# AN OVERVIEW OF MODELS FOR CONTACT CENTER RESOURCE PLANNING

DOI: 10.36724/2072-8735-2024-18-10-55-63

**Valentin G. Popov,**
*Moscow Technical University of Communications and Informatics, Moscow, Russia,*
**valentin1732@mail.ru**

**Mikhail S. Stepanov,**
*Moscow Technical University of Communications and Informatics, Moscow, Russia,*
**m.s.stepanov@mtuci.ru**

**Manuscript received** 18 September 2024;
**Accepted** 12 October 2024

***Keywords:*** *contact centers, service models, Poisson processes, machine learning, resource optimization, multichannel systems, artificial intelligence.*

**This article provides an in-depth examination of contemporary call handling models used in contact centers, emphasizing their evolution and adaptation to current technological advancements. The study starts with classical models, such as the Erlang models (Erlang B, Erlang C, and Erlang A), which are grounded in Poisson processes and Markov chains, and highlights their applicability in different contact center scenarios. These models are crucial for evaluating system performance, such as calculating the probability of service denial, waiting times, and the optimal number of agents needed. In addition to these traditional models, the article also explores more advanced approaches that incorporate artificial intelligence (AI) and machine learning (ML). These modern techniques are increasingly used for predicting incoming call loads, optimizing resource allocation, and staffing. The inclusion of AI and ML allows for more accurate forecasts of peak periods, better resource management, and improved service levels in a dynamic environment. The article thoroughly considers various unique characteristics of contact centers, such as multichannel communication systems, the segmentation of agents based on skill levels, and the presence of multiple service stages. These factors are essential in creating more sophisticated and realistic models that better reflect the complexities of modern contact center operations. The primary goal of the study is to conduct a comprehensive analysis of existing models to identify their strengths and weaknesses. The insights gained from this analysis will be instrumental in developing a new set of models tailored to describe modern and future information and referral services more accurately. These new models are expected to offer enhanced capabilities in managing the increasingly complex and technology-driven environments of contact centers.**

**Для цитирования:**
*Попов В.Г., Степанов М.С.* Обзор моделей для планирования ресурсов контакт-центра // T-Comm: Телекоммуникации и транспорт. 2024. Том 18. №10. С. 55-63.

**For citation:**

Popov V.G., Stepanov M.S. (2024). An overview of models for contact center resource planning. *T-Comm*, vol. 18, no.10, pp. 55-63.

## Introduction

Contact centers play a vital role in the modern world, providing a link between companies and their clients. They are used to resolve a wide range of issues: from providing information and technical support to conducting market research and handling customer complaints. In the context of increasing competition and rising customer expectations, the efficiency of contact centers becomes critically important for maintaining high service levels and strengthening companies' reputations.

One of the key aspects of contact center management is determining the optimal number of agents required to ensure quality service. Excessive agent numbers lead to unjustified costs, while a shortage can increase wait times, leading to customer dissatisfaction and potential losses for the company. Additionally, a lack of agents increases the workload on those already employed, which can result in professional "burnout" and increased turnover among trained personnel, negatively affecting the company's financial performance. Mathematical modeling is widely used to analyze and optimize contact center operations, allowing for the determination of the necessary number of agents and forecasting various scenarios based on changing conditions. Numerous studies are devoted to mathematical modeling.

The goal of this article is to analyze existing methods used to assess the characteristics of contact centers. By identifying their advantages and disadvantages, recommendations can be made for building a set of new models that adequately reflect the features of modern and prospective information and referral services. The results of the analysis will be further used to develop a set of new contact center models that most accurately describe modern and prospective information and referral services.

## Erlang Model

The classical Erlang model, also known as Erlang B, is used to analyze mass service systems where requests arrive according to a Poisson process, and service times are exponentially distributed. This model is applicable to systems where requests may be rejected if all agents are busy. The model was developed by Danish mathematician Agner Krarup Erlang and serves as the foundation for many modern contact center analysis models [4, 5].

Parameters:

- $\lambda$ – The intensity of incoming requests measured in the number of requests per unit of time.
- $\mu$ – The service intensity measured in the number of requests serviced per unit of time.
- $s$ – The number of agents or service devices in the system.
- $a$ – The system load indicator calculated as the ratio of incoming request intensity to service intensity.

Incoming requests enter the system with an intensity of $\lambda$ and are processed by $s$ agents at an intensity of $\mu$. If all agents are busy, new requests are denied.

To calculate the probability of service denial, the following formula is used:

$$B(s,a)=\frac{\frac{a^s}{s!}}{\sum_{k=0}^{s}\frac{a^k}{k!}}$$

where $a=\frac{\lambda}{\mu}$ — is the system load indicator, representing the ratio of incoming request intensity to service intensity.

System load indicator:

$$a=\frac{\lambda}{\mu}$$

Average number of busy agents:

$$L=a\times(1-B(s,a))$$

The advantage of this model is its simplicity and mathematical rigor, making it convenient for theoretical analysis and practical application to calculate key system characteristics. However, its simplicity is also its main drawback. The Erlang model does not account for the possibility of waiting in line and repeat attempts, making it less applicable to systems where these factors are important.

