THE OPEN UNIVERSITY OF SRI LANKA

FACULTY OF ENGINEERING TECHNOLOGY

THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

EEX5362 - Performance Modeling

Mini Project Report

A.M. Fathima Humaira

s22009993

222513186

2024/2025

Performance Modeling and Evaluation of a Monthly Outpatient Clinic Queue in a Government Hospital

Table of Contents

1. System Description and Performance Goals

This study focuses on the monthly outpatient clinic process in a government hospital. Although clinic dates are announced in advance during the previous month's clinic, patients continue to arrive very early on the clinic day to obtain tokens. Patients typically begin arriving as early as 6:00 a.m., while the doctor consultation session officially runs from 9:00 a.m. to 1:00 p.m.

During this period, patients wait in a queue and are seen by available doctors on a first-come-first-served basis. Each patient consultation takes between 3 to 10 minutes, depending on case complexity. Due to early arrivals, high patient volume, and limited doctor availability, long waiting times and congestion are frequently observed.

The primary performance goal of this system is to improve patient experience by reducing excessive waiting times.

2. Performance Objectives

The main **performance objective** of this study is:

To minimize the average patient waiting time between arrival at the clinic and the start of consultation.

To support this objective, the following secondary objectives are considered:

- Identify bottlenecks contributing to long waiting times
- Evaluate the impact of doctor availability on system performance
- Analyze congestion caused by early token-based arrivals
- Propose operational improvements to reduce queue buildup

These objectives provide measurable criteria for evaluating system performance using waiting time, queue length, and doctor utilization metrics.

3. Modeling Approach and Assumptions

Modeling Technique

The outpatient clinic was modeled as an M/M/c queuing system, where:

- Patient arrivals follow a Poisson process
- Service times follow an exponential distribution
- *c* represents the number of doctors (servers)

A discrete-event simulation approach was used to analyze system behavior over time. Simulation was selected because real-world clinic operations involve randomness and time-dependent congestion that are difficult to capture using purely analytical models.

Key Assumptions

- Patients arrive independently
- Early token collection creates bulk arrivals before 9:00 a.m.
- Doctors serve one patient at a time
- Service discipline is first-come-first-served
- No patient abandons the queue once a token is obtained

4. Data Description and Methodology

This study uses simulated data informed by direct observation and informal interviews with patients at a monthly government hospital outpatient clinic. Since detailed real-world queue records such as exact arrival times, waiting durations, and service times are not formally documented, a discrete-event simulation based on an M/M/c queuing model was implemented using Python to replicate the observed system behavior.

Patient arrival patterns were modeled using a Poisson process starting from 6:00 AM, reflecting early token-based arrivals reported by patients and observed during the clinic days. Doctors begin consultations at 9:00 AM and continue until 1:00 PM. Service times were randomly generated between 3 and 10 minutes per patient, based on patient feedback and observed consultation durations.

The simulation captures key performance metrics including waiting time, queue length over time, number of patients served, and doctor utilization. Multiple scenarios with varying numbers of doctors (up to 5) were simulated to evaluate how staffing levels affect clinic performance.

The generated data supports the analysis by realistically reflecting real-world congestion patterns and enabling meaningful comparisons across staffing scenarios, thereby helping identify operational bottlenecks and potential improvement strategies

5. Detailed Analysis and Findings

The simulation results indicate that early token-based patient arrivals cause a large number of patients to gather at the clinic well before consultations begin. This creates long queues from the early morning hours, even though doctors start seeing patients only at 9:00 a.m. As a result, patients experience significant waiting times before their consultations start.

When the number of patients arriving exceeds the service capacity of the available doctors, the system becomes congested. Under lower staffing levels, this leads to very high average waiting times, doctors being busy for most of the clinic hours, and queues that persist throughout the entire session. This clearly shows that the clinic is unable to cope with demand when staffing levels are insufficient.

Increasing the number of doctors improves system performance by reducing patient waiting times and shortening queue lengths. However, the results also show that adding more doctors alone does not fully solve the problem, as early mass arrivals continue to create congestion.

