



الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي



Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة الشاذلي بن جديد - الطارف
Université Chadli Bendjedid – El Tarf
كلية العلوم و التكنولوجيا
Faculté des Sciences et de la Technologie
قسم الرياضيات
Département de Mathématiques

Mémoire de fin d'études

En vue de l'obtention du diplôme de Master

Domaine : Mathématiques et Informatique

Filière : Mathématiques

Spécialité: Analyse fonctionnelle et calcul stochastique

Thème

**Analyse stochastique d'un centre d'appels par
La Troncature Finie.**

Présenté par:

Cheloufi Radhia

Devant le Jury :

Dr. Grine Razika	MCB	Univ Chadli Bendjedid El Tarf	Présidente
Dr. Zidani Nesrine	MCB	Univ Chadli Bendjedid El Tarf	Rapporteur
Dr. Benseghir Rym	MCB	Univ Chadli Bendjedid El Tarf	Examinatrice

Année Universitaire 2021-2022

Table des matières

Introduction	3
1 Les notions préliminaires	4
1.1 Processus stochastique	4
1.1.1 Définitions et propriétés de base	4
1.1.2 Processus de comptage	5
1.1.3 Processus de Poisson	6
1.1.4 Processus de naissance et de mort	8
1.2 Classification des systèmes de files d'attente	10
1.3 Quelques exemples de systèmes d'attente	12
1.4 Analyse mathématique	13
1.5 Système de files d'attente classique	14
1.5.1 Le modèle $M/M/1$	14
1.5.2 Le modèle $M/M/C$	15
2 Les systèmes de files d'attente avec rappels	19
2.1 Introduction	19
2.2 Terminologie et notation	19
2.3 Politique d'accès au serveur à partir de l'orbite	20
2.4 Exemples des systèmes de files d'attente avec rappels	21
2.5 Modèle général	25
2.6 Modèles markoviens	26
2.6.1 système de file d'attente avec rappels $M/M/1$	26

2.6.2	Système de files d'attente avec rappels $M/M/C$	28
2.7	Approche d'analyse des modèles à multiserveurs avec rappels	30
2.7.1	Introduction	30
2.7.2	Modèles tronqués	30
2.7.3	Méthode numérique	32
3	Application aux centres d'appels	36
3.1	Introduction	36
3.2	Les centres d'appels	37
3.3	Description mathématique du modèle	38
3.4	Condition d'existence d'un régime stationnaire	39
3.5	Régime stationnaire	43
3.6	Troncature finie	45
	Annexe	51
	Bibliographie	56

Liste des Figures

1.1.1 Processus de comptage	6
1.5.1 Schéma général d'un système classiques M/M/C	16
2.5.1 Schéma général d'un système avec rappels	26
2.6.1 Graphe des transitions du modèle M/M/1 avec rappels	27
2.6.2 Graphe des transitions du modèle M/M/C avec rappels	29

Dédicaces

Je dédie ce modeste travail

À ma chère mère,

À mon cher père,

À mon cher Yacine,

À tous les membres de ma famille,

À tous ceux qui m'aiment et me respectent...

Remerciement

*Je remercie **Allah** le tout puissant de m'avoir donné la santé et la volonté d'entamer et de terminer ce mémoire.*

*Tout d'abord, ce travail ne serait pas aussi riche et n'aurait pas pu voir le jour sans l'aide et l'encadrement de **Dr N. Zidani**, je la remercie pour la qualité de son encadrement exceptionnel, pour sa patience, sa rigueur et sa disponibilité durant ma préparation de ce mémoire.*

*Mes remerciements s'adressent à **Yacine** à qui était le mérite de réaliser mon rêve académique.*

*Mes sincères remerciements s'adressent à **Zina, Hadjer, Rym, Rawia** et toutes mes collègues de la promotion pour leurs aides.*

*Mes remerciements s'adressent également à mes collègues de travail **Rawda, Manel** et **Rokaya** pour leur soutien moral et leurs encouragements;*

Aux membres de jury qui m'ont honoré par leur participation à l'examen de ce travail.

Résumé

Dans ce travail, on s'intéresse à l'étude et l'analyse stochastique des performances du système d'attente au niveau d'un centre d'appels téléphoniques. À cet effet, nous analysons un système de files d'attente avec rappels et multiserveurs $M/M/C/K$.

En premier lieu, nous avons accordé une attention particulière aux éléments essentiels et quelques résultats concernant les modèles de files d'attente classiques markoviens.

En second lieu, nous avons pu présenter et étudier deux modèles de systèmes de files d'attente $M/M/1$ et $M/M/C$ avec rappels. Puis, nous avons décrit une des méthodes d'approximation la plus utilisée pour résoudre ce type de systèmes.

Enfin, nous avons analysé les performances du système d'attente au niveau d'un centre d'appels téléphoniques par la troncature Finie en résolvant un système multiserveurs, avec rappels, arrivées négatives, espace d'attente et abandons exponentiels.

Mots-clés: système de file d'attente, processus stochastique, Troncature Finie, centre d'appels.

Abstract

In this work, we are interested in the study and the stochastic analysis of the performance of the retrials queueing system at the level of a telephone call center. For this purpose, we analyze a retrials queueing system and multiservers $M/M/C/K$.

First, we paid particular attention to the essential elements and some results concerning classical Markovian queueing models.

Secondly, we were able to present and study two models of $M/M/1$ and $M/M/C$ retrials queueing systems. Then, we described one of the most used approximation methods to solve this type of system.

Finally, we analyzed the performance of the waiting system at the level of a telephone call center by finite truncation by solving a retrials multiserver system with waiting space, negative arrivals and exponential abandonment.

Keyword: queueing system, stochastic process, Finite Truncation, call center..

المخلص

في هذا العمل ، نهتم بالدراسة والتحليل العشوائي لأداء نظام الانتظار على مستوى مركز الاتصال الهاتفي. لهذا الغرض ، نقوم بتحليل نظام قائمة انتظار مع عمليات الاسترجاعات والخوادم المتعددة $M / M / C / K$.

أولاً ، أولينا اهتمامًا خاصًا بالعناصر الأساسية وبعض النتائج المتعلقة بنماذج طابور ماركوفيان الكلاسيكية.

ثانيًا ، تمكنا من تقديم ودراسة نموذجين لأنظمة انتظار $M / M / 1$ و $M / M / C$ مع عمليات الاسترجاعات. بعد ذلك ، وصفنا إحدى طرق التقريب الأكثر استخدامًا لحل هذا النوع من النظام.

أخيرًا ، قمنا بتحليل أداء نظام الانتظار على مستوى مركز الاتصال الهاتفي عن طريق الاقتطاع المحدود عن طريق حل نظام متعدد الخوادم ، مع عمليات الاسترداد ، والوصول السلبي ، ومساحة الانتظار ، والتسرب الأسي.

الكلمات المفتاحية: نظام الطابور ، العملية العشوائية ، الاقتطاع المحدود ، مركز الاتصال.

Introduction générale

Le monde contemporain est de plus en plus soucieux de satisfaire les attentes des clients, tellement il y a plusieurs choses dans la vie pour lesquelles il faut attendre. Un spécialiste de l'analyse d'attente a déterminé qu'une personne adulte passait au minimum le dixième de son temps à attendre. On attend les autobus, les ascenseurs, dans les magasins, les cinémas, les stations d'essence, chez le dentiste, etc. Toutefois, lorsqu'on parle d'attente, on pense souvent à des personnes, or, les clients en attente sont aussi des commandes en attente de traitement, des machines en attente de réparation, des programmes d'ordinateurs qui attendent d'être exécutés. Tout cela se passe en posant la fameuse question, pourquoi y a-t-il attente ? Sur cela, introduisons la notion des files d'attente.

La théorie des files d'attente est l'un des outils analytiques les plus puissants pour la modélisation de systèmes de logistiques et de télécommunications. Cette théorie a pour objet l'étude de systèmes et réseaux où des entités, appelées clients, cherchent à accéder à des ressources, généralement limitées, afin d'en obtenir un service. Cette théorie des files d'attente remontent à 1909 à l'époque où Danois Erlang a posé les bases dans ses recherches sur le trafic téléphonique de Copenhague. Ces travaux ont par la suite été intégrés à la recherche opérationnelle. Malheureusement, les publications sur la théorie des files d'attente ont adopté un langage de plus en plus mathématique, ce qui a freiné son utilisation. La situation a toutefois changé quand des chercheurs ont commencé à appliquer la théorie des files d'attente à l'évaluation des performances. Pour ce type d'applications, il est apparu que même des modèles relativement simples fournissaient des résultats qui correspondaient de près aux observations réelles. On assistera alors à une

évolution rapide de la théorie des files d'attente.

La théorie des files d'attente classique offre deux possibilités pour résoudre le conflit qui apparaît lorsqu'un client arrive dans le système à serveur occupé: soit le client quitte le système sans recevoir le service, soit il prend place dans une file d'attente. Une possibilité alternative est de permettre au client de répéter sa demande de service après une durée de temps aléatoire. Entre deux appels successives (rappels), le client en question est "en orbite". Un tel système est appelé système de files d'attente avec rappels.

Prenons par exemple le cas d'un centre d'appels téléphoniques, dans les agents sont les personnes qui répondent aux appels des clients. Lorsqu'un client passe un appel téléphonique, s'il y a un agent d'appel inactif, l'agent d'appel répond immédiatement à un client. Si tous les agents sont occupés, le client peut entendre des messages tels que "le système est occupé en ce moment, veuillez patienter un moment".

À ce moment, le client raccroche immédiatement ou continue d'entendre le message. Dans le premier cas, le client peut réessayer après un certain temps. Les clients qui décident d'attendre un agent d'appel libre peuvent revenir si le temps d'attente est trop long. Ces clients peuvent également passer un appel plus tard.

L'étude de tel système de files d'attente consiste à calculer ces paramètres de performance, afin d'évaluer son rendement et améliorer son fonctionnement (minimiser le temps d'attente et le temps d'inactivité de l'installation) de savoir par exemple si le nombre de serveurs dans le système est adéquat pour gérer le flux de demandes ou encore d'appréhender les effets d'une modification des conditions de fonctionnement, et ainsi prendre des décisions sur le nombre minimum de ressources nécessaires.

La littérature sur les files d'attente contient une quantité importante de travail consacré à l'analyse des performances des systèmes de files d'attente avec rappels à un

seul serveur et à multiserveurs. Une analyse exacte d'un tel système est complexe et peut être intraitable. C'est pourquoi, dans ce travail nous nous concentrons sur l'analyse d'un centre d'appels en tant que système de files d'attente avec rappels, multiserveurs, arrivées négatives, espace d'attente et abandons exponentiels. Pour le modèle à l'étude, nous trouvons la condition d'ergodicité et les caractéristiques de performance essentielles. À cette fin, nous appliquons la méthode de Troncation Finie.

La mémoire est composée d'une introduction générale, trois chapitres, d'une conclusion et annexe.

Dans le premier chapitre, nous présentons les notions de base des systèmes de files d'attente, certaines définitions et notations qui sont nécessaires dans l'étude des systèmes (Notation de Kendall, la loi de Little, ...). Dans le deuxième chapitre, nous donnons une synthèse sur la théorie de files d'attente avec rappels et nous étudions quelques modèles de files d'attente avec rappels ($M/M/1$ et $M/M/C$). Puis, nous nous intéressons aux modèles tronqués des systèmes de files d'attente avec rappels et multiserveurs. Le troisième chapitre est consacré à l'application de la méthode de troncation sur un centre d'appels.

En conclusion, nous résumons les principales contributions de la mémoire, et quelques idées de recherche sont suggérées et pourraient faire avancer le domaine des systèmes de file d'attente avec rappels.

1.1 Processus stochastique

Les processus stochastiques décrivent l'évolution d'une grandeur aléatoire en fonction du temps. Il existe de nombreuses applications des processus aléatoires notamment en physique statistique [14] (par exemple le ferromagnétisme, les transitions phases,...etc), en biologie (évolution, génétique et génétique des populations), médecine (croissance de tumeurs, épidémie), et bien entendu les sciences de l'ingénieur.

Dans ce dernier domaine, les applications principales sont pour l'administration des réseaux, de l'internet, des télécommunications et bien entendu dans les domaines économique et financier. L'étude des processus stochastiques s'insère dans la théorie des probabilités dont elle constitue l'un des objectifs les plus profonds. Elle soulève des problèmes mathématiques intéressants et souvent très difficiles.

1.1.1 Définitions et propriétés de base

Définition 1.1.1 *Un processus stochastique est une famille de variables aléatoires X_t , $t \in T$ ou chaque variable aléatoire X_t est indexé par le paramètre $t \in T$, si T est un ensemble de \mathbb{R}_+ , alors t signifie temps.*

Généralement X_t représente l'état du processus stochastique au temps t [23].

- Si T est dénombrable, i.e $T \subseteq \mathbb{N}$, alors nous disons que $X_t, t \in T$ est un processus à temps discret.
- Si T est un intervalle de $[0, \infty[$, alors le processus stochastique est dit un processus à temps continu.

L'ensemble des valeurs de X_t est appelé l'espace d'état, qui peut également être soit discret (fini ou infini dénombrable) ou continu (un sous-ensemble de \mathbb{R} ou \mathbb{R}^n), donc nous écrivons $(X_n)_{n \geq 0}$ pour le processus à temps discret et $(X_t)_{t \geq 0}$ pour le processus à temps continu.

1.1.2 Processus de comptage

Définition 1.1.2 Un processus stochastique $(N_t, t \in \mathbb{R}^+)$, est un processus de comptage si (N_t) représente le nombre total d'événements qui se sont produits entre 0 et t , il doit donc satisfaire:

- $(N_t) \geq 0$;
- (N_t) a des valeurs entières uniquement;
- pour $s < t, N(t) - N(s)$ est le nombre d'événements qui ont eu lieu entre s et t .

Définition 1.1.3 (Processus à accroissements indépendants)

Un processus $\{X_t\}$ tel que $X_0 = 0$ est à accroissements indépendants si pour tout suite finie $0 < t_1 < t_2 < t_3 \dots < t_n$ les variables aléatoires $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépendantes.

Définition 1.1.4 Un processus à accroissements indépendants est à accroissements stationnaires si la loi de l'accroissement $(X_{t+s} - X_t)$ ne dépend pas de t pour tout $t \geq 0$.

Définition 1.1.5 Un processus de comptage $\{N_t, t \geq 0\}$ est un processus de Poisson d'intensité $\lambda > 0$ si:

- $N(0) = 0$;
- le processus est à accroissements stationnaires;
- le processus est à accroissements indépendants;
- $\forall 0 < s < t$, la variable aléatoire $N(t) - N(s)$ suit une loi de Poisson de paramètre $\lambda(t - s)$.