## Erlang Model with Infinite Waiting

The Erlang model with infinite waiting, also known as Erlang C, extends the classical model by introducing the possibility of infinite waiting in the queue. This model is used for systems where requests can wait until an agent becomes available. Erlang C allows for a more accurate assessment of contact center parameters, especially under high load conditions when service denials are unacceptable [4, 6].

Main Parameters of the Erlang Model with Infinite Waiting:

- $\lambda$ – The intensity of incoming requests.
- $\mu$ – The service intensity.
- $s$ – The number of agents.
- $a$ – The system load indicator calculated as the ratio of incoming request intensity to service intensity.

Requests enter the system at an intensity of $\lambda$. If an agent is free, the request is immediately processed. If all agents are busy, the request is placed in a queue to await service.

The Erlang model with expectation is shown in Figure 1.

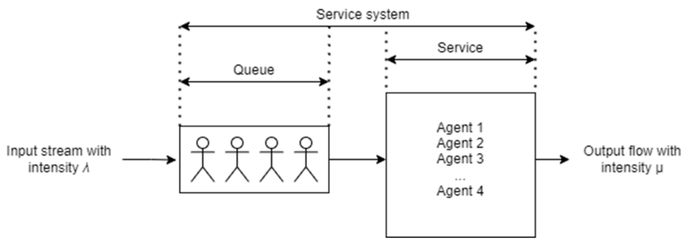


**Fig. 1.** The Erlang model with infinite expectation

The formulas for evaluating the main characteristics of the model are given below.

Probability of waiting:

$$C(s,a)=\frac{\frac{a^s}{s!}\times\frac{s}{s-a}}{\sum_{k=0}^{s-1}\frac{a^k}{k!}+\frac{a^s}{s!}\times\frac{s}{s-a}}$$

Average waiting time in queue:

$$W_q = \frac{C(s,a)}{s \times \mu - \lambda}$$

System load indicator:

$$a = \frac{\lambda}{\mu}$$

Average number of busy agents:

$$L = a$$

This model accounts for queue waiting time, making it more realistic for high-load systems. It also allows the estimation of average waiting time and the average number of busy agents. However, it assumes infinite waiting time, which may not match real-life scenarios where customers may leave the system if their wait is too long. Additionally, it requires more complex calculations compared to the classical Erlang model.

### Erlang Model with Finite Waiting

The Erlang model with finite waiting, also known as Erlang A, accounts for the possibility of limited waiting time in the queue. This model is suitable for systems where requests may leave the queue if the waiting time exceeds the allowable limit. Erlang A is more complex but more realistic, as it accounts for customers leaving the queue due to excessively long waits [2, 4].

Model parameters:

- $\lambda$ – The intensity of incoming requests measured in the number of requests per unit of time.
- $\mu$ – The service intensity measured in the number of requests serviced per unit of time.
- $s$ – The number of agents or service devices in the system.
- $a$ – The system load indicator calculated as the ratio of incoming request intensity to service intensity.
- $W$ – The maximum waiting time after which the request leaves the system.
- $\theta$ – The intensity of customer departure from the queue (the inverse of the maximum waiting time $W$).

Requests enter the system with an intensity of $\lambda$. If an agent is free, the request is immediately processed. If all agents are busy, the request is placed in a queue and waits for an agent to become available, but not longer than the time $W$. Requests that exceed the waiting time $W$, leave the system.

The diagram of the Erlang model with finite expectation is shown in Figure 2.

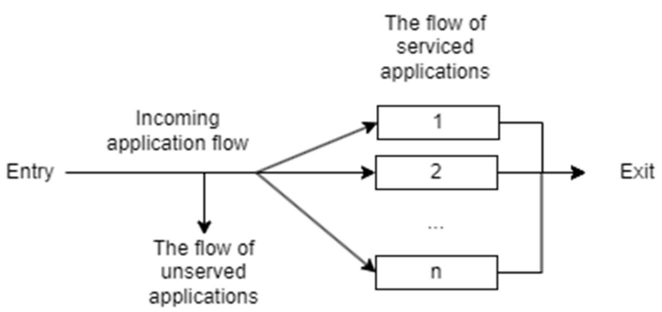


**Fig. 2.** The Erlang model with finite expectation

Probability of waiting:

$$P_w = \frac{\frac{a^s}{s!} \times \frac{s}{s-a}}{\sum_{k=0}^{s-1} \frac{a^k}{k!} + \frac{a^s}{s!} \times \frac{s}{s-a}}$$

Average waiting time in queue:

$$W_q = \frac{P_w}{s \times \mu - \lambda + \theta}$$

Average time in the system:

$$W_s = W_q + \frac{1}{\mu}$$

System load indicator:

$$a = \frac{\lambda}{\mu}$$

Average number of busy agents:

$$L = a$$

This model accounts for limited waiting time, making it more applicable in practice, as customers may leave the queue when waiting too long. This model allows the estimation of the probability of waiting, average waiting time, and probability of denial. The consideration of the additional parameter $\theta$ adds complexity to the calculations but allows for more accurate predictions of system behavior and resource optimization to improve customer service. However, it should be noted that the model assumes that service time and waiting time are exponentially distributed random variables, which may not always reflect the real situation.

