



Math-Net.Ru

Общероссийский математический портал

Я. М. Агаларов, Оптимизация объема буферной памяти узла коммутации при схеме полного разделения памяти, *Информ. и её примен.*, 2018, том 12, выпуск 4, 25–32

DOI: 10.14357/19922264180404

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.15.149.243

13 сентября 2024 г., 16:17:00



ОПТИМИЗАЦИЯ ОБЪЕМА БУФЕРНОЙ ПАМЯТИ УЗЛА КОММУТАЦИИ ПРИ СХЕМЕ ПОЛНОГО РАЗДЕЛЕНИЯ ПАМЯТИ*

Я. М. Агаларов¹

Аннотация: Рассматривается задача оптимизации объема буферной памяти (БП) узла коммутации (УК) сети с коммутацией пакетов (КП), в котором используется схема полного разделения памяти между выходными линиями передачи, с учетом потерь пакетов из-за переполнения памяти и средних задержек пакетов. Для модели УК с пуассоновскими входящими потоками, общими распределениями времени обслуживания и одноканальными выходными линиями передачи получены свойства решения задачи и разработан алгоритм поиска оптимального объема. В рамках указанной модели разработан алгоритм поиска оптимального плана распределения БП при заданном ее объеме. Приведены результаты вычислительных экспериментов с использованием разработанных алгоритмов. Показана применимость полученных в работе подходов и алгоритмов к решению рассматриваемой задачи в рамках модели УК с рекуррентными входящими потоками, экспоненциальными временами обслуживания и одноканальными линиями передачи.

Ключевые слова: узел коммутации; управление потоками; распределение буферной памяти; оптимальный объем буферной памяти

DOI: 10.14357/19922264180404

1 Введение

Технико-экономические характеристики сети КП в значительной степени зависят от управления ресурсами и потоками требований в узлах-источниках и УК сети. Управление потоками предназначено для эффективной загрузки основных ресурсов (БП и каналов связи) в условиях их ограниченного объема с целью обеспечения требований к качеству обслуживания (Quality of Service — QoS) в сети. Зависимость качества обслуживания от объема памяти УК возникает из того, что при ограниченной пропускной способности каналов связи и небольшом объеме памяти могут быть недопустимо большие потери пакетов, что вызывает рост числа повторных передач пакетов и, как следствие, резкий рост нагрузки на отдельных участках сети (возможно, и в сети в целом), а при большом объеме памяти возможны недопустимо большие задержки пакетов в УК. Отсюда и возникает задача выбора оптимального объема БП УК с точки зрения выполнения определенных требований к качеству обслуживания пакетов и затратам, связанным со стоимостью оборудования и технической эксплуатации УК [1, 2]. При этом следует учитывать зависимость качества обслуживания в узле коммутации (и во всей сети в целом) и от схем (планов) распределения БП, из которых наиболее часто применяемыми на прак-

тике являются статические схемы управления [3, 4]: CS (Complete Sharing); CP (Complete Partitioning); SMQ (Sharing with Maximum Queue Length); SMA (Sharing with Minimum Allocation); SMQMA (Sharing with Maximum Queue Length and Minimum Allocation). Зависимость качества обслуживания в УК от схем распределения БП вызвана тем, что при бесконтрольном допуске пакетов в УК происходит захват буферов пакетами более интенсивных потоков, что приводит к неэффективному использованию выходных линий.

С момента зарождения сетей разработано великое множество алгоритмов локального управления потоками, многие из которых успешно применяются и в современных сетях. Основные результаты по управлению потоками в УК получены методом математического моделирования с использованием различных типов систем массового обслуживания (СМО). Некоторые из этих результатов описаны в работах [4–6], которые посвящены сравнительному анализу различных схем распределения БП, и в работах [7–13], в которых получены явные или алгоритмические решения задач по оптимизации параметров управления потоками в УК (объема БП, схем распределения БП УК).

В работе [5] приводится подробное описание методов анализа и расчета некоторых наиболее известных схем распределения БП и анализ известных

* Работа выполнена при частичной финансовой поддержке РФФИ (проекты 18-07-00692, 17-07-000717 и 17-07-00851).

¹ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, agglar@yandex.ru

результатов оптимизации схем СР и SMQ в предположении пуассоновских входящих потоков, экспоненциального времени обслуживания и одноканальных линий передачи.

В работе [7] при этом же предположении рассмотрено динамическое управление распределением БП и получен для случая трех каналов вид допустимого пространства состояний, соответствующий оптимальному решению. В работе [8] рассмотрена двухступенчатая схема управления потоками в УК и приведен алгоритм оптимизации интенсивностей потоков пакетов, допускаемых на первой ступени в УК, при схеме SMQ на второй ступени и в предположении пуассоновских входных потоков, экспоненциальных времен обслуживания и многоканальных выходных линий передачи.

В работе [9] предложен эвристический алгоритм решения задачи совместного управления объемом БП и пропускной способностью линий передачи, который рассматривается в следующих аспектах:

- (1) при заданной степени загрузки каналов и заданной пропускной способности найти необходимый объем БП УК, обеспечивающий заданную среднюю задержку в сети и требуемую вероятность потери ячеек вследствие переполнения буферов;
- (2) при заданной степени загрузки каналов и известных объемах БП найти необходимую пропускную способность, обеспечивающую заданную среднюю задержку в сети и требуемую вероятность потери ячеек вследствие переполнения буферов.

