УДК 510.676, 519.217

# УСЛОВИЯ ЭРГОДИЧНОСТИ СМО С ОТНОСИТЕЛЬНЫМ ПРИОРИТЕТОМ[1]

**Мистрюков А.В.***, **Ушаков В.Г.***,**
*МГУ им. М.В. Ломоносова, г. Москва
**Федеральный исследовательский центр «Информатика и управление» Российской академии наук, г. Москва

*Поступила в редакцию 12.12.2018, после переработки 11.01.2019.*

Известные результаты по эргодичности приоритетных систем массового обслуживания получены в предположении, что входящие потоки требований всех приоритетов являются пуассоновскими. В данной работе это требование ослаблено, а именно: найдены достаточные условия эргодичности систем массового обслуживания с двумя классами приоритетов, в которых поток требований высшего приоритета является гиперэкспоненциальным, а низшего — рекуррентным. Исследована система с относительным приоритетом. Для получения искомых условий для последовательных времен ожидания в очереди требований каждого приоритета получены рекуррентные соотношения, известные как рекурсия Линдли. Опираясь на результаты, известные для гиперэкспоненциального распределения, полученная цепь Маркова исследуется методом пробных функций.

**Ключевые слова:** относительный приоритет, гиперэкспоненциальный входящий поток, эргодичность, метод пробных функций, рекурсия Линдли.

*Вестник ТвГУ. Серия: Прикладная математика. 2019. № 1. С. 5–14.*
https://doi.org/10.26456/vtpmk523

## 1. Введение

Проблема нахождения условий эргодичности традиционна для теории массового обслуживания. Эти условия важны для приложений, поскольку они определяют соотношения между параметрами модели, при выполнении которых не возникает бесконечно больших очередей. Существует обширная литература по эргодической теории случайных процессов (см., например, [3, 4]). Среди них особое место занимают марковские процессы. Большая часть условий эргодичности марковских процессов формулируется в терминах свойств переходной функции. Для

[1]Работа выполнена при финансовой поддержке РФФИ (проект №18-07-00678).

теории массового обслуживания, однако, нужны условия, выраженные через параметры исследуемой системы (входящие потоки, длительности обслуживания и т.п.), получение которых в качестве следствия из общих результатов (особенно для сложных, в частности, приоритетных, систем) является нетривиальной задачей. При изучении эргодичности приоритетных систем обычно накладывались ограничения либо на время между поступлениями требований в систему, либо на время их обслуживания. А именно, предполагалось, что эти времена имеют экспоненциальное распределение (см., например, [5]). Работ, в которых исследуется эргодичность приоритетных систем, не относящихся к классам M/G/1 и G/M/1, практически нет. При исследовании различных систем массового обслуживания часто удается найти случайный процесс $w_n$ с дискретным временем, характеризующий работу системы (длина очереди, время ожидания, число обслуженных требований и т. п.), который удовлетворяет рекуррентным соотношениям вида $w_{n+1} = \max(0; w_n + s_n - t_n)$. В теории массового обслуживания такие соотношения принято называть рекурсией Линдли. В статье [1] получены результаты, для случая, если приоритетный входящий поток является пуассоновским. В данной статье требование на приоритетный входящий поток ослаблены и получены достаточные условия эргодичности систем обслуживания с относительным приоритетом при условии, что приоритетный входящий поток является гиперэкспоненциальным.

## 2. Определения и вспомогательные утверждения

Пусть на измеримом пространстве $(X, \sigma(\mathrm{X}))$ задана марковская цепь $x(n),\ n \in N_0 = \{0, 1, 2...\}$ с переходными вероятностями $\mathbf{P}(x, A) = \mathbf{P}(x_{n+1} \in A | x_n = x),\ x \in X,\ A \in \sigma(X)$.

