

Санкт-Петербургский государственный университет

БАШЛАК Владислав Сергеевич

Выпускная квалификационная работа

Управление потоковыми переключательными сетями

Уровень образования: **бакалавриат**

Направление: **01.03.02 «Прикладная математика и информатика»**,

Основная образовательная программа: **СВ.5004.2017 «Прикладная математика и информатика»**

Профиль (при наличии) _____

Научный руководитель: кафедра
теоретической кибернетики СПбГУ,
профессор, доктор физико-
математических наук, Матвеев
Алексей Серафимович

Рецензент: Университет ИТМО,
профессор, доктор технических
наук. Фуртат Игорь Борисович

Санкт-Петербург
2021

Saint Petersburg State University

BASHLAK Vladislav Sergeevich

Diploma Thesis

Flow Switching Networks Management

Education level: **bachelor's degree**

Direction: **01.03.02 «Applied Mathematics and Informatics»**,

Main educational program: **CB.5004.2017 «Applied Mathematics and Informatics»**

Profile (if available) _____

Scientific Supervisor: Department of
Theoretical Cybernetics, St.
Petersburg State University,
Professor, Doctor of Physical and
Mathematical Sciences, Matveev
Alexey Serafimovich

Reviewer: ITMO University,
professor, doctor of technical
sciences. Furtat Igor Borisovich

Saint Petersburg
2021

Оглавление

Введение	5
Глава 1. Математическая модель и основные теоремы	8
1.1. Описание математической модели	8
1.2. Критерии стабилизируемости в общем случае	10
Глава 2. Доказательство теорем и стабилизирующий протокол управления	13
2.1. Доказательство Теоремы 1	13
2.2. Доказательство Теоремы 2	21
Глава 3. Критерии устойчивости в специальных случаях	23
3.1. Вспомогательная лемма	23
3.2. Случай с непересекающимися пакетами	25
3.3. Случай с пересечением двух пакетов (из двух буферов) в одном буфере с μ_0 , не зависящей от пакета . . .	27
3.4. Случай с пересечением двух пакетов в одном буфере с μ_0 , не зависящей от пакета	29
3.5. Случай с пересечением двух пакетов (из двух буферов) в одном буфере с $\mu_{b,k}$, зависящей от пакета	32
3.6. Случай с пересечением двух пакетов в одном буфере с $\mu_{b,k}$, зависящей от пакета	36
Глава 4. Проверка критерия стабилизируемости с помощью Matlab	41

4.1. Описание программы	41
4.2. Исследование стабилизируемости на примерах	43
Заключение	47
Список литературы	48

Введение

Центральным объектом исследований данной работы является некоторое обобщение детерминированной модели системы поллинга [1]. Стандартная модель такой системы состоит из независимых потоков некоторой субстанции, интерпретируемой как жидкость и именуемой "работой", вливающих в буферы неограниченного объема, и одного сервера, обслуживающего эти буферы индивидуально и поочередно, переключаясь между ними за ненулевое время. Обслуживание буфера состоит в изъятии его содержимого вовне системы. Рассматриваемое обобщение такой модели мотивировано ситуацией, когда сервер имеет возможность обслуживать одновременно некоторую совокупность буферов, называемую пакетом. В данной работе рассматриваются как достаточно общая модель системы поллинга, понимаемой в указанном широком смысле, так и несколько частных случаев системы, определяемых способом пересечения ее пакетов и ограничениями на скорости обработки буферов.

Применение систем поллинга простирается на многие сферы: химическая кинетика, транспортные сети, компьютерные и коммуникационные сети, гибкие производственные системы, системы ремонта и т. д. [1, 2] Сам же термин отсылает к аналогии из жизни, когда садовник перебегает от одного куста к другому для стрижки. При этом все кусты растут непрерывно, а задачей садовника является поддержание сада в ухоженном состоянии.

Приведем пример из области компьютерных сетей — рассмотрим систему разделенного времени (с англ. Time-sharing). Она состоит из центрального компьютера и нескольких терминалов, связанных с центральным компьютером (Рис. 1). Переключаясь между различ-

ными терминалами, центральный компьютер выделяет необходимые ресурсы на выполнение внутренних команд терминалов. В простом случае обслуживаться может только один терминал, в более сложном случае выделение ресурсов происходит сразу нескольким терминалам. Подробно такая сеть разбирается в [3].

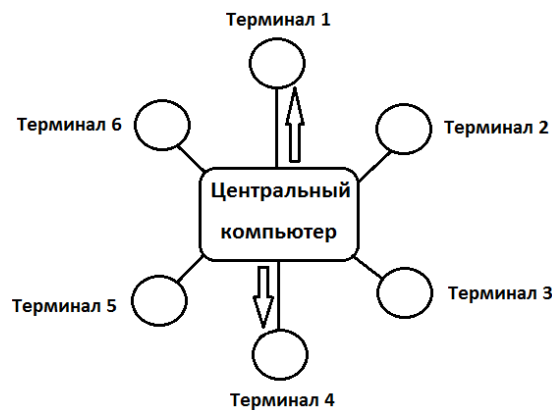


Рис. 1. Система разделения времени.

Динамика системы поллинга является гибридной, так как состояние такой системы описывается как вектором непрерывно меняющихся во времени содержимых буферов, так и дискретной переменной, описывающей изменчивое состояние сервера. Управление системой включает в себя два аспекта. Во-первых, это протокол переключения для сервера, который контролирует очередность перехода от пакета к пакету. А во-вторых, это протокол обслуживания, регулирующий время работы сервера с пакетом и скорость обработки содержимого каждого буфера в этом пакете. Таким образом, алгоритм управления диктует, когда и с какими пакетами работает сервер, и каким образом.

Существуют различные общие подходы к анализу и синтезу про-

токолов управления, они достаточно хорошо разобраны в [2] и [4]. Исследованию систем поллинга, частично с использованием этих подходов, посвящена обширная литература, с обзором которой можно ознакомиться в [1, 5, 6]. Вместе с тем ее подавляющая часть рассматривает классическую, "не-пакетную" схему поллинга, в то время, как исследование пакетной схемы находится в начальной стадии, и применительно к ней многие вопросы пока остаются открытыми. Среди них вопросы устойчивости и стабилизируемости системы.

Минимальное требование к протоколу управления — устойчивость замкнутой системы, понимаемая как ограниченность суммарной очереди в течение всего времени работы системы (в качестве которого в этом определении рассматривают интервал времени от его начального значения до бесконечности) независимо от начального состояния. В случае катастрофического роста содержимого буферов корректно использовать систему практически невозможно, поэтому так важен алгоритм управления, делающий систему поллинга устойчивой.

Целями работы являются:

- Вывод критериев стабилизируемости системы поллинга с сервером, обслуживающим целые пакеты, в общем виде и в частных случаях;
- Построение стабилизирующего протокола управления для стабилизируемой системы;
- Проверка критерия стабилизируемости с помощью программы на Matlab.

Глава 1

Математическая модель и основные теоремы

1.1. Описание математической модели

Теперь введем изучаемую далее модель *системы поллинга*. Имеем один *сервер*, обслуживающий n неограниченных *буферов*. λ_b — постоянная скорость поступающего потока работы извне для буфера b . Через x_b обозначим неотрицательное содержимое b -го буфера. Сервер подключается к буферам через *пакеты* P_1, P_2, \dots, P_s , каждый из которых — это совокупность нескольких буферов, и обслуживает все буферы одновременно в рамках одного пакета P_k ($P_k \neq \emptyset \forall k$). При этом тратится ненулевое время $\delta_{k' \rightarrow k''}$ на переключение между пакетами $P_{k'} \rightarrow P_{k''}$. При этом не существует буфера, который не принадлежит ни одному пакету, то есть $\bigcup_{k=1}^s P_k = [1 : n]$. Подключаясь к буферу b (из соответствующего пакета), сервер обслуживает его на скорости $u_b(t) \geq 0$ [7]. Эта управляющая переменная имеет свои ограничения в зависимости от структуры пакетов и системы поллинга. Далее рассмотрим следующие случаи:

- **Общий случай:** Пакеты пересекаются и максимальная скорость $\mu_{b,k}$ изъятия содержимого из буфера b зависит от пакета k ;
- **Частные случаи** (общие результаты будут конкретизированы для них):

- 1) Пакеты не пересекаются, и максимальная скорость изъятия работы из буфера — μ_b ;
- 2) Два пакета по два буфера в каждом пересекаются в одном общем, но максимальная скорость изъятия содержимого из общего буфера μ_b не зависит от подключенного пакета;
- 3) Два пакета с произвольным числом буферов в каждом пересекаются в одном общем буфере, но максимальная скорость изъятия содержимого из общего буфера μ_b не зависит от подключенного пакета;
- 4) Два пакета по два буфера в каждом пересекаются в одном общем, и максимальная скорость изъятия содержимого из общего буфера $\mu_{b,k}$ зависит от подключенного пакета;
- 5) Два пакета пересекаются, и максимальная скорость обслуживания общего буфера $\mu_{b,k}$ зависит от пакета.

Методика вывода более простых примеров будет способствовать получению более сложных результатов.

Суммарная очередь на обслуживание

$$0 \leq w(t) = \sum_{b=1}^n c_b x_b(t) \quad (1.1)$$

определяет качество обслуживания, где $c_b > 0$ — коэффициенты значимости буфера. Именно от этой переменной (1.1) отталкиваются в определении устойчивой системы поллинга. *Состояние* системы в момент времени t описывается парой функций $(X(t), q(t))$, где $X = \{x_b\}_{b=1}^n$ - вектор содержимого буферов ($x_b \geq 0$), а q - дискрет-

ная переменная состояния сервера

$$q = \begin{cases} k \in [1 : s], & \text{если сервер обслуживает пакет } P_k; \\ \otimes, & \text{если сервер переключается между пакетами.} \end{cases}$$

Процессом же назовем изменение состояния системы в течение некоторого времени, развитие сети. Процесс назовем *устойчивым*, если суммарная очередь ограничена с течением времени

$$\overline{\lim}_{t \rightarrow \infty} w(t) < \bar{w}.$$

Протокол или *политика управления* — это алгоритм, регулирующий работу сервера и отвечающий на следующие вопросы:

- На какой скорости обслуживать буфера текущего пакета;
- Когда останавливать обслуживание пакета;
- Какой пакет является следующим после текущего.

