Chapter 5 – M/M/1 Queuing Systems

Ioannis Glaropoulos

November 11, 2012

1 Exercise 5.1

In a computer network a link has a transmission rate of C bit/s. Messages arrive to this link in a Poisson fashion with rate λ messages per second. Assume that the messages have exponentially distributed length with a mean of $1/\mu$ bits and the messages are queued in a FCFS fashion if the link is busy.
a) Determine the minimum required C for given λ and μ such that the average system time (service time + waiting time) is less than a given time T_0.

Solution: System Description

- Single communication link: C bits per second
- Poisson arrivals: λ messages per second
- Exponential Service times: $E[T] = E[X]/C = 1/(\mu C)$, so the *exponential rate* is μC.
- First Come First Served policy
- Infinite Queue[1]

This is a typical M/M/1 System. We see the system diagram in Fig. 1. We first derive the state distribution (steady-state) of this system through the solution of the balance equations. We *define* $\rho = \lambda/(\mu C)$. For a no-loss system, ρ is the OFFERED and, at the same time, the ACTUAL load.

$$\lambda P_0 = (\mu C)P_1 \rightarrow P_1 = \rho P_0$$
$$\lambda P_1 = (\mu C)P_2 \rightarrow P_2 = \rho P_1 = \rho^2 P_0$$
$$\lambda P_2 = (\mu C)P_3 \rightarrow P_3 = \rho P_2 = \rho^3 P_0$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\lambda P_k = (\mu C)P_{k+1} \rightarrow P_{k+1} = \rho P_k = \rho^k P_0$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$

Then, we calculate the P_0 through the normalization equation:

$$\sum_{k=0}^{\infty} P_k = 1 \rightarrow \sum_{k=0}^{\infty} \rho^k P_0 = 1 \rightarrow P_0 \sum_{k=0}^{\infty} \rho^k = 1 \rightarrow P_0 \cdot \frac{1}{1-\rho} = 1 \rightarrow P_0 = 1 - \rho.$$

[1] If no buffer capacity is mentioned, we always assume that this is infinite.

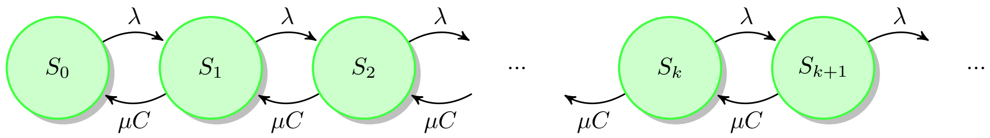

Figure 1: System diagram for the M/M/1 chain of exercise 5.1

Finally, the state distribution is given as

$$P_k = (1-\rho)\rho^k.$$

We, now, derive the average number of messages in the system, using the state distribution:

$$\overline{N} = \sum_{k=0}^{\infty} kP_k = \sum_{k=0}^{\infty} k(1-\rho)\rho^k = (1-\rho)\rho \sum_{k=0}^{\infty} k\rho^{k-1} =$$

$$= (1-\rho)\rho \sum_{k=0}^{\infty} \frac{d\rho^k}{d\rho} = (1-\rho)\rho \frac{d\left(\sum_{k=0}^{\infty} \rho^k\right)}{d\rho} = (1-\rho)\rho \frac{d(1/(1-\rho))}{d\rho} = \frac{\rho}{1-\rho}.$$

In order to solve the first question we can use the LITTLE's formula:

$$\overline{N} = \lambda_{\text{eff}} E[T_{total}] \to E[T_{\text{total}}] = \frac{\overline{N}}{\lambda} = \frac{\rho/(1-\rho)}{\lambda} = \frac{\lambda/(\mu C)/(1-\lambda/(\mu C))}{\lambda},$$

since $\lambda_{eff} = \lambda$, so, finally,

$$E[T_{total}] = \frac{1}{(\mu C) - \lambda}.$$

The minimum required C is determined by:

$$\frac{1}{\mu C - \lambda} \leq T_0 \to \mu C - \lambda \geq T_0^{-1} \to C \geq \frac{\lambda + T_0^{-1}}{\mu}.$$

2 Exercise 5.5

Consider a queuing system with a single server. The arrival events can be modeled with Poisson distribution, but two customers arrive at the system at each arrival event. Each customer requires an exponentially distributed service time.

