Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра системного анализа
и автоматического управления

СЕТИ МАССОВОГО ОБСЛУЖИВАНИЯ ПРОИЗВОЛЬНОЙ ТОПОЛОГИИ С ДЕЛЕНИЕМ И СЛИЯНИЕМ ТРЕБОВАНИЙ

АВТОРЕФЕРАТ НАУЧНО-КВАЛИФИКАЦИОННОЙ РАБОТЫ
(ДИССЕРТАЦИИ)

аспиранта 3 курса
направления 02.06.01 — Компьютерные и информационные науки
направленности — «Математическое моделирование, численные методы
и комплексы программ»
Осипова Олега Александровича

Научный руководитель
доцент, к. ф.-м. н. ________________ Тананко И. Е.
Зав. кафедрой
доцент, к. ф.-м. н. ________________ Тананко И. Е.

Саратов — 2018

Общая характеристика работы

Актуальность темы. Реальные системы, в которых имеет место параллельная и распределённая обработка (многопроцессорные системы, GRID-системы, распределённые базы данных, сети передачи данных), получают всё большее распространение. В таких системах поступающие для обработки задачи делятся на подзадачи более простые для выполнения, которые распределяются по системе, занимая выделенные для них ресурсы. После завершения своего выполнения подзадачи освобождают выделенные им ресурсы. При этом исходная задача считается выполненной только после завершения выполнения всех её подзадач.

Структурная и функциональная специфика систем такого класса требует разработки новых эффективных моделей и методов для использования при решении задач анализа, синтеза и оптимизации. Для описания и анализа таких систем используются разнообразные математические абстракции: модель акторов, сети Петри, потоки работ, модели теории массового обслуживания.

Сети массового обслуживания являются математическими моделями, используемыми для анализа дискретных стохастических систем с сетевой структурой, эффективность применения которых обусловила интенсивное развитие в течение последних пяти десятилетий теории сетей массового обслуживания.

Большой вклад в развитие теории, методов анализа, оптимизации и синтеза сетей массового обслуживания внесли А. А. Боровков, Г. П. Башарин, В. М. Вишневский, П. П. Бочаров, В. А. Ивницкий, Ю. И. Митрофанов, В. В. Рыков, А. В. Печинкин. Среди зарубежных специалистов необходимо отметить значительный вклад таких учёных как F. Baccelli, J. Jackson L. Kleinrock, F. Kelly, K. Chandy, D. Towsley, M. Reiser, J. Walrand, P. Harrison.

Сети массового обслуживания с делением и слиянием требований (fork-join queueing networks)являются математическими моделями, используемыми для анализа дискретных стохастических систем с параллельным и распределённым принципами функционирования. Ключевой особенностью в сетях обслуживания с делением и слиянием требований является деление поступающих требований на части — фрагменты, которые обслуживаются параллельно в системах сети обслуживания, и последующее объединение обслуженных фрагментов в исходные требования. Требование считается выполненным и покидает сеть только после окончания обслуживания всех его фрагментов.

В большинстве работ по сетям обслуживания с делением и слиянием требованийрассматриваются сети обслуживания, состоящие из множества параллельных систем массового обслуживания, а основным результатом является определение длительности пребывания требований в сети обслуживания. Поиск стационарного распределения осложнён тем, что оно не имеет простой мультипликативной формы даже в случае сетей, образованных множеством параллельных систем обслуживания.

В диссертационном исследовании рассматривается класс открытых сетей массового обслуживания произвольной топологии с делением и слиянием требований.

Целью данной работы является изучение сетей массового обслуживания произвольной топологии с делением и слиянием требований, получение стационарных характеристик этих сетей обслуживания, оптимизация их параметров, использование сетей указанного класса для моделирования реальных систем.

Для достижения поставленной цели были решены следующие **задачи**:

1. Формальное построение сетей массового обслуживания произвольной топологии с делением и слиянием требований.
2. Разработка и исследование методов анализа сетей массового обслуживания произвольной топологии с делением и слиянием требований.
3. Исследование зависимости характеристик сетей массового обслуживания от их параметров.
4. Разработка и реализация алгоритмов для расчёта стационарных характеристик и оптимизации сети обслуживания.

Научная новизна:

Постановка задач и полученные результаты являются новыми.

1. Доказано, что длительность пребывания требований в сети массового обслуживания произвольной топологии с делением и слиянием требований в случае бесконечноприборных базовых систем имеет фазовое распределение и определены его параметры.
2. Для элементарной сети обслуживания с бесконечноприборными базовыми системами получен метод нахождения стационарного распределения вероятностей состояний сети обслуживания.
3. Предложен метод аппроксимации для сетей обслуживания с одноприборными базовыми системами, исследована его точность посредством имитационного моделирования.
4. Решена задача оптимального распределения весов в сети обслуживания с делением и слиянием требований.

Теоретическая и практическая значимость Диссертация носит теоретический характер. Представленные в диссертационной работе результаты могут быть применены в задачах математического моделирования стохастических систем с параллельным и распределённым принципами функционирования, а именно, создания математических методов исследования систем параллельных и распределённых вычислений, сетей передачи данных. Рассматриваемые модели могут использоваться для решения задач оптимизации и синтеза реальных систем.

В основу диссертации положены результаты научных исследований, выполненных при участии автора в Саратовском государственном университете, по включённой в план НИР СГУ теме: «Развитие теории и методов анализа сетей массового обслуживания с групповыми переходами требований, распределением нагрузки и нестационарными структурами, разработка методов управления сетями и методов анализа сетей с управлением» (шифр «Ресурс», регистрационный № АААА-А17-117110220045-8).

Научные положения и методы, разработанные в диссертации, используются в учебном процессе Саратовского государственного университета.

