

Простейшие модели управления очередью в узлах сетей с коммутацией пакетов

Ю. В. Жерновий

Львовский национальный университет им. Ивана Франко, Львов, Украина

Поступила в редколлегию 23.06.2014

Аннотация—Рассмотрены системы обслуживания, моделирующие работу узлов сетей передачи пакетов данных с использованием алгоритмов вероятностного отбрасывания пакетов: одноканальная система с ограниченным буфером и многоканальная система без ограничений на длину очереди. Входящий поток моделируется с помощью обобщённого пуассоновского процесса, а время обслуживания показательно распределено. Предложен алгоритм определения стационарного распределения числа заявок и стационарных характеристик системы (средней длины очереди, среднего времени ожидания в очереди, вероятности обслуживания заявок). Полученные результаты проверены с помощью имитационных моделей, построенных с привлечением инструментальных средств GPSS World.

КЛЮЧЕВЫЕ СЛОВА: системы обслуживания, групповое поступление заявок, активное управление очередью, вероятностное отбрасывание заявок, стационарные характеристики.

1. ВВЕДЕНИЕ

С целью предотвращения перегрузок в узлах сетей с коммутацией пакетов (АТМ, TCP/IP и др.) используются алгоритмы активного управления очередью (англ. *Active Queue Management — AQM*). В соответствующей системе обслуживания, моделирующей работу узла сети, каждый поступающий пакет может быть отброшен с определённой вероятностью, даже если буфер ещё полностью не заполнен. Вероятность отбрасывания зависит от длины очереди в момент поступления пакета. Зависимость вероятности отбрасывания пакетов от длины очереди называют функцией отбрасывания [1].

В алгоритмах AQM используют различные функции отбрасывания, например в известном алгоритме RED (Random Early Detection — случайное раннее обнаружение) [2] эта функция является линейной. Благодаря профилактическому случайному отбрасыванию пакетов алгоритм AQM косвенно информирует отправителя, использующего протокол TCP, о приближающейся перегрузке. Применение такого алгоритма в маршрутизаторе может принести много полезных эффектов, в том числе сокращение очереди и времени сетевой задержки (больше подробностей можно найти в [3]).

Результаты исследований показывают [1], что механизм функции отбрасывания является мощным средством для регулирования параметров системы обслуживания. С его помощью можно регулировать не только длину очереди, вероятность потери заявок, время ожидания, дисперсию длины очереди, но и несколько из этих параметров одновременно. Модели с функцией отбрасывания имеют также свой глубокий универсальный смысл. Во всех случаях, когда нет возможности регулировать параметры системы обслуживания через изменения входящего потока либо процесса обслуживания, применение функции отбрасывания является простым и эффективным способом для обеспечения требуемых параметров системы обслуживания.

Использование теории массового обслуживания для анализа алгоритмов активного управления очередью началось в последние годы [1, 4–7]. Как правило, авторы ограничивались изучением одноканальных систем с ограниченным буфером.

Системы обслуживания с групповым поступлением заявок могут использоваться для моделирования процессов в сетях с пакетной коммутацией двумя способами: моделируя поступление групп пакетов одинаковых размеров или моделируя поступление отдельных пакетов разных размеров.

В настоящей статье мы рассматриваем многоканальную систему обслуживания с групповым поступлением заявок (пакетов) без ограничений на длину очереди с общей функцией отбрасывания пакетов. Изучаемая n -канальная система может служить моделью узла сети, в который поступают группы пакетов одинаковых размеров (одна заявка — один пакет). В то же время соответствующая одноканальная система (частный случай многоканальной системы при $n = 1$) может рассматриваться как модель узла с поступлением отдельных пакетов разных размеров.

Вместе с тем, мы отдельно изучаем одноканальную систему с ограниченным буфером, которая является моделью узла сети с поступлением отдельных пакетов разных размеров.

В настоящей работе мы ограничиваемся рассмотрением систем обслуживания с показательными распределениями промежутков времени между моментами поступления групп заявок и времени обслуживания заявок. Это позволяет получить простой алгоритм определения стационарных характеристик не только для одноканальной, но и для многоканальной системы, в которой реализован механизм функции отбрасывания пакетов.

2. ОДНОКАНАЛЬНАЯ СИСТЕМА С ОГРАНИЧЕННЫМ БУФЕРОМ КАК МОДЕЛЬ ПОСТУПЛЕНИЯ ПАКЕТОВ РАЗНЫХ РАЗМЕРОВ

Изучаемая в этом пункте система обслуживания может служить моделью узла сети в случае поступления отдельных пакетов разных размеров. Предполагается, что вероятность отбрасывания зависит от длины очереди в момент поступления пакета.

Рассмотрим одноканальную систему обслуживания $M^X/M/1/m$ с ограниченным буфером объёма m . Промежутки времени между моментами поступления групп заявок входящего потока — независимые случайные величины, распределённые по показательному закону с параметром λ . Время обслуживания каждой заявки распределено по показательному закону с параметром μ . С вероятностью a_k число заявок в группе входящего потока равно k ($k \geq 1$), причём

$$\sum_{k=1}^{\infty} a_k = 1, \quad a_{(1)} = \sum_{k=1}^{\infty} k a_k < \infty.$$

Переходя к интерпретации этой системы обслуживания как модели узла сети для передачи пакетов данных, можно, например, предположить, что объём буфера m выражается в гигабайтах. Тогда с вероятностью a_k размер поступающего пакета равен k гигабайт.