В работе [10] авторы решили задачу выбора объема БП для СМО типа $M/M/1/r$, сформулировав ее как задачу нелинейного программирования:

- (1) для фиксированной входной нагрузки найти объем накопителя (БП), при котором средняя задержка достигает минимума, а интенсивность потерь не превышает заданной величины;
- (2) для фиксированного объема накопителя (БП) найти значение входной нагрузки, при которой средняя задержка достигает минимума, а интенсивность потерь не превышает заданной величины.

Для СМО типа $G/M/1$ и $M/G/1$ с интегральной целевой функцией, учитывающей среднюю задержку в БП и вероятности потерь пакетов из-за переполнения БП, оптимизационная задача ограничения длины очереди решена в работах [11–13].

Ниже рассматривается экономическая модель УК с L выходными линиями передачи, представленная в виде L СМО типа $M/G/1/N$, в кото-

рой качество обслуживания регулируется методом штрафных функций. В рамках этой модели решается задача выбора оптимального объема БП УК и оптимальной схемы распределения типа СР.

2 Модель узла коммутации и постановка задачи

В качестве модели УК рассматривается многопотоковая СМО типа $M_L/G/L$ с накопителем емкости K и L приборами, на которую поступает $L \leq K$ пуассоновских потоков заявок $\lambda_j > 0$, $j = 1, \dots, L$, к каждому из которых для обслуживания заявок прикреплен отдельный прибор, время обслуживания на котором является независимой случайной величиной с заданной функцией распределения $G_j(t)$, $j = 1, \dots, L$. Накопитель СМО моделирует БП, приборы — линии передачи. Заявки каждого потока обслуживаются в порядке поступления, места в накопителе распределяются по схеме СР, т. е. за каждой j -й линией закрепляется $k_j > 0$ мест накопителя, где $\sum_{i=1}^L k_i = K$. В дальнейшем вектор $\bar{k} = (k_1, \dots, k_L)$, удовлетворяющий указанному условию, будем называть планом распределения БП. Поступившая заявка j -го потока занимает любое свободное место в накопителе, если в момент ее поступления число занятых мест в накопителе меньше k_j . Если заявка поступила в накопитель, она занимает любое свободное место в накопителе и обслуживается на приборе в порядке поступления. Заявка покидает систему только при завершении обслуживания, освободив одновременно прибор и накопитель, а на освободившийся прибор поступает очередная заявка соответствующего типа (если таковая есть в накопителе). Качество работы системы оценивается с помощью интегральной целевой функции (функцией дохода системы) со стоимостными параметрами (весовыми коэффициентами), зависящей от среднего времени ожидания и вероятности потерь заявок. Функция дохода имеет следующие стоимостные параметры:

$C_{0j} > 0$ — плата, получаемая системой за обслуживание единицы длины заявки j -го потока;

$C_{1j} > 0$ — величина штрафа, который платит система, если отклонена (потеряна) заявка j -го потока;

$C_{2j} > 0$ — величина штрафа системы за единицу времени задержки заявки j -го потока;

$C_{3j} > 0$ — величина штрафа системы за единицу времени простоя j -й линии;

$C_{4j} > 0$ — затраты системы в единицу времени на техническое обслуживание j -й линии.

Считается, что плату за обслуживание система получает в момент завершения обслуживания каждой заявки в зависимости от длины заявки (длительности занятия прибора).

Как следует из определения схемы СР, для каждого фиксированного значения K и фиксированного плана распределения БП рассматриваемая СМО представляет собой совокупность L независимых СМО типа $M/G_j/1/k_j$ (в работе используются обозначения СМО по классификации Кендалла). Это делает возможным использование в дальнейшем результатов работ [11–13], полученных для систем $M/G/1$ и $G/M/1$ с ограничением на длину очереди.

Введем обозначения:

$\pi_{ji}^{k_j}$, $0 \leq i \leq k_j = 1$, — стационарная вероятность того, что в системе находится i заявок потока j в момент завершения обслуживания заявки j -м прибором;

$g_j^{k_j}$ — значение предельного дохода подсистемы $M/G_j/1/k_j$, усредненного по числу обслуженных j -й линией заявок, $j = 1, \dots, L$;

$g^{\bar{k}}$ — значение суммарного предельного дохода всей системы, усредненного по числу обслуженных заявок при плане $\bar{k} = (k_1, \dots, k_L)$;

$Q_j^{k_j}$ — предельный средний доход, получаемый системой $M/G_j/1/k_j$ в единицу времени;

$Q^{\bar{k}}$ — предельный суммарный средний доход, получаемый всей системой в единицу времени при заданном плане \bar{k} ;

$\bar{v}_j = \int_0^\infty t dG_j(t)$ — среднее время обслуживания заявки потока j , $0 < \bar{v}_j < \infty$, $j = 1, \dots, L$;

$\rho_j = \lambda_j \bar{v}_j$ — нагрузка на линию j , $j = 1, \dots, L$.

