

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра системного анализа и
автоматического управления

**ИССЛЕДОВАНИЕ СЕТИ МАССОВОГО ОБСЛУЖИВАНИЯ С
УПРАВЛЕНИЕМ ИНТЕНСИВНОСТЯМИ ОБСЛУЖИВАНИЯ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 481 группы
направления 27.03.03 — Системный анализ и управление
факультета КНиИТ
Васильевой Анны Владиславовны

Научный руководитель
старший преподаватель

И. А. Люкшин

Заведующий кафедрой
к. ф.-м. н., доцент

И. Е. Тананко

Саратов 2024

ВВЕДЕНИЕ

Актуальность темы. Современный мир характеризуется высокой степенью конкуренции и постоянным ростом требований к эффективности и качеству продуктов, производства, обслуживания. Эффективное функционирование систем и сетей массового обслуживания становится неотъемлемой составляющей успеха во многих отраслях — телекоммуникации, транспортные системы, компьютерные сети, производственные процессы и другие. Для повышения качества обслуживания и улучшения производительности необходимо постоянно проводить исследование данного процесса с целью выявления оптимальных стратегий управления функционированием сети. Интенсивное развитие информационных технологий и расширение глобальных коммуникаций способствуют этому, позволяют использовать максимальное количество высокоуровневых инструментов, начиная с удобной реализацией сложных математических закономерностей и заканчивая использованием искусственного интеллекта и нейронных сетей.

Оптимизация, или повышение эффективности и качества при уменьшении затрат, является неотъемлемой частью эволюции человечества: чтобы быстрее выкапывать корешки из земли, люди создали палку-копалку; чтобы наносить больший урон убегающей добыче, изобрели копья и лук со стрелами; чтобы быстрее перемещаться по своим территориям — приручили лошадь, изобрели колесо, а потом и телеги; чтобы быстрее обмениваться информацией — изобрели различные способы коммуникации и в итоге — сеть Интернет. Все это примеры того, как человек усовершенствовал один из элементов большой системы, чтобы увеличить эффективность ее работы в целом. По такому принципу устроены и современные методы оптимизации — изменяя некоторые параметры сетей или систем получается улучшить необходимые характеристики их функционирования.

К нашему времени разработано множество методов оптимизации, которые успешно применяются для решения задач в области массового обслуживания. В них входят классические аналитические подходы, современные численные методы, например методы стохастической оптимизации и машинное обучение, а также методы оптимизации посредством разделения потоков или управления интенсивностями обслуживания, что наиболее часто применимо к замкнутым сетям массового обслуживания с дискретным временем

и переходами требований. Однако помимо применения к сетям метода оптимизации, за счет управления интенсивностями обслуживания, необходимо также исследовать влияние этого управления на качество функционирования сети — определить, какие характеристики наиболее подвержены переменам за счет внесения изменений, корректировку каких характеристик провести важнее всего, каково оптимальное время применения управления.

Целью данной выпускной квалификационной работы является исследование эволюции сети при использовании управления интенсивностями обслуживания, и сравнение ее характеристик с подобной сетью, но без применения управления. Предполагается, что использование управления интенсивностями обслуживания поспособствует улучшению стационарных характеристик сети, которые отображают загруженность систем сети, время обслуживания требований в них и общую пропускную способность сети в целом.

Цель бакалаврской работы — провести исследование эволюции сети при использовании управления интенсивностями обслуживания, и сравнить ее характеристики с подобной сетью, но без применения управления.

Поставленная цель определила **следующие задачи**:

1. Рассмотреть особенности функционирования замкнутых сетей массового обслуживания с дискретным временем и групповыми переходами требований;
2. Ознакомиться с методами оптимизации функционирования сети;
3. Создать программную реализацию модели функционирования сети массового обслуживания с дискретным временем и групповыми переходами требований, как без использования управления интенсивностями обслуживания, так и с его применением;
4. Сравнить полученные значения характеристик для составления выводов.

Методологические основы анализа сети с управлением интенсивностями обслуживания представлены в работах В. И. Долгова, Ю. И. Митрофанова, Е. С. Рогачко, Е. П. Станкевич.

