$M/M/\infty$ QUEUE WITH ON-OFF SERVICE SPEEDS

B. D'Auria [1]

UDC 519.2

Abstract

In this paper we show the analysis of a special kind of $M/M/\infty$ queueing system. We suppose that the rate of the servers alternates between two speeds according to an ON-OFF process whose ON periods are exponentially distributed and with general OFF periods. We look at the particular case when the OFF service rate is zero while the distribution of the OFF periods is heavy tailed and for this case we derive the tail of the number of the customers in the system. An application of this model could be related to transportation systems where the $M/M/\infty$ represents a delay line, s.a. a highway, whose crossing time may be unexpectedly interrupted, for a long time, by some accident.

1. Introduction and model description

The study of systems with interruptions has often received attention by queueing theorists. This kind of studies can indeed give useful information about the behavior of such systems when subject to special conditions. In this paper we are going to study an $M/M/\infty$ system whose server rate alternates between two speeds, μ and μ', according to an ON-OFF process. The ON-OFF process is a semi-markov process, its ON periods are exponentially distributed with parameter f while the OFF periods are generally distributed according to a distribution $L(\cdot)$ with mean r^{-1}, Laplace-Stieltjes transform (LST) $\mathcal{L}(s)$ and $L(0^+) = 0$. The arrival process is a Poisson Process with parameter $\lambda > 0$ and the service times are exponential distributed with parameter depending on the state of the system. During the ON periods the servers work at rate $\mu > 0$ while during the OFF periods they work at speed $\mu' \geq 0$. Generally we suppose that $\mu > \mu'$.

Our motivation for this study is twofold. Firstly we want to investigate the possibility to express the number of customers in the system with interruptions by the sum of two independent terms, the first one being the random variable associated to the system with no interruptions and the second one being a positive random variable. In this sense we extend some of the results of Baykal–Gursoy and Xiao [2] that looked at an $M/M/\infty$ system with markov-modulated service rate. They were able to get explicit formulas assuming that the OFF periods were exponentially distributed as well. Previously, many other authors looked at similar decompositions, but many of them generally studied the M/G/1 system such as in [9], [10] and [12]. Other authors looked at systems with a finite number of servers with vacations such as in [6] where it is analyzed a $G/M/K$ queue. For vacations however it is supposed that the system interrupts only at the end of a completed service and not at anytime. As for the $M/M/\infty$ model, Keilson and Servi [11] studied it when the servers had markov-modulated service rates by using a matrix-geometric approach.

The second motivation is to analyze the effect of the heaviness of the tail of the distribution of the OFF periods onto the stationary distribution of the number of customers in the system. In the result of [2] and [11], all the involved distributions being exponential or a mixture of exponentials, it turned out that the number of customers in the systems was not Poisson distributed but still its distribution had an exponential tail decay. In our case we relax the exponential assumption and we look at regularly varying distributions. This is in line with the most recent discoveries that show many natural phenomena are well described by heavy-tailed distributions. In fact similar studies were done, for example, about M/G/1 and M/M/1 systems [4] and [5] where they looked at a system whose speed followed, similarly to our case, an ON-OFF process. They showed that the heavy-tailed component of the OFF periods crucially influences the decay of tail of the buffer occupancy of the system. In our case we look at the tail of the number of customers in the system and we show that its power-decay exponent is related to the one of the OFF period when the low-speed rate $\mu' = 0$.

A possible application of our system is related to transportation and telecommunication systems. Usually the system with infinite server is used to model particular phenomena where some objects, that in our case are the customers of the system, experience some delay. As for transportation systems, this can be the case of a highway or some high traffic street and, for telecommunication systems, it can represent a long high-speed connection, e.g. transoceanic or satellite. In this real world application, the OFF periods may represent exceptional events such as

Proceedings of the Seminar on Stability Problems for Stochastic Models, Maiori, Italy, 2005, Part II.

1072-3374/05/1314-1

accidents in the case of the highways, or interruptions of the connection for the telecommunication scenarios. What we are than interested in is the amount of extra customers that we find in the system due to these interruptions.