Методология и методы исследования. В диссертационной работе использовались результаты теории вероятностей, случайных процессов, теории цепей Маркова, теории массового обслуживания, теории сетей массового обслуживания, методы имитационного моделирования.

Основные положения, выносимые на защиту:

1. Для предложенного класса сетей обслуживания получен вид функции распределения для длительности времени пребывания требований в сети обслуживания, а также алгоритм определения параметров этого распределения.
2. Метод нахождения стационарного распределения вероятностей состояний для элементарной сети обслуживания с бесконечноприборными базовыми системами.
3. Метод аппроксимации для сетей обслуживания с одноприборными базовыми системами, исследование его точности посредством имитационного моделирования.
4. Модель сети передачи данных с протоколом многопутевой маршрутизации и задача оптимизации, связанная с распределением субпотоков в сети передачи данных.

Достоверность полученных точных результатов обеспечивается корректными доказательствами всех приведённых в работе утверждений. Достоверность приближённых методов подтверждают результаты имитационного моделирования.

Апробация работы. Основные результаты работы докладывались и обсуждались на следующих научных семинарах и конференциях:

– научные семинары кафедры системного анализа и автоматического управления Саратовского государственного университета (2015–2018 г.),
– XXII Международная научная конференция студентов, аспирантов и молодых учёных «Ломоносов-2015» (2015 г., МГУ, Москва),
– XVI Всероссийский симпозиум по прикладной и промышленной математике (осенняя сессия), (2015 г., Сочинский государственный университет, Сочи – Дагомыс),

– Всероссийская конференция с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем» (2016, 2017 г., РУДН, Москва, Россия),
– Международная научная конференция «Компьютерные науки и информационные технологии» (2016 г., СГУ, Саратов),
– XVI Международная конференция имени А.Ф. Терпугова «Информационные технологии и математическое моделирование» (2017 г., Казанский национальный исследовательский технологический университет, Казань).

Соответствие паспорту специальности Диссертационная работа выполнена в соответствии с паспортом специальности 05.13.18 «Математическое моделирование, численные методы и комплексы программ» и включает оригинальные результаты в области математического моделирования, численных методов и комплексов программ.

Исследование, представленное в работе, соответствует следующим разделам паспорта специальности: п. 1 (Разработка новых математических методов моделирования объектов и явлений), п. 2 (Развитие качественных и приближенных аналитических методов исследования математических моделей), п. 3 (Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента), п. 8 (Разработка систем компьютерного и имитационного моделирования).

Личный вклад. Личное участие автора заключается в исследовании рассматриваемых моделей, получении аналитических результатов для них, разработке программ численного и имитационного моделирования.

Публикации. Основные результаты по теме диссертации изложены в 8 печатных изданиях, 3 из которых изданы в журналах, входящих в Перечень рецензируемых научных изданий, рекомендованных Высшей аттестационной комиссией при Министерстве образования и науки Российской Федерации для опубликования основных научных результатов диссертации, 5 — в тезисах и материалах докладов, также получено свидетельство о регистрации программы для ЭВМ.

В работах, опубликованных в соавторстве, вклад соискателя состоит в получении теоретических результатов и их интерпретации, а вклад научного руководителя — в постановке задач и обсуждении методов их решения.

Содержание работы

Во **введении** обосновывается актуальность исследований, проводимых в рамках данной диссертационной работы, определяется цель и ставятся задачи, формулируются научная новизна, а также теоретическая и практическая значимость представляемой работы.

Первая глава посвящена обзору основных опубликованных результатов исследования сетей массового обслуживания с делением и слиянием требований. В первом разделе приведены результаты работ, посвящённых сетям массового обслуживания, состоящим из параллельных систем обслуживания. Во втором разделе обсуждаются результаты для сетей обслуживания с произвольной топологией. Примеры использования сетей данного класса для моделирования реальных систем обсуждаются в третьем разделе. Четвёртый раздел посвящён методам анализа длительности пребывания требований в сетях массового обслуживания.

Вторая глава посвящена анализу открытых сетей обслуживания с делением и слиянием требований произвольной топологии в случае бесконечноприборных базовых систем обслуживания.

В *разделе 2.1* введено формальное описание для сетей обслуживания с произвольной топологией. Пусть $\mathcal{N}$ — открытая сеть массового обслуживания с одним классом требований и множеством $\{N_1, \ldots, N_L\}$ систем массового обслуживания. В сеть из источника N_0 поступает пуассоновский поток требований с интенсивностью Λ_0. Любое из требований может многократно делиться на части — фрагменты. Полученные фрагменты обслуживаются независимо друг от друга, переходят по системам сети обслуживания. Каждый из фрагментов может снова разделиться на новые фрагменты. Перед уходом из сети все фрагменты объединяются в исходное обслуженное требование. Далее будем использовать термин «фрагмент» для описания и требования, и его частей.

Пусть $X = \{1, \ldots, L\}$ — множество номеров систем обслуживания. Системы в сети обслуживания $\mathcal{N}$ поделены на три непустых, непересекающихся множества в соответствии с их номерами:

1. $X_{\mathcal{S}} = \left\{x_1^{\mathcal{S}}, \ldots, x_M^{\mathcal{S}}\right\}$ — множество номеров *базовых систем*;
2. $X_{\mathcal{F}} = \left\{x_1^{\mathcal{F}}, \ldots, x_K^{\mathcal{F}}\right\}$ — множество номеров *дивайдеров*;
3. $X_{\mathcal{J}} = \left\{x_1^{\mathcal{J}}, \ldots, x_K^{\mathcal{J}}\right\}$ — множество номеров *интеграторов*.

