OPTIMAL TRAFFIC ADMISSION CONTROL USING STOCHASTIC APPROXIMATIONS

REINFORCEMENT LEARNING – 2022/2023

INSTRUCTOR: F. DE PELLEGRINI

In this TP we shall learn how to perform optimal admission of packets of different classes to be transmitted over a priority channel of finite capacity using stochastic approximations as seen during lectures. You can use any programming language in order to produce the code that runs the algorithms of your task; however, you should implement the required algorithms from scratch.

System description. We consider a communication link which receives as input a stream of packets according to a Poisson process with rate λ packets per second. Packets can be either admitted to a priority queue, or sent in best effort. In particular, packets belong to R classes of jobs, where jobs of class $1 \leq i \leq R$ have a reward r_i. For the sake of notation, we consider $r_i \geq r_{i+1}$ for $i = 1, \dots, R-1$. The probability that a job is of class i is p_i, $\sum p_i = 1$. Upon arrival, jobs can admitted or not for transmission in priority and the objective is to maximize the average reward of priority packets.

Admission control. In order to perform admission control, it is sufficient to consider the state of the system at each time t when a job arrives, that is $s_t = r_t$ where r_t is the reward of the job t. The action set is $A = \{0, 1\}$ where 1 means "admit" and 0 means "do not admit".

You should consider the settings reported in Tab. 1 in order to perform your experiments. You should organize the episodes for the algorithm. One possibility is to use a time window T and sample the system ever T seconds in order to perform the stochastic approximation step using the input collected during the time window.

Task 1: Optimal policy. Determine the optimal stationary threshold policy for the underlying MDP by solving the fractional knapsack problem.

Task 2: Decreasing stepsizes. Implement the version of the Robinson-Monro algorithm tailored for the problem using decreasing stepsizes using the standard stepsize $\epsilon_n = c/n^\gamma$, $\gamma \in (1/2, 1]$ where $c \geq 0$; change the stepsize function, i.e., c and γ in order to trade-off convergence speed for noise rejection. What is important, γ or c in this trade-off? What is the effect of c?

Task 3: Constant stepsizes. Implement the version of the Robinson-Monro algorithm tailored for the problem using constant stepsizes; experiment with different

Table 1. Parameters of the assignment

Variable	*Value*
R: job classes	4
$\mathbf{r}$: job reward	$\mathbf{r} = (100, 50, 10, 1)$
λ: packet arrival rate	$10^5 pkts/s$
c: channel capacity	$0.34 \times 10^5 pkts/s$
$\mathbf{p}$: distribution	$\mathbf{p} = (1/3, 1/12, 1/4, 1/3)$

values of $\epsilon > 0$.
Task 4: Polyak Averages. Use the version developed for the decreasing stepsizes in combination with Polyak averages; show what happens with different window sizes. I leave to you how to choose the window for averaging.
(Optional): ODE. Provide numerical evidence that the sample paths follow the ODE (hint: choose the good timescale).
Note: Please read the text and try to answer all the questions!