République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2019/2020

Introduction aux files d'attente

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar
Discipline : MATHEMATIQUES
Spécialité : Analyse Stochastique, Statistique des rocessus et Application (ASSPA)

par

Mlle. Mokhtaria **T**AHRI [1]

Sous la direction de

Dr. Wahiba BENZATOUT

Soutenue le 15/09/2020 devant le jury composé de

Dr. F. Mokhtari	Université Dr Tahar Moulay - Saïda	Président
Dr. W.BENZATOUT	Université Dr Tahar Moulay - Saïda	Encadreur
Dr. M. Kadi	Université Dr Tahar Moulay - Saïda	Examinateur
Dr. L. Yahiaoui	Université Dr Tahar Moulay - Saïda	Examinateur

1. e-mail : tahri.mokhatria1995@gmail.com

Remerciements

Avant tout nous remercions "Allah" tout puissant qui nous a donné le courage, la volonté et la force afin d'accomplir ce modeste travail. Merci de nous avoir éclairé le chemin de la réussite.

Je adresse nos plus vifs remerciements à **Dr. BENZATOUT** d'avoir bien voulu accepter d'être notre encadreur, de nous avoir aidé et dirigé pour la réalisation de ce Rapport ainsi pour la bienveillance dont elle a su faire preuve, par son dévouement et sa patience.

Je remercie **Dr. MOKHTARI** d'avoir accepté à présider le jury de mon mémoire. Je souhaite également remercier **Dr. KADI** et **Dr. YAHIAOUI** de m'avoir fait l'honneur d'accepter d'ê examinateur de mom mémoir. Leur remarque et précieux conseils out abouti à une amélioration du ce document

Je remercie **Dr. KANDOUCI** pour leur aide précieuse, leur soutien moral et leur disponibilité à tout moment. Je souhaite qu'ils puissent trouver ici le témoignage de notre très vive et très respectueuse reconnaissance et de notre sincère gratitude.

Je remercie les plus sincères reviennent à nos parents, que dieu les protègent et que la réussite soit toujours à notre portée pour que nous puissions leur combler de bonheur.

Je tenons à remercier de tout noter cœur l'ensemble des professeurs pour leur patience, conseils et leurs précieux cours qui ont enrichi nos connaissances et de je ont guidé durant tous ces années.

Merci de nous avoir montré les clés du succès : avoir confiance en soi et toujours tenter de se dépasser.

Je remercie l'ensemble de notre famille et nos amies pour leurs encouragements et leur compréhension. Finalement, nous remercions tous ceux ou celles qui ont contribué de près ou de loin à l'accomplissement de ce mémoire.

A vous grand Merci.

Table des matières

Introduction

Les phénomrènes de files d'attente pruvent être observer dans plusieurs situations réelles, quand les équipements de service (compteurs, ascenseurs, lignes téléphoniques, feux de circulation) ne peuvent pas immédiatement rendre la quatité ou le genre de service exigé par leurs utilisateurs. Ea outre, au niveau de byte en technologies modernes de manipulation de données (systèmes de communication, réseaux informatiques), on peut se produire des phénomènes d'attente qui sont en générea mois évidents, mais les effets dont au niveau de l'utilisateur ne sont habiteuellement as moins sérieux.
Tout à fait souvent, de tels effets de congestion peuvent être proportionnellement étudiés par des méthodes matématiques de la théorie des files d'attente. Adoptant la terminologie abstraite de théorie, l'entité rincipale dans le moèles est la station, où les clients qui ont besoin d' une certaine quantité de service arrivent. Typiquement, les modès des files d'attente sont à caractère stochastique, dans le sens où la durée des temps entre arrivées et de service des clients successifs n' pas exactement indiquée, mais décrit en termes de distribustion de probabilité. La nature stochastique des modèles de files d'attente reflète le fait que dans la plupart d'application, il est intrinsèquement aléatoire ou incertain quand la demande se produit et puor quelle quantité de service.

Le modèle classique dans la théorie des files d'attente se comose d' une file d'attente simple servi par un serveur simple. Les modèle de file simple et serveur simple ont été étudiés intensivenment dans la littérateur [16] our un traitement rigoureux des résults analytique principaux. Dans plusieurs situation, les modèle traditionnels de file-simple serveur-simple se avérés trés résussis dans la prévision des temps d'attente, des longueurs de file d'attante, et des probabilités de débordement d'amortissur. Cependant, dans la plupart des applictions récentes. La théorie des files d'attente fournit un outil très puissant et efficace pour la modélisation des systèmes admettant un phénomène d'attente. Cette théorie datent du début du XXème siècle par les travaux de l'ingénieur danois Agner Krarup Erlang(1878, 1929). Ses études sur le trafic téléphonique de Copenhague pour le mieuxgérer afin de déterminer le nombre de circuits nécessaires pour fournir un service téléphonique acceptable, sont considérées comme la première brique dans cette théorie [8]. Ensuite, les files d'attente ont été intégrés dans la modélisation des systèmes informatiques et aux réseaux de communication. Cette intégration dans ces domaines et d'autre a permet une évolution de cette théorie surtout dans l'évaluation des paramètres de performances des systèmes informatiques et aux réseaux de communication. Actuellement ce sont les applications dans le domaine de l'analyse de performance des réseaux (téléphone mobile, Internet, multimédia,...) qui suscitent le plus de travaux.
Depuis les travaux d'Erlang [2] Un grand nombre d'applications dans tous les domaines ont été mis en oeuvre et publiées. En 1953, David G. Kendall [2] a introduit la notation de Kendall pour décrire les caractéristiques d'un système de file d'attente. en 1957 d'une manière particulièrement élégante et efficace Jackson a traité certains réseaux de files d'attente. En 1961, Thomas L. Saaty [3], auteur de l'un des premiers livres complets sur la théorie des files d'attente. Ensuite c'est les contributions des mathématiciens Khintchine, Palm, Pollaczek et Kolmogorov qui ont vraiment poussés la théorie des files d'attente.

Mon mémoire organisé comme suit : Le chapitre 1 présente les notions de bases de l'étude des systèmes de files d'attente, à savoir les processus stochastiques :Processus de comptage, processus de renouvellement, Processus de Poisson, et processus de naissance et de mort. Dans le chapitre 2, nous introduison : la terminologie de la théorie des files d'attente. Certaines définitions et notations qui sont nécessaires dans l'étude des systèmes de files d'attente comme

(Notation de Kendall, la loi de Little ,...etc) sont notamment données. Ensuite nous étudions quelque modèles de files d'attente (M/M/1, M/M/1/K, M/M/c, M/M/1) et l'évaluation de leurs paramètres de performance avec des exmples . Le chapitre 3 nous présentons Variations et extensions : Systèmes avec rejet, Systèmes avec retour, Une file d'attente avec deux serveurs, Files d'attente prioritaires de préemption.Enfin chapitre 4 nous présentons un'application :Traffic routier.

Chapitre 1

Définitions et propriétés de base

1.1 Introduction

Les processus stochastiques décrivent l'évolution d'une grandeur aléatoire en fonction du temps (ou de l'espace). Il existe de nombreuses applications des processus aléatoires notamment en physique statistique[22].Les processus stochastiques ont outil trèe puissant pour la modéliation des phénomèenes aléatoires évoluant dans le temps. leurs utilisations pour la description et l'analyse des propriétés des systémes dynamiques(files d'attente, réseaux informatique et téléphoniques, physique, boilogie ou économie...etc.)

Définition 1.1.1. ***(Processus stochastique).***
Un processus stochastique est une famille $(X(t); t \in T)$ de variables aléatoires définie sur un même espace de probabilité.
Généralement $X(t)$ représente l'état du processus stochastique au temps t [32].

- *Si T est dénombrable, i.e. $\mathbb{T} \subseteq \mathbb{N}$ alors nous disons que $(X(t); t \in \mathbb{T})$ 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'ensembele des valeurs de $X(t)$ est applé l'espace d'état, qui peut également être soit discert (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 prossus à temps continu.

1.2 Processus de comptage

Définition 1.2.1. ***(Processus de comptage).***
Un processus stochastique $N(t)$ est appelé processus de comptage si $N(t)$ représente le nombre total des événements qui sont arrivés avant l'instant t. Tout processus de comptage vérifie les propriétés suivantes :

1. *Pour tout $t \geq 0$ le nombre $N(t)$ est à valeurs entiéres positives ;*
2. *La fonction $t \longmapsto N(t)$ est croissante ;*

3. *Pour tout couple* $(s;t)(0 < s < t)$ *, la différence* $N(t) - N(s)$ *représente le nombre d'évènement se produisant dans l'intervalle de temps* $]s;t]$.

Un processus de comptage est un processus discret à temps continu. Un second processus peut être associé au processus des temps d'occurrence ; le processus des temps d'interarrivées $\{W_n; n \in N_0\}$ *ou* $\forall n \in N_0$ *la variable aléatoire* W_n *est le temps d'attente entre les*$(n-1)^{ieme}$ *,* n^{ieme} *occurrences[31], c-à-d :*

$$W_n = T_n - T_{n-1}$$

avec T_n *est le temps d'arrivé du* n^{ime} *client.*

Proposition 1.2.1. *Les relations suivantes sont triviales tel que* $T_0 = 0$ *:*

1. $T_n = W_1 + W_2 + \ldots + W_n \forall n \geq 1$;
2. $N(t) = \sup\{n \geq 0 : T_n \leq t\}$;
3. $\mathbb{P}[N(t) = n] = \mathbb{P}[T_n \leq t < T_{n+1}]$;
4. $\mathbb{P}[N(t) \geq n] = \mathbb{P}[T_n \leq t]$;
5. $\mathbb{P}[s < T_n < t] = \mathbb{P}[n(s) < n \leq N(t)]$.

Preuve (1) : *Ona* $W_n = T_n - T_{n-1}$

$$\begin{aligned} W_1 + W_2 + \ldots + W_n &= T_1 - T_0 + T_2 - T_1 + \ldots + T_{n-1} - T_{n-2} + T_n - T_{n-1} \\ &= -T_0 + T_n \\ &= T_n \quad car \quad T_0 = 0 \end{aligned}$$

Dans la suite nous intéresserons souvent aux deux propriétés suivantes des processus stochastique :

Définition 1.2.2. ***(Processus à accroissements indépendants).***
Le processus stochastique $\{X_t, t \in T\}$*est appelé processus à accroissents indépendanrs si* $\forall k \in N_0$ *et* $\forall t_0 < t_1 < t_2 < t_3 < \ldots < t_k$*, les variables aléatoires* $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_k} - X_{t_{k-1}}$ *sont indépendants.*

Définition 1.2.3. ***(Processus à accroissements stationnaires).***

Le processus stochastique $\{X_t, t \in T\}$ *est dit à accroissements stationnaires si* $\forall h > 0$ *les variables aléatoires* $X_{t+h} - X_t$ *ont la même distribution* $\forall t \in \mathbb{R}_+$.