### Generalized Model of Call Handling in Prospective Contact Centers

The next model discussed in the article includes several key elements and stages of request processing. The generalized structure of the model describes the process of receiving and processing requests in current and prospective contact centers using various communication channels such as telephone lines and the internet [4, 17-24].

Main Elements and stages of the model:

- Incoming Request Streams: Primary requests are received through access lines and serviced by IVR devices, agents, and consultants. Request streams follow a Poisson distribution with an intensity of $\lambda$.
- Service Stages: The first phase involves obtaining information from the IVR device, the second and third phases involve obtaining general and specialized information from the agent and consultant, respectively. The phase durations have an exponential distribution.
- Transition Probabilities: The transition probabilities between service stages are modeled by the parameters $q$ (probability of continuing service with the agent) and $p$ (probability of obtaining specialized information from the consultant).

• Service Denials: The model accounts for five types of denials, including lack of access lines and the occupation of all agents or consultants. Clients may repeat their requests after a random time interval or abandon further attempts.

• Waiting for Service Start: The possibility of waiting for service start if all agents or consultants are busy, with waiting time having an exponential distribution.

• Access Lines, Agents, and Consultants: The model considers the number of access lines, agents, and consultants, determining their interaction and workload during request processing.

• Repeated Requests: The interval between consecutive repeated attempts by a single subscriber is modeled with a probability *H*.

Some of the model's parameters:

• $j$ – The number of clients in a state of repeated request for information service.

• $i$ – The number of occupied access lines.

• $l$ – The total number of occupied agents and their waiting areas.

• $l_k$ – The total number of occupied consultants in the $k$-th group and their waiting areas.

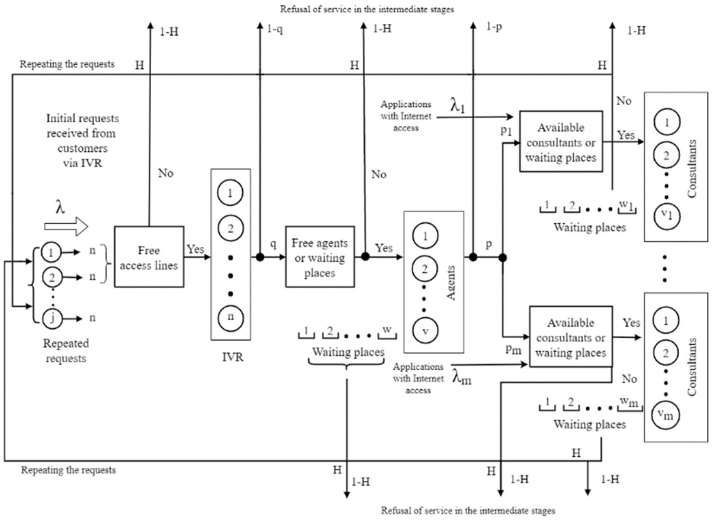


**Fig. 3.** The generalized model of contact center

The conditions of belonging to the states, the state vector of the contact center model $(j,i,l,l_1,...,l_m)$ must meet the following conditions:

$$j = 0,1,...$$

$$i = 0,1,...n$$

$$l = 0,1,...\upsilon + w$$

$$l_k = 0,1,...,\upsilon_k + w_k, k = 1,2,...m$$

Additional requirements are the fulfillment of the inequality:

$$i \geq \ell + \ell_1 + \ell_2 + \ldots + \ell_m$$

The change in the state of the model over time is described by a random process $r(t) = (j(t), i(t), \ell(t), \ell_1(t), \ldots, \ell_m(t))$, which takes values in the state space *S*. This process is Markovian, since all random variables that determine the duration of the model's stay in different states have an exponential distribution and do not depend on each other.

The following ratios are used to numerically evaluate the values of Internet query quality indicators:

$$p(j) = \frac{1}{j!}\left(\frac{\lambda H}{\nu(1-H)}\right)^j \times e^{-\left(\frac{\lambda H}{\nu(1-H)}\right)}, j = 0,1,2,...$$

It follows from this expression that the value of the average number of subscribers repeating a call is limited by:

$$\frac{\lambda H}{\nu(1-H)}$$

The model has analytical simplicity based on Markov processes, which simplifies the analysis and calculation of characteristics, flexibility allowing various request streams and service stages to be accounted for, and practical application for planning and optimizing contact center operations. However, the model has several drawbacks:

• Limited consideration of modern technologies (e.g., artificial intelligence and machine learning);

• Insufficient selection of communication channels relative to modern omnichannel contact centers;

• Simplified hypotheses (use of exponential distributions and Poisson law) that may not fully reflect the real complexity of request arrival and service processes;

• Incomplete consideration of customer behavior when interacting with different contact center elements.

### A queue management model based on customer priorities

The authors of such models are various researchers working on optimizing queuing systems. One example is the article "Priority Queueing Systems: Modeling and Analysis" by Smith J. and Zhang Y. [2].

The model is designed to improve queue management in contact centers, paying special attention to customer priorities. These priorities can be determined based on various criteria, such as the client's status (VIP, regular), the type of request (urgent, regular) or the expected time to resolve the issue. Incoming calls are queued according to priority, and agents serve customers in the order of their priority. For example, calls from VIP customers will be processed faster than calls from regular customers. The model also takes into account the dynamic change of queues and the possibility of redistributing agents between queues to ensure optimal service [2, 8].

