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 ТАНДЕМА ДВУХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С ДИНАМИЧЕСКИМ УПРАВЛЕНИЕМ ОБСЛУЖИВАНИЕМ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 481 группы
направления 27.03.03 — Системный анализ и управление
факультета компьютерных наук и информационных технологий
Татаринцева Святослава Сергеевича

Научный руководитель
доцент, к. ф.-м. н. ______________ Е. С. Рогачко

Заведующий кафедрой
к. ф.-м. н., доцент ______________ И. Е. Тананко

Саратов 2020

ВВЕДЕНИЕ

Актуальность темы. Рассматривается тандемная сеть массового обслуживания, состоящая из двух однеприборных систем массового обслуживания с неограниченным числом мест для ожидания в каждой системе. Такие системы массового обслуживания иначе называются двухфазными системами массового обслуживания [1–3]. Практическое применение тандемные сети массового обслуживания находят в качестве математических моделей систем транспортной логистики и производственных систем, например, сборочных линий. В тандемной сети массового обслуживания может использоваться управление обслуживанием[4–6]. Для рассматриваемой сети массового обслуживания предполагается, что в первой из двух системе требования могут обслуживаться по одному, или, в общем случае, группой либо с интенсивностью $\mu_1 = 0$, либо с $\mu_1 > 0$ в зависимости от состояния сети, определяемого числом требований в системах. Для описания функционирования сети массового обслуживания с управлением используется марковский процесс принятия решений с дискретным временем [2,7,8]. Для данного процесса известными методами [7] может быть найдена оптимальная стратегия управления, минимизирующая стоимость функционирования сети.

Цель бакалаврской работы. Целью бакалаврской работы является изучение метода анализа тандема двух систем массового обслуживания с динамическим управлением обслуживанием и исследование их характеристик.

Поставленная цель определила **следующие задачи**:

1. разработка алгоритма метода анализа тандемной сети массового обслуживания с динамическим управлением обслуживанием;
2. разработка программы для анализа тандема двух систем массового обслуживания с динамическим управлением обслуживанием;
3. исследование тандема двух систем массового обслуживания с динамическим управлением обслуживанием.

Практическая значимость бакалаврской работы. Была разработана программа, реализующая алгоритм метода анализа тандема двух систем массового обслуживания с динамическим управлением обслуживанием. Программа позволяет вычислять оптимальные средние затраты и оптимальную стратегию управления. С помощью программы можно исследовать реальные системы, моделируемые рассматриваемым тандемом двух систем массового обслуживания, а также изучать зависимость оптимальных средних затрат от изменения различных параметров систем.

Структура и объём работы. Бакалаврская работа состоит из введения, шести разделов, заключения, списка использованных источников и двух приложений. Общий объём работы — 61 страница, из них 45 страниц — основное содержание, включая 17 рисунков и 5 таблиц, список использованных источников информации — 20 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Описание тандемной сети массового обслуживания с управлением обслуживанием» посвящен описанию рассматриваемой тандемной сети массового обслуживания с управлением обслуживанием, а также формулировке задачи оптимального управления для описанной сети.

Рассматривается тандемная сеть массового обслуживания, состоящая из двух одноприборных систем массового обслуживания [9]. Поток поступления требований в первую систему обслуживания является пуассоновским с интенсивностью λ. Длительности обслуживания требований имеют экспоненциальное распределение. Интенсивность обслуживания в первой системе равна μ_1. После завершения обслуживания в первой системе требования поступают во вторую систему, в которой обслуживаются с интенсивностью μ_2. Для существования стационарного режима функционирования сети предполагается, что $\lambda < \mu_1$ и $\lambda < \mu_2$.

Формулируется задача оптимального управления для описанной сети массового обслуживания. Вводится постоянная стоимость ожидания c_{wait} для одного требования в единицу времени. Также для требований, находящихся во второй системе обслуживания, вводится дополнительная стоимость c_{loc}. Таким образом, стоимость ожидания в первой системе равна $c_1 = c_{wait}$ для одного требования в единицу времени, а во второй системе — $c_2 = c_{wait} + c_{loc}$. Предполагается, что $0 < c_1 < c_2$. В следствие того, что во второй системе стоимость ожидания выше, предпочтительным является пребывание требований в первой системе, чем во второй системе. Однако, необходимо избегать простоя второй системы обслуживания, когда в первой системе есть требования.