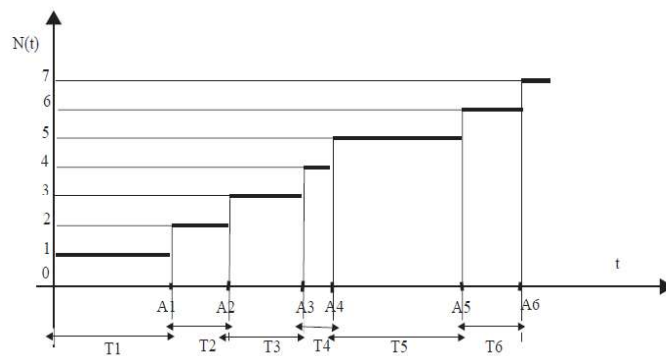


Figure 1.1.1 : Processus de comptage

1.1.3 Processus de Poisson

De nombreux phénomènes aléatoires se manifestent par des "arrivées" survenant une par une à des instants aléatoires successifs.[7, 13]

Exemple 1.1.1

- arrivées d'appels à un central téléphonique;
- impacts de micrométéorites sur un satellite;
- passage de véhicules à un péage d'autoroute;
- arrivée des clients à un guichet, occurrence d'accidents dans une ville, pannes de machines dans une usine...etc.

De tels phénomènes peuvent se définir par la famille $(A_n)_{n \in \mathbb{N}^*}$ des temps d'arrivées qui sont des variables aléatoires [16]. Mais on peut aussi le faire à partir du processus de comptage $(N_t)_{t \in \mathbb{R}}$, où N_t est le nombre d'événements apparus jusqu'à l'instant t . Où $N_{t+u} - N_u$ est le nombre d'événements apparus entre u et $u + t$.

L'espace des états du processus $(N_t)_{t \in \mathbb{R}_+}$ est $E = \mathbb{N}$ et l'espace des temps est $T = \mathbb{R}_+$. Le processus qui modélise convenablement les exemples cités est le processus de Poisson.

On conviendra que $N_0 = 0$.

On note :

- A_n l'instant de réalisation du $n^{\text{ième}}$ événement;
- T_n la durée séparant le $(n - 1)^{\text{ième}}$ événement du $n^{\text{ième}}$ événement pour $n \geq 2$ et $T_1 = A_1$.

On a :

- $A_n = T_1 + T_2 + \dots + T_n$ pour tout $n \in \mathbb{N}^*$;
- $T_1 = A_1$ et $T_n = A_n - A_{n-1}$ pour tout $n \geq 2$.

Ainsi, la connaissance de la famille $(A_n)_{n \in \mathbb{N}^*}$ équivaut à celle de la famille $(T_n)_{n \in \mathbb{N}^*}$.

D'autre part, $A_n \geq t$ signifie que le $n^{\text{ième}}$ événement a eu lieu à l'instant t ou avant, c'est-à-dire qu'à l'instant t , au moins n événements ont eu lieu, c'est-à-dire que $N_t \geq n$.

Ainsi

$$F_{A_n}(t) = P([A_n \leq t]) = P([N_t \geq n]) = 1 - \sum_{k=0}^{n-1} P([N_t = k])$$

et

$$P([N_t = n]) = P([N_t \geq n]) - P([N_t \geq n + 1]) = P([A_n \leq t]) - P([A_{n+1} \leq t]).$$

Par conséquent, la connaissance de $(N_t)_{t \in \mathbb{R}_+}$ équivaut à celle de la famille $(A_n)_{n \in \mathbb{N}^*}$.

Définition 1.1.6 Un processus est dit **stationnaire** (ou homogène dans le temps), si pour tout s et pour tout t , l'accroissement $N_{t+s} - N_s$ a même loi que N_t .

Définition 1.1.7 Un processus à accroissements indépendants stationnaire $(N_t)_{t \in \mathbb{R}_+}$ est dit à événements rares si:

$$\lim_{h \rightarrow 0^+} P([N_h > 0]) = 0 \text{ et si } \lim_{h \rightarrow 0^+} \frac{P([N_h > 1])}{P([N_h = 1])} = 0.$$

Définition 1.1.8 : Un processus de comptage $(N_t)_{t \in \mathbb{R}_+}$ tel que $N_0 = 0$ est un processus de Poisson si et seulement si:

- C_1 : $(N_t)_{t \in \mathbb{R}_+}$ est stationnaire;
- C_2 : $(N_t)_{t \in \mathbb{R}_+}$ est un processus à accroissements indépendants;
- C_3' : il existe $\lambda \geq 0$ tel que, pour tout $t \geq 0$, la variable aléatoire N_t suit la loi de Poisson de paramètre λt .

Théorème 1.1.1 [19]

Si un processus de comptage $(N_t)_{t \geq 0}$ satisfait C_1 , C_2 et C_3 alors:

$$P(N(t) = k) = p_k(t) = \frac{(\lambda t)^k}{k!} \exp^{-\lambda t}, \quad \lambda > 0, \quad k > 0,$$

$$E[N(t)] = \lambda t,$$

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

Théorème 1.1.2 [19]

Le temps T qui sépare un instant quelconque du prochain évènement est une variable exponentielle de paramètre λ .

1.1.4 Processus de naissance et de mort

Définition 1.1.9 On peut réaliser un processus de naissance et de mort de la façon suivante:

- Les arrivées et les départs d'entités obéissent à des lois exponentielles de taux respectifs $\lambda(n)$ et $\mu(n)$.

- A l'aide d'hypothèse de régularité : deux événements ne peuvent pas se produire en même temps, donc la probabilité que deux événements se produisent dans un intervalle de temps dt est négligeable.
- Il y a une transition vers un état voisin, soit par l'arrivée d'un client (naissance), soit par le départ d'un client (mort).

Alors, soit $\{N(t), t \geq 0\}$ un processus de naissance et de mort à états discrets et homogène dans le temps, c'est-à-dire:

$$P(N(t+s) = j / N(s) = j) = P_{ij}(t) \text{ ne dépend pas de } s.$$

Ce processus est de naissance et de mort si:

$$\begin{aligned} P_{i,i+1}(\Delta t) &= \lambda_i \Delta t + o(\Delta t); \quad i \geq 0; \\ P_{i,i-1}(\Delta t) &= \mu_i \Delta t + o(\Delta t); \quad i \geq 1; \\ P_{i,i}(\Delta t) &= 1 - (\lambda_i + \mu_{i-1}) \Delta t + o(\Delta t); \quad i \geq 0; \\ P_{i,j}(\Delta t) &= o(\Delta t); \quad |i - j| \geq 2; \\ P_{i,j}(0) &= \delta_{ij} = \begin{cases} 1, & \text{si } i = j. \\ 0, & \text{si } i \neq j. \end{cases} \end{aligned}$$

λ_i et μ_i sont appelés taux de transition (taux de naissance et de mort).

Exemple 1.1.2 :

Dans une file d'attente le temps s'écoulant entre deux arrivées consécutives est distribué exponentiellement ainsi que le temps de service.

Alors cette file d'attente peut être modélisée par un processus de naissance et de mort:

- naissance \longrightarrow arrivée du client.
- mort \longrightarrow départ du système après son service.

1.2 Classification des systèmes de files d'attente

La classification des systèmes des files d'attente se base principalement sur les trois éléments suivants: la nature stochastique du processus des arrivées, le mécanisme de service et la discipline d'attente.[11]

Le processus des arrivées spécifie les instants auxquels les clients arrivent dans les système. Ce processus peut être régulier c'est-à-dire à chaque unité du temps, on a une arrivée, ou aléatoire (processus de Poisson).

Le mécanisme de service comprend le nombre de serveurs et la distribution des durées de service. Dans la majorité des cas, on suppose que la populsion des clients est homogène, c'est-à-dire que les temps de service nécessaire au traitement des clients sont identiquement distribués selon une loi de probabilité connue. En particulier, on rencontre la distribution exponentielle qui est la plus simple à manipuler mathématiquement. Sa propriété la plus importante est l'absence de mémoire, qui pourrait être caractérisée par le fait que le temps résiduel d'un service est indépendant du temps déjà écoulé de ce dernier.

La capacité de l'espace d'attente peut être illimitée ou non. Dans le second cas, certains client qui arrivent vers le système n'ont pas la possibilité d'y entrer.

Pour la classification des systèmes d'attente, on a recours à la notation symbolique induite par Kendall au début des années cinquante. Cette notation comprend six symboles rangés dans l'ordre $T/Y/C/K/m/Z$,

où

1. T : indique le processus d'arrivée des clients. Les symboles utilisés sont:
 - M : Inter-arrivées des clients sont identiquement distribuées selon une loi exponentielle. Il correspond à un processus de Poisson ponctuel (propriété sans mémoire).

- D : Les temps inter-arrivées des clients ou les temps de service sont constants et toujours les mêmes.
 - GI : Inter-arrivées des clients ont une distribution générale (il n'y a aucune hypothèse sur la distribution mais les inter-arrivées sont indépendantes et identiquement distribuées).
 - G : Inter-arrivées des clients ont une distribution générale et peuvent être dépendantes.
 - E_k : Ce symbole désigne un processus où les intervalles de temps entre deux arrivées successives sont des variables aléatoires indépendantes et identiquement distribuées suivant une loi d'Erlang d'ordre k .
2. Y : décrit la distribution des temps de service d'un client. Les codes sont les mêmes que T .
 3. C : nombre de serveurs.
 4. K : capacité de la file. c'est le nombre de places dans le système en d'autre terme c'est le nombre maximal de clients dans le système y compris ceux en service.
 5. m : population des usagers.
 6. Z : discipline de service c'est la façon dont les clients sont ordonnés pour être servi. Les codes utilisés sont les suivants:
 - **FIFO** (first in, first out) ou FCFS (first come, first served): c'est la file standard dans laquelle les clients sont servis dans leur ordre d'arrivée. Notons que les disciplines FIFO et FCFS ne sont pas équivalentes lorsque la file contient plusieurs serveurs. Dans la première, le premier client arrivé sera le premier à quitter la file alors que dans la deuxième, il sera le premier à commencer son service. Rien n'empêche alors qu'un client qui commence son service après lui, dans un autre serveur, termine avant lui.
 - **LIFO** (last in, first out) ou LCFS (last come, first served). Cela correspond à une pile, dans laquelle le dernier client arrivé (donc posé sur la pile) sera le

premier traité (retiré de la pile). A nouveau, les disciplines LIFO et LCFS ne sont équivalentes que pour une file mono serveur.

- **SIRO** (Served In Random Order), les clients sont servis aléatoirement.
- **PNPN** (Priority service), les clients sont servis selon leur priorité. Tous les clients de la plus haute priorité sont servis premiers, puis les clients de priorité inférieur sont servis, et ainsi de suite.
- **PS** (Processor Sharing), les clients sont servis de manière égale. La capacité du système est partagée entre les clients.

Remarque 1.2.1 Dans sa version courte, seuls les trois premiers symboles $T/Y/C$ sont utilisés. Dans un tel cas, on suppose que la file est régie par une discipline FIFO et que le nombre de places d'attente ainsi que celui des clients susceptibles d'accéder au système sont illimités.

1.3 Quelques exemples de systèmes d'attente

- $T/Y/C$: 1 file et C stations offrant le même service; le client qui attend va à la première qui se libère.

Exemple 1.3.1 : guichets d'une poste...

- $T/Y/\infty$: infinité de serveurs; pas d'attente.

Exemple 1.3.2 : lignes téléphoniques.

- Systèmes à perte $T/Y/C/C$ pas d'attente.

Exemple 1.3.3 : Central téléphonique où les appels non traités sont rejetés.

- Systèmes $T/Y/C/K$ où $K > C$: longueur de la file limitée, dépendant de la capacité.

Exemple 1.3.4 : Station service, salle d'attente.

1.4 Analyse mathématique

L'étude mathématique d'un système d'attente se fait par l'introduction d'un processus stochastique qui décrit l'évolution temporelle du système pour cela on s'intéresse d'abord au nombre $X(t)$ de clients se trouvant dans le système à l'instant t ($t \geq 0$) et on cherche à calculer :

- les probabilités d'état $P_n(t) = P(X(t) = n)$ qui définissent le régime transitoire du processus.
- le régime stationnaire du processus est défini par $P_n = \lim_{t \rightarrow \infty} P_n(t)$.

Mesures de performance

À partir de la distribution stationnaire du processus $\{X(t); t \geq 0\}$, on peut calculer les caractéristiques du système telles que: le temps d'attente d'un client, le temps de séjour d'un client dans le système, le taux d'occupation des dispositifs de service, la durée de la période d'activité, et également les mesures de performance suivantes:

- nombre moyen de clients dans le système \bar{n} ;
- nombre moyen de clients dans la file d'attente \bar{n}_f ;
- temps moyen d'attente d'un client dans le système \bar{w} ;
- temps moyen de séjour d'un client dans le système \bar{w}_s .