Однако такая интерпретация имеет некоторые важные последствия. Например, если в буфере осталось 100 GB свободного места, то на обслуживание не может быть принят пакет размером 101 GB и больше. Поскольку пакеты нельзя делить, то нельзя принять в буфер фрагмент пакета.

Исходя из вышесказанного, будем использовать следующее правило принятия группы заявок в систему: прибывающая группа заявок получает отказ, если для всей группы не хватает места в буфере, даже если буфер ещё полностью не заполнен. Итак, если в момент прибытия группы заявок размером k в системе находится i заявок, то при условии $i + k > m + 1$ вся группа заявок получает отказ.

В качестве алгоритма управления очередью применим общую функцию отбрасывания пакетов, определяемую набором вероятностей β_i . Если в момент прибытия группы заявок раз-

мером k в системе находится i заявок и выполнено условие $i + k \leq m + 1$, то эта группа заявок принимается на обслуживание с вероятностью β_i ($0 \leq i \leq m$).

3. СТАЦИОНАРНЫЕ ХАРАКТЕРИСТИКИ ОДНОКАНАЛЬНОЙ СИСТЕМЫ

Пусть $p_i(t)$ — вероятность того, что в момент времени t в системе находится i заявок ($0 \leq i \leq m + 1$). Пределы $p_i = \lim_{t \rightarrow \infty} p_i(t)$ ($0 \leq i \leq m + 1$), очевидно, существуют. С целью упрощения системы уравнений для стационарных вероятностей p_i введём обозначения: $\nu_i = \mu/\beta_i$ ($1 \leq i \leq m$), $\nu_{m+1} = \mu$; $\alpha_i = \lambda/\nu_i$ ($1 \leq i \leq m + 1$); $\pi_i = \beta_i p_i/p_0$ ($1 \leq i \leq m$), $\pi_0 = 1$, $\pi_{m+1} = p_{m+1}/p_0$; $\alpha = \lambda/\mu$. Сделаем естественное предположение, что $\beta_0 = 1$.

Систему уравнений для стационарных вероятностей запишем как систему относительно π_i :

$$\begin{aligned} & -\lambda \sum_{k=1}^{m+1} a_k + \nu_1 \pi_1 = 0; \\ & -\lambda \pi_i \sum_{k=1}^{m+1-i} a_k + \lambda \sum_{k=0}^{i-1} \pi_k a_{i-k} - \nu_i \pi_i + \nu_{i+1} \pi_{i+1} = 0, \quad 1 \leq i \leq m; \\ & p_0 \left(1 + \sum_{k=1}^m \frac{\pi_k}{\beta_k} + \pi_{m+1} \right) = 1. \end{aligned} \tag{1}$$

После решения системы уравнений (1) приходим к утверждению.

Теорема 1. *Стационарные вероятности p_i определяются в виде*

$$p_0 = \frac{1}{1 + \sum_{k=1}^m \frac{\pi_k}{\beta_k} + \pi_{m+1}}; \quad p_i = \frac{p_0 \pi_i}{\beta_i}, \quad 1 \leq i \leq m; \quad p_{m+1} = p_0 \pi_{m+1}, \tag{2}$$

где π_i находим с помощью рекуррентных соотношений

$$\pi_0 = 1; \quad \pi_i = \alpha_i \sum_{j=0}^{i-1} \pi_j \sum_{k=1}^{m+1} a_{k-j}, \quad 1 \leq i \leq m + 1. \tag{3}$$

Доказательство. Равенство (2) для p_0 следует из последнего уравнения (1) (условия нормировки). Рекуррентные соотношения (3) легко доказать методом математической индукции. \square

Приведём формулы для стационарных характеристик системы: средней длины очереди $\mathbf{E}(Q)$, среднего числа заявок в системе $\mathbf{E}(N)$, вероятности обслуживания пакета \mathbf{P}_{sv} и среднего времени ожидания начала обслуживания $\mathbf{E}(W)$:

$$\mathbf{E}(Q) = \sum_{k=1}^m k p_{k+1}; \quad \mathbf{E}(N) = \sum_{k=1}^{m+1} k p_k = \mathbf{E}(Q) + 1 - p_0; \quad \mathbf{P}_{sv} = \frac{1 - p_0}{\alpha a_{(1)}}; \quad \mathbf{E}(W) = \frac{\mathbf{E}(Q)}{\lambda a_{(1)} \mathbf{P}_{sv}}.$$

Отметим, что формула для \mathbf{P}_{sv} получена как отношение стационарного среднего числа обслуженных пакетов за единицу времени $\mu(1 - p_0)/a_{(1)}$ к стационарному среднему числу поступивших пакетов за единицу времени λ .

4. ПРИМЕРЫ (ОДНОКАНАЛЬНАЯ СИСТЕМА С ОГРАНИЧЕННЫМ БУФЕРОМ)

Приведём результаты расчётов стационарных характеристик одноканальной системы, являющейся моделью узла сети в случае поступления отдельных пакетов разных размеров.

Предположим, что в одноканальный узел в среднем поступает 1,5 пакета информации за единицу времени (*e. в.*). В среднем за 1 *e. в.* непрерывно работающий узел обрабатывает 100 GB информации. Полная вместимость буфера — 400 GB (с учётом обрабатываемой в канале информации). Поступают пакеты трёх видов в зависимости от их размеров. С вероятностью $a_1 = 0,2$ прибывший пакет содержит 100 GB, с вероятностью $a_2 = 0,3$ — 200 GB и с вероятностью $a_3 = 0,5$ — 300 GB.

