Homework 3

Simulation and Performance Evaluation – University of Trento

You can solve the following assignments using any programming language. In doing so, make sure to implement the formulas explained in class, and try not to use functions made available by the languages to achieve the required tasks. However, you are allowed to use utility functions.

Exercise 1

Consider the following "weird" probability density function:

$$f(x) = \frac{|\mathrm{sinc}(x)|}{A} = \frac{1}{A}\left|\frac{\sin(\pi x)}{\pi x}\right|$$

where $-6 \le x \le 6$, and $A = 1.8988$ is a normalization factor such that $\int_{-6}^{6} f(t)\,\mathrm{d}t = 1$, so $f(x)$ is in fact a PDF. (Recall that $\mathrm{sinc}(x)$ is not defined for $x = 0$, but can be continuously extended to take the value $\mathrm{sinc}(0) = 1$.)

1. Employ rejection sampling to draw a large number of samples from the above PDF.
2. Plot the resulting empirical PDF (e.g., through a histogram) and compare it against $f(x)$.

Exercise 2

Consider a finite discrete-time Markov chain with the following states:

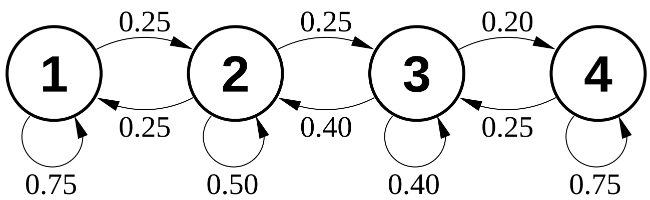

Figure 1: Markov chain considered in Exercise 1.

The states are numbered from 1 to 4. For each state, the arrows are labeled by the *transition probabilities*. These probabilities represent the likelihood that the chain will evolve from a given state (the tail of the arrow) to another of the states (the head of the arrow), upon two subsequent discrete time steps. The graph above can be translated into the following transition probability matrix P:

$$P = \begin{bmatrix} 0.75 & 0.25 & 0 & 0 \\ 0.25 & 0.50 & 0.25 & 0 \\ 0 & 0.40 & 0.40 & 0.20 \\ 0 & 0 & 0.25 & 0.75 \end{bmatrix}. \tag{1}$$

For example, the probability p_{23} that the chain will be in state 3 at time t given that it is in state 2 at a given time $t-1$ is 0.25. Markov chains like the one above are often used to model the performance of wireless communication channels. We will do something similar in this exercise.

1. Find the average fractions of time that the Markov chain spends in each of the four states, π_1, π_2, π_3, π_4. To do so you may want to simulate the Markov chain as follows:
   (a) Set the initial state at random;
   (b) Draw the next state: when you know the state you start from, say $i \in \{1, 2, 3, 4\}$, the probabilities to go to states 0, 1, . . . are given in row i of the matrix above; you can use CDF inversion for discrete distributions to draw the next state at random;
   (c) Repeat the previous step a large number of times (say $> 10^5$); each time, note down the state you end up in;

(d) Count the number of times that the chain is in states 1, 2, 3, and 4, and divide by the total number of state visits you simulated.

2. Plot the fractions computed in the previous point as a function of the number of steps (do not plot $> 10^5$ points! a subset of that is fine). Compare with the theoretical stationary probability distribution values, which are $\pi_1 = \pi_2 = 0.32$, $\pi_3 = 0.20$ and $\pi_4 = 0.16$.

Now let us use the results above to model a communication system. Assume that the chain above represents in fact the state of a wireless communication channel, where state 1 corresponds to the best channel conditions, and state 4 to the worst ones. In each state, the communication system can achieve the throughput listed in the table below:

State	**Throughput**
State 1	1.5 Gbit/s
State 2	1 Gbit/s
State 3	250 Mbit/s
State 4	50 Mbit/s

3. Estimate the average throughput attained by the communication system over the wireless channel modeled by the Markov chain. Plot the average throughput against the number of state transitions and the corresponding confidence intervals. Does the throughput converge to the values you expected?

Exercise 3

Consider the network scenario of Fig. 2. A source S wants to transmit a packet to destination D. A multihop network separates S from D. Specifically, there are r stages, each of which contains N relays. The source is connected to all nodes of stage 1. Each node of stage 1 is connected to all nodes of stage 2; each node of stage 2 is connected to all nodes of stage 3, and so on. Finally, all nodes of stage r are connected to D. The probability of error over every link in the whole network is equal to p.

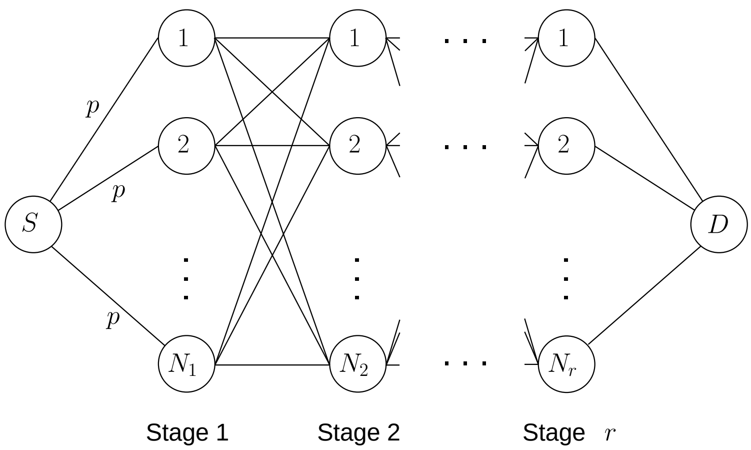

Figure 2: Reference scenario to simulate the flooding of a packet through a multihop network.

