

# Об экспоненциальной скорости сходимости распределения одной нерегенерирующей системы надёжности

**Г. А. ЗВЕРКИНА**

*Российский университет транспорта (МИИТ),  
Институт проблем управления им. В. А. Трапезникова РАН  
e-mail: zverkina@gmail.com*

УДК 519.21

**Ключевые слова:** восстанавливаемый элемент с тёплым резервом, скорость сходимости распределения, коэффициент надёжности, марковские процессы, метод склеивания.

## Аннотация

Получены оценки сверху для скорости сходимости распределения состояния восстанавливаемого элемента с тёплым резервом в случае, когда времена работы и восстановления обоих элементов могут быть зависимы и они ограничены сверху и снизу экспоненциальным распределением. Полученный результат позволяет оценивать скорость сходимости коэффициента готовности к стационарному значению.

## Abstract

*G. A. Zverkina, On the exponential convergence rate of the distribution for some nonregenerative reliability system, Fundamentalnaya i prikladnaya matematika, vol. 23 (2020), no. 1, pp. 145–160.*

Exponential upper bounds for the convergence rate of the distribution of restorable element with partially energized standby redundancy are found, in the case where all working and repair times are bounded by an exponential random variable (upper and lower), and working and repair times can be dependent. The convergence rate of the availability factor is estimated.

## 1. Введение

### 1.1. Мотивация

Как хорошо известно, математическая теория надёжности исследует поведение систем (элементов, приборов и т. п.), которые могут отказывать в случайные моменты времени (см., например, [1]). Важной частью теории надёжности является исследование поведения восстанавливаемого элемента, под которым может пониматься как отдельный прибор (устройство), так и сложная техническая система. Этот элемент безотказно работает в течение случайного времени, затем в течение случайного времени он восстанавливается и снова начинает работать.

В технических приложениях в ситуациях, когда безотказная работа исследуемого элемента очень важна, этот элемент имеет некую «страховку» в виде резервных элементов.

Это могут быть дополнительные элементы, «находящиеся на складе»: они подключаются к работе *через некоторое время* после отказа работающего элемента. В ряде случаев такие перерывы в работе недопустимы, поэтому резервные элементы могут также находиться в «горячем», «тёплом» или «холодном» резерве.

Если элементы находятся в «холодном» резерве, то они не работают и включаются в работу только после отказа основного элемента. Эта ситуация совпадает с поведением системы массового обслуживания с конечным числом мест для ожидания, и здесь такая задача рассматриваться не будет.

Если резервные элементы находятся в «горячем резерве», это означает, что они работают в полную силу (но вхолостую) и сразу после отказа основного элемента его работа переключается на один из резервных элементов — естественно, на тот, который в момент отказа основного элемента находится в рабочем состоянии. При этом обычно предполагается, что распределение времени безотказной работы и времени восстановления резервных элементов совпадает с соответствующими распределениями для основного элемента. Работа или восстановление любых элементов системы независимы. Понятно, что такое горячее резервирование излишне затратное. Поведение такой системы описывается поведением системы массового обслуживания с конечным количеством постоянно работающих приборов; здесь такая задача рассматриваться не будет.

Во многих технических системах резервные элементы находятся в «тёплом резерве», т. е. в случае исправности основного элемента они работают «не в полную силу». При этом время безотказной работы резервного элемента в холостом режиме существенно увеличивается, и в случае отказа основного элемента резервный элемент очень быстро переключается в рабочий режим. В этом случае, как правило, распределение времени безотказной работы (включая нахождение в резерве) и время восстановления резервного элемента из «тёплого» резерва отличается от соответствующих распределений основного элемента.

Важной характеристикой для восстанавливаемого элемента (или системы, состоящей из восстанавливаемых элементов) является *коэффициент надёжности*  $K_{\Gamma}(t)$  — вероятность того, что в момент  $t$  элемент (система) исправен. Для большинства восстанавливаемых систем существует предел

$$\lim_{t \rightarrow \infty} K_{\Gamma}(t) = K_{\Gamma}(\infty),$$

и важно знать, какова скорость этой сходимости.

Заметим, что

$$K_{\Gamma}(t) = \mathbb{P}\{X_t \in \mathcal{S}_0\},$$

где  $X_t$  — это состояние рассматриваемой системы, а  $\mathcal{S}_0$  — подмножество из пространства состояний системы. Поэтому для оценки скорости сходимости  $K_{\Gamma}(t)$

к стационарному значению достаточно оценить скорость сходимости распределения случайной величины  $X_t$  к стационарному распределению — в тех случаях, когда вычисление  $K_\Gamma(t)$  затруднительно или невозможно.