Пусть $A_K = \{\bar{k} > 0 : \sum_{i=1}^L k_i = K\}$, $U_{CP} = \{A_K, K \geq L\}$, $U_{CP,N} = \{\bar{k} : L \leq \sum_{j=1}^L k_j \leq N\}$,

$Q^k = \max_{\bar{k}: \bar{k} \in A_K} Q^{\bar{k}}$.

Ставятся следующие две задачи:

1. Для множества планов U_{CP} найти значение $K^* \geq L$ такое, что

$$\max_{K \geq L} Q^K = Q^{K^*}. \quad (1)$$

2. Найти $\bar{k}^0 \in U_{CP,N}$ такой, что

$$\max_{\bar{k} \in U_{CP,N}} k^{\bar{k}} = g^{\bar{k}^0}. \quad (2)$$

Ниже всюду через K^* будем обозначать решение задачи (1), через $\bar{k}^0 \in U_{CP,N}$ — решение задачи (2).

3 Решение задачи (1)

Так как при схеме СР рассматриваемая модель УК есть совокупность независимых СМО типа $M/G_j/1/k_j$, $j = 1, \dots, L$, то ее доход при заданном $\bar{k} \in U_{CP,K}$ представляется в виде суммы:

$$Q^{\bar{k}} = \sum_{j=1}^L Q_j^{k_j}.$$

Тогда, очевидно, верны равенства:

$$\begin{aligned} \max_{K \geq L} Q^K &= \max_{K \geq L} \max_{\bar{k}: \bar{k} \in A_K} Q^{\bar{k}} = \\ &= \max_{K \geq L} \max_{\bar{k}: \bar{k} \in A_K} \sum_{j=1}^L Q_j^{k_j} = \sum_{j=1}^L \max_{k_j > 0} Q_j^{k_j}, \quad (3) \end{aligned}$$

т. е. решением задачи (1) является $K^* = \sum_{j=1}^L k_j^*$, где k_j^* — решение задачи

$$\max_{k_j > 0} Q_j^{k_j}, \quad j = 1, \dots, L. \quad (4)$$

Таким образом, задача (1) сводится к задаче, аналогичной (4), рассмотренной в работе [12] для СМО типа $M/G/1$ с ограничением на длину очереди.

В работе [12] для дохода СМО типа $M/G/1$ при заданном пороговом значении длины очереди k получено выражение:

$$Q^k = \lambda (1 - \theta^k) g^k,$$

где g^k — значение предельного дохода системы $M/G/1/k$, усредненного по числу обслуженных линией заявок; θ^k — вероятность того, что поступившая заявка будет потеряна (для величин g^k и θ^k в [12] приведены явные аналитические выражения). Эта формула для системы $M/G_j/1/k_j$ ($j = 1, \dots, L$) в новых обозначениях записывается в виде:

$$Q_j^{k_j} = \lambda_j (1 - \theta_j^{k_j}) g_j^{k_j}.$$

Здесь (см. [12])

$$\theta_{k_j}^{k_j} = 1 - \frac{1}{\frac{k_j}{\pi_{j_0}^{k_j}} + \rho}, \quad g^{k_j} = \sum_{i=0}^{k_j-1} q_{j_i}^{k_j} \pi_{j_i}^{k_j},$$

где

$$\begin{aligned} q_{j_i}^{k_j} &= C_{0j} \bar{v} - C_{1j} \sum_{m=k_j-i+1}^{\infty} (m - k_j + i) r_{jm} - \\ &- \frac{C_{2j}}{\lambda_j} \left[\frac{1}{2} \sum_{m=2}^{k_j-i+1} (m-1) m r_{jm} + (k_j - i) \sum_{m=k_j-i+2}^{\infty} m r_{jm} - \right. \\ &- \left. \frac{1}{2} (k_j - i) (k_j - i + 1) \sum_{m=k_j-i+2}^{\infty} r_{jm} \right] - \\ &- C_{2j} (i-1) \bar{v}_j - C_4 \bar{v}_j, \quad 1 \leq i \leq k_j - 1; \end{aligned}$$

$$q_{j0}^{k_j} = q_{j1}^{k_j} - \frac{C_{3j} + C_{4j}}{\lambda_j};$$

$$r_{jm} = \int_0^\infty \frac{(\lambda_j v)^m}{m!} e^{-\lambda_j v} dG_j(v) \text{ при } m \geq 0.$$

Для системы $M/G_j/1/k_j$ функции $\tilde{G}(k)$ и $\tilde{F}(k)$, введенные в [12], в новых обозначениях примут вид:

$$H_j(k_j) = C_{0j} - \frac{C_{1j}}{\bar{v}_j} (\rho_j - 1) - C_{4j} - \frac{C_{2j}}{\rho_j} [\tilde{F}_j(k_j) + (k_j - 1)\rho_j]; \quad (5)$$

$$\tilde{F}(k_j) = \left(\sum_{i=1}^{k_j} \pi_{ji}^{k_j+1} \sum_{m=k_j-i+1}^{\infty} [m - (k_j - i + 1)] k_m + \pi_{j0}^{k_j+1} \sum_{m=k_j}^{\infty} (m - k_j) k_m \right) / \pi_{jk_j}^{k_j+1}.$$

Из (3), (5) и утверждений 1 и 2 работы [12] вытекают следующие два свойства решения задачи (1).

