МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра системного анализа и автоматического управления

МЕТОДЫ УПРАВЛЕНИЯ СЕТЯМИ МАССОВОГО ОБСЛУЖИВАНИЯ С ГРУППОВЫМ ОБСЛУЖИВАНИЕМ

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студентки 2 курса 271 группы
направления 09.04.01 — Информатика и вычислительная техника
факультета КНиИТ
Гурковой Виктории Марковны

Научный руководитель доцент, к. ф.-м. н.	____________	Е. П. Станкевич
Заведующий кафедрой к. ф.-м. н., доцент	____________	И. Е. Тананко

Саратов 2023

ВВЕДЕНИЕ

Актуальность темы. В настоящее время системы и сети массового обслуживания обширно используются в качестве моделей реальных систем при решении задач исследования, проектирования и оптимизации вычислительных систем и сетей передачи данных, производственных и транспортных систем. Широкое практическое применение систем и сетей массового обслуживания [1-5] обусловлено удобством и простотой отображения структуры систем, которые они моделируют.

Сети массового обслуживания также часто применяются для моделирования систем с дискретным временем, в частности, телекоммуникационных систем, компьютерных систем и сетей, а также при решении задач распределения ресурсов и задач анализа производственных систем и процессов, когда необходимо учитывать групповую обработку, производство и перемещение сырья, материалов и готовой продукции. В современных экономических условиях эффективная организация вышеперечисленных процессов невозможна без построения и анализа моделей систем и процессов в виде систем и сетей массового обслуживания с *групповым обслуживанием требований*. Использование соответствующих моделей способствует решению широкого класса задач анализа отдельных элементов сложных стохастических систем, таких как, например, рабочие станки и устройства с одновременной обработкой нескольких деталей, погрузочно-разгрузочные машины, транспортеры, грузовые автомобили и других.

Таким образом, исследование и разработка методов анализа и управления сетями массового обслуживания с групповым обслуживанием требований являются значимым направлением исследований [6-8], что подтверждает актуальность, теоретическую и практическую значимость настоящей работы.

Целью настоящей работы является разработка методов управления сетями массового обслуживания с групповым обслуживанием требований, а также разработка программного комплекса для исследования сети массового обслуживания с групповым обслуживанием и управлением.

Поставленная цель определила **следующие задачи**:

— изучить основные результаты исследований по системам и сетям массового обслуживания с групповым обслуживанием требований;

— описать математическую модель и метод анализа сети массового обслу-

живания с групповым обслуживанием требований;
- сформулировать алгоритмы разработанных методов управления сетью массового обслуживания с групповым обслуживанием требований;
- реализовать комплекс программ для численного анализа и имитационного моделирования сети с групповым обслуживанием и управлением;
- провести численные эксперименты с применением разработанного комплекса программ.

Методологические основы исследования систем и сетей массового обслуживания с групповым обслуживанием требований представлены в работах П. П. Бочарова, Е. П. Станкевич, Г. Ш. Цициашвили, N. T. J. Bailey, F. Downton, W. Henderson, R. J. Boucherie, X. Chao, A. Krishnamoorthy.

Теоретическая значимость магистерской работы. Рассмотренная модель сети массового обслуживания с групповым обслуживанием требований расширяет круг задач, решаемых в теории массового обслуживания, поскольку позволяет рассмотреть особенности структуры и функционирования сетей с групповым обслуживанием и управлением, в которых требования после завершения обслуживания перемещаются по сети независимо друг от друга. Помимо этого, в работе получен результат для системы массового обслуживания с групповым обслуживанием, состоящей из неограниченной очереди и одного обслуживающего прибора, а именно, выражение для интенсивности потока, когда группа состоит из двух требований, при которой математическое ожидание длительности пребывания требований в системе минимально.

Практическая значимость магистерской работы. Представленные в работе результаты могут быть применены для исследования, оптимизации и построения реальных стохастических сетевых систем с групповой обработкой объектов, таких как, например, телекоммуникационные и вычислительные системы, производственные и торговые системы, транспортные сети. Предложенная математическая модель может быть применена для расчёта характеристик телекоммуникационных систем и сетей, управления движением групп автомобилей в транспортной сети, описания перемещений и распределения партий товаров между складами в торговой или производственной системе с целью повышения эффективности её функционирования.