## 2. Восстанавливаемый элемент с одним элементом тёплого резерва: случай цепи Маркова в непрерывном времени

### 2.1. Восстанавливаемый элемент без резервирования

Для одного восстанавливаемого элемента с функцией распределения времени работы  $F(s) = 1 - e^{-\lambda s}$  и функцией распределения времени восстановления  $G(s) = 1 - e^{-\mu s}$ , как известно, в случае  $K_{\Gamma_1}(0) = 1$  (см., например, [1, § 2.3])

$$K_{\Gamma_1}(t) \stackrel{\text{def}}{=} P\{\text{в момент } t \text{ элемент исправен}\} = \frac{\mu + \lambda e^{-(\lambda+\mu)t}}{\mu + \lambda}. \quad (1)$$

Соответственно, скорость сходимости коэффициента готовности  $K_{\Gamma_1}(t)$  к его стационарному значению

$$K_{\Gamma_1}(\infty) \stackrel{\text{def}}{=} \frac{\mu}{\mu + \lambda}$$

легко находится, она экспоненциальна:

$$K_{\Gamma_1}(t) - K_{\Gamma_1}(\infty) = \frac{\lambda e^{-(\lambda+\mu)t}}{\mu + \lambda}. \quad (2)$$

### 2.2. Восстанавливаемый элемент с тёплым резервом

Теперь рассмотрим восстанавливаемый элемент в ситуации, когда время работы (и пребывания в состоянии тёплого резерва) и время восстановления резервного элемента также имеют экспоненциальное распределение. Это значит, что исправный резервный элемент имеет *интенсивность отказа*, зависящую от состояния основного восстанавливаемого элемента. А именно, первый (основной) элемент имеет функцию распределения времени безотказной работы  $F_1(s) = 1 - e^{-\lambda_1 s}$  и в случае отказа время его восстановления имеет функцию распределения  $G_1(s) = 1 - e^{-\mu_1 s}$ ; эти распределения не зависят от состояния резервного элемента. Таким образом, если первый (основной) элемент исправен в момент времени  $t$ , то

$$\mathbb{P}\{\text{основной элемент откажет за время } [t, t + \Delta]\} = \lambda_1 \Delta + o(\Delta). \quad (3)$$

Если в момент времени  $t$  основной элемент восстанавливается, то

$$\mathbb{P}\{\text{основной элемент станет исправен за время } [t, t + \Delta]\} = \mu_1 \Delta + o(\Delta). \quad (4)$$

Постоянные величины  $\lambda_1$  и  $\mu_1$  — это *постоянные интенсивности* отказов и восстановлений соответственно.

Второй (резервный) элемент имеет интенсивность отказа, зависящую от состояния основного элемента:

1) если основной элемент исправен, то интенсивность отказа резервного элемента равна  $\lambda_2$ ;

2) в противном случае интенсивность отказа резервного элемента равна  $\Lambda$ . Если резервный элемент неисправен, то интенсивность его восстановления равна  $\mu_2$  (для простоты здесь мы полагаем, что эта интенсивность не зависит от состояния основного восстанавливаемого элемента).

Предполагается, что в начальный момент времени  $t = 0$  оба элемента исправны.

Состояние описываемой системы надёжности из двух элементов в момент времени  $t$  обозначается парой  $X_t = (i_t, j_t)$ , где  $i_t, j_t \in \{0; 1\}$ ; первая компонента этой пары описывает состояние первого элемента, а вторая — второго. Исправное состояние каждого из элементов соответствует значению 0, а неисправное состояние обозначается 1. Процесс  $X_t$  — это цепь Маркова в непрерывном времени с пространством состояний  $\mathcal{X} = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$ . Коэффициент готовности этой системы — это

