Performance Analysis of Moments of Queue Length Distribution in M/M/1 and M/Er/1 Queues

Manveen Kaur
PIET, Patiala

Sonia
BBSBEC, Fatehgarh Sahib

Abstract

Knowledge of the moments of production and queue lengths are sufficient to carry out a great number of performance evaluation, optimization, and decision-support tasks. With the multiple points of view that exist in any queuing system, several measures are necessary in order to specify overall system Quality of service. In this paper, two methods of queue length moment distribution are discussed. Analysis of moments of queue length in M/M/1 and M/Er/1queue models is represented. The first order and second order moments of queue length of these two queuing models are also compared.

Introduction

Queuing theory is used to address number of questions regarding quality of service, which is a major concern of telecommunication of today. Quality is measured in variety of ways, which includes delay in gaining access to information. In various future technologies and faster internet services, based on packet transmission, where packets arrive, wait in various queues, receive service at various points, and exit after some time, packet delays in queues is considered. Performance evaluation of distributed systems and service-oriented architectures is often based on stochastic models, such as closed queuing networks which are commonly solved by the Mean Value Analysis (MVA) algorithm. However, the MVA is unable to solve models with hundreds or thousands of users accessing services of multiple classes, a configuration that is often useful to predict the performance of real-world applications Compared to the MVA algorithm, which is based on a recursive evaluation of mean queue-lengths, MoM (Method of Moments) defines a recursion on higher-order moments of queue-lengths that is solved at each step by a linear system of equations. This extends the feasibility of exact methods to a much larger family of multiclass performance models than those that can be solved by the MVA algorithm[8].

Moments of the queue length distribution have an important role in quantitative description of the behaviour of many communication networks modelled in terms of queuing theory [1]. For instance, Moments of the job size distribution can provide bounds on the mean waiting time. However, approximations based on the first two moments of the job size distribution cannot be accurate for all types of distributions. The third moment and other higher moments of the job size distribution have a significant impact on the mean waiting time [2]. Knowledge of the moments of production and queue lengths are sufficient to carry out a great number of performance evaluation, optimization, and decision-support tasks. For example, MR models are used to help optimize networks. These models may maximize the performance of a network (as measured by waiting times or inventory levels), minimize the cost of the network (in terms of capacity and cost), and maximize the expected throughput of the network. Control policies may also be defined to reduce production and queue length fluctuations (i.e., minimize the variances) – important in situations requiring output predictability such as Just-in-time manufacturing [3].

In this paper, importance of moments and methods to compute them are illustrated. Along with that analysis of moments of queue length of M/M/1 and M/Er/1with respect to utilization factor are also represented.

Moments of Queue Length Distribution

The moments of a distribution are a set of parameters that summarize various queuing models. Given a random variable X, its first moment about the origin, denoted M_0^1, is defined to be E[X]. Its second moment about the origin, denoted M_0^2, is defined as the expected value of the random variable X^2, or $E[X^2]$. In general, the rth moment of X about the origin, denoted M_0^r, is defined as $M_0^r = E[X^r]$. The third moment about the mean, M_μ^3, is used to construct a measure of skewness, which describes whether the probability mass is more to the left or the right of the mean, compared to a normal distribution. The fourth moment about the mean, M_μ^4, is used to construct a measure of peakedness, or kurtosis, which measures the "width" of a distribution [4].

There are two methods to find moments –

1. The Moment Generating Function (MGF) –

MGF M(t) of a random variable X is the exponential generating function of its sequence of moments. It provides the basis of an alternative route to analytical results compared with working directly with probability density functions or cumulative distribution functions. There are particularly simple results for the moment-generating functions of distributions defined by the weighted sums of random variables. The moment - generating function of a random variable X is

$$Mx(t) = E[e^{tX}]$$

Wherever this expectation exists. Mx(0) always exists and is equal to 1.

$$Mx(t) = 1 + tm_1 + \frac{t^2 m_2}{2!} + \ldots + \frac{t^n m_n}{n!} + \ldots\ldots$$

Where m_n is the nth moment.

2. Recursion Method-

Starting from the basic state transition equations of the queue, we get the alternative method to derive higher moments of queue length instead of taking derivatives. Moreover, complexity to compute moments is reduced to B^2, which is 2^B in MGF[7].