Структура и объем работы. Магистерская работа состоит из введения, пяти разделов, заключения, списка использованных источников и цифрового носителя в качестве приложения. Общий объём работы — 64 страницы, из них 56 страниц — основное содержание, включая 8 рисунков и 6 таблиц, 1 страница приложения, список использованных источников — 37 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Обзор результатов исследований систем и сетей массового обслуживания с групповым обслуживанием» посвящён обзору основных опубликованных результатов исследований систем и сетей массового обслуживания с групповым обслуживанием требований. В разделе приводится краткий обзор методов анализа, управления и расчёта характеристик систем и сетей массового обслуживания с групповым обслуживанием. Из обзора представленных работ, можно сделать вывод о том, что системы и сети массового обслуживания с групповым обслуживанием требований являются актуальным предметом исследований и представляют собой эффективный инструмент для анализа широкого класса реальных систем.

Второй раздел «Метод анализа сети массового обслуживания с групповым обслуживанием» посвящён описанию математической модели и метода анализа сети массового обслуживания с групповым обслуживанием требований без управления.

В подразделе 2.1 приведено описание математической модели и метода анализа сети массового обслуживания с групповым обслуживанием требований. Исследуемая сеть состоит из L систем массового обслуживания S_i, $i = 1, \ldots, L$. Каждая система обслуживания S_i состоит из одного обслуживающего прибора и очереди бесконечной длины. В сеть из источника S_0 поступает пуассоновский поток требований с интенсивностью λ_0. В случае, если обслуживающий прибор системы занят, то поступающие требования отправляются на ожидание в очередь системы.

Требования обслуживаются на приборах группами фиксированного размера. Для каждой системы S_i задан размер обслуживаемой группы b_i, $b_i \geq 1$, $i = 1, \ldots, L$. Так, если в очереди системы находится меньше b_i требований, то обслуживающий прибор будет простаивать до тех пор, пока в очереди не окажется b_i требований. После этого прибор системы незамедлительно начинает обслуживание группы, состоящей из b_i требований. Причём, выбираемая для обслуживания группа из b_i требований, формируется случайным образом.

Длительность обслуживания группы b_i требований в системе S_i есть экспоненциально распределенная случайная величина с параметром μ_i, $i = 1, \ldots, L$. Перемещения требований между системами обслуживания и источником производятся согласно маршрутной матрице $\Theta = (\theta_{ij})$, $i, j = 0, 1, \ldots, L$, где θ_{ij} — вероятность перехода требования из системы S_i в систему S_j. Важной особенностью рассматриваемой сети является тот факт, что после завершения обслуживания группы требований в системе S_i, каждое отдельное требование группы, независимо от остальных, направляется в систему S_j с вероятностью θ_{ij}, $i, j = 0, 1, \ldots, L$.

Предполагается также, что рассматриваемая сеть имеет большую размерность, и число смежных с системой S_i выходных систем значительно превышает размер обслуживаемой в системе S_i группы требований b_i.

Состояние сети определяется вектором $s = (s_1, \ldots, s_L)$, где s_i — число требований, находящихся в системе S_i.

Стационарные вероятности $\pi(s)$ состояний сети N, определяются выражением

$$\pi(s) = \prod_{i=1}^{L} \pi_i(s_i),\ s \in X,$$

где

$$\pi_i(s_i) = \pi_i(0) \prod_{n=1}^{s_i} \frac{\lambda_i}{\widetilde{\mu}_i(n)},$$

где $\widetilde{\mu}_i(n)$ — интенсивность перехода процесса размножения и гибели из состояния n в состояние $n-1$, $n \in \{1, 2, \ldots\}$.

Математическое ожидание (м. о.) числа требований в системе S_i, $i = 1, \ldots, L$, определяется выражением

$$\bar{s}_i = \sum_{n=1}^{\infty} n\pi_i(n),$$

а м. о. длительности пребывания требований в сети имеет вид

$$\bar{\tau} = \frac{1}{\lambda_0} \sum_{i=1}^{L} \bar{s}_i.$$

Подраздел 2.2 посвящён рассмотрению изолированной системы массового обслуживания S_i, $i = 1, \ldots, L$, с групповым обслуживанием требований, состоящей из неограниченной очереди и одного обслуживающего прибора в случае, когда в системе может обслуживаться группа, состоящая из двух требований. Для описанной системы в работе получено выражение для интенсивности входящего потока, когда группа состоит из двух требований, при которой математическое ожидание длительности пребывания требований в системе минимально в виде следующей теоремы [9].

Теорема 2.1. Для системы массового обслуживания S_i с обслуживанием группы, состоящей из $b_i = 2$ требований, математическое ожидание длительности пребывания требований в системе минимально при $\lambda_i = 0,75\mu_i$.

Третий раздел «Методы управления сетью массового обслуживания с групповым обслуживанием» посвящён описанию методов управления сетью массового обслуживания с групповым обслуживанием требований.

