

**Analyse des performances du
système $M/G/1$ avec rappels
et Bernoulli feedback**

M. BOUALEM & M. CHERFAOUI

**Laboratoire LAMOS, Université de
Béjaïa**

MSR'2013-Rennes, France

Introduction

- Le premier paramètre décrit le phénomène de rappels bien connu dans les applications liées :
 - Systèmes téléphoniques.
 - Systèmes informatiques.
 - Système informatique à temps réel.
- Le second paramètre décrit la notion de feedback introduite en général pour exploiter des situations d'attente où tous les arrivés demandent le principal service et seulement quelques uns parmi eux ont besoin de demander un autre service.

Le phénomène de feedback dans les systèmes d'attente avec rappels peut apparaître dans plusieurs situations pratiques :

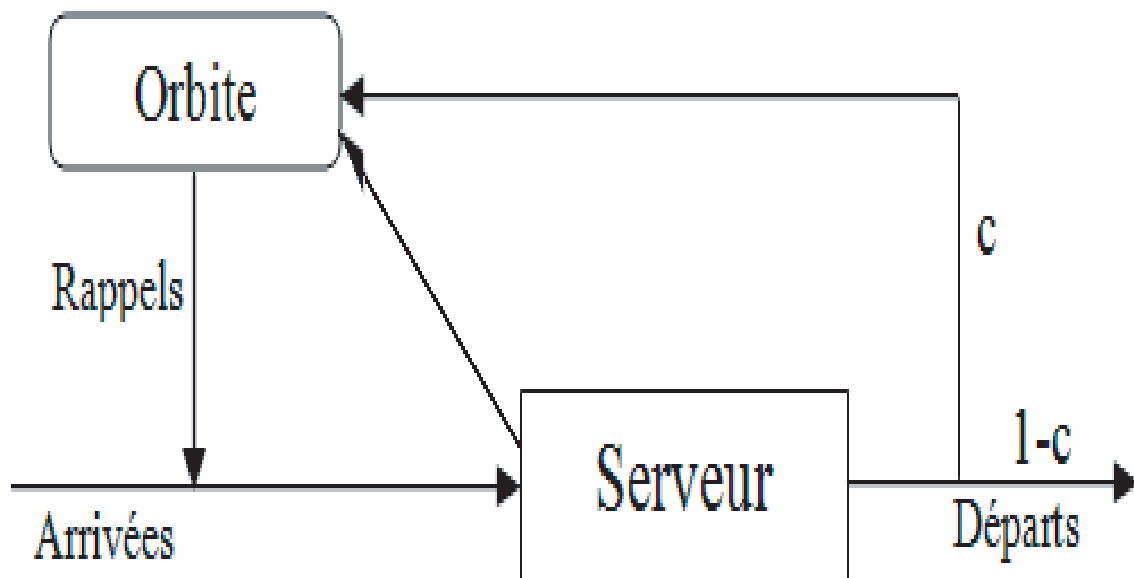
♣ Les systèmes MATS (Multiple Access Telecommunication Systems) où des messages s'avérant comme erreurs à la destination (messages perdus ou corrompus) sont renvoyés.

♣ Le protocole ARQ (Automatic Repeat Request) dans un réseau de communication à haute fréquence.

La notion de Feedback a été initialement introduite par **[Takacs (1963)]** pour l'étude de certains systèmes d'attente classiques, et depuis plusieurs papiers sont apparus sur ce sujet en considérant d'autres types de systèmes avec différentes variantes (rappels, vacances, pannes,...).

Pour plus de détails, voir **[Choi & Kim (2003); Choi & Kim & Lee (1998,2000,2003); Krishna Kumar & Raja (2006); Krishna Kumar, Arivudainambi & Krishnamoorthy (2006, 2010); Krishna Kumar, Pavai Madheswari & Vijayakumar (2002); Madan & Al-Rawwash (2005); Atencia & Moreno (2004); Atencia, Fortes & Sanchez (2009); Jau-Chuan Ke & Chang (2009); Zaiming, Jinbiao Wu & Gang Yang (2009); Boualem, Djellab & Aissani (2012)]**.