Formules de Little

La loi de Little est une relation qui s'applique à une grande classe de systèmes, la seule condition d'application de la loi de Little est que le système soit stable.

$$\bar{n} = \lambda_e \bar{w}_s;$$

$$\bar{n}_f = \lambda_e \bar{w};$$

$$\bar{w}_s = \bar{w} + \frac{1}{\mu};$$

$$\bar{n} = \bar{n}_f + \frac{\lambda_e}{\mu}.$$

Où λ_e et $1/\mu$ sont respectivement le taux d'arrivées des clients et la durée moyenne de service ($n > 0$). Une autre mesure importante d'un système de files d'attente, celle qui mesure le degré de saturation du système, est l'intensité du trafic ρ . Elle est définie par:

$$\rho = \frac{\text{temps moyen de service}}{\text{temps moyen entre deux arrivées}} = \frac{\lambda}{\mu}.$$

1.5 Système de files d'attente classique

1.5.1 Le modèle $M/M/1$

Le modèle $M/M/1$ est un modèle d'attente markovien le plus simple. Il permet en effet d'illustrer les concepts fondamentaux liées à l'attente devant un serveur, le flot des arrivées est Poissonien de taux λ et le temps qu'occupe le serveur pour satisfaire le client suit une loi exponentielle du paramètre $\mu > 0$, la discipline d'attente est FIFO. Dans ce cas la file d'attente $M/M/1$ est de capacité infinie, le processus du modèle est un processus de naissance et de mort dont les taux de transitions sont:

$$\lambda_n = \lambda, n \geq 0$$

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

Régime transitoire

Pour calculer les probabilités d'états $p_n(t) = p[X(t) = n]$ nous pouvons écrire d'après le théorème de probabilités totales et pour $n > 1$

$$\begin{aligned} P_n(\Delta t + t) &= \sum_{i \geq 0} P_i P_{in}(\Delta t) \\ &= P_{n-1} P_{n-1,n}(\Delta t) + P_n(t) P_{nn}(\Delta t) + P_{n+1}(t) P_{n+1,n}(\Delta t) \\ &= \lambda_{n-1} P_{n-1}(\Delta t) + (1 - (\lambda_n + \mu_n)) P_n(t) \Delta t + \mu_{n+1} P_{n+1}(\Delta t) + o(\Delta t) \\ &\Rightarrow \\ \frac{P_n(\Delta t + t) - P_n(t)}{\Delta t} &= \lambda_{n-1} P_{n-1}(t) - (\lambda_n + \mu_n) P_n(t) + \mu P_{n+1}(t) + \frac{o(\Delta t)}{\Delta t} \end{aligned}$$

On faisant tendre Δt vers 0, on obtient:

$$\begin{cases} P'_0(t) = -\lambda_0 P_0(t) + \mu_1 P_1(t) \\ P'_n(t) = -(\lambda_n + \mu_n) P_n(t) + \lambda_{n-1} P_{n-1}(t) + \mu_{n+1} P_{n+1}(t); n \in \mathbb{N} \end{cases}$$

Régime stationnaire

Lorsque $t \rightarrow \infty$, les limites $P_n = \lim_{t \rightarrow \infty} P_n(t)$ existent et sont indépendantes de l'état initial du processus, on a alors $\lim_{t \rightarrow \infty} P'_n(t) = 0$. Ceci se traduit par les équations dites de balance:

$$\begin{cases} \lambda_0 P_0(t) = \mu_1 P_1(t) \\ (\lambda_n + \mu_n) P_n(t) = \lambda_{n-1} P_{n-1}(t) + \mu_{n+1} P_{n+1}(t); n \in \mathbb{N} \end{cases}$$

Sous la condition d'ergodicité du système $\rho = \frac{\lambda}{\mu} < 1$ pour laquelle le régime stationnaire existe.

Mesures de performance

- le nombre moyen de clients dans le système $\bar{n} = \frac{\rho}{1-\rho}$;
- le nombre moyen de clients dans la file d'attente $\bar{n}_f = \frac{\lambda^2}{\mu(\mu-\lambda)}$;
- le temps moyen d'attente d'un client dans le système $\bar{w} = \frac{\bar{n}_f}{\lambda}$;
- le temps moyen de séjour d'un client dans le système $\bar{w}_s = \frac{\bar{n}}{\lambda} = \frac{1}{\mu-\lambda}$.

1.5.2 Le modèle $M/M/C$

Description du modèle

Les clients arrivent vers le système selon un processus de Poisson de taux $\lambda > 0$. Le service est assuré par $C \geq 1$ serveurs montés en parallèle. A l'arrivée d'un client, si l'un des serveurs est libre, le client commence immédiatement son service. Dans le cas contraire (tous les serveurs sont occupés par le service), le client prend place dans la file d'attente, commune pour tous les serveurs. La capacité d'attente est illimitée (le nombre de positions d'attente est infini). Lorsqu'un serveur se libère, le client en tête de la file d'attente occupe le serveur libéré. Par conséquent, la discipline d'attente est FIFO. Les

temps de service sont exponentiellement distribués de moyenne finie $1/\mu$. Les durées entre deux arrivées consécutives et les durées de service sont mutuellement indépendantes.[2]

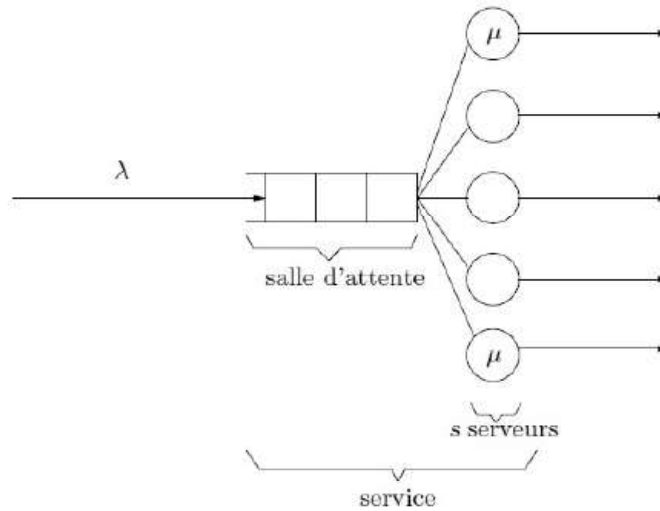


Figure 1.5.1 : Schéma général d'un système classiques M/M/C

Analyse du modèle

L'état du système à la date t peut être décrit à l'aide du processus $N(t)$, dont l'espace d'état est $S = \{0, 1, 2, \dots\}$. Ce dernier est un processus de naissance et de mort dont les taux de transition sont:

$$\lambda_n = \lambda, n \geq 0, \text{ et } \mu_n = \min \{n, c\} \times \mu, n \geq 1.$$

Régime transitoire

Le système d'équations de Kolmogorov pour les probabilités d'état

$$p_n(t) = P(N(t) = n), n \geq 0,$$

se présente de la manière suivante:

$$\begin{aligned} p'_0(t) &= -\lambda p_0(t) + \mu p_1(t); \\ p'_n(t) &= \lambda p_{n-1}(t) - (\lambda + n\mu)p_n(t) + (n+1)\mu p_{n+1}(t), \quad 1 \leq n < C; \\ p'_n(t) &= \lambda p_{n-1}(t) - (\lambda + C\mu)p_n(t) + C\mu p_{n+1}(t), \quad n \geq C. \end{aligned}$$

Régime stationnaire

Soit $p_n = \lim_{t \rightarrow \infty} p_n(t)$, $n \geq 0$. Cette distribution stationnaire satisfait les équations de balance:

$$\begin{aligned} 0 &= -\lambda p_0 + \mu p_1; \\ 0 &= \lambda p_{n-1} - (\lambda + n\mu)p_n + (n+1)\mu p_{n+1}, \quad 1 < n < C; \\ 0 &= \lambda p_{n-1} - (\lambda + C\mu)p_n + C\mu p_{n+1}, \quad n \geq C. \end{aligned}$$

Les résolutions du système d'équations ci-dessus nous donne:

$$\begin{aligned} p_n &= \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n p_0, \quad 1 \leq n \leq C; \\ p_n &= \frac{1}{C!} \frac{1}{C^{n-C}} \left(\frac{\lambda}{\mu}\right)^n p_0, \quad n \geq C; \end{aligned}$$

On remarque que pour $n = C$ les deux formules donnent la même valeur. Pour calculer la probabilité pour que le système est vide p_0 , on applique l'équation de normalisation

$$\sum_{n=0}^{\infty} p_n = 1.$$

En effet

$$p_0 = \left[\sum_{n=0}^{C-1} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n + \sum_{k=0}^{\infty} \frac{1}{C! C^k} \left(\frac{\lambda}{\mu}\right)^{C+k} \right].$$

La deuxième somme peut être réécrite de la manière suivante

$$\frac{1}{C!} \left(\frac{\lambda}{\mu}\right)^C \left[1 + \frac{\lambda}{\mu C} + \left(\frac{\lambda}{\mu C}\right)^2 + \dots \right].$$

La somme $\left[1 + \frac{\lambda}{\mu C} + \left(\frac{\lambda}{\mu C}\right)^2 + \dots \right]$ possède une limite égale à $\frac{1}{1 - \frac{\lambda}{\mu C}}$ si $\frac{\lambda}{\mu C} < 1$. Par conséquent, le système considéré est en régime stationnaire si $\rho = \frac{\lambda}{\mu C} < 1$, ρ est l'intensité globale du trafic. On obtient ainsi

$$p_0 = \left[\sum_{n=0}^{C-1} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n + \frac{(\lambda/\mu)^C}{C! \left(1 - \frac{\lambda}{\mu C}\right)} \right]^{-1}.$$

Encore,

$$\begin{aligned} p_0 &= \left[\sum_{n=0}^{c-1} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n + \frac{1}{c!} \left(\frac{\lambda}{\mu}\right)^c \sum_{n=c}^{\infty} \rho^{n-c} \right]^{-1} \\ &\text{et} \\ p_n &= \frac{1}{c!} \left(\frac{\lambda}{\mu}\right)^c \left(\frac{\lambda}{\mu c}\right)^{n-c} p_0 = \rho^{n-c} p_c. \end{aligned}$$

Mesures de performance

- le nombre moyen de clients dans le système:

$$\bar{n} = \sum_{n=0}^{\infty} n p_n = \sum_{n=1}^{c-1} \frac{n \left(\frac{\lambda}{\mu}\right)^n}{n!} p_0 + \sum_{n=c}^{\infty} \frac{n \left(\frac{\lambda}{\mu}\right)^n}{c! c^{n-c}} p_0 = \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^{c+1}}{c c! \left(1 - \frac{\lambda}{\mu c}\right)^2} p_0;$$

- le nombre moyen de clients dans la file d'attente:

$$\bar{n}_f = \sum_{k=0}^{\infty} k p_{c+k} = \frac{\left(\frac{\lambda}{\mu}\right)^c}{c!} \sum_{k=0}^{\infty} k \left(\frac{\lambda}{\mu c}\right)^k p_0 = \frac{(\lambda/\mu)^{c+1}}{c c! \left(1 - \frac{\lambda}{\mu c}\right)^2} p_0;$$

- le temps moyen d'attente d'un client dans le système

$$\bar{w} = \frac{\bar{n}_f}{\lambda} = \frac{c \mu (\lambda/\mu)^c}{c! (c \mu - \lambda)^2} p_0.$$

- le temps moyen de séjour d'un client dans le système

$$\bar{w}_s = \frac{\bar{n}}{\lambda} = \frac{1}{\mu} + \frac{\left(\frac{\lambda}{\mu}\right)^c}{c \mu c! \left(1 - \frac{\lambda}{\mu c}\right)^2} p_0;$$

Les systèmes de files d'attente avec rappels

2.1 Introduction

Dans la théorie des files d'attente classique, il est supposé qu'un client qui ne peut pas obtenir son service immédiatement dès son arrivée, rejoint la file d'attente ou quitte le système définitivement. Les systèmes de file d'attente développés tentent de prendre en considération des phénomènes de répétition des demandes de service, et ce-ci après une durée du temps aléatoire. Un tel système est connu comme système de files d'attente avec rappels".[12]

Dans ce chapitre nous définissons de manière générale un système de file d'attente avec rappels et nous donnons quelques exemples modélisés par ce genre de système, puis nous décrivons en particulier les systèmes de files d'attente $M/M/1$ et $M/M/C$ avec rappels et à la fin on décrit la méthode approximative la Troncature Finie.

2.2 Terminologie et notation

En utilisant la notation de Kendall établie pour la file d'attente classique, celle du modèle de file d'attente avec rappels s'écrit en ajoutant deux autres symboles aux six symboles précédement définis: $T/Y/C/K/m/Z/O/H$ avec:

- O : capacité de l'orbite;
- H : La probabilité de persévérance qui permet de définir le comportement du client devant une situation de blocage (serveurs occupés).

Remarque 2.2.1

1. O peut être supprimé si elle est infinie. H peut être également supprimé dans le cas d'un système sans perte (c'est-à-dire $H = 1$).
2. Le temps de rappel est défini comme l'intervalle du temps entre deux rappels consécutifs du même client secondaire.
3. La distribution de temps de rappel est supposé généralement exponentielle de taux θ .
4. Lorsque $\theta \rightarrow \infty$ le système d'attente avec rappels se rapporte à un système classique.
5. Lorsque $\theta \rightarrow 0$ le système d'attente avec rappels est un système d'Erlang avec perte.

2.3 Politique d'accès au serveur à partir de l'orbite

- **Politique de rappels classique :**

Le protocole le plus décrit dans la théorie de la file d'attente avec rappels est la politique classique dans laquelle chaque source dans l'orbite rappelle après un temps exponentiellement distribué avec un paramètre θ . Donc il y a une probabilité $n\theta dt + o(dt)$ d'un nouveau rappel dans le prochain intervalle $(t; t+dt)$, sachant que n clients sont en orbite à l'instant t . Une telle politique a été motivée naturellement par des applications dans la modélisation du comportement des abonnés dans les réseaux téléphoniques depuis les années 1940.

- **Politique de rappels constants:**

Durant les dernières années, les technologies ont considérablement évoluées. La littérature des files d'attente avec rappels décrit différents protocoles de rappels

spécifiques à certains réseaux informatiques et de communications modernes.

Dans les protocoles en question, le temps inter-rappels peut-être contrôlé par un dispositif électronique et par conséquent, est indépendant du nombre d'unités demandant le service. Ici, la probabilité d'un rappel durant $(t; t + dt)$ sachant que l'orbite n'est pas vide est $\beta dt + o(dt)$. Ce type de discipline de rappels est appelé politique de rappels constants. Le premier travail dans cette direction, considère une file d'attente $M/M/1$, où uniquement le client en tête de la file d'attente en orbite peut demander un service après un temps de rappels exponentiellement distribué avec un taux constant.

Cette sorte de politique de contrôle de rappels est bien connue par protocole ALOHA dans le système de communication.

- **Politique de rappels linéaires:**

Artalejo et Gomez-Correl [3] traitent les deux cas d'une manière unifiée en définissant une politique de rappels linéaire pour laquelle la probabilité d'un rappel durant $(t; t + dt)$, sachant que n clients sont en orbite à l'instant t est $(\beta(1 - \delta_{0n}) + n\theta) + o(dt)$. On mentionne aussi l'existence d'une autre politique dite politique de rappels quadratique.[4]

2.4 Exemples des systèmes de files d'attente avec rappels

Il est bien connu que les files d'attente avec rappels apparaissent naturellement comme des modèles pratiques dans la vie quotidienne et dans les réseaux de communication, par exemple en faisant des réservations, des réseaux de commutation de paquets, des accès CSMA non persistants (accès multiple du transporteur) et des systèmes en temps réel. Nous prétons quelques exemples de systèmes qui peuvent être modélisés comme une file d'attente avec rappels.[15]

1. Problème de réservation

C'est le plus simple exemple, soit un client qui veut prendre une réservation par téléphone dans un restaurant, il y a une seule ligne qui est employée à répondre aux réservations des clients. Ainsi, si un client appelle et trouve la ligne téléphonique occupée il rappellera après une certaine période de temps aléatoire avec la probabilité $H_k(H_k < 1)$ car le client ne peut rappeler indéfiniment.