Свойство 1. При любых значениях параметров $0 \leq C_{ij} < \infty$, $i = 0, 1, 3, 4$, и $0 < C_{2j} < \infty$, $j = 1, \dots, L$, существует оптимальное значение $0 < K^* < \infty$, причем

$$K^* = \sum_{j=1}^L k_j^*,$$

где $k_j^* = \min \{k_j > 0 : H_j(k_j) \leq Q_j^{k_j}\}$. Если $C_{2j} = 0$ и $Q_j^1 < H_j(1)$ хотя бы для одного $j \in \overline{1, L}$, то $k_j^* = \infty$, $K^* = \infty$.

Свойство 2. Значение $0 < K^* < \infty$ удовлетворяет условию $\max_{K>0} Q^K = Q^{K^*}$ тогда и только тогда, когда

$$K^* = \sum_{j=1}^L k_j^*$$

и $\bar{k}^* = (k_1^*, \dots, k_L^*)$ удовлетворяет одному из двух условий:

(а) для каждого $j \in \overline{1, L}$

$$k_j^* = \begin{cases} \min \{r_j > 0 : H_j(k_j) \leq Q_j^{k_j}\} \\ \text{при } 0 < C_{2j} < \infty \text{ и } H_j(1) > Q_j^1; \\ 1 & \text{при } H_j(1) \leq Q_j^1; \end{cases}$$

(б) для каждого $j \in \overline{1, L}$

$$Q_j^{k_j^*-1} < Q_j^{k_j^*}; \quad Q_j^{k_j^*+1} \leq Q_j^{k_j^*}.$$

Из условия (а) свойства 2 вытекает следующий алгоритм нахождения оптимальных значений \bar{k}^* и K^* [12].

1. Положить $j = 1$.
2. Положить $k_j = 1$.
3. Вычислить $H_j(k_j)$ и $Q_j^{k_j}$.
4. Если $C_{2j} = 0$ и $Q_j^{k_j} < H_j(k_j)$, то положить $k_j^* = \infty$ и перейти к п. 9.
5. Если выполняется неравенство $Q_j^{k_j} \geq H_j(k_j)$, то перейти к п. 8.
6. Увеличить k_j на единицу.
7. Вычислить $H_j(k_j)$ и $Q_j^{k_j}$ и перейти к п. 4.
8. Положить $k_j^* = k_j$.
9. Если $j < L$, то увеличить j на единицу и перейти к п. 2.
10. Положить $K^* = \sum_{j=1}^L k_j^*$.
11. Конец алгоритма.

Приведем некоторые выражения, упрощающие расчет величин $\tilde{F}_j(k_j)$ и $H_j(k_j)$, $j = 1, \dots, L$, и показывающие некоторые их свойства. Используя соотношения для $\pi_{ji}^{k_j}$ и $\bar{F}_j(k_j)$, полученные в [12] (см. (4) и (12)), найдем:

$$\tilde{F}_j(k_j) = \bar{F}_j(k_j) + \rho_j - 1 = \frac{\pi_{j0}^{k_j+1} + \rho_j - 1}{\pi_{jk_j}^{k_j+1}} - r_{j0}. \quad (6)$$

Отсюда получаем:

$$\begin{aligned} \tilde{F}_j(k_j) - \tilde{F}_j(k_j + 1) &= \\ &= \frac{\pi_{j0}^{k_j+1}}{\pi_{jk_j}^{k_j+1}} + \frac{\rho_j - 1}{\pi_{jk_j}^{k_j+1}} - \frac{\pi_{j0}^{k_j+2}}{\pi_{jk_j}^{k_j+2}} - \frac{\rho_j - 1}{\pi_{jk_j}^{k_j+2}} = \frac{\pi_{j0}^{k_j+1}}{\pi_{jk_j}^{k_j+1}} - \\ &- \frac{\pi_{j0}^{k_j+1} (1 - \pi_{jk_j}^{k_j+2})}{\pi_{jk_j}^{k_j+2}} + (\rho_j - 1) \left(\frac{1}{\pi_{jk_j}^{k_j+1}} - \frac{1}{\pi_{jk_j}^{k_j+2}} \right) = \\ &= \left(\pi_{j0}^{k_j+1} + \rho_j - 1 \right) \left(\frac{1}{\pi_{jk_j}^{k_j+1}} - \frac{1}{\pi_{jk_j}^{k_j+2}} \right) + \pi_{j0}^{k_j+1}. \end{aligned}$$

Из правой части последнего равенства следует, например, что величины $\tilde{F}_j(k_j) - \tilde{F}_j(k_j + 1) < \rho_j$ (т. е. $H_j(k_j)$ — убывающая функция по k_j) при условии $\pi_{j0}^{k_j+1} \leq \rho_j$ (заметим, что выполняются неравенства $\pi_{j0}^{k_j+1} + \rho_j - 1 > 0$, $1/\pi_{jk_j}^{k_j+1} - 1/\pi_{jk_j}^{k_j+2} < 0$, так как $\pi_{j0}^{k_j+1}$ и $\pi_{jk_j}^{k_j+1}$ убывают по $k_j \geq 0$ и $\lim_{k_j \rightarrow \infty} \pi_{j0}^{k_j+1} = 1 - \rho_j$ [14]).