À chaque fois qu'un client arrive et trouve le serveur occupé, quitte le service pour rejoindre un groupe de clients bloqués appelé orbite, sinon il reçoit immédiatement son service. Une fois servi, il décide, avec une probabilité c , de retourner en orbite pour demander un autre service ou de quitter le système avec une probabilité $\bar{c} = 1 - c$.



Système M/G/1 avec rappels et feedback

La discipline d'accès au serveur à partir de l'orbite est gouvernée par une loi exponentielle avec une intensité mixte donnée par $\alpha(1 - \delta_{0j}) + j\theta$, quand le nombre de clients en orbite est $j \in \mathbb{N}$, où δ_{0j} est la fonction de Kronecker.

★ Dans le cas $\alpha > 0$ et $\theta = 0$, on obtient la discipline de rappels constant (voir [**Artalejo (1997); Artalejo & Gomèz (2008); Aissani (2000, 2008, 2009); Boualem & Djellab & Aissani (2009)**]).

★ Alternativement, quand $\alpha = 0$ et $\theta > 0$, l'intensité de rappel devient la discipline de rappel classique (voir [**Yang Templeton (1987); Falin (1990); Gautam Choudhury (2006); Boualem, Djellab & Aissani (2012)**]).

OBJET

Faire une analyse mathématique du système d'attente $M/G/1$ avec rappels classiques et Bernoulli feedback. Pour cela, en utilisant la propriété de décomposition stochastique obtenue dans **[Boualem et al. (2012)]**.

Description mathématique

Les clients primaires arrivent suivant un flux poissonnien de taux λ .

Un client qui arrive et trouve le serveur occupé, quitte l'aire du service pour rejoindre un groupe de clients bloqués appelé *orbite*. Après un certain temps aléatoire, il renouvelle sa tentative d'entrer en service, une fois, deux fois, ..., jusqu'à ce qu'il le trouve disponible.

Une fois servi, le client doit décider, soit de rejoindre l'orbite pour un autre service avec une probabilité c ($0 < c < 1$) où de quitter le système définitivement avec une probabilité complémentaire \bar{c} .

Les intervalles de temps inter-rappels suivent une distribution exponentielle de taux $k\theta$. Comme cette politique de rappel dépend du nombre de clients dans l'orbite, on l'appelle politique de rappel classique.

Les temps de service sont supposés d'une loi arbitraire, de fonction de distribution $B(x)$, de transformée de laplace $\beta(s)$ et de deux premiers momments β_1 et β_2 .

Toutes les variables aléatoires introduites sont mutuellement indépendantes.

L'état du système à l'instant t peut être décrit par le processus :

$$X(t) = (C(t), N_o(t), \xi(t), t \geq 0), \quad (1)$$

où,

$$C(t) = \begin{cases} 0, & \text{si le serveur est oisif;} \\ 1, & \text{si le serveur est occupé;} \end{cases} \quad (2)$$

$N_o(t)$: le nombre de client en orbite à l'instant t .

Si $C(t) = 1$, $\xi(t)$ représente le temps de service écoulé du client en service.

Chaîne de Markov induite

Soit $\{t_n, n \in \mathbb{N}\}$ une suite d'instants de fin de service.

La suite de variables aléatoires $Y_n = \{q_n = N_o(t_n), n \in \mathbb{N}\}$ forme une chaîne de Markov induite, dont l'équation fondamentale est

$$q_{n+1} = q_n - \delta_{q_n} + v_{n+1} + \eta, \quad (3)$$

où,

- v_{n+1} : le nombre de clients qui arrivent pendant un temps de service qui se termine à l'instant t_{n+1} .