Функционирование рассмотренной сети массового обслуживания описывается марковским процессом принятия решений. Состояние сети наблюдается в моменты поступления и завершения обслуживания требований, то есть в дискретные моменты времени. Используется униформизация, чтобы перейти к марковскому процессу с дискретным временем [10]. Марковский процесс принятия решений с дискретным временем определяется набором $\{S, A, P, C\}$. Здесь S — множество состояний сети; состояние сети определяется вектором $i = (x_1, x_2) \in N^2$, где x_i — число требований в системе обслуживания i, $i = 1, 2$. A — множество управлений; $A = \{0, 1\}$, где управление 0 обозначает блокировку обслуживания в первой системе (то есть требования не могут перейти из первой системы во вторую), управление 1 обозначает, что требования обслуживаются в первой системе с интенсивностью обслуживания μ_1 и затем переходят из первой системы во вторую систему. Матрица P является матрицей вероятностей переходов из состояния i в состояние j при управлении $a \in A$; элементы матрицы P обозначаются $p^a(i, j)$. C обозначает функцию затрат; $c^a(i)$ — непосредственно ожидаемые затраты в единицу времени для каждого состояния $i \in S$ и управления $a \in \{0, 1\}$.

Оптимальная стратегия управления удовлетворяет уравнению Беллмана:

$$V^*(i) + g^* = min_{a\in A}(c^a(i) + \sum_{j\in S} p^a(i, j)V^*(j)), \quad i \in S, \tag{1}$$

где g^* — оптимальные средние затраты, а $V^*(i)$ — относительные веса. Оптимальные управления могут быть определены следующим образом:

$$f(i) = argmin_{a\in A}(c^a(i) + \sum_{j\in S} p^a(i, j)V^*(j)), \quad i \in S, \tag{2}$$

где $V^*(j)$ удовлетворяют уравнениям

$$V^*(i) + g^* = c^f(i) + \sum_{j\in S} p^f(i, j)V^*(j).$$

Таким образом, необходимо минимизировать средние затраты и определить оптимальное управление для каждого состояния.

Чтобы найти оптимальные управления для каждого состояния, минимизирующие средние затраты, используется метод итераций по значениям, в основе которого лежат следующие рекуррентные соотношения:

$$V_n(i) = \min_{a \in A}(c^a(i) + \sum_{j \in S} p^a(i,j)V_{n-1}(j)),$$

$$f_n(i) = argmin_{a \in A}(c^a(i) + \sum_{j \in S} p^a(i,j)V_{n-1}(j)).$$

Второй раздел «Структурные свойства оптимальной стратегии управления» посвящён описанию структурных свойств оптимальной стратегии управления.

Рассматривается пример тандемной сети массового обслуживания [9]. Описываемые далее структурные свойства оптимальной стратегии управления сохраняются для любых значений c_1 и c_2, удовлетворяющих условию $0 < c_1 < c_2$.

Оптимальная стратегия управления является пороговой и графически представляется так называемой, кривой переключения, разделяющей множество всех состояний рассматриваемой сети на две области и отделяющей область состояний сети, в которых применяется управление 0, от области состояний сети, в которых применяется управление 1.

Чтобы лучше понять динамику сети массового обслуживания, функционирующей при оптимальной пороговой стратегии управления, были построены графики траектории управления и показаны возможные отклонения от кривой переключения для двух случаев $\mu_1 < \mu_2$ и $\mu_1 > \mu_2$.

Линейность кривой переключения в случае $\mu_1 < \mu_2$ объясняется следующим образом. Оптимальная стратегия управления должна обеспечивать загруженность второй системы, когда в первой системе есть требования. Поэтому нежелательными являются состояния сети, при которых нет требований во второй системе. Если $\mu_1 < \mu_2$, то первая система не может "догнать"вторую систему и должна всегда обеспечивать достаточный входящий поток требований для второй системы.

Когда $\mu_1 > \mu_2$, первая система может "догнать"вторую систему, таким образом, снижается вероятность простоя второй системы. Все траектории ведут к кривой переключения, а затем продолжаются вдоль кривой переключения в направлении начала координат. Следовательно, число требований во второй системе может поддерживаться на низком уровне и, таким образом, кривая переключения приближается к горизонтальной оси при больших значениях числа требований в первой системе.

Третий раздел «Применение матрично-геометрического метода для поиска пороговой стратегии управления» посвящён описанию матрично-геометрического метода для поиска пороговой стратегии управления.

Чтобы сравнивать эффективность пороговых стратегий управления, определяется характеристика качества функционирования сети массового обслуживания как функция порогового значения K и затем выбирается наилучшее значение K. Функционирование рассматриваемой сети обслуживания описывается обобщенным процессом размножения и гибели. Инфинитезимальный оператор Q данного процесса является блочно-диагональной матрицей. Каждому макросостоянию обобщенного процесса размножения и гибели соответствует фиксированное число требований в первой системе обслуживания, а состояния процесса, соответствующие одному макросостоянию, определяются числом требований во второй системе обслуживания.