Число дивайдеров K совпадает с числом интеграторов, а $X = X_{\mathcal{S}} \cup X_{\mathcal{F}} \cup X_{\mathcal{J}}$. Обозначим через

– S_i — базовая система $N_{x_i^{\mathcal{S}}}$, $i = 1, \ldots, M$;
– F_j — дивайдер $N_{x_j^{\mathcal{F}}}$, $j = 1, \ldots, K$;
– J_k — интегратор $N_{x_k^{\mathcal{J}}}$, $k = 1, \ldots, K$.

Системы обслуживания, принадлежащие одному множеству, имеют сходное назначение и особенности.

1. $\mathcal{S} = \left\{N_{x_1^{\mathcal{S}}}, \ldots, N_{x_M^{\mathcal{S}}}\right\} = \{S_1, \ldots, S_M\}$ — множество базовых систем обслуживания — множество систем массового обслуживания с бесконечным числом одинаковых обслуживающих приборов.
   Длительность обслуживания фрагмента на любом приборе базо-

вой системы S_i имеет экспоненциальное распределение с параметром μ_i, $i = 1, \ldots, M$.

2. $\mathcal{F} = \left\{N_{x_1^{\mathcal{F}}}, \ldots, N_{x_K^{\mathcal{F}}}\right\} = \{F_1, \ldots, F_K\}$ — множество дивайдеров — множество одноприборных систем массового обслуживания. Процесс обслуживания в дивайдере F_k, $k = 1, \ldots, K$, заключается в мгновенном делении поступающего фрагмента на фиксированное число $d_k \geq 2$ новых фрагментов. Каждый из d_k фрагментов мгновенно покидает дивайдер.
3. $\mathcal{J} = \left\{N_{x_1^{\mathcal{J}}}, \ldots, N_{x_K^{\mathcal{J}}}\right\} = \{J_1, \ldots, J_K\}$ — множество интеграторов — множество систем массового обслуживания с бесконечным числом приборов. Процесс обслуживания в интеграторе заключается в следующем. Каждый поступающий фрагмент занимает свободный прибор и ожидает момента поступления в интегратор последнего из всех фрагментов, принадлежащих до разделения тому же самому ранее единому фрагменту. Сразу после этого эти фрагменты мгновенно освобождают обслуживающие приборы, объединяются в исходный фрагмент, который покидает интегратор.

Каждый из фрагментов в сети обслуживания характеризуется сигнатурой, которая идентифицирует фрагмент и отражает связи с другими фрагментами.

Определение 1. Пусть σ — некоторый произвольный фрагмент в сети обслуживания, который получен непосредственно в результате деления фрагмента σ^* в дивайдере F_k. Пусть i — номер фрагмента σ при делении, то есть $i \in \{1, \ldots, d_k\}$. *Сигнатурой фрагмента* σ назовём тройку (σ^*, k, i). В том случае, когда σ является поступающим из источника требованием, сигнатура имеет вид $(\sigma, 0, 1)$.

Если фрагмент σ имеет сигнатуру (σ^*, k, i), то будем записывать это $\sigma \sim (\sigma^*, k, i)$.

Предположим, что фрагмент σ поступает в произвольный дивайдер F_k, $k \in \{1, \ldots, K\}$, который делит фрагмент σ на d_k фрагментов $\{\sigma_1, \ldots, \sigma_{d_k}\}$, $\sigma_i \sim (\sigma, k, i)$, $i = 1, \ldots, d_k$. Полученные в результате деления фрагменты мгновенно покидают дивайдер и поступают в системы обслуживания. Базовые системы не меняют сигнатуру фрагмента, а каждый из полученных фрагментов может снова разделиться на новые фрагменты при поступлении в другой дивайдер. Для корректного объединения фрагментов $\{\sigma_1, \ldots, \sigma_{d_k}\}$ они все должны в итоге поступить в некоторый интегратор. Маршрутизация фрагментов задаётся так, что все фрагменты σ_i, $i = 1, \ldots, d_k$, полученные в дивайдере F_k, должны впоследствии поступить и быть объединены только в соответствующем интеграторе J_k. Переход данных фрагментов в интеграторы $J_l, l \neq k$, не допускается.

При поступлении фрагмента $\sigma_i \in \{\sigma_1, \ldots, \sigma_{d_k}\}$ в интегратор J_k он находится в этом интеграторе до тех пор, пока не поступят все оставшиеся фрагменты из данного множества. Как только это произошло, фрагменты $\{\sigma_1, \ldots, \sigma_{d_k}\}$ освобождают обслуживающие их приборы, мгновенно объединяются в исходный фрагмент σ, который сразу же покидает интегратор, переходя в другую систему обслуживания или источник.

Фрагмент σ, характеризуемый сигнатурой (σ^*, k, i), будем называть k-фрагментом. Фрагмент может вернуться в источник требований N_0 только в случае, когда он является требованием (0-фрагментом).

Для каждого фрагмента в сети обслуживания введём в рассмотрение вектор перемещений $\boldsymbol{v}$.

Определение 2. Пусть произвольный фрагмент σ характеризуется сигнатурой (σ_m, k_m, i_m), и при этом $\sigma_n \sim (\sigma_{n-1}, k_{n-1}, i_{n-1})$, $n = m, m-1, \ldots, 2$, $\sigma_1 \sim (\sigma_1, 0, 1)$. Вектор $\boldsymbol{v} = (k_0, k_1, \ldots, k_m) = (0, k_1, \ldots, k_m)$ назовём *вектором перемещений* для k_m-фрагмента σ, где k_m — *ведущий элемент* вектора перемещений.

Предполагается, что маршрутизация фрагментов такова, что вектор перемещений каждого фрагмента не может содержать одинаковых элементов, то есть $k_i \neq k_j,\ i \neq j,\ \ i,j = 1, \ldots, m$. Все векторы перемещений тогда образуют *множество допустимых векторов перемещений* $\mathcal{V}$, которое является конечным множеством, зависящим от топологии сети обслуживания. Пусть $\boldsymbol{v} = (0, k_1, \ldots, k_m)$, тогда через $(\boldsymbol{v}, k_{m+1})$ обозначим вектор $(0, k_1, \ldots, k_m, k_{m+1})$.