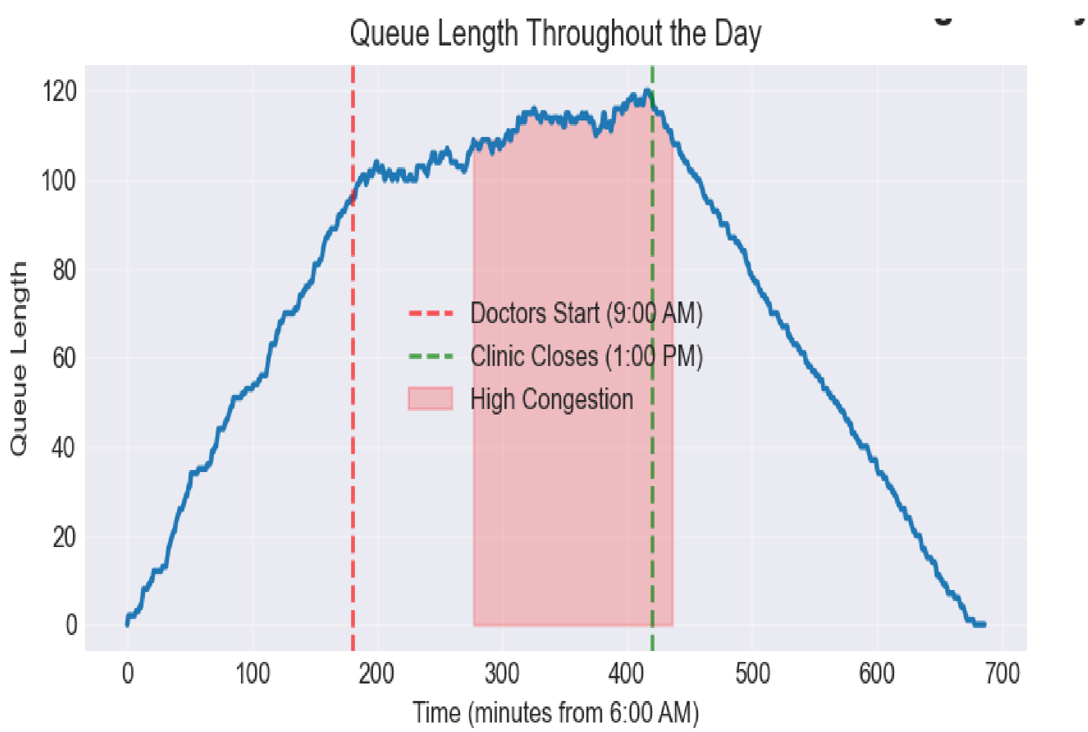
Overall, the findings highlight that in addition to increasing service resources such as doctors, effective patient arrival management methods including appointment scheduling or staggered token distribution are necessary to balance demand and significantly reduce patient waiting times.