1. Draw the state diagram
2. Determine p_k using local balance equations
3. Let $P(z) = \sum_{k=0}^{\infty} z^k p_k$. Calculate $P(z)$ for the system. Note, that $P(z)$ must be finite for $|z| < 1$, and we know $P(1) = 1$.

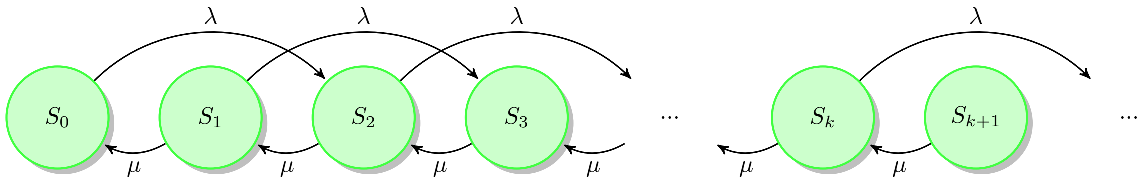

Figure 2: System diagram for the M/M/1 chain of exercise 5.5

4. Calculate the mean number of customers in the system with the help of $P(z)$ and compare it with the one of the M/M/1 system.

Solution: The system can be described by an M/M/1 model, since there is a single server, the service times are exponential service and the arrival process is Poisson. We must notice, however, that this Poisson Process models arrival events, but the events consist of two customer arrivals. (The departure events are still one-by-one, though.)

As always, for a Markovian System we must guarantee that all transitions are exponential. We define the usual state space: S_k : k customers in the system. Then, the state diagram is straightforward. Special care must be taken on determining the transitions and rates from state to state.

$$\textbf{Departure rate} = \mu$$
$$\textbf{Arrival Event rate} = \lambda$$

Clearly, the average customer arrival rate is 2λ and is NOT Poisson! What IS Poisson is the group arrival rate. We also DEFINE $\rho = \frac{\lambda}{\mu}$. This is neither the offered nor the actual load. We just use ρ to define this fraction.
The system diagram is given in Fig. 2.
Local Balance Equations:

$$\lambda P_0 = \mu P_1$$
$$\lambda P_{k-2} + \lambda P_{k-1} = \mu P_k, \quad k \geq 2$$

We can go ahead and solve them numerically. Alternatively, we can use the ZT methodology, since we only want to compute the average number of customers.

We consider the parametric local balance equation:

$$\mu P_k = \lambda P_{k-1} + \lambda P_{k-2} \rightarrow$$

$$\rightarrow \sum_{k=2}^{\infty} z^k \mu P_k = \sum_{k=2}^{\infty} z^k (\lambda P_{k-1} + \lambda P_{k-2})$$

$$\rightarrow \mu(P(z) - zP_1 - P_0) = \sum_{k=2}^{\infty} \lambda z^k P_{k-1} + \sum_{k=2}^{\infty} \lambda z^k P_{k-2}$$

$$\rightarrow \mu(P(z) - zP_1 - P_0) = \lambda z \sum_{k=2}^{\infty} \lambda z^{k-1} P_{k-1} + \lambda z^2 \sum_{k=2}^{\infty} \lambda z^{k-2} P_{k-2}$$

$$\rightarrow \mu(P(z) - zP_1 - P_0) = \lambda z(P(z) - P_0) + \lambda z^2 P(z)$$

We solve the equation with respect to $P(z)$

$$P(z) = \frac{\mu P_0 + \mu z P_1 - \lambda z P_0}{\mu - \lambda z - \lambda z^2} = \frac{P_0 + z P_1 - \rho z P_0}{1 - \rho z - \rho z^2}. \tag{1}$$

We need to apply two conditions that HOLD, in order to determine the unknown terms above. The first condition comes from the balance equation that we did not consider. We replace $P_1 = \rho P_0$ in (1), and obtain:

$$P(z) = \frac{P_0}{1 - \rho z - \rho z^2}. \tag{2}$$

The second condition comes from the NORMALIZATION in the probability or in the Z-domain:

$$\sum_{k=0}^{\infty} P_k = 1, \quad \text{or,} \quad P(z=1) = 1.$$

Replacing that in (2) we obtain $P_0 = 1 - 2\rho$, so finally

$$P(z) = \frac{1 - 2\rho}{1 - \rho z - \rho z^2} \tag{3}$$

Finally, we need to compute the mean number of customers. We have

$$\overline{N} = \left[\frac{dP(z)}{dz}\right]_{z=1}.$$

Proof:

$$\left[\frac{dP(z)}{dz}\right]_{z=1} = \left[\frac{d \sum_{k=0}^{\infty} z^k P_k}{dz}\right]_{z=1} = \left[\sum_{k=0}^{\infty} k z^{k-1} P_k\right]_{z=1} = \sum_{k=0}^{\infty} k P_k = \overline{N}.$$