Cet exemple peut être modélisé par une file d'attente $M/G/1$ avec rappel et avec perte en considérant que le processus d'arrivée des appels est poissonien.

2. Réseaux de communication par paquet

Considérons un réseau de communications d'ordinateurs dans lequel on trouve un ensemble d'interfaces "Interface Message Processors:IMP" (c'est le nœud de commutation de paquets utilisés pour connecter les ordinateurs à l'ARPANET) reliées entre elles par des câbles. Un ordinateur principal est connecté à l'une de ces interfaces. Si l'ordinateur envoie un message à un autre ordinateur il doit en premier lieu envoyer le message avec l'adresse de destination à l'interface à laquelle il est connecté. L'interface à son tour envoie le message à l'ordinateur destinataire directement si elle y est connectée, ou indirectement via d'autres interfaces. Considérons une interface à laquelle un ordinateur principal est connecté. Les messages arrivent de l'extérieur selon un processus aléatoire. Après la réception du message, l'ordinateur l'envoie immédiatement à l'interface. S'il y a un tampon libre, le message est accepté. Dans le cas contraire, le message est rejeté et l'ordinateur doit réessayer une autre fois après une période de temps. S'il existe des tampons libres, le message rejeté sera stocké dans un tampon de l'ordinateur principal. Dans le cas contraire, le message est considéré comme perdu. On peut poser les questions suivantes:

- Quelles sont les probabilités pour qu'un message soit rejeté par l'interface et par l'ordinateur principal?
- Quel est le nombre moyen de messages dans le tampon de IMP?

- Quel est le nombre moyen de messages dans le tampon de l'ordinateur principal?
- Quel est le temps d'attente d'un message dans le tampon de l'ordinateur principal?

Le problème présenté peut-être modélisé comme un système de file d'attente avec rappels à serveur unique (interface "IMP") possédant des tampons (positions d'attente). Le nombre de tampons de l'ordinateur principal constitue la capacité de l'orbite.

3. Système mémoire sur les disques magnétiques

Considérons un système mémoire où K unités disques partagent un contrôleur des disques et transmettent l'information quand elles trouvent ce dernier libre. Les demandes insatisfaites sont répétées après une rotation du disque. Ce système peut être présenté comme un système de files d'attente avec rappels. Le serveur est le contrôleur des disques. L'arrivée dans le système est quasi-aléatoire: le nombre de serveurs est fini et égal à K . La rotation du disque dans le cas de la répétition de demande peut être considéré comme un intervalle entre deux rappels consécutifs de durée constante.

4. La radio cognitive

La radio cognitive (CR) est actuellement l'une des technologies de transmission de l'information les plus prometteuses pour faire face aux problèmes de la rareté du spectre et de la sous-utilisation du spectre dans la communication sans fil. Les réseaux CR visent à améliorer l'efficacité du spectre pour répondre aux demandes toujours croissantes des utilisateurs finaux. Le principe est de donner la probabilité aux utilisateurs sans licence (utilisateurs secondaires, SU) d'accéder temporairement et dynamiquement à la bande passante inutilisée ou peu utilisée tout en veillant à ce qu'elle n'interfère ni ne dégrade les performances des titulaires de licence, communément appelés utilisateurs principaux (PU). Cela soulève plusieurs défis à relever dans les réseaux CR et la performance des utilisateurs secondaires. Nous

proposons de concevoir des réseaux CR comme un système de files d'attente avec rappels où les PU ont priorité sur les SU.

5. Systèmes téléphoniques

Il est bien connu qu'un abonné au téléphone qui obtient un signal occupé répète habituellement l'appel jusqu'à ce que la connexion requise soit établie. En conséquence, le flux des appels circulant dans un réseau téléphonique se compose de deux parties: le flux des appels primaires, qui reflète la véritable souhaits des abonnés au téléphones et le flux de répétitions des appels, ce qui est la conséquence du manque de succès des précédentes tentatives. Les modèles standard des systèmes téléphoniques des systèmes de file d'attente avec pertes, ne tiennent pas compte de cette structure de flux réel des appels et ne peut donc être appliqué à la réalisation d'un nombre de problèmes pratiquement importants.

Ces considérations mettent en évidence la nécessité des files d'attente avec rappels en tant que modélisation appropriée du comportement du client dans les systèmes téléphoniques classiques.

Cette classe de files d'attente est caractérisée par la caractéristique suivante: un client qui arrive et tous les serveurs accessibles pour lui sont occupés quitte la zone de service mais après un moment aléatoire répète sa demande. Notons que le but principal de tout centre d'appels est de fournir un service de qualité via un appareil, dans notre cas cet appareil étant un téléphone. De plus, lorsque la structure, la planification, la gestion et l'exécution d'un centre d'appels typique sont soigneusement étudiées, il est immédiatement clair comment la modélisation des systèmes téléphoniques en tant que files d'attente avec rappel clarifie le mieux le comportement des clients dans un tel scénario. Presque tous les principaux acteurs de l'industrie des télécommunication utilisent les centres d'appels comme leur principal moyen de communication et d'interaction avec leurs clients. Pour avoir une perspective, un centre d'appels peut être considéré comme une structure de file d'attente de base basée sur la file d'attente $M/M/c$ (l'un des modèles de file d'attente les plus utilisés) Shekhar et al.(2016).

2.5 Modèle général

En général, un système de file d'attente avec répétition d'appels peut être comme suit: Le système contient un espace de service composé de C ($C > 1$) serveurs identiques et indépendants, d'un espace d'attente ayant $N - C$ ($N > C$) positions d'attente.

Les clients arrivent dans le système selon un processus aléatoire avec une loi de probabilité donnée et forment un flux d'appels primaire. À l'arrivée d'un client, s'il ya un ou plusieurs serveurs libres et en bon états, le client sera servi immédiatement. Sinon, s'il y a des positions d'attente libre, le client rejoint la file selon la discipline de service adaptée dans le système.[4]

Lorsque tous les serveurs et toutes les positions d'attente du buffer sont occupés, le client quitte le système avec la probabilité $1 - H_0$ ou bien entre en "orbite" avec la probabilité de H_0 et devient source "d'appels répétés" ou "d'appels secondaires".

La capacité O peut être finie ou infinie. Dans le cas où k est finie et si l'orbite est plein, le client quitte le système pour toujours. Chaque client de l'orbite forme un processus "d'arrivées secondaires" de taux θ et il est traité de la même manière qu'un "client primaire" qui arrive de l'extérieur du système, trouve un serveur ou une position libre dans la file, ou quitte le système avec la probabilité $1 - H_k$ (s'il s'agit de la $k^{\text{ème}}$ tentative échouée) ou bien entre en orbite avec la probabilité H_k si l'orbite n'est pas plein. Le schéma général d'un système d'attente avec rappels est donné par la figure 2.5.1.

Remarque 2.5.1 *Le modèle de file d'attente avec rappels décrit ci-dessus est un modèle général. Plusieurs systèmes de file d'attente avec rappels peuvent être considérés comme des cas particuliers tels que: les systèmes sans buffer, les systèmes à un seul serveur, Les primaires ou secondaires qui arrivent durant un temps de service entrent en orbite sans aucune influence sur le processus de service.*

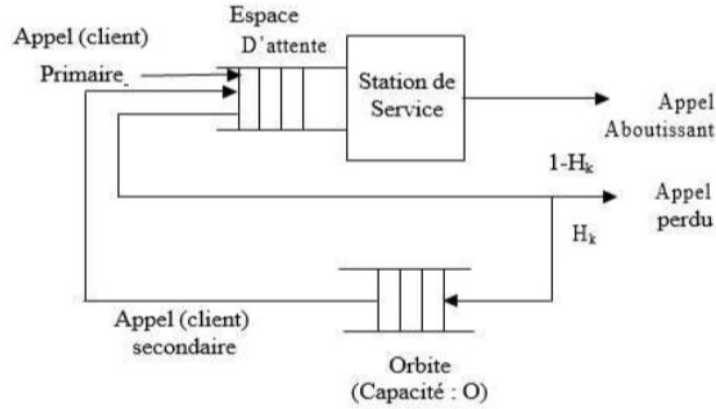


Figure 2.5.1 : Schéma général d'un système avec rappels

2.6 Modèles markoviens

2.6.1 système de file d'attente avec rappels $M/M/1$

Nous considérons un système de file d'attente avec un seul serveur dont lequel les clients primaires arrivent selon un processus de Poisson de taux $\lambda > 0$. Si un client primaire arrive et trouve le serveur libre, il commence son service. Sinon, il entre en orbite. Nous admettons que la durée de service et la durée entre deux rappels consécutifs sont exponentiellement distribuées de moyennes finies, respectivement $\frac{1}{\mu}$ et $\frac{1}{\theta}$.

L'état du système peut être décrit par le processus stochastique $\{C(t), N_\theta(t), t \geq 0\}$ où $C(t)$ est égal à 0 ou 1 selon le fait que le serveur est libre ou non, $N_\theta(t)$ est le nombre de client en orbite à l'instant t . Supposons que le régime stationnaire existe ($\rho = \frac{\lambda}{\mu} < 1$), le processus $\{C(t), N_\theta(t), t \geq 0\}$ est de Markov d'espace d'état $S = \{0, 1\} \times \mathbb{N}$. Les taux de transition sont donnés par:

$$q_{(0,j)(1,j)} = \lambda; j \in \mathbb{N};$$

$$q_{(1,j)(0,j)} = \mu; j \in \mathbb{N};$$

$$q_{(1,j)(1,j+1)} = \lambda; j \in \mathbb{N};$$

$$q_{(0,j)(1,j-1)} = j\theta; j \in \mathbb{N};$$

Les transformées inverses des (2.6.3) et (2.6.4) nous donnent les formules analytiques explicites suivantes:

$$P_{0j} = \frac{\rho}{j! \theta^j} \prod_{k=0}^{j-1} (1 + k\theta)(1 - \rho) \frac{\lambda}{\theta}{}^{+1};$$

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

2.6.2 Système de files d'attente avec rappels $M/M/C$

Description du modèle

Considérons un système de file d'attente avec rappels où l'espace de service comprend C serveurs entièrement disponibles dans lequel un flux des arrivées primaires est poissonien de taux $\lambda > 0$. Si un client primaire trouve au moins un serveur libre, il occupe immédiatement un serveur et quitte le système après le service. Sinon, si tous les serveurs sont occupés, il entre en orbite.

Nous admettons que la durée de service et la durée entre deux rappels consécutifs sont exponentiellement distribuées de moyennes finies, respectivement $\frac{1}{\mu}$ et $\frac{1}{\theta}$.

Le fonctionnement du système peut être décrit par un processus stochastique $\{C(t), N_\theta(t); t \geq 0\}$ où $C(t)$ est le nombre de serveurs occupés et $N_\theta(t)$ est le nombre de client en orbite. Sous les hypothèses ci-dessus, le processus $\{C(t), N_\theta(t); t \geq 0\}$ est markovien avec l'espace d'états $S = \{0, 1, \dots, c\} \times \mathcal{Z}^+$. Ses taux de transition $q_{(ij)(n,m)}$ sont donnés par:

1. Pour $0 \leq i \leq C - 1$

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

2. Pour $i = C$

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

Le graphe de transitions est présenté dans la Figure 2.6.2

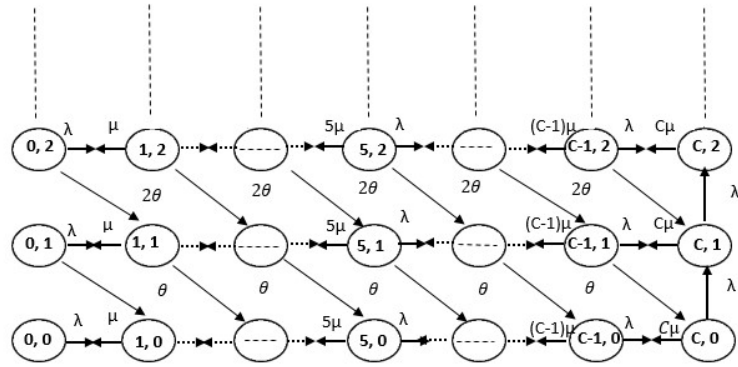


Figure 2.6.2 : Graphe des transitions du modèle M/M/C avec rappels

Supposons que la condition d'ergodicité $\lambda < C\mu$ est satisfaite. Soit

$$P_{ij} = \lim_{t \rightarrow \infty} P(C(t) = i, N_o(t) = j), \quad (i, j) \in S,$$

La distributin conjointe stationnaire du nombre de clients dans les zones de service et d'attente et du nombre de clients en orbite, pour laquelle les équations de Kolmogorov sont données par:

$$\begin{aligned} (\lambda + i\mu + j\theta)P_{ij} &= \lambda P_{i-1,j} + (j+1)\theta P_{i-1,j+1} + (i+1)\mu P_{i+1,j}, & (2.6.5) \\ 0 \leq i \leq C-1 \text{ et } j \geq 0; \end{aligned}$$

$$\begin{aligned} (\lambda + i\mu)P_{ij} &= \lambda P_{i,j-1} + (j+1)\theta P_{i-1,j+1} + \lambda P_{i-1,j}, & (2.6.6) \\ i = C \text{ et } j \geq 0. \end{aligned}$$

La distribution stationnaire P_{ij} contribue à l'obtention des caractéristiques de performance, telles que:

- La probabilité que tous les serveurs sont occupés $P_C = \lim_{t \rightarrow \infty} P(C(t) = C)$;
- Le nombre moyen de clients en orbite $\bar{N}_o = \lim_{t \rightarrow \infty} E[N_o(t)]$;
- Le nombre moyen de serveurs occupés $\bar{C} = \lim_{t \rightarrow \infty} E[C(t)]$.

Les caractéristiques essentielles de l'espace d'états du modèle considéré sont: sa dimension infinie (due à l'orbite illimitée) et la non-homogénéité le long de celui-ci (produite par la caractéristique des rappels). On ne peut donc pas obtenir la solution analytique exacte pour la distribution stationnaire P_{ij} . Dans ces circonstances, nous nous intéressons à l'approche approximative: La Troncature Finie.