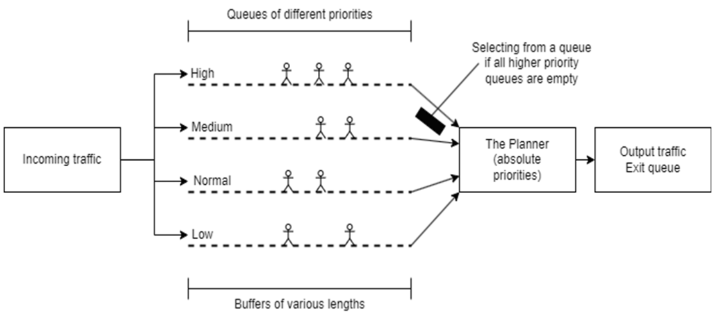


**Fig. 4.** Priority-based queue management scheme

Customers contacting the contact center receive priority based on pre-established criteria, such as the status of the customer or the type of request.

Customer calls are allocated in queues depending on their priority, and agents serve customers starting from the highest priority. If all agents are busy, clients remain in the queue until the agent is released. Depending on the current load on the contact center and the distribution of customer priorities, the system can dynamically redistribute agents between queues to ensure optimal service [2, 10].

Model parameters:

- $\lambda_i$ – The intensity of the incoming customer flow with priority *i*.
- $\mu_i$ – Average customer service time with priority *i*.

Average waiting time for priority customers *i*:

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

Probability of customer service with priority *i*:

$$P_i = \frac{W_i}{\sum W_j}$$

The model improves the service of priority customers and allows for more efficient allocation of resources, which increases the satisfaction of VIP customers. However, it does not take into account variations in agent productivity and assumes fixed priorities, which may not correspond to the dynamic reality of the contact center.

### A model for predicting the load on a contact center using machine learning

The goal was presented in the article [1]. This model is based on the use of machine learning algorithms to predict the future load on the contact center. The input data for the model includes the history of incoming calls, time of day, day of the week, marketing campaigns, holidays and other significant events. Using this data, the model learns to identify patterns and patterns, which allows it to accurately predict the number of calls in the future. For example, the model can predict an increase in workload during holidays or during special marketing campaigns. This allows the contact center to prepare in advance, allocating resources and staff in such a way as to cope with the expected workload.

At the first stage, historical data on incoming calls and significant events is collected, including the time of calls, day of the week, duration of calls and information about marketing campaigns and holidays. This data is preprocessed, including cleaning and normalization. The machine learning model is then trained on this data to identify patterns and dependencies, which allows you to predict the future number of calls. The model is used to predict the load, taking into account current and upcoming events, which allows the contact center to adapt its work plans, distributing agents and scheduling their shifts in accordance with forecasts [1, 14].

Model parameters:

- $X_t$ – Vector of input parameters (historical data, time of day, etc.).
- $f$ – A function approximating the relationship between the input parameters and the number of calls.
- $\epsilon_t$ – Model error.
- $\theta$ – Model parameters.
- $\widehat{y_t}$ – The predicted value of the number of calls.
- $y_t$ – The actual value of the number of calls.

The predicted number of calls at a given time *t*:

$$y_t = f(X_t) + \epsilon_t$$

Loss function:

$$L(\theta) = \frac{1}{N}\sum_{t=1}^{N}(y_t - \widehat{y_t})^2$$

The model allows you to accurately predict future workload, which contributes to effective planning and resource allocation. However, it requires a large amount of historical data for training and may not take into account unexpected events, which reduces the accuracy of forecasts in such situations.

### Optimization model of a multidisciplinary contact center using queue theory

The goal was presented in the article [3]. The model describes the optimization of personnel management in a multidisciplinary contact center. It takes into account three types of calls and four groups of agents with different skills. The purpose of the model is to minimize personnel costs while meeting the requirements of the service level. The model considers parameters such as the intensity of incoming calls for each type, the speed of agent service in each group, and routing rules based on agent skills and call priority.