Теоретическая значимость бакалаврской работы. Проведенное исследование сети массового обслуживания с управлением интенсивностями обслуживания расширяет круг задач, решаемых в теории массового обслуживания, поскольку позволяет вычислить показатели качества функционирования,

ния сети, в которой применяется динамическое управление интенсивностями обслуживания. Найдено стационарное распределение вероятностей состояний исследуемой сети, а также получены формулы для её основных стационарных характеристик

Практическая значимость бакалаврской работы. В ходе выполнения выпускной квалификационной работы было проведено исследование сети управлением интенсивностями обслуживания, которое доказало эффективность применения управления на малых длительностях тактов функционирования системы, что дает возможность применять данную модель сети в практических приложениях.

Структура и объем работы. Бакалаврская работа состоит из введения, 5 разделов, заключения, списка использованных источников и приложения. Общий объем работы — 61 страница, из них 45 страниц — основное содержание, включая 10 рисунков и 5 таблиц, 21 страницу приложения, список использованных источников информации — 20 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Системы массового обслуживания» посвящен описанию основных элементов теории систем массового обслуживания и рассмотрению их особенностей.

В подразделе 1.1 приводится небольшая историческая справка о развитии теории массового обслуживания.

Подраздел 1.2 посвящен описанию систем массового обслуживания и их характеристик.

В подразделе 1.3 описаны существующие классификации систем массового обслуживания.

Второй раздел «Сети массового обслуживания» посвящен описанию сетей массового обслуживания, рассмотрению их особенностей и приведению формул для вычисления их стационарных характеристик.

В подразделе 2.1 приводятся общие сведения о сетях массового обслуживания.

Подраздел 2.2 посвящен описанию замкнутой сети массового обслуживания с дискретным временем и групповыми переходами требований. Рассмотрим сеть N . Эта сеть замкнутая, в ней содержится L систем массового

обслуживания S_i , $i \in I = 1, \dots, L$, и H требований одного класса. Вероятности перехода требований между системами сети определяются неприводимой маршрутной матрицей $\Theta = (\theta_{ij})$, где $i, j = 1, \dots, L$. Система S_i , $i = 1, \dots, L$, включает κ_i одинаковых обслуживающих приборов и очередь емкости $H - \kappa_i$. Состояние сети N определяется вектором $s = (s_i)$, $i = 1, \dots, L$, где s_i — число требований, находящихся в системе S_i . Множество X состояний сети имеет мощность $c_X = |X|$, соответственно $s^{(n)} \in X$, для всех $n = 1, \dots, c_X$. В системе S_i число занятых обслуживающих приборов h_i , число свободных обслуживающих приборов g_i и число требований в очереди b_i зависят от числа требований s_i :

$$h_i = \min\{s_i, \kappa_i\}, \quad g_i = \kappa_i - h_i, \quad b_i = \max\{0, s_i - \kappa_i\}.$$

Все события в системах сети, связанные с завершением обслуживания требований, уходом и поступлением групп требований происходят в дискретные моменты времени, интервалы между которыми называются слотами (длительность слота — единица времени). Длительность обслуживания требований во всех системах S_i , $i = 1, \dots, L$, равная целому числу слотов, является дискретной случайной величиной ξ_i с геометрическим распределением с параметром μ_i , $0 < \mu_i < 1$, принимающей значения $k = 1, 2, \dots$ с вероятностями

$$P\{\xi_i = k\} = (1 - \mu_i)^{k-1} \mu_i.$$

Интенсивность обслуживания требований одним прибором в системе S_i , $i = 1, \dots, L$, также равна μ_i ; $\mu = (\mu_i)$, $i = 1, \dots, L$, — вектор интенсивностей обслуживания в системах сети N .

