



# UNIVERSAL INSIGHTS CONFERENCE ON SCIENCE AND SOCIETY

## international scientific conference

### МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ КЛАССИЧЕСКИХ МОДЕЛЕЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ

**Иниятова К.Х.**

Ассистент-преподаватель

Нукусский государственный технический университет

Нукус, Узбекистан

**Турумбетов Б.К.**

Доцент (PhD.)

Нукусский государственный технический университет

Нукус, Узбекистан

***Аннотация.** В данной работе приведены основные принципы классификации и математическое описание классических моделей системы массового обслуживания.*

***Ключевые слова:** Система, поток, узел, символика, интервал, время, обслуживание, устройство, распределение.*

Системы массового обслуживания (СМО) классифицируются по нескольким признакам: характеру входного потока, распределению времени обслуживания, количеству обслуживающих устройств, ёмкости очереди, дисциплине обслуживания и структуре взаимосвязей между узлами.

Символика Кендалла - это стандартная символика, используемая для описания систем массового обслуживания (queueing systems), которая позволяет кратко и понятно выразить структуру системы. Эта символика была впервые предложена английским ученым Дэвидом Кендаллом в 1953 году. Символика Кендалла позволяет представлять системы

## UNIVERSAL INSIGHTS CONFERENCE ON SCIENCE AND SOCIETY

### international scientific conference

массового обслуживания стандартным и понятным способом. Этот метод значительно упрощает определение, сравнение и анализ основных параметров системы.

В распространённом варианте она включает обозначения вида  $A/S/c/K/N/D$ , где:

- A: распределение интервала поступления заявок в систему (interarrival time);
- S: распределение времени обслуживания (service time);
- c: Количество обслуживающих устройств (серверов);
- K: Если в модели не приводятся значения, то максимальное количество очередей в системе считается бесконечным; размер очереди предполагается бесконечным относительно системы;
- N: Популяция клиентов (для ограниченных систем) - произвольная;
- D: Порядок очереди (discipline) - произвольная.

Основные символы распределений:

- M: Марковское (экспоненциальное распределение);
- D: Детерминированное (распределение с определенным временем);
- U: Uniform (равномерное распределение);
- G: General (произвольное распределение);
- Ek: Распределение Эрланга k-го порядка.

Символика Кендалла удобна для сопоставления моделей, но сама по себе не описывает все особенности системы. Для сложных объектов дополнительно указываются правила маршрутизации, приоритеты, ограничения буфера и зависимость параметров от состояния или времени.

*Модель M/M/1.* Модель M/M/1 описывает систему с пуассоновским входным потоком интенсивности  $\lambda$ , экспоненциальным временем

## UNIVERSAL INSIGHTS CONFERENCE ON SCIENCE AND SOCIETY

### international scientific conference

обслуживания и одним обслуживающим устройством с интенсивностью обслуживания  $\mu$ .

Коэффициент загрузки системы определяется выражением

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

Для существования стационарного режима необходимо выполнение условия

$$\lambda < \mu \text{ или } \rho < 1$$

Вероятность нахождения в системе  $n$  заявок определяется формулой

$$P_n = (1 - \rho)\rho^n, \quad n = 0, 1, 2, \dots$$

Среднее количество заявок в системе:

$$m = \frac{\rho}{1 - \rho}$$

Среднее количество заявок в очереди:

$$l = \frac{\rho^2}{1 - \rho}$$

Среднее время ожидания заявок в очереди:

$$w = \frac{\rho}{\mu(1 - \rho)} = \frac{\lambda}{\mu(\mu - \lambda)}$$

Среднее время пребывания заявки в системе:

$$u = \frac{1}{\mu - \lambda} = w + \frac{1}{\mu}$$

Данная модель позволяет оценивать влияние изменения входной нагрузки на задержку и длину очереди. По мере приближения коэффициента загрузки к единице среднее время ожидания быстро увеличивается, что указывает на необходимость резервирования дополнительных ресурсов.