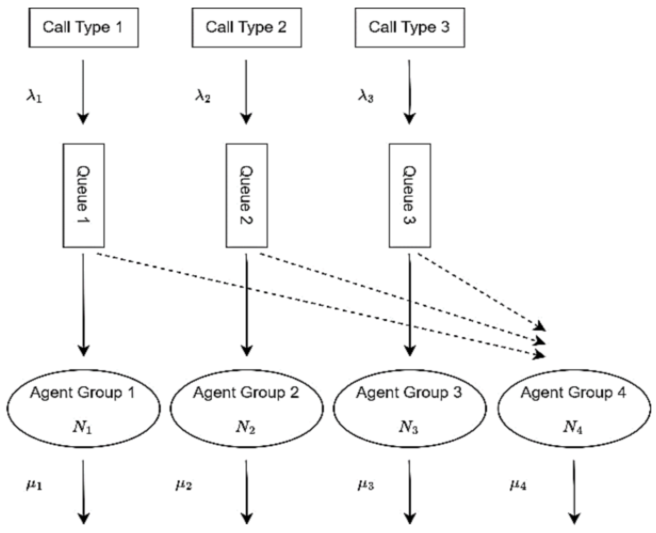


**Fig. 5.** The scheme of a multidisciplinary contact center

The model includes three types of calls (Type 1, Type 2, and Type 3) received according to the Poisson process with intensities $\lambda_1$, $\lambda_2$ and $\lambda_3$ accordingly. There are four groups of agents working in the contact center: Group 1, Group 2, Group 3 and Group 4, each of which has certain skills. Group 4 agents have all the necessary skills and can handle all types of calls. Incoming calls are allocated in queues depending on the type and priority of calls. Agents handle calls according to the routing policy, where calls with a higher priority are handled first. The model takes into account system states such as free, busy and overloaded, and transitions between these states depending on the receipt of new calls or the completion of service [1, 9, 15].

Probability of the system state:

$$P_i = \frac{p^{N_i}{}_i}{N_i!\sum_{j=0}^{N_i}\frac{p^j{}_i}{j!}}$$

where $p_i = \frac{\lambda_i}{\mu_i}, \lambda_i$ – the intensity of incoming calls of the type *i*, $\mu_i$ – The speed of call service type *i*, $N_i$ – the number of agents handling calls of the type *i*.

Transient probabilities:

$$q_{i-j} = \lambda_i P(n_i = N_i - 1)$$

where $P(n_i = N_i - 1)$ – the probability that the number of calls of the type *i*, waiting for service, equal to $N_i - 1$.

Service level:

$$SL_i = 1 - P_{ns,i}$$

where $SL_i$ – the level of call service type *i*, $P_{ns,i}$ – the probability that a call of the type *i*, will not be serviced at the specified time $T_i$.

The model allows you to efficiently allocate staff in the contact center, minimizing costs and meeting the requirements of the service level. It takes into account the different skills of agents and priorities of calls, which helps to improve the quality of service. However, the model does not take into account the variability of service time and possible changes in the load on the contact center in real time, which can reduce the accuracy of forecasts and management efficiency in a changing work environment.

## Development of a customer service automation system in the admissions office of universities and colleges

The article [16] presents a model of a contact center using a chatbot.

The mathematical model of the contact center functioning, special attention is paid to the use of a chatbot to automate the process of processing client requests. The model is based on queuing theory methods and includes a detailed description of the interaction between the chatbot and the contact center agents. The main purpose of this model is to provide a detailed analysis of the effectiveness of the contact center, taking into account various types of requests, as well as the time and resource costs of processing them.

The model allows not only to evaluate the overall performance of the system, but also to identify optimal functioning parameters, such as the number of agents required to provide a given level of service, and the allocation of resources between automated and manual request processing. Special attention is paid to the process of routing requests: starting from their entry into the system, through the stage of pre-processing by the chatbot, and, if necessary, transferring more complex requests to agents for consideration.

The model also takes into account the probabilities of different processing outcomes at each stage, which allows you to accurately predict the load on agents and optimize their number to minimize operating costs. The use of this model provides an opportunity not only for strategic planning of the contact center, but also for rapid response to changes in the flow of requests, which is especially important in conditions of dynamically changing customer requirements and market conditions. As a result, the proposed model contributes to a more efficient use of contact center resources, reduces customer waiting times, and increases overall user satisfaction through more prompt and accurate service of their requests.

The main elements of the model:

1) Types of requests:

• The model takes into account two main types of requests: routine (simple) ones that can be processed by a chatbot, and complex ones that require agent participation.

• The probability that a request can be processed at the first stage by a chatbot is determined through a set of probabilities $k = 1, 2, ..., n$. For each query category *k* there are two probabilities: the probability that the request will be routine and processed at the chatbot stage ($p_k$), and the probability that the request will require transmission to the agent ($q_k$).

2) Request flow:

• The flow of requests is described using a Poisson distribution with intensity $\lambda$, what does it mean that requests are randomly received with a certain average frequency.

• The request flow can be divided into several categories depending on the complexity and probability of their processing. Each category is characterized by its own probabilities $p_k$ and $q_k$, where $p_k$ – the probability that the request will be processed by the chatbot, a $q_k$ – the probability that the request will go to the agent.

3) Request service:

• Requests that cannot be processed by the chatbot are transmitted to agents. The agent's service process is modeled using a queuing system with waiting and a finite queue length.

• Agent maintenance is described by an exponential distribution with the parameter $\mu$, which characterizes the average service time of the request.

• The number of agents ($\nu$) and the length of the waiting queue ($\omega$) are key parameters that affect the performance of the contact center.

4) Maintenance stages:

• The first stage is the processing of requests by the chatbot, which can end either with the successful provision of information or with the transfer of the request to the agent.

• The second stage is the processing of requests by agents. If the waiting queue is full, the request may be lost.

The mathematical model with a chatbot is shown in Figure 6.

This model provides two stages of request processing. Requests that cannot be satisfied using the chatbot are redirected to the agents [16].

