International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

Mathematical Models and Algorithms of Patient Flow Management in Medical Institutions

N. I Sagidullaev[1], J. A. Orazbaev[2]

[1]PhD student at TUIT named after Muhammad al-Khwarizmi
[1,2]Assistant at Nukus State Technical University

***Abstract*— Efficient management of patient flow is one of the most critical challenges in modern healthcare systems, as increasing demand, limited medical resources, and uneven service loads often result in long waiting times and reduced service quality. This study presents a comprehensive mathematical and algorithmic approach to optimizing patient flow in medical institutions. A queueing theory–based M/M/m model is used to describe patient arrivals through a Poisson process and physician service times via exponential distributions. Dynamic queue boundaries are introduced to prevent uncontrolled queue growth under high load and ensure balanced workload distribution. An optimal patient allocation algorithm based on real-time physician utilization identifies the least-loaded doctor and automatically redistributes incoming patients. Additionally, a Minimax-based scheduling framework ensures balanced task distribution and minimizes the maximum service time across physicians. Simulation modeling, including discrete-event and agent-based approaches, is employed to evaluate system performance under varying load conditions. Results from clinical data demonstrate significant improvements: average waiting time reduced by 54%, maximum queue length decreased by 62%, workload variance dropped by 68%, and patient satisfaction increased from 68% to 89%. The proposed models and algorithms provide an effective and scalable solution for healthcare organizations aiming to enhance operational efficiency, reduce waiting times, and improve overall service quality**

***Keywords*— Patient flow management, queueing theory, Poisson process, service rate, system load, dynamic queue model, physician workload, scheduling optimization, simulation modeling, healthcare operations.**

I. INTRODUCTION

The flow of patients in modern medical institutions is increasing year by year. This process leads to long queues in polyclinics and specialized centers, overloading of doctors, deterioration of the quality of service, and a decrease in the overall level of patient satisfaction.

Therefore, mathematical modeling of the service process in medical institutions, the use of mass service systems, real-time load management, automated scheduling, and the creation of intelligent distribution algorithms are among the most pressing issues in modern medical management.

Medical Theoretical model of the service process. The patient admission process in a polyclinic usually consists of three main components:

1. random stream of visitors,
2. limited service resources,
3. queues and waiting processes.

These factors are mathematically represented by a public service system (Queueing System – QS) [1,4,6].

Mathematical model of visitor flow. In most medical facilities, patient flow closely follows a Poisson process:

$$\lambda = \frac{1}{E[T_k]}$$

here

λ - patient arrival intensity .

$E[T_k]$- Average patient arrival interval.

k - an index indicating the number of patient visits

This model practically means that load management is important for stable system operation during high load days.

Model of service speed and physician workload. The service speed of a physician is expressed as:

$$\mu = \frac{1}{E[T_x]}$$

μ— average duration of admission.

$E[T_x]$Time spent seeing 1 patient.

x - index indicating the number of patients

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

System load . A multi-doctor polyclinic is described by the M/M/m model:

$$\rho = \frac{\lambda}{m\mu}$$

ρ— system load .

m— number of service channels (number of doctors)

If $\rho > 1$:

- the system is overloaded,
- The queue grows endlessly,
- the quality of service decreases.

Practical research shows that:

- → 0.7 is normal,
- → 0.85 satisfactory,
- → 1 critical load,
- 1 emergency.

In traditional QS theory, queue length is assumed to be static. However, for medical facilities, a real-time dynamic queue model is more effective [7, 8, 9, 10].

Dynamic queue boundaries.

At minimum load:

$$L_{min} = \alpha \cdot m$$

At high load, the queue limit is limited:

$$L_{max} = \beta \cdot \frac{1}{1-\rho}$$

here:

L_{min}- minimum queue limit

α- load growth rate

L_{max}- r is the maximum allowed queue length

β- dynamic coefficient of variation related to the stability of the queue system

Through the dynamic queuing model:

- each doctor is managed according to his/her workload,
- when the system is overloaded, patients are referred to another doctor,
- uncontrolled growth of queues is stopped.

Algorithm for optimal allocation of patients to doctors. The distribution algorithm consists of three steps:

1. Assessing physician workload

$$U_i = \frac{\lambda_i}{\mu_i}$$

U_i- doctor's workload.

It is calculated individually for each doctor.

2. Identify the busiest doctor

$$i^* = arg\min_i(U_i)$$

3. Placing the newly arrived patient in the optimal channel. If $U_1, U_2, \ldots, U_m$the smallest of these is detected, the patient is referred to this doctor.

If the loading is high (ρ> 0.9):

- the queue is temporarily closed,
- the patient will be referred to another department or offered a re-registration.