Sa distribution est donnée par :

$$K_i = P(v_{n+1} = i) = \int_0^\infty \frac{(\lambda x)^i}{i!} e^{-\lambda x} dB(x). \quad (4)$$

- δ_{q_n} est la variable de Bernoulli définie par :

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

La distribution de δ_{q_n} est donnée par :

$$P(\delta_{q_n} = 1 / q_n = k) = \frac{k\theta}{\lambda + k\theta}, \quad (5)$$

$$P(\delta_{q_n} = 0 / q_n = k) = \frac{\lambda}{\lambda + k\theta}. \quad (6)$$

- La variable aléatoire η est définie par

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

En outre,

$$P[\eta = 1] = c \quad \text{et} \quad P[\eta = 0] = \bar{c} \quad (= 1 - c).$$

Condition d'ergodicité

Théorème 1. *La chaîne de Markov induite est ergodique si et seulement si*

$$\rho = \lambda\beta_1 + c < 1. \quad (7)$$

Probabilités de transition

$$\begin{aligned} p_{mn} &= P[q_{i+1} = n | q_i = m] \\ &= P[q_i - \delta_{q_i} + v_{i+1} + \eta = n | q_i = m] \\ &= K_{n-m} \frac{\lambda}{\lambda + m\theta} \bar{c} + K_{n-m+1} \frac{m\theta}{\lambda + m\theta} \bar{c} \\ &+ K_{n-m-1} \frac{\lambda}{\lambda + m\theta} c + K_{n-m} \frac{m\theta}{\lambda + m\theta} c. \quad (8) \end{aligned}$$

Notons que $p_{mn} \neq 0$ uniquement pour $m = 0, 1, 2, 3, \dots, n + 1$.

Propriété de décomposition stochastique [Boualem et al. (2012)]

$$\pi(z) = \left[\frac{(1 - \rho)\tilde{B}(\lambda - \lambda z)(1 - z)}{(\bar{c} + cz)\tilde{B}(\lambda - \lambda z) - z} \right] \quad (9)$$

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

Remarque :

Le premier est la fonction génératrice du nombre de clients dans le système $M/G/1$ avec feedback (voir [**Tackacs (1963)**]).

Le second facteur est la fonction génératrice du nombre de clients dans le système d'attente $M/G/1$ avec rappels et feedback étant donné que le serveur est libre (voir [**Djellab (2005)**]).

Nombre moyen de clients dans le système :

$$\bar{n} = \pi'(1) = \rho + c(1 - \lambda\beta_1) + \frac{\bar{c}\lambda\rho}{\mu(1 - \rho)} + \frac{c\lambda\beta_1}{(1 - \rho)} + \frac{\lambda^2\beta_2}{2(1 - \rho)}. \quad (10)$$

Nombre moyen de clients en orbite :

$$\bar{n}_0 = \bar{n} - \rho = c(1 - \lambda\beta_1) + \frac{\bar{c}\lambda\rho}{\mu(1 - \rho)} + \frac{c\lambda\beta_1}{(1 - \rho)} + \frac{\lambda^2\beta_2}{2(1 - \rho)}. \quad (11)$$

Temps moyen d'attente :

$$\bar{w} = \frac{c(1 - \lambda\beta_1)}{\lambda} + \frac{\bar{c}\rho}{\mu(1 - \rho)} + \frac{c\beta_1}{(1 - \rho)} + \frac{\lambda\beta_2}{2(1 - \rho)}. \quad (12)$$

Nombre moyen de rappels par client :

$$\bar{d} = \mu\bar{w} = \frac{c\mu(1 - \lambda\beta_1)}{\lambda} + \frac{\bar{c}\rho}{(1 - \rho)} + \frac{c\mu\beta_1}{(1 - \rho)} + \frac{\lambda\mu\beta_2}{2(1 - \rho)}. \quad (13)$$

Nombre moyen de clients qui demandent un service supplémentaire :

$$\bar{F} = c\bar{n} = c\rho + c^2(1 - \lambda\beta_1) + \frac{c\bar{c}\lambda\rho}{\mu(1 - \rho)} + \frac{c^2\lambda\beta_1}{(1 - \rho)} + \frac{\lambda^2c\beta_2}{2(1 - \rho)}. \quad (14)$$

Aspect pratique Afin de confirmer les résultats analytiques, obtenus dans la performance, nous avons élaboré un simulateur, sous environnement Matlab, basé sur la simulation à événements discrets, qui reproduira le comportement du modèle d'attente $M/G/1$ avec rappels et feedback et de faire les comparaisons entre les caractéristiques estimées à celles obtenues analytiquement. Pour cela, on a choisi deux lois de probabilités (loi exponentielle et de Cox d'ordre 2) pour les temps de service.

