Repositório ISCTE-IUL

Deposited in *Repositório ISCTE-IUL*:
2019-02-18

Deposited version:
Post-print

Peer-review status of attached file:
Peer-reviewed

Citation for published item:

Further information on publisher's website:
http://www.actaint.com/

Publisher's copyright statement:
This is the peer reviewed version of the following article: Ferreira, M. A. M. (2018). The M|M| queue transient probabilities. Acta Scientiae et Intellectus. 4 (6), 32-35. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

---

Use policy

Creative Commons CC BY 4.0
The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.
THE $M|M|\infty$ QUEUE TRANSIENT PROBABILITIES

Manuel Alberto M. Ferreira

Instituto Universitário de Lisboa (ISCTE-IUL), Business Research Unit (BRU-IUL) and Information Sciences, Technologies and Architecture Research Center (ISTAR-IUL), PORTUGAL

E-mail: manuel.ferreira@iscte.pt

ABSTRACT

In the theory and applications of queuing systems, transient probabilities play a key role. Often, due to the difficulty of their calculation they are replaced by the stationary probabilities. In the case of $M|G|\infty$ systems, those calculations are much friendlier than usual. We present in this work formulas for the transient probabilities of the $M|M|\infty$ system, where are considered service times with exponential distribution, showing that it is also possible to calculate them with initial conditions different from those usually considered.

Keywords: Transient probabilities, $M|M|\infty$, busy period, exponential distribution.

Mathematics Subject Classification: 60G99

INTRODUCTION

In the $M|G|\infty$ queue, customers arrive according to a Poisson process at rate $\lambda$, upon its arrival receive immediately a service with time length $d. f. G(\cdot)$ and mean $\alpha$. The traffic intensity is $\rho = \lambda \alpha$, see for instance [7].

Call $N(t)$ the number of occupied servers, at time $t, t \geq 0$ in the $M|G|\infty$ queue. Define the transient probabilities $p_{0n}(t)$ as

$$p_{0n}(t) = P\{N(t) = n | N(0) = 0\}, n = 0, 1, 2, ..., t \geq 0 \ (1.1)$$

So, see [7],

$$p_{0n}(t) = \frac{\lambda \int_0^t [1 - G(v)] \, dv \, n^n}{n!} e^{-\lambda \int_0^t [1 - G(v)] \, dv}, n = 0, 1, 2, ...$$

The stationary probabilities are the limit probabilities and so

$$p_{0n} = \lim_{t \to \infty} p_{0n}(t) = \frac{\rho^n}{n!} e^{-\rho}, n = 0, 1, 2, .. \ (1.2)$$

since $\alpha = \int_0^\infty [1 - G(v)] \, dv$ because $G(\cdot)$ is a positive distribution d.f.
As it happens for any queue, in the $M|G|\infty$ queue activity there is a sequence of idle and busy periods. For this queue the study of the busy period length distribution is very important since, as it is part of its definition, a customer must find immediately an available server upon its arrival. So, it is important for the manager to know how many, and how long, servers must be in prevention, see [1] and [5, 6].

Define now
\[ p_{1'n}(t) = P\{N(t) = n|N(0) = 1', n = 0,1,2, ..., t \geq 0 \} \] (1.3)
meaning $N(0) = 1'$ that the time origin is considered in an instant at which a customer arrives at the $M|G|\infty$ system finding it empty, that is: in the beginning of a busy period.

At the instant $t \geq 0$ it may occur:
- Either the customer that arrived at the initial instant abandoned the system, with probability $G(t)$, or goes on being served, with probability $1 - G(t)$,
- The other servers, that were unoccupied at the time origin, are either unoccupied or occupied with 1,2,... customers, with probability $p_{0n}(t), n = 0,1,2, ...$ given by (1.1).

As both systems, the one of the initial customer and the other of the initially unoccupied servers, are independent, consequently:
\[ p_{1'0}(t) = p_{00}(t)G(t) \]
\[ p_{1'n}(t) = p_{0n}(t)G(t) + p_{0n-1}(t)(1 - G(t)), n = 1,2, ... \] (1.4)

The stationary probabilities are also
\[ p_{1'n} = \lim_{t \to \infty} p_{1'n}(t) = \frac{\rho^n}{n!}e^{-\rho}, n = 0,1,2, ... \] (1.5).

