

УДК 004.9
DOI

ВЛИЯНИЕ НЕСТАЦИОНАРНОГО ПОТОКА НА ЭФФЕКТИВНОСТЬ МНОГОФАЗНОЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ПРИБОРАМИ РАЗЛИЧНОЙ ПРОИЗВОДИТЕЛЬНОСТИ

Габдрахманов Б.М., Самерханов И.З., Нуриев Н.К.

ФГБОУ ВО «Казанский национальный исследовательский технологический университет»,
Казань, e-mail: gbm445@yandex.ru

Современные системы массового обслуживания часто функционируют в условиях динамически изменяющейся нагрузки, что требует новых подходов к их проектированию. В статье исследуется влияние нестационарного входного потока на эффективность многофазных систем массового обслуживания с приборами различной производительности. Цель работы – выявление оптимальных соотношений между нагрузкой на систему и распределением производительности обслуживающих приборов для максимизации пропускной способности и минимизации вероятности отказов. Методологической основой исследования является имитационное моделирование в среде AnyLogic, позволяющее анализировать поведение системы в условиях, приближенных к реальным. В качестве входных данных использовались статистические показатели трафика Wikipedia, отражающие типичные суточные колебания нагрузки. Результаты исследования демонстрируют, что умеренная диспропорция в производительности приборов (20–30%) способствует повышению устойчивости системы при пиковых нагрузках. Установлено, что такие конфигурации обеспечивают значительное улучшение ключевых показателей эффективности по сравнению с однородными системами. Напротив, экстремальная неоднородность производительности приводит к формированию «узких мест» и снижению общей эффективности. Практическая значимость работы заключается в разработке рекомендаций для проектирования адаптивных систем массового обслуживания в IT-инфраструктуре, логистических комплексах и производственных системах, где критически важна минимизация простоев и отказов. Полученные результаты открывают перспективы для дальнейших исследований в области динамической балансировки нагрузки.

Ключевые слова: системы массового обслуживания, нестационарный поток, имитационное моделирование, приборы различной производительности, вероятность отказа, оптимизация нагрузки, динамическая адаптация

IMPACT OF NON-STATIONARY FLOW ON PERFORMANCE OF MULTIPHASE QUEUEING SYSTEMS WITH HETEROGENEOUS SERVERS

Gabdrakhmanov B.M., Samerkhanov I.Z., Nuriev N.K.

Kazan national research technological university, Kazan, e-mail: gbm445@yandex.ru

Modern queuing systems often operate under dynamically changing workloads, necessitating innovative design approaches. This study investigates the impact of non-stationary input flows on the performance of multiphase queueing systems with heterogeneous servers. The research aims to identify optimal relationships between system load and server performance distribution to maximize throughput and minimize rejection probability. The methodological framework is based on simulation modeling in the AnyLogic environment, enabling analysis of system behavior under real-world conditions. Input data were derived from Wikipedia traffic statistics, reflecting typical diurnal load fluctuations. The results demonstrate that moderate performance disparities (20–30%) between servers enhance system resilience during peak loads. Such configurations significantly improve key efficiency metrics compared to homogeneous systems. Conversely, extreme performance heterogeneity creates bottlenecks and reduces overall effectiveness. The practical significance of this work lies in its recommendations for designing adaptive queuing systems in IT infrastructure, logistics hubs, and production lines, where minimizing downtime and rejections is critical. The findings open new avenues for research in dynamic load balancing.

Keywords: queueing systems, non-stationary flow, heterogeneous service units, simulation modeling, rejection probability, load optimization, dynamic adaptation

Введение

Интенсивность входного потока представляет собой ключевую характеристику любой системы массового обслуживания (СМО), определяя количество заявок, поступающих в систему за единицу времени, и формируя тем самым ее нагрузку. В классических подходах к моделированию СМО традиционно предполагается стационарность входного потока, когда его интенсивность остается постоянной во времени ($\lambda = \text{const}$).