Возможные переходы фрагментов в сети обслуживания $\mathcal{N}$ заданы при помощи *набора матриц передач* $\boldsymbol{\Theta} = \left(\boldsymbol{\Theta}^{(0)}, \boldsymbol{\Theta}^{(1)}, \ldots, \boldsymbol{\Theta}^{(K)}\right)$, где $\boldsymbol{\Theta}^{(k)} = \left(\theta_{i,j}^{(k)}\right)$, $i,j = 0,1, \ldots, L,\ k = 0,1, \ldots, K$.

Элемент $\theta_{i,j}^{(k)}$ определяет

– вероятность перехода k-фрагмента из N_i в N_j, если $N_i \notin \mathcal{F}$;
– число k-фрагментов, полученных при делении фрагмента в дивайдере N_i, которые перейдут в N_j, если $N_i \in \mathcal{F}$.

Очевидно, что тогда

$$\theta_{x_i^{\mathcal{F}},j}^{(k)} = 0, \quad i = 1, \ldots, K, i \neq k,\ j = 0,1, \ldots, L;$$

$$\sum_{j=0}^{L} \theta_{x_k^{\mathcal{F}},j}^{(k)} = d_k.$$

Предполагается, что набор матриц передач $\boldsymbol{\Theta}$ обеспечивает корректную маршрутизацию фрагментов, то есть маршрутизация удовлетворяет условиям, указанным выше.

Раздел 2.2 посвящён методам анализа рассматриваемой сети обслуживания. Пусть $\lambda_{in}^{(\boldsymbol{v})}(N_i)$, $\lambda_{out}^{(\boldsymbol{v})}(N_i)$ — *интенсивности входящего и выходящего потоков фрагментов* с вектором перемещений $\boldsymbol{v} \in \mathcal{V}$ для системы обслуживания N_i, $i = 1, \ldots, L$. Тогда *суммарные интенсивности* входящего $\Lambda_{in}(N_i)$ и выходящего $\Lambda_{out}(N_i)$ потоков фрагментов для систем обслуживания N_i, $i = 1, \ldots, L$,

$$\Lambda_{in}(N_i) = \sum_{\boldsymbol{v}\in\mathcal{V}} \lambda_{in}^{(\boldsymbol{v})}(N_i), \qquad \Lambda_{out}(N_i) = \sum_{\boldsymbol{v}\in\mathcal{V}} \lambda_{out}^{(\boldsymbol{v})}(N_i).$$

Пусть k — ведущий элемент вектора перемещений $\boldsymbol{v} \in \mathcal{V}$, в стационарном режиме функционирования сети обслуживания интенсивности входящих и выходящих потоков связаны следующими соотношениями:

$$\lambda_{in}^{(0)}(N_j) = \Lambda_0\theta_{0,j}^{(0)} + \sum_{i\in X_{\mathcal{S}}\cup X_{\mathcal{J}}} \lambda_{out}^{(0)}(N_i)\theta_{i,j}^{(0)}, \quad k = 0; \tag{1}$$

$$\lambda_{in}^{(\boldsymbol{v})}(N_j) = \sum_{i\in X_{\mathcal{S}}\cup X_{\mathcal{J}}} \lambda_{out}^{(\boldsymbol{v})}(N_i)\theta_{i,j}^{(k)} + \frac{1}{d_k}\lambda_{out}^{(\boldsymbol{v})}(F_k)\theta_{x_k^{\mathcal{F}},j}^{(k)}, \quad k > 0; \tag{2}$$

где $j = 1, \ldots, L$.

Справедливо, что

$$\lambda_{out}^{(\boldsymbol{v})}(S_i) = \lambda_{in}^{(\boldsymbol{v})}(S_i), \quad i = 1, \ldots, M, \tag{3}$$

$$\lambda_{out}^{(\boldsymbol{v},k)}(F_k) = d_k\lambda_{in}^{(\boldsymbol{v})}(F_k), \quad \lambda_{out}^{(\boldsymbol{v})}(J_k) = \frac{1}{d_k}\lambda_{in}^{(\boldsymbol{v},k)}(J_k), \quad k = 1, \ldots, K. \tag{4}$$

$$\lambda_{in}^{(\boldsymbol{v})}(F_k) = \lambda_{out}^{(\boldsymbol{v})}(J_k), \quad k = 1, \ldots, K. \tag{5}$$

Интенсивности входящих и выходящих потоков фрагментов для всех $\boldsymbol{v} \in \mathcal{V}$ можно найти как решение уравнений (1), (2) с использованием равенств (3) и (4). Однако для вычисления интенсивностей потоков фрагментов удобнее использовать алгоритм, который представлен в диссертации и не приводится здесь в силу громоздкости.

Утверждение 1. Математическое ожидание $\mathsf{E}[x_i]$ числа фрагментов x_i в системе S_i определяется следующим выражением

$$\mathsf{E}[x_i] = \frac{\Lambda_{in}(S_i)}{\mu_i}, \quad i = 1, \ldots, M.$$

Определение 3. Множество $\mathcal{H}_k \subset \{N_1, \ldots, N_L\}$, $k \in \{1, \ldots K\}$, таких систем обслуживания, в которые могут поступить фрагменты с вектором перемещений, содержащим k, будем называть *подсетью порождённой дивайдером* F_k. Положим дополнительно, что $F_k \in \mathcal{H}_k$.