Взяв за единицу измерения объёма информации 100 GB и перейдя на язык теории массового обслуживания, получаем следующие значения параметров: $m = 3$; $\lambda = 1,5$; $\mu = 1$; $a_1 = 0,2$; $a_2 = 0,3$; $a_3 = 0,5$; $a_k = 0$ ($k \geq 4$).

В качестве алгоритма управления очередью применим общую функцию отбрасывания пакетов, определяемую набором вероятностей β_i ($0 \leq i \leq 3$). Рассмотрим 3 случая с разными наборами вероятностей β_i :

$\beta_i = 1, 0 \leq i \leq 3$ (случай I, вероятностное отбрасывание пакетов не применяется);

$\beta_0 = \beta_1 = 1, \beta_2 = 0,8, \beta_3 = 0,5$ (случай II);

$\beta_0 = \beta_1 = 1, \beta_2 = \beta_3 = 0,1$ (случай III).

Стационарные вероятности p_k ($0 \leq k \leq 4$) наличия в системе k заявок и стационарные характеристики \mathbf{P}_{sv} , $\mathbf{E}(Q)$, $\mathbf{E}(W)$ и $\mathbf{E}(N)$, вычисленные для различных наборов β_i , приведены в табл. 1. Анализируя полученные результаты, видим, что применение вероятностного отбрасывания пакетов приводит к незначительному уменьшению вероятности обслуживания (на 3,6%, если сравнивать случаи III и I) и в то же время к более ощутимому улучшению характеристик очереди: $\mathbf{E}(Q)$ уменьшается на 24,6%, $\mathbf{E}(W)$ — на 21,8%, а $\mathbf{E}(N)$ — на 17,2%.

Среднее стационарное число заявок в системе $\mathbf{E}(N)$ можно интерпретировать как среднее стационарное значение объёма информации, находящейся в буфере. Например, в случае II в буфере находится в среднем 253,12 GB информации, то есть в среднем он заполнен на 63,3%.

В табл. 2 приведено сравнение стационарных характеристик системы для случая II, вычисленных по формулам, предложенным в п. 3, с результатами имитационного моделирования, полученными с помощью системы GPSS World [8] (время моделирования $t = 10^6$) для различных распределений времени обслуживания: показательного распределения с параметром $\mu = 1$, равномерного распределения на отрезке $[0; 2]$, равномерного распределения на отрезке $[0,5; 1,5]$ и детерминированного значения, равного 1. Анализ данных таблицы показывает, что для показательного распределения времени обслуживания полученные аналитическим методом значения подтверждаются результатами имитационного моделирования. Уменьшение дисперсии распределения времени обслуживания приводит к незначительному улучшению показателей системы обслуживания: к увеличению вероятности обслуживания на 3,9% и уменьшению $\mathbf{E}(Q)$ и $\mathbf{E}(W)$ соответственно на 7,7% и на 11% (если сравнивать детерминированное и показательное распределения времени обслуживания).

Таблица 1. Стационарные характеристики системы для различных наборов β_i

№ набора β_i	p_0	p_1	p_2	p_3	p_4	\mathbf{P}_{sv}	$\mathbf{E}(Q)$	$\mathbf{E}(W)$	$\mathbf{E}(N)$
I	0,0653	0,0980	0,2254	0,3356	0,2756	0,2709	1,7236	1,8440	2,6582
II	0,0734	0,1101	0,2531	0,3390	0,2245	0,2686	1,6045	1,7316	2,5312
III	0,0988	0,1482	0,3408	0,2775	0,1348	0,2612	1,3001	1,4426	2,2013

Таблица 2. Сравнение стационарных характеристик системы для случая II с результатами имитационного моделирования для различных распределений времени обслуживания

Вид распределения	p_0	p_1	p_2	p_3	p_4	\mathbf{P}_{sv}	$\mathbf{E}(Q)$	$\mathbf{E}(W)$
показательное (аналитич. метод)	0,0734	0,1101	0,2531	0,3390	0,2245	0,2686	1,6045	1,7316
показательное (GPSS World)	0,0729	0,1104	0,2529	0,3385	0,2254	0,2690	1,6060	1,7330
равномерное на [0; 2] (GPSS World)	0,0562	0,1212	0,2912	0,3494	0,1821	0,2740	1,5360	1,6280
равномерное на [0,5; 1,5] (GPSS World)	0,0430	0,1320	0,3143	0,3520	0,1587	0,2780	1,4950	1,5620
детерминированное=1 (GPSS World)	0,0391	0,1344	0,3218	0,3534	0,1514	0,2790	1,4810	1,5410

5. МНОГОКАНАЛЬНАЯ СИСТЕМА БЕЗ ОГРАНИЧЕНИЙ НА ДЛИНУ ОЧЕРЕДИ

Изучаемая в этом пункте n -канальная система обслуживания $M^X/M/n$ может служить моделью узла сети в случае поступления групп пакетов одинаковых размеров. Превышение объёма буфера над поступающей нагрузкой считаем достаточным для того, чтобы не возникало необходимости отбрасывания пакетов по причине недостатка для них места в буфере.

