A $M/G/\infty$ Queue Busy Period Distribution Characterizing Parameter

Manuel Alberto M. Ferreira

1 Instituto Universitário de Lisboa, BRU - IUL, Lisboa, Portugal

Correspondence: Manuel Alberto M. Ferreira, Instituto Universitário de Lisboa, BRU - IUL, Lisboa, Portugal. E-mail: manuel.ferreira@iscte.pt

Abstract

In this work it is proposed a parameter, \( \eta \), to characterize the \( M/G/\infty \) Queue Busy Period distribution that is a modification of the *peakedness* proposed in Whitt (1984). As this distribution moments centered at the origin are given, in general, by very complex expressions it is relevant to search for parameters simpler but with good descriptive properties.

Keywords: \( M/G/\infty \), busy period, *peakedness*, \( \eta \)

1. Introduction

In the \( M/G/\infty \) queue system the customers arrive in accordance with a Poisson process at rate \( \lambda \), receive a service which length is a positive random variable with distribution function \( G(\cdot) \) and mean \( \alpha \) and, upon its arrival, find an available server. Each customer service is independent from the other customers’ services and from the arrivals process. The traffic intensity is given by \( \rho=\lambda \alpha \).

When operating, in the \( M/G/\infty \) system there is a sequence of busy and idle periods, as in any queue operation. As in the \( M/G/\infty \) busy period distribution the moments centered at the origin are given, with a few exceptions, by very complex expressions it is relevant to search for parameters analytically simpler but with good descriptive properties. The case of this queue with exponentially distributed service times is an example of moments complex expressions strange, because not even the evident markovian properties that it fulfills bring the expected simplifications, and important, because of its use in the applications.

In the sequence of this work it is developed a parameter, called \( \eta \), to characterize the \( M/G/\infty \) queue busy period distribution that is a modification of the *peakedness* proposed in the work of Whitt (1984). Its expression, for several service time distributions, in particular the exponential, is quite simple.

This work is organized as follows. In section 2 it is made a state of play on the \( M/G/\infty \) busy period distribution moments. Then, in section 3, it is introduced the *peakedness* for the \( M/G/\infty \) busy period distribution and are studied its properties. The parameter \( \eta \), the goal of this work, is introduced in section 4, as a result of the peakedness recognized properties in the former section. The parameter \( \eta \) is thought as an improvement of the *peakedness*. The paper ends with the conclusions and a short list of references.

2. An Overview on the \( M/G/\infty \) Busy Period Moments

The length of the \( M/G/\infty \) queue busy period is a random variable, designated \( B \), with Laplace transform given by, see Stadje (1985):

\[
\bar{B}(s) = 1 + \frac{s\alpha}{\lambda} \int_0^\infty e^{-st} - \lambda_s e^{-\lambda_s [1-\alpha(G(v))]} dv \, dt \tag{1}
\]

Its inversion is a complicate task except in the cases \( G_1(t) = \frac{e^{-\rho}}{e^{-\rho}+(1-e^{-\rho})e^{-\alpha t}} \), \( t \geq 0 \), for which \( B \) is exponentially distributed with a probability concentration, \( e^{-\rho} \), at the origin and \( G_2(t) = 1 - \frac{1}{1-e^{-\rho}+e^{-\rho+\lambda+\lambda e^{-\rho t}}} = 0 \), \( t \geq 0 \), for which \( B \) is exponentially distributed (Note 1).

For the \( M/D/\infty \) system it is possible to invert \( \bar{B}(s) \) (Ferreira, 1996a) but the resulting expression is extremely complex, although allows understanding the probabilistic structure of \( B \) (Note 2).

As for the moments of \( B \), after (1) it is obtained, see (Ferreira & Ramalhoto, 1994):

\[
E[B^n] = (-1)^{n+1} \left( \frac{e^\rho}{\lambda} n C^{(n-1)}(0) - e^\rho \sum_{p=1}^{n-1} (-1)^{n-p} \binom{n}{p} E[B^{n-p}] C^{(p)}(0) \right)
\]
where,
\[ C^{(n)}(0) = \int_0^\infty (-t)^n e^{-\lambda t^{\alpha} G_v(t)} dt, \quad n = 1, 2, ... \] (2)

So it is possible, with this recurrent formula, to write exact expressions for the moments centered at the origin, obtaining
\[ E[B] = \frac{e^\rho - 1}{\lambda} \]

whichever is the service time distribution and
\[ E[B^2] = \frac{2 e^\rho \int_0^\infty (e^{\lambda t^{\alpha} G_v(t)} - 1) dt}{\lambda^2} \]

that depends on the whole service time distribution probabilistic structure (Note 3).

