Lecture 3
Discrete-time Queues (2)

Yin Sun

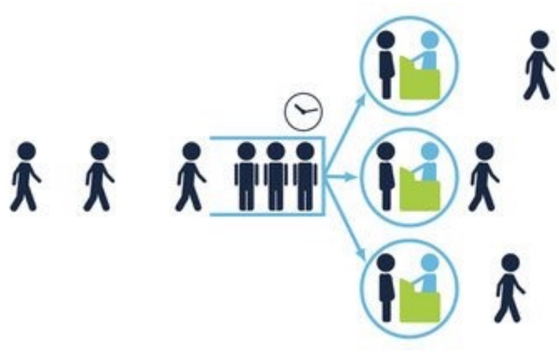
Dept. Electrical and Computer Engineering

Outline

- Discrete-time Markov Chains
- Little's Law
  - Reading: Section 3.4 of Srikant & Ying
- The Geo/Geo/1 Queue

- R. Srikant and Lei Ying, *Communication Networks: An Optimization Control and Stochastic Networks Perspective*, Cambridge University Press, 2014.

Little's Law



- **Little's Law:**

$$L = \lambda W$$

where L is the average number of customers in the system, W is the average time spent in the system, and λ is the arrival rate of the system.

- **Intuition:**
  - Suppose that the customers pay the system \$1 for each second spent in the system. Since a customer spends an average of W seconds in the system, the average rate of revenue for the system is λW \$/s.
  - The average number of customers in the system is L. Thus, the rate of revenue is also L \$/s.

Rigorous Statement

- Let $A(t)$ be the number of customer/packet arrivals up to and including time slot t
  - $q(t)$ be the queue length at time slot t
  - w_i be the waiting time of packet i

- Define $\lambda(T)$ as the average arrival rate by time slot T: $\lambda(T) = \frac{1}{T} A(T)$
  - $L(T)$ as the average queue length by time slot T: $L(T) = \frac{1}{T} \sum_{t=1}^{T} q(t)$
  - $W(n)$ as the average waiting time of the first n packets: $W(n) = \frac{1}{n} \sum_{i=1}^{n} w_i$

- Consider these three limits:

$$\lambda = \lim_{T \to \infty} \lambda(T),\ L = \lim_{T \to \infty} L(T),\ W = \lim_{n \to \infty} W(n)$$

  - So λ is the average arrival rate, L is the average queue length, and W is the average waiting time.

Little's Law

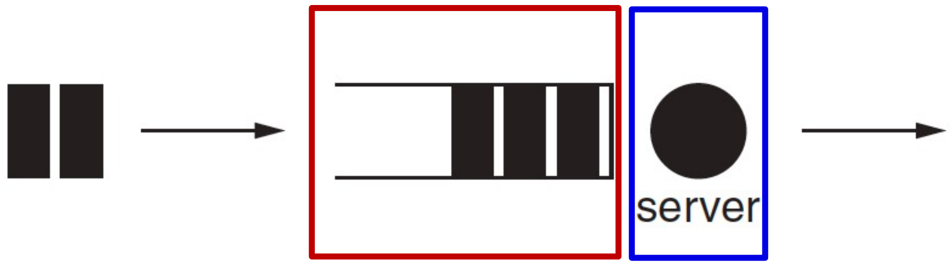
- **Theorem 3.4.1**

Assuming that λ and W exist and are finite, then L exists and $L = \lambda W$.

  - Proof: See Section 3.4 of Srikant & Ying.

- To use Little's law, we need to assume λ and W exist and are finite.

Applications of Little's Law



- Little's Law can be applied to
  - Just the waiting facility of a service center

    Mean number in the queue = Arrival rate x Mean waiting time
  - Customs currently receiving the service

    Mean number in service = Arrival rate x Mean service time

Outline

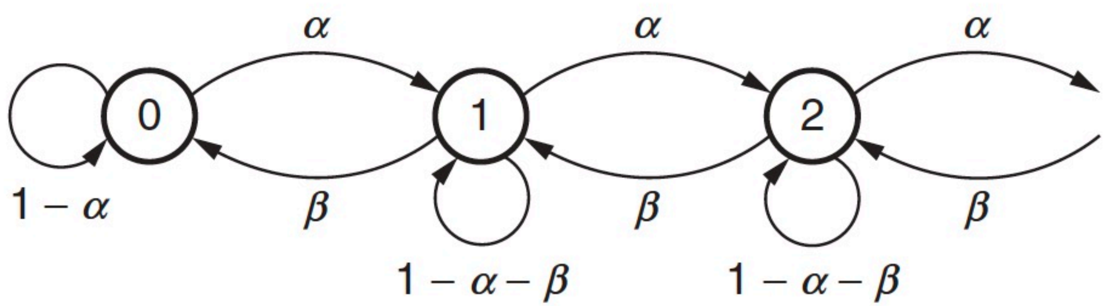
- Discrete-time Markov Chains
- Little's Law
- The Geo/Geo/1 Queue
  - Reading: Section 3.4 of Srikant & Ying
  - R. Srikant and Lei Ying, *Communication Networks: An Optimization Control and Stochastic Networks Perspective*, Cambridge University Press, 2014. https://www.academia.edu/35591665/Communication_Networks_An_Optimization_Control_and_Stochastic_Networks_Perspective