Подраздел 2.3 посвящен описанию функционирования замкнутой сети с дискретным временем и групповым переходом требований и вычислению ее стационарных характеристик. В каждом слоте в сети N выполняется следующая последовательность действий. В начале слота определяется состояние сети s , в котором сеть пребывает в течение слота. Требования, завершившие обслуживание в конце слота в системе S_i , $i = 1, \dots, L$, образуют группу d_i уходящих из этой системы требований. Из группы d_i , $i \in I$, требования в соответствии с маршрутной матрицей Θ формируют направляемые из S_i в S_j подгруппы d_{ij} требований, $j \in V_i$, V_i — множество номеров выходных смеж-

ных с S_i систем. Из подгрупп d_{ij} , $j \in I$, $i \in U_j$, U_j — множество номеров входных смежных с S_j систем, образуется группа a_j требований, поступающих в S_j . Таким образом, в конце слота формируется вектор $d = (d_i)$, $i = 1, \dots, L$, требований, выходящих после завершения обслуживания из всех систем, после чего он преобразуется в вектор $a = (a_j)$, $j = 1, \dots, L$, требований, поступающих в системы обслуживания сети в следующем слоте. Получается, для каждого состояния сети создаются два множества:

$$D = \left\{ d = (d_1, \dots, d_L) : 0 \leq d_i \leq \kappa_i, i = 1, \dots, L, \sum_{i=1}^L d_i \leq \min\{K, H\} \right\},$$

$$A = \left\{ a = (a_1, \dots, a_L) : 0 \leq a_j \leq \min\{K, H\}, j = 1, \dots, L, \sum_{j=1}^L a_j \leq \min\{K, H\} \right\}.$$

При переходе сети в следующее состояние, новый набор требований в каждой системе будет зависеть от векторов d и a :

$$s^{(n)} = s^{(m)} - d + a,$$

где

— $s^{(n)}$ — следующее состояние системы (в течение следующего слота);

— $s^{(m)}$ — предыдущее состояние системы (в текущем слоте);

Множества векторов выходящих D и входящих A требований указывают на то, что каждый вектор, содержащийся в них, будет формироваться в конце очередного слота с какой-либо вероятностью. Обозначим $p_{s,d}$ — вероятность формирования вектора $d \in D$ уходящих требований при пребывании сети N в состоянии $s \in X$, а ρ_{da} — вероятность преобразования вектора уходящих требований d в вектор поступающих требований $a \in A$. Соответствующие формулы будут выглядеть следующим образом:

$$p_{s,d} = \prod_{i=1}^L \binom{h_i}{d_i} \mu_i^{d_i} (1 - \mu_i)^{h_i - d_i},$$

$$\rho_{da} = \sum_{d_{ij} \in E} \prod_{i=1}^L \binom{d_i}{d_{i1}, \dots, d_{iL}} \prod_{j=1}^L \theta_{ij}^{d_{ij}},$$

при этом:

– множество E представляет собой множество требований из выходных групп каждой системы S_i , $i = 1, \dots, L$, которые попадают в систему с номером j :

$$E = \left\{ d_{ij}, i = 1, \dots, L, j \in V_i : \sum_{i=1}^L d_{ij} = a_j \right\}.$$

Эволюция сети N описывается цепью Маркова с дискретным временем и множеством состояний B . Матрица вероятностей перехода между состояниями цепи $P = (p_{mn})$, $m, n = 1, \dots, c_X$ содержит элементы, определяемые выражением:

$$p_{mn} = \sum_{\substack{d \in D, a \in A: \\ s^{(n)} = s^{(m)} - d + a}} p_{s, da},$$

где

– $p_{s, da} = p_{s, d} \rho_{da}$ — вероятность того, что при состоянии сети s в конце слота сформируется вектор выходящих требований d , а далее он преобразуется в вектор входящих требований a ;

– начальное состояние $s^{(m)}$ отображается в формулах под знаком s .

Стационарное распределение описанной цепи является стационарным распределением сети N : $\pi = (\pi_n)$, $n = 1, \dots, c_X$. Из конечности множества состояний B следует, что распределение π существует и является единственным решением уравнения $\pi = \pi P$ с условием нормировки $\sum_{n \in B} \pi_n = 1$.