Протокол назовем *корректным*, если из любого начального состояния он порождает единственный процесс, определенный при $t \geq 0$. Протокол называется *стабилизирующим*, если он корректен и из любого начального состояния он порождает устойчивый процесс. Соответственно систему поллинга назовем *стабилизируемой*, если для нее существует стабилизирующий протокол [8].

1.2. Критерии стабилизируемости в общем случае

Пакет P_{k^*} назовем *исчерпывающе-достаточным*, если

- 1) он охватывает все буфера $P_{k^*} = \{1, \dots, n\}$;
- 2) $\mu_{b,k^*} \geq \lambda_b \forall b$.

Тривиальная ситуация: *Существует исчерпывающе-достаточный пакет P_{k^*} .* Тогда система стабилизируема, причем сервер может не переключаться, постоянно пребывая в режиме этого пакета. Все дальнейшие рассуждения данной работы проводим, предполагая следующее.

Предположение 1. *В системе поллинга не существует исчерпывающе-достаточного пакета.*

Сформулируем две теоремы, связывающие стабилизируемость системы со специфическими соотношениями для заданных параметров λ_b и $\mu_{b,k}$. Первая теорема сводит вопрос о стабилизируемости системы поллинга к разрешимости (существованию решения) специальной системы линейных неравенств. Тест на разрешимость таких систем — стандартная тема линейного программирования. В материалах [9], [10] рассматриваются способы решения таких задач в среде Matlab.

Теорема 1. *Система поллинга стабилизируема тогда и только тогда, когда сопоставив каждому пакету P_k число τ_k , можно получить решение следующей системы соотношений*

$$\sum_{k=1}^s \tau_k < 1, \quad \tau_1 \geq 0, \dots, \tau_s \geq 0, \quad (1.2)$$

$$\sum_{k=1}^s \mu_{b,k} \tau_k \geq \lambda_b \quad \forall b = 1, \dots, n. \quad (1.3)$$

Следующая теорема предлагает эквивалентный критерий, который апеллирует не к разрешимости системы линейных неравенств, а к решению задачи выпуклого программирования. При решении некоторых вопросов такая форма оказывается более удобной.

Теорема 2. Система линейных неравенств (1.2), (1.3) имеет решение тогда и только тогда, когда

$$\min_{\chi \in \mathbb{S}_n} \max_{k=1, \dots, s} \sum_{b=1}^n \mu_{b,k} \chi_b > 1, \quad (1.4)$$

где

$$\mathbb{S}_n = \left\{ \chi = (\chi_1, \dots, \chi_n) \in \mathbb{R}^n : \chi_b \geq 0 \forall b, \sum_{b=1}^n \lambda_b \chi_b = 1 \right\}$$

— стандартный симплекс (стандартный в переменных $z_b := \lambda_b \chi_b$).

Глава 2

Доказательство теорем и стабилизирующий протокол управления

2.1. Доказательство Теоремы 1

Необходимость условий (1.2) и (1.3). Предположим, что система поллинга стабилизируема, то есть для нее существует стабилизирующий протокол. По определению из любого начального состояния он порождает устойчивый процесс. Выберем какое-либо начальное состояние и рассмотрим соответствующий устойчивый процесс; по определению устойчивости, для него существует такая константа $\bar{w} < \infty$, что $w(t) \leq \bar{w} \forall t \geq 0$. Тогда согласно (1.1),

$$x_b(t) \leq \bar{x}_b := \frac{\bar{w}}{c_b} \quad \forall b, \forall t \geq 0. \quad (2.1)$$

Для любого интервала $I \subset [0, \infty)$ вводим следующие обозначения:

- $t_k(I)$, суммарное время, которое сервер потратил на работу в режиме пакета P_k на интервале I ;
- $\text{sw}(I)$, общее время, которое затратил сервер на переключение между пакетами на этом интервале;
- $|I|$, длина интервала I .

Здесь первые две величины строятся исходя из выбранного устойчивого процесса.

Лемма 1. Пусть I — временной интервал, на котором нет переключений, то есть сервер работает только с одним пакетом R_{k^*} . Существует такая константа \bar{p} , что

$$|I| \leq \bar{p}. \quad (2.2)$$

Доказательство. Действительно, пакет R_{k^*} не является исчерпывающе-достаточным по **Предположению 1**. Это означает, что на интервале I один из буферов b либо 1) не обслуживается либо 2) обслуживается на скорости $u_b(t) \leq \mu_{b,k^*} < \lambda_b$. Полагая

$$\Delta^* := \min_{b=1,\dots,n} \left[\lambda_b - \max_{k:\mu_{b,k} < \lambda_b} \mu_{b,k} \right] > 0$$

(где \max по пустому множеству считаем равным нулю), заключаем, что на интервале I содержимое буфера b растёт со скоростью не ниже Δ^* и в целом за интервал возрастает не менее, чем на $|I|\Delta^*$. Отсюда в силу (2.1) получаем (2.2) с $\bar{p} := (\Delta^*)^{-1} \max_b \bar{x}_b$. \square

Теперь заметим, что любой интервал времени, полностью занятый переключением сервера между двумя пакетами (назовем такой интервал *интервалом переключения*), имеет длительность не более $\delta_+ := \max_{k \neq k'} \delta_{k \rightarrow k'} < \infty$, где максимум достигается, так как пакетов конечное число. Из изложенного следует, что любой интервал длительности $|I| > \bar{p} + 2\delta_+$ включает в себя не менее

$$\left\lfloor \frac{|I| - \delta_+}{\bar{p} + \delta_+} \right\rfloor$$

интервалов переключения. Следовательно,

$$\text{sw}(I) \geq \delta_- \left[\frac{|I| - \delta_+}{\bar{p} + \delta_+} \right], \quad \text{где } \delta_- := \min_{k \neq k'} \delta_{k \rightarrow k'} > 0. \quad (2.3)$$

Рассмотрим неограниченно возрастающую последовательность интервалов $I_j := [0, T_j]$, где $T_j \rightarrow \infty$ при $j \rightarrow \infty$. При любом j все компоненты вектора $[t_1(I_j)/|I_j|, \dots, t_s(I_j)/|I_j|, \text{sw}(I_j)/|I_j|] \in \mathbb{R}^{s+1}$ лежат в $[0, 1]$. Следовательно из последовательности этих векторов можно выбрать сходящуюся подпоследовательность. Присваивая ей исходное обозначение, можно таким образом без ограничения общности считать, что существуют пределы

$$\tau_k := \lim_{j \rightarrow \infty} \frac{t_k(I_j)}{|I_j|} \quad k = 1, \dots, s, \quad \text{sw}^\infty := \lim_{j \rightarrow \infty} \frac{\text{sw}(I_j)}{|I_j|}. \quad (2.4)$$

При этом из (2.3) вытекает, что

$$\text{sw}^\infty \geq \frac{\delta_-}{\bar{p} + \delta_+} > 0. \quad (2.5)$$

Лемма 2. Предел τ_k из (2.4) удовлетворяет неравенствам (1.2) и (1.3).

Доказательство. Неотрицательность τ_k очевидна. Так как интервал времени I полностью распределен между обработкой пакетов и переключениями между ними, и одновременная обработка разных пакетов невозможна, имеем

$$|I_j| = \text{sw}(I_j) + \sum_{k=1}^s t_k(I_j).$$

Поделив на $|I_j|$ и перейдя к пределу при $i \rightarrow \infty$ с учетом (2.5), убеждаемся в справедливости (1.2).

Распишем изменение содержимого для одного буфера на интервале I_j :

$$\begin{aligned} \bar{x}_b &\geq x_b(T_j) = |I_j|\lambda_b + x_b(0) - \sum_{k:b \in P_k} \int_{I_j} u_{b,k}(t) dt \geq \\ &\geq |I_j|\lambda_b + x_b(0) - \sum_{k:b \in P_k} \mu_{b,k} t_k(I_j) \Rightarrow |I_j|\lambda_b \leq \sum_{k:b \in P_k} \mu_{b,k} t_k(I_j) + \bar{x}_b - \\ &- x_b(0) \Rightarrow \lambda_b \leq \sum_{k:b \in P_k} \mu_{b,k} \frac{t_k(I_j)}{|I_j|} + \frac{\bar{x}_b - x_b(0)}{|I_j|} \xrightarrow{j \rightarrow \infty} \sum_{k:b \in P_k} \mu_{b,k} \tau_k. \end{aligned}$$

Таким образом, выполнено (1.3). \square

Необходимость доказана.

Достаточность. Имеем решение $\{\tau_k\}$ системы (1.2), (1.3). Рассмотрим взаимно-однозначную нумерацию пакетов $Z_\sigma = \{P_0, \dots, P_{s-1}\}$ и соответствующую перестановку компонент τ_k решения системы. Для любого буфера b , индивидуально скорректируем решение $\{\tau_k\}$: уменьшая надлежащим образом τ_k до некоторого значения $\tau_{b,k} \in [0, \tau_k]$, добьемся преобразования (1.3) из неравенства в равенство, то есть обеспечим выполнение соотношения

$$\sum_{k=0}^{s-1} \mu_{b,k} \tau_{b,k} = \lambda_b \quad \forall b = 1, \dots, n. \quad (2.6)$$

Введем протокол управления, определяемый параметром $T > 0$ и подчиняющий работу сервера следующим правилам.