Для рассмотренной выше одноканальной системы предполагалось, что "одна прибывающая группа заявок = одному поступающему пакету информации". Для данной системы считаем, что "одна прибывающая группа заявок = одной поступающей группе пакетов информации то есть "одна заявка = одному пакету".

Предположения п. 2 относительно распределений для входящего потока и времени обслуживания заявки оставим без изменений, сделав дополнительное предположение, что

$$a_{(2)} = \sum_{k=1}^{\infty} k^2 a_k < \infty.$$

В качестве алгоритма управления очередью, как и выше, применяем общую функцию отбрасывания пакетов, определяемую набором вероятностей β_i . Зафиксировав натуральное число h ($h > n$), определим набор вероятностей β_i с помощью равенств

$$\beta_i = \begin{cases} \beta_i, & 0 \leq i \leq h; \\ \beta, & i \geq h + 1. \end{cases}$$

Если в момент прибытия группы заявок в системе находится i заявок, то эта группа заявок принимается на обслуживание с вероятностью β_i ($i \geq 0$). Предположим, что $\beta \leq \beta_i$ ($0 \leq i \leq h$).

Уточним введённые в п. 2 обозначения:

$$\begin{aligned} \nu_i &= \mu/\beta_i, & \alpha_i &= \lambda/\nu_i, & i &\geq 1; \\ \pi_i &= \beta_i p_i/p_0, & i &\geq 0; & \alpha &= \lambda/\mu, & \rho &= \alpha \beta a_{(1)}. \end{aligned}$$

Введём нумерацию состояний системы: s_0 — в системе нет заявок; s_i ($i \geq 1$) — в системе находится i заявок. Пусть $p_i(t)$ — вероятность того, что в момент времени t система находится в состоянии s_i . Предполагая, что существуют пределы $p_i = \lim_{t \rightarrow \infty} p_i(t)$ ($i \geq 0$) (условия их

существования мы получим ниже), запишем систему уравнений для определения π_i

$$\begin{aligned}
 & -\lambda\pi_0 + \nu_1\pi_1 = 0; \\
 & -(\lambda + i\nu_i)\pi_i + \lambda \sum_{k=0}^{i-1} \pi_k a_{i-k} + (i+1)\nu_{i+1}\pi_{i+1} = 0, \quad 1 \leq i \leq n-1; \\
 & -(\lambda + n\nu_i)\pi_i + \lambda \sum_{k=0}^{i-1} \pi_k a_{i-k} + n\nu_{i+1}\pi_{i+1} = 0, \quad i \geq n; \\
 & p_0 \sum_{k=0}^{\infty} \frac{\pi_k}{\beta_k} = 1.
 \end{aligned} \tag{4}$$

Введём обозначения:

$$A_i = 1 - \sum_{k=1}^i a_k, \quad i \geq 0, \quad A_0 = 1.$$

Теорема 2. Если

$$a_{(1)} < \infty; \quad \rho < n; \quad \beta \leq \beta_i \quad (0 \leq i \leq h), \tag{5}$$

то стационарные вероятности p_i ($i \geq 0$) существуют и определяются в виде

$$p_0 = \frac{n - \rho}{n + \sum_{k=1}^{n-1} \frac{(n-k)\pi_k}{\beta_k} + \alpha a_{(1)} \left(\beta_0 - \beta + \sum_{k=1}^h \left(1 - \frac{\beta}{\beta_k} \right) \pi_k \right)}; \quad p_i = \frac{p_0 \pi_i}{\beta_i}, \quad i \geq 1, \tag{6}$$

где π_i находим с помощью рекуррентных соотношений

$$\begin{aligned}
 \pi_0 &= \beta_0; \quad \pi_{i+1} = \frac{\alpha_{i+1}}{i+1} \sum_{k=0}^i A_{i-k} \pi_k, \quad 0 \leq i \leq n-1; \\
 \pi_{i+1} &= \frac{\alpha_{i+1}}{n} \sum_{k=0}^i A_{i-k} \pi_k, \quad i \geq n.
 \end{aligned} \tag{7}$$

Доказательство. Пользуясь уравнениями (4), рекуррентные соотношения (7) легко доказать методом математической индукции. Для доказательства первого равенства (6) с помощью условия нормировки (4) находим

$$\begin{aligned}
 \sum_{i=0}^{\infty} \sum_{k=0}^i A_{i-k} \pi_k &= \sum_{i=0}^{\infty} \pi_i \sum_{k=0}^{\infty} A_k = \frac{a_{(1)}}{p_0} \sum_{i=0}^{\infty} \beta_i p_i = \frac{a_{(1)}}{p_0} \left(\beta + \sum_{i=0}^h (\beta_i - \beta) p_i \right); \\
 \sum_{i=1}^n \frac{i\pi_i}{\alpha_i} + n \sum_{i=n+1}^{\infty} \frac{\pi_i}{\alpha_i} &= \frac{1}{\alpha} \left(\sum_{i=1}^n \frac{i\pi_i}{\beta_i} + n \left(\frac{1}{p_0} - \sum_{i=0}^n \frac{\pi_i}{\beta_i} \right) \right).
 \end{aligned}$$

Учитывая соотношения (7), получаем равенство

$$\alpha a_{(1)} \left(\frac{\beta}{p_0} + \beta_0 - \beta + \sum_{i=1}^h (\beta_i - \beta) \frac{\pi_i}{\beta_i} \right) = \sum_{i=1}^n \frac{i\pi_i}{\beta_i} + n \left(\frac{1}{p_0} - 1 - \sum_{i=1}^n \frac{\pi_i}{\beta_i} \right),$$

приводящее к выражению (6) для p_0 .