S employs a flooding policy to send its packet through the network. This means that every node that receives the packet correctly will re-forward it exactly once. Note that this has implications for the probability that a given relay receives a packet correctly.

For example, at relay stage 1, the probability that any node will fail to receive the packet from S is p. However, say that k nodes at stage i receive the packet correctly: because of the flooding policy, all k nodes will retransmit the packet. Therefore, the probability that a node at stage $i+1$ fails to receive the packet is not p, but rather p^k (i.e., the probability that no transmissions from any of the k relays at stage i are received by the node at stage $i+1$).

1. Use Monte-Carlo simulation to estimate the probability that a packet transmitted by the source S *fails to reach* the destination D. Consider two different cases: $r = 2, N = 2$, and $r = 5$, $N = 10$. For each Monte-Carlo trial, simulate the transmission of the packet by S, the correct or incorrect reception by the

relays at stage 1, the retransmission of the packet towards the next stages, and so forth until the packet reaches D or is lost in the process.
(*Hint*: remember that the probability to fail the reception of a packet is p^k, where k is the number of nodes that hold a copy of the packet at the previous stage.)

2. Repeat the above process for different values of the link error probability p. Plot the probability of error at D against p for the two cases $\{r = 2, N = 2\}$, and $\{r = 5, N = 10\}$. Plot also the confidence intervals for each simulation point.

3. Compare your results against the theoretical error probability curve provided in the file `theory_ex3.csv` (column 1: values of p; column 2: probability of error at D for $\{r = 2, N = 2\}$; column 3: probability of error at D for $\{r = 5, N = 10\}$).

4. Draw conclusions on the behavior of the network for the chosen values of r and N.

5. Plot the average number of successful nodes at each stage, and the corresponding confidence intervals. What can you say about the relationship between the number of successful nodes and the probability of error at D?

Exercise 4

Consider two network nodes i and j. Node i is located at a random position within the rectangle whose bottom-left and top-right corners are located at the coordinates $(0, 0)$ and $(20, 60)$, respectively. Node j is located at a random position within the rectangle whose bottom-left and top-right corners are located at $(60, 0)$ and $(80, 60)$, respectively. Call (x_i, y_i) and (x_j, y_j) the coordinates of nodes i and j, and call d_{ij} the Euclidean distance between the two nodes.

To transmit a packet to node j, node i employs a wireless communication system, having a transmit power $P_T = 5$ W. The propagation of the signal from i to j is subject to a deterministic attenuation and to a random fading process that may attenuate or amplify the received signal power. Under these assumptions, the signal-to-noise ratio (SNR) γ_{ij} characterizing the signal received by j can be written as

$$\gamma_{ij} = \frac{\rho_{ij}\, P_T\, d_{ij}^{-k}}{P_N} \ , \tag{2}$$

where:

- $k = 2$ is the path loss exponent;
- $P_N = 3.2 \cdot 10^{-5}$ W is the noise power;
- ρ_{ij} is the coefficient that describes the power gain due to the fading process; in this exercise, we model ρ_{ij} an exponentially-distributed random variable of average value 1.

We consider the transmission from i to j to be erroneous if the SNR is lower than a threshold θ. We want to compute the probability that the transmission is erroneous.

Formally, we want to compute $p = \mathbb{P}[\gamma_{ij} < \theta]$. Since the positions of i and j are random and the value of the fading gain ρ_{ij} is also random, this probability can be found by averaging over the distribution of these parameters as follows:

$$p = \int_{x_1=0}^{x_i=20} \int_{y_i=0}^{y_i=60} \int_{x_j=60}^{x_j=80} \int_{y_j=0}^{y_j=60} \int_{\rho_{ij}=0}^{\rho_{ij}\to+\infty} \frac{\mathrm{e}^{-\rho_{ij}}}{A_i\, A_j} \mathbb{1}\left[\frac{\rho_{ij}\, P_T\, d_{ij}^{-k}}{P_N} < \theta\right] \mathrm{d}\rho_{ij}\, \mathrm{d}y_j\, \mathrm{d}x_j\, \mathrm{d}y_i\, \mathrm{d}x_i \ , \tag{3}$$

where $A_i = A_j = 20 \times 60$, and $\mathrm{e}^{-\rho_{ij}}/(A_i\, A_j)$ is the joint distribution of fading and of the locations of the two nodes i and j. The function $\mathbb{1}[\,\cdot\,]$ is the "indicator" function, and returns 1 whenever the argument is true. Note: the average value of the indicator function $\mathbb{1}[\,\cdot\,]$ is exactly the probability that the event indicated by the function takes place. The event under examination is "$\gamma_{ij} < \theta$," hence this average computes exactly $\mathbb{P}[\gamma_{ij} < \theta]$, which is equal to p. Initially, you can set $\theta = 10$.

1. Use Monte-Carlo integration to compute p in Equation (3). You may consider these steps:
   (a) Extract the position of nodes i and j uniformly at random in the rectangles mentioned above, and compute the distance d_{ij};

(b) Extract a random fading realization ρ_{ij};

(c) Compute the SNR γ_{ij} from Equation (2), and test if it is lower than θ.

(d) Repeat the test for several (say 1000) different draws of the positions of i and j, and for each position consider at least 50 different fading realizations, and compute p.

(e) Plot the variation of p vs. θ. Take $\theta \in [1, 320]$.

2. Play with the size of the areas where you draw the positions of i and j. Play also with the number of realizations, both of fading and of the node positions. What happens if you decrease these realizations? How many realizations are needed for the results to be insensitive to an increase in the number of realizations?