2.7 Approche d'analyse des modèles à multiserveurs avec rappels

2.7.1 Introduction

Due à l'absence des formules analytiques explicites pour les caractéristiques probabilistes principales des modèles de files d'attente avec rappels à multiserveurs, la seule méthode pour obtenir des données numériques précises consiste à résoudre les équations de Kolmogorov numériquement. Mais dès que ce système d'équations est infini, il ne peut pas être résolu directement même sur ordinateur. Les transformations qui réduisent ces équations à une solution d'un petit problème fini dans le cas général ne sont pas disponibles. Donc, on a besoin d'une méthode pour approximer la solution numérique de ce système.[5]

2.7.2 Modèles tronqués

Dans ce modèle par opposition au modèle principal, la taille de l'orbite est limitée par une constante donnée M . Si le nombre de sources est égal à M , alors les appels bloqués

sont perdus et n'ont aucune influence sur le fonctionnement du système.[24]

La dynamique stochastique du système peut être décrite par le processus bidimensionnel $\{C^{(M)}(t), N_\theta^{(M)}(t)\}$ ou $C^{(M)}(t)$ est le nombre des serveurs occupés et $N_\theta^{(M)}(t)$ est le nombre de clients en orbite à la date t . Sous les suppositions ci-dessus, le processus $\{C^{(M)}(t), N_\theta^{(M)}(t)\}$ est markovien et possède l'espace d'états $S^{(M)} = \{0, 1, \dots, c\} \times \{0, 1, \dots, M\}$. es taux de transition $q_{(ij)(nm)}^{(M)}$ sont donnés par:

1. Si $0 \leq i \leq C - 1, 0 \leq j \leq M$

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

2. Si $i = C, 0 \leq j \leq M - 1$

$$q_{Cj}^{(M)}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (C, j + 1) \\ C\mu & \text{si } (k, l) = (C - 1, j) \quad ; \\ -(\lambda + C\mu) & \text{si } (k, l) = (i, j) \\ 0 & \text{sinon} \end{cases}$$

3. Si $i = C, j = M$

$$q_{CM}^{(M)}(k, l) = \begin{cases} C\mu & \text{si } (k, l) = (C - 1, M) \\ -C\mu & \text{si } (k, l) = (C, M) \quad . \\ 0 & \text{sinon} \end{cases}$$

Remarque 2.7.1 Vu que l'espace d'états est fini, le régime stationnaire existe toujours.

La distribution stationnaire $P_{ij}^{(M)} = P\{C^{(M)}(t) = i, N^{(M)}(t) = j\}$ satisfait le système d'équations suivant:

$$\begin{aligned} (\lambda + i\mu + j\theta)P_{ij}^{(M)} &= \lambda P_{i-1,j}^{(M)} + (j+1)\theta P_{i-1,j+1}^{(M)} + (i+1)\mu P_{i+1,j}^{(M)}, & (2.7.1) \\ 0 \leq i \leq C - 1, 0 \leq j \leq M - 1; & \end{aligned}$$

$$(\lambda + i\mu + M\theta)P_{iM}^{(M)} = \lambda P_{i-1,M}^{(M)} + (i+1)\mu P_{i+1,M}^{(M)} \quad (2.7.2)$$

$$0 \leq i \leq C-1, j = M;$$

$$\begin{aligned} (\lambda + C\mu)P_{ij}^{(M)} &= \lambda P_{C-1,j}^{(M)} + (j+1)\theta P_{C-1,j+1}^{(M)} \\ &\quad + \lambda P_{C,j-1}^{(M)}, \end{aligned} \quad (2.7.3)$$

$$i = C, 0 \leq j \leq M-1;$$

$$C\mu P_{C,M}^{(M)} = \lambda P_{C-1,M}^{(M)} + \lambda P_{C,M-1}^{(M)}, \quad (2.7.4)$$

$$i = C, j = M;$$

et satisfont la condition de normalisation

$$\sum_{i=0}^C \sum_{j=0}^M P_{ij}^{(M)} = 1. \quad (2.7.5)$$

2.7.3 Méthode numérique

Pour obtenir une précision correcte de la solution dans le cas d'un trafic fort ou intensité de rappels faible, la valeur M doit être vraiment large (environ les centaines). Si le nombre de serveurs est aussi large, alors le nombre d'inconnues peut atteindre quelques centaines de milliers et, ainsi l'application de la méthode de troncation rencontre des difficultés considérables. Donc il est important pour les modèles essentiels de bien simplifier les calculs. Comment ceci peut être réalisée? Une idée est due à Wilkinson [22]: on introduit les nouvelles inconnues $r_{ij}^{(M)} = \frac{P_{ij}^{(M)}}{P_{0M}^{(M)}}$. Si on peut trouver $r_{ij}^{(M)}$ alors on peut calculer $P_{ij}^{(M)}$:

$$P_{ij}^{(M)} = \frac{r_{ij}^{(M)}}{\sum_{i=0}^C \sum_{j=0}^M r_{ij}^{(M)}}. \quad (2.7.6)$$

Les variables $r_{ij}^{(M)}$ satisfont l'ensemble d'équations suivant (obtenu à partir des équations (2.7.1)-(2.7.5) pour les probabilités $P_{ij}^{(M)}$) [Zidani.N et Djellab.N, JNMA'14 (2014)]:

$$r_{0M}^{(M)} = 1; \quad (2.7.7)$$

$$(\lambda + i\mu + j\theta)r_{ij}^{(M)} = \lambda r_{i-1,j}^{(M)} + (j+1)\theta r_{i-1,j+1}^{(M)} + (i+1)\mu r_{i+1,j}^{(M)}, \quad (2.7.8)$$

$$0 \leq i \leq C-1, 0 \leq j \leq M-1;$$

$$(\lambda + i\mu + M\theta)r_{iM}^{(M)} = \lambda r_{i-1,M}^{(M)} + (i+1)\mu r_{i+1,M}^{(M)}, \quad (2.7.9)$$

$$0 \leq i \leq C-1, \quad j = M;$$

$$(\lambda + C\mu)r_{Cj}^{(M)} = \lambda r_{C-1,j}^{(M)} + (j+1)\theta r_{C-1,j+1}^{(M)} + \lambda r_{C,j-1}^{(M)}, \quad (2.7.10)$$

$$i = C, \quad 0 \leq j \leq M-1;$$

$$C\mu r_{CM}^{(M)} = \lambda r_{C-1,M}^{(M)} + \lambda r_{C,M-1}^{(M)}, \quad (2.7.11)$$

$$i = C, \quad j = M.$$

Calculons les variables $r_{ij}^{(M)}$ par groupe, chacun de taille $C+1$. Au début, on calcule $r_{0M}^{(M)}, \dots, r_{CM}^{(M)}$, ensuite $r_{0,M-1}^{(M)}, \dots, r_{C,M-1}^{(M)}$ et ainsi de suite, jusqu'à ce qu'on trouve $r_{00}^{(M)}, \dots, r_{C0}^{(M)}$.

1. Mettre $j = M$.

(a) Pour trouver le groupe $r_{0M}^{(M)}, \dots, r_{CM}^{(M)}$, réécrire l'équation (2.7.9) comme suit:

$$r_{i+1,M}^{(M)} = \frac{(\lambda + i\mu + M\theta)r_{iM}^{(M)} - \lambda r_{i-1,M}^{(M)}}{(i+1)\mu}, \quad 0 \leq i \leq C-1,$$

ce qui est équivalent à

$$r_{i,M}^{(M)} = \frac{(\lambda + (i-1)\mu + M\theta)r_{i-1,M}^{(M)} - \lambda r_{i-2,M}^{(M)}}{i\mu}, \quad 1 \leq i \leq C.$$

D'après (2.7.7) $r_{0M}^{(M)} = 1$. De ce fait, il est possible de calculer récursivement les variables $r_{1M}^{(M)}, \dots, r_{CM}^{(M)}$.

2. Mettre $j = j-1$. Calculons les variables $r_{0j}^{(M)}, \dots, r_{Cj}^{(M)}$.

(a) La dernière variables, $r_{Cj}^{(M)}$, peut être trouvée à partir de l'équation (2.7.11) (si $j = M-1$):

$$r_{C,M-1}^{(M)} = \frac{C\mu r_{C,M}^{(M)} - \lambda r_{C-1,M}^{(M)}}{\lambda}, \quad (2.7.12)$$

où à partir de l'équation (2.7.13) avec j remplacé par $j+1$ (si $j < M-1$):

$$r_{Cj}^{(M)} = \frac{(\lambda + C\mu)r_{C,j+1}^{(M)} - \lambda r_{C-1,j+1}^{(M)} - (j+2)\theta r_{C-1,j+2}^{(M)}}{\lambda}. \quad (2.7.13)$$

- (b) Pour trouver les variables $r_{0j}^{(M)}, \dots, r_{C-1,j}^{(M)}$, nous utilisons l'équation (2.7.8) pour $i = 0, \dots, C - 1$. Cet ensemble d'équations a la forme:

$$\alpha_i x_{i-1} + \beta_i x_i + \gamma_i x_{i+1} = \delta_i, \quad 0 \leq i \leq C - 1, \quad (2.7.14)$$

où

$$x_i = r_{ij}^{(M)},$$

$$\alpha_i = -\lambda,$$

$$\beta_i = \lambda + i\mu + j\theta,$$

$$\gamma_i = -(i + 1)\mu,$$

$$\delta_i = (j + 1)\theta r_{i-1,j+1}^{(M)},$$

et les valeurs

$$x_{-1} = 0, \quad x_C = r_{Cj}^{(M)}.$$

sont connues.

Ces équations, qui sont des équations aux différences, produisent des solutions numériques comme les équations différentielles de second ordre. L'algorithme le plus efficace pour leur résolution peut être trouvé dans plusieurs livres sur les méthodes numériques (voir, par exemple [8]). D'après cet algorithme, on calcule en premier lieu les variables $B_i, D_i, 0 \leq i \leq C - 1$, par des formules récursives

$$B_0 = \beta_0, \quad D_0 = \delta_0,$$

$$B_i = \beta_i - \frac{\alpha_i \gamma_{i-1}}{B_{i-1}}, \quad D_i = \delta_i - \frac{\alpha_i D_{i-1}}{B_{i-1}}, \quad 1 \leq i \leq C - 1,$$

et puis, à partir de l'équation

$$B_i x_i + \gamma_i x_{i+1} = D_i, \quad 0 \leq i \leq C - 1,$$

on calcule récursivement (dans l'ordre inverse) les inconnues x_{C-1}, \dots, x_0 .

Dans le cas présent, nous avons la procédure suivante:

- Calculer les variables B_{ij} , D_{ij} , $0 \leq i \leq C - 1$, à l'aide des équations

$$\begin{aligned}
 B_{0j} &= \lambda + j\theta, \\
 B_{ij} &= \lambda + i\mu + j\theta - \frac{\lambda i}{B_{i-1,j}}, \text{ pour } 1 \leq i \leq C - 1, \\
 D_{0j} &= 0, \\
 D_{ij} &= (j + 1)\theta r_{i-1,j+1}^{(M)} + \frac{\lambda D_{i-1,j}}{B_{i-1,j}}, \text{ pour } 1 \leq i \leq C - 1.
 \end{aligned} \tag{2.7.15}$$

- Puis calculer récursivement $r_{ij}^{(M)}$, $0 \leq i \leq C - 1$, (dans l'ordre inverse, commencer par $r_{Cj}^{(M)}$ connu à partir de l'étape 2.(a) à l'aide de l'équation

$$r_{ij}^{(M)} = \frac{D_{ij} + (i + 1)\mu r_{i+1,j}^{(M)}}{B_{ij}}, \quad i = C - 1, C - 2, \dots, 1, 0.$$

- (c) La répétition de l'étape 2 tant que $j \geq 0$ permet d'obtenir toutes les variables $r_{ij}^{(M)}$ (ce qui est, successivement pour $j = M - 2, M - 3, \dots, 0$).

Puisque $P_{ij}^{(M)} = r_{ij}^{(M)} \cdot P_{0M}^{(M)}$, alors on aura:

$$P_{0M}^{(M)} = \frac{1}{\sum_{i=0}^C \sum_{j=0}^M r_{ij}^{(M)}}.$$

A présent, on peut calculer les probabilités $P_{ij}^{(M)} = r_{ij}^{(M)} \cdot P_{0M}^{(M)}$ et les principales caractéristiques probabilistes du système tronqué $S^{(M)}$ [Zidani.N et Djellab.N, JJC (2014)]:

- La probabilité de blocage

$$P_C^{(M)} = \sum_{j=0}^M r_{Cj}^{(M)} \cdot P_{0M}^{(M)};$$

où $C^{(M)}(t)$ est le nombre de serveurs occupés.

- Le nombre moyen de serveurs occupés

$$\bar{C}^{(M)} = \sum_{i=0}^C \sum_{j=0}^M i r_{ij}^{(M)} \cdot P_{0M}^{(M)};$$

- Le nombre moyen de clients en orbite

$$\bar{N}_o^{(M)} = \sum_{i=0}^C \sum_{j=0}^M j r_{ij}^{(M)} \cdot P_{0M}^{(M)}.$$

Application aux centres d'appels

3.1 Introduction

Lorsqu'une personne compose un numéro de téléphone d'un centre d'appels pour solliciter un service quelconque, la première question qui lui passe par l'esprit est : "est ce que je vais attendre beaucoup avant d'être servi si jamais j'arrive à joindre le centre d'appels ?". Le meilleur scénario serait d'avoir directement une personne au bout de la ligne au lieu d'un répondeur automatique. En pratique, aucun centre d'appels n'est capable de satisfaire les clients de la sorte. C'est de la qu'un système de file d'attente s'est imposé pour créer un certain équilibre entre des gestionnaires de centres qui veulent fournir un bon service et des clients qui veulent être servis sans avoir à attendre beaucoup de l'autre côté.[1]

Dans ce chapitre, nous allons analyser les performances du système d'attente au niveau d'un centre d'appels téléphoniques. À cet effet, nous donnons une description mathématique du modèle, trouvons la condition d'ergodicité, présentons les étapes d'application de la méthode de Troncation, nous évaluons quelques mesures de performances pour certaines paramètres du système.

3.2 Les centres d'appels

Un centre d'appels se définit comme étant un ensemble de ressources, généralement du personnel, des ordinateurs et des équipements de télécommunication, qui permettent d'offrir des services par téléphone. Les centres d'appels sont aujourd'hui les éléments clés de presque toutes les grandes organisations, Ils sont utilisés, par exemple, pour fournir de l'information ou du support aux clients. Ils permettent aux compagnies de faire des ventes et aux clients de payer leurs factures par téléphone, ... etc. Ils sont aussi utilisés par les services d'information gouvernementaux et les services d'urgence (police, ambulance...). Ainsi, il existe un besoin pratique d'analyser ce type de modèle en tant que des systèmes de files d'attente avec rappels et multiserveurs.