**Определение 1.** *Инвариантная мера марковской цепи с переходными вероятностями $P(x, A)$ есть любая вероятностная мера $\pi$ на $X$, удовлетворяющая при любом $A \in \sigma(X)$ равенству*

$$\pi(A) = \int\limits_X \pi(dx) P(x, A).$$

**Определение 2.** *Будем говорить, что марковская цепь эргодична, если она имеет единственную инвариантную меру. В этом случае марковская цепь будет подчиняться эргодической теореме Биркгофа — Хинчина [3].*

Для того, чтобы сформулировать условия эргодичности, введем следующие обозначения:

$$Pf(x) = \mathbf{E}\left(f(x_{n+1}) | x_n = x\right) = \int\limits_X f(y) P(x, dy),$$

где $f$ — функция, определенная на $X$, а $\tau_A(x) = \inf\limits_{n \geqslant 1}(n : x(n) \in A | x(0) = x), A \in \sigma(X)$ — время первого попадания в множество A из точки x,

$$\mathbf{E}_x(\tau_A) = \mathbf{E}\,(\inf_{n \geqslant 1}(n\ :\ x(n) \in A | x(0) = x))$$

— среднее время первого попадания в множество A из точки x. Сформулируем далее теоремы, дающие достаточные условия эргодичности.

**Теорема 1.** *I. Пусть марковская цепь удовлетворяет следующим условиям: существуют $A \subset X$, $p > 0$ и вероятностная мера $\nu$ на $X$ такие, что:*

*a)* $\mathbf{P}(\tau_A(x) < \infty) = 1$ *для любого* $x \in A^c$;

*b)* $\sup\limits_{x \in A} \mathbf{E}_x(\tau_A) < \infty$;

*c) существует* $m \in N$ *такое, что* $P_m(x, B) \geqslant p\nu(B)$ *для любых* $B \in \sigma(X)$ *и* $x \in A$.

*Тогда она имеет единственную инвариантную меру* $\pi$.

*II. Если к тому же марковская цепь апериодична, то переходные вероятности сходятся к инвариантной мере по полной вариации:* $||P^n(x, \cdot) - \pi(\cdot)||_{TV} \to 0$ *для любого* $x \in X$.

Более подробно см. [1]. Марковскую цепь, удовлетворяющую условиям теоремы 1, называют Харрисовой.

**Теорема 2** (Критерий Фостера-Ляпунова)**.** *Пусть существуют* $w(x) : X \to R^+$ *и* $\varepsilon > 0$ *такие, что*

*a)* $Pw(x) \leqslant w(x) - \varepsilon$ *для любого* $x \in A^c$ , $A \subset X$;

*b)* $\sup\limits_{x \in A} Pw(x) < \infty$ *для любого* $x \in A$.

*Тогда* $\sup\limits_{x \in A} \mathbf{E}_x(\tau_A) < \infty$.

## 3. Результаты

### *3.1 Описание системы и обозначения*

Рассматривается одноканальная система массового обслуживания с неограниченным числом мест для ожидания и двумя потоками требований. Первый поток — гиперэкспоненциальный, второй — рекуррентный с абсолютно непрерывной функцией распределения интервалов между поступлениями требований. Требования первого потока обладают относительным приоритетом перед требованиями второго потока.

При гиперэкспоненциальном входящем потоке время между поступлениями требований приоритетной очереди описывается гиперэкспоненциальным распределением с параметрами $(c_j, a_j)$, $j = 1, \ldots, N$:

$$a(t) = \begin{cases} \sum\limits_{j=1}^{N} c_j \, a_j \, \exp(-a_j t), & t \geqslant 0, \\ 0, & t < 0, \end{cases} \tag{1}$$

где

$$c_i > 0, \; i = 1, \ldots, N, \;\; \sum_{j=1}^{N} c_j = 1.$$

Рекуррентный входящий поток, задаваемый плотностью распределения (1), эквивалентен следующему: интервалы между поступлениями требований независимы в совокупности и показательно распределены со случайным параметром $a$, принимающим значения $a_j$ с вероятностью $c_j$, $j = 1, \ldots, N$.

Обозначим через $s_1^{(i)}, s_2^{(i)}, \ldots$ и $t_1^{(i)}, t_2^{(i)}, \ldots$ последовательные времена обслуживания и интервалы между поступлениями требований i-го потока, $s^{(i)} = \mathbf{E}\, s_1^{(i)}$, $t^{(i)} = \mathbf{E}\, t_1^{(i)}$. Пусть, далее, $w_n^{(2)}$ — время ожидания до начала обслуживания n-м требованием второго потока (нумерация требований производится для каждого потока отдельно в порядке поступления требований в систему).