The paper is organized as follows. In Section 2 we describe the mathematical model and derive the fundamental equations governing the hidden Markov process. In Section 3 we show our first result about the stochastic decomposition of the stationary number of the customers in the system. Then, in Section 4 we show the second main result on the tail of the distribution of the stationary number of customers in the system and in Section 5 we finally give our conclusions.

2. Mathematical model

Let $N(t)$ be the number of customers at time t present in the system and let $U(t)$ denote the state of the system, then $U(t)$ is a binary random variable that has value 'H' (High rate) during an ON period and value 'L' (Low rate) during an OFF period. Due to the fact that the OFF periods are not exponentially distributed, the process $(N(t), U(t))$ is not Markovian and therefore in order to have a Markov process we need to add the additional information of the elapsed time $L_{\text{past}}(t)$ for the OFF period at time t.

In order to describe the state of the system completely we introduce the following distribution functions:

$$F_H(n,t) := \Pr\{N(t)=n, U(t)=H\}, \tag{1}$$

$$F_L(n,x,t)dx := \Pr\Big\{N(t)=n, x < L_{\text{past}} \le x+dx, U(t)=L\Big\}. \tag{2}$$

It is easy to check that the ergodicity conditions are always satisfied (cf. [1]) for any values of $\mu+\mu' > 0$, in that case $N(t) \to N$ as $t \to \infty$ where N is the required stationary distribution for the number of customers in the system.

By using a total probability argument we can write down the following relation for the distribution $F_H(n, t+\Delta t)$ related to the H state

$$\begin{aligned} F_H(n,t+\Delta t) = {} & F_H(n,t)(1-\lambda\Delta t)(1-n\mu\Delta t)(1-f\Delta t) \\ & + (n>0)F_H(n-1,t)(\lambda\Delta t)(1-(n-1)\mu\Delta t)(1-f\Delta t) \\ & + F_H(n+1,t)((n+1)\mu\Delta t)(1-\lambda\Delta t)(1-f\Delta t) \\ & + (1-\lambda\Delta t)(1-n\mu\Delta t)\Delta t\int_0^t \frac{F_L(n,x,t)}{1-L(x)}dL(x), \end{aligned}$$

where we used $(\cdot)$ as the indicator function of the set $\{\cdot\}$, and the following relation for $F_L(n, x+\Delta t, t+\Delta t)$, $x > 0$, about the L state

$$\begin{aligned} F_L(n,x+\Delta t,t+\Delta t) = {} & F_L(n,x,t)(1-\lambda\Delta t)(1-n\mu'\Delta t)\frac{1-L(x+\Delta t)}{1-L(x)} \\ & + (n>0)F_L(n-1,x,t)\,\lambda\Delta t\,(1-(n-1)\mu'\Delta t)\frac{1-L(x+\Delta t)}{1-L(x)} \\ & + F_L(n+1,x,t)(n+1)\,\mu'\Delta t\,(1-\lambda\Delta t)\frac{1-L(x+\Delta t)}{1-L(x)}, \end{aligned}$$

plus the boundary condition for the case $x = 0$

$$F_L(n,\Delta t,t+\Delta t) = f\Delta t F_H(n,t)(1-\lambda\Delta t)(1-n\mu\Delta t).$$

Using straightforward computations and supposing to be in stationary condition, when the time derivatives are all null, we get the following equilibrium equations

$$\begin{aligned} F_H(n)(\lambda+n\mu+f) &= \lambda F_H(n-1)+(n+1)\mu F_H(n+1)+\int_0^\infty \frac{F_L(n,x)}{1-L(x)}dL(x) \\ \frac{d}{dx}F_L(n,x) &= -\left(\lambda+n\mu'+\frac{L'(x)}{1-L(x)}\right)F_L(n,x)+\lambda F_L(n-1,x) \\ & \quad + (n+1)\mu' F_L(n+1,x) \\ F_L(n,0) &= fF_H(n). \end{aligned}$$

