IoT final project - smart bins

Consider N smart bins deployed in a urban area.
Assume the maximum capacity of a bin is 100 items. Each bin is equipped with a sensor that says how full it is, in percentage.
For each bin B_i and at each time step, people can throw a maximum of M^i_{max} items. This means that at each time t, the number of items in the bin can increment of $0, 1, \ldots, M^i_{max}$. We can assume that this number follows the uniform distribution, i.e., the probability that at each time step, k items are thrown ($k = 0, 1, \ldots, M^i_{max}$) is $\frac{1}{M^i_{max}+1}$. For each bin, we can define a Markov Chain where the state is the filling percentage of the bin. It holds that, for each bin B_i, given a generic state $s_i = p$ ($p \in \{0, 1, \ldots, 100\}$), the next reachable states are the ones in the following set:

$$S_i := \{s'_i : s'_i = \min\{p + k, 100\}\}$$

where $k = 0, 1, \ldots, M^i_{max}$, each reachable with probability $\frac{1}{M^i_{max}+1}$.
Bins publish their results via MQTT in a topic ".../{bin_ID}/fill_level". A central station subscribes to the fill_level topic and reads the data published by the bins.

Project 1 Consider $N = 5$, $M^i_{max} = 5$, $\forall i = 1, \ldots, 5$. Consider a Markov Decision Process where a state is given by the state (i.e., the fill percentage) of each bin, as in the Markov chain described above, i.e., $s = (s_1, \ldots, s_5)$. As the central station reads the data published by the bins, it can decide to take one of the six possible actions: $\mathcal{A} := \{\text{wait, collect trash from } B_i, \forall i = 1, \ldots, 5\}$. Collecting trash comes with a cost $c = 1$, while waiting has no cost. Every time a bin fill level is beyond 95%, the reward is -5, otherwise it is 0. Hence, the reward associated with taking an action a and reaching a state s' is given by the sum of the rewards of each state s'_i, minus the cost of the action. Every time the trash in bin is collected, its fill level goes to 0. Simulate the environment with MQTT as a communication protocol. Use any algorithm to find the optimal policy π to optimize bin collection and show your results. (The space state is huge and it is taking forever to solve? Consider discretising the fill level, which can assume values in $\{0\%, 10\%, 20\%, \ldots, 100\%\}$.)

Project 2 Define the cost of traveling between two smart bins B_i and B_j as $c_{ij} = c_{ji}$. At each time step, as it reads the data published by the bins, the central station decides whether to send workers to collect the trash or not. The central station can decide to send workers to collect just from some bins. Let N_t the number of trash bins that the central station decides to collect at time t, $N_t = 0, 1, \ldots, N$. Simulate the communication with MQTT, and design an heuristic to decide the trash bins to collect at each time step t, avoiding bins to get full, and minimising the travel cost. Every time the trash in bin is collected, its fill level goes to 0. Show your results. To minimise the travel cost, you might use the greedy algorithm for the Travelling Salesman Problem (`https://www.geeksforgeeks.org/travelling-salesman-problem-greedy-approach/`). Notice: you can assume that the system does not change while you are collecting the trash bin.