$$K_{\Gamma_2} \stackrel{\text{def}}{=} \mathbb{P}\{X_t \neq (1, 1)\}.$$

Обозначим

$$p_{i,j} \stackrel{\text{def}}{=} \mathbb{P}\{X_t = (i, j)\}.$$

Аналогично (3), (4), учитывая формулу полной вероятности, имеем

$$\begin{cases} p_{0,0}(t + \Delta) = p_{0,0}(1 - \lambda_1\Delta - \lambda_2\Delta) + p_{1,0}\mu_1\Delta + p_{0,1}\mu_2\Delta + o(\Delta), \\ p_{1,0}(t + \Delta) = p_{1,0}(1 - \mu_1\Delta - \Lambda\Delta) + p_{0,0}\lambda_1\Delta + p_{1,1}\mu_2\Delta + o(\Delta), \\ p_{0,1}(t + \Delta) = p_{0,1}(1 - \mu_2\Delta - \lambda_1\Delta) + p_{0,0}\lambda_2\Delta + p_{1,1}\mu_1\Delta + o(\Delta), \\ p_{1,1}(t + \Delta) = p_{1,1}(1 - \mu_1\Delta - \mu_2\Delta) + p_{1,0}\lambda_2\Delta + p_{0,1}\lambda_1\Delta + o(\Delta). \end{cases}$$

Учитывая, что  $p_{0,0} + p_{1,0} + p_{0,1} + p_{1,1} = 1$ , получаем уравнения Колмогорова:

$$\begin{cases} \dot{p}_{0,0}(t) = p_{0,0}(-\lambda_1 - \lambda_2) + p_{1,0}\mu_1 + p_{0,1}\mu_2, \\ \dot{p}_{1,0}(t) = p_{0,0}(\lambda_1 - \mu_2) + p_{1,0}(-\mu_1 - \Lambda - \mu_2) - p_{0,1}\mu_2 + \mu_2, \\ \dot{p}_{0,1}(t) = p_{0,0}(\lambda_2 - \mu_1) - p_{1,0}\mu_1 + p_{0,1}(-\mu_2 - \lambda_1 - \mu_1) + \mu_1. \end{cases} \quad (5)$$

Поскольку  $p_{0,0}(t) + p_{0,1}(t) = K_{\Gamma_1}(t)$  — коэффициент готовности основного элемента,

$$\alpha_1 \stackrel{\text{def}}{=} -(\lambda_1 + \mu_1)$$

является собственным значением системы (5) (см. (1)).

Оставшиеся два собственных значения — это решения квадратного уравнения:

$$\alpha_{2,3} = \frac{-\Lambda}{2} - \frac{\lambda_1}{2} - \frac{\lambda_2}{2} - \frac{\mu_1}{2} - \mu_2 \pm \frac{\sqrt{(\Lambda + \mu_1)^2 + (\lambda_1 + \lambda_2)^2 + 2\mu_1(\lambda_1 - \lambda_2 + 2\mu_2 - 8) - 2\Lambda(\lambda_1 + \lambda_2)}}{2}. \quad (6)$$

Заметим, что в рассмотренном случае скорость сходимости коэффициента готовности  $K_{\Gamma_2}(t) = p_{0,0}(t) + p_{0,1}(t) + p_{1,0}(t)$  к стационарному значению

$$K_{\Gamma_2}(\infty) = \frac{\mu_1(\mu_1\lambda_2 + \Lambda(\lambda_1 + \lambda_2) + 2(\mu_2^2 + \mu_1 - \mu_2))}{(\lambda_1 + \mu_1)(\mu_1(2 + \lambda_2) + \mu_2(\mu_2 + \Lambda + \lambda_1 + \lambda_2) + \Lambda(\lambda_1 + \lambda_2))} + \\ + \frac{\mu_2(\lambda_1(\lambda_1 + \lambda_2 + \mu_2) + \mu_1\mu_2(2\lambda_1 + \lambda_2 + \Lambda))}{(\lambda_1 + \mu_1)(\mu_1(2 + \lambda_2) + \mu_2(\mu_2 + \Lambda + \lambda_1 + \lambda_2) + \Lambda(\lambda_1 + \lambda_2))}$$

может быть меньше, чем в (2).

Действительно,

$$K_{\Gamma_2}(t) = K_{\Gamma_2}(\infty) + C_1e^{\alpha_1 t} + C_2e^{\alpha_2 t} + C_3e^{\alpha_3 t},$$

где  $C_1, C_2, C_3$  зависят от начального состояния системы. Если  $\Lambda, \lambda_2$  и  $\mu_2$  достаточно малы, то в соответствии с (6)

$$\alpha_2 \stackrel{\text{def}}{=} \frac{-\Lambda}{2} - \frac{\lambda_1}{2} - \frac{\lambda_2}{2} - \frac{\mu_1}{2} - \mu_2 + \frac{\sqrt{* * *}}{2} > \alpha_1,$$

и скорость сходимости  $K_{\Gamma_2}(t)$  к предельному значению не менее  $C_2e^{\alpha_2 t}$  — естественно, в том случае, когда  $C_2 \neq 0$ .