### *3.2 Условия эргодичности*

Как показано в [1], для эргодичности очереди приоритетного потока достаточно, чтобы существовали первые моменты $s^{(1)}$, $t^{(1)}$, $s^{(2)}$, $t^{(2)}$ и $\mathbf{E}\,(s^{(1)} - t^{(1)}) < 0$. Для получения условий эргодичности для очереди неприоритетного потока рассмотрим двумерную марковскую цепь $(\nu,\ w_n^{(2)})$, где последовательность $w_n^{(2)}$ задается рекурсией Линдли

$$w_{n+1}^{(2)} = I(w_n^{(2)} + s_n^{(2)} + T(s_n^{(2)}, \nu) - t_{n+1}^{(2)} > 0)(w_n^{(2)} + s_n^{(2)} + T(s_n^{(2)}, \nu) - t_{n+1}^{(2)}) + \\ I(w_n^{(2)} + s_n^{(2)} + T(s_n^{(2)}, \nu) - t_{n+1}^{(2)} < 0) \cdot T^*(t_{n+1}, w_n^{(2)}, T(s_n^{(2)}, \nu)), \quad (2)$$

а $\nu$ — компонента гиперэкспоненциального распределения приоритетных требований, активная на момент поступления $n$-го требования второго потока на прибор, $T(s_n^{(2)}, \nu)$ — интервал времени с момента поступления на обслуживание $n$-го требования второго приоритета до первого после этого момента освобождения системы от этого требования и требований более высокого приоритета, $I(A)$ — индикатор события A, а случайная величина $T^*(t_{n+1}, w_n^{(2)}, T(s_n^{(2)}, \nu))$ имеет такое же распределение, как время до первого освобождения системы после момента времени $w_n^{(2)} + s_n^{(2)} + T(s_n^{(2)}, \nu) - t_{n+1}^{(2)}$, если в систему поступает только поток приоритетных требований и в начальный момент система свободна от них. Если второй момент $\mathbf{E}\,((s_n^{(2)})^2)$ существует и конечен, то для получения условий эргодичности, как показано в [1], достаточно получить условия, при которых

$$s^{(2)} + \mathbf{E}\,(T(s_n^{(2)}, \nu)) - t^{(2)} < 0.$$

Таким образом, нужно оценить $\mathbf{E}\,(T(s_n^{(2)}, \nu))$. Для этого нужно оценить сколько требований придет за $s_n^{(2)}$, при условии, что на момент начала обслуживания активна компонента $\nu$, а также оценить $\mathbf{E}\,(\Pi(n, \nu))$ — математическое ожидание периода занятости приоритетной очереди, при условии, что он начинается с $n$ требований.

**Лемма 1.** *Пусть* $s^{(1)} - t^{(1)} < 0$, *тогда*

$$\mathbf{E}\,(\Pi(n, \nu)) = n \cdot s^{(1)} \cdot (1 - \mu_1^{*'}(0)) - \sum_{l \neq \nu} \frac{\mu_1^{*'}(0)}{a_l} - \sum_{j=2}^{N} \frac{\mu_1^{*'}(0)}{\mu_j^*(0)},$$

*где*

$$\mu_1^{*'}(0) = \frac{s^{(1)}}{\left(s^{(1)} - \sum\limits_{j=1}^{N} \frac{1}{a_j} + \sum\limits_{j=1}^{N} c_j \sum\limits_{l \neq j} \frac{1}{a_l}\right)},$$

*а* $\mu_j^*(s),\ j = 1, \dots, N,\ \mu_1^*(0) = 1$, *являются решениями уравнения*

$$\frac{\prod\limits_{j=i}^{N} \mu_k^*(s) + a_i}{\sum\limits_{j=1}^{N} c_j a_j \prod\limits_{i \neq j} \mu_k^*(s) + a_i} = \beta(s - \mu_k^*(s)), \tag{3}$$

