Equity in Access to Care: Appointment Scheduling across Patient Groups with Varied No-Show Rates

Background and Motivation. The issue of equity in access to healthcare services has gained significant attention. For example, a recent study reveals a significant diagnostic delay disparity in amyotrophic lateral sclerosis (ALS) patients, with black patients experiencing a 24-month delay compared to 16 months for white patients (Raymond et al., 2024). The disparity could be attributed to algorithmic biases due to a higher prevalence of ALS among the white population. Additionally, there could be scheduling preferences for avoiding appointment no-shows, which are found to be relevant to ethnic, racial, and socioeconomic characteristics (Samorani et al., 2022). Our study aims to design appointment scheduling (AS) schemes that address equity issues and evaluate how patient-group information affects both schedule efficiency and equity in access to medical appointment services (i.e., fairness). We consider patient heterogeneity characterized by their no-show behaviors. It is well known in the literature that patient demographics and socioeconomic status may influence no-show rates (Cayirli & Veral, 2003). This variability in no-show behaviors complicates the efficient allocation of time and resources, particularly when the clinic must strategically reserve some of the mid-day time slots for the possibility of re-entrance. In our problem, appointment requests come in gradually, requiring the scheduler to account for stochastic re-entrance and patient group-dependent no-show rates. While studies in the literature typically focus on the fairness in the outcome, we propose AS schemes that come with different levels of procedural fairness, and each proposed scheme leads to a different optimization problem. We investigate how different types of scheduling fairness influence schedule efficiency.

Problem and Model Description. We consider an AS problem consisting of the appointment process and the following service process (Feng et al., 2023). During the service process of a given clinic session, several randomnesses are considered: patients may show up or not show up for their appointments, service time is stochastic, and the need to re-enter the service after examination is also stochastic. During the appointment process for a given service session, appointment requests from different patient groups arrive gradually as a stochastic process. As a result, the schedule is built up gradually. At each decision epoch, the scheduler must decide:

1. Whether to accept the request:
    - If yes, the scheduler determines the appointment time and updates the schedule.

- If no, the request is re-direct to another service session (on a different date or time) that is still open for appointment.

2. Whether to close the schedule book for the session.

We propose three different schemes based on how patient-group information, related to the no-show rate, is revealed and utilized:

- Group-Unaware Scheme (GUS): In this scheme, the scheduler has no knowledge of the no-show group information. Decisions are made based solely on aggregated information from the overall population.
- Group-Aware Scheme (GAS): Here, the scheduler is fully aware of the no-show group information for each incoming request, as well as for all past requests within the same service session.
- Semi-Aware Scheme (SAS): In this case, the scheduler uses the no-show group information from the existing schedule when making decisions regarding the current request, but does not have information about the current request's no-show group when deciding whether to accept the request.

The AS problem is modeled as a finite-horizon Markov Decision Process (MDP), where the reward is defined as the utility derived from serving the patients on the schedule, minus the disutility incurred from factors such as wait times, overtime penalties, and the loss of goodwill associated with each redirection request.

Methodology and Results. In theory, the MDP for each scheme can be solved using standard dynamic programming. However, due to the curse of dimensionality, such a solution becomes computationally infeasible for problems of practical size. Therefore, we turn to reinforcement learning (RL). Specifically, we design a simulation-based actor-critic framework for all three schemes. To ensure proper training of the RL model, we first compare the solutions obtained through dynamic programming with those derived from RL for smaller problem sizes. The comparisons show that, for these smaller problem sizes, the solution quality of the RL model closely matches that obtained via dynamic programming.

Subsequently, we use RL to conduct numerical studies on problems of practical scale. In these studies, we generate a large number of request sequences following the estimated mixture probabilities of patient groups. To evaluate our decision-making policies and compare the performance of different schemes, we compute the average reward for each scheme's policy across the same set of generated request sequences. Our results demonstrate that GAS consistently outperforms the other two schemes. This outcome is quite intuitive, as GAS fully leverages patient category information, enabling more forward-looking scheduling decisions.

Perhaps more intriguing is that we identify specific operational environments—characterized by model parameters—in which SAS performs exceptionally well. In these environments, we achieve high operational efficiency by utilizing only the information from the existing schedule when making decisions regarding upcoming requests without needing the group identity of those requests. This finding suggests that, under certain conditions, SAS offers a competitive alternative to GAS. An important implication of this result is that we can achieve both efficiency and perceived fairness simultaneously. By reducing the reliance on patient group information, SAS minimizes the potential for biases while still optimizing the operational efficiency. Moreover, with increased system-wide productivity, SAS potentially allows for accommodating more patients in each session, thus enhancing overall access to care for all groups. This highlights the value of developing scheduling policies that not only improve operational efficiency but also promote equitable access to healthcare services.

References

Cayirli, T., & Veral, E. (2003). Outpatient scheduling in health care: A review of literature. *Production and Operations Management*, *12*(4), 519–549.

Feng, H., Alvarado, M., Konda, S., & Lawley, M. (2023). Sequential clinical scheduling with stochastic patient re-entrance: Case of mohs micrographic surgery. *Computers & Industrial Engineering*, *184*, 109589.

Raymond, J., Nair, T., Gwathmey, K. G., Larson, T., Horton, D. K., & Mehta, P. (2024). Racial disparities in the diagnosis and prognosis of als patients in the united states. *Journal of Racial and Ethnic Health Disparities*, 1–7.

Samorani, M., Harris, S. L., Blount, L. G., Lu, H., & Santoro, M. A. (2022). Overbooked and overlooked: Machine learning and racial bias in medical appointment scheduling. *Manufacturing & Service Operations Management*, *24*(6), 2825–2842.