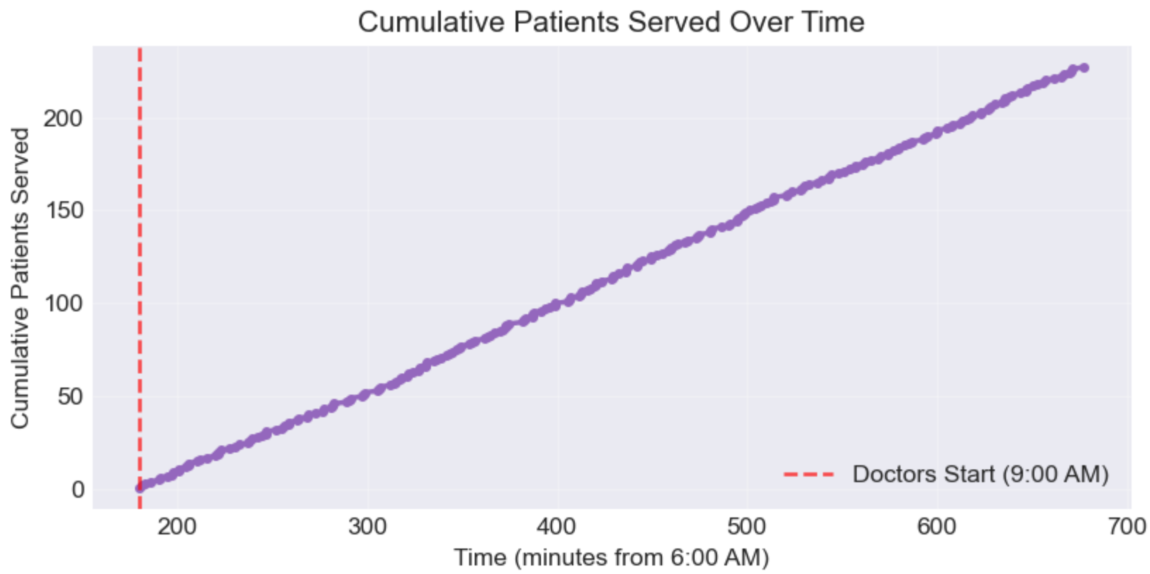
6. Visualizations

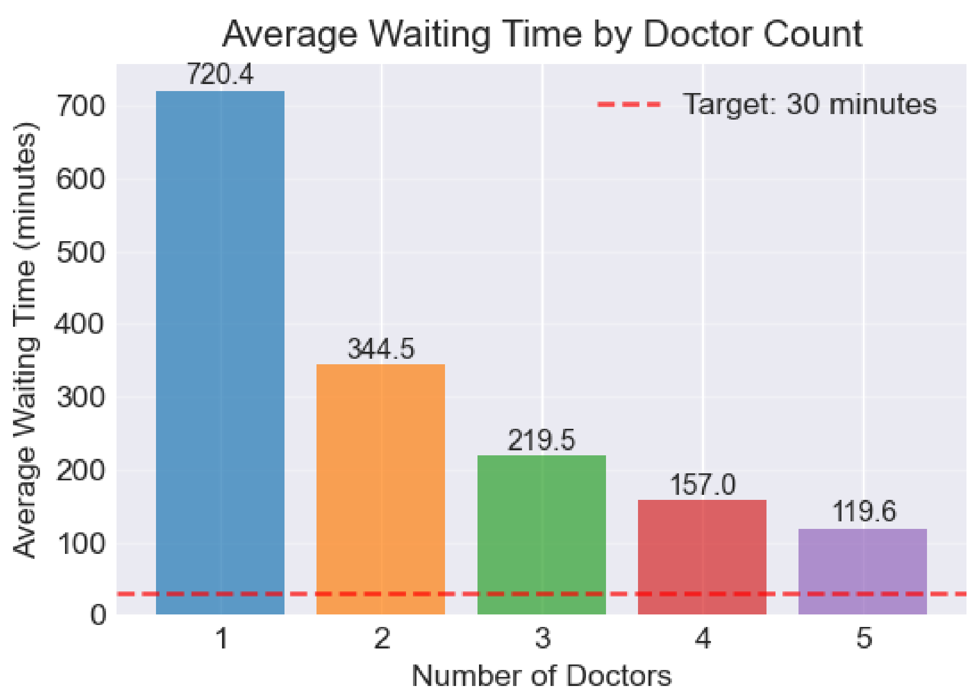
The analysis includes:

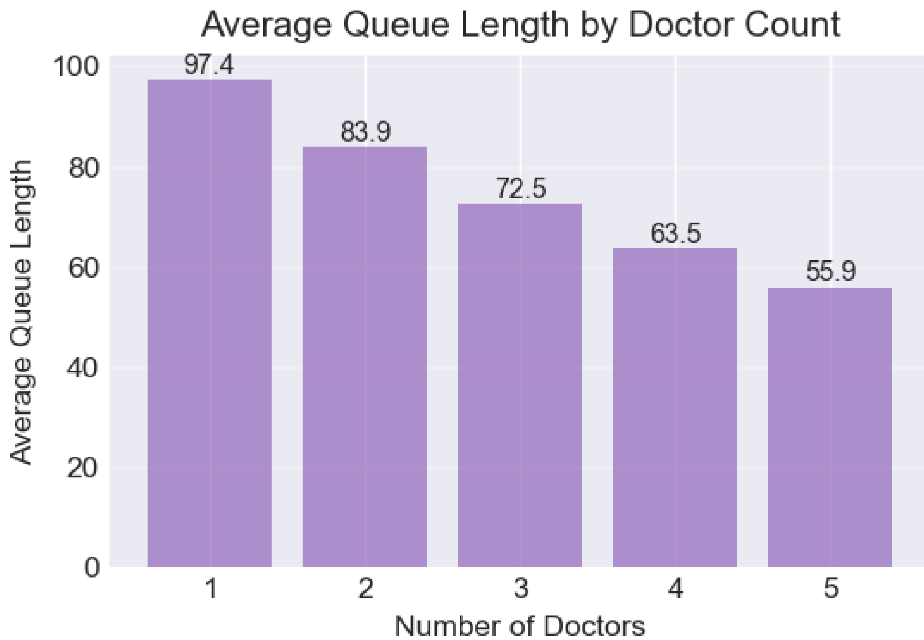
- Queue length versus time graphs showing congestion patterns
- Comparative charts illustrating waiting times under different doctor configurations