Стабилизирующий протокол управления.

- r1)** На интервале $[0, T]$, сервер переключается в режим пакета P_0 , после чего простаивает (обслуживает буфера на нулевой скорости) до момента T , если переключение завершено раньше (этот шаг связан с циклическим обходом сервера по пакетам — начать нужно с последующим пакетом);
- r2)** На любом интервале вида $[jT, (j+1)T]$ ($j \geq 1$) сервер выполняет следующие операции:

- A)** Осуществляет цепочку переключений между всеми пакетами, следуя циклическому порядку

$$P_0 \mapsto \dots \mapsto P_{s-1} \mapsto P_0, \quad (2.7)$$

начиная и заканчивая обход в пакете P_0 ;

- B)** В режиме очередного пакета k , сервер изымает содержимое каждого буфера b пакета на максимально возможной скорости $\mu_{b,k}$ в течение $\tau_{b,k}T$ единиц времени, а затем на нулевой скорости до истечения $\tau_k T$ единиц времени с начала обслуживания пакета;
- C)** После этого сервер переключается в режим следующего пакета, то есть пакета P_{j+1} , если $j < s - 1$, и P_0 , если $j = s - 1$ (можно кратко записать в виде $P_j \mapsto P_{j \oplus 1} \forall j$, где \oplus — сложение по модулю s);
- D)** В первом случае сервер работает по операции **(B)**, а во втором случае он простаивает (обслуживает буфера на нулевой скорости) до момента $(k + 1)T$, если переключение завершено раньше.

Покажем, что данный протокол корректен (сформулированные инструкции выполнимы) и стабилизирует систему, если параметр T выбран достаточно большим. Докажем следующую лемму.

Лемма 3. *Для всех достаточно больших $T > 0$ верны следующие утверждения:*

- а) *Длительности начального интервала $[0, T]$ достаточно для того, чтобы сервер завершил запрошенное в **r1**) переключение;*
- б) *При выполнении операции **r2B**) ни в одном из обслуживаемых буферов не создается ситуация, когда запрошенное извлечение содержимого на максимальной скорости невозможно выполнить ввиду опустошения буфера;*
- с) *на каждом из интервалов $[jT, (j+1)T]$, $j \geq 1$, сервер успевает завершить полный цикл требуемых переключений.*

Доказательство. Для выбора T заметим, что общее время, которое сервер должен потратить на переключения между пакетами в пределах одного цикла (2.7), не зависит от параметра T и равно

$$\Delta_{\text{sw}} := \sum_{k=0}^{s-1} \delta_{k \rightarrow k \oplus 1}. \quad (2.8)$$

В силу (1.2) для всех достаточно больших T выполнены неравенства

$$T \sum_{k=0}^{s-1} \tau_k + \Delta_{\text{sw}} \leq T, \quad T > \max_{k' \neq k''} \delta_{k' \rightarrow k''}. \quad (2.9)$$

Именно такие T далее и рассматриваем.

(a) Данное утверждение верно в силу второго неравенства из (2.9).

(c) В пределах одного цикла серверу требуется $T \sum_{k=0}^{s-1} \tau_k$ единиц времени для выполнения "активной" работы, предусмотренной п. **r2B**) и еще Δ_{sw} единиц времени на переключения пакетов. Поэтому c) верно в силу первого неравенства (2.9).

(b) Покажем вначале, что если для некоторого $j \geq 1$

$$x_b[jT] \geq \lambda_b T \quad \forall b, \quad (2.10)$$

то утверждение **b**) верно на интервале $[jT, (j+1)T]$. При работе в режиме пакета P_0 , серверу предстоит извлечь из каждого буфера $b \in P_0$ объем работы, равный $\mu_{b,0}\tau_{b,0}T$. Согласно (2.6) и (2.10), в начале этого извлечения содержимое буфера b не меньше $\mu_{b,0}\tau_{b,0}T$ и поэтому извлечение на максимальной скорости запрошенного объема осуществимо.

Полагая $\mu_{b,k} := 0$, если $b \notin P_k$, замечаем, что в момент окончания работы в режиме пакета P_0 содержимое каждого буфера b не меньше $\lambda_b - \mu_{b,0}\tau_{b,0}T \geq \mu_{b,1}\tau_{b,1}T$, где второе неравенство вытекает из (2.6). В момент начала работы в режиме пакета P_1 это неравенство очевидно сохраняется. Оно позволяет повторить проделанные рассуждения для пакета P_1 и убедиться, что в связанный с ним период работы утверждение **b**) верно, а в момент окончания работы в режиме пакета P_1 содержимое каждого буфера b не меньше $\lambda_b - \mu_{b,0}\tau_{b,0}T - \mu_{b,1}\tau_{b,1}T \geq \mu_{b,2}\tau_{b,2}T$, где последнее неравенство снова вытекает из (2.6). Продолжая рассуждать аналогично, убеждаемся в справедливости **b**) на интервале $[jT, (j+1)T]$.

Заметим также, что

$$(2.10) \Rightarrow x_b[(j+1)T] = x_b[jT] \quad \forall b. \quad (2.11)$$

Действительно, так как на интервале $[jT, (j+1)T]$ операция **r2B)** корректна, в режиме пакета P_k из буфера b извлекается объем работы, равный $\mu_{b,k}\tau_{b,k}T$. В целом на интервале $[jT, (j+1)T]$ в результате обхода по всем пакетам цикла P_0, \dots, P_{s-1} извлекается объем работы, равный $\sum_{k=0}^{s-1} \mu_{b,k}\tau_{b,k}T$. Остается заметить, что он равен объему работы, поступившей в буфер b на интервале $[jT, (j+1)T]$.

В заключение заметим, что (2.10) верно при $j = 1$ в силу **r1)**. Тогда применяя индукцию по $j = 1, 2, \dots$ и опираясь на установленные факты, убеждаемся, что утверждение **b)** верно на интервале $[jT, (j+1)T]$ при любом $j \geq 1$ и значит в целом на интервале времени $[T, \infty)$, где в соответствии с протоколом происходит обработка буферов. Кроме того заключение из (2.11) также верно при любом $j = 1, 2, \dots$ \square

Завершим доказательство последней леммой.

Лемма 4. *Протокол управления **r1)**, **r2)** корректен и стабилизирует систему.*

Доказательство. Корректность очевидна. В заключительной части доказательства предыдущей леммы было показано, что содержимое буферов остается постоянным в моменты jT ($j = 1, 2, \dots$). При этом содержимое буфера возрастает со скоростью не выше λ_b , и поэтому за время T возрастает не более, чем на $\lambda_b T$. При этом на любом промежутке $[jT, (j+1)T]$ при любом $j \geq 1$ объем пришедшей работы в

систему равен общему объему работы, изъятой из нее на том же промежутке, как следует из (2.11). А значит можно оценить изменение суммарного содержимого и по определению устойчивости завершить доказательство:

$$\bar{w} = w(0) + T \sum_{b=1}^n c_b \lambda_b \geq w(t) \quad \forall t.$$

□

2.2. Доказательство Теоремы 2

Перепишем систему (1.2), (1.3) в форме $Ax \leq a, Bx < b$:

$$x = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_s \end{bmatrix} \in \mathbb{R}^s, \quad Ax := \begin{bmatrix} -\sum_{k=1}^s \mu_{1,k} \tau_k \\ -\sum_{k=1}^s \mu_{2,k} \tau_k \\ \vdots \\ -\sum_{k=1}^s \mu_{n,k} \tau_k \\ -\tau_1 \\ \vdots \\ -\tau_s \end{bmatrix} \leq a := \begin{bmatrix} -\lambda_1 \\ -\lambda_2 \\ \vdots \\ -\lambda_n \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$Bx := \sum_{k=1}^s \tau_k < b := 1.$$

В этом случае $y = \mathbf{col}(\chi_1, \dots, \chi_n, \zeta_1, \dots, \zeta_s) \in \mathbb{R}^{n+s}$ и $z \in \mathbb{R}$. По следствию теоремы Моцкина [11] и [12], система (1.2), (1.3) неразрешима

в том и только том случае, если

$$\begin{aligned} \exists \chi_b \geq 0, \zeta_k \geq 0 : \sum_{b=1}^n \mu_{b,k} \chi_b + \zeta_k = 0 \quad \forall k = 1, \dots, s, \quad \sum_{b=1}^n \lambda_b \chi_b > 0 \quad \text{или} \\ \exists \chi_b \geq 0, \zeta_k \geq 0, z > 0 : - \sum_{b=1}^n \mu_{b,k} \chi_b - \zeta_k + z = 0, - \sum_{b=1}^n \lambda_b \chi_b + z \leq 0. \end{aligned}$$

Так как ζ_k однозначно определяются из первого уравнения после ":", эти условия можно переписать в следующем эквивалентном виде

$$\begin{aligned} \exists \chi_b \geq 0 : \sum_{b=1}^n \mu_{b,k} \chi_b \leq 0 \quad \forall k = 1, \dots, s, \quad \sum_{b=1}^n \lambda_b \chi_b > 0 \quad \text{или} \\ \exists \chi_b \geq 0, z > 0 : \sum_{b=1}^n \mu_{b,k} \chi_b \leq z, z \leq \sum_{b=1}^n \lambda_b \chi_b. \end{aligned}$$

Второе из этих условий равносильно требованию

$$\exists \chi_b \geq 0 : \sum_{b=1}^n \mu_{b,k} \chi_b \leq \sum_{b=1}^n \lambda_b \chi_b \quad \forall k = 1, \dots, s, \quad \sum_{b=1}^n \lambda_b \chi_b > 0. \quad (2.12)$$

Оно включает в себя первое условие и поэтому необходимо и достаточно для неразрешимости системы (1.2), (1.3). Умножая χ_b на общий положительный множитель, условию (2.12) можно придать следующий вид

$$\exists \chi_b \geq 0 : \sum_{b=1}^n \mu_{b,k} \chi_b \leq 1 \quad \forall k = 1, \dots, s, \quad \sum_{b=1}^n \lambda_b \chi_b = 1.$$

Остается заметить, что отрицание этого утверждения равносильно (1.4). \square

Глава 3

Критерии устойчивости в специальных случаях

3.1. Вспомогательная лемма

В этом разделе будут встречаться громоздкие формулы. Для упрощения введем обозначение *коэффициента утилизации буфера*

$$\gamma_{b,k} = \frac{\lambda_b}{\mu_{b,k}}. \quad (3.1)$$

Если же максимальная скорость не зависит от подключаемых пакетов, то индекс k будем опускать: $\gamma_b = \frac{\lambda_b}{\mu_b}$.