Напомним, что важной задачей теории надёжности является оценка скорости сходимости коэффициента готовности к стационарному состоянию не только в описанном выше простейшем (экспоненциальном) случае, но и в более сложных ситуациях.

В этой работе мы рассмотрим случай, когда распределения времени работы и восстановления основного и резервного элементов могут зависеть друг от друга и их распределение описывается *ограниченными сверху и снизу интенсивностями*.

Для оценки скорости сходимости коэффициента готовности к стационарному значению будет оценена скорость сходимости распределения *полного* состояния восстанавливаемой системы в метрике полной вариации. Из оценки этой скорости сходимости автоматически вытекает оценка скорости сходимости коэффициента готовности.

Для получения оценки скорости сходимости распределения полного состояния восстанавливаемой системы в метрике полной вариации будет использован метод склеивания и марковизация полумарковского процесса, описывающего поведение восстанавливаемой системы с помощью пополнения пространства состояний системы. Напомним, что вычисления будут делаться в предположении, что распределения работы и восстановления обоих элементов описываются *ограниченными сверху и снизу интенсивностями*. В этом случае будет показано, что скорость сходимости параметров изучаемой восстанавливаемой системы экспоненциальна. Такого рода результаты имеют много важных приложений (см., например, [10, 11]).

### 3. Восстанавливаемый элемент с тёплым резервом: полумарковский случай

Итак, рассматривается восстанавливаемая система, состоящая из двух элементов. Предполагается, что теперь распределения времени работы и отказа могут быть неэкспоненциальными, но они являются абсолютно непрерывными.

Под *полным состоянием системы* мы будем понимать две пары:

$$X_t \stackrel{\text{def}}{=} ((i_t, x_t), (j_t, y_t)), \quad (7)$$

где  $i_t = 0$  или  $i_t = 1$ , если первый (основной) элемент работает или нет соответственно в момент времени  $t$ . Величина  $x_t$  равна времени, прошедшему с последнего изменения состояния  $i_t$  первого (основного) элемента до момента времени  $t$ . Пара  $(j_t, y_t)$  описывает состояние резервного элемента в момент времени  $t$  точно таким же образом.

Заметим, что, как правило, поведение восстанавливаемой системы с резервированием описывается полумарковским (линейчатым) процессом (см., например, [2]). Однако использование вместо функций распределения их интенсивностей позволяет рассматривать этот процесс как марковский процесс  $X_t$  на пространстве состояний

$$\mathcal{X} \stackrel{\text{def}}{=} \{0, 1\} \times R_+ \times \{0, 1\} \times R_+.$$

#### 3.1. Описание поведения процесса с помощью интенсивностей

В значительном количестве публикаций по теории массового обслуживания и теории надёжности используется *постоянная интенсивность* поступления требований, окончания обслуживания или отказов. Постоянная интенсивность соответствует экспоненциальному распределению (времени между поступлениями требований, времени обслуживания, времени безотказной работы или восстановления). Поведение моделей теории массового обслуживания и теории надёжности с постоянными интенсивностями описывается цепью Маркова в непрерывном времени, как было показано в разделе (2.2). Тем не менее во многих прикладных задачах возникает необходимость рассматривать произвольные распределения описанных выше величин, и в этом случае используется понятие *интенсивности завершения случайного периода времени* (см., например, [1, 2]).

**Определение 1.** Пусть  $\mathcal{T}$  — это некоторый случайный период времени (с функцией распределения  $\Phi(x)$ ), начавшийся в момент времени  $t = 0$ . Пусть к моменту времени  $s > 0$  этот период  $\mathcal{T}$  не закончился. Тогда

$$\mathbb{P}\{\mathcal{T} \in (s, s + \Delta] \mid \mathcal{T} > s\} = \frac{\Phi(s + \Delta) - \Phi(s)}{1 - \Phi(s)}.$$

Если функция распределения  $\Phi(x)$  абсолютно непрерывна, то

$$\phi(s) \stackrel{\text{def}}{=} \lim_{\Delta \downarrow 0} \frac{\mathbb{P}\{T \in (s, s + \Delta] \mid T > s\}}{\Delta} = \frac{\Phi'(s+0)}{1 - \Phi(s)}.$$

Функция  $\phi(s)$  называется *интенсивностью* окончания случайного периода  $T$  в момент  $s$  при условии, что  $T \geq s$ .