## UNIVERSAL INSIGHTS CONFERENCE ON SCIENCE AND SOCIETY international scientific conference

*Модель M/M/c.* Модель M/M/c используется для анализа системы с несколькими параллельно работающими обслуживающими устройствами. При одинаковой интенсивности обслуживания каждого устройства условие устойчивости имеет вид

$$\rho = \frac{\lambda}{c\mu} < 1$$

Здесь  $c$  - количество обслуживающих устройств,  $\lambda$  - интенсивность поступления заявок,  $\mu$  - интенсивность обслуживания одним устройством.

Введем параметр  $b$

$$b = \frac{1}{\mu}$$

Вероятность отсутствия заявок в системе определяется выражением

$$P_0 = \left[ \frac{(c\rho)^c}{c! (1 - \rho)} + \sum_{i=0}^{c-1} \frac{(c\rho)^i}{i!} \right]^{-1}$$

Вероятность того, что поступившая заявка будет вынуждена ожидать обслуживания, может быть определена по формуле Эрланга С:

$$P = \frac{(c\rho)^c}{c! (1 - \rho)} P_0$$

Среднее время ожидания заявок в очереди составляет

$$w = \frac{Pb}{c(1 - \rho)}$$

Модель применяется для анализа многоканальных систем, в которых несколько ресурсов обслуживают общий поток заявок. В телекоммуникациях подобный подход может использоваться для предварительной оценки нагрузки на группы серверов или параллельные вычислительные ресурсы.

*Модель M/G/1.* Модель M/G/1 характеризуется пуассоновским входным потоком, одним обслуживающим устройством и произвольным распределением времени обслуживания.

## UNIVERSAL INSIGHTS CONFERENCE ON SCIENCE AND SOCIETY international scientific conference

Обозначим через  $S$  случайную величину, характеризующую время обслуживания одной заявки. Тогда коэффициент загрузки равен

$$\rho = \lambda E[S]$$

здесь  $E$  - математическое ожидание.

При условии  $\rho < 1$  среднее время ожидания в очереди определяется формулой Поллачека-Хинчина:

$$w = \frac{\lambda E[S]^2}{2(1 - \rho)}$$

здесь  $E[S]^2$  - математическое ожидание квадрата времени обслуживания, а не квадрат его математического ожидания.

Среднее время пребывания заявки в системе:

$$u = w + E[S]$$

Преимущество модели состоит в возможности учитывать различные распределения времени обслуживания. Это важно в системах, где продолжительность обработки заявок зависит от их объема, типа или вычислительной сложности.

*Модель M/M/1/K.* В модели M/M/1/K количество заявок в системе ограничено значением  $K$ , включающим как обслуживаемую заявку, так и заявки, находящиеся в очереди.

Коэффициент загрузки системы

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

Вероятность нахождения  $n$  заявок в системе при  $\rho \neq 1$  равна

$$P_n = \frac{(1 - \rho)\rho^n}{1 - \rho^{K+1}}, \quad n = 0, 1, \dots, K$$

Здесь  $P_n$  - вероятность того, что в системе находится  $n$  заявок, а  $K$  - максимальное количество заявок в системе, включая обслуживаемую заявку.

## UNIVERSAL INSIGHTS CONFERENCE ON SCIENCE AND SOCIETY international scientific conference

Вероятность отказа в обслуживании (вероятность потери заявки) при поступлении новой заявки равна вероятности заполнения системы:

$$P_{loss} = P_K$$

Поскольку система вмещает максимум  $(K)$  заявок, новая заявка получает отказ, если все места заняты.

Раскрытая формула:

$$P_{loss} = \frac{(1 - \rho)\rho^K}{1 - \rho^{K+1}}$$

Эффективная интенсивность поступления обслуживаемых заявок составляет

$$\lambda_{eff} = \lambda(1 - P_K)$$

Эта формула показывает, сколько заявок в среднем поступает в систему и принимается на обслуживание за единицу времени с учётом потерь.

Эта модель позволяет оценивать влияние емкости буфера на потери заявок. Она особенно полезна при анализе пакетных сетей, где объем буферной памяти ограничен.