Poisson flow of requests with intensity:

$$\Lambda = \lambda \sum_{k=1}^{n} p_k (1 - f_k)$$

The intensity of the Poisson flow coming to the agents and expressed in erlangs:

$$a = \frac{\Lambda}{\mu}$$

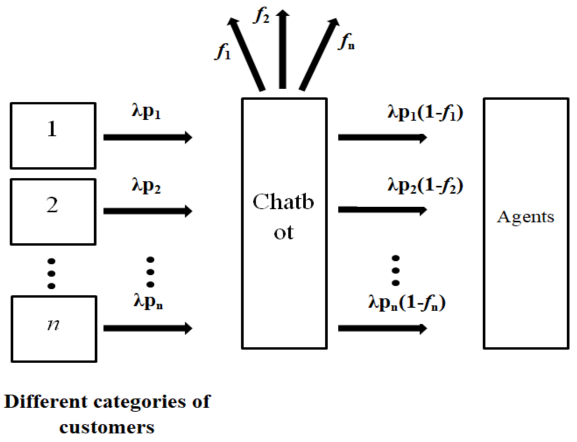


**Fig. 6.** Mathematical model of a contact center with a chatbot

The ratio of lost requests is the same as the proportion of time spent by the model in the state:

$$\pi_c = \pi_t = p(\upsilon + w)$$

In addition to the ratio of lost requests, important characteristics of the model are the following: the average waiting time for a request before the start of service, calculated based on all rejected requests that entered the queue; the average number of requests in the queue, determined by analyzing its length for those requests that were rejected; the probability that the request will enter the waiting queue; $y$ – the average number of employed agents. These indicators can be conveniently calculated using the following recursive formulas:

$$W_d = \left( \frac{1}{\upsilon - a} - \frac{w\left(\frac{a}{\upsilon}\right)^w}{\upsilon\left(1 - \left(\frac{a}{\upsilon}\right)^w\right)} \right) \frac{1}{\mu}$$

$$L_b = \left( \frac{\Lambda}{\upsilon - a} - \frac{w\left(\frac{a}{\upsilon}\right)^{w+1}}{1 - \left(\frac{a}{\upsilon}\right)^w} \right) \left( 1 - \frac{\left(1 - \frac{a}{\upsilon}\right)\left(\frac{a}{\upsilon}\right)^w}{1 - \left(\frac{a}{\upsilon}\right)^{w+1}} \right)$$

$$p\{W > 0\} = \frac{\upsilon\left(1 - \left(\frac{a}{\upsilon}\right)^w\right)}{E(\upsilon, a) + a\left(1 - \left(\frac{a}{\upsilon}\right)^w\right)}$$

$$p(\upsilon + w) = \frac{(\upsilon - a)\left(\frac{a}{\upsilon}\right)^w}{\frac{\upsilon - a}{E(\upsilon, a)} + a\left(1 - \left(\frac{a}{\upsilon}\right)^w\right)}$$

$$y = a(1 - p(\upsilon + w))$$

$$L_q = \Lambda W_d p\{W > 0\}$$

$$W_q = \frac{W_d p\{W > 0\}}{1 - p(\upsilon + w)}$$

The main advantages of the presented mathematical model are the consideration of various categories of requests, the possibility of optimizing the number of agents and reducing maintenance costs, as well as calculating the probability of failures at the stage of processing requests by agents. However, the model does not take into account the dynamic change in the intensity of requests, the impact of multi-channel communication, complex service scenarios and the lack of self-learning ability of the chatbot, which is important in modern contact centers.

## Conclusions

The study examined the evolution of mathematical models used to analyze and optimize the work of contact centers, as well as their adaptation to modern requirements. Classical models such as Erlang B and Erlang C and their modern modifications, including models that take into account customer priorities and the possibility of dynamic resource management, are analyzed. The article also focuses on new approaches based on the use of machine learning for load forecasting and resource optimization.

The results of the analysis show that the development of technology and the changing needs of customers require the creation and implementation of more flexible and adaptive models that can take into account multichannel, variability in customer behavior and the use of modern technologies such as artificial intelligence. The models developed on the basis of these approaches will ensure more efficient management of contact centers, improving the quality of service and customer satisfaction.

Thus, for the successful operation of contact centers in a rapidly changing market and increasing requirements for the quality of service, it is necessary to develop and implement current models that are able to adapt to modern conditions and technologies.

## References


[1] W. Li, H. Chen, "Forecasting Call Center Workload Using Machine Learning." *Journal of Artificial Intelligence Research*, 2023, pp. 35-48, 120 p.

[2] J. Smith, Y.Zhang, "Priority Queueing Systems: Modeling and Analysis," *Operations Research*, 2022, pp. 112-126.

[3] N.A. Hassan, N.M.S. Abdallah, R.A. Attwa, "Optimizing multi-skill call center staffing using queuing models: A study of service level," *Journal of Applied Research and Technology*, 2024, pp. 88-101.

[4] M.S. Stepanov, "Development and Analysis of a Generalized Model for Call Handling in Prospective Contact Centers." PhD dissertation, Moscow: MTUCI, 2016.

[5] Y. Xu, Z. Li, "Multi-channel Call Center Efficiency: An Empirical Study," *Journal of Service Research*, 2021, pp. 203-219.

[6] R. Kumar, A.Singh, "Advances in Call Center Analytics: A Comprehensive Review," *Telecommunication Systems*, 2022, pp. 78-92.