Используя теоремы из предыдущих разделов, выведем упрощенные критерии устойчивости для описанных ранее частных случаев. При этом будет полезен следующий технический факт.

Лемма 5. *Обозначим $\Theta = \left\{ \{\theta_i\}_{i=1}^m : \theta_i \geq 0, \sum_i \theta_i = 1 \right\}$. Предположим, что заданы некоторые константы $\omega_i > 0 \forall i$. Введем функцию $J(\Theta) = \max_k \omega_k^{-1} \theta_k$. Тогда*

$$\min_{\theta_i \in \Theta} J(\theta) = \frac{1}{\omega_1 + \dots + \omega_m},$$

при этом минимум достигается в точке с координатами

$$\theta_i = \frac{\omega_i}{\omega_1 + \dots + \omega_m}.$$

Доказательство. Перейдем к вспомогательной эквивалентной задаче, вычитая из минимизируемого функционала значение, заявленное в лемме как минимально возможное. Достаточно показать, что в модифицированной, таким образом, задаче минимальное значение функционала равно нулю и достигается в точке, указанной в формулировке леммы. Модифицированный функционал равен:

$$\min_{\theta} \max_k \left(\omega_k^{-1} \theta_k - \frac{1}{\sum_i \omega_i} \right) = \min_{\theta} \max_k \left(\frac{\theta_k \sum_i \omega_i - \omega_k}{\omega_k \sum_i \omega_i} \right).$$

Предположим, что внутреннее значение в некоторой точке $\theta \in \Theta$ для любого k отрицательно:

$$\frac{\theta_k \sum_i \omega_i - \omega_k}{\omega_k \sum_i \omega_i} < 0 \Rightarrow \theta_k \sum_i \omega_i - \omega_k < 0 \quad \forall k.$$

Просуммируем все такие выражения, сумма должна также оставаться отрицательна. Но по построению θ получим противоречие:

$$\sum_k \theta_k \sum_i \omega_i - \sum_i \omega_i < 0 \Rightarrow \sum_i \omega_i - \sum_i \omega_i < 0 \Rightarrow 0 < 0.$$

Значит все одновременно внутренности быть отрицательными не могут. Следовательно и минимум по θ не может быть отрицательным, поскольку из всех значений сначала выбирается наибольшее по k . Таким образом, наилучшая ситуация для минимума это 0. А этого можно достичь, приравняв все значения к нулю с помощью

ПОДСТАНОВКИ

$$\theta_i = \frac{\omega_i}{\omega_1 + \dots + \omega_m}.$$

□

3.2. Случай с непересекающимися пакетами

Теорема 3. *Предположим, что пакеты попарно не пересекаются и их число не меньше двух. Система поллинга стабилизируема в том и только в том случае, если*

$$\sum_{k=1}^s \max_{b \in P_k} \gamma_b < 1, \quad (3.2)$$

где γ_b — коэффициент утилизации (3.1).

Доказательство. Так как число пакетов ≥ 2 и они не пересекаются, нет исчерпывающе-достаточного пакета и следовательно к системе применима **Теорема 2**. Распишем сумму из определения стандартного симплекса в **Теореме 2**, где введем новые переменные $\theta = \{\theta_i\}_{i=1}^s$ как общую сумму каждого пакета P_i :

$$\sum_{k=1}^s \theta_k = 1, \quad \theta_k \geq 0, \quad \theta_k = \sum_{b \in P_k} \lambda_b \chi_b, \quad P_k \sim \theta_k, \quad \forall k = 1 : s.$$

Разобьем минимум по симплексу:

$$\min_{\chi \in \mathbb{S}_n} \sim \min_{\theta \in \Theta} \min_{\{\chi_b\} \in \Lambda_k(\theta)},$$

где Θ определено в **Лемме 5** (где сейчас $m := s$), а

$$\Lambda_k(\theta) = \left\{ \{\chi_b\}_{b \in P_k} : \chi_b \geq 0, \sum_{b \in P_k} \lambda_b \chi_b = \theta_k \right\}.$$

И используем это разбиение для преобразования выражения (1.4):

$$\begin{aligned} \min_{\chi \in \mathbb{S}_n} \max_k \sum_{b \in P_k} \mu_b \chi_b &= \min_{\theta \in \Theta} \min_{\{\chi_b\} \in \Lambda_k(\theta)} \max_k \sum_{b \in P_k} \mu_b \chi_b \stackrel{[1]}{=} \min_{\theta \in \Theta} \max_k \min_{\{\chi_b\} \in \Lambda_k(\theta)} \sum_{b \in P_k} \mu_b \chi_b \stackrel{[2]}{=} \\ &= \min_{\theta \in \Theta} \max_k \theta_k \min_{b \in P_k} \frac{\mu_b}{\lambda_b} \stackrel{[3]}{=} \min_{\theta \in \Theta} \max_k \omega_k^{-1} \theta_k > 1 \stackrel{[4]}{\iff} \\ \iff \max_k \omega_k^{-1} \theta_k > 1 \quad \forall \{\theta_k\} : \sum_k \theta_k = 1 &\stackrel{[5]}{\iff} \frac{1}{\omega_1 + \dots + \omega_s} > 1 \iff \\ \iff \sum_{k=1}^s \max_{b \in P_k} \frac{\lambda_b}{\mu_b} < 1. &\tag{3.3} \end{aligned}$$

- [1] Перестановка максимума и минимума возможна, поскольку нет общих элементов в разных пакетах;
- [2] Так как $\Lambda_k(\theta)$ — выпуклый многогранник, линейная функция достигает рассматриваемого минимума в одной из его вершин, и можно перейти к минимуму по вершинам. В свою очередь для вершины все координаты χ_b равны нулю кроме одной, которая находится из определения $\Lambda_k(\theta)$ и равна θ_k/λ_b ;
- [3] Введем обозначение $\omega_k = \max_{b \in P_k} \frac{\lambda_b}{\mu_b}$, тогда $\omega_k^{-1} = \min_{b \in P_k} \frac{\mu_b}{\lambda_b}$;
- [4] Перепишем основное условие из **Теоремы 2**, избавляясь от минимума;

[5] Применяем Лемму 5.

□

Далее будем рассматривать случаи, когда пакеты пересекаются. Сначала разберем ситуацию с двумя пакетами, в каждом из которых по два буфера, а один буфер является общим, и максимальная скорость его обслуживания (т.е. изъятия содержимого) не зависит от пакета. Затем рассмотрим более сложные и интересные случаи, используя в том числе приемы, отработанные на текущем простом случае.

3.3. Случай с пересечением двух пакетов (из двух буферов) в одном буфере с μ_0 , не зависящей от пакета

Теорема 4. Пусть имеем два пакета, в каждом из которых по два буфера. При этом у них имеется общий буфер с максимальной скоростью обслуживания, не зависящей от подключаемого пакета (на Рис. 3.1 схематически изображена интересующая нас система, где b_0 — общий буфер.). Система поллинга стабилизируема в том и только в том случае, если

$$\gamma_0 < 1, \quad \gamma_1 + \gamma_2 < 1. \quad (3.4)$$

Доказательство. В рассматриваемом случае условие (1.4) примет вид:

$$\max \{ \mu_1 \chi_1 + \mu_0 \chi_0; \mu_2 \chi_2 + \mu_0 \chi_0 \} > 1, \quad \forall \chi \in \mathbb{S}_3.$$

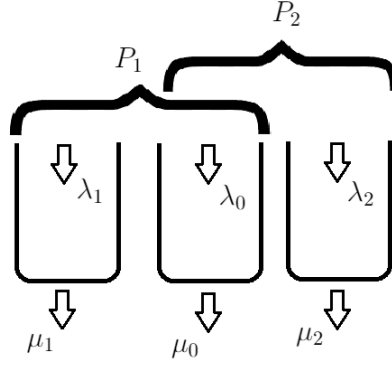


Рис. 3.1. Частный случай пересечения пакетов из 3-х буферов.