Итак, при выполнении условий (5) система уравнений равновесия (4) имеет нетривиальное решение. Пусть $\xi(t)$ — число заявок в системе в момент времени t . Тогда $\{\xi(t), t \geq 0\}$ — случайный процесс с дискретными состояниями s_i ($i \geq 0$). Поскольку все состояния этого неприводимого марковского процесса сообщаются, то из критерия регулярности [9, с. 40] следует его регулярность, а из эргодической теоремы Фостера [9, с. 48] вытекает, что условия теоремы 2 являются достаточными для существования стационарных вероятностей p_i ($i \geq 0$). Теорема доказана. \square

6. СТАЦИОНАРНЫЕ ХАРАКТЕРИСТИКИ МНОГОКАНАЛЬНОЙ СИСТЕМЫ

Обозначим через $N(F)(z)$ и $D(F)(z)$ соответственно числитель и знаменатель в выражении для некоторой функции $F(z)$. Рассмотрим производящие функции

$$\pi(z) = \sum_{i=0}^{\infty} \pi_i z^i; \quad A(z) = \sum_{i=1}^{\infty} a_i z^i.$$

Теорема 3. *Производящая функция $\pi(z)$ определяется в виде*

$$\begin{aligned} \pi(z) &= \frac{N(\pi)(z)}{D(\pi)(z)}; \quad D(\pi)(z) = n\mu(1-z) - \lambda\beta z(1-A(z)); \\ N(\pi)(z) &= \mu(1-z) \left(n \left(\beta_0 + \sum_{i=1}^h \pi_i z^i - \beta \sum_{i=n}^h \frac{\pi_i z^i}{\beta_i} \right) - \beta \sum_{i=1}^{n-1} \frac{i\pi_i z^i}{\beta_i} \right). \end{aligned} \tag{8}$$

Доказательство. Умножая i -е уравнение системы (4) на z^i ($i \geq 0$) и суммируя, с учётом равенства

$$\sum_{i=n}^{\infty} \nu_i \pi_i z^i = \mu \left(\sum_{i=n}^h \frac{\pi_i z^i}{\beta_i} + \frac{1}{\beta} \left(\pi(z) - \sum_{i=0}^h \pi_i z^i \right) \right)$$

получаем

$$\begin{aligned} & - \lambda(1-A(z))\pi(z) - \frac{n\mu}{\beta} \left(1 - \frac{1}{z} \right) \pi(z) \\ & - n\mu \left(1 - \frac{1}{z} \right) \left(\sum_{i=n}^h \frac{\pi_i z^i}{\beta_i} - \frac{1}{\beta} \sum_{i=0}^h \pi_i z^i \right) - \left(1 - \frac{1}{z} \right) \sum_{i=1}^{n-1} i\nu_i \pi_i z^i = 0. \end{aligned}$$

Отсюда находим $\pi(z)$ в виде (8). Теорема доказана. \square

Приступим к вычислению производной функции $\pi(z)$.

Лемма 1. *Производная функции $\pi(z)$ при $z = 1$ вычисляется по формуле:*

$$\begin{aligned} \pi'(1) &= \frac{1}{2(n-\rho)^2} \left(2(n-\rho) \left(n \sum_{i=1}^h i\pi_i - n\beta \sum_{i=n}^h \frac{i\pi_i}{\beta_i} - \beta \sum_{i=1}^{n-1} \frac{i^2\pi_i}{\beta_i} \right) \right. \\ & \left. + \alpha\beta(a_{(1)} + a_{(2)}) \left(n \sum_{i=0}^h \pi_i - n\beta \sum_{i=n}^h \frac{\pi_i}{\beta_i} - \beta \sum_{i=1}^{n-1} \frac{i\pi_i}{\beta_i} \right) \right). \end{aligned} \tag{9}$$

Доказательство. При вычислении производной $\pi'(1)$ будем учитывать равенства

$$A'(1) = \sum_{k=1}^{\infty} k a_k = a_{(1)}; \quad A''(1) = \sum_{k=2}^{\infty} k(k-1)a_k = a_{(2)} - a_{(1)}.$$

Поскольку $A(1) = \sum_{k=1}^{\infty} a_k = 1$, то для функции $\pi(z)$, определяемой в виде (8), выполняются соотношения

$$N(\pi)(1) = D(\pi)(1) = N(\pi')(1) = D(\pi')(1) = N'(\pi')(1) = D'(\pi')(1) = 0; \quad D''(\pi')(1) \neq 0. \quad (10)$$

Из (10) следует равенство

$$\pi'(1) = \frac{N''(\pi')(1)}{D''(\pi')(1)}. \quad (11)$$

Поскольку

$$\begin{aligned} N(\pi')(z) &= N'(\pi)(z) \cdot D(\pi)(z) - N(\pi)(z) \cdot D'(\pi)(z); \\ N'(\pi')(z) &= N''(\pi)(z) \cdot D(\pi)(z) - N(\pi)(z) \cdot D''(\pi)(z); \\ N''(\pi')(z) &= N'''(\pi)(z) \cdot D(\pi)(z) + N''(\pi)(z) \cdot D'(\pi)(z) \\ &\quad - N'(\pi)(z) \cdot D''(\pi)(z) - N(\pi)(z) \cdot D'''(\pi)(z), \end{aligned}$$

то с учётом (10) получим формулу

$$N''(\pi')(1) = N''(\pi)(1) \cdot D'(\pi)(1) - N'(\pi)(1) \cdot D''(\pi)(1),$$

пользуясь которой с помощью (11) приходим к равенству (9). Лемма доказана. \square

Используя стационарное распределение числа заявок и производную $\pi'(1)$, можно получить простые формулы для основных стационарных характеристик системы.

