Week 5: Markov chains
Random access in communication networks

A pervasive situation in communication systems is to have a common infrastructure known as an *access point* (AP) serving a group of physically distributed devices. For instance, this is the situation of your cellphone sending information to a cellular base station, your laptop transmitting to the 802.11 wireless router, or a group of satellites communicating with a ground station. In these examples, there is a common AP—the base station, the router, or the ground station—serving a set of devices—cellphones, laptops or satellites. One problem that occurs during the "uplink" communication—i.e., communication from the distributed terminals to the AP—is how to separate the information transmitted by different devices. One way is to assign different times or frequencies to different terminals. These systems are know as *time* or *frequency division multiplexing.* Another solution is to let terminals transmit at random and hope that luck will avoid simultaneous transmissions. This is called random access (RA). The main advantage of RA is that it requires almost no coordination between terminals. It is not clear, however, that RA is a viable communication strategy: how often will communication be successful? What type of protocol should we put in place to maximize the probability of successfully communicating? In what follows, we define a stochastic model for an RA protocol and study some its key performance indicators.

Consider J distributed devices that try to communicate with the AP during time slots $n = 1, 2, \ldots$ Each device maintains a buffer in which packets are stored to await transmission to the AP. Denote as Q_{jn} the number of packets in the buffer of the j-th terminal at time n. In each time slot n, device j produces a new packet with probability λ and, if the transmission buffer is not empty, attempts to transmit a packet with probability p. If the communication is successful—something that for now we say happens with probability q—the packet is removed from the device's buffer since it was delivered to the AP. If not, the packet stays in the buffer to await retransmission at a later time.

To simplify the analysis, we introduce the following two assumptions:

(A) **No concurrence**: we assume that packet transmission and creation of a new packet never occur in the same time slot.

(B) **Dominant system**: if the transmission queue of device j is empty—i.e., if $Q_{jn} = 0$—terminal j still transmits a dummy packet with probability p. This packet carries no information and is transmitted with the sole purpose of interfering with other terminals. This dummy packet makes it so that the probability of interference occurring due to any terminal j is independent of the number of packets in its queue Q_{jn}.

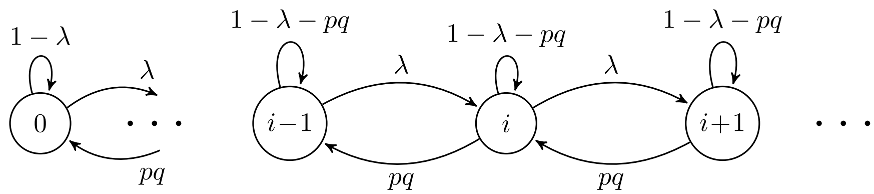

Figure 1: State transition diagram for queue lengths.

A Markov chain model. The number Q_{jn} of packets in queue is a Markov chain (MC). Assuming no concurrence as per (A), its transition probabilities are given by

$$\begin{aligned}
\mathbb{P}\left[Q_{j(n+1)} = k+1 \mid Q_{jn} = k\right] &= \lambda, \quad \text{for all } k \\
\mathbb{P}\left[Q_{j(n+1)} = k-1 \mid Q_{jn} = k\right] &= pq, \quad \text{for } k \neq 0 \\
\mathbb{P}\left[Q_{j(n+1)} = k \mid Q_{jn} = k\right] &= 1-\lambda-pq, \quad \text{for } k \neq 0 \\
\mathbb{P}\left[Q_{j(n+1)} = 0 \mid Q_{jn} = 0\right] &= 1-\lambda
\end{aligned}$$

A transition diagram of this MC is shown in Fig. 1. Explain these transition probabilities were obtained. For which combinations of parameters is this MC recurrent? For which parameters is it ergodic? Explain.