Также, в подразделе выведены формулы для вычисления стационарных характеристик сети и систем внутри нее: м.о. числа требований в каждой системе \bar{s}_i , м.о. числа занятых приборов в системах \bar{h}_i , м.о. длительности обслуживания в системе \bar{u}_i , м.о. длительности реакции системы $\bar{\zeta}_i$, интенсивность входящего потока в систему λ_i , коэффициент использования системы ψ_i , пропускная способность Λ . Соответствующие формулы приведены ниже:

$$\bar{s}_i = \sum_{k=1}^H k P\{s_i = k\}, \text{ при этом } P\{s_i = k\} = \sum_{n \in B: s_i^{(n)} = k} \pi_n,$$

$$\bar{h}_i = \sum_{k=1}^H \min\{k, \kappa_i\} P\{s_i = k\}, \quad \bar{u}_i = \frac{\bar{s}_i}{\lambda_i}, \quad \bar{\zeta}_i = \frac{1}{\lambda_i} \sum_{j=1, j \neq i}^L \lambda_j \bar{u}_i, \quad \lambda_i = \bar{h}_i \mu_i,$$

$$\psi_i = \frac{\bar{h}_i}{\kappa_i}, \quad \Lambda = \sum_{i=1}^L \lambda_i.$$

Третий раздел «Оптимизация функционирования СеМО» посвящен описанию методов оптимизации функционирования СеМО с помощью методов динамического управления, которые отличаются тем, к чему именно применяется динамическое управление.

В подразделе 3.1 описываются методы оптимизации, которые могут быть применены для улучшения функционирования сетей массового обслуживания.

В подразделе 3.2 представлен метод оптимизации посредством управления интенсивностями обслуживания. Цель метода - приблизить м.о. требований в системах к базовому распределению s^o , поэтому в множестве состояний сети определяется два подмножества: Y — множество доминантных состояний, или таких, которые отличаются от базового не более чем на граничные значения, а так же Z — множество ординарных состояний. В зависимости от того, к какому множеству относится состояние сети в очередном такте, динамическое управление может применяться (тогда такт называется коррективным), или не использоваться (тогда такт называют нормальным).

Основными действиями, выполняемыми в сети с управлением в момент начала очередного такта, являются:

1. идентификация состояния $s^{(n,k)}$, $n \in B$, сети;
2. проверка принадлежности состояния $s^{(n,k)}$ к множеству доминантных состояний Y посредством проверки выполнения неравенств

$$|s_i^{(n,k)} - s_i^o| \leq q_i, i = 1, \dots, L;$$

3. если данные неравенства выполняются для всех систем сети, то следующий такт является нормальным, и в течение такта $x^{(k)}$ используется вектор μ , после чего выполняется очередной нормальный такт;

4. если хотя бы для одной системы не выполнилось данное неравенство, то такт $x^{(k)}$ является коррективным, формируется вектор $\mu^{(J,k)} = (\mu_i^{(J,k)})$, $i = 1, \dots, L$, $J \in \{1, \dots, c_Z\}$, который соответствует состоянию сети $s^{(n,k)} \in Z$, $n = c_Y + J$, в момент начала такта; этот вектор используется в течение такта $x^{(k)}$, после чего выполняется очередной коррективный такт.

Четвертый раздел «Анализ сети с управлением» посвящен описанию метода исследования сети с управлением интенсивностями обслуживания.

Для описания эволюции сети N^c в течение нормальных и коррективных тактов также используются цепи Маркова \hat{C} и \tilde{C}^J , $J \in \{1, \dots, c_Z\}$, с множеством состояний B и дискретным временем. При этом для цепи, описывающей нормальные такты, множеством начальных состояний будет являться множество состояний $\{1, \dots, c_Y\}$, а для цепи, описывающей коррективные такты, — $\{c_Y + J\}$. Матрицы вероятностей перехода, соответствующие цепям с различными тактами, называются $\hat{P} = (\hat{p}_{mn})$ и $\tilde{P}^J = (\tilde{p}_{mn}^J)$, $m, n = 1, \dots, c_X$. Элементы матрицы \hat{P} вычисляются аналогично элементам матрицы сети без управления. Для вычисления элементов матрицы \tilde{P}^J в этом выражении происходит замена $p_{s,d}$ на $\tilde{p}_{s,d}^J$, в которой вместо заданного изначально вектора интенсивностей обслуживания $\mu = (\mu_i)$ используется $\mu^J = (\mu_i^J)$:

$$\tilde{p}_{s,d}^J = \prod_{i=1}^L \binom{h_i}{d_i} (\mu_i^J)^{d_i} (1 - \mu_i^J)^{h_i - d_i}.$$