[7] T. Wang, L.Zhou, "AI-Powered Call Centers: Enhancing Customer Experience," *IEEE Transactions on Services Computing*, 2023, pp. 45-59.

[8] K. Brown, J. Taylor, "Modeling Call Center Performance: New Trends and Challenges," *Journal of Telecommunications Management*, 2024, pp. 119-135.

[9] S. Gupta, M. Sharma, "Queue Management in Contact Centers: A Simulation Approach," *International Journal of Information Technology & Decision Making*, 2023, pp. 67-82.

[10] J. Lee, S. Park, "Dynamic Resource Allocation in Multi-Skill Call Centers," *Journal of Operational Research Society*, 2022, pp. 150-164.

[11] I.A. Ivanov, P.M. Sidorov, "Modeling Mass Service in Contact Centers: Methods and Applications," *Bulletin of Moscow State Technical University*, 2023, pp. 200-215.

[12] P. Zhang, Y. Liu, "Customer Satisfaction in Call Centers: A Review of Key Factors," *Service Science*, 2021, pp. 89-103.

[13] A. Andersson, P. Svensson, "Telecommunications and Service Quality: Case Studies in Modern Call Centers," *European Journal of Operational Research*, 2022, pp. 240-256.

[14] E.N. Melnikova, A.V. Petrov, "Predicting Workload in Contact Centers Using Neural Networks," *Information Technology and Modeling*, 2022, pp. 120-136.

[15] R. Johnson, D. Williams, "Emerging Technologies in Contact Centers: Impact on Operations," *Journal of Service Management*, 2024, pp. 100-115.

[16] M.S. Stepanov, V.G. Popov, N.S. Fedorova, F.S. Kroshin, and A.R. Muzata, "The development of a chatbot for university admissions office: A mathematical model of a contact center," *T-Comm*, 2024, no. 16(10), pp. 51-56.

[17] S.N. Stepanov, M.S. Stepanov, "Construction and analysis of a generalized model of a contact center," *Automation and Remote Control*. 2014. No. 11, pp. 55-69.

[18] S.N. Stepanov, M.S. Stepanov, "Algorithms for estimating throughput characteristics in a generalized call center model," *Automation and Remote Control*. 2016. No. 7, pp. 86-102.

[19] S.N. Stepanov, M.S. Stepanov, M.O. Shishkin, "Performance Measures of Emergency Services in Case of Overload," *Lecture Notes in Computer Science*. 2020. Vol. 12563 LNCS, pp. 436-449.

[20] *M.S. Stepanov, A.R. Muzata, V.D. Zyuzin, N.S. Kostina, M.O. Shishkin*, "Estimation of Contact Center Performance Measures in Case of Overload and Chatbot Implementation," 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, 2021, P. 9415983.

[21] S.N. Stepanov, M.S. Stepanov, H.M. Zhurko, "The Modeling of Call Center Functioning in Case of Overload," *Lecture Notes in Computer Science*. 2019. Vol. 11965 LNCS, pp. 391-406.

[22] S. Stepanov, M. Stepanov, "Estimation of the Performance Measures of a Group of Servers Taking into Account Blocking and Call Repetition before and after Server Occupation," *Mathematics*, 2021, vol. 9. No. 21, 2811.

[23] A.P. Buslaev, D.A. Kuchelev, M.V. Yashina, "Dynamic systems and mathematical models of information traffic," *T-Comm*. 2018. Vol. 12. No. 3, pp. 22-38.

[24] I.V. Bogachkov, "Detection of strained sections in optical fibers based on the Brillouin reflectometry method," *T-Comm*. 2016. Vol. 10. No. 12, pp. 85-91.

## ОБЗОР МОДЕЛЕЙ ДЛЯ ПЛАНИРОВАНИЯ РЕСУРСОВ КОНТАКТ-ЦЕНТРА

**Попов Валентин Геннадьевич,** *Московский технический университет связи и информатики, Москва, Россия,* ***valentin1732@mail.ru***

**Степанов Михаил Сеергевич,** *Московский технический университет связи и информатики, Москва, Россия,* ***m.s.stepanov@mtuci.ru***

**Аннотация**

В данной статье проводится углубленный анализ современных моделей обработки вызовов в контакт-центрах, с акцентом на их эволюцию и адаптацию к текущим технологическим достижениям. Исследование рассматривает классические модели, такие как модели Эрланга (Эрланг B, Эрланг C и Эрланг A), которые основаны на пуассоновских процессах и цепях Маркова, и подчеркивает их применимость в различных сценариях работы контакт-центров. Эти модели важны для оценки эффективности системы, например, расчета вероятности отказа в обслуживании, времени ожидания и оптимального числа операторов. В дополнение к традиционным моделям, в статье также рассматриваются более продвинутые подходы, включающие искусственный интеллект (ИИ) и машинное обучение (МО). Эти современные техники все чаще используются для прогнозирования нагрузки вызовов, оптимизации распределения ресурсов и управления персоналом. Использование ИИ и МО позволяет более точно предсказывать периоды пиковых нагрузок, оптимизировать управление ресурсами и повышать уровень обслуживания клиентов. В статье рассмотрены способы учета в математических моделях таких особенностей современных контакт-центров, как омниканальность, сегментация операторов по уровню навыков и наличие нескольких этапов обслуживания. Эти факторы являются важными при создании более сложных и реалистичных моделей, которые лучше отражают сложность современных справочных служб. Основная цель исследования – провести комплексный анализ существующих моделей для выявления их сильных и слабых сторон. Полученные данные используются для разработки нового комплекса моделей, более точно описывающих современные и перспективные справочно-информационные службы.

