IERG 6120: Homework 6

Due December 2nd, 2024. Please turn it in at my office SHB708.

Consider a scenario where two wireless sensors send status updates to a base-station. Assume that time is slotted. Due to interference, only one sensor can transmit in each time-slot. If sensor i transmits ($i = 1, 2$), its transmission succeeds with probability p_i, independently of other transmissions. Define the "age-of-information" (AoI) $h_i(t)$ of sensor i at time t as the elapsed time since the last successful transmission before time t. Thus, if a sensor i successfully transmits an update at time-slot t, its AoI is reset to $h_i(t+1) = 1$ in the next time-slot. If a sensor is not transmitting, or its transmission is unsuccessful, its AoI is increased by 1 in the next time-slot, i.e., $h_i(t+1) = h_i(t) + 1$. Our goal is to decide, based on the current AoI of the two sensors, which sensor should transmit at each time-slot, in order to minimize the long-term total discounted AoI across the two sensors, i.e.,

$$\min \mathbf{E} \left[\sum_{t=1}^{\infty} \gamma^{t-1} (h_1(t) + h_2(t)) \right],$$

where $\gamma < 1$ is the discounting factor.

1. Identify the state space, action space, transition probabilities and costs for this MDP.
2. Derive the Bellman optimality equation for this infinite-horizon discounted-cost problem.
3. Write a program that implements the value iteration and finds the optimal policy. Please provide the pseudo-code in your homework. Your simulation should use the following parameters: $p_1 = 0.9, p_2 = 0.5$, and $\gamma = 0.9$. Your program will likely need to restrict the state-space to a finite set. In that case, limit the AoI of each sensor to be within $[1, 100]$. (In other words, if the AoI of a sensor increases to 101, make it 100 instead.)
4. From the numerical results in part 3, verify that the optimal policy is a threshold policy, i.e., given the current AoI of sensor 2, sensor 1 will be scheduled to transmit if and only if its AoI is above a threshold that depends on the AoI of sensor 2. Plot this threshold as a function of the AoI of sensor 2.
5. Assume that $p_1 > p_2$. Prove directly from the Bellman equation that the optimal policy is of the above threshold form. (Hint: prove first that the value function is non-decreasing in each AoI. Then, assume that at a given pair of AoI, sensor 1 is already scheduled to transmit. Show that if the AoI of sensor 1 is increased by 1, it should still be scheduled to transmit.)

6. Change the objective slightly. Suppose that we want to give preference to sensor 1, and hence we define its cost as the *square* of its AoI, while the cost of sensor 2 is still its AoI. In other words, the objective of the MDP changes to

$$\min \mathbf{E}\left[\sum_{t=1}^{\infty} \gamma^{t-1}(h_1^2(t) + h_2(t))\right].$$

Repeat part 3 and part 4, and plot the new thresholds as a function of the AoI of sensor 2. What difference do you see compared to part 4? Can you still prove that a threshold structure holds for the optimal policy? You can assume $p_1 > p_2$. (The opposite case, which is not required for this homework, is also true, but it requires additional effort to be proved.)