Такое допущение действительно для многих процессов на ограниченных временных интервалах – например, для производственных линий [1] в течение рабочей смены. Однако в более широком временном контексте большинство реальных систем демонстрирует выраженную нестационарность поведения, обусловленную влиянием внешних факторов: сезонными колебаниями спроса, изменением деловой активности, человеческим фактором и другими динамическими процессами.

Современные исследования [2–4] все чаще обращаются к моделям СМО с нестационарными потоками, где интенсивность поступления заявок является функцией времени $\lambda(t)$. Подобные модели находят широкое применение при проектировании и анализе: телекоммуникационных сетей (суточные колебания трафика), транспортных систем (пассажиропоток в метро), медицинских учреждений (сезонность обращений), логистических центров (предпраздничный рост заказов), производственных процессов (изменение спроса на продукцию).

Особый практический интерес представляют системы, сочетающие два важных свойства:

- 1) нестационарный характер входного потока;
- 2) неоднородность и многофазность обслуживающих приборов по производительности.

В данной работе для моделирования нестационарного потока использовались реальные данные о запросах к серверам Wikipedia за 16 апреля 2025 года. Эти данные демонстрируют типичные суточные колебания трафика: минимальная нагрузка составила 428 553 запроса в час (ночные часы), а пиковая – 1 417 977 запросов в час (утренний период активности). Такой подход позволяет более точно оценить эффективность СМО в условиях, приближенных к реальным. Как показывают исследования [5–7], различие в скорости обработки заявок между приборами может существенно влиять на эффективность системы в целом. При этом оптимальное распределение производительности приборов обслуживания позволяет улучшить ключевые показатели эффективности СМО. Однако большинство существующих работ рассматривает либо нестационарные потоки в однородных системах, либо стационарные потоки в системах с разнородными приборами [8–10].

Настоящее исследование направлено на восполнение этого пробела путем комплексного анализа многофазных СМО, функционирующих в условиях:

- динамически изменяющегося входного потока;
- существенной разницы в производительности обслуживающих приборов;
- необходимости оптимального распределения заявок между фазами.

Цель работы – выявление оптимальных соотношений между нагрузкой на систему и соотношением производительностей приборов в многофазных СМО, обеспечивающих максимизацию пропускной способности системы и минимизацию вероятности отказов. Полученные результаты могут

иметь важное прикладное значение для проектирования адаптивных систем обслуживания в различных отраслях экономики.

Материалы и методы исследования

Аналитическое исследование многофазных систем массового обслуживания с нестационарным входным потоком и приборами различной производительности представляет значительные математические сложности. Для таких СМО не существует точных аналитических решений в явном виде, что обусловлено: нелинейностью дифференциальных уравнений Колмогорова [1]; зависимостью параметров от времени; сложным взаимодействием разнородных обслуживающих приборов [10–12].

В связи с этим в 2025 году на базе Казанского национального исследовательского технологического университета был проведен комплекс имитационных экспериментов с использованием программного комплекса AnyLogic. Данная среда обладает рядом ключевых преимуществ: поддержка дискретно-событийного подхода; возможность задания нестационарных параметров; гибкая настройка логики обслуживания; встроенные средства сбора статистики; визуализация процессов в реальном времени. На рисунке 1 представлена модель многофазной СМО, состоящая из стандартных объектов среды AnyLogic [12].

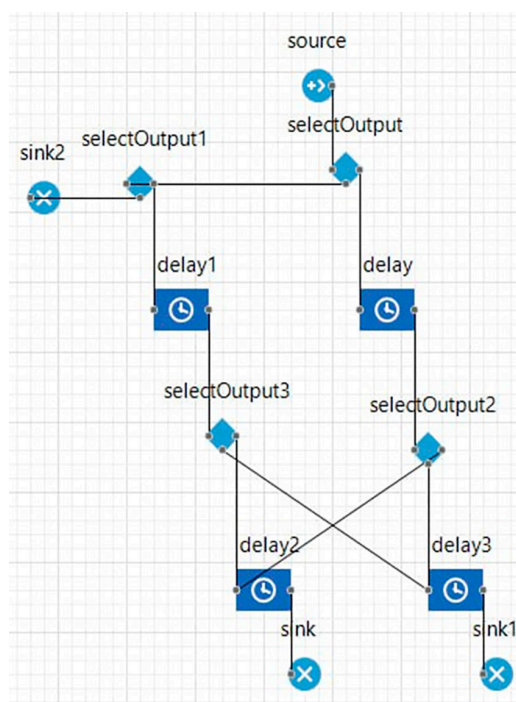


Рис. 1. Модель многофазной СМО
Примечание: составлено авторами по результатам моделирования