Définition 1.2.4. *Un processus de comptage*$\{N(t), t \geq 0\}$ *est un processus de Poisson d'intensité* $\lambda > 0$ *si*

- $N(0) = 0$;
- *le processus stochastique à accroissements indépendants ;*
- *le processus stochastique à accroissements stationnaires ;*
- $\forall 0 \leq s < t$ *les variables aléatoires* $N(t) - N(s)$*suit une loi de Poisson de paramètre* $\lambda(t-s)$.

1.3 Processus de renouvellement

Un processus de renouvellement a pour fonction le dénombrement des occurrences d'un phénomène donné, lorsque les délais entre deux occurrences consécutives sont des variables aléatoires indépendantes et identiquement distribuées.

Définition 1.3.1. ***(Processus de renouvellement).***
Un processus de comtage dont la suit des inter-arrivées forme une suit de variables aléatoires indépendantes et indentiquement distribuées s'appelle processus de renouvellement.

Définition 1.3.2. ***(Processus de renouvellement).***
Un processus de comptage pour lequel les temps entre deux arrivés consécutives sont des variables aléatoires i.i.d. s'appelle processus de renouvellement. Les temps de renouvellement (ou les temps de la n-ième arrivée) sont :

$$A_n = \sum_{i=1}^{n} a_i \quad n = 1, 2, ...$$

avec $a_i, i = 1, 2, \ldots$ *est le temps entre deux arrivées consecutives. Il est facile de voir que le nombre d'arrivées avant le tempst, i.e. le processus*

$$(N_t)_{t \in R_+} = \sup_k \{k : A_k \leq t\}$$

est un processus de comptage.

1.4 Processus de Poisson

Beaucoup de phénomènes naturels peuvent présenter des changements de valeurs à n'importe quel moment plutôt qu'à des dates fixes. On aura besoin pour modéliser cela de processus en temps continu c.à.d. des processus $\{X_t : t > 0\}$ indicés par la demi droite positive $[0; +\infty)$ Les processus que l'on analyse ici sont à valeurs entières. Le processus de Poisson sert à modéliser l'occurrence d'événements successifs. Chaque évènement est tel que dans un intervalle de temps $(t; t + \triangle t)$ avec $\triangle t$ petit Il peut servir à modéliser par exemple :

- les appels téléphoniques arrivant dans une centrale ;
- les temps d'arrivée de clients à une caisse ;
- les temps d'occurrence de sinistres à dédommager par une compagnie d'assurance . . . etc.

1.4.1 la distribution de poisson

la loi de Poisson est attribuée à Simeon D. Poisson, mathématician français $(1781 - 1840)$. Cette loi fut proposée par Poisson dans un ouvrage qu'il publia en 1837 sous le titre " Recherche sur la probabilité de hugements en matiére criminelle et en matière civile".

la distribution de Poisson de paramatère $\lambda > 0$ et donnée par :

$$\mathbf{P}_k = \frac{\lambda^k}{k!} e^{-\lambda} \quad pour \quad k = 0, 1, 2 \cdots \tag{1.1}$$

Soit X une variable aléatoire ayant la distribution de poisson dans(1.1) nous évaluons la moyenne, où premier, par :

$$\mathrm{E}[X] = \sum_{k=0}^{+\infty} k\mathrm{P}_K$$

$$= \lambda.$$

Pour trouver la variance, on détermine d'abord

$$\mathrm{E}[X(X-1)] = \sum_{k=2}^{+\infty} k(k-1)\mathrm{P}_k$$

$$= \lambda^2;$$

alors

$$\mathrm{E}[X^2] = \mathrm{E}[X(X-1)] + \mathrm{E}[X]$$

$$= \lambda^2 + \lambda.$$

tandis que

$$\sigma_x^2 = Var[X]$$

$$= \lambda^2 + \lambda - \lambda^2 = \lambda.$$

1.4.2 la distribution de exponentielle

Soit t une variable aléatoire avec $t > 0$ qui suit une distribution exponentielle. La densité de probabilité de t est $f(t) = \mu e^{-\mu t}$ et la distribution cumulé corresponante est $F(t) = 1 - e^{-\mu t}$. L'esprérance et la variance de t sont :

$$\mathrm{E}[X] = \frac{1}{\mu} \quad et \quad V(t) = \frac{1}{\mu^2}$$

respectivement.

1.4.3 Le lien entre la distribution de exponentielle et la distribution de poisson

La densité de probabilité d'une distribution exponentielle $f(t) = \mu e^{-\mu t}$ Supposons τ est exponentielle avec espérance $\frac{1}{\mu}$, et n est poisson de moyenne μ on a :

$$\begin{aligned}\mathbb{P}(\tau > t) &= 1 - F(t) \\ &= e^{-\mu t} \\ &= \mathbb{P}(n = 0 \quad en \quad t) \\ &= \mathbb{P}(0, t).\end{aligned}$$

Nous $\mathbb{P}(n, t)$ la probabilité d'avoir n unités dans le temps t.

$$\begin{aligned}\mathbb{P}(0, t) &= e^{-\mu t} \\ \mathbb{P}(1, t) &= \int_{\tau=0}^{t} \mathbb{P}(0, \tau) f(1 - \tau) d\tau = \mu t e^{-\mu t} \\ \mathbb{P}(2, t) &= \int_{\tau=0}^{t} \mathbb{P}(1, \tau) f(1 - \tau) d\tau = (\mu t)^2 e^{-\mu t}/2! \\ \ldots \\ \mathbb{P}(n, t) &= \int_{\tau=0}^{t} \mathbb{P}(n - 1, \tau) f(1 - \tau) d\tau = \mu t)^n e^{-\mu t}/n!\end{aligned}$$

1.4.4 Propriété sans mémoire de la distribution exponentielle

Une variable aléatoire X est dite sans mémoire (ou sans usure) si :
$\forall s, t \geq 0$

$$\mathbb{P}(X > t + s / X > t) = \mathbb{P}(X > s)$$

Si X est la durée de vie d'un matériel quelconque l'équation précédante s'interprète de la manière suivante, sachant le matériel en état de bon fonctionnement au temps t, la loi de probabilité de sa durée de vie future est la même que celle de sa durée de vie initiale.
En d'autres termes, le matériel ne s'use pas.

1.4.5 loi Poisson et loi exponentielle

Définition 1.4.1. *Une variable aléatoire X à valeur entières suite loi d Poisson de paramètre $\lambda > 0$ si :*

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda k}$$

Définition 1.4.2. *Une variable aléatoire Y à valeur réelles strictement positives suit une loi exponentille de paramètre $\mu > 0$ si :*

$$\forall t > 0, \quad \mathbb{P}(Y = T) = \mu e^{-\mu t}$$

1.4.6 Processus de Poisson

De tels phénomènes peuvent se définir la famille $(T_n)_{n\in\mathbb{N}^*}$ des d'arrivées qui sont des variables aléatoires. Mais on peut aussi le faire a partir du processus de comptage $(N_t)_{n\in\mathbb{R}}$. N_t est le nombre d'événements apparus jusqu'a l'instant t.

Définition 1.4.3. ***(Processus de Poisson).***

Soit (X_n) une suite de variables aléatoires indépendantes et indentiquement distribuées de loi exponentielle $\varepsilon(\lambda)$ avec $\lambda > 0$ si
$S_n = X_1 + X_2 + \cdots + X_n$, $N_0 = 0$ *et pour tout $t > 0$*

$$N_t = \sum_{n=1}^{\infty} \mathbb{1}_{S_n < t}$$

(N_t) *est un processus de Poisson d'intensité λ.*

Définition 1.4.4. ***(Processus de Poisson).***

Un processus de poisson $N = (N_t)$ d'ntensité λ est un pocessus de comptage à trajectoires continues à droit [14] tel que :

1. $N(0) = 0$,
2. N *est un processus à accroissement indépendant et stationnaires,*
3. *pour tout $t > 0$, $N_t$suit la loi de poisson $\mathcal{P}(\lambda t)$.*

Proposition 1.4.1. $(N_t)_{t\geq 0}$ *est un processus de Poisson si es seulement si les quater propriétés suivantes sont vérivifiées :*

1. $N0 = 0$,
2. *pour tout $t > 0$, N_t une loi de poisson d'espérance λt,*
3. *pour tout $t > s > 0$, $N_t - N_s$ suit une loi de poisson d'espérance $\lambda(t - s)$,*
4. *pour tout $0 < t_1 < t_2 < \cdots < t_n$ les variables aléatoire* $N_{t_1}, N_{t_2} - N_{t_1}, \cdots, N_{t_n} - N_{t_{n-1}}$ *sont indépendantes.*

Caractérisation d'un procssus de Poisson par ses temps d'arrivé :

Soit T_n l'instant de la n^{ime} arrivé : $T_n = inf\{T \geq 0, N_t = n\}$ et W_n est le temps séparant le $(n-1)^{ieme}$ évènement du n^{ieme} évènement pour $n \in \mathbb{N}^* W_n = T_n - T_{n-1}$ (en convenant T_0) on a :

$$T_n = \sum_{i=1}^{n} W_i,$$

$$N_t = max\{n \geq 0, T_{n \leq t}\}.$$

1.5 Processus de naissance et de mort :

Ces processus permettent de façons générale décrire l'évolution temporelle de la taille d'une population d'un type donné.
Dans le cas d'systéme d'attente, on considère par exemple des population comprenant tous les clients qui sont dans le systèmeà l'instant t.
Les processus de de naissance et de mort sont des processus stochastique à temps continu et à espace d'étars discret $n = 0, 1, 2, \ldots$ Ils sont sans mémoire, et à partir d'un état donné n, des transitions ne sont possibles que vers l'un ou l'outre des états voisins

$$(n+1) \ \ et \ \ (n-1) \ \ pour \ \ n \geq 1$$

Les files d'attente de type MarKovien(M/M) sont des cas particuliers très importas de processus de nissance et de mort.