Pour notre étude, on se focalise sur les trois situations suivantes :

- 1° **cas** : on considère la file $M/M/1$ avec rappels et feedback avec tel que : $\lambda = 2$, $\beta_1 = 1/4$, $\mu = 2$ et $c = 0,3$.
- 2° **cas** : on considère la durée de temps de service suit une distribution de Cox d'ordre 2, tel que $\nu_1 = 4$, $\nu_2 = 7,5$ et $\gamma = 0,05$. Et on fixe $\lambda = 2$, $\mu = 2$ et $c = 0,3$.
- 3° **cas** : on considère le cas particulier $c = 0$ (pas de demande d'un service supplémentaire) ; avec : $\lambda = 2$, $\beta_1 = 1/4$ et $\mu = 5$.

Pour un temps de simulation $t_{max} = 1\ 000$ unités de temps et $n = 100$ (nombre de répliquations). Les résultats comparatifs obtenus, pour les trois cas présentés, sont résumés dans le tableau suivant :

	Caractéristiques	\bar{n}	\bar{n}_0	\bar{w}	\bar{d}	\bar{F}
1	Théoriques	3,725	2,925	2,1 625	4,325	1,117
	Simulation	3,731	2,931	1,9 716	3,9 432	1,119
	$t_{(n,\alpha)}$	0,363	0,408	0,628	0,973	0,082
2	Théoriques	4,077	3,266	0,726	1,452	1,223
	Simulation	4,060	3,246	0,835	1,669	1,218
	$t_{(n,\alpha)}$	0,804	0,927	0,931	1,872	0,236
3	Théoriques	1,250	0,750	0,325	1,625	—
	Simulation	1,261	0,761	0,3 663	1,206	—
	$t_{(n,\alpha)}$	1,930	1,890	1,035	1,734	—