Вынесем слагаемое общего буфера за максимум и введем новые переменные $\xi_k = \lambda_k \chi_k$ для $k = 0, 1, 2$:

$$\frac{\mu_0}{\lambda_0} \xi_0 + \max \left\{ \frac{\mu_1}{\lambda_1} \xi_1; \frac{\mu_2}{\lambda_2} \xi_2 \right\} > 1, \quad \forall \xi_k \geq 0 : \xi_0 + \xi_1 + \xi_2 = 1.$$

При $\xi_0 = 1$ итоговый критерий перейдет в условие $\mu_0 > \lambda_0 > 0$. Теперь пусть $\xi_0 < 1$. Вводя новые переменные $\theta_k = \frac{\xi_k}{1 - \xi_0} \geq 0$, где $k = 1, 2$ и $\theta_1 + \theta_2 = 1$, перепишем условие стабилизируемости:

$$\begin{aligned} & \frac{\mu_0}{\lambda_0} \xi_0 + \max \left\{ \frac{\mu_1}{\lambda_1} (1 - \xi_0) \theta_1; \frac{\mu_2}{\lambda_2} (1 - \xi_0) \theta_2 \right\} = \\ & = \frac{\mu_0}{\lambda_0} \xi_0 + (1 - \xi_0) \max \left\{ \frac{\mu_1}{\lambda_1} \theta_1; \frac{\mu_2}{\lambda_2} \theta_2 \right\} > 1 \quad \forall \theta_k \geq 0 : \\ & \theta_1 + \theta_2 = 1, \quad \forall \xi_0 \in [0, 1) \xLeftrightarrow[1] \frac{\mu_0}{\lambda_0} \xi_0 + (1 - \xi_0) \frac{1}{\frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2}} > 1, \quad \forall \xi_0 \in [0, 1) \xLeftrightarrow[2] \\ & \iff \frac{\lambda_0}{\mu_0} < 1, \quad \frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2} < 1. \end{aligned} \quad (3.5)$$

[1] Обращаемся к Лемме 5;

[2] Получили линейную функцию относительно ξ_0 на промежутке $[0, 1)$. Для выполнения неравенства на всем интервале необходимо и достаточно, чтобы оно выполнялось на концах интервала.

□

3.4. Случай с пересечением двух пакетов в одном буфере с μ_0 , не зависящей от пакета

Теорема 5. Пусть имеется два пакета P_1 и P_2 с n_1 и n_2 буферами соответственно (Рис. 3.2). Пусть также они пересекаются только в одном буфере b_0 , а максимальная скорость обслуживания этого общего буфера не зависит от пакета. Система поллинга стабилизируема тогда и только тогда, когда

$$\gamma_0 < 1, \quad \max_{P'_1} \gamma_b + \max_{P'_2} \gamma_b < 1, \quad (3.6)$$

где $P'_i = P_i \setminus b_0$ для $i = 1, 2$.

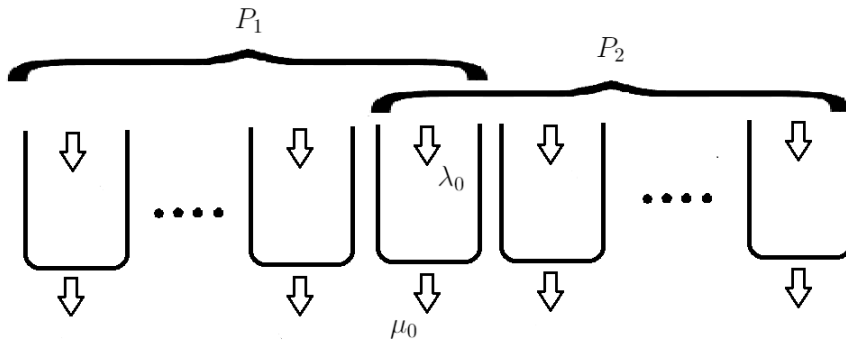


Рис. 3.2. Частный случай пересечения пакетов в одном буфере.

Доказательство. По **Теореме 2** система стабилизируема тогда и только тогда, когда

$$\max \left\{ \sum_{b \in P_1} \mu_b \chi_b, \sum_{b \in P_2} \mu_b \chi_b \right\} > 1 \quad \forall \chi \in \mathbb{S}_n.$$

Выделим слагаемое с общим буфером из суммы, а потом вынесем его за максимум:

$$\max \left\{ \mu_0 \chi_0 + \sum_{b \in P_1, b \neq b_0} \mu_b \chi_b, \mu_0 \chi_0 + \sum_{b \in P_2, b \neq b_0} \mu_b \chi_b \right\} > 1 \quad \forall \chi \in \mathbb{S}_n,$$

$$\mu_0 \chi_0 + \max \left\{ \sum_{b \in P_1, b \neq b_0} \mu_b \chi_b, \sum_{b \in P_2, b \neq b_0} \mu_b \chi_b \right\} > 1 \quad \forall \chi \in \mathbb{S}_n.$$

Обозначим $\theta_i = \sum_{P'_i} \lambda_b \chi_b$, $\theta_0 = \lambda_0 \chi_0$. Зафиксируем χ_0 . Если $\theta_0 = 1$, то критерий примет вид: $\mu_0 > \lambda_0$ - запомним этот вариант. Пусть $\Lambda_k(\theta) = \{ \{ \chi_b \}_{b \in P'_k} : \chi_b \geq 0, \sum_{b \in P'_k} \lambda_b \chi_b = \theta_k \}$, где $\theta := (\theta_0, \theta_1, \theta_2)$. Перепишем условие через минимум для остальных значений θ_0 и разобьем минимум как было в доказательстве предыдущей теоремы:

$$\min_{\theta_1 + \theta_2 = 1 - \theta_0} \min_{\Lambda_k(\theta)} \max \left\{ \sum_{P'_1} \mu_b \chi_b, \sum_{P'_2} \mu_b \chi_b \right\} > 1 - \theta_0 \frac{\mu_0}{\lambda_0} \quad \forall \theta_0 \in [0, 1).$$

Будем работать с левой частью неравенства. Переставим максимум и минимум (это возможно, так как P'_1 и P'_2 не пересекаются):

$$\begin{aligned}
&= \min_{\theta_1+\theta_2=1-\theta_0} \min_{\Lambda_k(\theta)} \max \left\{ \sum_{P'_1} \mu_b \chi_b, \sum_{P'_2} \mu_b \chi_b \right\} = \\
&= \min_{\theta_1+\theta_2=1-\theta_0} \max \left\{ \min_{\Lambda_1(\theta)} \sum_{P'_1} \mu_b \chi_b, \min_{\Lambda_2(\theta)} \sum_{P'_2} \mu_b \chi_b \right\} = \\
&= \min_{\theta_1+\theta_2=1-\theta_0} \max \left\{ \theta_1 \min_{\Lambda_1(\theta)} \frac{\mu_b}{\lambda_b}, \theta_2 \min_{\Lambda_2(\theta)} \frac{\mu_b}{\lambda_b} \right\}.
\end{aligned}$$

Обозначим $\min_{\Lambda_k} \frac{\mu_b}{\lambda_b} = \omega_k^{-1}$. Также переобозначим старые переменные по новому $\theta_k = \frac{\theta_k}{1-\theta_0}$. Тогда для нового выражения применим **Лемму 5**:

$$(1 - \theta_0) \min_{\theta_1+\theta_2=1} \max \{ \theta_1 \omega_1^{-1}, \theta_2 \omega_2^{-1} \} > 1 - \theta_0 \frac{\mu_0}{\lambda_0} \quad \forall \theta_0 \in [0, 1),$$

$$(1 - \theta_0) \frac{1}{\omega_1 + \omega_2} + \theta_0 \frac{\mu_0}{\lambda_0} > 1 \quad \forall \theta_0 \in [0, 1).$$

Получили ограничение на линейную функцию относительно θ_0 . Оно верно тогда и только тогда, когда

$$\frac{1}{\omega_1 + \omega_2} > 1, \quad \frac{\mu_0}{\lambda_0} > 1$$

или

$$\frac{\lambda_0}{\mu_0} < 1, \quad \max_{P'_1} \frac{\lambda_b}{\mu_b} + \max_{P'_2} \frac{\lambda_b}{\mu_b} < 1. \quad (3.7)$$

Случай с $\theta_0 = 1$ уже включен здесь. \square

3.5. Случай с пересечением двух пакетов (из двух буферов) в одном буфере с $\mu_{b,k}$, зависящей от пакета

Усложним второй случай.

Теорема 6. Пусть имеем два пакета, в каждом из которых по два буфера. При этом у них имеется общий буфер с максимальной скоростью обслуживания $\mu_{0,k}$, зависящей от подключаемого пакета (без ограничения общности можно считать, что $\mu_{0,1} > \mu_{0,2}$). Система поллинга стабилизируема в том и только в том случае, если

$$\gamma_{0,1} < 1, \quad \gamma_{0,2}(\gamma_{0,1} + \gamma_2 - 1) < \gamma_2\gamma_{0,1}, \quad \gamma_1 + \gamma_2 < 1. \quad (3.8)$$

Доказательство. Постараемся вновь упростить критерий для рассматриваемой ситуации:

$$\max\{\mu_{0,1}\chi_0 + \mu_1\chi_1, \mu_{0,2}\chi_0 + \mu_2\chi_2\} > 1 \quad \forall \chi \in \mathbb{S}_3.$$

Введем новые переменные: $\xi_k = \lambda_k\chi_k$ ($k = 0 : 2$). Зафиксируем ξ_0 . Если $\xi_0 = 1$, тогда получим:

$$\frac{\lambda_0}{\max\{\mu_{0,1}, \mu_{0,2}\}} < 1, \quad \gamma_{0,1} < 1.$$

Позднее вспомним про это ветвление. А сейчас перепишем в новых

переменных неравенство:

$$\max \{ \xi_0 \gamma_{0,1}^{-1} + \xi_1 \gamma_1^{-1}, \xi_0 \gamma_{0,2}^{-1} + \xi_2 \gamma_2^{-1} \} > 1$$

для $\xi_1 + \xi_2 = 1 - \xi_0$, $\xi_1, \xi_2 \geq 0$, $\xi_0 \in [0, 1)$.