Introducing the following characteristic functions,

$$G_H(z) := \sum_{n=0}^{\infty} z^n F_H(n), \quad |z| \leq 1,$$
$$G_L(z,x) := \sum_{n=0}^{\infty} z^n F_L(n,x), \quad |z| \leq 1.$$

with the additional auxiliary function $H_L(z,x) := \frac{G_L(z,x)}{1-L(x)}$, we finally get the following system of partial differential equations

$$\mu(z-1)\,\partial_z G_H(z) = (\lambda(z-1) - f)\,G_H(z) + \int_0^\infty H_L(z,x)dL(x), \tag{3a}$$
$$\partial_x H_L(z,x) = \lambda(z-1)H_L(z,x) + \mu'(1-z)\,\partial_z H_L(z,x), \tag{3b}$$
$$H_L(z,0^+) = fG_H(z). \tag{3c}$$

This system turns out to be quite difficult to analyze in complete generality and therefore in the sequel we are going to focus on the more simple case when $\mu' = 0$ that corresponds to a pure ON-OFF switching of the servers.

3. Stochastic decomposition

In this section we are going to prove that generally, independently of the distribution $L(\cdot)$, we can decompose, in the case $\mu' = 0$, the number of customers in the system at equilibrium in two independent terms. The first one is the number of customers at equilibrium in an ordinary $M/M/\infty$ system with no interruptions while the second one is a positive random variable that depends on all the parameters of the system. This extends the results of [2] for the case $\mu' = 0$ and also for that case we give a different decomposition of the second term. Before stating the theorem let the function $\hat{\mathcal{L}}(\cdot)$ be the LST of the function $\hat{L}(\cdot)$ that we define as the residual life time of the OFF period, i.e.

$$\frac{1}{r}\hat{\mathcal{L}}(s) = \frac{1-\mathcal{L}(s)}{s}.$$

Theorem 1. *The number of customers in the system, N, at equilibrium has the form*

$$N =_d N_\phi + B \cdot Y_1 + (1-B)(Y_2 + X), \tag{4}$$

where N_ϕ, B, Y_1 and Y_2, and X are five positive and independent random variables. N_ϕ is Poisson distributed with parameter λ/μ, B is Bernoulli distributed with parameter $r/(r+f)$, Y_1 and Y_2 have characteristic function $R(z) = e^{-\frac{f}{\mu}\frac{1}{r}\int_0^{\lambda(1-z)} \hat{\mathcal{L}}(y)dy}$ while X has characteristic function $\hat{\mathcal{L}}(\lambda(1-z))$.

Proof. By considering $\mu' = 0$ we have that equation (3b) simplifies in

$$\partial_x H_L(z,x) = \lambda(z-1)H_L(z,x),$$

that gives, using (3c), the following expression for $H_L(z,x)$

$$H_L(z,x) = fG_H(z)e^{\lambda(z-1)x}. \tag{5}$$

Having the value of $H_L(z,x)$ we can then write

$$G_L(z,x) = fG_H(z)(1-L(x))e^{\lambda(z-1)x}$$

and also

$$G_L(z) := \int_0^\infty G_L(z,x)dx = \frac{f}{r}G_H(z)\hat{\mathcal{L}}(\lambda(1-z)).$$

By substituting the expression (5) in (3a), after having simplified, we get the following differential equation

$$\mu(z-1)\,\partial_z G_H(z) = (\lambda(z-1) + f(\mathcal{L}(\lambda(1-z)) - 1))\, G_H(z),$$

whose solution is

$$G_H(z) = Ce^{\frac{\lambda}{\mu}z} e^{\frac{\lambda}{\mu}\frac{f}{r}\int_0^z \hat{\mathcal{L}}(\lambda(1-t))dt},$$

with C being a constant whose value can be computed by imposing $G(1) = 1$ where $G(z) := G_H(z) + G_L(z)$. Finally setting

$$R(z) = e^{-\frac{f}{\mu}\frac{1}{r}\int_0^{\lambda(1-z)} \hat{\mathcal{L}}(y)dy}$$

we can summarize the previous results by the following expressions

$$G(z) = e^{-\frac{\lambda}{\mu}(1-z)} \left[\left(\frac{r}{r+f} \right) R(z) + \left(\frac{f}{r+f} \right) R(z)\hat{\mathcal{L}}(\lambda(1-z)) \right] \tag{6a}$$

$$G_H(z) = \frac{r}{r+f} e^{-\frac{\lambda}{\mu}(1-z)} R(z) \tag{6b}$$

$$G_L(z) = \frac{f}{r+f} e^{-\frac{\lambda}{\mu}(1-z)} \hat{\mathcal{L}}(\lambda(1-z)) R(z). \tag{6c}$$