В подразделе 3.1 описана постановка задачи управления входящим потоком, заключающаяся в определении последовательности интенсивностей входящего потока, которые обеспечивают максимальную пропускную способность сети при заданном ограничении на математическое ожидание длительности пребывания требований в сети. Идея данного метода управления основана на результатах работы [10].

Первым этапом решения поставленной задачи является переход от открытой сети массового обслуживания к замкнутой. Так, полученная замкнутая сеть содержит большое число N требований и состоит из двух систем. Первая система представляет собой систему типа $M|M|1$ с интенсивностью обслуживания λ_{N-n}, где N — общее число требований в сети, а n — число требований, находящихся во второй системе. Вторая система представляет собой сеть массового обслуживания с групповым обслуживанием требований.

Следующим шагом необходимо агрегировать сеть массового обслуживания с групповым обслуживанием требований. Так, описанную выше сеть массового обслуживания с групповым обслуживанием требований можно свести к системе типа $M|M|1$ с интенсивностью обслуживания μ_n, зависящей от числа требований n, находящихся в системе, методом агрегирования, основанном на теореме Нортона. После применения метода агрегирования, имеем замкнутую сеть массового обслуживания, состоящую из двух систем S_0 и S_1, в которой находится большое число N требований.

Согласно теореме Нортона, для полного определения эквивалентной сети необходимо вычислить (аналитически или с помощью имитационной модели) или измерить интенсивности обслуживания $\mu_1, \mu_2, \ldots, \mu_n$ агрегированной сети S_1.

После агрегирования сети массового обслуживания с групповым обслуживанием в систему $M|M|1$ и определения её интенсивностей обслуживания $\mu_1, \mu_2, \ldots, \mu_n$, можно перейти к определению последовательности интенсивностей входящего потока $\lambda_{N-1}, \lambda_{N-2}, \ldots, \lambda_1, \lambda_0$, которые обеспечивают максимальную пропускную способность сети при заданном ограничении на математическое ожидание длительности пребывания требований в сети $\overline{\tau}$.

Одним из простейших способов нахождения последовательности λ_{N-1}, $\lambda_{N-2}, \ldots, \lambda_1, \lambda_0$ является *метод перебора*. Помимо метода перебора, для решения поставленной задачи также может быть применён *метод деления отрезка пополам*, алгоритм которого подробно описан в работе.

Подраздел 3.2 посвящён описанию метода управления, при котором осуществляется изменение маршрутной матрицы сети исходя из оптимального коэффициента использования для каждой системы сети. Идея данного метода управления основана на результатах работ [11,12]. Целью данного метода управления является достижение минимального значения математического ожидания длительности пребывания требований в сети $\bar{\tau}$. Алгоритм предлагаемого метода управления содержит следующие шаги:

1. Вычислить значения $\psi_{i(opt)}$ для каждой системы S_i, $i = 1, \ldots, L$, сети и начальное $\bar{\tau}$.
2. Сформировать матрицу $D = (d_{ij})$, $i = 1, \ldots, L$, $j = 0, 1, \ldots, L$, элементы которой вычисляются согласно выражению
$$d_{ij} = |\psi_{i(opt)} - \psi_i|,$$
для всех $i = 1, \ldots, L$, для которых $\psi_i \geq \psi_{i(opt)}$.
3. Вычислить новые элементы маршрутной матрицы Θ^*, используя выражение
$$\theta^*_{ij} = \theta_{ij} \frac{d_{ij}}{\sum_{=1}^{L} d_{ij}}.$$
4. Провести процедуру нормализации полученной маршрутной матрицы

$\Theta^* = (\theta^*_{ij})$, $i, j = 0, 1, \ldots, L$.

5. Если число систем, коэффициент использования ψ_i которых превышает оптимальное значение $\psi_{i(opt)}$, больше 1 — вернуться к шагу 2. Иначе перейти к шагу 6.
6. Завершить работу алгоритма и вычислить итоговое значение $\bar{\tau}^*$.

После того, как получена новая матрица Θ^* и проведена процедура нормализации, вычисляется значение математического ожидания длительности пребывания требований в сети $\overline{\tau}^*$ и сравнивается со значением $\overline{\tau}$ при предыдущей матрице Θ.

В подразделе 3.3 описан метод управления сетью массового обслуживания с групповым обслуживанием, основанный на определении количества требований в очереди каждой смежной системы. Алгоритм работы предлагаемого метода представлен в двух вариантах.