Множеством состояний обобщенного процесса размножения и гибели, описывающего функционирование рассматриваемой сети массового обслуживания, является множество

$$S = ((x_1, x_2) : x_1 \in Z^+, 0 \leqslant x_2 \leqslant K).$$

Здесь x_1 обозначает число требований в первой системе и определяет номер макросостояния, а x_2 обозначает число требований во второй системе и определяет номер состояния внутри макросостояния. Максимальное число требований во второй системе ограничено пороговым значением K.

Блочная повторяющаяся структура матрицы Q приводит к матрично-геометрической форме стационарного распределения. Пусть $\underline{\pi_i} = (\pi_{i0}, ..., \pi_{iK})$ $(K+1)$-мерные векторы стационарных вероятностей состояний сети массового обслуживания; $\pi_{x_1 x_2}$ — стационарная вероятность пребывания x_1 требований в первой системе и x_2 требований во второй системе. Тогда из уравнений равновесия:

$$\underline{\pi}_{i-1}\Lambda + \underline{\pi}_i D + \underline{\pi}_{i+1} M = 0, \quad i \geqslant 1,$$

определяется

$$\underline{\pi_i} = \underline{\pi_0} R^i,$$

где матрица R — минимальное неотрицательное решение следующего квадратного матричного уравнения:

$$R^2 M + RD + \Lambda = R.$$

Для того чтобы вычислить значение функции стоимости, определяется математическое ожидание числа требований для каждой системы обслуживания:

$$E[x_1] = \underline{\pi}_0 R(I - R)^{-2} e,$$

$$E[x_2] = \underline{\pi}_0 (I - R)^{-1} J,$$

где J — вектор-столбец $(0, 1, ..., K)^T$, e — $(K + 1)$-мерный вектор, все компоненты которого равны 1.

Чтобы определить оптимальное пороговое значение, минимизируются затраты по всем возможным пороговым значениям:

$$\min_{K \in N} \{\underline{\pi}_0 (I - R)^{-1} \cdot (c_1 R(I - R)^{-1} e + c_2 J)\}. \tag{3}$$

Можно определить наилучшее пороговое значение по отношению к затратам на функционирование сети массового обслуживания и сравнить его с результатами вычисления оптимальной стратегии управления для марковского процесса принятия решений. Найденное пороговое значение соответствует оптимальной пороговой стратегии. Определение оптимального порогового значения в вычислительном отношении существенно проще, чем вычисление оптимальной стратегии управления с использованием соответствующих методов для марковских процессов принятия решений.

Четвертый раздел «Управление групповым обслуживанием в тандемной сети массового обслуживания» посвящён описанию группового обслуживания в тандемной сети массового обслуживания.

Рассматривается обобщение тандемной сети массового обслуживания с управлением обслуживанием, описанной в первом разделе. В первой системе одним обслуживающим прибором могут одновременно обслуживаться N требований, в отличие от сети обслуживания, рассмотренной в первом разделе, в которой в первой системе требования обслуживаются по одному.

Так как в первой системе разрешается групповое обслуживание, множество управлений сети расширяется от $A = \{0, 1\}$ до $A = \{0, ..., x_1\}$, где x_1 — число требований в первой системе (то есть множество возможных управлений теперь зависит от текущего состояния сети). Значение $a \in A$ соответствует размеру выбранной группы требований и ограничено, очевидно, числом требований в первой системе. Вероятности переходов между состояниями сети определяются аналогично, как в первом разделе.

При групповом обслуживании оптимальная стратегия управления также, как и в случае, описанном во втором разделе, является пороговой и определяется кривой переключения. Оптимальная стратегия управления может быть охарактеризована единственной кривой переключения, отделяющей область состояний сети, в которых выбирается управление 0, от области всех остальных состояний, в которых выбираются другие управления.

Функционирование тандемной сети массового обслуживания с групповым обслуживанием требований в первой системе описывается цепью Маркова типа $GI/M/1$, что позволяет применить матрично-геометрический метод для нахождения стационарного распределения сети.

После вычисления стационарных распределений сети для различных пороговых значений K, можно определить наилучшее пороговое значение на основе минимизации средних затрат с использованием (3).

Пятый раздел «Описание алгоритма и программы для анализа тандема двух систем массового обслуживания с динамическим управлением обслуживанием» посвящён описанию алгоритма и программы для анализа тандема двух систем массового обслуживания с динамическим управлением обслуживанием.

В подразделе 5.1 описывается алгоритм метода анализа тандема двух систем массового обслуживания с динамическим управлением обслуживанием и приводится блок-схема алгоритма. Алгоритм состоит из пяти блоков, которые выполняются последовательно.

В первом блоке задаются исходные параметры сети.

Во втором блоке формируется множество состояний рассматриваемой тандемной сети массового обслуживания.

В третьем блоке формируется множество управлений.