Using this method, the graphical representation of moments of M/M/1 and M/Er/1 with respect to utilization factor is shown.

A. M/M/1 queuing system

In the M/M/1 queue, the arrival process follows a Poisson process with parameter μ and service times are assumed to exponentially distribute with parameter λ and are independent of the arrival process. It is the simplest Markovian queue that has only a single server and an infinite buffer. In this queue with mean arrival rate λ and mean service time 1/ μ, when the offered load ρ= λ /μ = < 1, the kth moment M k (k ≥ 1) of queue length satisfies [7]-

$$M_k = \frac{\sum_{l=2}^{k+1} \binom{k+1}{l}[(-1)^l + \rho] M_{k+1-l} + (-1)^{k+1}[1-\rho]}{(k+1)(1-\rho)}$$

Where
$$M_0 = \sum_{n=0}^{-} P_n = 1.$$

Figure-1 demonstrates the variation of moments of queue length with the increase of utilization factor as it reaches at higher order.
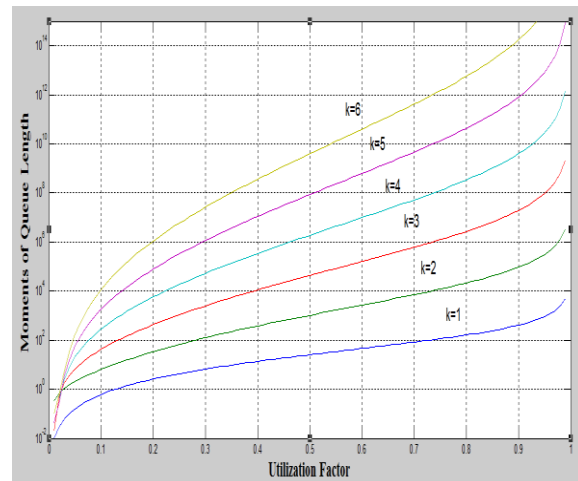

Figure-1 Moments of queue length of M/M/1

B. M/Er/1 queueing system

The Erlang distribution can be used to model service times with a low coefficient of variation (less than one), but it can also arise naturally. For instance, if a job has to pass, stage by stage, through a series of r independent production stages, where each stage takes a exponentially distributed time.

In an M/Er/1 queue with mean arrival rate λ and mean service time $1/\mu$, when the offered load per server $\rho = \lambda/c\mu < 1$, then the kth moment Mk (k≥ 1) of queue length satisfies [7] –

$$M_k = \frac{\sum_{l=2}^{k+1}\binom{k+1}{l}[(-r)^l + r\rho]M_{k+1-l} + (-r)^{k+1}P_0}{r(k+1)(1-\rho)}$$

$$- \frac{\sum_{n=1}^{r-1} n^{k+1}P_0\left(\frac{(\rho r)^{n+r}}{(n+r)!} + (1+r\rho\frac{(r\rho)^n}{n!} - r\rho\frac{\rho r^{n-1}}{(n-1)!}\right)}{r(k+1)(1-\rho)}$$
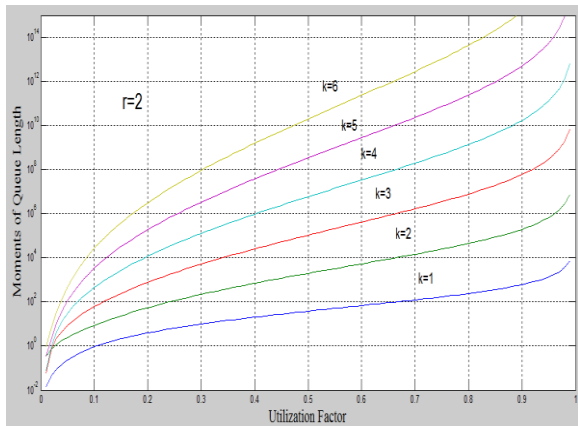

Figure -2 Moments of queue length of M/Er/1 (r=2)

