THE OPEN UNIVERSITY OF SRILANKA

FACULTY OF ENGINEERING TECHNOLOGY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

BACHELOR OF SOFTWARE ENGINEERING HONOURS

EEX5362

Performance Modelling

Mini Project

Performance Modeling and Evaluation of a Telemedicine Call Center System

Student Name:

Edwin Sanjeeveni Aishwarya Devi

Student SID:

S92068920

Registration No:

421428920

Repository Link:
[https://github.com/sanjeevi83/Telemedicine_Call_Center_MiniProject.git]

1 SYSTEM DESCRIPTION AND PERFORMANCE GOALS

This mini project presents a performance modeling and evaluation study of a telemedicine call center system. The system was initially identified and described in Deliverable-01, where the problem context and dataset were introduced. Telemedicine services allow patients to consult healthcare professionals remotely using telephone communication, reducing physical hospital visits and improving access to medical care.

In the telemedicine call center, patient calls arrive continuously throughout the operating period. Each incoming call is handled by a medical support agent. Since the number of available agents is limited, not all calls can be served immediately. When all agents are busy, incoming calls are placed in a waiting queue until an agent becomes available.

System performance is particularly critical in healthcare environments. Excessive waiting times may reduce patient satisfaction and may pose risks for patients with urgent medical needs. Therefore, evaluating system performance and identifying potential bottlenecks is essential for ensuring efficient and reliable service delivery.

2 PERFORMANCE OBJECTIVES

The primary objective of this mini project is to evaluate the operational performance of the telemedicine call center using analytical performance modeling techniques. The specific objectives are:

- To determine the average waiting time experienced by patients
- To analyze the utilization of service agents
- To evaluate the system's ability to handle incoming call demand
- To identify performance bottlenecks during high arrival periods

Waiting time is an important performance metric because long delays may discourage patients from using telemedicine services. Agent utilization indicates how efficiently healthcare resources are being used, while throughput reflects the system's overall service capacity.

3 MODELING APPROACH AND ASSUMPTIONS

Queueing theory was selected as the modeling approach because it is well suited for systems characterized by random arrivals and limited service capacity. The telemedicine call center was modeled using an M/M/c queueing model, where multiple agents serve incoming calls in parallel.

In this model

- The arrival rate (λ) represents the average number of calls arriving per minute.
- The service rate (μ) represents the average number of calls handled by one agent per minute.
- The parameter c represents the number of available service agents.

The following assumptions were made to simplify the analysis:

- Call arrivals occur randomly and independently
- Service times follow an exponential distribution
- Calls are served on a first-come, first-served basis
- Call abandonment and priority handling are not considered

These assumptions simplify the analysis while still capturing the essential behavior of the system.

4 DATA DESCRIPTION AND METHODOLOGY

The dataset used in this study consists of telemedicine call records stored in CSV format. Each row in the dataset represents an individual patient call and includes the arrival time of the call, the service duration, and the identifier of the handling agent.

Arrival times are measured in minutes from the beginning of the observation period. Service time represents the duration required to complete each call. The dataset contains 30 call records collected over a total observation period of approximately 58 minutes.

4.1. PARAMETER ESTIMATION

The arrival rate (λ) was calculated by dividing the total number of calls by the total observation time:

$$\boldsymbol{\lambda = \frac{30}{57.9} = 0.52\ \mathit{call\ per\ minute}}$$

The average service time calculated from the dataset is approximately 7.41 minutes per call. The service rate (μ) was obtained as the reciprocal of the average service time:

$$\mu = \frac{1}{7.41} = 0.135\ \mathit{calls\ per\ minute\ per\ agent}$$

For performance evaluation, the system was analyzed, assuming seven service agents (c = 7). These estimated parameters were used as inputs for the M/M/c queueing model.

5 PERFORMANCE ANALYSIS AND RESULTS

Using the estimated arrival and service rates, the performance of the telemedicine call center was evaluated analytically.