The expression (6a) gives us the required stochastic decomposition for the random variable N in five independent random variables. Indeed we have

$$N =_d N_\phi + BY_1 + (1-B)(Y_2 + X),$$

where N_ϕ is a Poisson random variable with parameter λ/μ that refers to the $M/M/\infty$ system with no interruptions, B is Bernoulli distributed with parameter $r/(r+f)$, Y_1 and Y_2 are random variables with characteristic function $R(z)$ and X is a random variable with characteristic function $\hat{\mathcal{L}}(\lambda(1-z))$, and the thesis follows. □

Remark 1. *In the exponential case, i.e.* $L(x) = 1 - e^{-rx}$, $R(z)$ *simplifies to*

$$R(z) = \left(\frac{r}{r + \lambda(1-z)} \right)^{\frac{f}{\mu}},$$

so that $Y_1, Y_2 \sim NB(f/\mu, r/(r+\lambda))$ *where* $NB(\phi, \delta)$ *refers to a generalized negative binomial distribution with parameters* ϕ *and* δ *and finally* $X \sim NB(1, r/(r+\lambda))$*. These results are in agreement with what was proved in [2].*

4. Heavy-tailed OFF periods

In this section we are going to look at the tail of the distribution of N when the distribution $L(\cdot)$ is regular varying, more precisely when

$$\bar{L}(t) = l(t)t^{-\alpha} \text{ as } t \to \infty,$$

with $l(t)$ a slowly varying function at infinity (see [8]), and $\bar{L}(t) := 1 - L(t)$. As noticed in the previous section, in the exponential case N preserves the exponential decay of the tail similarly to the original N_ϕ. In the next theorem we show that this is not the case when $L(\cdot)$ is regular varying. In the following we use $f(t) \sim g(t)$ as $t \to \infty$ to mean that $\lim_{t\to\infty} f(t)/g(t) = 1$, and for LST we write $F(s) \sim G(s)$ as $s \to 0$ to mean that $\lim_{s\to 0} F(s)/G(s) = 1$.

Theorem 2. *Let*

$$\bar{L}(t) = l(t)t^{-\alpha} \text{ as } t \to \infty,$$

with $1 < \alpha < 2$*, then*

$$\Pr\{N(t) > n\} \sim \frac{rf}{r+f} \frac{\lambda^{\alpha-1}}{1-\alpha} l'(n) n^{1-\alpha} \text{ as } n \to \infty, \tag{7}$$

where $l'(n) = l(1/\lambda(1 - e^{-1/n}))$ *is a slowly varying function.*

Proof. In the sequel we will use $Y(s) := R(e^{-s})$ and $X(s) := \hat{\mathcal{L}}(\lambda(1-e^{-s}))$, the LST of the r.v.s Y and X. By applying Karamata's Tauberian theorem (see Theorem 8.1.6 in [3]) we have that

$$1 - \hat{\mathcal{L}}(s) \sim r s^{\alpha-1} \frac{\Gamma(2-\alpha)}{\alpha-1}\, l\left(\frac{1}{s}\right).$$

This directly gives us the first term expansion of $X(s)$

$$1 - X(s) \sim r\frac{\lambda^{\alpha-1}}{\alpha-1} s^{\alpha-1}\Gamma(2-\alpha)\, l'\left(\frac{1}{s}\right),$$

where $l'(t) = l(1/\lambda(1-e^{-1/t}))$ is a slowly varying function at infinity. Applying once more Karamata's theorem we get

$$\Pr\{X > n\} \sim r\frac{\lambda^{\alpha-1}}{\alpha-1} n^{1-\alpha}\, l'(n).$$

More involved is the computation of the decay of $\Pr\{Y > n\}$. We have that

$$\int_0^{\lambda(1-e^{-s})} \hat{\mathcal{L}}(y)dy \sim \lambda(1-e^{s}) + r\frac{\Gamma(1-\alpha)}{\alpha}\left(\lambda\left(1-e^{-s}\right)\right)^{\alpha} l'\left(\frac{1}{s}\right),$$