Для описания эволюции сети в течение нескольких слотов возведем матрицы вероятностей переходов между состояниями в степень, равную количеству слотов в такте. Тогда верно сказать, что:

$$\hat{P}^{(t)} = (\hat{P})^t, \quad \tilde{P}^{(t),J} = (\tilde{P}^J)^t.$$

Значит, для описания эволюции конкретной сети N^c будем использовать цепь Маркова с дискретным временем с множеством состояний B и матрицей вероятностей перехода $P^c = (p_{mn}^c)$, $m, n = 1, \dots, c_X$, вычисляемой по следующей формуле с учетом заданного значения длительности тактов φ :

$$p_{mn}^c = \begin{cases} \hat{p}_{mn}^{(\varphi)}, & \text{при } m \in \{1, \dots, c_Y\}, \\ \tilde{p}_{mn}^{(\varphi),J}, & \text{при } m = c_Y + J, J \in \{1, \dots, c_Z\}. \end{cases}$$

По аналогии с расчетом вероятностей состояний сети без управления, стационарное распределение $\pi^c = (\pi_n^c)$, $n = 1, \dots, c_X$, вероятностей состояний сети N^c при заданном значении φ является единственным решением уравнения $\pi^c = \pi^c P^c$ с условием нормировки $\sum_{n \in B} \pi_n^c = 1$.

Основные стационарные характеристики систем $S_i, i \in I$, которые включены в сеть N^c , вычисляются по формулам для сети без управления при замене μ_i на $\bar{\mu}_i^c$:

$$\bar{\mu}_i^c = \mu_i \sum_{n=1}^{c_Y} \pi_n^c + \sum_{J=1}^{c_Z} \mu_i^J \pi_{c_Y+J}^c.$$

Пятый раздел «Пример вычисления стационарных характеристик СеМО с управлением интенсивностями обслуживания» посвящен созданию программы на языке Python для вычисления стационарных характеристик сети, и анализу результатов, которые были получены при исследовании замкнутых сетей массового обслуживания с дискретным временем и групповыми переходами требований с управлением интенсивностями обслуживания с различными входными параметрами. Также в этом разделе сформулированы выводы относительно положительного эффекта применения управления.

В подразделе 5.1 описываются две сети N_1 и N_1^c без управления и с ним соответственно. Проводится вычисление стационарных характеристик этих сетей при различном значении длительности тактов, с дальнейшим сравнением. В том числе были вычислены вероятности пребывания сети в множестве доминантных состояний и в базовом состоянии — π_Y и π_1 соответственно. Значения этих вероятностей для сети с управлением при уменьшении длительности тактов увеличиваются и при малой длительности тактов $\varphi = 10$ несколько раз превосходят значения данных характеристик сети без управления. Таким образом, качество функционирования сети N_1^c при $\varphi < 200$ выше, чем сети N_1 , что говорит о возможности достижения высокой эффективности метода динамического управления интенсивностями обслуживания при малых значениях φ .

Также, в данном разделе была найдена закономерность, что при $\varphi \rightarrow \inf$ значения стационарных характеристик стремятся к значениям соответствующих характеристик систем сети без управления.

Также, применение управления интенсивностями обслуживания влечет за собой увеличение интенсивностей потоков требований λ_i в системы сети и пропускной способности сети Λ в целом. Таким образом, исследование данной пары сетей позволяет сделать вывод о том, что управление интен-

сивностями обслуживания улучшает качество функционирования сети N_1^c и дает возможность сбалансировать нагрузку на системы при малых значениях φ . Однако, при многократном увеличении длительности тактов эффект от управления интенсивностями обслуживания уменьшается. Это происходит вследствие уменьшения влияния управления на эволюцию сети N_1^c .