В четвертом блоке описывается итерационный цикл, являющийся реализацией метода итераций по значениям для поиска оптимальных средних затрат и оптимальной стратегии управления.

В пятом блоке описывается вывод результатов.

В подразделе 5.2 описывается программа для анализа тандема двух систем массового обслуживания с динамическим управлением обслуживанием. Описывается назначение программы, её интерфейс и правила использования, приводится список идентификаторов, используемых в программе, производится описание функций, имеющихся в программе, а также приводится пример использования программы.

Шестой раздел «Результаты исследования тандема двух систем массового обслуживания с динамическим управлением обслуживанием» посвящён описанию результатов исследования тандема двух систем массового обслуживания с динамическим управлением обслуживанием.

На числовых примерах исследуется зависимость оптимальных средних затрат g от:

- интенсивности потока требований поступающих в первую систему λ;
- интенсивности обслуживания требований во второй системе μ_2;
- стоимости ожидания для требований в первой системе c_1;
- соотношения интенсивностей обслуживания в системах.

С помощью программы, описанной в разделе 5, для каждого примера были найдены оптимальные средние затраты сети и соответствующая оптимальная стратегия управления.

Как следует из результатов исследования, при увеличении интенсивности поступления требований в первую систему увеличиваются оптимальные средние затраты. Это объясняется тем, что в этом случае увеличивается число требований, пребывающих в сети, и соответственно, увеличиваются затраты на их пребывание в сети.

Также при увеличении стоимости ожидания требований в первой системе увеличиваются оптимальные средние затраты. При большем значении c_1 блокировка прибора в первой системе происходит при большем значении числа требований во второй системе. Это объясняется тем, что становится более невыгодным останавливать обслуживание требований в первой системе из-за высокой стоимости содержания требований в ней.

ЗАКЛЮЧЕНИЕ

В бакалаврской работе был рассмотрен метод анализа тандема двух систем массового обслуживания с динамическим управлением обслуживанием и проведено исследование его характеристик. Данный метод позволяет найти оптимальные средние затраты сети и соответствующую оптимальную стратегию управления. Был разработан алгоритм этого метода анализа и выполнено его подробное описание. Алгоритм реализован в виде программы на ЭВМ на языке $C\#$ в среде разработки Microsoft VisualStudio 2019.

С помощью разработанной программы были проведены исследования тандема двух систем массового обслуживания с динамическим управлением обслуживанием. На основе полученных результатов исследования было установлено, что при большей интенсивности обслуживания требований в первой системе в сравнении с интенсивностью обслуживания требований во второй системе, значение оптимальных средних затрат увеличивается, а оптимальная стратегия управления сохраняет свои структурные свойства. Также была исследована зависимость оптимальных средних затрат от изменения различных параметров тандема систем массового обслуживания.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Саати, Т. Л. Элементы теории массового обслуживания и ее приложения / Т. Л. Саати; пер. с англ. Е. Г. Коваленко; под ред. И. Н. Коваленко, Р. Д. Когана. – М. : Сов. радио, 1965. – 511 с.

2 Таха, Х. Введение в исследование операций. В 2-х книгах. Кн. 2. / Х. Таха; пер. с англ. – М. : Мир, 1985. – 496 с.

3 Новиков, О. А. Прикладные вопросы теории массового обслуживания / О. А. Новиков, С. И. Петухов. – М. : Сов. радио, 1969. – 400 с.

4 Gajrat, A. Large-deviations analysis of the fluid approximation for a controllable tandem queue / A. Gajrat, A. Hordijk, A. Ridder // The Annals of Applied Probability. Vol. 13, No.4, - P. 1423-1448.

5 Rosberg, Z. Optimal control of service in tandem queues / Z. Rosberg, P.P. Varaiya, J.C. Walrand // IEEE Transactions on automatic control. - 1982. - Vol. AC-27, No.3. - P. 600–610.
6 Weber, R. Optimal Control Of Service Rates In Networks Of Queues / R. Weber // Applied Probability Trust. - 1987. - Prob. 19. - P. 202-218.
7 Ховард, Р. А. Динамическое программирование и марковские процессы / Р. А. Ховард; пер. с англ. В. В. Рыкова; под ред. Н. П. Бусленко. – М. : Сов. радио, 1964. – 189 с.
8 Майн, Х. Марковские процессы принятия решений /Х. Майн, С. Осаки; пер. с англ. В. В. Калашникова; под ред. Н. П. Бусленко. – М. : Наука, 1977. – 176 с.
9 Boucherie, R. J. Markov Decision Processes in Practice / R. J. Boucherie, N. M. van Dijk. – Cham : Springer, 2017.– 550 pp.
10 Puterman, M. L. Markov Decision Processes: Discrete Stochastic Dynamic Programming / M. L. Puterman. – Hoboken : Wiley, 1994. – 649 pp.