Вводя новые переменные $z_0 := \xi_0$, $z_i = \frac{\xi_i}{1 - \xi_0}$ ($i = 1, 2$), где $z_i \geq 0$ и $z_1 + z_2 = 1$, преобразуем условие стабилизируемости:

$$\max \{ z_0 \gamma_{0,1}^{-1} + z_1 (1 - z_0) \gamma_1^{-1}; z_0 \gamma_{0,2}^{-1} + z_2 (1 - z_0) \gamma_2^{-1} \} > 1$$

для $\forall z_1, z_2 \geq 0$, $z_1 + z_2 = 1$, $\forall z_0 \in [0, 1]$.

Зафиксировав значения z_1 и z_2 , получаем в левой части неравенства максимум от двух линейных по z_0 функций на интервале $[0, 1]$. На этом интервале максимум как функция от z_0 принимает наименьшее значение либо в концах отрезка, либо в точке в точке пересечения $z_0^0 \in (0, 1)$ линейных функций. Рассмотрим все три варианта точки, где достигается минимальный максимум.

Пусть $z_0 = 0$ — левый конец, тогда критерий стабилизируемости переписывается:

$$\max \{ z_1 \gamma_1^{-1}; z_2 \gamma_2^{-1} \} > 1, \quad \forall z_1, z_2 \geq 0, z_1 + z_2 = 1.$$

Перепиывая (1.4) через минимум и применяя **Лемму 5**, получим условия стабилизируемости для этого случая:

$$\min_{z_1+z_2=1, z_i \geq 0} \max \{ z_1 \gamma_1^{-1}; z_2 \gamma_2^{-1} \} > 1,$$

$$\frac{1}{\gamma_1 + \gamma_2} > 1, \quad \gamma_1 + \gamma_2 < 1.$$

Перейдем теперь к правому концу $z_0 = 1$. Условие стабилизируемости примет вид:

$$\max \left\{ \frac{\mu_{0,1}}{\lambda_0}; \frac{\mu_{0,2}}{\lambda_0} \right\} > 1, \quad \gamma_{0,1} < 1.$$

Наконец, вариант с пересечением линейных функций на интервале $[0, 1]$. Вычислим сначала точку пересечения, приравняв две линейные функции, а затем вычислим само значение, которое должно быть больше 1:

$$\begin{aligned} z_0 \gamma_{0,1}^{-1} + z_1 (1 - z_0) \gamma_1^{-1} &= z_0 \gamma_{0,2}^{-1} + z_2 (1 - z_0) \gamma_2^{-1}, \\ z_0 (\gamma_{0,1}^{-1} - \gamma_{0,2}^{-1}) + (1 - z_0) (z_1 \gamma_1^{-1} - z_2 \gamma_2^{-1}) &= 0. \end{aligned}$$

Полученная линейная функция имеет корень на $(0, 1)$ при условии, что она принимает на концах значения противоположных знаков. Так как $\mu_{0,1} > \mu_{0,2}$, на правом конце имеем положительное значение. Соответственно на левом конце должно быть отрицательное значение, а значит

$$z_1 \gamma_1^{-1} - z_2 \gamma_2^{-1} < 0, \quad z_1 \gamma_2 < z_2 \gamma_1. \quad (3.9)$$

Сам корень равен

$$z_0^0 = \frac{\gamma_2^{-1} z_2 - \gamma_1^{-1} z_1}{\gamma_{0,1}^{-1} - \gamma_{0,2}^{-1} + \gamma_2^{-1} z_2 - \gamma_1^{-1} z_1}.$$

Теперь подставим эту точку в одну из функций в максимуме, вынося за скобки громоздкий знаменатель, обозначенный временно

за a (во всех неравенствах $z_1, z_2 \geq 0$, $z_1 + z_2 = 1$):

$$\begin{aligned} \frac{1}{a} \left((\gamma_2^{-1} z_2 - \gamma_1^{-1} z_1) \gamma_{0,1}^{-1} + (a - \gamma_2^{-1} z_2 + \gamma_1^{-1} z_1) \gamma_1^{-1} z_1 \right) &> 1, \\ (\gamma_2^{-1} z_2 - \gamma_1^{-1} z_1) \gamma_{0,1}^{-1} + (\gamma_{0,1}^{-1} - \gamma_{0,2}^{-1}) \gamma_1^{-1} z_1 &> a, \\ \gamma_2^{-1} \gamma_{0,1}^{-1} z_2 - \gamma_{0,2}^{-1} \gamma_1^{-1} z_1 &> a, \\ \gamma_2^{-1} \gamma_{0,1}^{-1} z_2 - \gamma_{0,2}^{-1} \gamma_1^{-1} z_1 &> \gamma_{0,1}^{-1} - \gamma_{0,2}^{-1} + \gamma_2^{-1} z_2 - \gamma_1^{-1} z_1, \\ z_1 \gamma_1^{-1} (1 - \gamma_{0,2}^{-1}) + z_2 \gamma_2^{-1} (\gamma_{0,1}^{-1} - 1) &> \gamma_{0,1}^{-1} - \gamma_{0,2}^{-1}. \end{aligned} \quad (3.10)$$

Полученное неравенство должно выполняться для любых $z = (z_1, z_2)$, удовлетворяющих $z_i \geq 0$ ($i=1, 2$), $z_1 + z_2 = 1$ и (3.9) $z_1 \gamma_2 < z_2 \gamma_1$. На двумерной плоскости множество этих точек представляет собой отрезок Q прямой линии с концами $(0, 1)$ и $(\bar{z}, 1 - \bar{z})$ — выколота точка, где \bar{z} — корень уравнения

$$\bar{z} \gamma_1^{-1} = (1 - \bar{z}) \gamma_2^{-1} \iff \bar{z} = \frac{\gamma_2^{-1}}{\gamma_2^{-1} + \gamma_1^{-1}} \in (0, 1).$$

В левой части неравенства (3.10) наблюдаем линейную функцию от z . Таким образом, для выполнения этого соотношения на всем отрезке Q необходимо и достаточно выполнение строгого неравенства в точке $(0, 1)$ и нестрогого в точке $(\bar{z}, 1 - \bar{z})$ (после упрощения):

$$\gamma_2^{-1} \gamma_{0,1}^{-1} - \gamma_2^{-1} > \gamma_{0,1}^{-1} - \gamma_{0,2}^{-1}, \quad \gamma_1 + \gamma_2 \leq 1.$$

При домножении первого неравенства на $\gamma_2 \gamma_{0,1} \gamma_{0,2}$ слева и справа получим требуемое второе неравенство в (3.8).

Объединяя результаты всех трех случаев расположения мини-

мального максимума от двух линейных функций, приходим к требуемому. \square

Заметим, что второе неравенство из (3.8) заведомо выполнено, если третье неравенство верно после замены в нем γ_1 на $\gamma_{0,1}$ и, более того, если левая часть полученного таким образом неравенства не превосходит единицы (то есть неравенство нестрогое). В противном случае второе неравенство из (3.8) принимает вид

$$\gamma_{0,2} < \frac{\gamma_2 \gamma_{0,1}}{\gamma_2 + \gamma_{0,1} - 1}$$

и означает, что коэффициент утилизации общего буфера во втором пакете должен быть достаточно мал, то есть максимальная скорость его обслуживания в этом пакете должна быть достаточно велика по сравнению со скоростью входящего потока.

3.6. Случай с пересечением двух пакетов в одном буфере с $\mu_{b,k}$, зависящей от пакета

Хотим объединить два предыдущих примера в один.

Теорема 7. *В системе поллинга два пакета P_1 и P_2 с n_1 и n_2 буферами соответственно. Также имеется общий буфер b_0 , максимальная скорость обработки которого зависит от пакета (без потери общности можно считать $\mu_{0,1} > \mu_{0,2}$). Такая система поллинга стабилизируема тогда и только тогда, когда*

$$\gamma_b(\gamma_{0,1} + \gamma_2 - 1) < \gamma_b \gamma_{0,1} \quad \forall b \in P_2', \quad (3.11)$$

$$\gamma_{0,1} < 1, \quad \max_{P'_1} \gamma_b + \max_{P'_2} \gamma_b < 1,$$

где $P'_i = P_i \setminus b_0$ для $i = 1, 2$.

Доказательство. Критерий стабилизируемости по **Теореме 2** принимает вид:

$$\max \left\{ \mu_{0,1} \chi_0 + \sum_{b \in P'_1} \mu_b \chi_b, \mu_{0,2} \chi_0 + \sum_{b \in P'_2} \mu_b \chi_b \right\} > 1 \quad \forall \chi \in \mathbb{S}_n.$$

Аналогично предыдущему случаю проводим замену переменных $z_0 = \lambda_0 \chi_0$, $z_b = \frac{\lambda_b \chi_b}{1 - z_0}$ и переходим к рассмотрению максимума относительно двух линейных по z_0 функций:

$$\max \left\{ z_0 \gamma_{0,1}^{-1} + (1 - z_0) \sum_{b \in P'_1} z_b \gamma_b^{-1}, z_0 \gamma_{0,2}^{-1} + (1 - z_0) \sum_{b \in P'_2} z_b \gamma_b^{-1} \right\} > 1,$$

где $\forall b \ z_b \geq 0$, $\sum_{b \in P'_1 \cup P'_2} z_b = 1$, $z_0 \in [0, 1]$.

