math. Comput.*, 44(1–2), 187–213.
- Reith III, H. C. (2017). System and method for collision detection and avoidance for network communications. US Patent 9,553,828.
- Roszik, J. (2004). Homogeneous finite-source retrial queues with server and sources subject to breakdowns and repairs. *Ann. Univ. Sci. Budap. Rolando Eötvös, Sect. Comput.*, 23, 213–227.

- Rubinstein, R. Y., Kroese, D. P. (2016). *Simulation and the Monte Carlo Method*. John Wiley & Sons, Hoboken.
- Stewart, W. J. (2009). *Probability, Markov Chains, Queues, and Simulation: The Mathematical Basis of Performance Modeling*. Princeton University Press, Princeton.
- Sztrik, J. (2005). Tool supported performance modelling of finite-source retrial queues with breakdowns. *Publicationes Mathematicae*, 66, 197–211.
- Sztrik, J., Almási, B., Roszik, J. (2006). Heterogeneous finite-source retrial queues with server subject to breakdowns and repairs. *Journal of Mathematical Sciences*, 132, 677–685.
- Takeda, T., Yoshihiro, T. (2017). A distributed scheduling through queue-length exchange in CSMA-based wireless mesh networks. *Journal of Information Processing*, 25, 174–181.
- Tóth, Á., Bérczes, T., Sztrik, J., Kuki, A. (2019). Comparison of two operation modes of finite-source retrial queueing systems with collisions and non-reliable server by using simulation. *Journal of Mathematical Sciences*, 237, 846–857.
- Tóth, Á., Bérczes, T., Sztrik, J., Kvach, A. (2017). Simulation of finite-source retrial queueing systems with collisions and a non-reliable server. *International Conference on Distributed Computer and Communication Networks*, 146–158. Springer.
- Wang, J., Zhao, L., Zhang, F. (2010). Performance analysis of the finite source retrial queue with server breakdowns and repairs. *Proceedings of the 5th International Conference on Queueing Theory and Network Applications*, ACM, 169–176.
- Wang, J., Zhao, L., Zhang, F. (2011). Analysis of the finite source retrial queues with server breakdowns and repairs. *Journal of Industrial and Management Optimization*, 7(3), 655–676.
- Wentink, M. (2017). Collision avoidance systems and methods. US Patent App. 15/401,606.
- Wüchner, P., Sztrik, J., de Meer, H. (2010). Finite-source retrial queues with applications. *Proceedings of 8th International Conference on Applied Informatics*, Eger, Hungary, 275–285.
- Yao, J. (2016). *Asymptotic Analysis of Service Systems with Congestion-Sensitive Customers*. Columbia University, New York City.
- Yeo, G. M., Kim, Y. I., Park, D. G., Song, S. Y., Lee, Y. T., Lee, H. W. (2017). Control method and apparatus for collision avoidance in low-power wireless sensor communication. US Patent App. 15/361,194.
- Zhang, F., Wang, J. (2013). Performance analysis of the retrial queues with finite number of sources and service interruptions. *Journal of the Korean Statistical Society*, 42(1), 117–131.