Рис. 2. Гистограмма распределения интенсивности входного потока ($\times 10^3$ заявок/час)

Примечание: составлено авторами на основе данных о часовой нагрузке серверов Wikipedia за 16 апреля 2025 года

Конфигурации многофазной СМО с распределением производительности приборов ($\mu_1, \mu_2, \mu_3, \mu_4$)

		μ_1	μ_2	μ_3	μ_4
Приборы одинаковой производительности	СМО-0	500	500	700	700
Приборы различной производительности	СМО-1	450	550	650	750
	СМО-2	400	600	600	800
	СМО-3	350	650	550	850
	СМО-4	300	700	450	950
	СМО-5	200	800	400	1000
	СМО-6	150	850	350	1050
	СМО-7	100	900	250	1150
	СМО-8	50	950	100	1300

Примечание: составлено авторами по результатам вычислительных экспериментов.

В качестве основы для моделирования нестационарного потока использовались реальные данные о нагрузке на серверы Wikipedia за 16 апреля 2025 года [13]. Расписание интенсивности входного потока представлено на рисунке 2.

Таким образом, в исследуемой системе интенсивность входного потока изменяется в течение суток в диапазоне от 428,553 до 1417,977 заявок в час. Подобное поведение характерно для многих реальных систем, таких как серверные кластеры крупных интернет-ресурсов (на примере Wikipedia) или высоконагруженные call-центры [14; 15]. С разработанной моделью проведена серия вычислительных экспериментов, направленных на сравнительный анализ эффективности многофазных СМО с различными распределениями производительностей обслуживающих приборов.

В качестве ключевых показателей эффективности были выбраны: P_0 – вероятность свободного состояния системы и $P_{отк}$ – вероятность отказа в обслуживании.

Для однозначности сравнения введены следующие обозначения:

- P_{01} и $P_{отк1}$ – показатели для базовой конфигурации (СМО-0) с приборами одинаковой производительности;

- P_{02} и $P_{отк2}$ – показатели для тестируемых конфигураций (СМО-1...СМО-8) с различной производительностью приборов.

При моделировании работы систем с неоднородными приборами применялся алгоритм динамического распределения нагрузки, согласно которому входящие заявки направляются на свободный прибор с максимальной производительностью. Данный подход доказал свою эффективность в предыдущих исследованиях для систем без очереди.

Результаты исследования и их обсуждение

Проведенные вычислительные эксперименты охватывали анализ девяти различных конфигураций многофазных СМО, характеристики которых представлены в таблице.

Каждая конфигурация отличается уникальным распределением производительностей обслуживающих приборов при сохранении постоянной суммарной мощности системы.

Результаты анализа вероятности свободного состояния (P_0)

Для оценки влияния неоднородности производительности приборов на эффективность системы был проведен сравнительный анализ изменения вероятности свободного состояния P_0 относительно базовой конфигурации (СМО-0). Результаты представлены на рисунке 3.

Методика расчета

Для каждой конфигурации (от СМО-1 до СМО-8) вычислялось отклонение по формуле:

$$\Delta P_0 = P_{0_i} - P_{0_0},$$

где $P_{0_0} = 0,341$ (базовое значение для СМО-0).

Полученные значения были визуализированы в виде столбчатой диаграммы. На оси X отображены номера конфигураций (СМО-1 – СМО-8), а на оси Y – значения ΔP_0 . Для наглядности использовано цветовое разделение: зеленый обозначает положительные значения, а красный – отрицательные.

Ключевые результаты (рис. 2)

Наибольшая положительная эффективность ($\Delta P_0 > 0$) наблюдается в конфигурациях СМО-1 – СМО-4, где диспропорция составляет 10–40%. Максимальный прирост достигнут в СМО-2 и составляет +0,024.