Снова имеем три случая для точки максимума: левый и правый конец, внутренняя точка на интервале $[0, 1]$. Для левого конца ($z_0 = 0$) по аналогичным выкладкам в случае 3 получаем:

$$\max_{P'_1} \gamma_b + \max_{P'_2} \gamma_b < 1.$$

Для правого конца также все выведено ранее в случае 4:

$$\gamma_{0,1} < 1.$$

Перейдем к внутренней точке интервала $[0, 1]$.

$$z_0\gamma_{0,1}^{-1} + (1 - z_0) \sum_{b \in P'_1} z_b \gamma_b^{-1} = z_0\gamma_{0,2}^{-1} + (1 - z_0) \sum_{b \in P'_2} z_b \gamma_b^{-1},$$

$$z_0(\gamma_{0,1}^{-1} - \gamma_{0,2}^{-1}) + (1 - z_0) \left(\sum_{b \in P'_1} z_b \gamma_b^{-1} - \sum_{b \in P'_2} z_b \gamma_b^{-1} \right) = 0.$$

Чтоб корень последней функции был в промежутке $[0, 1]$, нужно

$$\sum_{b \in P'_1} z_b \gamma_b^{-1} - \sum_{b \in P'_2} z_b \gamma_b^{-1} < 0, \quad \sum_{b \in P'_1} z_b \gamma_b^{-1} < \sum_{b \in P'_2} z_b \gamma_b^{-1}. \quad (3.12)$$

Сам корень равен:

$$z_0 = \frac{\sum_{b \in P'_2} z_b \gamma_b^{-1} - \sum_{b \in P'_1} z_b \gamma_b^{-1}}{\gamma_{0,1}^{-1} - \gamma_{0,2}^{-1} + \sum_{b \in P'_2} z_b \gamma_b^{-1} - \sum_{b \in P'_1} z_b \gamma_b^{-1}}.$$

После подстановки в первую линейную функцию в максимуме и упрощений получаем условие для внутренней точки:

$$\gamma_{0,1}^{-1} \sum_{b \in P'_2} z_b \gamma_b^{-1} - \gamma_{0,2}^{-1} \sum_{b \in P'_1} z_b \gamma_b^{-1} > \sum_{b \in P'_2} z_b \gamma_b^{-1} - \sum_{b \in P'_1} z_b \gamma_b^{-1} + \gamma_{0,1}^{-1} - \gamma_{0,2}^{-1},$$

$$\sum_{b \in P'_1} z_b \gamma_b^{-1} (1 - \gamma_{0,2}^{-1}) + \sum_{b \in P'_2} z_b \gamma_b^{-1} (\gamma_{0,1}^{-1} - 1) > \gamma_{0,1}^{-1} - \gamma_{0,2}^{-1}. \quad (3.13)$$

Полученное неравенство должно выполняться при $\sum_{P'_1 \cup P'_2} z_b = 1$, $z_b \geq 0$ и $\sum_{b \in P'_1} z_b \gamma_b^{-1} < \sum_{b \in P'_2} z_b \gamma_b^{-1}$. Данная область представляет собой n -мерный многогранник в пространстве \mathbb{R}^n , и получается он отсе-

чением некоторой части от симплекса по (3.12). Если все вершины этого многогранника удовлетворяют неравенству (3.13), то и весь многогранник удовлетворяет этому соотношению. Этому и хотим достичь, наложив необходимые условия на вершины.

Все ребра симплекса лежат в двумерных плоскостях, зависящих только от двух переменных; остальные же переменные равны нулю на этих ребрах. Неравенство (3.12) задает полупространство, отделенное некоторой гиперплоскостью. На самом деле, имеем систему линейных уравнений и неравенств, решением которой является выпуклая оболочка базисных планов, а значит необходимый многогранник будет составлен из базовых решений [13]: из точек пересечения гиперплоскости с ребрами симплекса и вершин самого симплекса, удовлетворяющих (3.12).

Встает вопрос о том, где искать эти точки пересечения. Следует понять какие пары переменных (с соответствующим ребром) нужно выбирать. Если брать ребро с вершинами, соответствующим одному пакету, то при подстановке этого ребра в уравнение гиперплоскости будем получать равенство выпуклой оболочки двух переменных нулю. Таких точек (z_1, z_2) на симплексе не существует, а следовательно можно не рассматривать вершины, связанные с одним и тем же пакетом. Целиком ребро лежать в гиперплоскости также не может, потому что ни одна вершина не может принадлежать этой гиперплоскости. Поскольку точка пересечения будет делить ребро на две части, удовлетворяющую (3.12) и нет, то стоит перебрать все пары (z_a, z_b) , где для вершин $z_a = 1 \forall a \in P'_1$ соотношение не выполняется, так как $\sum_{P'_1 \cup P'_2} z_b = 1$, а для вершин $z_b = 1 \forall b \in P'_2$ наоборот верно. Вычислим все эти точки пересечения.

Рассмотрим некоторое ребро симплекса. Все переменные равны 0 кроме двух $z_a + z_b = 1$. Проведем замену переменных $z_a = 1 - z_b$ в (3.12) и рассчитаем точку пересечения ребра симплекса с гиперплоскостью, введенной ранее:

$$(1 - z_b)\gamma_a^{-1} = z_b\gamma_b^{-1},$$

$$z_b = \frac{\gamma_a^{-1}}{\gamma_a^{-1} + \gamma_b^{-1}}, \quad z_a = 1 - \frac{\gamma_a^{-1}}{\gamma_a^{-1} + \gamma_b^{-1}} = \frac{\gamma_b^{-1}}{\gamma_a^{-1} + \gamma_b^{-1}}.$$

Для этой точки должно выполняться неравенство (3.13):

$$\frac{\gamma_b^{-1}}{\gamma_a^{-1} + \gamma_b^{-1}}\gamma_a^{-1}(1 - \gamma_{0,2}^{-1}) + \frac{\gamma_a^{-1}}{\gamma_a^{-1} + \gamma_b^{-1}}\gamma_b^{-1}(\gamma_{0,1}^{-1} - 1) \geq \gamma_{0,1}^{-1} - \gamma_{0,2}^{-1},$$

$$\frac{\gamma_b^{-1}\gamma_a^{-1}}{\gamma_a^{-1} + \gamma_b^{-1}} \geq 1, \quad \gamma_a + \gamma_b \leq 1.$$

Это неравенство уже включено в ранее выведенное соотношение для левого конца. Вместе с условием на вершины $z_b = 1$ $b \in P'_2$ (из него выводится первое неравенство (3.11) также как и в случае 4)

$$\gamma_b^{-1}(\gamma_{0,1}^{-1} - 1) > \gamma_{0,1}^{-1} - \gamma_{0,2}^{-1},$$

справедливо желаемое утверждение про n -мерный многогранник (его удовлетворение условию (3.13)). А значит теорема доказана. \square

Глава 4

Проверка критерия стабилизируемости с помощью Matlab

4.1. Описание программы

В рамках выпускной квалификационной работы была разработана программа на Matlab, позволяющая из начальных данных $(P_k, \mu_{b,k}$ и $\lambda_b \forall b, k)$ установить или опровергнуть стабилизируемость системы по **Теореме 1**. В первом случае существует решение τ_k системы (1.2), (1.3). Тогда оно находится и подставляется в уже рассмотренный стабилизирующий протокол для проверки ограниченности суммарной очереди и оценки величины очереди. Во втором случае та же система неравенств неразрешима. В данной главе приводим результаты численного исследования иллюстративных примеров с помощью разработанного программного обеспечения. Оно использует несколько программных модулей, разработанных автором.

Функция Solver.m либо устанавливает неразрешимость системы неравенств (1.2), (1.3), возвращая пустой массив, либо находит ее решение τ_k . Вычисления происходят с помощью пакетной функции Матлаба linprog(), которая решает задачу линейного программирования с такими ограничениями как линейные уравнения, линейные нестрогие неравенства и двойные нестрогие неравенства. За целевую функцию берется сумма τ_k . Остальные аргументы переписываются под требования linprog().

В функции Work_time.m для вычисления $\tau_{b,k}$ (решение системы

(2.6)) используется также функция `linprog()`. Подходящее решение находится для каждого буфера отдельно с учетом всех пакетов. Результатом `Work_time.m` является матрица всех времен $\tau_{b,k}$, требуемых серверу для обслуживания буфера b в пакете P_k .

Расчет освобожденного содержимого из всех буферов на промежутке $[0, t]$ происходит в функции `Free_work.m`. Сначала рассчитывается общий объем работы, вышедший из буферов за целое число полных периодов. Далее оставшееся время, меньшее периода, тратится на последовательный обход пакетов. Когда оставшегося времени не хватает уже для полного обслуживания пакета, то сервер работает с ним только оставшуюся часть времени.

`PollingSystem.m` — центральный файл программы. Здесь задаются начальные данные: пакеты и буферы, скорости входных потоков в буферы, максимальные скорости обслуживания буферов, начальное содержимое в каждом буфере и время остановки процесса. Вся информация по буферам, пакетам и максимальной скорости обслуживания задается матрицей $\mu_{b,k}$ размерности $m \times n$, где m — число буферов в системе поллинга, а n — число пакетов. Каждая строка отвечает за один конкретный буфер, а каждый столбец за определенный пакет. Если буфер b не принадлежит пакету P_k , то соответствующий элемент в матрице задается нулем $\mu_{b,k} = 0$. Иначе записывается значение максимально возможной скорости обработки буфера b относительно пакета P_k . Вектор скорости притока работы λ_b и вектор начального содержимого x_{beg} имеют размерность m . На основе выхода функции `Solver.m` принимается вердикт о стабилизируемости построенной системы. Не нашлось подходящих τ_k — устойчивости нет, программа завершается. Если же нашлось решение, то вычисляется изменение суммарной работы на $[0, t]$, вычитая пришедшую и

освобожденную с помощью описанных ранее функций, а результат выводится на графике — наблюдается устойчивый процесс. Стабилизирующий протокол проходит по пакетам от 1 до n , рассматривая поочередно каждый столбец матрицы $\mu_{b,k}$ слева направо.

Также следует уточнить, что в данной модели для простоты пренебрегаем временем на переключение между пакетами. При этом на стабилизируемость системы эти значения не влияют, если период обхода по всем пакетам достаточно большой по сравнению со временем переключения. Если суммарное переключение за один период сможет уложиться в этот временной промежуток, то устойчивость процесса сохранится.