Учитывая, что

$$\pi'(1) = \sum_{k=1}^{\infty} k\pi_k = \sum_{k=1}^h k\pi_k + \frac{\beta}{p_0} \sum_{k=h+1}^{\infty} kp_k, \quad \sum_{k=h+1}^{\infty} kp_k = \frac{p_0}{\beta} \left(\pi'(1) - \sum_{k=1}^h k\pi_k \right),$$

формулы для стационарной средней длины очереди в системе и стационарного среднего числа заявок (пакетов) в системе

$$\mathbf{E}(Q) = \sum_{k=n+1}^{\infty} (k-n)p_k, \quad \mathbf{E}(N) = \sum_{k=1}^{\infty} kp_k$$

приводим к виду

$$\begin{aligned} \mathbf{E}(Q) &= \sum_{k=n+1}^h kp_k - n \left(1 - \sum_{k=0}^n p_k \right) + \frac{p_0}{\beta} \left(\pi'(1) - \sum_{k=1}^h k\pi_k \right), \\ \mathbf{E}(N) &= \sum_{k=1}^h kp_k + \frac{p_0}{\beta} \left(\pi'(1) - \sum_{k=1}^h k\pi_k \right). \end{aligned}$$

Формула для стационарного среднего числа обслуженных заявок (пакетов) за единицу времени после использования условия нормировки принимает вид

$$\bar{N}_{sv} = \mu \left(\sum_{k=1}^{n-1} kp_k + n \sum_{k=n}^{\infty} p_k \right) = \mu \left(n(1-p_0) - \sum_{k=1}^{n-1} (n-k)p_k \right).$$

Стационарная вероятность обслуживания заявки (пакета) \mathbf{P}_{sv} и стационарное среднее время ожидания начала обслуживания $\mathbf{E}(W)$ определяются в виде

$$\mathbf{P}_{sv} = \frac{\bar{N}_{sv}}{\lambda a_{(1)}} = \frac{n(1-p_0) - \sum_{k=1}^{n-1} (n-k)p_k}{\alpha a_{(1)}}, \quad \mathbf{E}(W) = \frac{\mathbf{E}(Q)}{\bar{N}_{sv}}.$$

7. СЛУЧАЙ ОРДИНАРНОГО ВХОДЯЩЕГО ПОТОКА

Поскольку для ординарного потока заявок выполняются равенства

$$a_1 = 1; \quad a_k = 0 \quad (k \geq 2); \quad A_k = 0 \quad (k \geq 1),$$

то рекуррентные соотношения (7) упрощаются, что позволяет получить выражения для π_k ($k \geq 0$) в явном виде

$$\begin{aligned} \pi_k &= \frac{\alpha^k}{k!} \prod_{i=0}^k \beta_i, \quad 0 \leq k \leq n; & \pi_k &= \frac{\alpha^k}{n!n^{k-n}} \prod_{i=0}^k \beta_i, \quad n+1 \leq k \leq h; \\ \pi_k &= \frac{\alpha^k \beta^{k-h}}{n!n^{k-n}} \prod_{i=0}^h \beta_i, \quad k \geq h+1. \end{aligned}$$

Стационарное распределение числа заявок в системе существует при выполнении условия $\alpha\beta < n$ и определяется равенствами

$$p_0 = \left(1 + \sum_{k=1}^n \frac{\alpha^k}{k!} \prod_{i=0}^{k-1} \beta_i + \frac{1}{n!} \sum_{k=n+1}^h \frac{\alpha^k}{n^{k-n}} \prod_{i=0}^{k-1} \beta_i + \frac{\alpha^{h+1}}{n!n^{h-n}(n-\alpha\beta)} \prod_{i=0}^h \beta_i \right)^{-1}; \quad p_k = \frac{p_0 \pi_k}{\beta_k}, \quad k \geq 1.$$

Для стационарных характеристик системы получаем формулы

$$\begin{aligned} \mathbf{E}(Q) &= \sum_{k=n+1}^h (k-n)p_k + \frac{p_0 \alpha^{h+1} (n + (h-n)(n-\alpha\beta))}{n!n^{h-n}(n-\alpha\beta)^2} \prod_{i=0}^h \beta_i; & \mathbf{E}(W) &= \frac{\mathbf{E}(Q)}{\lambda P_{sv}}; \\ P_{sv} &= \frac{1}{\alpha} \left(\sum_{k=1}^{n-1} k p_k + n \sum_{k=n}^h p_k + \frac{p_0 \alpha^{h+1}}{n!n^{h-n}(n-\alpha\beta)} \prod_{i=0}^h \beta_i \right). \end{aligned}$$

8. ПРИМЕРЫ (МНОГОКАНАЛЬНАЯ СИСТЕМА БЕЗ ОГРАНИЧЕНИЙ НА ДЛИНУ ОЧЕРЕДИ)

Приведём результаты расчётов стационарных характеристик многоканальной системы, выполненных с помощью аналитических соотношений, полученных в п. 6.