Dans cette mémoire, nous considérons un centre d'appels avec intégration de téléphonie informatique (CTI) composé de plusieurs lignes réseau. Ces lignes relient les appels vers le centre via l'autocommutateur privé (PABX). Le PABX identifie automatiquement le numéro de téléphone de tous les appels entrants en utilisant l'identification automatique de numéro (ANI). Un certain nombre de serveurs répondent aux appels des clients. Lorsqu'un appel (client ordinaire) arrive, il sera immédiatement servi si un serveur est disponible. Si tous les serveurs sont occupés avec d'autres appels, nous avons deux possibilités selon le fait que le seuil pour le nombre de clients en attente est atteint ou non. Dans le premier cas, l'appel arrivant sera automatiquement déconnecté (par PABX) et le client entendra un message lui demandant d'essayer d'accéder au centre d'appels plus tard; alors que dans le second cas le client en question sera tenu d'attendre jusqu'à ce qu'un serveur devient inoccupé. En outre, certains clients sont impatients d'attendre un serveur et abandonnent après avoir attendu un certain temps. Le système peut être sujet à des interruptions de service ou des pannes (en raison des arrivées négatives). Dans ce cas, un ou plusieurs clients seront supprimés du service. Les intervalles de temps entre les interruptions de service sont aléatoires et peuvent dépendre de l'état du système et du processus d'arrivée des appels ordinaires.

3.3 Description mathématique du modèle

Nous considérons un système de files d'attente avec C serveurs et $K - C$ positions d'attente vers lequel les clients primaires positifs et négatifs arrivent selon des processus de Poisson avec des taux $\lambda > 0$ et $\mu > 0$, respectivement. Un client positif reçoit un service immédiat s'il y a un serveur inactif, sinon nous avons deux possibilités selon le fait que le nombre de clients dans l'espace d'attente est égal à $K - C$ ou inférieur à $K - C$. Dans le premier cas, il quitte temporairement la zone de service pour rejoindre le groupe de clients bloqués (orbite); dans le second cas, il est tenu d'attendre dans l'espace d'attente.

Tout client en orbite va répéter à plusieurs reprises ses tentatives jusqu'au moment où il trouve un serveur inactif ou une position d'attente libre. Les temps de rappels sont répartis de façon exponentielle avec la fonction de répartition $T(x) = 1 - e^{-\theta x}$, $x \geq 0$ et le taux $\theta > 0$. Il s'agit d'une politique des rappels classique: les tentatives répétées sont faites individuellement par chaque client bloqué et le taux total des rappels est $j\theta$ lorsque la taille de l'orbite est $j \geq 0$. Les clients positifs en attente (les clients dans l'espace d'attente) abandonneront le système si leur seuil de patience est dépassé. Nous supposons que les clients abandonnent après avoir attendu un temps exponentiellement distribué avec la fonction de répartition $W(x) = 1 - e^{-\gamma x}$, $x \geq 0$ et le taux $\gamma > 0$.

Une arrivée négative a pour effet de retirer un certain nombre de clients du service. Soit q_s la probabilité de supprimer $1 \leq s \leq C$ clients lorsqu'une arrivée négative se produit. Ainsi, le taux d'effacement est définie par $\eta_s = \nu q_s$. Les temps de service suivent une loi exponentielle de fonction de répartition $B(x) = 1 - e^{-\mu x}$, $x \geq 0$ et de moyenne finie $1/\mu$. Enfin, nous admettons l'hypothèse de l'indépendance mutuelle entre toutes les variables aléatoires définies ci-dessus.

L'état du système à l'instant t peut être décrit au moyen du processus $\{C(t), N_o(t), t \geq 0\}$, où $C(t)$ est le nombre de clients dans les zones de service et d'attente, $N_o(t)$ représente le nombre de clients en orbite à l'instant t . Il s'agit d'une chaîne de Markov continue avec

3.4. Condition d'existence d'un régime stationnaire

un espace d'états $S = \{0, 1, \dots, C, C + 1, \dots, K\} \times \mathbb{N}$. Ses taux de transition $q_{ij}(k, l)$ sont donnés par:

1. Pour $0 \leq i \leq C - 1$

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

2. Pour $C \leq i \leq K - 1$

$$q_{ij}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (i + 1, j) \\ C\mu + (i - C)\gamma & \text{si } (k, l) = (i - 1, j) \\ \eta_s & \text{si } (k, l) = (i - s, j), 1 \leq s \leq i \\ j\theta & \text{si } (k, l) = (i + 1, j - 1) \\ -(\lambda + C\mu + (i - C)\gamma + j\theta + \eta_s) & \text{si } (k, l) = (i, j) \\ 0 & \text{sinon} \end{cases} ;$$

3. Pour $i = K$

$$q_{Kj}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (K, j + 1) \\ C\mu + (K - C)\gamma & \text{si } (k, l) = (K - 1, j) \\ -(\lambda + C\mu + (K - C)\gamma) & \text{si } (k, l) = (K, j) \\ 0 & \text{sinon} \end{cases} .$$

Puisque le modèle à étudier a une structure complexe incluant les arrivées négatives, les rappels et les abandons, son analyse mathématique présente un intérêt particulier ainsi qu'une difficulté analytique particulière.

3.4 Condition d'existence d'un régime stationnaire

Condition suffisante pour l'ergodicité

Nous utilisons un critère basé sur les accroissements moyens (théorie des fonctions de

Lyapunov [18]), en particulier un résultat (nommé théorème de Tweedie dans [6]) donné dans [21], qui définit les hypothèses pour qu'un processus de Markov soit régulier et ergodique [9, 10].

Théorème 3.4.1 *Soit $X(t)$ un processus de Markov avec espace d'états discret S et taux de transition q_{sp} , $s, p \in S$, $\sum_p q_{sp} = 0$. Supposons qu'il existe:*

- une fonction $\varphi(s)$, $s \in S$ bornée ci-dessous (cette fonction est appelée fonction test ou fonction de Lyapunov):
- un nombre positif ϵ tel que:

(a) variables $y_s = \sum_{p \neq s} q_{sp} (\varphi(p) - \varphi(s)) < \infty$, pour tous $s \in S$;

(b) $y_s \leq -\epsilon$ pour tous $s \in S$ sauf peut être un nombre fini d'états.

Alors le processus $X(t)$ est régulier et ergodique.

Nous considérons la fonction de test suivante $\varphi(s) \equiv \varphi(i, j) = ai + j$, où a est un paramètre inconnu et $s \in S$. Les accroissements moyens $y_s \equiv y_{ij}$ sont donnés par:

$$y_{ij} = \begin{cases} \lambda a + i\mu(-a) + s\eta_s(-a) + j\theta(a-1), & \text{si } 0 \leq i \leq C-1 \\ \lambda a + (C\mu + (i-C)\gamma)(-a) + s\eta_s(-a) + j\theta(a-1), & \text{si } C \leq i \leq K-1 \\ \lambda + (C\mu + (K-C)\gamma)(-a), & \text{si } i = K \end{cases} .$$

Puisque pour tout $i = 0, 1, \dots, K$, il existe

$$\lim_{j \rightarrow \infty} y_{ij} = L_i = \begin{cases} (a-1) \cdot \infty, & \text{si } 0 \leq i \leq C-1 \\ (a-1) \cdot \infty, & \text{si } C \leq i \leq K-1 \\ \lambda - a[C\mu + (K-C)\gamma], & \text{si } i = K \end{cases} ;$$

les hypothèses du théorème de Tweedie restent vraies si et seulement si toutes les variables L_i sont négatives. Ainsi,

$$\begin{aligned} a-1 &< 0; \\ a-1 &< 0; \\ \lambda - a[C\mu + (K-C)\gamma] &< 0. \end{aligned}$$

3.4. Condition d'existence d'un régime stationnaire

Ces conditions représentent un ensemble d'inégalités linéaires pour le paramètre inconnu a . Ces conditions peuvent être écrites sous la forme

$$\frac{\lambda}{[C\mu + (K - C)\gamma]} < a < 1.$$

Un tel a peut être trouvé si l'intervalle $((\lambda/(C\mu + (K - C)\gamma)), 1)$ est non vide, c'est-à-dire si $\lambda < C\mu + (K - C)\gamma$.

La dernière expression est une condition suffisante pour l'ergodicité du modèle étudié.

Condition nécessaire pour l'ergodicité

Pour prouver que la condition ci-dessus est aussi une condition nécessaire, on applique le théorème suivant.

Théorème 3.4.2 [17]: *Soit $\{\zeta_n\}$ une chaîne de Markov avec espace d'états S et probabilités de transition r_{sp} . Supposons qu'il existe:*

- Une fonction non négative $\varphi(s)$, $s \in S$, telle que pour un certains d on a

$$r_{sp} \neq 0 \Rightarrow |\varphi(s) - \varphi(p)| \leq d;$$

- Un nombre positif b tel que :

(a) L'ensemble $A_b \equiv \{s \in S / \varphi(s) > b\} \neq \emptyset$;

(b) $\inf_{s \in A_b} x_s \equiv \inf_{s \in A_b} E(\varphi(\zeta_{n+1}) - \varphi(s) / \zeta_n = s)$ est non négatif (ou positif).

Alors la chaîne $\{\zeta_n\}$ est non-ergodique (n'admet pas un régime stationnaire ou elle est transitoire).

Considérons $\{\zeta_n\}$ une chaîne de Markov induite pour le processus de Markov défini ci-dessus $\{C(t), N_o(t), t \geq 0\}$. Ses probabilités de transition sont données par

1. Pour $0 \leq i \leq C - 1$

$$r_{ij}(k, l) = \begin{cases} \frac{\lambda}{\lambda+i\mu+j\theta+\eta_s} & \text{si } (k, l) = (i + 1, j) \\ \frac{i\mu}{\lambda+i\mu+j\theta+\eta_s} & \text{si } (k, l) = (i - 1, j) \\ \frac{\eta_s}{\lambda+i\mu+j\theta+\eta_s} & \text{si } (k, l) = (i - s, j), 1 \leq s \leq i \ ; \\ \frac{j\theta}{\lambda+i\mu+j\theta+\eta_s} & \text{si } (k, l) = (i + 1, j - 1) \\ 0 & \text{sinon} \end{cases}$$

2. Pour $C \leq i \leq K - 1$

$$r_{ij}(k, l) = \begin{cases} \frac{\lambda}{\lambda+C\mu+(i-C)\gamma+j\theta+\eta_s} & \text{si } (k, l) = (i + 1, j) \\ \frac{C\mu+(i-C)\gamma}{\lambda+C\mu+(i-C)\gamma+j\theta+\eta_s} & \text{si } (k, l) = (i - 1, j) \\ \frac{\eta_s}{\lambda+C\mu+(i-C)\gamma+j\theta+\eta_s} & \text{si } (k, l) = (i - s, j), 1 \leq s \leq i \ ; \\ \frac{j\theta}{\lambda+C\mu+(i-C)\gamma+j\theta+\eta_s} & \text{si } (k, l) = (i + 1, j - 1) \\ 0 & \text{sinon} \end{cases}$$

3. Pour $i = K$

$$r_{Kj}(k, l) = \begin{cases} \frac{\lambda}{\lambda+C\mu+(K-C)\gamma} & \text{si } (k, l) = (K, j + 1) \\ \frac{C\mu+(K-C)\gamma}{\lambda+C\mu+(K-C)\gamma} & \text{si } (k, l) = (K - 1, j) \\ 0 & \text{sinon} \end{cases} .$$

Supposons que $\lambda \geq C\mu + (K - C)\gamma$ et considérons la fonction de test $\varphi(i, j) = i + j$.

Alors,

- Si $r_{ij}(n, m) \neq 0$, alors $|i - n| \leq 1$, $|j - m| \leq 1$. Par conséquent:

$$|\varphi(i, j) - \varphi(n, m)| = |i - n + j - m| \leq |i - n| + |j - m| \leq 2.$$

- Puisque la variable b peut prendre un nombre positif quelconque. Alors,

(a) $A_b = \{(i, j) / i = 0, 1, \dots, K; j > \max(b - i, 0)\} \neq \emptyset$;

(b)

$$x_{ij} = \begin{cases} \frac{\lambda-i}{\lambda+i\mu+j\theta+\eta_s} & \text{si } 0 \leq i \leq C - 1 \\ \frac{\lambda-i}{\lambda+C\mu+(i-C)\gamma+j\theta+\eta_s} & \text{si } C \leq i \leq K - 1 \ . \\ \frac{\lambda-K}{\lambda+C\mu+(K-C)\gamma} & \text{si } i = K \end{cases}$$

Ainsi, la variable x_{ij} est toujours non négative et de plus $x_{ij} \geq 0$ pour $(i, j) \in A_b$. Par conséquent, le processus de Markov $\{C(t), N_o(t); t \geq 0\}$ est non-ergodique.

A présent, considérons le cas où $\lambda > C\mu + (K - C)\gamma$ et la fonction test est $\varphi(i, j) = ai + j$ avec $a \in (1, (\lambda / (C\mu + (K - C)\gamma)))$. Alors:

- Si $r_{ij}(n, m) \neq 0$ alors $|i - n| \leq 1, |j - m| \leq 1$.
Par conséquent, $|\varphi(i, j) - \varphi(n, m)| = |a(i - n) + j - m| \leq a|i - n| + |j - m| \leq a + 1$, c'est à dire $d = a + 1$.
- L'ensemble A_b est non vide quelque soit $b > 0$. Alors, on doit choisir b de telle sorte que $\inf_{(i,j) \in A_b} x_{ij} > 0$. Mais pour $i = K$, $x_{Kj} = \frac{\lambda - aK}{\lambda + C\mu + (K - C)\gamma} > 0$ (pour tous j); et pour $0 \leq i \leq K - 1$, $\lim_{j \rightarrow \infty} x_{ij} = a - 1 > 0$.

Par conséquent pour tout $i = 0, 1, \dots, K$, il existe N_i tel que $x_{ij} > \frac{a-1}{2}$ pour tout N_i . Ici, N_i est le nombre de i -clients en orbite. Si $\epsilon = \min \left\{ \frac{\lambda - aK}{\lambda + C\mu + (K - C)\gamma}, \frac{a-1}{2} \right\}$, alors $x_{ij} \geq \epsilon$ pour tout $(i, j) \in S$ sauf pour (i, j) tels que $i = 0, 1, \dots, K - 1$ et $0 \leq j \leq N_i - 1$. Supposons que b est assez grand pour que l'ensemble A_b ne contienne pas ces états. Alors $x_{ij} \geq \epsilon$, pour tout $(i, j) \in A_b$, ou $\inf_{(i,j) \in A_b} x_{ij} \geq \epsilon > 0$. En conséquence, la chaîne de Markov à temps continu $\{C(t), N_o(t), t \geq 0\}$ est non-ergodique.