В работе рассмотрены основные принципы классификации и математического описания классических моделей систем массового обслуживания (СМО). Изучены особенности символики Кендалла, позволяющей компактно представлять структуру систем с различными характеристиками входного потока, распределением времени обслуживания, количеством обслуживающих устройств и ограничениями на ёмкость очереди.

Проведённый анализ показал, что каждая из рассмотренных моделей предназначена для решения определённого круга задач. Практическая значимость рассмотренных моделей заключается в возможности их применения при анализе телекоммуникационных сетей,



## UNIVERSAL INSIGHTS CONFERENCE ON SCIENCE AND SOCIETY international scientific conference

вычислительных систем, серверных комплексов и других объектов, функционирующих в условиях случайного поступления и обслуживания заявок. Использование математических моделей СМО позволяет обосновывать выбор количества обслуживающих устройств, оценивать потребность в буферной памяти и исследовать влияние изменения нагрузки на основные показатели качества обслуживания.

Таким образом, классические модели систем массового обслуживания составляют важную математическую основу для анализа производительности и оптимизации систем с очередями. Дальнейшие исследования целесообразно направить на применение данных моделей к неоднородным потокам заявок, нестационарной нагрузке и многоклассовым системам, а также на сопоставление аналитических результатов с данными имитационного моделирования. Это позволит повысить точность оценки характеристик реальных систем и расширить возможности использования теории массового обслуживания при решении современных задач управления сетевыми и вычислительными ресурсами.

### СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Baiocchi A. Network Traffic Engineering: Stochastic Models and Applications. Wiley, 2020. 816 p. DOI: <https://doi.org/10.1002/9781119632498>.
2. Giambene G. Queuing Theory and Telecommunications: Networks and Applications. 3rd ed. Springer, Cham, 2021. DOI: <https://doi.org/10.1007/978-3-030-75973-5>.
3. Anisimov V., Limnios N. (eds.). Queueing Theory 1: Advanced Trends. ISTE / Wiley, 2021. DOI: <https://doi.org/10.1002/9781119755432>.

## UNIVERSAL INSIGHTS CONFERENCE ON SCIENCE AND SOCIETY international scientific conference

4. Anisimov V., Limnios N. (eds.). Queueing Theory 2: Advanced Trends. Wiley, 2021. DOI: <https://doi.org/10.1002/9781119755227>.
5. Kim K. Finite-Buffer M/G/1 Queues with Time and Space Priorities // Mathematical Problems in Engineering. 2022. Article ID 4539940. DOI: <https://doi.org/10.1155/2022/4539940>.
6. Wan S., Lan S., Tang Y. Performance Analysis and Capacity Optimization Design of an M/G/1 Queueing System with Vacations and Patient Server // Journal of Systems Science and Mathematical Sciences. 2023. Vol. 43, No. 9. P. 2292–2309. DOI: <https://doi.org/10.12341/jssms22707>.
7. Aalto S., Scully Z. Minimizing the Mean Slowdown in the M/G/1 Queue // Queueing Systems. 2023. Vol. 104. P. 187–210. DOI: <https://doi.org/10.1007/s11134-023-09888-6>.
8. Wu H., He Q.-M., Xia L. The Impact of the Correlation Coefficient of Interarrival and Service Times on Queueing Performance: The M/M/1 Case // European Journal of Operational Research. 2024. Vol. 319, No. 2. P. 625–637. DOI: <https://doi.org/10.1016/j.ejor.2024.07.032>.
9. Thompson Y. L. E., Levine G. M., Chen W., et al. Applying Queueing Theory to Evaluate Wait-Time-Savings of Triage Algorithms // Queueing Systems. 2024. Vol. 108. P. 579–610. DOI: <https://doi.org/10.1007/s11134-024-09927-w>.
10. Алиев Т.И. Основы моделирования дискретных систем. – СПб: СПбГУ ИТМО, 2009. – 363 с.
11. Gross D., Shortle J. F., Thompson J. M., Harris C. M. Fundamentals of Queueing Theory. 5th ed. Hoboken: Wiley, 2018. 576 p. DOI: <https://doi.org/10.1002/9781119453765>.