Définition 1.5.1. ***(Processus de naissance et de mort).***
On peut réaliser un processus de naissance et de mort de façon suivante :

- ◇ *Les arrivées et les départs d'entités obéissent à des lois exponentielle de taus 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).*

  C'est le point de départ de la théorie des files d'attent. On introduit les données suivantes :

  - λ_n *: taux de naissance quand le nombre de population à* n.
  - μ_n *: taux de mort quand le nombre de population à* n*[32]. .*

Chapitre 2

Les modèles de file d'attente

2.1 Système de files d'attente :

La théorie des files d'attente a commencé en 1909 avec les travaux de recherches de l'ingénieur danois Agner Krarup Erlang (1878, 1929) sur le trafic téléphonique de Copenhague pour déterminer le nombre de circuits nécessaires afin de fournir un service téléphonique acceptable. Par la suite, les files d'attente ont été intégrés dans la modélisation de divers domaines d'activité[12]. On assista alors à une évolution rapide de la théorie des files d'attente qu'on appliqua à l'évaluation des performances des systèmes informatiques et aux réseaux de communication. Les chercheurs oeuvrant dans cette branche d'activité ont élaboré plusieurs nouvelles méthodes qui ont été ensuite appliquées avec succés dans d'autres domaines, notamment dans le secteur de la fabrication. On a aussi constaté une résurgence des applications pratiques de la théorie des files d'attente dans des secteurs plus traditionnels de la recherche opérationnelle, un mouvement mené par Peter Kolesar et Richard Larson[6]. Grace à tous ces développements, la théorie des files d'attente est aujourd'hui largement utilisé et ses applications sont multiples.

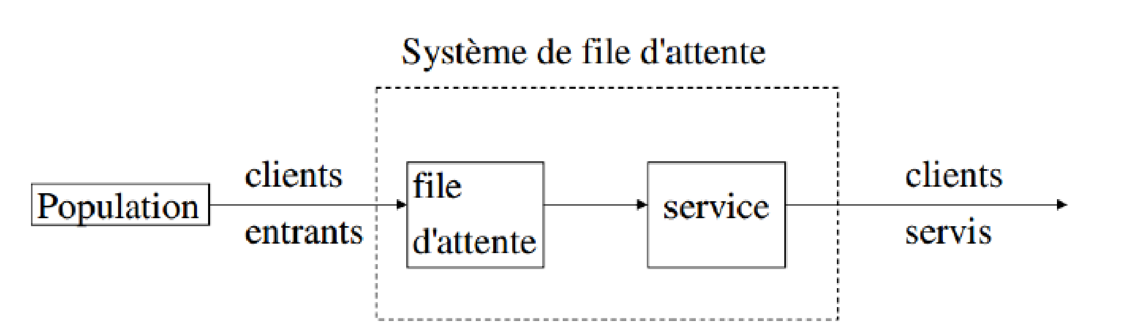
2.1.1 Définitions :

Définition 2.1.1. ***(File d'attente).***
l'ensemble des clients qui attendent d'être servis, à l'exclusion de celui qui est en train de se faire servir.

Définition 2.1.2. ***(Système d'attente).***
l'ensemble des clients qui font la queue, y compris celui qui se fait servir.
Le phénomène d'attente s'étend à tous les clients possibles (dans le cas de systèmes bouclés, où les mêmes clients reviennent plus tard à l'entrée par exemple les machines qui tombent en panne dans un atelier, le nombre des clients est, en général, fini). Ces appellations se généralisent et prennent surtout leur intérêt dans les situations où il existe plusieurs stations et plusieurs files d'attente.

2.2 Structure de base :

- Population : La population constitue la source de clients potentiels. Elle est caractérisée par son nombre d'élément (fini ou infini).
- File d'attente : La file d'attente est caractérisée par le nombre maximum permis de clients en attente (fini ou infini)



- Clients : Les clients (issus de la population) se joignent au système avec un taux moyen d'arrivée.
- Service : Le service peut être assuré par un ou plusieurs serveurs. Le temps qui s'écoule entre le début et la fin de service d'un client est dénoté le temps de service suivant une loi de probabilité. Le taux de service est une autre caractéristique du système.
- Interarriée : est l'itervalle de temps séparant l'arrivée de desux clients successifs.
  On suppose que ces Inter-arrivées sont indépendantes les une des autres et suivent la même loi de probabilité
- Stratégie de service : La stratégie de service réfère à l'ordre selon laquelle les clients sont servis : premier arrivée premier servi ; au hasard, selon des priorités. . .

Définition 2.2.1. ***(File d'ttente simple).***
Une file simple (ou station) est un système constitué d'un ou plusieurs serveurs et d'un espace d'attente. Les clients arrivent de l'extérieur, patientent éventuellement dans la file d'attente, reçoivent un service, puis quittent la station[19]. Afin de spécifier complètement une file simple, on doit caractériser le processus d'arrivé des clients, le temps de service ainsi que la structure et la discipline de service de la file d'attente.

- Processus d'arrivé :
  L'arrivée des clients à la station sera décrite àl'aide d'un processus stochastique de comptage $(N_t)_{t\geq 0}$
  Si A_n désigne la variable aléatoire mesurant l'instant d'arrivée du n^{ieme} client dans le système, on aura ainsi :

$$A_0 = 0 \ \ et \ \ A_n = inf\{t; N_t = n\}.$$

  Si T_n désigne la variable aléatoire mesurant le temps séparant l'arrivée du $(n-1)^{ieme}$ et du n^{ieme} client [20] , on a alors :

$$T_n = A_n - A_{n-1}.$$

- Le processus de service :
  qui compris (le nombre de serveus etla loi probabiliste décervant la durée des services)
- Temps de service :
  On considère une file à serveur unique.
  On note D_n la variable aléatoire mesurant l'instant de départ du n^{ieme} client du système et Y_n la variable aléatoire mesurant le temps de service du n^{ieme} client (temps séparant

le début et la fin du service).
Un instant de départ correspond toujours à une fin de service, mais ne correspond pas forcément à un début de service. Il se peut en effet qu'un client qui quitte la station laisse celle-ci vide. Le serveur est alors inoccupé jusqu'à l'arrivée du prochain client.
On considèrera uniquement des stations dont les temps de service consécutifs sont décrits par des variables Y_n indépendantes et identiquement distribuées (i.i.d.).

2.3 Classifcation des systèmes d'attentes :

2.3.1 Notation de Kendall :

Pour la classification des systèmes de files d'attente, on a recours à une notation symbolique introduite par Kendall, comprenant six symboles [7] :

$$A/B/s/N/D/O$$

avec

- A : indique le processus d'arrivée des clients. Les codes utilisés sont :
  - M (Markov) : inter-arrivées des clients sont indépendamment, identiquement distribuées selon une loi exponentielle. Il correspond à un processus de Poisson ponctuel (propriété sans mémoire),
  - D (Répartition déterministe) : les temps inter-arrivées des clients ou les temps de service sont constants et toujours les mêmes.
  - GI (général indépendant) : Les inter-arrivées des clients ont une distribution générale (il n'y a aucune hypothèse sur la distribution mais les interarrivées sont indépendentes et identiquement distribuées),
  - G (général) : Inter-arrivées de 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.
- B : décrit la distribution des temps de service d'un client. les codes sont les mêmes que A,
- s : nombre de serveurs,
- N : capacité de la file (c'est le nombre de places dans le système en d'autre tèrme c'est le nombre maximal de clients permis dans le système Bcompris ceux en service.
- D : population des usagers.
- O : 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 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.

2.3.2 Loi de Little :

La loi de little (1961) est une relation très général qui s'applique à une grande classe de sysémes. Elle ne concerene que le régime premanet du système. Aucune hypothèse sur les variables aléatoires qui caractérisent le systéme (temps d'inter aarivées, temps de service,...) n'est nécessaire. La seule condition d'application de la loi de Little est que système soit stable.

$$d_e = d_e = d$$

La loi de Little s'exprime dans le théorème suivant :

Théorème 2.1. ***(Formule de Little).*** *Le nombre mouyen de clients L, le temps moyen passé dand le système W et le débit moyen d'un système stable d en régime prmanent se relient de la façon suivants :*

$$L = W \times D$$

Remarque 2.3.1. *La loi de Little s'applique à tous les modéles de file d'attente rencontrés en pratique (pas seulement à file $M/M/1$)*

2.4 Terminologie et notations :

En lien avec la loi exponentielle[4] :

- λ_n : :Le taux moyen d'arrivées (espérance mathématique de nomber d' arrivées par unité de temps).
- $\frac{1}{\lambda_n}$: Le temps moyen entre les arrivées lorque les clients sont dans le système.
- μ_n :Taux moyen de service d'un client
- $\frac{1}{\mu_n}$:Temps moyen de service du n^{ieme} client.

L'analyse d'un système de file d'attente dépends de l'état initial et du temps écoulé. C'est la situation transitoire où l'étude est très complexe. Dans la théorie des files d'attente l'étude se fait une fois que le système atteint sa situation d'équilibre ; où les états du système sont essentiellement indépendants de l'état initial et du temps déjà écoulé. On suppose que le système est en opération depuis un très long moment.

2.5 En situation d'équilibre on note :

- P_n : Probabilité qu'il y ait n clients dans le système.
- L_s : Nombre moyen (espérance mathématique) de client dans le système.
- L_q : Nombre de clients dans la file d'attente (excluant ceux qui sont dans le service).
- W_s : Le temps moyen passée par un client dans le système.
- W_q :Le temps moyen passée par un client dans la file (excluant le temps de service).
- c :Nombre de serveurs[24].

2.6 Modèle d'attente M/M/1

La file d'attente $M/M/1$ est un modèle caractérisé par des arrivées suivants un processus de Poisson de taux λ, Temps de service exponentielle de paramètre μ et un seul serveur. Les clients arrivent à la station selon un processus de Poisson de taux λ, si le serveur est vide le client est pris en charge immédiatement sinon il rejoint la file d'attente (de capacité illimitée et discipline FIFO), les temps des inter-arrivées sont indépendants. Ces clients reçoivent le service selon une distribution exponentielle de paramètre μ. Ces temps de service sont également supposés indépendants[21]. En outre, toutes les variables aléatoires concernés sont censés être indépendants les uns des autres. Les clients sont servis dans leur ordre d'arrivée. Cette file est un cas particulier du processus de naissance et de mort

$$\forall n \in \mathbb{N},\ \lambda_n = \lambda, \mu_n = \mu.$$

ie indépendants du nombre de client dans le système. sous l'hypothèse que $\lambda < \mu$ (le taux d'arrivée est plus petit que le taux de service)
on a :

$$\rho = \frac{\lambda}{\mu}.$$

c'est l'intensité du trafic, avec $\rho < 1$:

$$\begin{cases} \pi_n = \pi_0 \rho^n, & ; \\ \\ \pi_0 = \frac{1}{\sum_{n=0}^{+\infty} \rho^n} = 1 - \rho, & . \end{cases}$$