Из формул (5), (6) данной статьи и (4) из [12] следует, что при известных значениях величин r_{ji} , $i = 1, \dots, k_j$, $j = 1, \dots, L$, для вычисления значений \bar{k}^* и K^* предложенный алгоритм требует порядка $\sum_{j=1}^L k_j^{*2}$ арифметических операций.

4 Решение задачи (2)

Для решения задачи (2) предлагается алгоритм, состоящий из следующих пунктов.

1. Если $N = L$, то положить $k_j^0 = 1, j = 1, \dots, L, K^0 = L$ и перейти к п. 12.
2. Положить $K = L, m = 0, k_j = 1, \varphi_j = 0, j = 1, \dots, L$.
3. Вычислить $g_j^{k_j}, j = 1, \dots, L$.
4. Для $j = 1, \dots, L$ выполнить:
 - 1.1 Вычислить $g_j^{k_j+1}, \Delta_j(k_j) = g_j^{k_j+1} - g_j^{k_j}$;
 - 2.1 Если $\Delta_j(k_j) \leq 0$, то положить $\varphi_j = 1, m = m + 1$.
5. Если $m = L$, то перейти к п. 11.
6. Найти $j_0 : \Delta_{j_0}(k_{j_0}) = \max\{\Delta_j(k_j), \varphi_j \neq 1, j = 1, \dots, L\}$.
7. Положить $K = K + 1, k_j = k_j + 1$ для $j = j_0$.
8. Если $K = N$, то перейти к п. 11.
9. Вычислить $g_j^{k_j+1}, \Delta_j(k_j) = g_j^{k_j+1} - g_j^{k_j}$ для $j = j_0$.
10. Если $\Delta_{j_0}(k_{j_0}) \leq 0$ то положить $\varphi_{j_0} = 1, m = m + 1$ и перейти к п. 5.
11. Положить $k_j^0 = k_j, j = 1, \dots, L, K^0 = \sum_{j=1}^L k_j^0$.
12. Конец алгоритма.

Справедливо следующее

Утверждение 1. *Результат работы алгоритма является решением задачи (2).*

Доказательство. Справедливость утверждения при $N = \infty$ следует непосредственно из условия (б) свойства 2 решения задачи (1). Действительно, согласно алгоритму в этом случае для каждого $j = 1, \dots, L$ величина k_j увеличивается до тех пор, пока не выполняются неравенства упомянутого выше условия свойства 2. В этом случае $K^0 = K^*$.

Докажем для случая $N < \infty$. Если $K^* = \sum_{j=1}^L k_j^* < N$, то, как легко видеть, алгоритм работает так же, как и при $N = \infty$, и, следовательно, утверждение справедливо и $K^0 = K^*$.

Пусть $K^* \geq N$. В работах [11, 12] для $g_j^{k_j}, j = 1, \dots, L$, приводится равенство:

$$g_j^{k_j} - g_j^{k_j+1} = \pi_{j k_j}^{k_j+1} \left[g_j^{k_j} - G_j(k_j) \right], \quad (7)$$

где $G_j(k_j) = H_j(k_j) \bar{v}_j$.

Согласно следствию теоремы 1 из [11] (см. пп. 1 и 3) верны неравенства $g_j^{k_j} < g_j^{k_j+1}$ при $0 < k_j < k_j^*, j = 1, \dots, L$. Тогда, так как $\pi_{j k_j}^{k_j+1} < 1$ и $H_j(k_j) \bar{v}_j -$

убывающие по k_j функции, из (6) получаем неравенства:

$$\Delta_j(k_j) < \Delta_j(k_j + 1) \quad \text{для } 0 < k_j \leq k_j^* - 2 \text{ при } k_j^* > 2. \quad (8)$$

Рассмотрим последовательности $\{\Delta_j(k_j), k_j > 0\}, j = 1, \dots, L$, и объединим их в одну последовательность $\{\Delta(k), k > 0\}$, упорядоченную по убыванию элементов. Обратим внимание, что алгоритм каждый раз, перед тем как увеличить K на единицу, в п. 6 находит элемент $\Delta(K + 1) = \Delta_{j_0}(k_{j_0})$ объединенной последовательности и выполняет п. 7. Заметим также, что

$$\begin{aligned} g^{\bar{k}} &= \sum_{j=1}^L g_j^{k_j} = \sum_{j=1}^L \left[\sum_{m=1}^{k_j-1} \Delta_j(m) + g_j^1 \right] = \\ &= \sum_{j=1}^L \sum_{m=1}^{k_j-1} \Delta_j(m) + \sum_{j=1}^L g_j^1, \end{aligned}$$