*где* $\beta(s)$ *— преобразование Лапласа-Стилтъеса времени обслуживания требования приоритетной очереди.*

*Доказательство.* Введем обозначения: $\Pi(n)$ — длительность периода занятости, начавшегося со случайного числа требований $n$, $j(t)$ — активная компонента входящего приоритетного потока:

$$\Pi_{j\nu}(n,t)dt = \mathbf{P}(\Pi(n) \in (t+dt), j(t) = \nu \mid j(0) = j)$$

$$\pi_{j\nu}(n,s) = \int e^{-st}\Pi_{j\nu}(n,t)dt.$$

Тогда, используя результаты из [2], имеем:

$$\pi_\nu(n,s) = \sum_{j=1}^{N} \pi_{j\nu}(n,s) = \beta(s - \mu_1^*(s))^n \cdot \prod_{l \neq \nu} \frac{\mu_1^*(s) + a_l}{a_l} \cdot \prod_{j=2}^{N} \frac{\mu_j^*(s)}{\mu_j^*(s) - \mu_1^*(s)}, \tag{4}$$

где $\mu_k^*(s)$ удовлетворяют уравнению (3).

Из (4) находим

$$\mathbf{E}\left(\Pi(n,\nu)\right) = -\pi_\nu^{'}(n,0) = -\left(n \cdot \beta^{'}(0) \cdot (1 - \mu_1^{*'}(0)) + \sum_{l \neq \nu} \frac{\mu_1^{*'}(0)}{a_l} + \sum_{j=2}^{N} \frac{\mu_1^{*'}(0)}{\mu_j^*(0)}\right) =$$
$$= n \cdot s^{(1)} \cdot (1 - \mu_1^{*'}(0)) - \sum_{l \neq \nu} \frac{\mu_1^{*'}(0)}{a_l} - \sum_{j=2}^{N} \frac{\mu_1^{*'}(0)}{\mu_j^*(0)}.$$

Из (3) имеем

$$\mu_1^{*'}(0) = \frac{s^{(1)}}{\left(s^{(1)} - \sum\limits_{j=1}^{N} \frac{1}{a_j} + \sum\limits_{j=1}^{N} c_j \sum\limits_{l \neq j} \frac{1}{a_l}\right)}.$$

Лемма доказана. □

Обозначим через $\mathbf{E}\left(N(s_n^{(2)}, \nu)\right)$ среднее число требований, пришедших за $s_n^{(2)}$.

**Лемма 2.** *Для любого $\nu$ справедливо неравенство*

$$\mathbf{E}\left(N(s_n^{(2)},\nu)\right) \leqslant \int \frac{\partial(\sum\limits_{i=1}^{N}\pi_i(z,s_n^{(2)}))}{\partial z}\Big|_{z=1} dF(s_n^{(2)})\ ,$$

*где*

$$\pi_i(z,t)=\sum_{j=1}^{N}\frac{c_i\cdot(C^{-1}\cdot b)_j}{c_1(\mu_j(z)+a_i)}\,e^{\mu_j(z)\cdot t},\ i=1,\ldots,N, \tag{5}$$

$$C_{ij}^{-1}=(a_j-\mu_i(z))\cdot\frac{\prod\limits_{k=1}^{N}(\mu_i(z)-a_k)}{\sum\limits_{k=1}^{N}\prod\limits_{l\neq k}(a_j-a_l)(\mu_i(z)-a_j)}\cdot\frac{\prod\limits_{k=1}^{N}(a_j-\mu_k(z))}{\sum\limits_{k=1}^{N}\prod\limits_{l\neq k}(\mu_i(z)-\mu_l(z))(a_j-\mu_i(z))}, \tag{6}$$