So, this is what we will do. We differentiate the derived ZT in (3):

$$\frac{dP(z)}{dz} = \frac{(-1)(1 - 2\rho)(-\rho - 2\rho z)}{(1 - \rho z - \rho z^2)^2}$$

Replacing $z = 1$ we obtain

$$\overline{N} = \frac{3\rho}{1 - 2\rho} = \frac{3\lambda}{\mu - 2\lambda}.$$

The typical M/M/1 system with the same average customer arrival rate (2λ) and service rate (μ) has $\overline{N}_{M/M/1} = \frac{\rho}{1-\rho}$, where ρ is its offered load, and is equal to $\rho = 2\lambda/\mu$. So, finally,

$$\overline{N}_{M/M/1} = \frac{2\lambda}{\mu - 2\lambda}$$

so it is different, and, actually, less. Why?

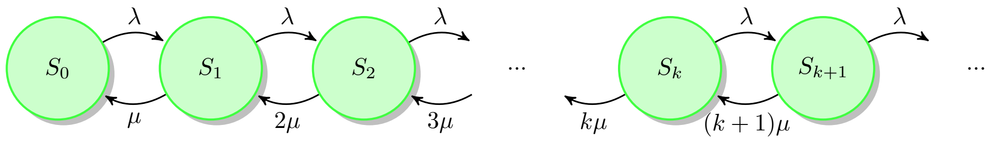

Figure 3: System diagram for the M/M/1 chain of exercise 5.6

3 Exercise 5.6

A queuing system has one server and infinite queuing capacity. The number of customers in the system can be modeled as a birth-death process with $\lambda_k = \lambda$ *and* $\mu_k = k\mu$, $k = 0, 1, 2, ...$ *thus, the server increases the speed of the service with the number of customers in the queue. Calculate the average number of customers in the system as a function of* $\rho = \lambda/\mu$.

Solution: The system is an M/M/1 queue, since it has infinite buffer, 1 server, and Markovian arrival and departure process. However, as we can see, it is not a typical M/M/1 case, as the service rates depend on the current system state. The system diagram is shown in Fig. 3. We need to solve the system of balance equations:

$$\begin{aligned}
&\lambda P_0 = \mu P_1 \rightarrow P_1 = \rho P_0 \\
&\lambda P_1 = 2\mu P_2 \rightarrow P_2 = \tfrac{1}{2}\rho P_1 = \tfrac{1}{2}\rho^2 P_0 \\
&\lambda P_2 = 3\mu P_3 \rightarrow P_3 = \tfrac{1}{3}\rho P_2 = \tfrac{1}{2\cdot 3}\rho^3 P_0 \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&\lambda P_{k-1} = k\mu P_k \rightarrow P_k = \tfrac{1}{k}\rho P_{k-1} = \dots = \tfrac{1}{k!}\rho^k P_0 \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&\textstyle\sum_{k=0}^{\infty} P_k = 1 \quad \text{(normalization)}
\end{aligned}$$

From the last general equation and the normalization equation we obtain the state distribution:

$$\sum_{k=0}^{\infty} \frac{\rho^k}{k!} P_0 = 1 \rightarrow P_0 e^{\rho} = 1 \rightarrow P_0 = e^{-\rho}.$$

so finally, for each k

$$P_k = \frac{\rho^k}{k!} e^{-\rho}$$

so the state distribution is POISSON! Then, we can calculate the average number of customers from the state distribution

$$\overline{N} = \sum_{k=0}^{\infty} k P_k = \rho$$

or simply say that the average is ρ, from the Poisson distribution.
From LITTLE we can, also, calculate the average system time

$$E[T_{total}] = \frac{\overline{N}}{\lambda} = \frac{1}{\mu}.$$

This means that the arriving customers only stay in the system for an average time equal to the service time![2]

4 Exercise 5.7

Customers arrive to a single server system in groups of 1,2,3 and 4 customers. The number of customers per group is i.i.d. There are in total 4 places in the system. If a group of customers does not fit into the system, none of the members of the group joins the queue. 10% of the customers arrive in a group of 1, 20% of the customers arrive in groups of 2, 30% in a group of 3 and 40% in a group of 4 customers. The average number of arriving customers is 75 customers per hour, the interarrival time between groups is exponentially distributed. The service time is exponentially distributed with a mean of 0.5 minutes.