и, если $\bar{k}' = (k'_1, \dots, k'_L) \in U_{\text{СР}, K}$ удовлетворяет условию $g^{\bar{k}'} = \max_{\bar{k} \in U_{\text{СР}, K}} g^{\bar{k}}$, то $\max_{\bar{k} \in U_{\text{СР}, K+1}} g^{\bar{k}'} = g^{\bar{k}'} + \Delta_{j_0}(k'_{j_0})$, где j_0 вычисляется согласно п. 6 алгоритма. Тогда, так как $K^* \geq N$ и имеет место (8), согласно алгоритму по завершении п. 11 имеют место равенства:

$$K^0 = \sum_{j=1}^L k_j^0 = N;$$

$$\begin{aligned} g^{\bar{k}^0} &= \sum_{j=1}^L g_j^{k_j^0} = \sum_{j=1}^L \sum_{m=1}^{k_j^0-1} \Delta_j(m) + \sum_{j=1}^L g_j^1 = \\ &= \sum_{k=L+1}^N \Delta(k) + \sum_{j=1}^L g_j^1 = \\ &= \max_{\bar{k} \in U_{\text{СР}, N}} \left[\sum_{j=1}^L \sum_{m=1}^{k_j-1} \Delta_j(m) + \sum_{j=1}^L g_j^1 \right] = \max_{\bar{k} \in U_{\text{СР}, N}} g^{\bar{k}}. \end{aligned}$$

Следовательно, утверждение 1 справедливо.

5 Примеры

Рассмотрим СМО $M_2/G/2$ с функциями распределения времени обслуживания

$$\begin{aligned} G_1(t) &= f_{11} (1 - e^{-\mu_{11}t}) + f_{12} (1 - e^{-\mu_{12}t}); \\ G_2(t) &= f_{21} (1 - e^{-\mu_{21}t}) + f_{22} (1 - e^{-\mu_{22}t}), \end{aligned}$$

стоимостными параметрами $C_{10} = C_{20} = 30, C_{11} = C_{21} = 15, C_{12} = C_{22} = 4, C_{13} = C_{23} = 0, C_{14} = C_{24} = 0,4$ и одинаковыми входными потоками

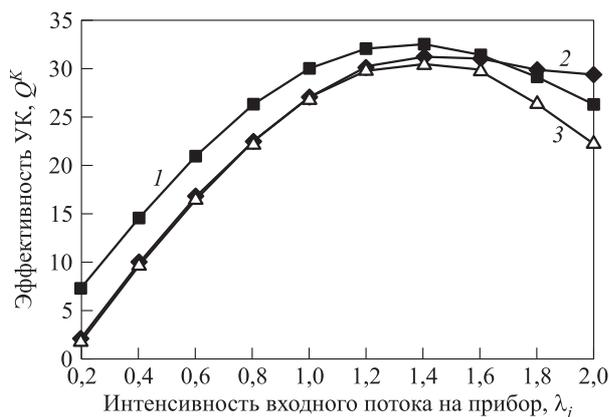


Рис. 1 Зависимость величины дохода системы от интенсивности входного потока: 1 — оптимальная схема СР; 2 — схема квадратного корня; 3 — схема СС

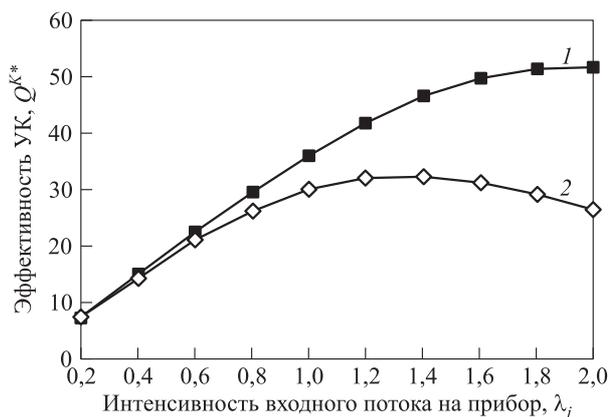


Рис. 2 Зависимость дохода системы от интенсивности входного потока и времени обслуживания приборами: 1 — $\mu_{11} = 1,5, \mu_{12} = 3, \mu_{21} = 2$ и $\mu_{22} = 4$; 2 — $\mu_{11} = 1, \mu_{12} = 2, \mu_{21} = 1,5$ и $\mu_{22} = 3$

на приборы (выходные линии передачи) $\lambda_1 = \lambda_2$. На графиках рис. 1 проиллюстрированы зависимости величины дохода Q^k рассматриваемого СМО от интенсивности входного потока для трех схем распределения БП: оптимальной схемы СР, схемы квадратного корня при $K = K^*$, схемы СС при $K = K^*$. На всех трех графиках доход обозначен через Q^k . Значения параметров функций $G_1(t)$ и $G_2(t)$ соответственно равны: $f_{11} = f_{12} = f_{21} = f_{22} = 0,5; \mu_{11} = 1; \mu_{12} = 2; \mu_{21} = 1,5; \mu_{22} = 3$.

На рис. 2 изображена зависимость значения дохода Q^{k*} системы от интенсивности входного потока и времени обслуживания приборами.

На рис. 3 приведена диаграмма зависимости оптимального значения K^* от интенсивности входного потока.