The Geo/Geo/1 Queue

- Consider a discrete-time single-server queue with infinite buffer space
  - Arrival: i.i.d. Bernoulli process with parameter λ
  - Service: a packet is served with prob. μ, not served with prob. $1 - \mu$

- Under these assumptions, the inter-arrival and departure times are geometrically distributed, so the queue is called a Geo/Geo/1 queue.

As a Birth–Death Process



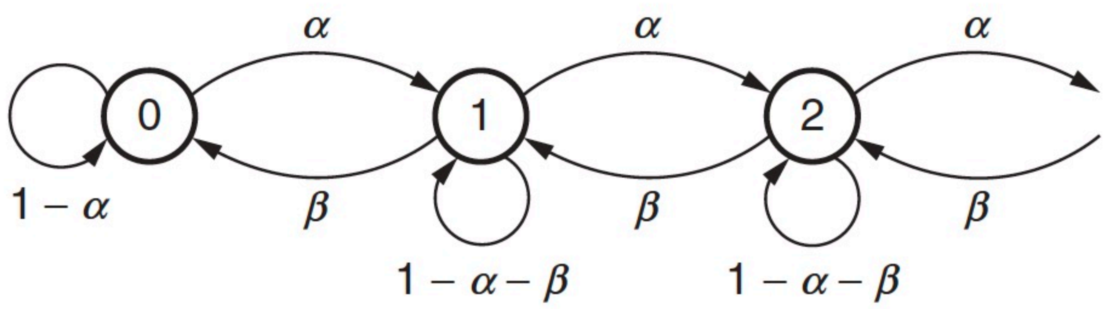
- Probability that the queue length increases by one:

$$\alpha = \Pr(1\ arrival, no\ departure) = \lambda(1 - \mu)$$

- Probability that the queue length decreases by one:

$$\beta = \Pr(no\ arrival, 1\ departure) = (1 - \lambda)\mu$$

Transition Matrix



- Transition matrix P:

$$\mathbf{P} = \begin{bmatrix} 1-\alpha & \alpha & 0 & 0 & 0 & \dots \\ \beta & 1-\alpha-\beta & \alpha & 0 & 0 & \ddots \\ 0 & \beta & 1-\alpha-\beta & \alpha & 0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Steady-state Distribution

- Solve $\pi = \pi$ P, yields

$$\pi_0 = (1-\alpha)\pi_0 + \beta\pi_1$$
$$\pi_i = \alpha\pi_{i-1} + (1-\alpha-\beta)\ \pi_i + \beta\pi_{i+1}$$

  - From this, we get

$$\beta\pi_{i+1} = \alpha\ \pi_i$$

- Define $\rho\ =\ \alpha/\beta\ =\ \lambda(1\ -\ \mu)/\mu(1\ -\ \lambda)$
  - Then, $\pi_i = \rho^i\ \pi_0$

$$\sum_{i=1}^{\infty} \pi_i = \pi_0 \sum_{i=1}^{\infty} \rho^i = 1$$

Steady-state Distribution (2)

- If $\rho < 1$, i.e., $\lambda/\mu < 1$, then

$$\pi_0 = 1 - \rho$$

$$\pi_i = \rho^i(1 - \rho)$$

  - The Markov chain is positive recurrent.
- If $\rho \geq 1$, then

$$\sum_{i=1}^{\infty} \pi_i = \infty$$

  - The Markov chain is not positive recurrent.

Average Queue Length

- If $\rho < 1$, the average queue length is

$$
\begin{aligned}
E[q] &= \sum_{i=0}^{\infty} \rho^i (1-\rho) i \\
&= (1-\rho)\rho \sum_{i=1}^{\infty} i \rho^{i-1} \\
&= (1-\rho)\rho \frac{1}{(1-\rho)^2} \\
&= \frac{\rho}{1-\rho}.
\end{aligned}
$$

Average Waiting Time

- By Little's law, the average waiting time of a packet is

$$W = \frac{L}{\lambda} = \frac{\rho}{\lambda(1-\rho)}$$

Summary

- **Little's Law**

  In a queueing system, if the mean arrival rate and expected waiting time exist, the mean queue length is equal to the product of the mean arrival rate and the expected waiting time.

- **The Geo/Geo/1 queue**
  - The mean queue length equals $\rho/(1 - \rho)$ and the expected waiting time equals $\rho/\lambda(1 - \rho)$.

- **The Geo/Geo/1/B queue**
  - Packet dropping, read Srikant & Ying by yourself

- Reading: Section 3.4 of Srikant & Ying