Определение 4. Пусть сеть обслуживания $\mathcal{N}$ функционирует в стационарном режиме. Обозначим через τ_k, $k = 1, \ldots, K$, случайную величину, равную длительности интервала времени от момента разделения в дивайдере F_k произвольного фрагмента до момента объединения всех его частей в интеграторе J_k. Будем называть τ_k *длительностью реакции подсети* $\mathcal{H}_k$.

Определение 5. Матрицу $\mathbf{\Theta}^{(k)}$ из набора матриц передач $\mathbf{\Theta}$ назовём *элементарной матрицей*, если столбцы, соответствующие переходам во все дивайдеры, состоят из нулей, то есть

$$\theta^{(k)}_{i,x^{\mathcal{F}}_j} = 0, \quad i = 0,1, \ldots, L,\ j = 1, \ldots, K.$$

Определение 6. Пусть для некоторого $k \in \{1, \ldots, K\}$, $\mathbf{\Theta}^{(k)}$ — элементарная матрица, тогда будем называть подсеть $\mathcal{H}_k$ *элементарной подсетью.*

Пусть $\mathcal{H}_k$, $k \in \{1, \ldots, K\}$, есть элементарная подсеть. Через $\mathcal{A}(k)$ обозначим множество номеров базовых систем обслуживания, в которые переходят фрагменты непосредственно после деления в дивайдере F_k, то есть

$$\mathcal{A}(k) = \left\{ j \in \{1, \ldots, M\} \colon \theta^{(k)}_{x^{\mathcal{F}}_k, x^{\mathcal{S}}_j} > 0 \right\}. \tag{6}$$

Пусть $\theta^{(k)}_{x^{\mathcal{F}}_k, x^{\mathcal{S}}_i} > 0$, тогда определим (k,i)-*ветвь подсети* как упорядоченную последовательность $\mathcal{B}(k,S_i)$ номеров всех базовых систем S_{b_j}, $j = 2, \ldots, c_{\boldsymbol{b}}$, в которые возможно поступление k-фрагментов из системы S_i,

$$\mathcal{B}(k,S_i) = (b_1, b_2, \ldots, b_{c_{\boldsymbol{b}}}),$$
$$b_1 = i, \quad b_2 < b_3 < \cdots < b_{c_{\boldsymbol{b}}}.$$

Пусть ξ есть непрерывная случайная величина с фазовым распределением[1,2] и параметрами $(\boldsymbol{\alpha},\boldsymbol{A})$, обозначим это как $\xi \sim \mathrm{PH}(\boldsymbol{\alpha},\boldsymbol{A})$, известно что если $\xi \sim \mathrm{PH}(\boldsymbol{\alpha},\boldsymbol{A})$ и $\eta \sim \mathrm{PH}(\boldsymbol{\beta},\boldsymbol{B})$ суть независимые случайные величины, то случайная величина $\max\{\xi,\eta\}$ имеет фазовое распределение $\mathrm{PH}(\boldsymbol{\gamma},\boldsymbol{C})$, где $(\boldsymbol{\gamma},\boldsymbol{C}) = (\boldsymbol{\alpha},\boldsymbol{A}) \vee (\boldsymbol{\beta},\boldsymbol{B})$, а операция “ $\vee$ ” определена следующим образом

$$\boldsymbol{\gamma} = (\boldsymbol{\alpha} \otimes \boldsymbol{\beta}, (1 - \boldsymbol{\beta}\mathbf{1})\boldsymbol{\alpha}, (1 - \boldsymbol{\alpha}\mathbf{1})\boldsymbol{\beta})\,, \tag{7}$$

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{A} \oplus \boldsymbol{B} & \boldsymbol{I} \otimes (-\mathbf{B1}) & (-\mathbf{A1}) \otimes \boldsymbol{I} \\ \mathbf{0} & \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{B} \end{pmatrix}, \tag{8}$$

здесь $\mathbf{1}$ — вектор-столбец из единиц, $\oplus$, $\otimes$ — символы операций суммы и произведения Кронекера, $\boldsymbol{I}$ — единичная матрица.

[1] *He Q.-M.* Fundamentals of Matrix-Analytic Methods. New York : Springer, 2014. P. 349.

[2] *Neuts M.* Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach. Baltimore : The Johns Hopkins University Press, 1981.

Теорема 1. *Длительность реакции τ_k элементарной подсети $\mathcal{H}_k$, $k \in \{1,\ldots,K\}$, имеет фазовое распределение* $\mathrm{PH}\left(\boldsymbol{\alpha}^{(k)}, \boldsymbol{A}^{(k)}\right)$, *где*

$$(\boldsymbol{\alpha}^{(k)}, \boldsymbol{A}^{(k)}) = \bigvee_{j \in \mathcal{A}(k)} \left(\bigvee_{l=1}^{\theta^{(k)}_{x^{\mathcal{F}}_k, x^{\mathcal{S}}_j}} \left(\hat{\boldsymbol{\alpha}}^{(j)}, \hat{\boldsymbol{A}}^{(j)}\right) \right),$$

$$\hat{\boldsymbol{\alpha}}^{(j)} = \left(\hat{\alpha}^{(j)}_n\right), \quad n = 1,\ldots,c_{\boldsymbol{b}},$$
$$\hat{\alpha}^{(j)}_1 = 1,$$

$$\hat{\boldsymbol{A}}^{(j)} = \left(\hat{a}^{(j)}_{n,m}\right), \quad n,m = 1,\ldots,c_{\boldsymbol{b}},$$
$$\hat{a}^{(j)}_{n,m} = \mu_{b_n}\theta^{(k)}_{n^*,m^*}, \quad n \neq m,$$
$$\hat{a}^{(j)}_{n,n} = -\mu_{b_n}.$$

Здесь $(b_1,\ldots,b_{c_{\boldsymbol{b}}}) = \mathcal{B}(k,S_j)$,

$$n^* = x^{\mathcal{S}}_{b_n}, \quad m^* = x^{\mathcal{S}}_{b_m}.$$

Замечание 1. Элементарная подсеть $\mathcal{H}_k$ может быть представлена как система типа $\cdot/\mathrm{PH}(\boldsymbol{\alpha}^{(k)}, \boldsymbol{A}^{(k)})/\infty$.