Similarly, figure-2 & figure-3 illustrates the variation in number of moments of queue length at higher order with respect to Utilization Factor. In addition to this, figure- 4 & figure-5 compare the first order and second order moments of queue length in these two queuing models.
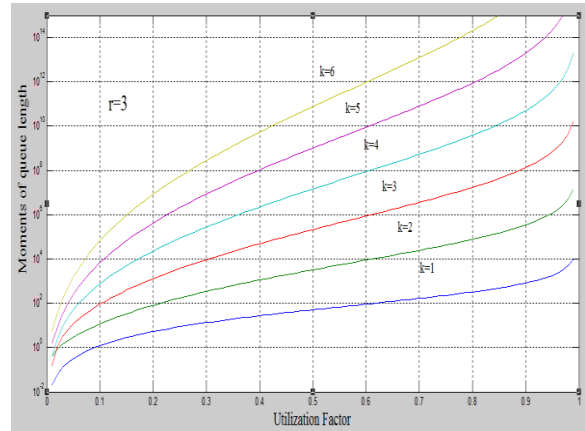

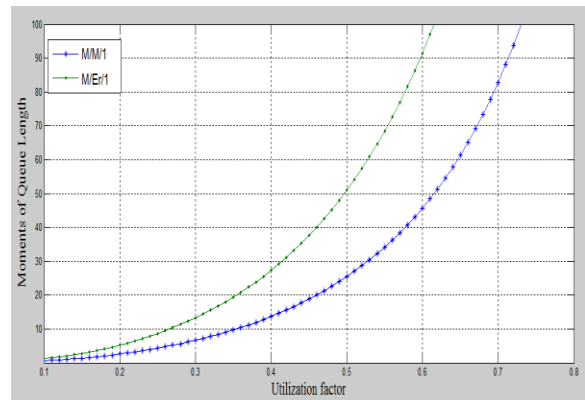**Figure-3** Moments of queue length of M/Er/1 (r=3)

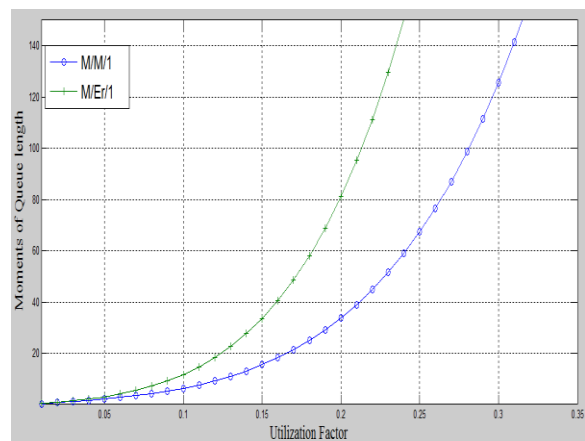**Figure-4 Comparing first moments of M/M/1 and M/Er/1**

Figure-5 Comparing second moments of M/M/1 and M/Er/1

Conclusion

The method of moment (MoM) distribution has been discussed for the performance analysis of various queuing models The basic idea of MoM is to solve queueing models by recursively computing a set of higher-order moments of the stations' queue-length distributions. The advantage of higher-order moments is that they can be computed efficiently in a recursive manner using a system of linear equations involving normalizing constants. The comparative analysis of moments of queue length in M/M/1 and M/Er/1queue models is represented. The first and second moment of queue length of these two queuing models are also compared with respect to utilization.

References

[1] Dudin A., Klimenok V. And Lee M. H., "Recursive formulas for the moments of queue length in the MAP/G/1 queue" , *Communications Letters, IEEE,* Volume: 13 , 2009 , Page(s): 351 – 353

[2] V. Gupta, J.Dai, M. Harchol-Balter and B. Zwart, "The effect of higher moments of job size distribution on the performance of an M=G=K queueing system*", School of Computer Science, Carnegie Mellon University, Pittsburgh*, PA, USA, February 2008

[3] web.mit.edu/sgraves/www/jsh_thesis.pdf

[4] http://en.wikipedia.org/wiki/Moment- generating _function

[5] Adan I., Resing J., "*Queueing Theory*", Eindhoven University of Technology 2004

[6] Cooper R.B., "*Introduction to queueing theory*" North Holland; 2nd edition (1981) | ISBN-10: 0444003797

[7] Liu J., Jiang X, and Horiguchi S., "Recursive Formula for the Moments of Queue Length in the M/M/1 Queue", *IEEE Communication Letters,*vol-12,no.-9 sep. 2008.

[8] Casale G." Exact analysis of performance models by the Method of Moments", *Elsevier ,* Feb. 2011.