$$b_{\nu_{max}}=\frac{c_1}{c_{\nu_{max}}},\ b_i=0,\ i\neq\nu_{max},$$

*а $\mu_j(z)$ находятся из уравнения*

$$\prod_{i=1}^{N}(\mu+a_i)=z\cdot\sum_{i=1}^{N}c_ja_j\prod_{i\neq j}(\mu+a_i),$$

*$\nu_{max}$ — номер компоненты смеси с минимальным средним.*

*Доказательство.* Чтобы оценить сверху $E(N(s,\nu))$ предположим, что на момент начала обслуживания неприоритетного требования активна компонента $\nu_{max}$, имеющая минимальное среднее. Обозначим через $P_j(k,t)$ вероятность того, что в момент времени $t$ активна компонента с номером $j$ и в приоритетной очереди $k$ требований. Система прямых уравнений Колмогорова для $P_j(k,t)$ имеет вид:

$$\frac{\partial P_j(k,t)}{\partial t}=-a_j\cdot P_j(k,t)+(1-\delta_{k,0})\sum_{l=1}^{N}P_l(k-1,t)\cdot a_l\cdot c_j \tag{7}$$

с начальными условиями

$$P_j(k,0)=\delta_{k,0}\delta_{j,\nu_{max}},$$

где $\delta_{i,j}=1$, при $i=j$, и $=0$ при $i\neq j$. Переходя в (7) к производящим функциям, получаем

$$\frac{\partial\pi_j(z,t)}{\partial t}=-a_j\pi_j(z,t)+c_jz\sum_{l=1}^{N}a_l\pi_l(z,t) \tag{8}$$

с начальными условиями

$$\pi_j(z,0)=\delta_{j,\nu_{max}}. \tag{9}$$

Матрица этой системы имеет вид:

$$\begin{cases} b_{i,j} = c_i z a_j, & i \neq j, \\ b_{i,j} = -a_i + c_i z a_i, & i = j. \end{cases}$$

Характеристический многочлен есть

$$\prod_{i=1}^{N}(\mu + a_i) - z \cdot \sum_{i=1}^{N} c_j a_j \prod_{i \neq j}(\mu + a_i).$$

В [2] показано, что он имеет $N$ различных корней, которые обозначим $\mu_i(z)$, $i = 1, \ldots, N$. Отсюда следует, что решение (8) с начальными условиями (9) имеет вид (5), где $C_{ij}$ определяются из (6). □

Объединяя результаты лемм, получаем следующую теорему.

**Теорема 3.** *Пусть $s^{(1)}$, $t^{(1)}$, $s^{(2)}$, $t^{(2)}$ существуют, и $s^{(1)} - t^{(1)} < 0$. Пусть также существует второй момент $E((s_n^{(2)})^2)$. Тогда для эргодичности очереди второго приоритета достаточно, чтобы*

$$\int \frac{\partial(\sum\limits_{i=1}^{N} \pi_i(z, s_n^{(2)}))}{\partial z}|_{z=1} dF(s_n^{(2)}) \cdot s^{(1)} \cdot (1 - \mu_1^{*'}(0)) -$$
$$- \sum_{l \neq \nu_{max}} \frac{\mu_1^{*'}(0)}{a_l} - \sum_{j=2}^{N} \frac{\mu_1^{*'}(0)}{\mu_j^{*}(0)} + E(s^{(2)}) - E(t^{(2)}) < 0.$$

Доказательство следует из лемм 1 и 2 и критерия Фостера-Ляпунова.

## Заключение

Методы исследования эргодичности приоритетных систем массового обслуживания, развитые в статье, могут быть применены и для других типов входящих потоков (в частности, эрланговских) и приоритетных дисциплин без прерывания обслуживания.

## Список литературы

[1] Мистрюков А.В., Ушаков В.Г. Достаточные условия эргодичности приоритетных систем массового обслуживания // Информатика и ее применения. 2018. Т. 12, № 2. С. 24–26.

[2] Матвеев В.Ф., Ушаков В.Г. Системы массового обслуживания. М.: Изд-во МГУ, 1984. 240 с.

[3] Боровков А. Эргодичность и устойчивость случайных процессов. М.: Эдиториал УРСС, 1999.