Enfin, le processus de Markov considéré $\{C(t), N_o(t), t \geq 0\}$ est ergodique si et seulement si $\lambda < C\mu + (K - C)\gamma$. ■

3.5 Régime stationnaire

Supposons que la condition d'ergodicité $\lambda < C\mu + (K - C)\gamma$ est satisfait. Soit

$$P_{ij} = \lim_{t \rightarrow \infty} P(C(t) = i, N_o(t) = j), \quad (i, j) \in S,$$

la distribution conjointe stationnaire du nombre de clients dans les zones de service et d'attente et du nombre de clients en orbite, pour laquelle les équations de Kolmogorov

sont données par:

$$(\lambda + i\mu + j\theta + \eta_s)P_{ij} = \lambda P_{i-1,j} + (j+1)\theta P_{i-1,j+1} + ((i+1)\mu + \eta_s)P_{i+1,j},$$

$$0 \leq i \leq C-1 \text{ et } j \geq 0;$$

$$(\lambda + C\mu + (i-C)\gamma + j\theta + \eta_s)P_{ij} = \lambda P_{i-1,j} + (j+1)\theta P_{i-1,j+1}$$

$$+(C\mu + (i+1-C)\gamma + \eta_s)P_{i+1,j},$$

$$C \leq i \leq K-2 \text{ et } j \geq 0;$$

$$(\lambda + C\mu + (i-C)\gamma + j\theta + \eta_s)P_{ij} = \lambda P_{i-1,j} + (j+1)\theta P_{i-1,j+1}$$

$$+(C\mu + (i+1-C)\gamma)P_{i+1,j},$$

$$i = K-1 \text{ et } j \geq 0;$$

$$(\lambda + C\mu + (K-C)\gamma)P_{K,j} = \lambda P_{K-1,j} + (j+1)\theta P_{K-1,j+1} + \lambda P_{K,j-1},$$

$$i = K \text{ et } j \geq 0;$$

$$\sum_{i=0}^K \sum_{j=0}^{\infty} P_{ij} = 1.$$

La distribution stationnaire P_{ij} contribue à l'obtention des mesures de performance, telles que:

- probabilité de blocage $P_K = \lim_{t \rightarrow \infty} P(C(t) = K)$;
- nombre moyen de clients en orbite $\bar{N}_o = \lim_{t \rightarrow \infty} E[N_o(t)]$;
- nombre moyen de clients dans les zones de service et d'attente $\bar{C} = \lim_{t \rightarrow \infty} E[C(t)]$.

Les caractéristiques essentielles de l'espace d'états du modèle considéré sont: sa dimension infinie (due à l'orbite illimitée) et la non-homogénéité le long de celui-ci (produite par la caractéristique des rappels). On ne peut donc pas obtenir la solution analytique exacte pour la distribution stationnaire P_{ij} . Dans ces circonstances, nous nous intéressons à l'approche approximative: la troncature finie.

3.6 Troncature finie

La manière intuitive d'exécuter la procédure de troncature consiste à inclure dans la description du modèle la condition selon laquelle le nombre maximal de clients en orbite ne dépasse pas une valeur M (choisie suffisamment grande). Cependant, la mise en oeuvre de ce schéma de troncature peut conduire à prendre en compte un grand nombre d'états avec des probabilités négligeables. Vu la méthodologie, qui est proposée dans [20] et décrite dans le paragraphe 2.2.2, cet inconvénient peut être atténué en explorant toutes les directions dans l'espace d'états considéré où les probabilités d'état diminuent, puis en déterminant des frontières de troncature. Nous avons donc la possibilité de fournir l'erreur relative donnée de l'estimation des mesures de performance.

Supposons que la taille de l'orbite de notre modèle soit bornée par une constante donnée M , choisie de manière appropriée. Dans le cadre de l'hypothèse ci-dessus, le nouveau processus $\{C^{(M)}(t), N_o^{(M)}(t), t \geq 0\}$ est un Markovien avec l'espace d'états $S^{(M)} = \{0, \dots, C, C + 1, \dots, K\} \times \{0, 1, \dots, M\}$. En supposant que le taux de service $\mu = 1$ (ceci est toujours possible sans perte de généralité), nous avons les taux de transition infinitésimaux suivants $q_{ij}^{(M)}(k, l)$:

1. Pour $0 \leq i \leq C - 1, 0 \leq j \leq M - 1$

$$q_{ij}^{(M)}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (i + 1, j) \\ i & \text{si } (k, l) = (i - 1, j) \\ \eta_s & \text{si } (k, l) = (i - s, j), 1 \leq s \leq i \\ j\theta & \text{si } (k, l) = (i + 1, j - 1) \\ -(\lambda + i + j\theta + \eta_s) & \text{si } (k, l) = (i, j) \\ 0 & \text{sinon} \end{cases} ;$$

2. Pour $0 \leq i \leq C - 1, j = M$

$$q_{iM}^{(M)}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (i + 1, M) \\ i & \text{si } (k, l) = (i - 1, M) \\ \eta_s & \text{si } (k, l) = (i - s, M), 1 \leq s \leq i \\ M\theta & \text{si } (k, l) = (i + 1, M - 1) \\ -(\lambda + i + M\theta + \eta_s) & \text{si } (k, l) = (i, M) \\ 0 & \text{sinon} \end{cases} ;$$

3. Pour $C \leq i \leq K - 1, 0 \leq j \leq M - 1$

$$q_{ij}^{(M)}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (i + 1, j) \\ C + (i - C)\gamma & \text{si } (k, l) = (i - 1, j) \\ \eta_s & \text{si } (k, l) = (i - s, j), 1 \leq s \leq i \\ j\theta & \text{si } (k, l) = (i + 1, j - 1) \\ -(\lambda + C + (i - C)\gamma + j\theta + \eta_s) & \text{si } (k, l) = (i, j) \\ 0 & \text{sinon} \end{cases} ;$$

4. Pour $C \leq i \leq K - 1, j = M$

$$q_{iM}^{(M)}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (i + 1, M) \\ C + (i - C)\gamma & \text{si } (k, l) = (i - 1, M) \\ \eta_s & \text{si } (k, l) = (i - s, M), 1 \leq s \leq i \\ M\theta & \text{si } (k, l) = (i + 1, M - 1) \\ -(\lambda + C + (i - C)\gamma + M\theta + \eta_s) & \text{si } (k, l) = (i, M) \\ 0 & \text{sinon} \end{cases} ;$$

5. Pour $i = K, 0 \leq j \leq M - 1$

$$q_{Kj}^{(M)}(k, l) = \begin{cases} \lambda & \text{si } (k, l) = (K, j + 1) \\ C + (K - C)\gamma & \text{si } (k, l) = (K - 1, j) \\ -(\lambda + C + (K - C)\gamma) & \text{si } (k, l) = (K, j) \\ 0 & \text{sinon} \end{cases} ;$$

6. Pour $i = K, j = M$

$$q_{KM}^{(M)}(k, l) = \begin{cases} C + (K - C)\gamma & \text{si } (k, l) = (K - 1, M) \\ -(C + (K - C)\gamma) & \text{si } (k, l) = (K, M) \\ 0 & \text{sinon} \end{cases} .$$

Comme l'espace d'états $S^{(M)}$ est fini, le processus $\{C^{(M)}(t), N_o^{(M)}(t), t \geq 0\}$ est toujours ergodique. Sa distribution en régime stationnaire

$P_{ij}^{(M)} = \lim_{t \rightarrow \infty} P(C^{(M)}(t) = i, N_o^{(M)}(t) = j)$ satisfait aux équations suivantes:

$$(\lambda + i + j\theta + \eta_s)P_{ij}^{(M)} = \lambda P_{i-1,j}^{(M)} + (j + 1)\theta P_{i-1,j+1}^{(M)} + ((i + 1) + \eta_s)P_{i+1,j}^{(M)}, \quad (3.6.1)$$

$$0 \leq i \leq C - 1, \quad 0 \leq j \leq M - 1;$$

$$\begin{aligned} (\lambda + C + (i - C)\gamma + j\theta + \eta_s)P_{ij}^{(M)} &= \lambda P_{i-1,j}^{(M)} + (j + 1)\theta P_{i-1,j+1}^{(M)} \\ &+ (C + (i + 1 - C)\gamma + \eta_s)P_{i+1,j}^{(M)}, \end{aligned} \quad (3.6.2)$$

$$C \leq i \leq K - 2, \quad 0 \leq j \leq M - 1;$$

$$\begin{aligned} (\lambda + C + (i - C)\gamma + j\theta + \eta_s)P_{ij}^{(M)} &= \lambda P_{i-1,j}^{(M)} + (j + 1)\theta P_{i-1,j+1}^{(M)} \\ &+ (C + (i + 1 - C)\gamma + \eta_s)P_{i+1,j}^{(M)}, \end{aligned} \quad (3.6.3)$$

$$i = K - 1, \quad 0 \leq j \leq M - 1;$$

$$(\lambda + i + M\theta + \eta_s)P_{iM}^{(M)} = \lambda P_{i-1,M}^{(M)} + ((i + 1) + \eta_s)P_{i+1,M}^{(M)}, \quad (3.6.4)$$

$$0 \leq i \leq C - 1, \quad j = M;$$

$$\begin{aligned} (\lambda + C + (i - C)\gamma + M\theta + \eta_s)P_{iM}^{(M)} &= (C + (i + 1 - C)\gamma + \eta_s)P_{i+1,M}^{(M)} \\ &+ \lambda P_{i-1,M}^{(M)}, \end{aligned} \quad (3.6.5)$$

$$C \leq i \leq K - 2, \quad j = M;$$

$$\begin{aligned} (\lambda + C + (i - C)\gamma + M\theta + \eta_s)P_{iM}^{(M)} &= (C + (i + 1 - C)\gamma)P_{i+1,M}^{(M)} \\ &+ \lambda P_{i-1,M}^{(M)}, \end{aligned} \quad (3.6.6)$$

$$i = K - 1, \quad j = M;$$

$$(\lambda + C + (K - C)\gamma)P_{K,j}^{(M)} = \lambda P_{K-1,j}^{(M)} + (j + 1)\theta P_{K-1,j+1}^{(M)} + \lambda P_{K,j-1}^{(M)}, \quad (3.6.7)$$

$$i = K, 0 \leq j \leq M - 1;$$

$$(C + (K - C)\gamma)P_{K,M}^{(M)} = \lambda P_{K-1,M}^{(M)} + \lambda P_{K,M-1}^{(M)}, \quad (3.6.8)$$

$$i = K, j = M;$$

$$\sum_{i=0}^K \sum_{j=0}^M P_{ij}^{(M)} = 1. \quad (3.6.9)$$

Cet ensemble d'équations peut être résolu par un algorithme récursif, décrit dans le paragraphe 2.7.3 (Voir Annexe A), et nous permet de trouver certaines variables $r_{ij}^{(M)} = \frac{P_{ij}^{(M)}}{P_{0M}^{(M)}}$, $0 \leq i \leq K$ et $0 \leq j \leq M$. Ensuite, les probabilités $P_{ij}^{(M)} = \frac{r_{ij}^{(M)}}{\sum_{i=0}^K \sum_{j=0}^M r_{ij}^{(M)}}$. En prenant en considération les spécificités de notre modèle, nous avons adapté cet algorithme pour calculer les mesures de performance introduites ci-dessus pour le système tronqué:

$$P_K^{(M)} = \sum_{j=0}^M r_{Kj}^{(M)} P_{0M}^{(M)};$$

$$\bar{N}_o^{(M)} = \sum_{i=0}^K \sum_{j=0}^M j r_{ij}^{(M)} P_{0M}^{(M)};$$

$$\bar{C}^{(M)} = \sum_{i=0}^K \sum_{j=0}^M i r_{ij}^{(M)} P_{0M}^{(M)}.$$

Conclusion générale

Dans ce travail, nous nous sommes intéressés à l'étude de la propriété de troncature des systèmes de files d'attente avec rappels, en particulier des modèles à multiserveurs, espace d'attente, abandons exponentielle et arrivées négative.

Par conséquent, nous avons eu recours à une analyse approximative, basée sur le principe de troncature de l'espace d'états infini, pour obtenir la distribution conjointe en permanent le nombre de clients sur l'orbite.

- Nous avons commencé par un rappel sur les processus stochastiques, qui sont un outil très puissant pour la modélisation des phénomènes dynamiques, et les files d'attente classiques.
- Ensuite, nous avons fait une étude du modèles avec rappels à multiserveurs. Vu la complexité du modèle nous avons présenté l'une des approches permettant de faciliter les investigations sur les systèmes avec rappels, celle basée sur le principe de la troncature de l'espace d'états illimité.
- Finalement, nous avons étudié le modèle de files d'attente avec rappels M/M/C dans des cas particuliers: abandons exponentiels, arrivées négatives à l'aide de la méthode de Troncature Finie.

Le travail présenté dans cette mémoire nous permet d'envisager plusieurs axes de recherche tels que:

3.6. Troncature finie

- IL serait intéressant de considérer une structure plus complexe en prenant une distribution générale pour les temps de rappels.
- IL serait, aussi, intéressant de définir une nouvelle approche de résolution des Modèle à multiserveurs.

Annexe

Un algorithme pour le calcul numérique de la distribution stationnaire dans le système tronqué $S^{(M)}$

En utilisant la forme spécifique des équations (3.6.1)-(3.6.9), nous proposons un algorithme récursif pour le calcul des probabilités $P_{ij}^{(M)}$.