and handling the series expansion of the various terms in the expression of $Y(s)$ we finally get the following

$$1 - Y(s) \sim \frac{\lambda}{r}\frac{f}{\mu}s - \frac{f}{\mu}\frac{\lambda^{\alpha}}{\alpha}\Gamma(1-\alpha)s^{\alpha}\, l'\left(\frac{1}{s}\right)$$

and again by Karamata's theorem we get

$$\Pr\{Y > n\} \sim \frac{f}{\mu}\frac{\lambda^{\alpha}}{\alpha} n^{-\alpha}\, l'(n).$$

As $\Pr\{Y > n\} \sim o\left(\Pr\{X > n\}\right)$ by using a result in [7] we have that $\Pr\{Y + X > n\} \sim \Pr\{X > n\}$ and applying it also to N_{ϕ}, we get

$$\Pr\{N > n\} \sim \frac{f}{f+r}\Pr\{X > n\}$$

and the thesis follows. □

Remark 2. *In the case $\mu' \neq 0$ the decay of the tail of the distribution of N is no longer heavy-tail but exponential. Indeed the system has as lower-bound the behavior of a classical $M/M/\infty$ system with constant service rates equal to μ' whose stationary distribution for N is Poisson with parameter λ/μ'.*

5. Conclusion

In this paper we presented a preliminary analysis of the $M/M/\infty$ system with two service rates alternating according to an ON-OFF process whose OFF periods are generally distributed. When the lowest service speed is zero, we have, by Theorem 1, a stochastic decomposition of the distribution function of the number of customers in the system. This decomposition allows us also to compute, by Theorem 2, the asymptotic decay of the distribution in case the OFF periods are regular varying distributed. It turns out that the distribution of the number of customers in the system is regular varying as well. The decay exponent is related to the one of the residual lifetime tail distribution of the OFF periods. The more general system with positive low speed seems to be quite complicated to analyze in explicit form but as we pointed out in Remark 2 the distribution of the number of customers in the system will be no longer heavy-tail.

REFERENCES

1. F. Baccelli and P. Brémaud, *Elements of Queueing Theory. Palm–Martingale Calculus and Stochastic Recurrences*, Springer–Verlag, Berlin (2003).
2. M. Baykal-Gursoy and W. Xiao, "Stochastic Decomposition in $M/M/\infty$ Queues with Markov-Modulated Service Rates" in: *Queueing Systems* (2004), 48:75–88.
3. N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular Variation*, Cambridge University Press (1987).
4. O. Boxma and I. Kurkova, "The M/M/1 queue in a heavy-tailed random environment" in: *Statist.Neerlandica* (2000), 54:221–236.
5. O. Boxma and I. Kurkova, "The M/G/1 queue with two service speeds" in: *Adv.Appl.Prob.* (2001), 33:520–540.
6. X. Chao and Y. Zhao, "Analysis of multi-server queues with station and server vacations" in: *European J.Oper.Res.* (1998), 110:392–406.
7. P. Embrechts, C. Goldie and N. Veraverbeke, "Subexponentiality and infinite divisibility" in: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* (1979), 49:335–347.
8. W. Feller, *An Introduction to Probability Theory and its Applications*, Wiley, New York, 2nd edition (1971).
9. S. Fuhrmann, "A note on the M/G/1 queue with server vacations" in: *Oper. Res.* (1984), 32:1368–1373.

10. S. Fuhrmann and R. Cooper, "Stochastic decomposition in the M/G/1 queue with generalized vacations" in: *Oper. Res.* (1985), 33:1117–1129.

10. J. Keilson and I. Servi, "The matrix $M/M/\infty$ system: Retrials models and Markov modulated sources" in: *Adv. in Appl. Probab.* (1993), 25:453–471.

11. Y. Levy and U. Yechiali, "Utilization of idle time in an M/G/1 queuing system" in: *Managm. Sci.* (1975), 22:202–211.

[1] *EURANDOM, P.O. Box 513 - 5600 MB Eindhoven, The Netherlands. E-mail: bdauria@eurandom.tue.nl*