Remarque 2.6.1. *On note π_0 la probabilité d'être servi immédiatement a l'arrivée. La condition $\rho < \mu$ est bien entendu équivalente à la condition $\rho < 1$ Ainsi, un régime stationnaire peut exister si (et seulement si) l'intensité de trafic est inférieure a cent pour cent. on voit aussi le rôle de ce paramètre lorsque l'on calcule le nombre moyen de clients en régime stationnaire.*
avec

$$\sum_{n=0}^{\infty} (\frac{\lambda}{\mu})^n = \frac{1}{1 - \frac{\lambda}{\mu}}.$$

π_n *: c'est la probabilité d'avoir n client dans le système.*

$$\pi_n = \rho^n \pi_0$$

avec

$$\pi_0 = 1 - \rho$$

- Débit :
  Ici $d = \lambda$ car $\lambda_n = \lambda$ pour tout $n \geq 0$.

$$d = \sum_{n=0}^{\infty} \pi_n \mu = (1 - \pi_0)\mu = \rho\mu = \lambda;$$

  donc :

$$d = d_e = d_s.$$

- Taux d'utilisation du serveur U :

$$U = 1 - \pi_0 = \rho.$$

- Nombre moyen de clients dans le système L :

$$\sum_{n=0}^{\infty} n\pi_n = \frac{\rho}{1-\rho} = \frac{\frac{\lambda}{\mu}}{1-\frac{\lambda}{\mu}} = \frac{\lambda}{\mu - \lambda}.$$

  Remarque 2.6.2. *Plus que l'intensité du trafic se rapproche de un, plus que la longueur moyenne de la file d'attente L tend vers l'infinie.*

- Temps moyen de séjour W :
  Ce paramètre est obtenu en utilisant la loi de Little :

$$W = \frac{L}{d} = \frac{1}{\mu(1-\rho)}.$$

  qui peut se décomposer en :

$$W = \frac{1}{\mu} + \frac{\rho}{\mu(1-\rho)}$$

- La durée moyenne d'attente en régime stationnaire :

$$\mathrm{E}(W) = \frac{\rho}{\lambda(1-\rho)}.$$

Proposition 2.6.1. *Soit $0 < \lambda < \mu$. Dans le système $M/M/1$ la variable aléatoire W égale à la durée de séjour des clients dans le système en régime stationnaire suit une loi exponentielle de paramètre :*

$\mu(1-\rho) = \mu - \lambda$

- *Démonstration[11] : Temps moyen passé dans la file d'attente W_q :*

$$W_q = \frac{\rho}{\mu(1-\rho)}.$$

- *Nombre moyen de clients dans la file d'attente L_q :*

$$L_q = \lambda W_q = \frac{\rho^2}{1-\rho}.$$

2.7 Modèle d'attente M/M/c

On considère un système identique à la file M/M/1 excepté qu'il comporte C serveurs identiques et indépendants les uns des autres. On conserve les hypothèses :
le processus d'arrivée des clients dans la file est un processus de Poisson de taux λ et le temps de service d'un client est une variable aléatoire exponentielle de taux μ (pour chacun des serveurs). Ce système est connu sous le nom de file M/M/c [13]. L'espace d'états E est comme pour la M/M/1 infini. Dans ce cas aussi, le processus modélisant le nombre de clients dans le système est un processus de naissance et de mort avec [1] :

$$\lambda_n = \lambda$$

$$\mu_n = \begin{cases} n\mu & \text{pour : } \quad n = 1, 2, ..., c-1; \\ \\ c\mu & \text{pour.} \qquad\quad n \geq c. \end{cases}$$

La condition de stabilité de ce modèle est :

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

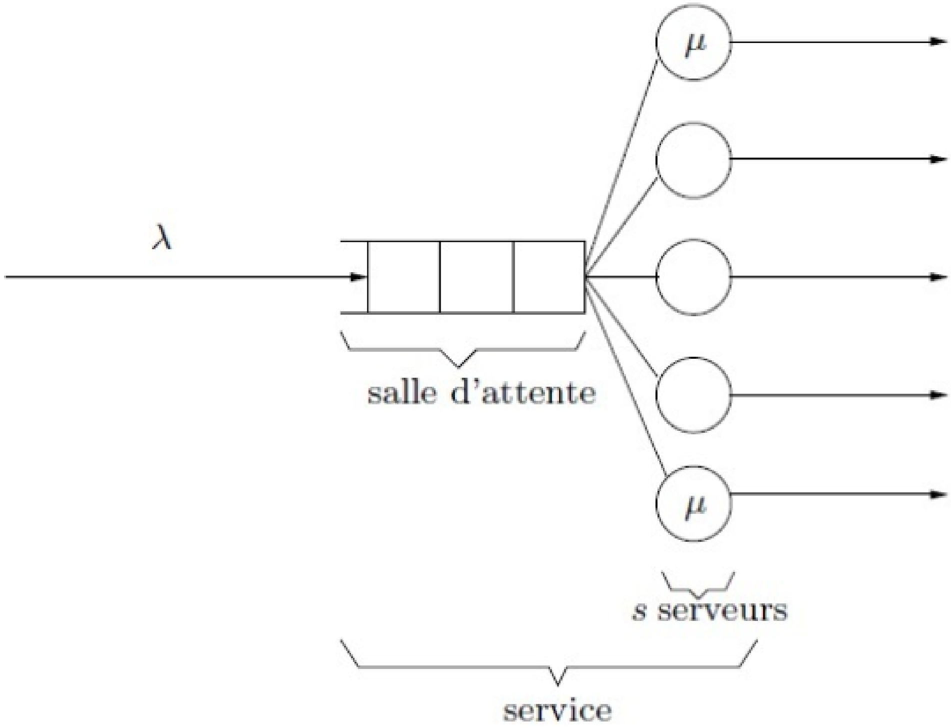
FIGURE 2.7 – Reprsentation schmatique d'une file M/M/c

Théorème 2.2. *pour calculer π_n le système doit être stable ($\lambda < c\mu$ donc $\rho < c$) qui exprime le fait que le nombre moyen de clients qui arrivent à la file par unité de temps doit être inférieur au nombre moyen de clients que les serveurs de la file sont capables de traiter par unité de temps, de la manière suivante :*

- Temps moyen de séjour W :

  Le temps moyen de séjour d'un client se décompose en un temps moyen dans la file d'attente, plus un temps moyen de service. Il suffit alors d'appliquer la loi de Little à la seule file :

$$W = W_q + W_c = \frac{L_q}{d} + \frac{1}{\mu} = \frac{L_q}{\lambda} + \frac{1}{\mu}$$



Avec : $d = \lambda$
Il reste alors à calculer le nombre moyen de clients en attente dans la file, L_q :

$$\begin{aligned} L_q = \sum_{n=c}^{+\infty}(n-c)\pi_n &= \sum_{n=c}^{+\infty}(n-c)\frac{\rho^n}{c!c^{n-c}}\pi_0; \\ &= \frac{\rho^{c+1}}{c!c}\sum_{n=c}^{+\infty}(n-c)(\frac{\rho}{c})^{n-c-1}\pi_0; \\ &= \frac{\rho^{c+1}}{c!c}\frac{1}{(1-\frac{\rho}{c})^2}\pi_0 = \frac{\rho^{c+1}}{(c-1)}. \end{aligned}$$

On en déduit le temps moyen de séjour

$$W = \frac{\rho^c}{\mu(c-1)!(c-\rho)^2\pi_0} + \frac{1}{\mu}.$$

- Nombre moyen de clients dans le système L :

Le nombre moyen de clients s'obtient alors par application de la loi de Little à l'ensemble de la file :

$$L = W \times d = W \times \lambda = \frac{\rho^{c+1}}{\mu(c-1)!(c-\rho)^2}\pi_0 + \rho.$$

2.8 Modèle d'attente M/ M/1/ k

On considère un système à serveur simple identique à la file $M/M/1/k$ excepté que la capacité de la file d'attente est finie. On a donc toujours les hypothèses suivantes :

– le processus d'arrivée des clients dans la file est un processus de Poisson de taux λ.

– le temps de service d'un client est une variable aléatoire exponentielle de taux μ.

– Soit K la capacité de la file d'attente : c'est le nombre maximal de clients qui peuvent être présents dans le système, soit en attente, soit en service.

Quand un client arrive alors qu'il y a déjà Kclients présents dans le système, il est perdu. Ce système est connu sous le nom de file $M/M/1/K$. L'espace d'états E est maintenant fini : $E = \{0, 1, 2, ..., K\}$. La capacité de la file étant limitée, même si les clients arrivent en moyenne beaucoup plus vite que ce que le serveur de la file est capable de traiter, dés que celle-ci est pleine, les clients qui se présentent sont rejetés [26]. Le nombre de clients dans la file ne peut donc jamais "partir" à l'infini. De plus, dés qu'un client est autorisé à entrer, il sortira un jour et son temps de séjour dans la file est fini, puisqu'il correspond au temps de service de tous les clients devant lui et que ce nombre est limité par K . Sur un très long, le débit de sortie sera donc bien égal au débit d'entrée, ce qui correspond bien à la stabilité inconditionnelle du système. Donc ce processus est considéré comme un processus de naissance et de mort avec :

◇ un taux de naissance $\lambda_n = \lambda$, pour tout $n < K$

◇ et le taux de mortalité $\mu_n = \mu$, pour tout $i \neq 0$.

soit $\pi_{(n)}, n = 0, 1, 2, ..., K$, la probabilité pour qu'il ait n clients dans le système à l'instant t.

$$\lambda \pi_0 = \mu \pi_K$$

$$(\lambda + \mu)\pi_n = \lambda \pi_{n-1} + \mu \pi_{n+1}$$

pour $n = 0, 1, 2, ..., K - 1$

$$\lambda \pi_{K-1} = \mu \pi_K.$$

puisque la capacité est limitée, nous obtenons un régime stationnaire indépendant des conditions initiales quelle que soit la valeur de l'intensité de trafic ρ

Théorème 2.3. *Le calcul de*π_n *se fait de la manière suivante [18] :*

$$\pi_n = \begin{cases} \frac{(1-\rho)\rho^n}{1-\rho^{K+1}} & pour : \quad n \leq K; \\ \\ 0 & pour. \quad n > K. \end{cases}$$

Cas particulier $\rho = 1$

dans ce cas

$$\pi_n = \frac{1}{K+1}$$

$$\pi_n = 0 \ \ pour \ \ n > K$$

- Taux d'utilisation du serveur $U(K)$:
  Le taux d'utilisation du serveur est la probabilité pour que le serveur de la file soit occupé donc au moins il y a un client dans la file

$$U(K) = \sum_{n=1}^{K} \pi_n = 1 - \pi_0$$

$$= \rho \frac{1-\rho^k}{1-\rho^{K+1}}$$

Remarque 2.8.1.