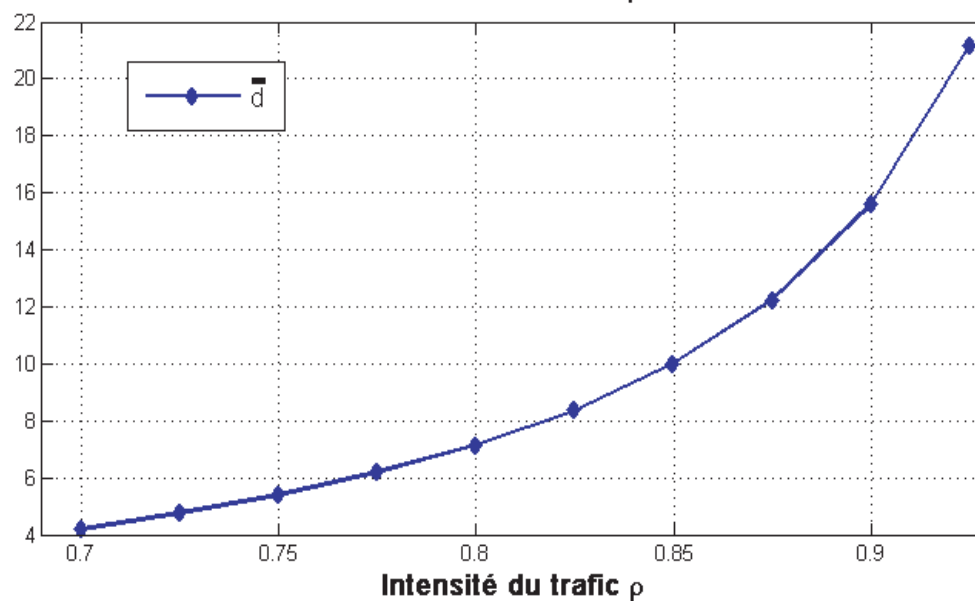
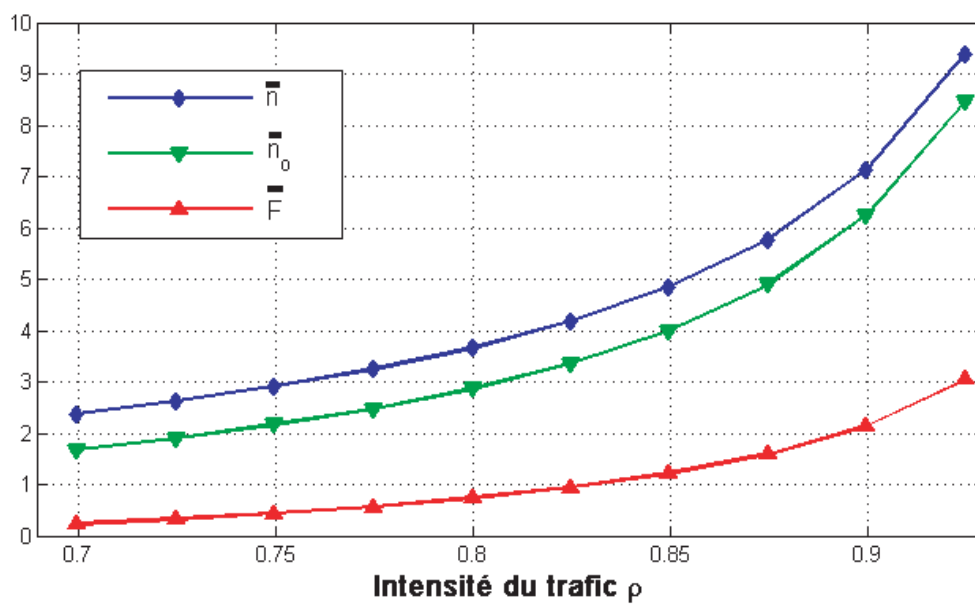
Comparaison des résultats de simulation avec les résultats théoriques

On constate que toutes les statistiques de Student calculées, pour un seuil $\alpha = 0,05$, sont inférieures à $t(0,025 ; 99) = 1,96$. Du ce fait, on ne rejette pas H_0 au seuil $\alpha = 0,05$, ce qui signifie que les résultats théoriques ne sont pas contradictoires avec ceux de simulation au seuil $\alpha = 0,05$.