Теорема 1 позволяет рассмотреть процесс обслуживания фрагмента в элементарной подсети как эволюцию некоторой поглощающей цепи Маркова. Данную цепь Маркова можно рассматривать также как модель процесса обслуживания фрагмента в подсети обслуживания, состоящей только из базовых систем.

Рассмотрим процедуру модификации исходной сети обслуживания $\mathcal{N}$, при которой элементарная подсеть $\mathcal{H}_k$ замещается на другую подсеть, не содержащую дивайдеров и интеграторов, с таким же распределением длительности пребывания фрагментов в ней. Указанную процедуру будем называть *редукцией сети обслуживания $\mathcal{N}$ относительно элементарной подсети $\mathcal{H}_k$.*

Последовательное применение процедуры редукции позволяет найти распределение для длительности пребывания требований в сети обслуживания.

Теорема 2. *Длительность пребывания требований в сети массового обслуживания $\mathcal{N}$ с делением и слиянием требований имеет фазовое распределение.*

Раздел 2.3 содержит подробный пример использования разработанных методов.

В *разделе 2.4* рассматривается задача нахождения стационарного распределения для элементарной сети обслуживания с делением и слиянием требований. Пусть $\mathcal{N}$ — открытая сеть массового обслуживания с делением и слиянием требований, состоящая из дивайдера F, интегратора J, базовых систем $S_1, \ldots, S_M$. Предполагается, что каждое требование из источника непосредственно поступает в дивайдер F, где делится на d фрагментов, объединение фрагментов происходит затем в интеграторе J, а полученное требование возвращается в источник. Будем называть такую сеть *элементарной* сетью обслуживания.

Состояние сети $\mathcal{N}$ определяется вектором $\boldsymbol{x} = (x_1, \ldots, x_M)$, где x_i — число фрагментов в системе S_i. Процесс $\{\boldsymbol{x}(t), t > 0\}$ является цепью Маркова с непрерывным временем и счётным множеством состояний.

Определение 7. Пусть фрагмент σ делится в дивайдере F на фрагменты $\{\sigma_1, \ldots, \sigma_d\}$, тогда под *способом размещения* требования в сети обслуживания будем понимать мультимножество $\boldsymbol{l}$, состоящее из d элементов и определяемое следующим образом:

$$\boldsymbol{l} = \{N(\sigma_1), \ldots, N(\sigma_d)\},$$

где $N(\sigma_i)$ обозначает систему, в которой находится фрагмент σ_i.

Процесс $\{\boldsymbol{l}(t), t > 0\}$, описывающий переход фрагментов некоторого требования от момента деления в дивайдере до момента объединения в интеграторе, является поглощающей цепью Маркова на множестве $\boldsymbol{L} = \{\boldsymbol{l}_0, \ldots, \boldsymbol{l}_X\}$ всех размещений, с поглощающим состоянием $\boldsymbol{l}_0 = \{J, \ldots, J\}$ и начальным $\boldsymbol{l}_1 = \mathcal{A}(F)$.

Определение 8. Сопоставим цепи Маркова $\{\boldsymbol{l}(t), t > 0\}$ сеть Джексона $\mathcal{N}^{\boldsymbol{L}}$ с бесконечным числом обслуживающих приборов в системах обслуживания так, что источнику требований будет соответствовать поглощающее состояние $\boldsymbol{l}_0$, а каждой системе в сети $\mathcal{N}^{\boldsymbol{L}}$ поставим во взаимнооднозначное соответствие некоторое непоглощающее состояние из $\boldsymbol{L}$. Будем называть сеть $\mathcal{N}^{\boldsymbol{L}}$ *сетью размещений.* Далее будем использовать элементы из $\boldsymbol{L}$ для обозначения соответствующих им систем обслуживания или источника в сети $\mathcal{N}^{\boldsymbol{L}}$.

Сеть обслуживания $\mathcal{N}^{\boldsymbol{L}}$ содержит X систем. Обозначим через y_j число требований в системе обслуживания сети $\mathcal{N}^{\boldsymbol{L}}$, соответствующей способу размещения $\boldsymbol{l}_j$. Пусть сеть $\mathcal{N}^{\boldsymbol{L}}$ находится в состоянии $\boldsymbol{y} = (y_1, \ldots, y_X)$, тогда число фрагментов x_i в системе S_i сети обслуживания $\mathcal{N}$

$$x_i = \sum_{j=1}^{X} y_j \, \varphi(\boldsymbol{l}_j, S_i), \quad i = 1, \ldots M, \tag{9}$$

где $\varphi(\boldsymbol{l}_j, S_i)$ обозначает число вхождений элемента S_i в мультимножество $\boldsymbol{l}_j$.

Известно, что для сети $\mathcal{N}^{\boldsymbol{L}}$, стационарное распределение вероятностей состояний имеет мультипликативную форму[3],

$$\pi(y_1, \ldots, y_X) = \prod_{i=1}^{X} \pi_i(y_i), \tag{10}$$

где $\pi_i(y_i)$ — стационарная вероятность пребывания y_i требований в системе с номером i.

Элементы множества $\boldsymbol{L}_i \subset \boldsymbol{L}$ будем называть значимыми способами размещения для системы S_i, если для любого $\boldsymbol{l} \in \boldsymbol{L}_i$, выполняется $\varphi(\boldsymbol{l}, S_i) > 0$.