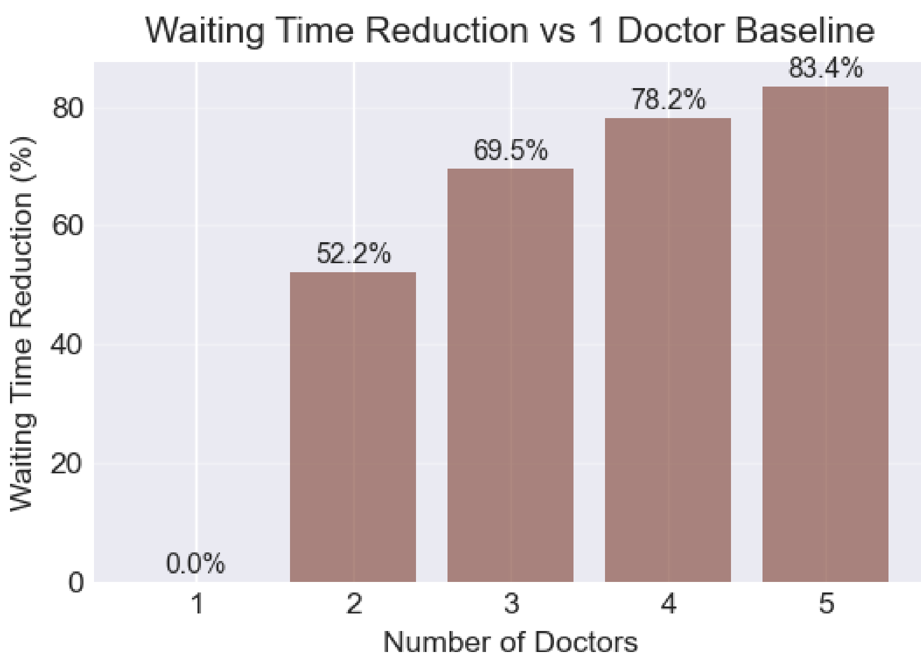
These visualizations clearly demonstrate how early arrivals and limited capacity affect system performance.




Cumulative Patients Served Over Time
Cumulative Patients Served
0
50
100
150
200
200
300
400
500
600
700
Time (minutes from 6:00 AM)
Doctors Start (9:00 AM)


Average Waiting Time by Doctor Count
Average Waiting Time (minutes)
0
100
200
300
400
500
600
700
720.4
344.5
219.5
157.0
119.6
1
2
3
4
5
Number of Doctors
Target: 30 minutes


Average Queue Length by Doctor Count
97.4
83.9
72.5
63.5
55.9
Average Queue Length
Number of Doctors


Waiting Time Reduction vs 1 Doctor Baseline
0.0%
52.2%
69.5%
78.2%
83.4%
Waiting Time Reduction (%)
Number of Doctors

7. Limitations and Future Extensions

Limitations

- The study relies on simulated data rather than complete real-world queue records, which may not capture all nuances.
- Some simplifying assumptions like patients arriving independently, no one leaving the queue early, and strict first-come-first-served might differ from real patient behavior.
- Service times are randomly generated between 3-10 minutes, which may not fully reflect very complex or unusually short consultations.
- The doctor schedule is fixed, ignoring any variations that might occur in actual clinic operations.
- Early morning bulk arrivals are modeled simply and may not fully capture the range of patient arrival patterns.

Future Extensions

- Include real-time clinic data to make simulations more accurate and reflective of actual operations.
- Explore arrival management strategies such as staggered token distribution, pre-booked appointments, or priority queuing to reduce congestion.
- Test flexible doctor schedules and include other staff like nurses or administrative personnel to understand their effect on patient flow.
- Use more realistic service time distributions and consider patient behaviors such as no-shows or late arrivals.
- Perform a cost-benefit analysis to determine the best combination of staffing and operational improvements for better patient experience.

8. Reference

[1] GeeksforGeeks, *Queuing Models in Operating System*, GeeksforGeeks. [Online]. Available: https://www.geeksforgeeks.org/operating-systems/queuing-models-in-operating-system/

[2] ILS-02, *Performance Modeling*, Lecture slides, EEX5362, The Open University of Sri Lanka.