$$\lim_{K \longrightarrow +\infty} U_K = \begin{cases} \rho \quad pour: \quad \rho < 1; \\ 1 \quad pour. \quad \rho > 1. \end{cases}$$

qui représente le taux d'utilisation du serveur Dans le cas $M/M/1$

- Nombre moyen de clients L :

$$L = \sum_{n=0}^{K} i\pi_n = \frac{1-\rho}{1-\rho^{K+1}} \sum_{n=1}^{K} n\rho^n;$$

$$= \frac{\rho(1-\rho)}{1-\rho^{K+1}} \sum_{n=1}^{K} n\rho^{n-1};$$

$$\frac{\rho}{1-\rho} \frac{1-(K+1)\rho^K + K\rho^{K+1}}{1-\rho^{K+1}};$$

Pareille que le taux d'utilisation du serveur, lorsque$K \longrightarrow +\infty$ on obtient

$$L = \frac{\rho}{1-\rho}.$$

qui représente le nombre moyen de clients pour la file $M/M/1$

- Temps moyen de séjour W :
  Le calcul du temps moyen de séjour W ce fait on applicant la formule de Little qui relie Nombre moyen de client L et débit d de la file :
$$W = \frac{L}{d}.$$

2.9 Modèle d'attente M/M/ ∞

Description du modèle :

Pour ce modèle de file d'attente, le système est composé d'un nombre illimité de serveurs identiques et indépendants les uns des autres. Dés qu'un client arrive, il est immédiatement servi (c'est le cas ou il n'y a pas d'attente). Dans cette file les clients arrivent à des instants $0 < t_1 < t_2 < \ldots$ formant un processus de Poisson de taux λ et les temps de service sont exponentiels de taux μ . le taux de transition d'un êtat n quelconque vers l'état $n-1$ est égal à $n\mu$ et correspond au taux de sortie d'un client parmi les n clients en service [5]. De même, le taux de transition d'un état n vers l'état $n+1$ est égal à λ qui correspond au taux d'arrivée d'un client, donc c'est un processus de naissance est de mort avec :

$$\forall n \in \mathbb{N}, \lambda_n = \lambda, \mu_n = n\mu$$
$$\pi_n = \frac{\rho^n}{n!}\pi_0.$$

avec

$$\pi_0 = \frac{1}{\sum_{n=0}^{+\infty} \frac{\rho^n}{n!}} = \frac{1}{e^\rho}.$$

Notons que la série $\sum_{n=0}^{+\infty} \frac{\rho^n}{n!}$ converge pour toutes valeurs de ρ , ce qui est cohérent avec la stabilité inconditionnelle de la file. On obtient finalement :

$$\pi_n = \frac{\rho^n}{n!}e^{-\rho} \ \ pour \ \ n = 1, 2, \ldots$$

- Le débit d :
  Le service s'effectue avec un taux $n\mu$ dans chaque état où le système contient n clients :
$$d = \sum_{n=0}^{+\infty} n\pi_n\mu = e^{-\rho}\sum_{n=1}^{+\infty} \frac{\rho^n}{(n-1)!}\mu = e^{-\rho}\rho e^\rho \mu = \lambda.$$
  On retrouve la stabilité inconditionnelle de la file.
- Taux d'utilisation du serveur U :
$$U = 1 - \pi_0.$$

– Nombre moyen de clients dans le système L :

$$d = \sum_{n=1}^{+\infty} n\pi_n = e^{-\rho} \sum_{n=1}^{+\infty} \frac{\rho^n}{(n-1)!} = e^{-\rho}\rho e^{\rho} = \rho.$$

– Temps moyen de séjour W :
Ce paramètre est obtenu en utilisant la loi de Little :

$$W = \frac{L}{d} = \frac{\rho}{\lambda} = \frac{1}{\mu}.$$

2.9.1 Exemples

Exemple 2.9.1. *Taux de service d'un système informatique*

- *Objectif : choisir taux de service* μ *pour que temps systè moyen d'une tâche inférieur à* $0,5$ *seconde*
- *Taux d'arrivée de* $\lambda = 1$ *tâches /seconde*
- *Arrivée selon loi de poisson et temps de service selon loi exponentielle (système M/M/1)*
- *En utilisant* $W = \frac{L}{D} = \frac{1}{\mu(1-\rho)}$ *avec* $\rho = \lambda/\mu = 1/\mu$

$$W = \frac{1}{\mu(1 - 1/\mu)} < 0,5 \;\Rightarrow\; 0,5\mu < 1,5$$

- *Donc, il faut que*$\mu > 3$ *tâches /seconde.*

Exemple 2.9.2. *Teaux d'arrivée maximum d'un système de télécommunication*
- *Objectif : choisir taux d'arrivée* λ *de messages maximum pour que temps d'attente moyen d'un message inférieur à* 1 *seconde*
- *Messages de longeur de* K *bits, distribué selon loi exponentialle de moyenne de* 600 *bits*
- *Vitesse de transmission du système de* $C1200$ *bits/s*
- *Messages requiert enmoyenne* K/C *secondes pour être transmis, donc* $\mu = K/C = 600/1200 = 2$ *messages/seconde*
- *En utilisant* $W_q = \frac{\rho}{\mu(1-\rho)}$ *avec* $\rho = \lambda/\mu = \lambda/2$

$$W_q = \frac{\lambda/2}{\mu(1 - \lambda/2)} < 1 \;\Rightarrow\; \lambda < 4/3$$

- *Il faut donc* $\lambda < 4/3$ *messages/seconde, ce qui correspond à utilisation de* $\rho = 2/3$.

Exemple 2.9.3. *Système manufacturier avec une seule machine et capacité finie de stocage de pièces (système M/M/1/K)*
- *Obhectif, mois de* 10% *des pièces seront blouées :*

$$\pi_n = (1 - \rho)\frac{\rho^K}{1 - \rho^{K+1}} < 0,1$$

- *Taux d'arrivée de* λ *pices/minute*
- *Trois chois de machines*
  1. $\mu = 0,5$ *pièces / minute au coûte de* 100 *euro*
  2. $\mu = 1,2$ *ièces / minute au coûte de* 300 *euro*
  3. $\mu = 2$ *pièces / minute au coûte de* 500 *euro*
- *coûte de stocage de* 80 *euro par emplacement (pièce)*

1. *Premier choix* ($\mu = 0,5,\ \rho = 2$)
   - $\pi_n(K) = \frac{2^K}{2^{K+1}-1}$
   - *Avec* $K = 1, \pi_n(1) = 2/3 > 0,1$
   - $\lim_{K\to\infty} \pi_n(K) = 0,5 \;\Rightarrow$ *solution non faisable comme* $\pi_n(K)$ *toujour plus grand que* $0,1$.
2. *Deuxième choix* ($\mu = 1,2,\ \rho = 5/6$)
   - $\pi_n(K) = \frac{5^K}{6^{K+1}-5^{K+1}}$
   - $\pi_n(1) = 5/9 > 0,1$

- $\pi_n(5) \simeq 0,1007 > 0,1, \ \pi_n(6) \simeq 0,0774 < 0,1$
- *Avec* $K = 6$, *coûte total de* $300 + 6 \times 80 = 780$ *euro*

3. *Troisième choix* $(\mu = 2, \ \rho = 0,5)$
   - $\pi_n(K) = \frac{1}{2^{K+1}-1}$
   - $\pi_n(1) = 1/3 > 0,1$
   - $\pi_n(2) = 1/7 > 0,1, \ \pi_n(3) \simeq 0,0667 < 0,1$
   - *Avec* $K = 3$, *coûte total de* $500 + 3 \times 80 = 740$ *euro*

Troisième option avec $K = 3$ *espaces de stocage est solution optimale*

Chapitre 3

Variations et extensions

Dans ce chapiter nous considérons quelques variantes des modèles simples étudiés jusqu'à présent , Nous nous limitons à des arrivées de Poisson et des temps de service exponentielles.

3.1 Systèmes avec rejet

Supposons qu'un client qui arrive quand il y a n clients dans le système entre avec une probabilité p_n et part avec la probabilité $q_n = 1 - p_n$ Si une longue file d'attente décourage les clients, alors p_n est une fonction décroissante de n .

Comme un cas particulier s'il y a une salle d'attente fini de capacité K nous pouvons supposer que :

$$p_n = \begin{cases} 1 \quad \text{pour :} \quad n < K; \\ \\ 0 \quad \text{pour.} \quad n \geq K. \end{cases}$$

Indiquant que, une fois la salle d'attente est remplie de capacité K, aucun plus de clients ne peuvent entrer dans le système .
Soit à présent, $X(t)$ le nombre de clients dans le système à l'instant t .
Nous admettons que si le processus d'arrivées est Poisson de taux λ et les clients qui arrivent quand il y a n clients dans le système entrent avec la probabilité p_n, alors les paramètres de naissances sont :

$$\lambda_n = \lambda p_n \ \ pour \ \ n = 1, 2, ...$$

Dans le cas d'un seul serveur,

$$\mu_n = \mu \ \ pour \ \ n = 1, 2, ...$$

Nous pouvons évaluer la loi stationnaire π_k de la longueur de la file d'attente par les moyens habituels.
Dans des systèmes avec rejet, pas tous les clients arrivants entrent dans le système et certains

sont perdus .
Alors le taux d'entrée est le taux au quel les clients entrent effectivement au systéme à l'état stationnaire et est donné par

$$\lambda_I = \lambda \sum_{n=0}^{\infty} \pi_n p_n$$

le taux au quel les clients sont perdus est :$\lambda \sum_{n=0}^{\infty} \pi_n q_n$
Examinons en détail le cas d'un système $M/M/c$ dans lequel ;le client arrivant entre dans le système si et seulement si un serveur est libre, alors

$$\lambda_k = \begin{cases} \lambda \quad \text{pour :} \quad k = 0, 1, ..., c-1; \\ \\ 0 \quad \text{pour.} \qquad\quad k = c. \end{cases}$$

et

$$\mu_k = k\mu \quad pour \quad k = 0, 1, ..., c$$

Pour déterminer la distribution stationnaire on a :

$$\rho_k = \frac{\lambda_0 \lambda_1 ... \lambda_{k-1}}{\mu_0 \mu_2 ... \mu_k}$$

$$= \frac{1}{k!} (\frac{\lambda}{\mu})^k \;\; pour \;\; k = 0, 1, ..., c$$

et comme :