A model based on scheduling theory. The medical facility is modeled as follows:

- “patient” → work,
- “doctor” → car,
- “acceptance” → deadline for completion of work.

The main rule is the Minimax principle:

$$T_{final} = \min\max(T_i)$$

The most effective planning rules:

- The longest appointment time is for the busiest doctor.
- Small reception times are distributed according to the queuing strategy.
- Limits are placed on the number of admissions to avoid overloading the doctor.

Patient flow simulation model. To assess the effectiveness of patient flow management in a medical facility, it is necessary to create a simulation model outside of the real system [2,3,5]. Simulation allows for a preliminary assessment of the system's performance under various load conditions.

The following simulation approaches are typically used:

– Discrete-time simulation. In this approach, the time course of a system is divided into small intervals. At each step, the following processes are observed:

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

- whether a new patient arrives or not (according to Poisson),
- the doctor's busy or free status,
- end of reception,
- change in queue length.

- Discrete simulation has the following state map:

- Arrival process — random(λ)
- Service process — random(μ)
- Queue management — dynamic constraint
- Redistribution — referral to the freest doctor
- Total system load — $\rho(t)$

- Agent-based simulation. In this approach:

- every patient is an agent,
- Every doctor is considered a resource.

Each agent has its own status:

- "came"
- "in turn"
- "in reception"
- "it came out"

Doctors manage:

- available load,
- queue number,
- real-time relaxation.

- Simulation results. The following is calculated through simulation:

1) Average waiting time

$$W_q = \frac{L_q}{\lambda}$$

L_q- average number of patients in the queue

2) Average queue length

$$L_q = f(\rho, m)$$

3) Employment rate of doctors

$$B_i = \frac{total\ reception\ time}{working\ hours}$$

4) Service Quality (KPI)

- "waiting < 15 minutes" indicator,
- "patient satisfaction level".

Analysis based on clinical examples and real data . The analysis is based on statistical data obtained at the Multidisciplinary Medical Center of the Republic of Karakalpakstan (in general terms)[2].

- Current density
- Number of visitors per day: 450–700 patients
- "Peak hours": 09:00–12:00
- Minimum flow: 14:00–16:00
- Doctors' workload
- Therapists: 0.92–0.98 (critical)
- Pediatricians: 0.73–0.85 (average)
- Surgeons: 0.35–0.45 (low)

This shows that:

- additional resources are needed for therapists,
- and surgeons are not optimally loaded.
- The effectiveness of the dynamic queuing model

Indicator	Traditional system	Dynamic model	Difference
Average waiting time	28 minutes	13 minutes	−54%
Maximum queue length	37 patients	14 patients	−62%
Physician workload variance	0.22	0.07	−68%
Patient satisfaction	68%	89%	+21%

These results prove the effectiveness of the proposed model. Table 1.

II. CONCLUSION AND SCIENTIFIC RESULTS

In this scientific study, the issue of optimizing patient flow in medical institutions was comprehensively studied based on mathematical models, algorithms, and software tools. The analyses, models, and simulations conducted allowed us to draw the following scientific and practical conclusions.

The practical effectiveness of the proposed models is as follows:

- Queues are reduced by 40–60%
- The workload of doctors is distributed evenly
- Average patient wait time drops from 28 minutes to 13 minutes

- Congestion occurring in the system during "peak hours" will be eliminated
- Population satisfaction level increases by 21%
- The capacity of the medical center will increase

Socio-economic impact of the system

- The number of patients dissatisfied with medical services is decreasing
- Excessive human factor (administrator error) is reduced
- The acceptance time is clear and transparent.
- Medical center labor productivity increases
- Used as a universal model for public and private clinics

REFERENCES

[1] Оплюшкина О. "Модели и алгоритмы управления обслуживанием пациентов медицинского учреждения". Penza, 2020.

[2] Sagidullaev NI "Automation of the activities of the Republican Multidisciplinary Medical Center". EJMTCS, 2021.

[3] Uteuliev NU, Dzhaikov GM, Sagidullaev NI "Automation of medical institution management and creation of a mathematical model". APAMIT-2024.

[4] Gross, D., Shortle, JF "Fundamentals of Queuing Theory". Wiley, 2018.

[5] Peterson, J. "Petri Net Theory and the Modeling of Systems." Prentice-Hall, 1981.

[6] Kleinrock L. Queuing Systems, Volume 1: Theory. Wiley-Interscience, 1975.

[7] Banks, J. "Discrete Event Simulation". Wiley, 2014.

[8] Hillier, F., Lieberman G. "Introduction to Operations Research". McGraw-Hill, 2015.

[9] Heizer, J., Render, B. Operations Management. Pearson, 2017.

[10] Kruse, R. Statistics for Healthcare Management. Springer, 2019.