В первом варианте, который далее для краткости обозначен как «Кратчайшая очередь», для системы S_i, в которой происходит завершение обслуживания требования, с помощью маршрутной матрицы Θ определяется множество номеров смежных систем $J = \{j \in J : \theta_{ij} \neq 0,\ j = 1, \ldots, L\}$. Далее для каждой системы S_j, $j \in J$ определяется множество длин очередей $Q = \{q_j \in Q : j \in J\}$, где q_j есть количество требований в очереди смежной системы S_j. После определения количества требований, находящихся в очереди каждой смежной системы, осуществляется выбор очереди минимальной длины $q_{min} = \min Q$. Таким образом, после завершения обслуживания в системе S_i требование отправится на обслуживание в систему S_j с наименьшим количеством q_{min} требований в очереди среди смежных систем.

Во втором варианте предлагаемого метода, который для краткости обозначен как «Умный выбор», помимо множества номеров смежных с системой S_i систем J и множества длин очередей Q определяется множество значений интенсивностей обслуживания $M = \{\mu_j \in M : j \in J\}$, где μ_j есть значение интенсивности обслуживания в смежной системе S_j. Далее формируется рейтинг смежных систем S_j, исходя из значений интенсивностей μ_j в виде убывающей последовательности $\{\mu_k\} \Leftrightarrow \forall k \in \{1, \ldots, |J| - 1\} : \mu_k \geq \mu_{k+1}$. После определения последовательности $\{\mu_k\}$ находим среди смежных систем с наибольшими значениями интенсивности обслуживания требований систему S_k с минимальным числом требований в очереди q_{min}. Таким образом, мы исключаем возможность того, что смежная система с наименьшим количеством требований в очереди окажется самой «непроизводительной» из имеющихся систем.

Четвёртый раздел «Комплекс программ для численного анализа и имитационного моделирования сети массового обслуживания с групповым обслуживанием» посвящён подробному описанию программ, разработанных для численного анализа и имитационного моделирования сети массового обслуживания с групповым обслуживанием требований. Комплекс программ разработан на языке программирования *Python 3.10*.

В подразделе 4.1 приведено описание программы для численного анализа сети массового обслуживания с групповым обслуживанием, в которой представлены методы, реализующие вычисление основных стационарных характеристик сети, а также метода управления, основанного на преобразовании элементов маршрутной матрицы сети.

Подраздел 4.2 посвящён подробному описанию программы для имитационного моделирования рассматриваемой сети обслуживания. При реализации имитационной модели сети использовался дискретно-событийный подход с методом продвижения модельного времени от события к событию. Требования, системы, группа и источник в имитационной модели представлены объектами некоторого класса, наделённого существенными для данной модели значениями. В программе также реализованы обработчики событий и часы модельного времени для дискретно-событийного имитационного моделирования.

Пятый раздел «Численные примеры» посвящён описанию численных экспериментов, суть которых заключается в вычислении математического ожидания длительности пребывания требований в сети и его оптимизации посредством применения методов управления, основанных на преобразовании маршрутной матрицы сети и определении смежных систем с наименьшим количеством требований в очереди.

В подразделе 5.1 приведено описание результатов работы программы для численного анализа сети массового обслуживания с групповым обслуживанием требований до и после применения метода управления, основанного на преобразовании элементов маршрутной матрицы сети. На основании проведённых экспериментов можно сделать вывод о том, что предложенный метод управления, основанный на изменении элементов маршрутной матрицы сети, оптимизирует математическое ожидание длительности пребывания требований в сети в среднем на 27%.

Подраздел 5.2 посвящён описанию результатов работы имитационной модели сети массового обслуживания с групповым обслуживанием требований до и после применения метода управления, основанного на информации о длинах очередей смежных систем. Длительность моделирования составляет $n = 1\ 000\ 000$ единиц модельного времени. Обозначим количество требований, поступивших в сеть за время моделирования как T. Тогда, при заданном количестве единиц модельного времени n в сеть поступит около $T = 800\ 440$ требований.

На основании полученных результатов можно сделать вывод о том, что метод «Кратчайшая очередь» оптимизирует значение математического ожидания длительности пребывания требований в сети в среднем на 26.9%, а метод «Умный выбор» — на 30.6%. В сравнении с методом управления, основанном на преобразовании элементов маршрутной матрицы, наименее эффективным оказался метод «Кратчайшая очередь». Следует отметить, что под эффективностью понимается степень уменьшения математического ожидания длительности пребывания требований в сети по сравнению со значением без применения управления.