5.1. PERFORMANCE METRICS

The calculated performance metrics are summarized in Table 1

Summary of Performance Metrics

Metric	**Value**
Arrival Rate (λ)	0.52 calls/min
Service Rate (μ)	0.135 calls/min/agent
Number of Agents (c)	7
Utilization (ρ)	0.552
Average Waiting Time (Wq)	0.276 minutes

System utilization was calculated

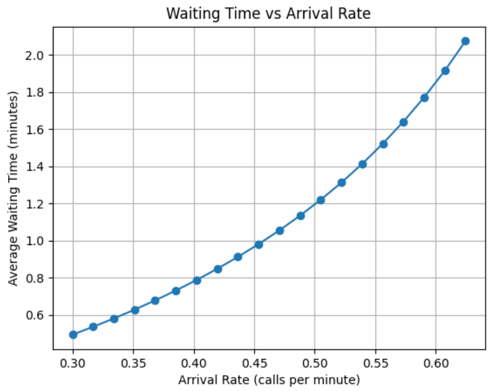
$$\rho = \frac{\lambda}{c\mu} = \frac{0.52}{7 \text{ x } 0.135} = 0.552$$

This result indicates that service agents are busy approximately 55.2% of the time, suggesting that the system operates under stable conditions.

The average waiting time in the queue (Wq), calculated using Erlang-C formulas, is approximately 0.276 minutes. This indicates that under the observed workload, patients experience minimal waiting delays.

6 VISUAL ANALYSIS

A performance diagram was generated using Python to analyze how the system behaves under increasing load. The graph plots arrival rate versus average waiting time while keeping service capacity constant.



- Very low waiting times at low arrival rates
- Rapid growth in waiting time as arrival rates increase
- Clear congestion behavior as system utilization approaches capacity

This performance diagram effectively represents the system's operational behavior and highlights the importance of capacity planning in telemedicine services.

7 PERFORMANCE IMPROVEMENT STRATEGIES

Based on the analysis, the following improvements are recommended

- Increasing the number of agents during peak hours
- Introducing priority handling for urgent medical calls
- Implementing scheduled callbacks to smooth demand
- Using predictive staffing based on historical call patterns

These strategies can significantly reduce waiting times and improve patient satisfaction.

8 LIMITATIONS AND FUTURE WORK

This study has several limitations. The M/M/c model assumes exponential service times and does not consider call abandonment or variations in call complexity. The dataset represents a limited observation period and may not capture long-term demand patterns.

Future work may include simulation-based analysis, priority queue models, abandonment modeling, and evaluation using larger real-world datasets.

9 REFERENCES

1 Internet Archive. (2022). *Fundamentals of queueing theory : Gross, Donald : Free Download, Borrow, and Streaming : Internet Archive*. [online] Available at: https://archive.org/details/fundamentalsofqu0000gros/page/n5/mode/2up.

2 Internet Archive. (2021). *The art of computer systems performance analysis : techniques for experimental design, measurement, simulation, and modeling : Jain, Raj : Free Download, Borrow, and Streaming : Internet Archive*. [online] Available at: https://archive.org/details/artofcomputersys0000jain/page/n3/mode/2up? [Accessed 14 Dec. 2025].

3 Green, L. (n.d.). *Chapter 11 QUEUEING ANALYSIS IN HEALTHCARE*. [online] Available at: https://business.columbia.edu/sites/default/files-efs/pubfiles/4386/chapter%2011%20QueueingAnalysis.pdf?.

10 ANALYSIS CODE STRUCTURE AND EXPLANATION

The performance evaluation was implemented using Python and organized into modular files.

- **data_loader.py**: Loads the CSV dataset
- **arrival_rate_calculation.py**: Computes the arrival rate
- **service_rate_calculation.py**: Computes the service rate
- **queue_metrics_mm_c.py**: Calculates utilization and waiting time using Erlang-C formulas
- **performance_plots.py**: Generates performance diagrams
- **main_analysis.py**: Integrates all modules and executes the analysis