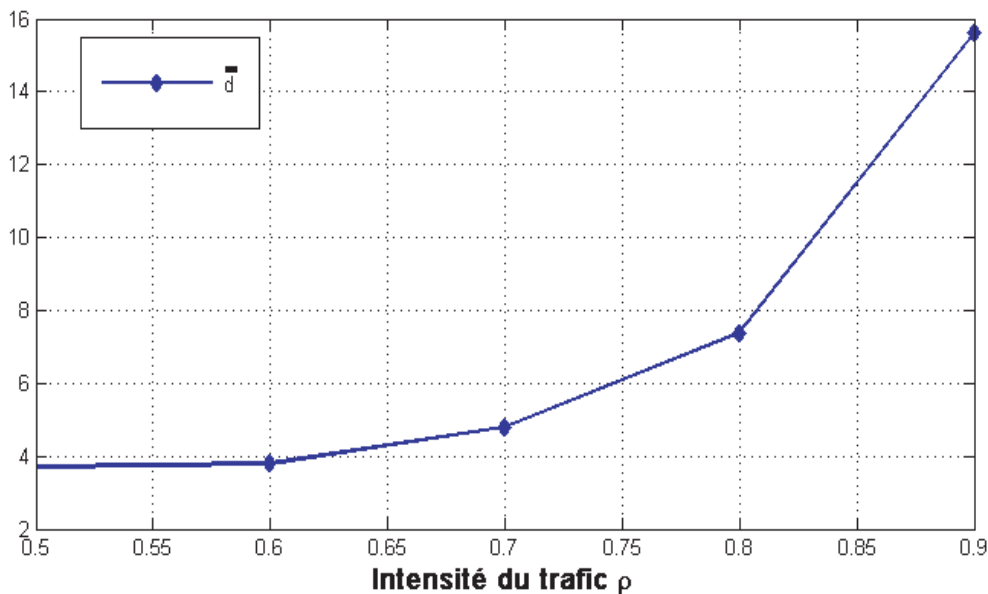
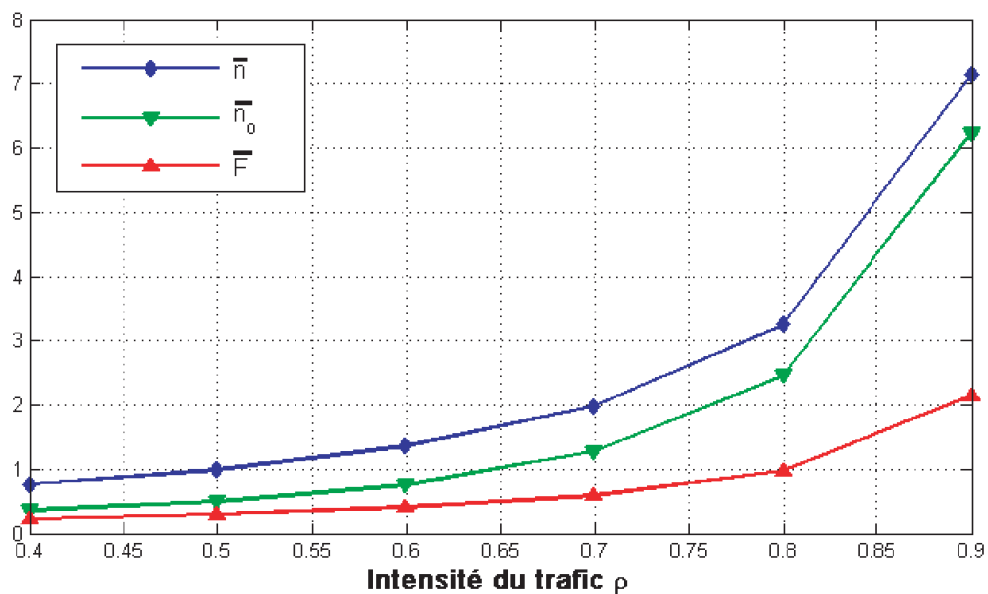
Influence de l'intensité du trafic sur les mesures de performance Nous analysons l'influence de l'intensité du trafic sur les caractéristiques stationnaires du système et cela dans le but d'avoir une tendance générale des résultats.