1. *Give the Kendall notation of the system and draw the state transition diagram.*
2. *Calculate the average number of customers in the queue and the mean waiting time per customer.*
3. *Calculate the probability that the system is full and the probability that a customer arriving in a group of k customers can not join the queue.*
4. *Calculate the probability that an arriving customer in general can not join the queue and the probability that an arriving group of customers can not join the queue.*
5. *What is the average waiting time for a customer arriving in a group of 3 customers?*

Solution: This is a very interesting problem that reveals the problems when the arriving process is complex and not straightforward so it must be derived.

First, we give the Kendall notation of the system. We have:

- Exponential GROUP inter-arrival times, so the arrival time will be Markovian.[3]
- The service times are exponentially distributed
- The system has a single server
- The total capacity is 4

[2]This is equivalent to the case where there is no queue and each customer is served in parallel with the others, so actually this system is equivalent to an M/M/∞ system!

[3]It is our task to find an appropriate state space where the event arrival process is Poisson.

Consequently, the Kendall notation is M/M/1/4.
We must draw the state transition diagram. We consider the typical state space where S_k means "k customers in the system". As a result the system has 5 states in total. The service rates are always the same, with

$$\mu = \frac{1}{E[T_s]} = \frac{1}{0.5/60} = 120h^{-1}.$$

The difficulty lies in deriving the arrival transition rates. We are given that the number of customers per group is i.i.d. We assume, naturally, that the GROUP arrival process is HOMOGENEOUS, that is, groups arrive in each state with the same rate! Also, since the inter-arrival times between GROUP arrivals are exponential we conclude that the GROUP arrivals is a Poisson process.

We are given that the number of customers per group is an i.i.d. process, but we are NOT given the distribution. Let

$$q_1, \quad q_2, \quad q_3, \quad q_4$$

denote the probabilities that a random arriving group contains 1,2,3,4 customers respectively.

Let λ_G denote the Poisson **group** arrival rate. Then, the individual rates for each groups is ALSO a Poisson process, based on the *Poisson split* property, with rates

$$\lambda_G q_1, \quad \lambda_G q_2, \quad \lambda_G q_3, \quad \lambda_G q_4.$$

For any $i = 1, 2, 3, 4,$

$$\lambda_G \cdot q_i$$

defines the (average) rate of arrivals for group's of type i, and, consequently,

$$\lambda_G \cdot q_i \cdot i$$

defines the (average) rate of arrivals of customers belonging to group of type i.

Based on the given data from the exercise regarding the ratio of customers arriving in any of the groups, we obtain the following equations:

$$\lambda_G q_1 \cdot 1 = 10\% \cdot 75$$

$$\lambda_G q_2 \cdot 2 = 20\% \cdot 75$$

$$\lambda_G q_3 \cdot 3 = 30\% \cdot 75$$

$$\lambda_G q_4 \cdot 4 = 40\% \cdot 75$$

From the above it is clear that $q_1 = q_2 = q_3 = q_4 \rightarrow q_i = \frac{1}{4}, \quad \forall i = 1, 2, 3, 4.$
Finally, using any of the above equations we compute the group arrival rate:

$$\lambda_G = 30 \text{ groups / hour}$$

We can now complete the state diagram (Fig. 4). Then, we solve the LOCAL balance equations, to define the state probabilities.