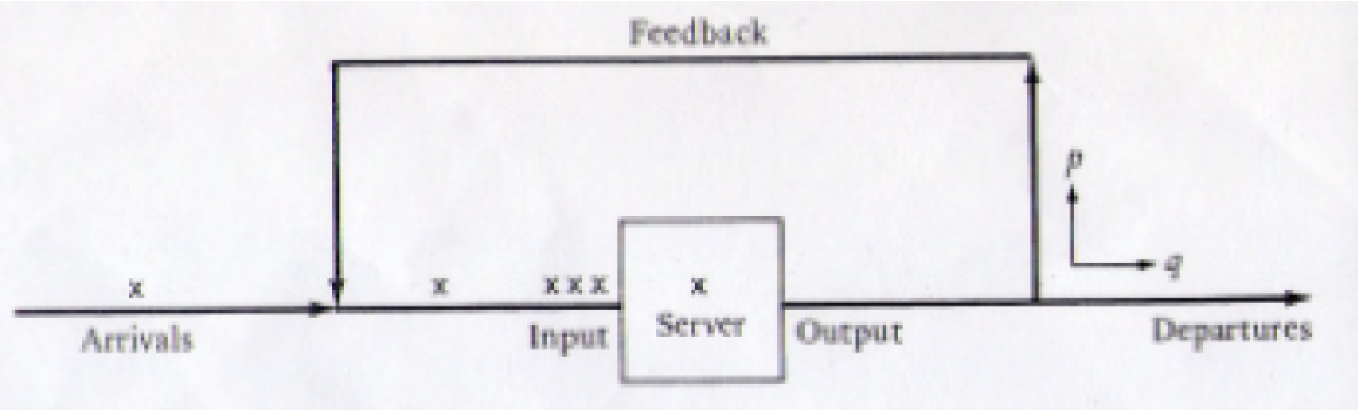
$$\pi_k = \frac{\rho_k}{\sum_{k=0}^{\infty} \rho_k}$$

Alors,

$$\pi_k = \frac{\frac{1}{k!}(\frac{\lambda}{\mu})^k}{\sum \frac{1}{j!}(\frac{\lambda}{\mu})^j} \;\; pour \;\; k = 0, 1, ..., c \tag{3.1}$$

3.2 Systèmes avec retour

Considérons un système avec un seul serveur avec des arrivées de Poisson et des temps de service exponentielles mais supposons qu'un certain client en départ du serveur retourne à la fin de la file d'attente pour un service supplémentaire. Supposons en particulier qu'un client quittant le serveur s'écarte du système avec la probabilité q et retourne à la file d'attente pour un service supplémentaire avec la probabilité $p = 1 - q$. Comme le montre le schéma suivant :



Ainsi que toutes telles décisions sont statistiquement indépendantes et que les demandes d'un client retournant au service sont statistiquement les mêmes que ceux d'un client arrivant de l'extérieur le système.
Soient,

- λ : le taux d'arrivée au système.
- μ : le taux de service.
- $X(t)$: le nombre de clients dans le système à l'instant t.

Il s'en suit que $X(t)$ est bien un processus de naissance et de mort. Avec ;

$$\lambda_n = \lambda \;\; pour \;\; n = 0, 1, ...$$

$$\mu_n = q\mu \;\; pour \;\; n = 0, 1, ...$$

$$\rho_k = \frac{\lambda_0 \lambda_1 ... \lambda_{k-1}}{\mu_0 \mu_2 ... \mu_k}$$

$$= (\frac{\lambda}{q\mu})^k$$

$$\pi_0 = \frac{1}{\sum_{k=0}^{\infty} \rho_k}$$

$$= \frac{1}{\sum_{k=0}^{\infty} (\frac{\lambda}{q\mu})^k}$$

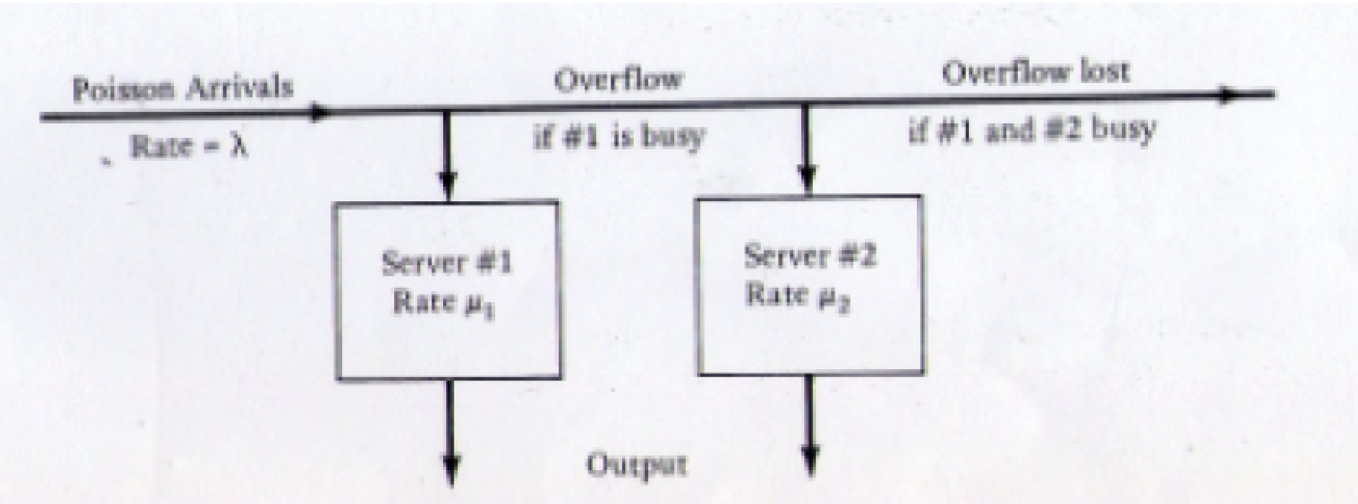
$$= (1 - \frac{\lambda}{q\mu})$$

Donc, quand $\lambda < k\mu$

$$\pi_k = (1 - \frac{\lambda}{q\mu})(\frac{\lambda}{q\mu})^k$$

3.3 Une file d'attente avec deux serveurs

Considérons un système à deux serveurs où le serveur i a le taux μ_i pour $i = 1, 2$. Les arrivées dans le système suivent un processus de Poisson de taux λ. Un client qui arrive au systéme quand il est vide va directement au premier serveur.
Si un client arrive au système quand le premier serveur est occupé s'oriente au deuxième serveur. comme le montre le schéma suivant



Et si les deux serveurs sont occupés, le client est perdu.
Le système est décrit par la paire $(X(t), Y(t))$ où,

$$X_t = \begin{cases} 1 & \text{si le serveur 1 est occupé} \quad ; \\ 0 & \text{si le serveur 1 est inoccupé} \quad . \end{cases}$$

et

$$Y_t = \begin{cases} 1 & \text{si le serveur 2 est occupé} \quad ; \\ 0 & \text{si le serveur 2 est inoccupé} \quad . \end{cases}$$

Les quatre états du système sont : $\{f(0,0), (1,0), (0,1), (1,1)\}$ et les transitions entre ces états sont données dans le tableau suivant :

De l'état	à l'état	taux de transition	Description
$(0,0)$	$(1,0)$	λ	arrivée lorsque le système est vide
$(1,0)$	$(0,0)$	μ_1	la fin du service 1 quand le service 2 est libre
$(1,0)$	$(1,1)$	λ	arrivée lorsque le service 1 est occupé
$(1,1)$	$(1,0)$	μ_2	la fin du service 2 quand 1 est occupé
$(1,1)$	$(0,1)$	μ_1	la fin du service 1 lorsque 2 est occupé
$(0,1)$	$(1,1)$	λ	arrivée quand 2 est occupé et 1 libre
$(0,1)$	$(0,0)$	μ_2	la fin du service 2 quand 1 est libre

Le processus $(X(t), Y(t))$à états finis, est une chaîne de Markov continue et les taux de transitions donnent la matrice infinitésimal de la chaîne de Markov

$$A = \begin{Vmatrix} -\lambda & 0 & \lambda & 0 \\ \mu_2 & -(\lambda+\mu_2) & 0 & \lambda \\ \mu_1 & 0 & -(\lambda+\mu_1) & \lambda \\ 0 & \mu_1 & \mu_2 & -(\mu_1+\mu_2) \end{Vmatrix} \tag{3.2}$$

On trouve la loi stationnaire

$$\pi = (\pi_{(0,0)}, \pi_{(0,1)}, \pi_{(1,0)}, \pi_{(1,1)})$$

par résolvant l'équation :

$$\pi A = 0$$

qui donne naissance au système d'équation :

$$\begin{array}{ccccc} -\lambda\pi_{(0,0)} & +\mu_2\pi_{(0,1)} & +\mu_1\pi_{(1,0)} & & = 0 \\ & -\lambda(\lambda+\mu_2)\pi_{(0,1)} & & +\mu_1\pi_{(1,1)} & = 0 \\ \lambda\pi_{(0,0)} & & -\lambda(\lambda+\mu_1)\pi_{(1,0)} & +\mu_2\pi_{(1,1)} & = 0 \\ & \lambda\pi_{(0,1)} & \lambda\pi_{(1,0)} & -(\mu_1+\mu_2)\pi_{(1,1)} & = 0 \end{array}$$

Avec

$$\pi_{(0,0)} + \pi_{(0,1)} + \pi_{(1,0)} + \pi_{(1,1)} = 1$$

Par un calcul d'algèbre élémentaire nous obtenons la solution :

$$\begin{aligned} \pi_{(0,0)} &= \frac{\mu_1\mu_2(2\lambda+\mu_1+\mu_2)}{D} \\ \pi_{(0,1)} &= \frac{\lambda^2\mu_1}{D} \\ \pi_{(1,0)} &= \frac{\lambda\mu_2(\lambda+\mu_1+\mu_2)}{D} \\ \pi_{(1,1)} &= \frac{\lambda^2(\lambda+\mu_2)}{D} \end{aligned} \tag{3.3}$$

où

$$D = \mu_1\mu_2(2\lambda + \mu_1 + \mu_2) + \lambda^2\mu_1 + \lambda\mu_2(\lambda + \mu_1 + \mu_2) + \lambda^2(\lambda + \mu_2)$$

3.4 Files d'attente prioritaires de préemption

Envisageons un processus de file d'attente de serveur unique qui a deux catégories de clients ;prioritaires et non prioritaires, formant un processus d'arrivée de Poisson,indépendantes avec des taux α etβ respectivement.
Les temps de services sont indépendantes et exponentiellement distribués avec les paramètres γ etδ respectivement.
Dans cette étude la discipline du système est la suivante :

- Premier arrivé premier servi pour les clients non prioritaires.

- Le service des clients prioritaires n'est jamais interrompu.

- Si un client prioritaire arrive lors d'un service d'un client non prioritaire, alors le service de ce dernier est immédiatement arrêté en faveur au client prioritaire.

- Le service interrompu du client est repris quand il n'y a pas de clients prioritaires actuels.