[4] Meyn S., Tweedie R. Markov Chains and Stochastic Stability / eds. by B.W. Dickinson, E.D. Sontag, M. Thoma, A. Fettweis, J.L. Massey, J.W. Modestino. Series: Communications and Control Engineering. London: Springer-Verlag, 1993. https://doi.org/10.1007/978-1-4471-3267-7

[5] Гнеденко Б.В., Даниелян Э.А., Димитров Б.Н., Климов Г.П., Матвеев В.Ф. Приоритетные системы обслуживания. М.: Изд-во МГУ, 1973. 448 с.

### Образец цитирования

Мистрюков А.В., Ушаков В.Г. Условия эргодичности СМО с относительным приоритетом // Вестник ТвГУ. Серия: Прикладная математика. 2019. №1. С. 5–14. https://doi.org/10.26456/vtpmk523

### Сведения об авторах

1. **Мистрюков Андрей Вадимович**
аспирант кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М.В. Ломоносова.

*119991, г. Москва, ГСП-1, Ленинские горы, д. 1, стр. 52, ВМК МГУ.*
*E-mail: unf08@rambler.ru*

2. **Ушаков Владимир Георгиевич**
профессор кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М.В. Ломоносова; старший научный сотрудник Института проблем информатики ФИЦ ИУ РАН.

*119991, г. Москва, ГСП-1, Ленинские горы, д. 1, стр. 52, ВМК МГУ.*
*E-mail: vgushakov@mail.ru*

## SUFFICIENT ERGODICITY CONDITIONS FOR QUEUEING SYSTEMS WITH NON-PREEMPTIVE PRIORITY

**Mistryukov Andrei Vadimovich**
Post-graduate student at Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University
*Russian Federation, 119991, Moscow, GSP-1, 1-52, Leninskiye Gory, CMC MSU.*
*E-mail:* *unf08@rambler.ru*

**Ushakov Vladimir Georgievich**
Professor at Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University
Senior scientist at Institute of Computer Science Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences
*Russian Federation, 119991, Moscow, GSP-1, 1-52, Leninskiye Gory, CMC MSU.*
*E-mail:* *vgushakov@mail.ru*

*Received 12.12.2018, revised 11.01.2019.*

Known results in ergodicity of priority queues are based on the assumption, that interarrival times in each queue have exponential distribution. This paper relaxes this assumption, providing sufficient conditions for queues with two priority classes under assumption, that interarrival times in high priority class queue have hyperexponential distribution. Queues with non-preemptive priority are considered. To formulate desired conditions, we use Lindley's recursion for waiting times of each priority class queue. Using Lyapunov-Foster criteria, we obtain sufficient conditions for given recursion to be Harris-ergodic markov chain.

**Keywords:** nonpreemtive queues, hyperexponential interarrival times, ergodicity, Lyapunov-Foster criteria, Lindley recursion.

### Citation

Mistryukov A.V., Ushakov V.G., "Sufficient ergodicity conditions for queueing systems with non-preemptive priority", *Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2019, № 1, 5–14 (in Russian). https://doi.org/10.26456/vtpmk523

### References

[1] Mistryukov A.V., Ushakov V.G., "Sufficient conditions for the ergodicity of priority systems of mass service", *Informatika i ee Primeneniya*, **12**:2 (2018), 24–26 (in Russian).

[2] Matveev V.F., Ushakov V.G., *Sistemy Massovogo Obsluzhivaniya [Queuing systems]*, MSU Publ., Moscow, 1984 (in Russian), 240 pp.

[3] Borovkov A., *Ergodichnost i Ustoichivost Sluchajnykh Protsessov [Ergodicity and steadiness of random processes]*, Editorial URSS Publ., Moscow, 1999 (in Russian).

[4] Meyn S., Tweedie R., *Markov Chains and Stochastic Stability*, Communications and Control Engineering, eds. B.W. Dickinson, E.D. Sontag, M. Thoma, A. Fettweis, J.L. Massey, J.W. Modestino, Springer-Verlag, London, 1993, https://doi.org/10.1007/978-1-4471-3267-7.

[5] Gnedenko B.V., Danielyan E.A., Dimitrov B.N., Klimov G.P., Matveev V.F., *Prioritetnye sistemy obsluzhivaniya [Priority Service Systems]*, MSU Publ., Moscow, 1973 (in Russian), 448 pp.