On introduit de nouvelles variables $r_{ij}^{(M)}$, $0 \leq i \leq K$, $0 \leq j \leq M$, comme suit:

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

Il est clair que, si on trouve les variables $r_{ij}^{(M)}$, alors on peut calculer les probabilités $P_{ij}^{(M)}$ comme suit:

$$P_{ij}^{(M)} = \frac{r_{ij}^{(M)}}{\sum_{i=0}^K \sum_{j=0}^M r_{ij}^{(M)}}.$$

Les variables $r_{ij}^{(M)}$ satisfont l'ensemble d'équations suivant (obtenu à partir des équations (3.6.1)-(3.6.8) pour les probabilités $P_{ij}^{(M)}$):

$$r_{0M}^{(M)} = 1; \tag{0.0.10}$$

$$(\lambda + i + j\theta + \eta_s)r_{ij}^{(M)} = \lambda r_{i-1,j}^{(M)} + (j+1)\theta r_{i-1,j+1}^{(M)} + ((i+1) + \eta_s)r_{i+1,j}^{(M)}, \quad (0.0.11)$$

$$0 \leq i \leq C-1, \quad 0 \leq j \leq M-1;$$

$$\begin{aligned} (\lambda + C + (i-C)\gamma + j\theta + \eta_s)r_{ij}^{(M)} &= \lambda r_{i-1,j}^{(M)} + (j+1)\theta r_{i-1,j+1}^{(M)} \\ &+ (C + (i+1-C)\gamma + \eta_s)r_{i+1,j}^{(M)}, \end{aligned} \quad (0.0.12)$$

$$C \leq i \leq K-2, \quad 0 \leq j \leq M-1;$$

$$\begin{aligned} (\lambda + C + (i-C)\gamma + j\theta + \eta_s)r_{ij}^{(M)} &= \lambda r_{i-1,j}^{(M)} + (j+1)\theta r_{i-1,j+1}^{(M)} \\ &+ (C + (i+1-C)\gamma + \eta_s)r_{i+1,j}^{(M)}, \end{aligned} \quad (0.0.13)$$

$$i = K-1, \quad 0 \leq j \leq M-1;$$

$$(\lambda + i + M\theta + \eta_s)r_{iM}^{(M)} = \lambda r_{i-1,M}^{(M)} + ((i+1) + \eta_s)r_{i+1,M}^{(M)}, \quad (0.0.14)$$

$$0 \leq i \leq C-1, \quad j = M;$$

$$\begin{aligned} (\lambda + C + (i-C)\gamma + M\theta + \eta_s)r_{iM}^{(M)} &= (C + (i+1-C)\gamma + \eta_s)r_{i+1,M}^{(M)} \\ &+ \lambda r_{i-1,M}^{(M)}, \end{aligned} \quad (0.0.15)$$

$$C \leq i \leq K-2, \quad j = M;$$

$$\begin{aligned} (\lambda + C + (i-C)\gamma + M\theta + \eta_s)r_{iM}^{(M)} &= (C + (i+1-C)\gamma)r_{i+1,M}^{(M)} \\ &+ \lambda r_{i-1,M}^{(M)}, \end{aligned} \quad (0.0.16)$$

$$i = K-1, \quad j = M;$$

$$(\lambda + C + (K-C)\gamma)r_{K,j}^{(M)} = \lambda r_{K-1,j}^{(M)} + (j+1)\theta r_{K-1,j+1}^{(M)} + \lambda r_{K,j-1}^{(M)}, \quad (0.0.17)$$

$$i = K, \quad 0 \leq j \leq M-1;$$

$$(C + (K-C)\gamma)r_{K,M}^{(M)} = \lambda r_{K-1,M}^{(M)} + \lambda r_{K,M-1}^{(M)}, \quad (0.0.18)$$

$$i = K, \quad j = M.$$

Calculons les variables $r_{ij}^{(M)}$ par groupe, chacun de taille $K+1$. Au début on calcule $r_{0M}^{(M)}, \dots, r_{KM}^{(M)}$, ensuite $r_{0,M-1}^{(M)}, \dots, r_{K,M-1}^{(M)}$ et ainsi de suite, jusqu'à ce qu'on trouve $r_{00}^{(M)}, \dots, r_{K,0}^{(M)}$.

1. Mettre $j = M$

1.1 Pour trouver le groupe $r_{0M}^{(M)}, \dots, r_{KM}^{(M)}$, réécrire les équations (0.0.14), (0.0.15) et (0.0.16) comme suit respectivement:

$$r_{i+1,M}^{(M)} = \frac{(\lambda + i + M\theta + \eta_s)r_{iM}^{(M)} - \lambda r_{i-1,M}^{(M)}}{((i+1) + \eta_s)}, \quad 0 \leq i \leq C-1;$$

$$r_{i+1,M}^{(M)} = \frac{(\lambda + C + (i-C)\gamma + M\theta + \eta_s)r_{iM}^{(M)} - \lambda r_{i-1,M}^{(M)}}{(C + (i+1-C)\gamma + \eta_s)}, \quad C \leq i \leq K-2;$$

$$r_{i+1,M}^{(M)} = \frac{(\lambda + C + (i-C)\gamma + M\theta + \eta_s)r_{iM}^{(M)} - \lambda r_{i-1,M}^{(M)}}{(C + (i+1-C)\gamma)}, \quad i = K-1.$$

Ce qui est équivalent à

$$r_{i,M}^{(M)} = \frac{(\lambda + (i-1) + M\theta + \eta_s)r_{i-1,M}^{(M)} - \lambda r_{i-2,M}^{(M)}}{(i + \eta_s)}, \quad 1 \leq i \leq C;$$

$$r_{i,M}^{(M)} = \frac{(\lambda + C + (i-1-C)\gamma + M\theta + \eta_s)r_{i-1,M}^{(M)} - \lambda r_{i-2,M}^{(M)}}{(C + (i-C)\gamma + \eta_s)}, \quad C+1 \leq i \leq K-1;$$

$$r_{i,M}^{(M)} = \frac{(\lambda + C + (i-1-C)\gamma + M\theta + \eta_s)r_{i-1,M}^{(M)} - \lambda r_{i-2,M}^{(M)}}{(C + (i-C)\gamma)}, \quad i = K.$$

D'après (0.0.10) $r_{0M}^{(M)} = 1$. De ce fait il est possible de calculer récursivement les variables $r_{1M}^{(M)}, \dots, r_{KM}^{(M)}$.

2. Mettre $j = j - 1$. Calculer les variables $r_{0j}^{(M)}, \dots, r_{Kj}^{(M)}$.

2.1 La dernière variable $r_{Kj}^{(M)}$, peut être trouvée à partir de l'équation (0.0.18) (si $j = M - 1$)

$$r_{K,M-1}^{(M)} = \frac{(C + (K-C)\gamma)r_{KM}^{(M)} - \lambda r_{K-1,M}^{(M)}}{\lambda}.$$

Où à partir de l'équation (0.0.17) avec j remplacé par $j + 1$ (si $j < M - 1$):

$$r_{Kj}^{(M)} = \frac{(\lambda + C + (K-C)\gamma)r_{K,j+1}^{(M)} - \lambda r_{K-1,j+1}^{(M)} - (j+2)\theta r_{K-1,j+2}^{(M)}}{\lambda}.$$

2.2 Pour trouver les variables $r_{0j}^{(M)}, \dots, r_{K-1,j}^{(M)}$, nous utilisons les équations (0.0.11), (0.0.12) et (0.0.13) pour $i = 0, \dots, C, \dots, K - 1$. Cet ensemble d'équations à la forme:

$$\alpha_i x_{i-1} + \beta_i x_i + \sigma_i x_{i+1} = \delta_i, \quad 0 \leq i \leq K - 1;$$

où

$$x_i = r_{ij}^{(M)}, \quad \alpha_i = -\lambda, \quad \delta_i = (j+1)\theta r_{i-1,j+1}^{(M)},$$

$$\sigma_i = -((i+1) + \eta_s), \quad \beta_i = \lambda + i + j\theta + \eta_s, \quad \text{pour } C \leq i \leq K-2;$$

$$\sigma_i = -(C+(i+1-C)\gamma+\eta_s), \quad \beta_i = \lambda+C+(i-C)\gamma+j\theta+\eta_s, \quad \text{pour } C \leq i \leq K-2;$$

$$\sigma_i = -(C+(i+1-C)\gamma), \quad \beta_i = \lambda+C+(i-C)\gamma+j\theta+\eta_s, \quad \text{pour } i = K-1;$$

et les valeurs $x_{-1} = 0$, $x_K = r_{Kj}^{(M)}$ sont connues.

Ces équations, qui sont des équations aux différences, produisent des solutions numériques comme les équations différentielles de second ordre. L'algorithme le plus efficace pour leur résolution peut être trouvé dans plusieurs livres sur les méthodes numériques (voir, par exemple [?]). D'après cet algorithme, on calcule en premier lieu les variables B_i , D_i , $0 \leq i \leq K-1$, par des formules récursives

$$B_0 = \beta_0, \quad D_0 = \delta_0,$$

$$B_i = \beta_i - \frac{\alpha_i \sigma_{i-1}}{B_{i-1}}, \quad D_i = \delta_i - \frac{\alpha_i D_{i-1}}{B_{i-1}}, \quad 1 \leq i \leq K-1;$$

et puis, à partir de l'équation

$$B_i x_i + \sigma_i x_{i+1} = D_i, \quad 0 \leq i \leq K-1;$$

on calcule récursivement (dans l'ordre inverse) les inconnues $x_{K-1}, \dots, x_C, \dots, x_0$.

Dans le cas présent, nous avons la procédure suivante:

- Calculer les variables B_{ij} , D_{ij} , $0 \leq i \leq K-1$, à l'aide des équations

$$B_{0j} = \lambda + j\theta + \eta_s;$$

$$B_{ij} = \lambda + i + j\theta + \eta_s - \frac{\lambda(i + \eta_s)}{B_{i-1,j}}, \quad \text{pour } 1 \leq i \leq C-1;$$

$$B_{ij} = (\lambda + C + (i-C)\gamma + j\theta + \eta_s) - \frac{\lambda(C + (i-C)\gamma + \eta_s)}{B_{i-1,j}}, \quad \text{pour } C \leq i \leq K-2;$$

$$B_{ij} = (\lambda + C + (i-C)\gamma + j\theta + \eta_s) - \frac{\lambda(C + (i-C)\gamma)}{B_{i-1,j}}, \quad \text{pour } i = K-1;$$

$$D_{0j} = 0;$$

$$D_{ij} = (j+1)\theta r_{i-1,j+1}^{(M)} + \frac{\lambda D_{i-1}}{B_{i-1,j}}, \quad \text{pour } 1 \leq i \leq K-1.$$

- Puis calculer récursivement $r_{ij}^{(M)}$, $0 \leq i \leq K - 1$, (Dans l'ordre inverse, commencer par $r_{Kj}^{(M)}$ connu à partir de l'étape 2.1) à l'aide des équations

$$r_{ij}^{(M)} = \frac{D_{ij} + (C + (i + 1 - C)\gamma)r_{i+1,j}^{(M)}}{B_{ij}}, \quad i = K - 1;$$

$$r_{ij}^{(M)} = \frac{D_{ij} + (C + (i + 1 - C)\gamma + \eta_s)r_{i+1,j}^{(M)}}{B_{ij}}, \quad i = K - 2, \dots, C;$$

$$r_{ij}^{(M)} = \frac{D_{ij} + ((i + 1) + \eta_s)r_{i+1,j}^{(M)}}{B_{ij}}, \quad i = C - 1, C - 2, \dots, 0.$$

3. La répétition de l'étape 2 tant que $j \geq 0$ permet d'obtenir toutes les variables $r_{ij}^{(M)}$ (ce qui est, successivement pour $j = M - 2, M - 3, \dots, 0$).

Puisque $P_{ij}^{(M)} = r_{ij}^{(M)} \cdot P_{0M}^{(M)}$, alors on aura:

$$P_{0M}^{(M)} = \frac{1}{\sum_{i=0}^K \sum_{j=0}^M r_{ij}^{(M)}}.$$

A présent, on peut calculer les probabilités $P_{ij}^{(M)} = r_{ij}^{(M)} \cdot P_{0M}^{(M)}$ et les principales caractéristique probabiliste du système tronqué $S^{(M)}$.

Bibliographie

- [1] M. S. Aguir. Modèles stochastiques pour l'aide à la décision dans les centres d'appel. Ecole Centrale, Paris, (2004).
- [2] N. K. Arrar. Problèmes de convergence, optimisation d'algorithme et analyse stochastique de systèmes de file d'attente avec rappels. Université Badji Mokhtar, Annaba, Algérie, Université Paris 1, Panthéon-Sorbonne, France, (2012).
- [3] J. Artalejo and A. Gómez-Corral. Advances in retrial queues. European Journal of Operation Research, (2008).
- [4] K. Belmadi, A. A. Diouane. Analyse d'un système $M/M/1$ avec rappels, serveur non fiable et clients impatientes. Université Blida 1, (2020).
- [5] M. El Haddad. Approximate analysis of $M/M/c$ retrial queue. Université Djillali liabes, Sidi Bel Abbas, (2020).
- [6] G.I. Falin and J.G.C. Templeton. Retrial queues. Chapman and Hall, (1997).
- [7] K. Fezza. Système de files d'attente avec dérobade, modèle $M/M/s$ avec dérobade. Université Dr Tahar Moulay, Saida, (2017).
- [8] Godunov and V.S. Ryabenkii. Difference scemes. North Holland, 47-50, (1987).
- [9] O. Ibe. Markov processes for stochastic modelling. Elseiver academy press, (2009).
- [10] D. L. Isaacson and R.W. Madsen. Markov chains, theory and application. Jhon Wiley sons, (1976).

-
- [11] S. Kacimi, N. Sadoudi. Système de files d'attente $M/G/1$ avec rappels. Université M'hamed Bougara, Boumerdès, (2016).
- [12] M. Kadi. Estimation et analyse d'un système de files d'attente avec dérobade. Université Djillali Liabes, Sidi Bel Abbas, (2017).
- [13] A. Lagnoux. Processus stochastique et modélisation. Université de Toulouse, (2013).
- [14] P. Le Gall. Les systèmes avec ou sans attente et les processus stochastiques. Tome I. Dunod, Paris, (1962).
- [15] F. Louis. Approximation dans le système de file d'attente avec rappels a serveur non fiable et clients impatientes. Université Mouloud Mammeri, Tizi-ouzou, (2017).
- [16] J. Loius Poss. Probabilité et statistique. Version 2.1 Mai 2003 p 74.
- [17] V.A. Malyshev and M.V. Menshikov. Ergodicity, continuity and analyticity of countable Markov chains. Proceeding of the Moscow Mathematical Society, 39, (1979).
- [18] A. G. Pakes. Some conditions for ergodicity and recurrence of Markov chains. Operations Research, 17: 1058-1061, (1969).
- [19] A. Ruegg. Processus stochastique. Press Polytechnique Romandes, Lausanne, Suisse, (1989).
- [20] S.N. Stepanov. Markov model with retrials: the calculation of stationary pergormance measures based on the concept of truncated. Mathematical and Computer Modelling, 30: 207-228, (1999).
- [21] R.L. Tweedie. Sufficient conditions for regularity, recurrence and ergodicity of Markov processes. Mathematical Proceeding of Cambridjge Philosophical Society, 78(1), (1975).
- [22] R.I. Wilkinson. Theories for tolltraffic engineering in USA. Bell System tech. J., 35:421-507, (1956).

- [23] K. ZaKKar. Stochastic processes (Stochastic II), University of Ulm, Institute of Stochastics, (2014).
- [24] N. Zidani. Les approximations dans l'analyse stochastique des systèmes de files d'attente avec rappels et multiserveurs. Université Badji Mokhtar, Annaba, (2017).