Pour notre étude nous introduisons les outils suivants :

1. Le taux d'arrivée du système est :

$$\lambda = \alpha + \beta$$

2. La fraction

$$p = \frac{\alpha}{\lambda}$$

Dont sont les clients prioritaires dans le système.

3. La fraction

$$q = \frac{\beta}{\lambda}$$

Dont sont les clients non prioritaire dans le système.

4. Dans le système le temps moyen de service est donnée par les moyens pondérées de manière appropriée :

$$\frac{1}{\gamma}$$

Pour les clients prioritaire et

$$\frac{1}{\delta}$$

Pour les clients non prioritaires. et pour la somme est :

$$\begin{aligned}\tfrac{1}{\mu} &= p(\tfrac{1}{\gamma}) + q(\tfrac{1}{\delta}) &(3.4)\\ &= \tfrac{1}{\lambda}(\tfrac{\alpha}{\gamma} + \tfrac{\beta}{\delta}) &(3.5)\end{aligned}$$

Oùμ est le taux du service du système Finalement nous introduisons

5. l'intensité du trafic pour le système :

$$\rho = \frac{\lambda}{\mu}$$

6. pour les clients prioritaires :

$$\sigma = \frac{\alpha}{\gamma}$$

7. pour les clients non prioritaires :

$$\tau = \frac{\beta}{\delta}$$

Par la formule [3.5]On remarque que :

$$\rho = \sigma + \tau$$

L'état du système est décris par la paire $((X(t), Y(t))$où $X(t)$ est le nombre de clients prioritaires dans le système et $Y(t)$ le nombre de clients non prioritaires dans le système.
Observant à présent que les clients prioritaires regardent le système comme une simple file d'attente $M/M/1$;Par conséquence on a la distribution stationnaire donnée par la l'équation [2.6]

$$\begin{aligned}\pi_m &= \lim_{t \longrightarrow \infty} \mathbb{P}(X(t) = m) &(3.6)\\ &= (1-\sigma)\sigma^m &(3.7)\end{aligned}$$

Ainsi que pour la longueur de la fie :

$$L_p = \frac{\sigma}{1-\sigma}$$

Le temps moyen de service

$$W = \frac{L_p}{\alpha} = \frac{1}{\gamma - \alpha}$$

Mais ce n'est pas le cas pour les clients non prioritaire et n'est pas aussi facile d'obtenir des informations puisque ces arrivées sont fortement affectés par les clients prioritaires. Cependant $(X(t), Y(t))$est un état discret, alors la chaine de Markov en temps continu et ses techniques nous permettra de décrire la distribution limite quand elle existe.
Les transitions sont décrites dans la table suivante :

De l'état	à l'état	taux de transition	Description
(m,n)	(m+1,n)	α	arrivée d'un prioritaire
(m,n)	(m,n+1)	β	arrivée d'un non prioritaire
$(0, n) n \geq 1$	$(0, n-1)$	δ	service d'un non prioritaire est complet
(m,n)$m \geq 1$	(m-1,n)	γ	service d'un prioritaire est complet

Considérons

$$\pi_{m,n} = \lim_{t \longrightarrow \infty} \mathbb{P}(X(t) = m, Y(t) = n)$$

La loi stationnaire du processus.
Par un raisonnement analogue

$$\begin{array}{rclllll}
(\alpha+\beta)\pi_{1,0} & = & \gamma\pi_{1,0} & +\delta\pi_{0,1} & & & (1) \\
(\alpha+\beta+\gamma)\pi_{m,1} & = & \gamma\pi_{m+1,0} & & & +\alpha\pi_{m-1,0} & (2) \\
(\alpha+\beta+\gamma)\pi_{0,n} & = & \gamma\pi_{1,n} & \delta\pi_{0,n-1} & +\beta\pi_{0,n-1} & & (3) \\
(\alpha+\beta+\gamma)\pi_{m,n} & = & \gamma\pi_{m+1,n} & & +\beta\pi_{m,n-1} & +\alpha\pi_{m-1,n} & (4)
\end{array}$$

Nous allons nous contenter de déterminer le nombre moyen L_n des clients non prioritaires dans le système àl'état d'équilibre donné par :

$$L_n \;=\; \textstyle\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} n\pi_{m,n} \tag{3.8}$$

Nous introduisons la notation :

$$M_n \;=\; \textstyle\sum_{n=0}^{\infty} n\pi_{m,n} \;=\; \sum_{n=0}^{\infty} n\pi_{m,n} \tag{3.9}$$

et donc :

$$L_n \;=\; M \;\; +M \;\; +... \tag{3.10}$$

Par [3.7]nous obtenons :

$$p_n \;=\; \mathbb{P}\{X(t) = m\} \;=\; \textstyle\sum_{n=0}^{\infty} \pi_{m,n} \;=\; (1-\sigma)\sigma^m \tag{3.11}$$

et

$$\pi_n \quad = \quad \mathbb{P}\{Y(t)=n\} \quad = \quad \sum_{n=0}^{\infty} \pi_{m,n} \tag{3.12}$$

Nous commençons par additionner les deux cotés de (1) et (2) pour $m = 0, 1, ...$ Pour obtenir :

$$(\alpha+\beta)\pi_{0,0} \quad +\gamma\sum_{n=0}^{\infty}\pi_{m,0} \quad = \quad \gamma\sum_{n=0}^{\infty}\pi_{m,0} \quad +\delta\pi_{0,1} \quad +\alpha\pi_0$$

Après simplification ;

$$\beta\pi_0 \quad = \quad \delta\pi_{0,1} \tag{3.13}$$

Maintenant nous sommons (3)et(4) sur$m = 0, 1, ...$
Pour obtenir :

$(\alpha+\beta+\delta)\pi_{0,n} + (\alpha+\beta+\gamma)\sum_{m=0}^{\infty}\pi_{m,n} = \gamma\pi_{1,n} + \delta\pi_{0,n+1} + \beta\pi_{0,n-1} +$
$\gamma\sum_{m=0}^{\infty}\pi_{m+1,n} + \beta\sum_{m=0}^{\infty}\pi_{m,n-1} + \alpha\sum_{m=0}^{\infty}\pi_{m-1,n}$

Par suite :

$$\begin{aligned}(\alpha+\beta)\pi_n + \delta\pi_{0,n} + \gamma\sum_{m=1}^{\infty}\pi_{m,n} \quad &= \quad \gamma\sum_{m=1}^{\infty}\pi_{m,n} + \delta\pi_{0,n+1} + \beta\pi_{n-1} \\ +\alpha\pi_n\end{aligned}$$

Par suite :

$$\beta\pi_n + \delta\pi_{0,n} \quad = \quad \beta\pi_{n-1} + \delta\pi_{0,n+1}$$

Par induction avec [3.13], nous obtenons :

$$\beta\pi_n \quad = \quad \delta\pi_{0,n+1} \quad pour \quad n = 0, 1, ... \tag{3.14}$$

Sommant [3.14] et utilisant le fait que :$\sum_n \pi_n = 1$
Nous avons

$$\beta = \quad \delta\sum_{n=0}^{\infty}\pi_{0,n+1} \quad = \quad \delta\mathbb{P}\{X(t)=0, Y(t)>0\}$$

Or

$$\mathbb{P}\{X(t)=0, Y(t)>0\} \quad = \quad \sum_{n=0}^{\infty}\pi_{0,n} \quad = \quad \frac{\beta}{\delta} = \tau \tag{3.15}$$

Comme [3.5] affime que

$\mathbb{P}\{X(t)=0\} \quad = \quad 1-\frac{\alpha}{\gamma} = 1-\sigma$
On a :
$\pi_{0,0} = \mathbb{P}\{X(t)=0, Y(t)>0\} \quad = \quad \mathbb{P}\{X(t)=0\}\mathbb{P}\{X(t)=0, Y(t)>0\}$

$$= \quad 1-\frac{\alpha}{\gamma}-\frac{\beta}{\delta} = 1-\sigma-\tau$$

où$\sigma+\tau<1$

Avec les résultats préliminaires, nous nous tournons vers la détermination de

$$M_m = \sum_{n=1}^{\infty} n\pi_{m,n}$$

et ceci en multipliant (3) par n et en sommant, nous tirons :

$$\begin{aligned}(\alpha+\beta+\delta)M_0 &= \gamma M_1 + \delta\sum_{n=1}^{\infty} n\pi_{0,n+1} + \beta\sum_{n=1}^{\infty} n\pi_{0,n-1} \\ &= \gamma M_1 + \delta M_0 - \delta\sum_{n=1}^{\infty} n\pi_{0,n+1} + \beta M_0 + \beta\sum_{n=1}^{\infty} n\pi_{0,n-1} \\ &= \gamma M_1 + \delta M_0 - \delta(\tfrac{\beta}{\delta}) + \beta M_0 + \beta(1-\sigma)\end{aligned}$$

où la dernière ligne est un résultat de [3.9]et[3.15].
Après simplification et réarrangement, le résultat et :

$$M_1 = \sigma M_0 + \tfrac{\beta}{\gamma}\sigma \tag{3.16}$$

Par suite nous multiplions (4) par n et nous sommons ;
Alors

$$\begin{aligned}(\alpha+\beta+\delta)M_m &= \gamma M_1 + \beta\sum_{n=1}^{\infty} n\pi_{m,n-1} + \alpha M_{m-1} \\ &= \gamma M_{m+1} + \beta M_m + \beta\sum_{n=1}^{\infty} n\pi_{m,n-1} + \alpha M_{m-1}\end{aligned}$$

Nous nous référons à [3.11]et nous simplifions la formule ;

$$(\alpha+\delta)M_m = \gamma M_{m+1} + \alpha M_{m-1} + \beta(1-\sigma)\sigma^m \quad pour \quad m = 1, 2, ... \tag{3.17}$$

Après quelques calculs, nous arrivons à la résolution des équation (3.20)et(3.21) qui donnent :

$$M_m = M_0\sigma^m + \tfrac{\beta}{\gamma}\sigma^m \quad pour \quad m = 1, 2, ...$$

nous sommons sur m pour obtenir le résultat que nous désirons :

$$L_n = \sum_{m=0}^{\infty} M_m = \tfrac{1}{1-\sigma}[M_0 + \tfrac{\beta}{\gamma}\tfrac{\sigma}{(1-\sigma)}] \tag{3.18}$$

Ce résultat détermine L_n en fonction de M_0
Pour obtenir la seconde relation, nous multiplions (3.18) par n et nous sommons :

$$\begin{aligned}\beta\sum_{m=0}^{\infty} n\pi_{0,n+1} &= \beta L_n = \delta\sum_{m=0}^{\infty} n\pi_{0,n+1} \\ &= \delta M_0 - \sum_{m=0}^{\infty} n\pi_{0,n+1} \\ &= \delta M_0 - \delta(\tfrac{\beta}{\delta})\end{aligned}$$

Du coup :

$$M_0 = \tfrac{\beta}{\delta}(L_n+1) = \tau(L_n+1) \tag{3.19}$$

Nous remplaçons [3.18]dans [3.19] :

$$
\begin{aligned}
L_n &= \tfrac{1}{1-\sigma}[M_0 + \tfrac{\beta}{\gamma}\tfrac{\sigma}{(1-\sigma)}] \\
&= \tfrac{1}{1-\sigma}[\tau(L_n+1) + \tfrac{\beta}{\gamma}\tfrac{\sigma}{(1-\sigma)}] \\
L_n - \tfrac{\tau}{1-\sigma}L_n &= \tfrac{1}{1-\sigma}[\tau + \tfrac{\beta}{\gamma}\tfrac{\sigma}{(1-\sigma)}] \\
(1-\tfrac{\tau}{1-\sigma})L_n &= \tfrac{1}{1-\sigma}[\tau + \tfrac{\beta}{\gamma}\tfrac{\sigma}{(1-\sigma)}]
\end{aligned}
$$

$$L_n = \tfrac{\tau}{1-\sigma-\tau}[1 + (\tfrac{\delta}{\gamma}\tfrac{\sigma}{(1-\sigma)})] \tag{3.20}$$

Donc la condition pour queL_n soit fini(et la loi stationnaire existe) est

$$\rho = \sigma + \tau < 1$$

Autrement dit, l'intensité du trafic du système doit être inférieure à 1.