$$\rho = c + \lambda\beta_1, \quad (15)$$

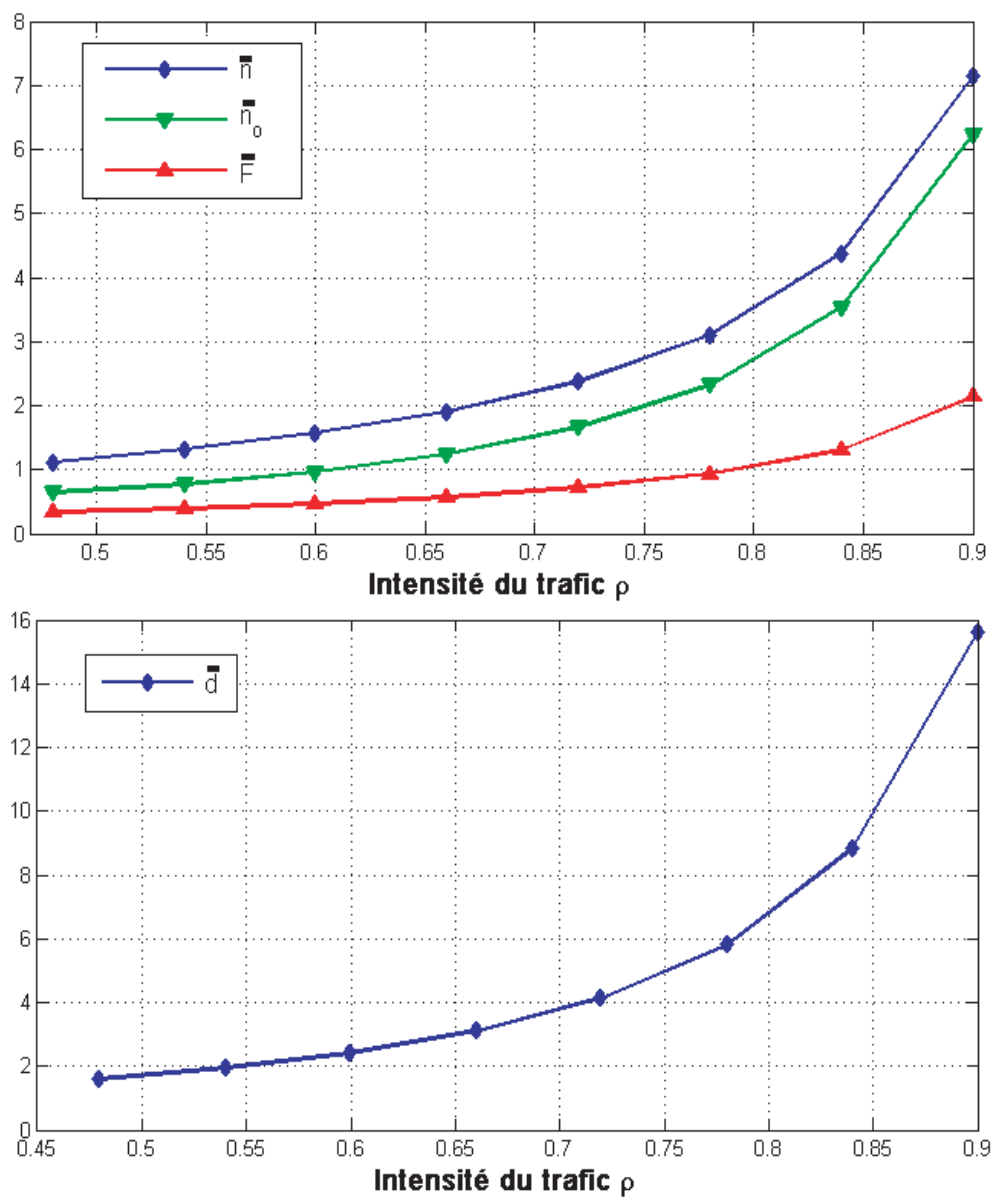
Première situation : Variation des caractéristiques du système en fonction de ρ (cas de variation de c) pour $1/\beta_1 = 1$, $\lambda = 0,6$ et $\mu = 1,5$.



Deuxième situation : Variation des caractéristiques du système en fonction de ρ (cas de variation de λ) pour $1/\beta_1 = 1$, $c = 0,3$ et $\mu = 1,5$.



Troisième situation : Variation des caractéristiques du système en fonction de ρ (cas de variation de β_1 , $\lambda = 0,6$, $c = 0,3$ et $\mu = 1,5$).



Conclusion

À partir des résultats obtenus (Théoriques et Numériques), reflétant les caractéristiques du système étudié, on constate une dépendance exponentielle entre les caractéristiques du système et l'intensité de trafic. Plus précisément, cette dépendance est plus sensible, pour les deux cas de variations : taux des arrivées et le taux de service, par rapport à la variation de la probabilité de feedback.

Nous envisageons d'analyser d'autres aspects de systèmes d'attente avec rappels et feedback. En effet, notre étude peut être généralisée dans d'autres directions plus intéressantes. Par exemple, nous pouvons étendre notre analyse pour discuter le contrôle optimal de notre modèle. De plus, il serait utile de considérer une structure plus complexe en prenant une distribution générale pour les temps de rappels et pannes du serveur.