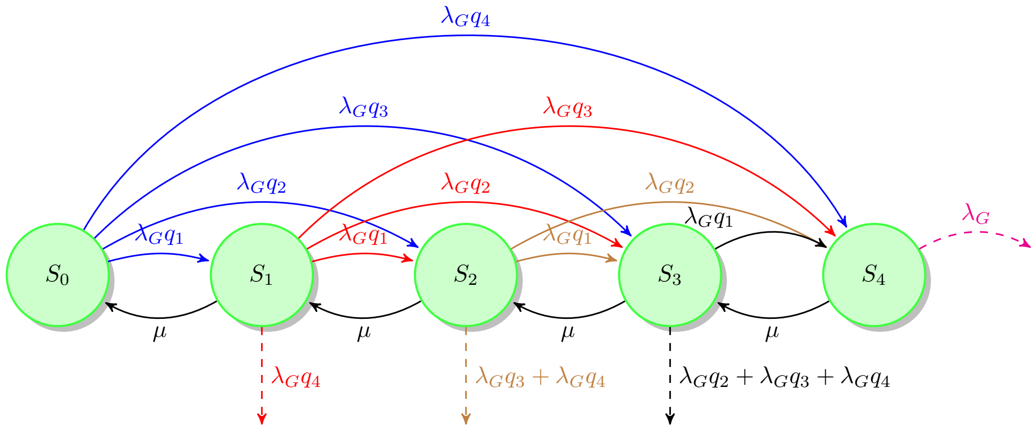

Figure 4: System diagram for the M/M/1 chain of exercise 5.6

We can compute BOTH the average number of customers in the system, and the average number of customers in the queue:

$$\overline{N}_{queue} = 1 \cdot P_2 + 2 \cdot P_3 + 3 \cdot P_4$$
$$\overline{N} = 1 \cdot P_1 + 2 \cdot P_2 + 3 \cdot P_3 + 4 \cdot P_4$$

For the average waiting time, we could apply the LITTLE result. For that we need the effective customer arrival rate, which is different from the nominal, since there are losses in the system.

It is important to notice again that the system is HOMOGENEOUS in group arrivals but not in customer arrivals.

We have

$$\lambda_{eff} = P_0\lambda_G \cdot (q_1+2q_2+3q_3+4q_4)+P_1\lambda_G \cdot (q_1+2q_2+3q_3)+P_2\lambda_G \cdot (q_1+2q_2)+P_3\lambda_G \cdot (q_1).$$

Then, from LITTLE we compute first the system time, $\overline{N} = \lambda_{eff}E[T]$, and then the average waiting time will be $\overline{W} = E[T] - E[T_s]$.

The probability that the system is full (as seen by an independent observer) is simply P_4.

The probability that a customer of a group k does not join the queue is, actually, the probability that the whole particular group does not join the queue. Since the system is homogeneous in group arrivals (a random group SEES state distribution),

$$\begin{aligned}
\Pr(\text{a random group 1 is blocked}) &= P_4 \\
\Pr(\text{a random group 2 is blocked}) &= P_4 + P_3 \\
\Pr(\text{a random group 3 is blocked}) &= P_4 + P_3 + P_2 \\
\Pr(\text{a random group 4 is blocked}) &= P_4 + P_3 + P_2 + P_1
\end{aligned}$$

An arriving customer in general, belongs to groups 1,2,3,4 with probabilities 10%, 20%, 30% and 40%. So given these probabilities, he follows the group blocking probabilities:

$$\Pr(\text{a random customer is blocked}) =$$
$$= 40\% \cdot (P_1 + P_2 + P_3 + P_4) + 30\% \cdot (P_2 + P_3 + P_4) + 20\%2 \cdot (P_3 + P_4) + 10\% \cdot P_4$$

An arriving group of customers is blocked with probability

$\Pr(\text{a random group is blocked}) =$
$= \sum_{i=1}^{4} \Pr\{\text{a random group has } i \text{ customers}\} \cdot \Pr\{\text{a random group } i \text{ is blocked}\} =$
$= P_4 + P_3 \cdot 3/4 + P_2 \cdot 1/2 + P_1 \cdot 1/4.$

A customer that arrives in group of 3 customers MEANS that the arrived group sees either state 0 or state 1, otherwise there is no mean of WAITING time, since the group is rejected. The arrivals are homogeneous in groups, so the groups see state 0, 1 with probabilities P_0, P_1, respectively. So

$$\overline{W}_3 = \frac{P_0}{P_0 + P_1} \cdot W_3^0 + \frac{P_1}{P_0 + P_1} \cdot W_3^1$$

where the two waiting times are

$$W_3^0 = \frac{1}{3}(0.5 + 1 + 0), \qquad W_3^1 = \frac{1}{3}(0.5 + 1 + 1.5)$$