Вычисления проводились для трёхканальной системы обслуживания ($n = 3$), в которую заявки (пакеты информации одинаковых размеров) прибывают группами численностью от одной до трёх с распределением числа заявок в группе, задаваемым вероятностями: $a_1 = 0, 2$; $a_2 = 0, 3$; $a_3 = 0, 5$; $a_k = 0$ ($k \geq 4$). Значения параметров показательных распределений интервалов времени между моментами поступления групп заявок и времени распределения соответственно равны: $\lambda = 3$; $\mu = 1$.

Для заданных значений параметров условие $\alpha a_{(1)} < n$ не выполняется, поэтому стационарное распределение числа заявок в системе не существует. Существование стационарного распределения обеспечим с помощью вероятностного отбрасывания заявок.

В качестве алгоритма управления очередью применим общую функцию отбрасывания групп заявок с пороговым значением $h = 5$. Рассмотрим 4 случая с разными наборами вероятностей β_i :

- $\beta_i = 1$ ($0 \leq i \leq 3$); $\beta_4 = 0, 8$; $\beta_5 = 0, 4$; $\beta_i = \beta = 0, 1$ ($i \geq 6$) (случай I);
- $\beta_i = 1$ ($0 \leq i \leq 3$); $\beta_4 = 0, 8$; $\beta_5 = 0, 4$; $\beta_i = \beta = 0, 2$ ($i \geq 6$) (случай II);
- $\beta_i = 1$ ($0 \leq i \leq 3$); $\beta_4 = 0, 8$; $\beta_5 = 0, 4$; $\beta_i = \beta = 0, 3$ ($i \geq 6$) (случай III);

$\beta_i = 1$ ($0 \leq i \leq 3$); $\beta_4 = 0,8$; $\beta_5 = 0,4$; $\beta_i = \beta = 0,4$ ($i \geq 6$) (случай IV).

Стационарные вероятности p_k ($0 \leq k \leq 8$) наличия в системе k заявок и стационарные характеристики \mathbf{P}_{sv} , $\mathbf{E}(Q)$ и $\mathbf{E}(W)$, вычисленные для различных наборов β_i , приведены в табл. 3. Анализируя полученные результаты, видим, что увеличение вероятности β в 4 раза (если сравнивать случаи IV и I) приводит к незначительному увеличению вероятности обслуживания (на 3,9%) и в то же время к более ощутимому ухудшению характеристик очереди: $\mathbf{E}(Q)$ увеличивается в 8,2 раза, а $\mathbf{E}(W)$ — в 7,9 раза.

В табл. 4 приведено сравнение стационарных характеристик системы для случая II, вычисленных по формулам, предложенным в п. 6, с результатами имитационного моделирования, полученными с помощью системы GPSS World (время моделирования $t = 10^6$) для различных распределений времени обслуживания: показательного распределения с параметром $\mu = 1$, равномерного распределения на отрезке $[0; 2]$, равномерного распределения на отрезке $[0,5; 1,5]$ и детерминированного значения, равного 1. Анализ данных таблицы показывает, что для показательного распределения времени обслуживания полученные аналитическим методом значения подтверждаются результатами имитационного моделирования. Уменьшение дисперсии распределения времени обслуживания приводит к незначительному улучшению показателей системы обслуживания: к увеличению вероятности обслуживания на 3,8% и уменьшению $\mathbf{E}(Q)$ и $\mathbf{E}(W)$ соответственно на 9,9% и на 13,2% (если сравнивать детерминированное и показательное распределения времени обслуживания). Программа для GPSS World приведена в Приложении.

Таблица 3. Стационарные характеристики трёхканальной системы для различных наборов β_i

№ набора β_i	p_0	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	\mathbf{P}_{sv}	$\mathbf{E}(Q)$	$\mathbf{E}(W)$
I	0,0095	0,0285	0,0541	0,0816	0,1390	0,2035	0,2112	0,1419	0,0718	0,4146	2,5179	0,8802
II	0,0079	0,0236	0,0448	0,0676	0,1153	0,1687	0,1751	0,1351	0,0888	0,4180	3,3481	1,1608
III	0,0055	0,0166	0,0315	0,0475	0,0809	0,1185	0,1229	0,1071	0,0853	0,4230	5,5720	1,9090
IV	0,0018	0,0055	0,0104	0,0158	0,0269	0,0393	0,0408	0,0397	0,0368	0,4309	20,7619	6,9833

Таблица 3. Сравнение стационарных характеристик трёхканальной системы для случая II с результатами имитационного моделирования для различных распределений времени обслуживания

Вид распределения	p_0	p_1	p_2	p_3	p_4	p_5	\mathbf{P}_{sv}	$\mathbf{E}(Q)$	$\mathbf{E}(W)$
показательное (аналит. метод)	0,0079	0,0236	0,0448	0,0676	0,1153	0,1687	0,4180	3,3481	1,1608
показательное (GPSS World)	0,0080	0,0234	0,0447	0,0679	0,1154	0,1688	0,4180	3,3470	1,1610
равномерн. на $[0; 2]$ (GPSS World)	0,0004	0,0041	0,0183	0,0538	0,1282	0,2105	0,4310	3,1530	1,0600
равномерн. на $[0,5; 1,5]$ (GPSS World)	0,0001	0,0012	0,0114	0,0519	0,1365	0,2259	0,4330	3,0630	1,0250
детерминированное=1 (GPSS World)	0,0002	0,0063	0,0475	0,1448	0,2369	0,2323	0,4340	3,0160	1,0080