B Limit distribution. Assuming the communication system has been operating for a long time—large n—, find the probability distribution of the queue length Q_{jn}. Formally, we seek

$$\pi_k = \lim_{n\to\infty} \mathbb{P}\left[Q_{jn} = k\right], \quad \text{for all } k. \tag{1}$$

To find the limit distributions in (1), notice that the system is fundamentally different if $\lambda < pq$, $\lambda = pq$, and $\lambda > pq$. Discuss these three cases and show that for $\lambda < pq$ the limit distribution has the form $\pi_k := c\alpha^k$. Find the constants α and c.

C Probability of empty queue and probability of minimal wait. An important performance metric for communication systems is the probability of a terminal's queue becoming empty. What is the value of this probability for sufficiently large n? Two other related figures of merit that quantify the communication delay are (i) the probability T_1 of a packet being transmitted in the first time slot after its arrival and (ii) probability S_1 of a packet being *successfully* transmitted in the first slot after its arrival. Compute T_1 and S_1 for large n.

D Expected queue length. Yet another performance metric is the expected queue length $\mathbb{E}[Q_{jn}]$. This measure is related to the memory a terminal must allocate for its queue. Compute this metric for large n, i.e., report

$$\bar{Q}_j = \lim_{n\to\infty} \mathbb{E}\, Q_{jn}. \tag{2}$$

E Probability of successful transmission and optimal transmission probability $p^\star$. Compute the probability of successful transmission pq under the dominant system hypothesis (B). Prove that making $p = p^\star = 1/J$ maximizes the probability of successful transmission pq and, consequently, both maximizes the asymptotic probability of empty queues $\mathbb{P}\left[Q_{nj} = 0\right]$ and minimizes the expected queue length $\bar{Q}_j$. For $p^\star$, write down the corresponding probability of a transmitted packet reaching the AP ($q^\star$) and show that as the number of terminals $J \to \infty$, the number of successful transmissions converges to $1/e \approx 0.36$. This is remarkable! RA is able to use 36% of the available resources without any coordination overhead between terminals.

F Average time occupancies. We could argue that the probabilities from Parts B and C, as well as the expected value computed in Part D, are of little practical value. Indeed, these probabilities and expectations express averages across all possible paths of the communication system, i.e., across realizations ("ensemble averages"). So if we run the system and obtain a certain path $Q_{jn}^{(1)}$, then run it again to yield a path $Q_{jn}^{(2)}$, and so on. These probabilities measure how likely different events are across these different realizations of the stochastic process. In a practical implementation, however, we need performance measures that hold for each run of the system. One such metrics is, for instance, the following time average

$$\bar{p}_k = \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \mathbb{I}\left(Q_{jm} = k\right), \quad \text{for all } k \geq 0, \tag{3}$$

which represents the fraction of time there were k packets awaiting transmission in the j-th queue. Different from the π_k in (1), the $\bar{p}_k$ in (3) are a figure of merit associated with each experiment. Find $\bar{p}_k$ and explain why, despite our criticisms, π_k is actually useful. What useful performance metric can you define that is equivalent to the expected value in (2)?

G System simulation. Write a function that simulates this system without assuming the dominant system hypothesis (B). To simplify the implementation of the no concurrence hypothesis (A), suppose that packet arrivals have precedence over transmissions, i.e., only if no new packet arrives at the terminal can it try to transmit. The function should take as inputs the number of users J, the transmission probability p, the packet arrival rates λ, and a simulation time parameter N. It must then return a history of the number of packets in each of the J queues ($\{R_{jn}\}$) between times 1 and N. All queue lengths should be initialized at 0. Run your function for $J = 16$ users using the optimal transmission probability $p^\star$ from Part E, $\lambda = 0.9(pq)$, and $N = 10^4$. Show a graph with the path followed by terminals 1 through 4.

H Compare numerical and analytical results. Define the limit distribution of the simulated system without the dominant system hypothesis as

$$\xi_k = \lim_{n\to\infty} \mathbb{P}\left[R_{jn} = k\right]. \tag{4}$$

These probabilities cannot be computed in closed form. Use a *single run* of your function from Part G to estimate the distribution of (4). Explain why this is possible. Plot your estimate of the limit distribution ξ_k and compare it with π_k. Comment on the relevance and usefulness of the dominant system hypothesis.