THE $M|M|\infty$ CASE

For the $M|M|\infty$ queue, that is exponential service times, (1.1) becomes:
\[ p_{0n}^M(t) = \frac{\left(\rho \left(1 - e^{-\frac{t}{\alpha}}\right)\right)^n}{n!} e^{-\rho \left(1 - e^{-\frac{t}{\alpha}}\right)}, n = 0,1,2, ... \] (2.1)

and (1.4)
\[ p_{1'0}^M(t) = p_{00}^M(t) \left(1 - e^{-\frac{t}{\alpha}}\right) \]
\[ p_{1'n}^M(t) = p_{0n}^M(t) \left(1 - e^{-\frac{t}{\alpha}}\right) + p_{0n-1}(t)e^{-\frac{t}{\alpha}}, n = 1,2, ... \] (2.2),

owing to the exponential distribution lack of memory, that allows to consider as initial instant anyone at which there is only one customer in the system, not necessarily demanding it is the beginning of a busy period.
Analogously to the procedure followed to deduce (1.4), considering two independent subsystems, one with the clients present at the initial instant and another with the infinite servers vacated at that instant, it is concluded, taking into account the memory shortage of the exponential distribution that for the $M|M|\infty$ queue:

\[
p_{mn}^M(t) = \sum_{k=0}^{\infty} \binom{m}{k} \frac{\rho \left(1-e^{-\frac{t}{\alpha}}\right)^n}{(n-k)!} e^{-kt/\alpha-\rho \left(1-e^{-\frac{t}{\alpha}}\right)} \left(1-e^{-\frac{t}{\alpha}}\right)^{m-k} I(n-k), \quad m = 0,1,2, ..., n = 0,1,2, ... \text{ and } I(n-k) = \begin{cases} 1, & n \geq k \\ 0, & n < k \end{cases} \quad (2.3) .
\]

In particular, for $n = 0$:

\[
p_{m0}^M = e^{-\rho \left(1-e^{-\frac{t}{\alpha}}\right)} \left(1-e^{-\frac{t}{\alpha}}\right)^m, m = 0,1,2, ... \quad (2.4) .
\]

**Theorem 2.1**

For the $M|M|\infty$ queue:

i) \( m \geq \rho \)

\( p_{m0}^M(t), t > 0 \) is an increasing function,

ii) \( 0 < m < \rho \)

a) \( p_{m0}^M(t), \) is an increasing function in \( ]0,t_m[ \)

b) \( p_{m0}^M(t), \) is a decreasing function in \( ]t_m,0[ \)

c) The \( p_{m0}^M(t) \) maximum is

\[
p_{m0}^M(t_m) = \left(\frac{m}{\rho e}\right)^m \quad (2.5)
\]

being

\[
t_m = -\alpha \ln \left(1 - \frac{m}{\rho}\right) \quad (2.6),
\]

iii) \( m=0 \)

\( p_{m0}^M(t), t > 0 \) is a decreasing function.

**Dem:** It is enough to have in mind that \( \frac{d}{dt} p_{m0}^M(t) = e^{-\rho \left(1-e^{-\frac{t}{\alpha}}\right)} \begin{pmatrix} \frac{m}{\alpha} - \frac{1}{1-e^{-\frac{t}{\alpha}}} - \frac{1}{\alpha} \end{pmatrix} e^{-\frac{t}{\alpha}} \begin{pmatrix} \frac{m}{\alpha} \frac{1}{1-e^{-\frac{t}{\alpha}}} - \frac{1}{\alpha} \end{pmatrix} .\)
Obs: If $\rho < 1$, $p_{m0}^M(t)$, $m = 1,2, ...$ always can be normalized in order to behave as d.f.. ■

CONCLUSIONS

In this text, the transient probabilities for the $M|G|\infty$ queue are reviewed, and clarified for the $M|M|\infty$ queue. Based on this are determined transient probabilities for the $M|M|\infty$ system, with initial conditions different from the usual, where play a key role the exponential distribution lack of memory.

REFERENCES


17) S. Skiena, “Polya's Theory of Counting.” §1.2.6 in Implementing Discrete Matyhematics: Combinatorics and Graph Theory with Mathematica.