4.2. Исследование стабилизируемости на примерах

Пример 1.

Рассмотрим систему из 4 буферов и 4 пакетов с начальными данными:

$$\mu_{b,k} = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 0 & 10 & 0 & 2 \\ 2 & 0 & 0,1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \lambda_b = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.3 \\ 0.4 \end{pmatrix}, \quad x_{beg} = \begin{pmatrix} 2 \\ 1 \\ 3 \\ 3 \end{pmatrix}. \quad (4.1)$$

Программа нашла решение для системы (1.2), (1.3) ($\tau_k = (0.0485, 0.0138, 0.2189, 0.1811)$), а значит система поллинга стабилизируема. Можно исследовать поведение устойчивого процесса. Посмотрим на

работу системы с периодом обхода буферов $\text{period}=5$ (сек.) на промежутке $[0, 10]$ (Рис. 4.1).

Как можно заметить по кусочной линейности функции за один период, сервер использует все 4 пакета для работы с буферами: на каждом линейном участке сервер обслуживает необходимый пакет. В момент времени $t \approx 2.3$ (сек.) сервер отключается от пакетов и переходит в режим ожидания. На оставшемся участке периода наблюдается только линейный приток работы в буферы.

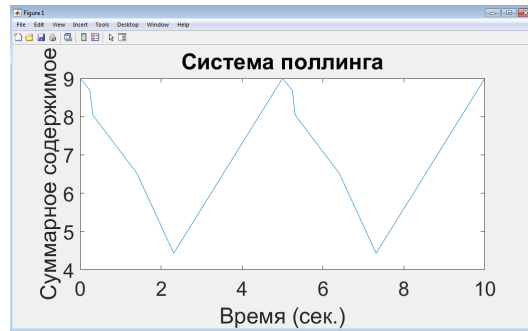


Рис. 4.1. Изменение суммарного содержимого с начальными данными (4.1) на $[0, 10]$.

Теперь рассмотрим поведение графика на промежутке $[0, 1000]$ (Рис. 4.2). Он действительно ограничен, не происходит избытка работы.

Изменим параметр периода и сравним два процесса системы с разными периодами (Рис. 4.3). Пусть $\text{period}=1$ (сек.) для второго протокола. Заметим, что свойство ограниченности сохранилось. А значит выбор периода не влияет на устойчивость процесса.

Пример 2.

Обратимся к системе из предыдущего примера, но скорости входных потоков увеличим в 10 раз: $\lambda_b = (5, 5, 3, 4)$. В этом случае про-

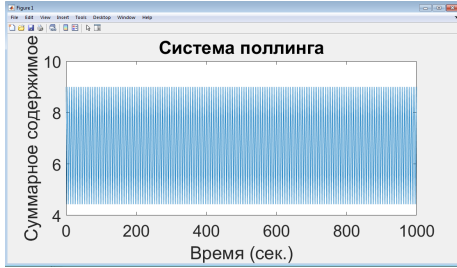


Рис. 4.2. Изменение суммарного содержимого с начальными данными (4.1) на $[0, 1000]$.

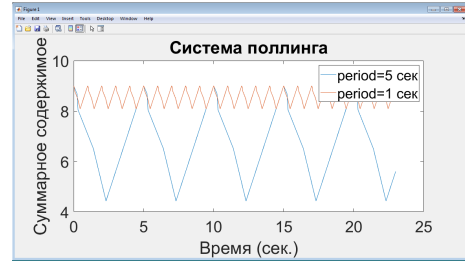


Рис. 4.3. Изменение суммарного содержимого с начальными данными (4.1) на $[0, 23]$ с разными периодами.

грамма выдала сообщение: "Система неравенств решений не имеет. Система поллинга не стабилизируема." Мощностей обработки буферов не хватает для поддержания ограниченности суммарной работы — приток содержимого слишком велик.

Пример 3.

Теперь рассмотрим систему с большим числом пакетов и буферов:

$$\mu_{b,k} = \begin{pmatrix} 0 & 1 & 2 & 0 & 0 & 2 & 0 & 2 \\ 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0,1 & 0 & 0 & 3 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 2 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 5 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 1 & 1 \\ 9 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 9 & 0 & 0 \end{pmatrix}, \lambda_b = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.3 \\ 0.4 \\ 0.7 \\ 0.5 \\ 0.6 \\ 0.7 \\ 0.1 \\ 0.3 \end{pmatrix}, x_{beg} = \begin{pmatrix} 2 \\ 1 \\ 3 \\ 3 \\ 5 \\ 4 \\ 7 \\ 8 \\ 2 \\ 4 \end{pmatrix}. \quad (4.2)$$

Построенная система имеет решение для (1.2), (1.3) ($\tau_k = (0.0111, 0.0500, 0.1375, 0.1500, 0, 0.0906, 0, 0.1500)$), а значит система поллинга стабилизируема. Рассмотрим некоторый процесс. Для начала обратимся к двум графикам: пришедшей работы и освобожденной (Рис. 4.4). Уже здесь заметна близость графиков. Обратившись же к самому процессу, действительно убеждаемся в устойчивости процесса (Рис. 4.5).

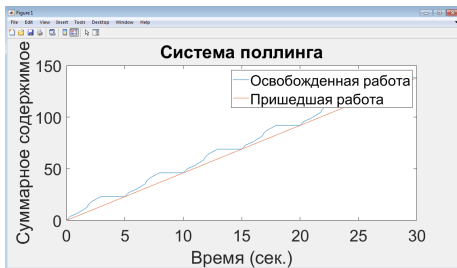


Рис. 4.4. Приход и отток работы из буферов с начальными данными (4.2) на $[0, 30]$.

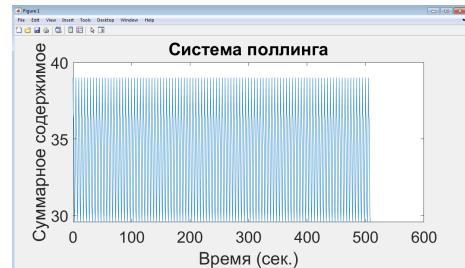


Рис. 4.5. Изменение суммарного содержимого с начальными данными (4.2) на $[0, 508]$.

Заметим, что суммарная очередь никогда не превосходит начального суммарного содержимого системы. Такое свойство наблюдается и в **Примере 1**.

Заключение

В дипломной работе была исследована детерминированная (жидкостная) модель системы поллинга с пакетной обработкой буферов (входящих потоков). Применительно к общему случаю такой системы получены два (необходимых и достаточных) критерия стабилизируемости. В случае стабилизируемости построен стабилизирующий протокол управления. Этот протокол децентрализован и использует только локальные данные, то есть информацию о текущем состоянии буферов обслуживаемого пакета.

Были выведены упрощенные версии критерия устойчивости для частных случаев: без пересечения пакетов и с пересечением пакетов, где максимальная скорость могла зависеть от подключаемого пакета или быть одинаковой во всех пакетах.

Для проверки критерия стабилизируемости была написана программа. Работа этой программы была продемонстрирована на трех примерах — в одном случае система оказалась не стабилизируемой. Для двух других примеров была смоделирована работа стабилизирующего протокола. На соответствующих графиках наблюдалась ограниченность суммарного содержимого — оно не превосходило начального суммарного содержимого системы.

Список литературы

1. Вишнеvский В. М., Семенова О. В. Математические методы исследования систем поллинга // Автоматика и Телемаханика. — 2006. — №2. — С. 3–56.
2. Baker K. R. Introduction to Sequencing and Scheduling. — New York: John Wiley & Sons, 1974. — 305 p.
3. Klimov. G. P. Time-sharing service systems // Theory of Probability & Its Applications. — 1975. — Vol. 19, no. 3 — P. 532–551.
4. Pinedo M. L. Scheduling Theory: Algorithms and Systems. — 3rd ed. — New York: Prentice Hall, 2008. — 662 p.
5. Вишнеvский В. М., Семенова О. В. Системы поллинга: Теория и применения в широкополосных беспроводных сетях. — М.: Техносфера, 2007. — 312 с.
6. Вишнеvский В. М., Семенова О. В. Обзор моделей систем поллинга и их применение в телекоммуникационных сетях // Проблемы Информатики. — 2020 — № 3. — С.29–59.
7. Matveev A., Feoktistova V., Bolshakova K. On global optimality of special periodic protocols for fluid polling systems with setups // Journal of Optimization Theory and Applications. — 2015. — Vol. 171 — P. 1055–1070
8. Matveev A. S. On stabilizability of fluid multi-server polling systems with setups // Cybernetics and physics. — 2012. — Vol. 2, no. 1. — P. 26–34.
9. Решение инженерных задач в SCILAB: Учебное пособие / А. Б. Андриевский и др. — СПб: Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, 2013. — 97 с.

10. Федюков А.А. Применение средств пакета МАТЛАВ для численного решения задач стабилизации по выходу динамических систем с фазовыми ограничениями: Учебно-методическое пособие — Нижний Новгород: Нижегородский госуниверситет, 2014 — 37 с.
11. Schrijver A. Theory of linear and integer programming. — New York: John Wiley & Sons, 1999. — 484 p.
12. Зоркальцев В.И., Киселева М.А. Системы линейных неравенств: Учебное пособие — Иркутск: ИГУ, 2007. — 129 с.
13. Банди Б. Основы линейного программирования — Москва: Радио и связь, 1989 — 145 с.
14. Designs of optimal switching feedback decentralized control policies for fluid queueing networks / V. Feoktistova et al. // Mathematics of Control, Signals, and Systems. — 2012. — no. 24. — P. 477–503.
15. Eekelen J. A. W. M. Modelling and control of discrete event manufacturing flow lines. — Eindhoven: Technische Universiteit Eindhoven, 2008. — 266 p.