Dynamic Staffing Control for a Triage-Based Service System

Real-Life Problem

In 2050, humans develop a new autonomous technology capable of treating patients with any illness. However, it is costly to activate and maintain these unit, and only 1 new unit can be activated each hour. The Emergency Department is trying to decide the use of active unit to adapt to the current situation. The objective here is finding the optimal policy to minimizing the cost while serving patients.

Model Overview

Description and Explaination

We consider a service system (e.g., an Emergency Department) with Poisson arrivals of rate μ. Each arrival is assigned a triage level $i \in \{1, 2, 3\}$ with probabilities

$$p_1 = 0.1, \qquad p_2 = 0.3, \qquad p_3 = 0.6,$$

corresponding to Critical, Urgent, and Non-urgent.

The queue can hold at most M patients; arrivals have to leave if it exceeds the queue capacity.

Each level-i patient has an exponential service time with rate λ_i. Calling each unit a staff, if a staff becomes free, it immediately treats the highest-acuity waiting patient.

Let S_t denote the number of active staff at time t.

For each hour interval t, we have:

- Services time for each patient follows exponential distribution
  $\Rightarrow$ we only need to consider the number of currently served patient for each triage level (Memoryless).
- Interarrival time follows exponential distribution with rate μ

Therefore, we only need to consider the number of patients in each triage in queue and service, and then simulate the system.

Optimization

In order to derive the optimal policy for this model, we use a simulation-based Reinforcement learning. Here, we will describe the model, ultilizing the simulation above.

Dynamic Staffing Decisions

Let S_t denote the number of active staff at time t. Then, the controller chooses an action

$$a_t \in \{-S_t, \dots, -1, 0, +1\},$$

representing a request to decrease or increase staff. Staffing is bounded:

$$0 \leq S_t \leq S_{\max}.$$

Increases take effect immediately. Decreases will only be effective when staff finish service; if no staff become free, staffing remains unchanged. The next staffing level is

$$S_{t+1} = \min\{S_{\max},\, \max\{0,\, S_t + a_t^{\text{effective}}\}\}.$$

Reward Structure

Let $C_{i,t}$ denote the number of level-i patients complete treatment during t. During interval t, we have:

- Complete treatment for a level-i patient yields reward $r_1 > r_2 > r_3 > 0$.
- Maintain S_t staff cost $c_m S_t$.
- Activate new staff cost c_a.
- Block P_t patient receive $c_p P_{\text{ovf},t}$ penalty.

Then, the reward will be:

$$R_t = -c_m S_t - c_a \mathbf{1}_{\text{new activation}} - c_p P_{\text{ovf},t} + \sum_{i=1}^{3} r_i C_{i,t}$$

We will calculate this using the simulation above.

MDP Formulation

Because arrivals are Poisson and service times are exponential, the system is Markov once we track, for each triage level $i \in \{1, 2, 3\}$:

$$\begin{aligned} Q_{i,t} &= \text{number of level-}i\text{ patients waiting in queue at time }t, \\ B_{i,t} &= \text{number of level-}i\text{ patients in service at time }t. \end{aligned}$$

These satisfy

$$\sum_{i=1}^{3} Q_{i,t} \le M, \qquad \sum_{i=1}^{3} B_{i,t} \le S_t.$$

Thus the state is

$$X_t = (Q_{1,t}, Q_{2,t}, Q_{3,t},\ B_{1,t}, B_{2,t}, B_{3,t},\ S_t).$$

State transitions are determined by arrivals, service completions, capacity limits, and the priority service rule (highest-acuity first).