Переходная зона включает СМО-4 с $\Delta P_0 = +0,020$ и СМО-5, где отклонение становится отрицательным ($\Delta P_0 = -0,013$).

При дальнейшем увеличении диспропорции (более 50%) наблюдается резкое снижение эффективности. Наихудший результат зафиксирован в СМО-8 с минимальным значением $\Delta P_0 = -0,186$.

Результаты анализа вероятности отказа ($P_{отк}$)

На рисунке 4 представлены результаты сравнительного анализа изменения вероятности отказа для различных конфигураций системы. График отражает зависимость $\Delta P_{отк}$ от степени неоднородности производительности приборов.

Методика расчета

Для каждой конфигурации определялось отклонение по формуле:

$$\Delta P_{отк} = P_{отк_0} - P_{отк_i},$$

где $P_{отк_0} = 0,226$ (базовое значение для СМО-0).

Результаты были визуализированы на диаграмме, где по оси X указаны номера конфигураций (СМО-1 – СМО-8), а по оси Y – значения $\Delta P_{отк}$. Положительные значения выделены синим цветом, а отрицательные – оранжевым.

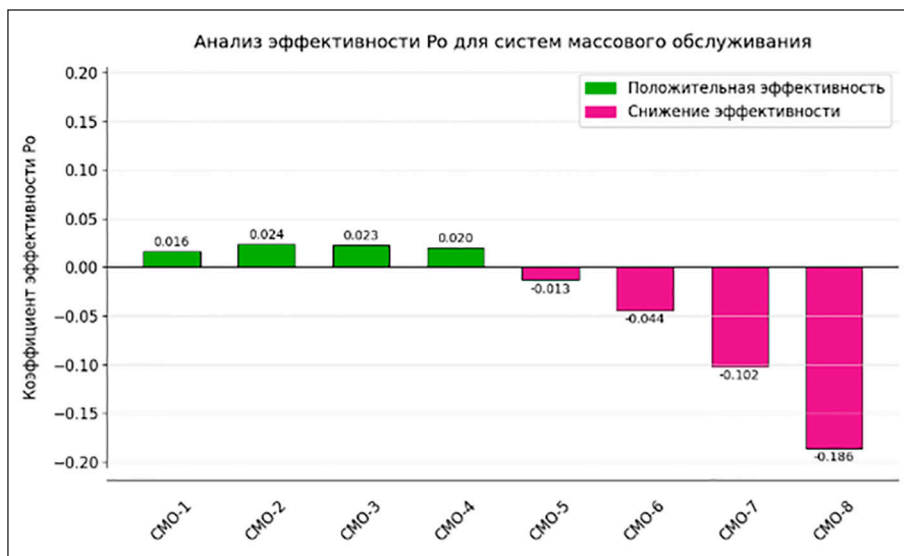


Рис. 3. Зависимость изменения вероятности свободного состояния (ΔP_0) от степени неоднородности производительности приборов в многофазной СМО
Примечание: составлено авторами по результатам данного исследования