The Expression (2) has evidently no interest in the cases of the \( M/G_1/\infty \) and \( M/G_2/\infty \) systems. But for the \( M/D/\infty \) system
\[ C^{(0)}(0) = 1 - e^{-\rho} \quad \text{and} \quad C^{(n)}(0) = e^{-\rho} (-\alpha)^n \int_0^\infty \lambda^{n-1} C^{(n-1)}(0) dt, \quad n = 1, 2, ... \]

and then it is possible to write, by a process of recurrence, simple expressions for the \( E[B^n] \), \( n = 1, 2, ... \)

So, the situation is easy in the cases of the \( M/G_1/\infty \) and \( M/G_2/\infty \) systems. For the \( M/D/\infty \) system is good only for the moments’ computation. In other systems, in particular for the \( M/M/\infty \) system, the situation is complicated (Note 4).

In fact it is strange that for the \( M/M/\infty \) system, very important in the applications, see for instance (Ferreira, Andrade, Filipe & Coelho, 2011), despite its evident markovian properties, since both the inter-arrival times and the service time are exponentially distributed, it is not possible to have simple exact expressions for the busy period moments.

3. The Peakedness

In Whitt (1984), where it is conducted a search for queues parameters extremal distributions, it is proposed that in the probability distributions related with queues study, the peakedness, that is: the value of the Laplace transform in \( 1/\alpha \), is used to their characterization instead of the variance. That author stands on the works of Holtzman (1973), Eckberg (1977) and Rolski (1972, 1974 and 1976) that make similar proposes for parameters and extremal distributions (Note 5).

For the situation under consideration in this work, designating the peakedness by \( p \), after (1):
\[ p = B \left( \frac{1}{\alpha} \right) = 1 + \frac{1}{\rho} - \frac{1}{\lambda} \int_0^\infty e^{-\lambda t^{\alpha} G_v(t)} dt \] (3)

and, also,
\[ p = \sum_{n=0}^\infty (-1)^n \frac{E[B^n]}{n! \lambda^n} \] (4)

So, \( p \) incorporates information about the whole moments of \( B \). Then, for the \( M/G/\infty \) queue busy period it is poor to consider it a mere option to the variance. It may be considered as a composite parameter that globally characterizes the distribution (Note 6).

For instance:

- \( M/G_1/\infty \)
\[ p^{G_1} = \frac{\rho + 1}{e^p + p} \] (5)

- \( M/G_2/\infty \)
\[ p^{G_2} = \frac{\rho}{e^p + 1} \] (6)

- \( M/D/\infty \)
\[ p^D = \frac{\rho + 1}{e^p + 1 + p} \] (7)

- \( M/M/\infty \)
\[ p^M = \frac{e^{\rho - p - 1}}{\rho (e^p - 1)} \] (8)

Note that, in these four examples, \( p \) depends only on \( \rho \), being the expressions obtained quite simple. With the exception of \( E[B] \) it is the only parameter with a simple expression for the \( M/M/\infty \) system \( B \) distribution, confer with (Ferreira, Andrade, & Filipe, 2012).
For service times distributions related with the exponential, it also holds:

- If the service time distribution is $\text{NBUE}$ - new better than used in expectation – with mean $\alpha$,
  \[
  \int_0^\infty [1 - G(v)] dv \geq \int_0^\infty e^{-\frac{v}{\alpha}} dv, \text{ see (Ross, 1983, p. 273) and, so,}
  p^{\text{NBUE}} \leq \frac{e^{\rho - 1}}{\rho (e^{\rho - 1})} \tag{9}
  \]

- If the service time distribution is $\text{NWUE}$ - new worse than used in expectation – with mean $\alpha$,
  \[
  \int_0^\infty [1 - G(v)] dv \leq \int_0^\infty e^{-\frac{v}{\alpha}} dv, \text{ see (Ross, 1983, p. 273) and, so,}
  p^{\text{NWUE}} \geq \frac{e^{\rho - 1}}{\rho (e^{\rho - 1})} \tag{10}
  \]

And also, in general,

- If $G(.)$ and $H(.)$ are service time distributions such that $1 - G(t) \leq 1 - H(t)$
  \[p^G \geq p^H \tag{11}\]

4. The Parameter $\eta$

Put the Expression (4) in the form

\[p = 1 - \frac{e^\rho - 1}{\rho} + \sum_{n=2}^\infty (-1)^n \frac{E[B^n]}{n! \alpha^n}\]

As $1 - \frac{e^\rho - 1}{\rho}$ is insensible to the service time distribution (Note 7) it is defined:

\[\eta = p - \frac{\rho}{e^\rho - 1} + 1 \tag{12}\]

It was retired from $p$ what does not distinguish the various service time distributions and then the result was normalized dividing for that common part.

After what was seen in the former section, it is possible to obtain simple expressions for $\eta^{G_1}$, $\eta^{G_2}$, $\eta^D$ and $\eta^M$ depending only on $\rho$. And evidently the properties of $p$ extend to $\eta$.