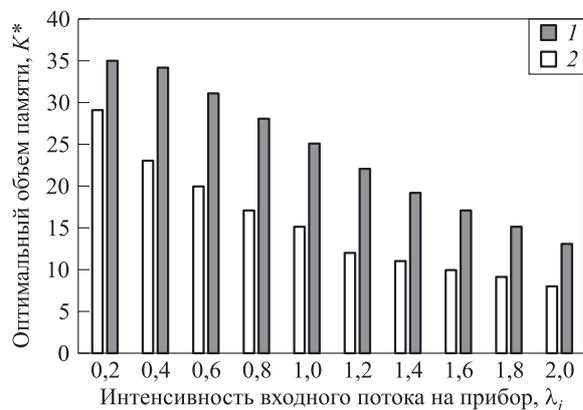


Рис. 3 Зависимость оптимального объема БП от интенсивности входного потока и времени обслуживания приборами: 1 — $\mu_{11} = 1,5, \mu_{12} = 3, \mu_{21} = 2$ и $\mu_{22} = 4$; 2 — $\mu_{11} = 1, \mu_{12} = 2, \mu_{21} = 1,5$ и $\mu_{22} = 3$

6 Заключение

На основании проведенных выше исследований можно сделать следующие выводы:

- при схеме СР, если параметр стоимости для времени задержки — положительная величина, то существует конечный объем БП, при котором функция дохода Q^k достигнет максимального значения на множестве планов $U_{СР}$;
- функция Q^k является унимодальной функцией от переменной $K > 0$;
- функция дохода УК зависит от объема БП, интенсивностей входных потоков и распределения БП между выходными линиями (см. рис. 1 и 2), т.е. оптимальная процедура локального управления потоками в УК должна управлять всеми тремя указанными параметрами;
- оптимальный объем БП существенно зависит от интенсивности входной нагрузки, механизма распределения БП и пропускной способности каналов связи (см. рис. 1 и 3).

Отметим, что соотношение вида (7), использованное при доказательстве приведенного выше утверждения 1, получено также и в работе [13] для СМО типа $G/M/1$ с пороговым значением длины очереди. Указанное соотношение (см. (12) в [13]) имеет вид:

$$g^k - g^{k+1} = \pi_0^{k+1} [g^k - G(k)], \quad k > 0,$$

где k соответствует объему БП; π_0^{k+1} и $G(k)$ обладают аналогичными свойствами величин $\pi_{jk}^{k_j+1}$ и $G_j(k_j)$, использованным выше в доказательстве

утверждения 1. Следовательно, алгоритм и утверждение 1, приведенные в разд. 3, применимы для рассмотренной выше модели УК также при предположениях: входящие потоки — рекуррентные; времена обслуживания — экспоненциальные. Однако заметим, что в этом случае, в отличие от рассмотренного выше, согласно работе [13] g^k — предельный доход, усредненный по числу поступивших в систему заявок.

Отметим также, что каждая из схем CP, CS, SMQ, SMA и SMQMA предпочтительней в смысле эффективности, чем остальные, только в соответствующей области значений параметров системы и выбранного критерия эффективности [5]. Полученные с помощью предложенных выше алгоритмов оптимальные при схеме CP значения объема БП и плана распределения БП могут рассматриваться для других схем как оценочные.

Литература

1. Клейнрок Л. Вычислительные системы с очередями / Пер. с англ. — М.: Мир, 1979. 600 с. (*Kleinrock L. Queueing systems. Vol. II: Computer applications.* — New York, NY, USA: Wiley, 1976. 549 p.)
2. Дэвис Д., Барбер Д., Прайс У., Соломонидес С. Сети связи для вычислительных систем / Пер. с англ. — М.: Мир, 1982. 563 с. (*Davies D. W., Barber D. L. A., Price W. L., Solomonides C. M.* 1979. *Computer networks and their protocols.* — New York, NY, USA: Wiley. 487 p.)
3. Irland M. Buffer management in a packet switch // IEEE T. Commun., 1978. Vol. 26. No. 3. P. 328–337.
4. Kamoun F., Kleinrock L. Analysis of shared finite storage in a computer networks node environment under general traffic conditions // IEEE T. Commun., 1980. Vol. 28. No. 7. P. 992–1003.
5. Башарин Г. П., Самуйлов К. Е. Об оптимальной структуре буферной памяти в сетях передачи данных с коммутацией пакетов. — М., 1982. 70 с.
6. Михеев П. А. Анализ стратегий разделения конечной буферной памяти маршрутизатора между выходными каналами // Автоматика и телемеханика, 2014. № 10. С. 125–128.
7. Forschini G., Gopinath B. Sharing memory optimally // IEEE T. Commun., 1983. Vol. 31. No. 3. P. 352–359.
8. Агаларов Я. М. Двухступенчатое адаптивное управление потоками в узле коммутации телекоммуникационной сети // Вестник РУДН. Сер. Прикладная математика и информатика, 2003. № 1. С. 134–141.
9. Линец Г. И. Управление объемом буферной памяти и пропускной способностью каналов в мультисервисных сетях // Инфокоммуникационные технологии, 2008. Т. 6. № 2. С. 62–64.
10. Жерновский Ю. В. Решение задач оптимального синтеза для некоторых марковских моделей обслуживания // Информационные процессы, 2010. Т. 10. № 3. С. 257–274.
11. Агаларов Я. М., Агаларов М. Я., Шоргин В. С. Об оптимальном пороговом значении длины очереди в одной задаче максимизации дохода СМО типа $M/G/1$ // Информатика и её применения, 2016. Т. 10. Вып. 2. С. 70–79.
12. Агаларов Я. М. Максимизация среднего стационарного дохода СМО типа $M/G/1$ // Информатика и её применения, 2017. Т. 11. Вып. 2. С. 25–32.
13. Агаларов Я. М., Шоргин В. С. Об одной задаче максимизации дохода системы массового обслуживания типа $G/M/1$ с пороговым управлением очередью // Информатика и её применения, 2017. Т. 11. Вып. 4. С. 55–64.
14. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. — М.: РУДН, 1995. 529 с.