Chapitre 4

Application :Traffic routier

Nous avons travaillé sur un document réalisé par Daniel FLIPO en (2013) qui traite une situation routière basique se rapprochant des systèmes $M/M/1$. Le modèle est le suivant :

1. La route est une voie à deux directions en ligne droite
2. Un entrepreneur souhaite efectuer des réparations sur une des voies, il va devoir faire un chantier qui bloquera la voie concernée sur une certaine longueur
3. il installe des feux tricolores de part et d'autre du chantier, permettant à chaque file de passer alternativement le chantier sur la file restante Les paramètres du modèle sont :

▷ L : La longueur du chantier (en m).

▷ λ_1, λ_2 : Les paramètres des processus de Poisson modèlisant les arrivées de véhicules dans un sens et dans l'autre.

▷ v : La vitesse des véhicules $(en\ m/s)$.

▷ d_1, d_2 :La durée durant laquelle chaque feu est au vert/ au rouge $(en\ s)$.

▷ K_1, K_2 :Le nombre de voitures que l'on veut laisser passer dans chaque sens.

▷ a : Le temps qu'une voiture met pour démarrer $(en\ s)$.

4.0.1 Cycle des feux, engorgement, contraintes

Cycle des feux

Un cycle des feux se déroule de la façon suivante :

▷ à l'instant 0 le feu 1 passe au vert.

▷ à l'instant d_1 il passe au rouge.

▷ à l'instant $d_1 + \frac{L}{v}$ (le temps qu'il faut pour évacuer la dernière voiture passée au feu vert de la zone du chantier) le feu 2 passe au vert.

▷ à l'instant $d_1 + \frac{L}{v} + d_2$ le feu 2 passe au rouge.

▷ à l'instant $d_1 + \frac{L}{v} + d_2 + \frac{L}{v}$le cycle est fini, on revient à l'instant zéro.

Engorgement, contraintes

Si on veut faire passer K véhicules dans un sens, on doit laisser le feu au vert pendant un laps de temps au moins égal à $a.K$, ce qui donne ici :

$$\begin{cases} d_1 \geq aK_1, & ; \\ \\ d_2 \geq aK_1, & . \end{cases}$$

Le risque principal d'un mauvais ajustement des paramètres des feux est la situation où la longueur de la file d'attente d'un des feux augmente plus qu'elle ne réduit ; sa longueur augmente sans borne et le système s'engorge. Dans ce modèle M. Flipo suggère que l'engorgement peut s'éviter en respectant ces inégalités :

$$\begin{cases} d_1 \geq a\lambda_1(d_1 + d_2 + 2\frac{L}{v}), & ; \\ \\ d_2 \geq a\lambda_2(d_1 + d_2 + 2\frac{L}{v}), & . \end{cases}$$

Il considère ici que la durée pendant laquelle un feu est au vert doit être au moins égale à la durée laissant passer le nombre moyen de véhicules arrivés pendant un cycle. Il n'y a donc théoriquement que très peu de chance que le système s'engorge, il faudrait que le flux d'arrivée se comporte pendant longtemps largement au dessus sa moyenne.
On peut discuter le fait que ces différentes minorations de d_1 et d_2 ne prennent pas en compte la distance entre le feu et le véhicule qui démarre, l'auteur considérant que chaque voiture démarre directement du feu.
Ceci est contrebalancé par le fait que le coeffcient appliqué à a est la durée totale d'un cycle (le feu reste au vert assez longtemps pour laisser passer le nombre théorique moyen d'arrivants pendant la durée qu'il passe au rougeET au vert). Néanmoins dans certaines simulations on a rencontré des (rares) cas où ces inégalités s'avèrent insuffsantes.
n pourra aussi (par exemple) inffuencer le fonctionnement du système dans l'autre sens en augmentant artiffciellement le temps a de réaction d'un automobiliste avant de démarrer afin de prendre en compte la distance moyenne qui le sépare du feu dans la file d'attente.

Conclusion

Dans ce mémoire nous illustrons l'utilité de la théorie des files d'attente . Ce travail nous a également permis de montrer l'importance de cette étude qui s'agit de prédire le comportement des systémes d'attente.

Comme nous venons de voir ; les phénomènes d'attente sont retrouvés dans certains systèmes tels les réseaux téléphoniques, les systèmes informatiques, dans les banques ,dans la route, etc...

Si jamais vous avez vu une caricature dans votre journal local, vous savez ce que vous avez regardé n'est pas la photo d'une personne particulière, mais malgré cela vous reconnaissez de qui il s'agit ; parce que l'artiste pour renseigner le lecteur a représenter par quelques coups de crayons bien choisis les trais caractéristiques de son visage.

C'est exactement ce que fait un bon modèle des files d'attente .

Bibliographie

[1] Abou El-Ata, M. O. and Hariri, A. M. A. The M/M/c/N queue with balking and reneging. Computers and Operations Research. 19 (1992), No. 13 713-716.

[2] Abdel-Karim Aboul-Hassan, Sherif I. Rabia and Ahmed Kadry, Analytical study of a discrete time retrial queue with balking customers and early arrival scheme, Alexandria Engineering Journal, 44 (2005), No. 6, 911-917.

[3] Allen, A. O. 1990. Probability, Statistics, and Queueing Theory with Computer Science Applications. Second edition, Academic Press, New York (First edition :1978).

[4] Anisimov, V. V., Zakusilo, O. K., and Donchenko, V. S. 1987. Elements ofQueueing Theory and Asymptotic System Analysis. Vishcha Shkola, Kiev (in Russian).

[5] A. Gomez-Corral and M.F. Ramalhoto. On the waiting time distribution and the busy period of a retrial queue with constant retrial rate. Stochastic Modelling and Applications, 3, 37-47, (2000).

[6] Baccelli, F., and Bremaud, P. 1990. Mathematical Theory of Queues. Springer- Verlag, Berlin.

[7] Claudie Hasseforder CHABRIAC , Eléments de Théorie des files d'attente,page05, Janvier 2008.

[8] Chaudhry, M. L., and Templeton, J. G. C. 1983. A First Course in Bulk. Queues.John Wiley and Sons, New York.

[9] Différents modèles ont différentes caractéristiques, avec performances évaluées pour le régime permanent

[10] Fondation théorique pour une grande variété de systémes

[11] Frédéric Sur. Programmation dynamique, chaînes de Markov, files d'attente. école des Mines de Nancy $(2013) - (2014)$.

[12] Gelenbe, E.,and Pujolle, G. 1987. Introduction to Queuing Networks. John Wiley and Sons, Chichester, England (French Original : Introduction aux réseaux de files d'attente, Edition Eyrolles ; Paris).

[13] Gross,D.and C.M,Harris,Fundamentals of Queueing Theory ,Willcy,New york,1975 1984.

[14] G. Rubino , Processus Stochastiques , Février 2006.

[15] HIDAKI TAKAGI , QUEUEING ANALYSIS, Tokyo Research Laboratory, North HOLLAND,Volume $1 - (1991)$.

[16] J.W. Cohen,The single server queue2nd end, North-Holland, Amsterdam, 1982.

[17] KARLIN QUEUEING SYSTEMS ,John Wiley and Sons $(1990), 346 - 386$

[18] Kashyap, B. R. K., and Chaudhry, M. L. 1988. An Introduction to Queueing Theory. Aarkay, Calcutta, India.

[19] K.B, GK, Techniques de modélisation : Méthodes analytiques.

[20] Khintchine, A. Y 1969. Mathematical Methods in the Theory of Queueing. Secondedition, Hafner Publishing Company, New York (First edition : Griffin, London, 1960 ; Russian original :1955). $859-866$.

[21] Kumar, R. and Sharma, S.K. An M/M/1/N Queuing Model with Retention of reneged cus- tomers and Balking, American Journal of Operational Research.$2(2012), No.1, 1-5$.

[22] LeGall, P. (1962). Les systemes avec ou sans attente et les processus stochastiques, Tome I. Dunod, Paris.

[23] LEONARD KLEINROCK , QUEUEING SYSTEMS ,Volume J :Theory,John Wiley and Sons $-(1975), 89-106$.

[24] Moulay Hachemi , Files d'attente et applications ,page 12 , 2014/2015.

[25] MTH2302D ,S. Le Digabel, éEcole Polytechnique de Montréal,A2017,(v1).

[26] Newell, G. F. 1982. Applications of Queueing Theory. Second edition, Chapman and Hall, London (First edition : 1971).

[27] Présentation basée sur chapitre 8 de Cassandras et Lafortune (2008)

[28] PHILIPPE ROBERT, Réseaux et files d'attente ,Mathématique et Applications,Springer$(1983), 25-39$.

[29] PERRUT,THIRARD COLIN ,AUGER VINCENT,Mèmoire en Phénomènes d'attente et application à la gestion des feux de circulation$-(2013)$.

[30] VLADIMIR ANISIMOV, Switching Processes in Queueing Models, Mathematical models 2,John Wiley$-(1988)$, chapitre 12.

[31] Yves Caumel. Probabilités et processus stochastiques. Springer. Verlag France, (2011).

[32] Zakhar Kabluchko, Stochastic Processes (Stochastik II), University of Ulm Institute of Stochastics, (2013-2014).