It is easy to show that

\[1 \leq \eta \leq \frac{e^{\rho - 1}}{e^\rho - 1} \tag{13}\]

for any service time distribution, concluding so that

\[\lim_{\rho \to \infty} \eta = 1 \tag{14}\]

Then the parameter $\eta$ is computed for some service time distributions and values of $\rho$, see Table 1 below, being $P$ the power function with parameter $c=.5$ (Note 8).

Table 1. Values of $\eta$ for some service times distribution functions

|   | $G$   | $\rho=5$ | $\rho=10$ | $\rho=15$ | $\rho=20$ |
|---|-------|----------|-----------|-----------|-----------|
| $G_1$ | 3.3469730 | 1.7488465 | 1.0013731 | 1.0000000 | 1.0000000 |
| $G_2$ | 2.4633636 | 1.5121659 | 1.0011518 | 1.0000000 | 1.0000000 |
| $D$   | 2.0123054 | 1.3319113 | 1.0005158 | 1.0000000 | 1.0000000 |
| $M$   | 2.5414941 | 1.5819767 | 1.0067837 | 1.0000454 | 1.0000000 |
| $P$   | 2.4721612 | 1.5747122 | 1.0120525 | 1.0001526 | 1.0000000 |
Note that for great values of $\rho$, after $\rho=10$ for some service times distributions and $\rho=20$ for others, under certain conditions, $B$ is approximately exponentially distributed, see (Ferreira & Ramalhoto, 1994) and Ferreira (1996b). In the same conditions $\eta$ assumes the value 1.

5. Conclusions

The parameter $\eta$, defined in Expression (12), incorporates information on the whole moments of $B$ centered at the origin.

It assumes different values when the distributions of $B$ differ as functions of the service distribution.

And there are simple bounds for $\eta$ depending only on the traffic intensity $\rho$.

Even for the $M/M/\infty$ system, with analytically intractable parameters in the case of the distribution of $B$, $\eta$ is given for a quite simple expression.

It is even admissible that close values of $\eta$, for different service time’s distributions, indicate similar behaviors of the respective $B$ distributions.

So $\eta$ is a parameter that can help to characterize the distribution of $B$, discriminating by different service time distributions.

This parameter, $\eta$, may be included in the class of other composite parameters, as for instance the kurtosis or the skewness, that are used to characterize probability distributions putting an emphasis in their probability density functions graphic characteristics. But it is more complete and complex than the two examples here called, since incorporates the whole distribution moments. Because of that, its interpretation in probability density functions graphic characteristics terms is not so obvious and it is an interesting open field of investigation.

Acknowledgements

The author thanks very much the work of the Editor and the Reviewers on this paper that resulted in very valuable suggestions for its improvement.

This work was financially supported by FCT through the Strategic Project PEst-OE/EGE/UI0315/2011.

References

Eckberg, Jr, A. E. (1977). Sharp bounds on Laplace-Stieltjes transforms, with applications to various queuing problems. *Mathematics of Operations Research*, 2(2), 135-142. http://dx.doi.org/10.1287/moor.2.2.135

Eckberg, A. E. (1983). Generalized peakedness of teletraffic processes. *Tenth International Teletraffic Congress, Montreal, 4(4b)*, 3.

Ferreira, M. A. M. (1996a). Distribuição do período de ocupação da fila de espera M/D/\infty. *Investigação Operacional*, 16, 43-45.

Ferreira, M. A. M. (1996b). Distribuição do comprimento do período de ocupação da fila de espera M/G/\infty em situações de “heavy-traffic” (distribuições de serviço potência e de Pareto). *Communication presented at the IV Congresso Anual da S. P. E. Funchal*.

Ferreira, M. A. M., & Andrade, M. (2009). The ties between the M/G/\infty queue system transient behavior and the busy period. *International Journal of Academic Research, I*(1), 84-92.

Ferreira, M. A. M., & Ramalhoto, M. F. (1994). Estudo dos parâmetros básicos do período de ocupação da fila de espera M/G/\infty. *A Estatística e o Futuro e o Futuro da Estatística. Actas do I Congresso Anual da S. P. E. Edições Salamandra. Vimeiral*.

Ferreira, M. A. M., Andrade, M., & Filipe, J. A. (2012). The age or excess of the M/G/\infty queue busy cycle mean value. *Computer and Information Science, 5*(5), 93-97. http://dx.doi.org/10.5539/cis.v5n5p93

Ferreira, M. A. M., Andrade, M., Filipe, J. A., & Coelho, M. P. (2011). Statistical queuing theory with some applications. *International Journal of Latest Trends in Finance and Economic Sciences, I*(4), 190-195.

Holtzman, J. M. (1973). The accuracy of the equivalent random method with renewal inputs. *Bell System Technical Journal*, 52(9), 1673-1679.