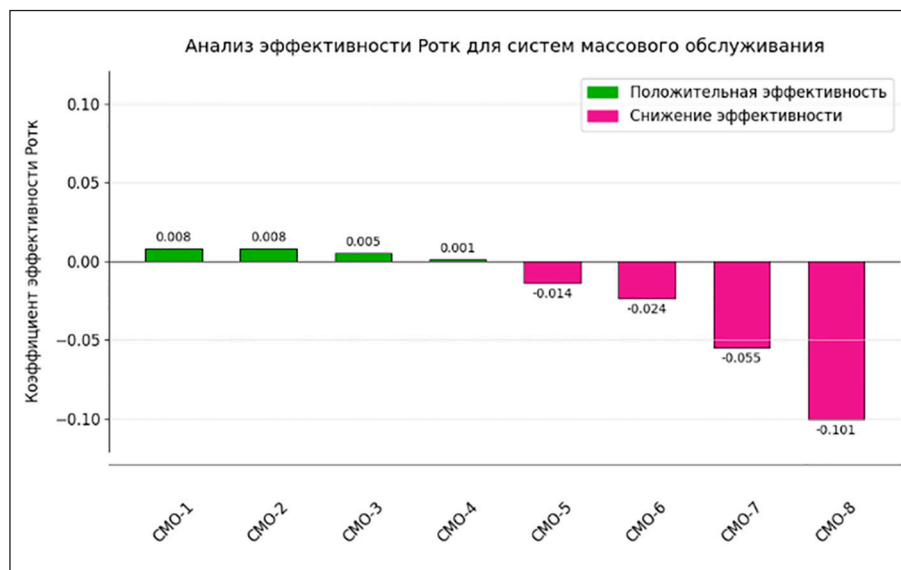


Рис. 4. Зависимость изменения вероятности отказа ($\Delta P_{отк}$) от степени неоднородности производительности приборов
Примечание: составлено авторами по результатам данного исследования

Выводы

Проведенное исследование многофазных систем массового обслуживания с нестационарным входным потоком и приборами различной производительности позволило сделать следующие ключевые выводы.

Оптимальная конфигурация

Умеренная неоднородность производительности приборов (20-30%) демонстрирует максимальную эффективность. Для исследованных параметров наилучшие показатели достигнуты в конфигурациях СМО-2 и СМО-3, где наблюдается увеличение вероятности нулевой очереди (P_0) на 4.7-7.0% и снижение вероятности отказа ($P_{отк}$) на 3.5-5.2%.

Критические ограничения

Диспропорция производительности свыше 50% приводит к резкому ухудшению характеристик систем: снижению P_0 до 18.6% и росту вероятности отказа до 10.1%. В таких условиях формируются «узкие места» в фазах с минимальной производительностью, что негативно влияет на общую эффективность системы.

Практические рекомендации

Для систем с нестационарной нагрузкой рекомендуется обеспечивать адаптивное распределение нагрузки и избегать экстремальных диспропорций в производительности устройств. Оптимальная суммарная производительность системы должна на 15-20% превышать пиковую нагрузку для обеспечения устойчивой работы.

Практическая ценность

Полученные результаты имеют важное значение для проектирования IT-инфраструктуры (серверные кластеры, облачные системы), производственных и логистических комплексов, телекоммуникационных сетей нового поколения, а также систем обработки данных в реальном времени.

Исследование подтвердило гипотезу о существовании оптимального уровня неоднородности производительности приборов, обеспечивающего максимальную эффективность многофазных СМО в условиях нестационарных нагрузок.

Список литературы

1. Самерханов И.З. Нестационарный поток в системе массового обслуживания с приборами различной производительности // Современные наукоемкие технологии. 2025. № 4. С. 76-81. URL: <https://top-technologies.ru/ru/article/view?id=40368> (дата обращения: 19.04.2025). DOI: 10.17513/snt.40368.
2. Грачев В.В., Моисеев А.Н., Назаров А.А., Ямпольский В.З. Многофазная модель массового обслуживания системы распределенной обработки данных // Доклады Томского государственного университета систем управления и радиоэлектроники. 2012. № 2-2 (26). С. 248-251. EDN: PXPJMZ. URL: <https://www.elibrary.ru/item.asp?edn=lbtpbc> (дата обращения: 18.04.2025). DOI: 10.17513/snt.38467.
3. Печеный Е.А., Самерханов И.З., Нурiev Н.К. Модель управления системой массового обслуживания с неэквивалентными каналами // Современные наукоемкие технологии. 2022. № 4. С. 83-88. URL: <https://top-technologies.ru/ru/article/view?id=39112> (дата обращения: 18.04.2025). DOI: 10.17513/snt.39112.
4. Самерханов И.З. О влиянии разделения каналов различной производительности на показатели эффективности системы массового обслуживания // Сборник трудов одиннадцатой всероссийской научно-практической конферен-

ции по имитационному моделированию и его применению в науке и промышленности «Имитационное моделирование. Теория и практика» (ИММОД-2023). Казань: Издательство АН РТ, 2023. С. 698-703. URL: <https://www.elibrary.ru/item.asp?edn=qxlynz> (дата обращения: 19.04.2025). EDN: QXLYNZ.