9. ЗАКЛЮЧЕНИЕ

Таким образом, в настоящей статье изучены системы обслуживания, моделирующие работу узлов сетей передачи пакетов данных с использованием алгоритмов вероятностного отбрасывания пакетов. Приведено сравнение результатов тестовых расчётов, выполненных с помощью полученных аналитических соотношений для случая показательного распределения времени обслуживания, с результатами имитационного моделирования для различных распределений времени обслуживания. Незначительное расхождение в значениях стационарных характеристик для различных распределений времени обслуживания свидетельствует о возможности

использования полученных аналитических соотношений для проведения оценочных расчётов при разработке оптимальных алгоритмов активного управления очередью.

10. ПРИЛОЖЕНИЕ. ПРОГРАММА ДЛЯ GPSS WORLD

```

Lam EQU 3 ; значение  $\lambda$ 
Myu EQU 1 ; значение  $\mu$ 
TIME EQU 1000000 ; время моделирования
DISTR TABLE (Q$QUE+S$SYS) 0,1,1500 ; параметры распределения числа заявок
Prob VARIABLE N$LB2/(2.3#N$LB0) ; вероятность обслуживания
SYS STORAGE 3 ; трёхканальная система
GENERATE 1
TABULATE DISTR ; вычислить распределение числа заявок
TERMINATE
GENERATE (Exponential(5,0,(1/Lam))) ; поток заявок
Lb0 TEST E (S$SYS+Q$QUE),4,Lt5 ; проверка условия для задания  $\beta_4$ 
TRANSFER 800,Lbg ; задание  $\beta_4$ 
TRANSFER ,OUT
Lt5 TEST E (S$SYS+Q$QUE),5,Lt6 ; проверка условия для задания  $\beta_5$ 
TRANSFER 400,Lbg ; задание  $\beta_5$ 
TRANSFER ,OUT
Lt6 TEST GE (S$SYS+Q$QUE),6,Lbg ; проверка условия для задания  $\beta$ 
TRANSFER 200,Lbg ; задание  $\beta$ 
TRANSFER ,OUT
Lbg TRANSFER 200,LQ ; задание  $a_1$ 
TRANSFER 375,L2 ; задание  $a_2$ 
SPLIT 3,LQ ; задание  $a_3$ 
TRANSFER ,OUT
L2 SPLIT 2,LQ
TRANSFER ,OUT
LQ QUEUE QUE
ENTER SYS
DEPART QUE
ADVANCE (Exponential(5,0,(1/Myu))) ; распределение времени обслуживания
Lb2 LEAVE SYS
OUT TERMINATE
GENERATE TIME
SAVEVALUE Prob,V$Prob
TERMINATE 1
START 1

```

СПИСОК ЛИТЕРАТУРЫ

1. Chydziński A. *Nowe Modele Kolejowe Dla Węzłów Sieci Pakietowych*. Gliwice: Pracownia Komputerowa Jacka Skalmierskiego, 2013.
2. Floyd S., Jacobson V. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1993, vol. 1, № 4, pp. 397–413.
3. Bohacek S., Shah K., Arce G. R., Davis M. Signal processing challenges in active queue management. *IEEE Signal Processing Magazine Transactions on Networking*, 2004, vol. 21, № 5, pp. 69–79.
4. Bonald T., May M., Bolot J.-C. Analytic evaluation of RED performance. *Proceedings of INFOCOM*, 2000, vol. 3, pp. 1415–1424.
5. Tikhonenko O., Kempa W. M. The generalization of AQM algorithms for queueing system with bounded capacity. *Lecture Notes in Computer Sciences*, 2012, vol. 7204, pp. 242–251.
6. Tikhonenko O., Kempa W. M. Queue-size distribution in M/G/1-type system with bounded capacity and packet dropping. *Communications in Computer and Information Science*, 2013, vol. 356, pp. 177–186.
7. Kempa W. M. A direct approach to transient queue-size distribution in a finite-buffer queue with AQM. *Applied Mathematics and Information Sciences*, 2013, vol. 7, № 3, pp. 909–915.
8. Боев В. Д. *Моделирование систем. Инструментальные средства GPSS World*. Санкт-Петербург: БХВ-Петербург, 2004.
9. Бочаров П. П., Печинкин А. В. *Теория массового обслуживания*. М.: РУДН, 1995.

The simplest models of queue management in the nodes of packet-switched networks

Zhernovyi Yu. V.

We consider queueing systems modeling the work of the nodes packet data networks using probabilistic algorithms of packet discarding. This is a single-channel system with limited buffer and multi-channel system with unlimited queue. We simulate an input flow using a generalized Poisson process. The service time is exponentially distributed. An algorithm for finding of the stationary distribution of the number of customers and stationary characteristics of the system (the mean queue length, the mean waiting time in the queue, the probability of customer service) is proposed. The obtained results are verified with the help of simulation models constructed with the assistance of GPSS World tools.

KEYWORDS: queueing systems, batch arrival of customers, active queue management, probabilistic discarding of customers, stationary characteristics.