ЗАКЛЮЧЕНИЕ

Таким образом, в настоящей работе были получены следующие основные результаты:

1. Описана математическая модель и метод анализа сети массового обслуживания с групповым обслуживанием требований без управления.
2. Получен результат для системы массового обслуживания с групповым обслуживанием, состоящей из неограниченной очереди и одного обслуживающего прибора, а именно, выражение для интенсивности потока, когда группа состоит из двух требований, при которой математическое ожидание длительности пребывания требований в системе минимально.
3. Описана постановка задачи управления входящим потом, заключающаяся в определении последовательности интенсивностей входящего потока, которые обеспечивают максимальную пропускную способность сети

при заданном ограничении на математическое ожидание длительности пребывания требований в сети.

4. Сформулированы алгоритмы управления сетью массового обслуживания с групповым обслуживанием требований, основанные на преобразовании элементов маршрутной матрицы сети и информации о длинах очередей смежных систем.
5. Разработан комплекс программ имитационного моделирования и численного анализа сети массового обслуживания с групповым обслуживанием требований и управлением.
6. Проведены численные эксперименты, иллюстрирующие эффективность предложенных методов управления для сети массового обслуживания с групповым обслуживанием требований.

Отдельные части магистерской работы были представлены на конференциях:

1. Гуркова, В.М. Оптимальные интенсивности потоков в системы $M|M^{[2]}|1$ открытой сети массового обслуживания / В. М. Гуркова, Е. П. Станкевич, И. Е. Тананко // *Системы управления, информационные технологии и математическое моделирование: Материалы Междунар. науч. конф.* – Омск: ОмГТУ. — 2022. – С. 135-139.
2. Гуркова, В. М. Использование сетей массового обслуживания с групповым обслуживанием требований в качестве моделей информационных интернет порталов / В. М. Гуркова // *Материалы Международного молодежного научного форума «Ломоносов-2022».* [Электронный ресурс] – М.: МАКС Пресс. – 2022.

Основные источники информации:

1. Baetens, J. System occupancy in a multiclass batch-service queueing system with limited variable service capacity / J. Baetens, B. Steyaert, D. Claeys, H. Bruneel // *Annals of Operations Research.* – 2020. – Vol. 293, no 1. – P. 3-26.
2. Bellalta, B. A space-time batch-service queueing model for multi-user MIMO communication systems / B. Bellalta, M. Oliver // *Proceedings of the 12th ACM international conference on modeling, analysis and simulation of wireless and mobile systems.* – 2009. – P. 357-364.
3. Hanschke, T. Queueing networks with batch service / T. Hanschke, H. Zisgen //

European Journal of Industrial Engineering 4. – 2011. – Vol. 5, no 3. – P. 313-326.

4. Santhi, K. Performance analysis of cloud computing bulk service using queueing models / K. Santhi, R. Saravanan // *International Journal of Applied Engineering Research.* – 2017. – Vol. 12, no. 17. – P. 6487-6492.
5. Mitici, M. On a tandem queue with batch service and its applications in wireless sensor networks / M. Mitici, J. Goseling, J.-K. van Ommeren, M. de Graaf, R. J. Boucherie // *Queueing systems.* – 2017. – Vol. 87, no. 1-2. – P. 81-93.
6. Станкевич, Е. П. Анализ системы массового обслуживания с групповым обслуживанием требований / Е. П. Станкевич, И. Е. Тананко, М. Пагано // *Компьютерные науки и информационные технологии: Материалы Междунар. науч. конф.* — Саратов: Научная книга. – 2021. — С. 148-151.
7. Gupta, G. K. Steady state analysis of system size-based balking in $M/M^b/1$ queue / G. K. Gupta, A. Banerjee // *International Journal of Mathematics in Operational Research.* – 2019. – Vol. 14, no. 3. – P. 319-337.
8. Klünder, W. Decomposition of open queueing networks with batch service / W. Klünder // *Operations Research Proceedings.* – 2016. – P. 575-581.
9. Гуркова, В. М. Оптимальные интенсивности потоков в системы $M|M^{[2]}|1$ открытой сети массового обслуживания / В. М. Гуркова, Е. П. Станкевич, И. Е. Тананко // *Системы управления, информационные технологии и математическое моделирование: Материалы Междунар. науч. конф.* — Омск: ОмГТУ. — 2022. — С. 135-139.
10. Lazar, A. Optimal flow control of a class of queueing networks in equilibrium / A. Lazar // *IEEE transactions on Automatic Control.* – 1983. – Vol. 28, no. 11. – P. 1001-1007.
11. Stankevich, E. Analysis of open queueing networks with batch services / E. Stankevich, I. Tananko, M. Pagano // *Information Technologies and Mathematical Modelling. Queueing Theory and Applications.* – 2022. – P. 40-51.
12. Stankevich, E. Optimization of open queuing networks with batch services / E. Stankevich, I. Tananko, M. Pagano // *Mathematics.* — 2022. — Vol. 10, no. 16. – P. 3027.