В подразделе 5.2 проводится анализ пары систем N_2 без управления и N_2^c с управлением, в которых увеличено количество систем. При анализе стационарных характеристик сети N_2^c , было отмечено, что управление интенсивностями обслуживания позволило сократить время реакции одной из систем в 2 раза, что в совокупности с остальными изменениями привело к увеличению пропускной способности сети. Исследование данной пары сетей позволяет сделать вывод о том, что увеличение количества систем значительно увеличивает сложность вычислений программы, за счет преумножения количества состояний сети, но не отменяет положительного эффекта от применения управления интенсивностями обслуживания.

В подразделе 5.3 проводится анализ пары систем N_3 без управления и N_3^c с управлением, в которых одинаковое количество приборов в системах. Исследование данной пары сетей позволило сделать вывод о том, что управление интенсивностями обслуживания способно улучшить значения стационарных характеристик сети, которая и без применения управления показывает высокое качество функционирования.

ЗАКЛЮЧЕНИЕ

На данный момент разработано и используется достаточно много математических моделей для реальных сетей и методов их оптимизации, а выбор конкретного метода для решения поставленной задачи зависит от особенностей сети и цели оптимизации. При этом метод управления интенсивностями обслуживания является одним из наиболее применимых методов для оптимизации функционирования сетей с дискретным временем и групповыми переходами требований, так как основывается на анализе и изменении параметров, которые изначально известны и легко изменяемы в реальной сети.

Представленные в данной работе методы вычисления стационарных характеристик сети массового обслуживания с дискретным временем и групповыми переходами требований, а также метод управления интенсивностями

ми обслуживания, позволяют более полно понять принципы функционирования сетей данного вида и рассмотреть алгоритм применения динамического управления к интенсивностям обслуживания требований в системах сети. При проведении анализа нескольких сетей с дискретным временем и групповыми переходами требований, было определено, с достаточной для практических приложений точностью, что использование управления интенсивностями обслуживания способствует улучшению стационарных характеристик и общей пропускной способности сети при малой длительности тактов, а при увеличении длительности тактов влияние управления на эволюцию сети уменьшается.

Таким образом, при написании выпускной квалификационной работы поставленная цель была достигнута, посредством выполнения всех необходимых для этого задач: изучены особенности функционирования замкнутых сетей массового обслуживания с дискретным временем и групповыми переходами требований, подробно рассмотрен метод оптимизации интенсивностей обслуживания, разработана программная реализация модели эволюции исследуемой сети и сделаны выводы о положительном влиянии управления интенсивностями обслуживания на качество функционирования сети.

Основные источники информации:

1. Долгов, В. И. Исследование замкнутых сетей массового обслуживания с дискретным временем, групповыми переходами требований и управлением интенсивностями обслуживания методом имитационного моделирования / В. И. Долгов // Компьютерные науки и информационные технологии : матер. Междунар. науч. конф., Саратов: Издат. центр «Наука», 2016. — С. 145–148.
2. Черушева, Т. В. Теория массового обслуживания : учебное пособие / Т. В. Черушева, Н. В. Зверовщикова. — Пенза: Изд-во ПГУ, 2021. — 224 с.
3. Рогачко, Е. С. Метод анализа замкнутых сетей массового обслуживания с дискретным временем и групповыми переходами требований / Е. С. Рогачко // Компьютерные науки и информационные технологии : матер. Междунар. науч. конф. — Саратов: Издат. центр «Наука», 2016. — С. 342–344.
4. Митрофанов, Ю. И. Метод анализа замкнутых сетей массового обслу-

живания с дискретным временем, групповыми переходами требований и динамическим управлением интенсивностями обслуживания / Ю. И. Митрофанов, В. И. Долгов, Е. С. Рогачко, Е. П. Станкевич // Известия Саратов. ун-та. Нов. сер. Серия Математика. Механика. Информатика. — 2017. — Т. 17, вып. 1. — С. 96–108.

5. Станкевич Е. П. Анализ замкнутых сетей массового обслуживания с дискретным временем, групповыми переходами требований и управлением интенсивностями обслуживания / Е. П. Станкевич // Компьютерные науки и информационные технологии: Материалы Междунар. науч. конф. — Саратов: Издат. центр «Наука». — 2016. — С. 400–403.
6. Маккини, У. Python и анализ данных / У. Маккини; пер. с англ. А. А. Слинкина. — М.: ДМК Пресс, 2020. — 540 с.: ил.