Поступила в редакцию 10.06.18

OPTIMIZATION OF BUFFER MEMORY SIZE OF SWITCHING NODE IN MODE OF FULL MEMORY SHARING

Ya. M. Agalarov

Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The problem of optimizing the volume of the buffer memory of the network switching node with the CP (Complete Partitioning) is considered, which uses a scheme for the complete memory sharing between the output transmission lines, taking into account packet loss due to memory overflow and average packet delays. The properties of the solution of the problem are obtained, and an algorithm for searching for the optimal volume is developed for the switching node model with Poisson incoming flows, common service time distributions, and single-channel output transmission lines. Within the given model, an algorithm for finding the optimal buffer

memory allocation plan for a given volume is developed. The results of computational experiments using the developed algorithms are presented. The applicability of the proposed algorithms and approaches is obtained in the framework of the switching node model with recurrent incoming flows, exponential service times, and single-channel transmission lines.

Keywords: switching node; flow management; buffer memory allocation; optimal buffer memory capacity

DOI: 10.14357/19922264180404

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (projects 18-07-00692, 17-07-000717, and 17-07-00851).

References

1. Kleinrock, L. 1976. *Queueing systems. Vol. II: Computer applications*. New York, NY: Wiley. 549 p.
2. Davies, D.W., D.L.A. Barber, W.L. Price, and C.M. Solomonides. 1979. *Computer networks and their protocols*. New York, NY: Wiley. 487 p.
3. Irland, M. 1978. Buffer management in a packet switch. *IEEE T. Commun.* 26(3):328–337.
4. Kamoun, F., and L. Kleinrock. 1980. Analysis of shared finite storage in a computer networks node environment under general traffic conditions. *IEEE Trans. Commun.* 28(7):992–1003.
5. Basharin, G. P., and K. E. Samuilov. 1982. *Ob optimal'noy strukture bufernoy pamyati v setyakh peredachi dannykh s kommutatsiy paketov* [On the optimal structure of BP in data transmission networks with packet commutation]. — Moscow. 70 p.
6. Mikheev, P.A. 2014. Analyzing sharing strategies for finite buffer memory in a router among outgoing channels. *Automat. Rem. Contr.* 75(10):1814–1825.
7. Forschini, G., and B. Gopinath. 1983. Sharing memory optimally. *IEEE T. Commun.* 31(3): 352–359.
8. Agalarov, Ya. M. 2003. Dvukhstupenchatoe adaptivnoe upravlenie potokami v uzle kommutatsii telekommunikatsionnoy seti [Two-level flow control for a switching node of communication network]. *Vestnik RUDN. Ser. Prikladnaya matematika i informatika* [PFUR Bull., Applied Mathematics and Computer Science] 1:134–141.
9. Linets, G. I. 2010. Upravlenie ob'emom bufernoy pamyati i propusknoy sposobnost'yu kanalov v mul'tiservisnykh setyakh [Volume management of buffer memory and throughput of channels in multiservice networks]. *Infokommunikatsionnye tekhnologii* [Information Communication Technologies] 6(2):62–64.
10. Zhernovyy, Ju. V. 2010. Reshenie zadach optimal'nogo sinteza dlya nekotorykh markovskikh modeley obsluzhivaniya [Solution of optimum synthesis problem for some Markov models of service]. *Informatsionnye protsessy* [Information Processes] 10(3):257–274.
11. Agalarov, Ya.M., M. Ya. Agalarov, and V. S. Shorgin. 2016. Ob optimal'nom porogovom znachenii dliny ocheredi v odnoy zadache maksimizatsii dokhoda SMO tipa $M/G/1$ [About the optimal threshold of queue length in particular problem of profit maximization in $M/G/1$ queueing system]. *Informatika i ee Primeneniya — Inform. Appl.* 10(2):70–79.
12. Agalarov, Ya. M. 2017. Maksimizatsiya srednego statsionarnogo dokhoda SMO tipa $M/G/1$ [Maximization of average stationary profit in $M/G/1$ queueing system]. *Informatika i ee Primeneniya — Inform. Appl.* 11(2):25–32.
13. Agalarov, Ya.M., and V.S. Shorgin. 2016. Ob odnoy zadache maksimizatsii dokhoda sistemy massovogo obsluzhivaniya tipa $G/M/1$ s porogovym upravleniem ochered'yu [Profit maximization in $G/M/1$ queueing system on a set of threshold strategies with two switch points]. *Informatika i ee Primeneniya — Inform. Appl.* 11(4): 55–64.
14. Bocharov, P.P., and A.V. Pechinkin. 1995. *Teoriya massovogo obsluzhivaniya* [Queueing theory]. Moscow: RUDN. 529 p.

Received June 10, 2018

Contributor

Agalarov Yaver M. (b. 1952) — Candidate of Science (PhD) in technology, associate professor; leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; agglar@yandex.ru