***Ключевые слова:*** *контакт-центры, модели обслуживания, пуассоновские процессы, машинное обучение, оптимизация ресурсов, многоканальные системы, искусственный интеллект.*

**Литература**

1. *Li W., Chen H.* Forecasting Call Center Workload Using Machine Learning // Journal of Artificial Intelligence Research, 2023, pp. 35-48, 120 p.
2. *Smith J., Zhang Y.* Priority Queueing Systems: Modeling and Analysis // Operations Research, 2022, pp. 112-126.
3. *Hassan N.A., Abdallah N.M.S., Attwa R.A.* Optimizing multi-skill call center staffing using queuing models: A study of service level // Journal of Applied Research and Technology, 2024, pp. 88-101.
4. *Степанов М.С.* Разработка и анализ обобщенной модели обслуживания вызовов в перспективных контакт-центрах. Диссертация на соискание ученой степени кандидата технических наук. Москва: МТУСИ, 2016.
5. *Xu Y., Li Z.* Multi-channel Call Center Efficiency: An Empirical Study // Journal of Service Research, 2021, pp. 203-219.
6. *Kumar R., Singh A.* Advances in Call Center Analytics: A Comprehensive Review // Telecommunication Systems, 2022, pp. 78-92.
7. *Wang T., Zhou L.* AI-Powered Call Centers: Enhancing Customer Experience // IEEE Transactions on Services Computing, 2023, pp. 45-59.
8. *Brown K., Taylor J.* Modeling Call Center Performance: New Trends and Challenges // Journal of Telecommunications Management, 2024, pp. 119-135.
9. *Gupta S., Sharma M.* Queue Management in Contact Centers: A Simulation Approach // International Journal of Information Technology & Decision Making, 2023, pp. 67-82.
10. *Lee J., Park S.* Dynamic Resource Allocation in Multi-Skill Call Centers // Journal of Operational Research Society, 2022, pp. 150-164.
11. *Иванов И.А., Сидоров П.М.* Моделирование массового обслуживания в контакт-центрах: методы и приложения // Вестник Московского государственного технического университета, 2023, с. 200-215.
12. *Zhang P., Liu Y.* Customer Satisfaction in Call Centers: A Review of Key Factors // Service Science, 2021, pp. 89-103.
13. *Andersson A., Svensson P.* Telecommunications and Service Quality: Case Studies in Modern Call Centers // European Journal of Operational Research, 2022, pp. 240-256.
14. *Мельникова Е.Н., Петров А.В.* Прогнозирование нагрузки в контакт-центрах с использованием нейронных сетей // Информационные технологии и моделирование, 2022. С. 120-136.
15. *Johnson R., Williams D.* Emerging Technologies in Contact Centers: Impact on Operations // Journal of Service Management, 2024, pp. 100-115.
16. *Stepanov M.S., Popov V.G., Fedorova N.S., Kroshin F.S., Muzata A.R.* The development of a chatbot for university admissions office: A mathematical model of a contact center // T-Comm: Телекоммуникации и транспорт, 2024. № 16(10). С. 51-56.
17. *Степанов С.Н., Степанов М.С.* Построение и анализ обобщенной модели контакт-центра // Автоматика и Телемеханика. 2014. №11. С. 55-69.
18. *Степанов С.Н., Степанов М.С.* Алгоритмы оценки показателей пропускной способности обобщенной модели контакт-центра // Автоматика и Телемеханика. 2016. № 7. С. 86-102.
19. *Stepanov S.N., Stepanov M.S., Shishkin M.O.* Performance Measures of Emergency Services in Case of Overload // Lecture Notes in Computer Science. 2020. Vol. 12563 LNCS, pp. 436-449.
20. *Stepanov M.S., Muzata A.R., Zyuzin V.D., Kostina N.S., Shishkin M.O.* Estimation of Contact Center Performance Measures in Case of Overload and Chatbot Implementation // 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, 2021, P. 9415983.
21. *Stepanov S.N., Stepanov M.S., Zhurko H.M.* The Modeling of Call Center Functioning in Case of Overload // Lecture Notes in Computer Science. 2019. Vol. 11965 LNCS, pp. 391-406.
22. *Stepanov S., Stepanov M.* Estimation of the Performance Measures of a Group of Servers Taking into Account Blocking and Call Repetition before and after Server Occupation // Mathematics, 2021, vol. 9. № 21, 2811.
23. *Буслаев А.П., Кучелев Д.А., Яшина М.В.* Динамические системы и математические модели трафика информации // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 3. С. 22-38.
24. *Богачков И.В.* Обнаружение натяжённых участков в оптических волокнах на основе метода бриллюэновской рефлектометрии // T-Comm: Телекоммуникации и транспорт. 2016. Т. 10. № 12. С. 85-91.