5. Назаров А.А., Рожкова С.В., Титаренко Е.Ю. Исследование системы с обратной связью, рекуррентным обслуживанием и неординарным пуассоновским входящим потоком // Информационные технологии и математическое моделирование (ИТММ-2020): Материалы XIX Международной конференции имени А.Ф. Терпугова. Томск: Изд-во науч.-техн. лит., 2021. С. 223-227. URL: <https://www.elibrary.ru/item.asp?id=45668279> (дата обращения: 19.04.2025). EDN: QFYARU.

6. Xu J., Liu L., Zhu T. Transient Analysis of Two-Heterogeneous Server Queue with Impatient Behavior and Multiple Vacations // J. Systems Science and Information. 2018. Vol. 6, Is. 1. P. 69-84. URL: <https://ideas.repec.org/a/bpj/jossai/v6y2018i1p69-84n5.html> (дата обращения: 20.04.2025). DOI: 10.21078/JSSI-2018-069-16.

7. Саати Т.Л. Элементы теории массового обслуживания и ее приложения. М.: Советское радио, 1971. 520 с. С. 120-145. ISBN: 978-5-397-01283-6.

8. Ершов Д.С., Хайруллин Р.З. Математическая модель рабочего места проверки средств измерений как нестационарная система обслуживания // Известия высших учебных заведений. Приборостроение. 2022. Т. 65. № 10. С. 701-711. URL: <https://www.elibrary.ru/item.asp?edn=csghkq> (дата обращения: 20.04.2025). DOI: 10.17586/0021-3454-2022-65-10-701-711. EDN: CSGHKQ.

9. Friedman J.H. Stochastic Gradient Boosting // Computational Statistics and Data Analysis. 2002. Vol. 38. № 4. P. 367-

378. URL: <https://www.bohrium.com/paper-details/stochastic-gradient-boosting/813056516323016704-964> (дата обращения: 20.04.2025). DOI: 10.1016/S0167-9473(01)00065-2.

10. Khayyati S., Tan B. Supervised-Learning-Based Approximation Method for Multi-server Queueing Networks under Different Service Disciplines with Correlated Interarrival and Service Times // International Journal of Production Research. 2022. Vol. 60. № 17. P. 5176-5200. URL: <https://www.tandfonline.com/doi/full/10.1080/00207543.2021.1951448> (дата обращения: 19.04.2025). DOI: 10.1080/00207543.2021.1951448.

11. Kumar B.K., Sankar R., Krishnan R.N., Rukmani R. Performance Analysis of Multi-processor Two-Stage Tandem Call Center Retrial Queues with Non-Reliable Processors // Methodology and Computing in Applied Probability. 2022. Vol. 24. № 1. P. 95-142. URL: https://ideas.repec.org/a/spr/metcap/v24y2022i1d10.1007_s11009-020-09842-6.html (дата обращения: 20.04.2025). DOI: 10.1007/s11009-020-09842-6.

12. AnyLogic 8.7 Professional: User Guide. Saint Petersburg: The AnyLogic Company, 2024. 415 p. P. 210-235.

13. Wikipedia Traffic Statistics. April 2025 Dataset. Wikimedia Foundation, 2025. URL: <https://stats.wikimedia.org> (дата обращения: 18.04.2025).

14. Sridhar A., Pitchai R. Analysis of a Markovian Queue with Two Heterogeneous Servers and Working Vacation // Intern. J. Applied Operational Research. 2015. Vol. 5. Is. 4. P. 1-15. URL: <https://ijorlu.liau.ac.ir/article-1-458-en.html> (дата обращения: 18.04.2025).

15. Язвенко М.Р., Морозков А.Г. Моделирование морского грузового порта как системы массового обслуживания в среде AnyLogic // Системный анализ и логистика. 2020. № 4 (26). С. 59-66. URL: <https://www.elibrary.ru/item.asp?edn=pcabyr> (дата обращения: 20.04.2025). DOI: 10.31799/2007-5687-2020-4-59-66.