Rolski, T. (1972). Some inequalities for GI/M/n queues. *Zast. Mat., 13*(1), 43-47.

Rolski, T. (1974). Some inequalities in queuing theory. *Colloquia Math. Soc. Janos Bolyai, 9*, 653-659.

Rolski, T. (1976). Order relations in the set of probability distributions functions and their applications. *Dissertationnes Mathematicae*, Polish Scientific Publishers, Warsaw.
Ross, S. (1983). *Stochastic Processes*. Wiley. New York.

Sathe, Y. S. (1985). Improved bounds for the variance of the busy period of the M/G/∞ queue. *Advances in Applied Probability*, 17, 913-914. http://dx.doi.org/10.2307/1427096

Stadje, W. (1985). The busy period of the queueing system M/G/∞. *Journal of Applied Probability*, 22, 697-704. http://dx.doi.org/10.2307/3213872

Shanthikumar, J. K., & Buzakott, J. A. (1980). On the approximations to the single server queue. *International Journal of Production Research*, 19, 761-773. http://dx.doi.org/10.1080/00207548008919705

Whitt, W. (1982). Refining diffusion approximation for queues. *Operations Research Letters*, 1(5), 165-169. http://dx.doi.org/10.1016/0167-6377(82)90033-5

Whitt, W. (1984). On approximations for queues, I: extremal distributions. *A T &T Bell Laboratories Technical Journal*, 63(1), 115-138.

**Notes**

Note 1. Both expressions result from, see (Ferreira & Andrade, 2009),

\[
G(t) = 1 + \frac{1}{\lambda} \int_0^\infty e^{-\lambda \int_0^t \beta(u) \, du} \, dt, \quad t \geq 0, \quad \lambda \leq \frac{\int_0^\infty \beta(u) \, du}{t} \leq \frac{\lambda}{e^\rho - 1},
\]

making \( \beta(t) = \beta \) (constant), being \( \beta = 0 \) for \( G_1(t) \) and \( \beta \frac{\lambda}{e^\rho - 1} \) for \( G_2(t) \).

Note 2. That is: the M/G/∞ queue with constant (Deterministic) service times.

\[
B^\beta(s) = 1 + \frac{1}{\lambda} \left( s - \frac{(s + \lambda)s}{\lambda e^{-(s + \lambda)\rho} + s} \right)
\]

which leads to the probability density function, see Ferreira (1996a)

\[
b(t) = \sum_{n=0}^\infty \left( \frac{d^c(t)}{dt^{-\rho}} \right)^n \rho^n (1 - e^{\rho t})^n
\]

where * is the convolution operator, \( \frac{d^c(t)}{dt^{-\rho}} \) and \( \frac{d^{1-d(t)}}{dt^{-\rho}} \) are defined by \( \int_0^t e^{-\rho \tau} \, d\tau \) and \( \int_1^t e^{-\rho \tau} \, d\tau \) respectively.

Note 3. In Sathe (1985), very simple bounds for the variance of the busy period are presented depending only on \( \rho, \lambda \) and \( \gamma_2 \) (service variation coefficient):

\[
\lambda^2 \max[\rho^2 + \rho^2 \gamma_2^2 - 2 \rho \rho^2 - 1; 0] \leq \text{VAR}[B] \leq \lambda^2 [2\rho^2 (\gamma_2^2 + 1)(\rho^2 - 1 - \rho) - (\rho^2 - 1)^2]
\]

Note 4. That is: exponential service times. The \( M \) to designate the exponential distribution is due to that it fulfills the Memorable property:

\[
P(X > t + y | X > t) = P(X > y)
\]

Note 5. For Holtzman (1973) and Eckberg (1983) the peakedness is the ratio of the variance to the mean of the steady-state number of busy servers in a GI/M/∞ system associated to a GI/M/k system. And Eckberg (1977) showed that this parameter, together with the mean, is a much better second parameter than the variance to characterize the inter-arrival time or service time distribution for the GI/M/k system. Then it is recommended in (Whitt, 1984) to use the peakedness instead of the variance. As for a renewal process, to know the ratio of the variance to the mean is equivalent to know the Laplace transform evaluated in \( 1/\lambda \), follows the definition of the peakedness in Equation (3).

Note 6. Many approximations for the equilibrium mean queue length in the GI/G/1 queue are based on the first two moments of the interarrival-time and service-time distributions; see Shanthikumar and Buzakott (1980) and
Whitt (1982). In Whitt (1984) it is suggested to consider the set-valued function that maps the four moment parameters into the set of possible values of the mean queue length. Here the parameter $ρ$ is composed of infinite moments.

Note 7. In theory of queues, a quantity is said to be insensible to the service time distribution when it depends on it only through its mean.

Note 8. $\alpha = \frac{c}{c+1}$. In this case, $\eta^ρ$ was computed using directly the Expression (3).