Теорема 3. *Для сети $\mathcal{N}$ стационарная вероятность $p_i(x_i)$ того, что система S_i содержит x_i фрагментов*

$$p_i(x_i) = \sum_{u_i((y_{r_1}, \ldots, y_{r_m})) = x_i} \pi_{r_1}(y_{r_1}) \ldots \pi_{r_m}(y_{r_m}), \tag{11}$$

где r_j, $j = 1, \ldots, m$, есть множество номеров значимых состояний для системы S_i, $u_i(\boldsymbol{y}) = \sum_{j=1}^{X} y_j \, \varphi(\boldsymbol{l}_j, S_i)$.

Следствие 1. Для сети $\mathcal{N}$ стационарная вероятность $p(x_1, \ldots, x_M)$ определяется как

$$p(x_1, \ldots, x_M) = \sum_{\substack{u_1(\boldsymbol{y}) = x_1, \\ \ldots \\ u_M(\boldsymbol{y}) = x_M}} \pi_1(y_1) \ldots \pi_X(y_X). \tag{12}$$

Таким образом задача нахождения стационарного распределения для элементарной сети с делением и слиянием требований сводится к определению соответствующей сети размещений, метод построения которой представлен в диссертации и здесь не приводится.

Третья глава посвящена исследованию открытых сетей массового обслуживания произвольной топологии с делением и слиянием требований, зависимых от нагрузки.

Рассматривается открытая бесконечноприборная сеть обслуживания $\mathcal{N}$ с делением и слиянием требований, в которой интенсивность обслуживания фрагментов одним прибором базовой системы сети *зависит от интенсивности поступающего в эту систему потока*. Пусть для базовых

[3] *Вишневский В.* Теоретические основы проектирования компьютерных сетей. Москва : Техносфера, 2003. С. 512.

систем известны интенсивности входящего потока $\Lambda_{in}(S_i)$, $i = 1, \ldots, M$, тогда интенсивность обслуживания $\hat{\mu}_i$ фрагментов прибором базовой системы S_i определяется следующей зависимостью

$$\hat{\mu}_i = \mu_i - \Lambda_{in}(S_i), \quad i = 1, \ldots, M, \tag{13}$$

где μ_i есть заданные константы. Длительность пребывания фрагментов в базовой системе S_i имеет экспоненциальное распределение с параметром $\hat{\mu}_i$. Для анализа сети обслуживания $\mathcal{N}$ можно использовать методы, изложенные в предыдущей главе.

Пусть топология и маршрутизация в сети обслуживания $\mathcal{N}^f$ совпадают с сетью $\mathcal{N}$, но *все базовые системы являются одноприборными*, а интенсивность обслуживания фрагментов прибором базовой системы S_i не зависит от интенсивности поступающего в неё потока и равна μ_i.

Задачу анализа сети обслуживания $\mathcal{N}^f$ с одноприборными базовыми системами, можно свести к известной задаче анализа сети $\mathcal{N}$ с бесконечноприборными базовыми системами и соответствующим образом выбранными интенсивностями обслуживания приборов, которые будут зависеть от нагрузки. Данный метод анализа является приближённым, а проведённые эксперименты на разработанной имитационной модели позволили установить, что полученное таким образом приближение для математического ожидания длительности пребывания требований в сети обслуживания является допустимым для реальных технических задач. Максимальная относительная погрешность при этом составила не более 6%, средние относительные погрешности составили 4%.

В *разделе 3.2* рассматривается задача оптимизации *квазиэлементарной* сети обслуживания — сети с делением и слиянием требований, в которой все подсети, порождаемые дивайдерами, являются элементарными, а требования из источника поступают непосредственно в дивайдеры и после объединения соответствующих фрагментов в интеграторах сразу возвращаются в источник.

Предполагается, что каждое поступающее в сеть обслуживания требование имеет *вес* $w = 1$. Требование, поступившее на дивайдер F_k, $k = 1, \ldots, K$, мгновенно делится на фрагменты, которые распределяются по базовым системам в соответствии с распределением весов $\mathcal{W}$,

$$\mathcal{W} = \{w_k(l)\colon k = 1, \ldots, K, l \in \mathcal{A}(k)\}.$$

Фрагмент с весом $w_k(l) \in (0,1]$ поступает в базовую систему S_l, $l \in \mathcal{A}(k)$, где множество $\mathcal{A}(k)$ определяется выражением (6).

Выполняется *закон сохранения веса*

$$\sum_{l \in \mathcal{A}(k)} w_k(l) = 1, \quad k = 1, \ldots, K.$$

Фрагменты, полученные при делении в дивайдере F_k, будем так же называть k-фрагментами, если же дополнительно известно, что k-фрагмент поступил из F_k *непосредственно в* S_l, то будем называть его (k,l)-*фрагментом.*

Длительность пребывания фрагмента в базовой системе S_i, $i = 1, \ldots, M$, имеет *экспоненциальное распределение с параметром* $\hat{\mu}_i$,

$$\hat{\mu}_i = \mu_i - \Lambda^*_{in}(S_i),$$

где $\Lambda^*_{in}(S_i)$ есть *суммарная взвешенная интенсивность* входящего в S_i потока. Предполагается, что выполнено условие

$$\mu_i - \Lambda^*_{in}(S_i) > 0, \quad i = 1, \ldots, M.$$

Зависимость параметра $\hat{\mu}_i$ длительности обслуживания может быть выбрана исходя их технических особенностей функционирования реальной системы.

Для систем обслуживания N_i, $i = 1, \ldots, L$, обозначим через

– $\lambda^{(k,l)}_{in}(N_i)$, $\lambda^{(k,l)}_{out}(N_i)$ *интенсивности входящего и выходящего потоков* (k,l)-*фрагментов*;
– $\lambda^{(0)}_{in}(N_i)$, $\lambda^{(0)}_{out}(N_i)$ *интенсивности входящего и выходящего потоков* (0)-*фрагментов.*

Суммарная взвешенная интенсивность $\Lambda^*_{in}(N_i)$ входящего в N_i потока фрагментов определяется выражением

$$\Lambda^*_{in}(N_i) = \lambda^{(0)}_{in}(N_i) + \sum_{k=1}^{K} \sum_{l \in \mathcal{A}(k)} \lambda^{(k,l)}_{in}(N_i) w_k(l). \tag{14}$$

Рассмотрим задачу нахождения для каждого дивайдера F_k, $k = 1, \ldots, K$, такого *распределения весов* $\{w^*_k(l) : l \in \mathcal{A}(k)\}$, которое *минимизирует математическое ожидание длительности пребывания требований в сети обслуживания.* Предполагается, что найденное распределение может иметь нулевые компоненты, что будет означать запрет поступления фрагментов в соответствующие базовые системы.

Для решения поставленной задачи минимизации приемлемым способом является использование методов внутренней точки, предназначенных для решения задач нелинейной оптимизации.

Моделирование сетей передачи данных с многопутевой маршрутизацией на примере протокола MPTCP рассмотрено в *разделе 3.3.* При данном виде маршрутизации поток от источника к получателю разделяется на несколько субпотоков, которые передаются по различным маршрутам, что позволяет эффективно распределить нагрузку в сети и тем самым увеличить пропускную способность. Однако при этом возникает необходимость осуществления синхронизации всех субпотоков. Если соединение использует протокол MPTCP, то возможен обмен пакетами с несколькими адресами/интерфейсами одновременно, в рамках одного соединения. Таким образом, при использовании многопутевой маршрутизации передача данных происходит одновременно по нескольким маршрутам, которые могут быть не полностью изолированными (совместно использовать несколько маршрутизаторов).

Для моделирования процессов, которые происходят с субпотоками в сетях передачи данных с многопутевой маршрутизацией, предлагается использовать сети массового обслуживания с делением и слиянием требований.

В **четвертой главе** приведено описание разработанного комплекса программ для анализа сетей обслуживания произвольной топологии с делением и слиянием требований. Приводится краткое описание структуры программ, примеры их использования, обсуждаются полученные результаты.

Представленный проблемно-ориентированный комплекс состоит из программы имитационного моделирования сетей обслуживания, а также программы для численных расчётов вероятностных характеристик сетей обслуживания на основе полученных в работе результатов. Интерфейс пользователя позволяет быстро описывать исследуемые модели, проводить моделирование, выполнять анализ результатов.

С использованием программного комплекса исследуется возможность использования бесконечноприборных моделей с зависимой от нагрузки интенсивностью обслуживания для приближённого анализа аналогичных сетей с одноприборными базовыми системами.

В **заключении** приведены основные результаты работы, которые заключаются в следующем:

1. Дано формальное описание сетей массового обслуживания произвольной топологии с делением и слиянием требований.
2. Разработаны методы анализа сетей массового обслуживания произвольной топологии с делением и слиянием требований в случае бесконечноприборных базовых систем обслуживания. Доказано, что длительность пребывания требования в сети массового обслуживания произвольной топологии с делением и слиянием требований имеет фазовое распределение и определены его параметры.
3. Для элементарной сети обслуживания с бесконечноприборными базовыми системами получен метод нахождения стационарного распределения.
4. Предложен метод аппроксимации для сетей обслуживания с одноприборными базовыми системами, исследована его точность посредством имитационного моделирования.

5. Предложен метод оптимизации распределения весов в сети обслуживания с делением и слиянием требований.
6. Разработан комплекс программ имитационного моделирования и численного анализа сетей массового обслуживания с делением и слиянием требований.

Публикации автора по теме диссертации

1. Анализ сетей массового обслуживания с разделением и слиянием требований и управлением потоками : Свидетельство о государственной регистрации программы для ЭВМ: 2014662807 / О. А. Осипов.
2. *Осипов О. А.*, *Тананко И. Е.* Сети массового обслуживания произвольной топологии с делением и слиянием требований: случай бесконечноприборных систем обслуживания // Вестник Тверского государственного университета. Серия: Прикладная математика. — 2017. — № 4. — С. 43—58.
3. *Осипов О. А.* Система обслуживания с делением и слиянием требований, в которой требование занимает все свободные обслуживающие приборы // Вестник Российского университета дружбы народов. Серия: Математика. Информатика. Физика. — 2018. — Т. 26, № 1. — С. 28—38.
4. *Осипов О. А.* Анализ *RQ*-сети массового обслуживания с делением и слиянием требований // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. — 2018. — № 43. — С. 49—55.
5. *Тананко И. Е.*, *Осипов О. А.* Методология имитационного моделирования открытых сетей массового обслуживания произвольной топологии с делением и слиянием требований // Компьютерные науки и информационные технологии. Материалы Международной научной конференции. — 2016. — С. 408—411.
6. *Осипов О. А.* Исследование сетей массового обслуживания с делением и слиянием требований // Международная научная конференция студентов, аспирантов и молодых учёных Ломоносов-2015. — МАКС Пресс. 2015.
7. *Осипов О. А.* О сети массового обслуживания с делением и слиянием требований с ограничением на фрагментацию // Шестнадцатый Всероссийский симпозиум по прикладной и промышленной математике (осенняя сессия). Т. 22. — Обозрение прикладной и промышленной математики. 2015. — С. 490—491.

8. *Осипов О. А.*, *Тананко И. Е.* Моделирование сетей передачи данных с многопутевой маршрутизацией сетями массового обслуживания с делением и слиянием требований // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. — РУДН. 2016. — С. 110—112.

9. *Осипов О. А.* Построение модели системы распределённых вычислений в виде системы массового обслуживания с делением и слиянием требований // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. — РУДН. 2017. — С. 135—136.