**Замечание 1.** Из определения 1 следует, что

$$\mathbb{P}\{T \in (s, s + \Delta] \mid T > s\} = \phi(s)\Delta + o(\Delta).$$

**Замечание 2.** Функция распределения и плотность распределения случайной величины  $T$  могут быть определены по интенсивности окончания периода  $T$ :

$$\Phi(s) = 1 - \exp\left(-\int_0^s \phi(u) \, du\right), \quad \Phi'(s) = \phi(s) \exp\left(-\int_0^s \phi(u) \, du\right). \quad (8)$$

Легко видеть, что если функция распределения случайной величины  $T$  экспоненциальна, то интенсивность окончания этого периода постоянна, и наоборот.

Если функция распределения  $\Phi(s)$  не является непрерывной, то также можно рассматривать интенсивность для функции распределения  $\Phi(s)$ , но формулы (8) принимают другой вид, учитывающий скачки функции распределения.

**Определение 2.** Случайная величина  $\eta$  не превосходит случайную величину  $\theta$  по распределению, если для всех  $s \in \mathbb{R}$  верно неравенство

$$F_\eta(s) = \mathbb{P}\{\eta \leq s\} \geq \mathbb{P}\{\theta \leq s\} = F_\theta(s).$$

Иначе говоря, функция распределения случайной величины  $\theta$  не превосходит функции распределения случайной величины  $\eta$  по распределению. Это отношение порядка. Обозначим его через  $\eta \prec \theta$  (см. [5]).

**Замечание 3.** Если интенсивность  $\phi(s)$  окончания случайного периода  $T$  удовлетворяет неравенству

$$0 < c < \phi(s) < C < \infty,$$

то

$$\mathcal{T}_- \prec T \prec \mathcal{T}_+, \quad (9)$$

где  $\mathbb{P}\{\mathcal{T}_- \leq s\} = 1 - e^{-Cs}$  и  $\mathbb{P}\{\mathcal{T}_+ \leq s\} = 1 - e^{-cs}$ . Соответственно,  $\mathbb{E}T \leq 1/c$ .

Действительно, в соответствии с (8)

$$\mathbb{P}\{\mathcal{T}_- \leq s\} = 1 - \exp\left(-\int_0^s C \, du\right), \quad \mathbb{P}\{\mathcal{T}_+ \leq s\} = 1 - \exp\left(-\int_0^s c \, du\right),$$

откуда и следует (9).

После этих подготовительных сведений опишем поведение рассматриваемой восстанавливаемой системы на языке интенсивностей.

### 3.2. Предположения и обозначения

Предполагаем, что первый (основной) элемент имеет интенсивность отказа (в работающем режиме) и интенсивность восстановления (в неработающем режиме), зависящие от полного состояния системы:  $\lambda_1(X_t)$  и  $\mu_1(X_t)$  соответственно. Точно так же второй элемент имеет интенсивность отказа  $\lambda_2(X_t)$ , а интенсивность восстановления второго (резервного) элемента в случае отказа равна  $\mu_2(X_t)$ . Предполагаем, что при восстановлении первый (основной) элемент сразу начинает работать, а резервный элемент, даже если он был исправен, переходит в состояние тёплого резерва. Иначе говоря,

$$\left\{ \begin{array}{l} \mathbb{P}\{i_{t+\Delta} = 1, x_{t+\Delta} \in [0, \Delta), j_{t+\Delta} = j_t, y_{t+\Delta} = y_t + \Delta \mid i_t = 0\} = \\ \quad = \lambda_1(X_t)\Delta + o(\Delta), \\ \mathbb{P}\{i_{t+\Delta} = 0, x_{t+\Delta} \in [0, \Delta), j_{t+\Delta} = j_t, y_{t+\Delta} = y_t + \Delta \mid i_t = 1\} = \\ \quad = \mu_1(X_t)\Delta + o(\Delta), \\ \mathbb{P}\{j_{t+\Delta} = 0, y_{t+\Delta} \in [0, \Delta), i_{t+\Delta} = i_t, x_{t+\Delta} = x_t + \Delta \mid j_t = 0\} = \\ \quad = \lambda(X_t)\Delta + o(\Delta), \\ \mathbb{P}\{j_{t+\Delta} = 1, y_{t+\Delta} \in [0, \Delta), i_{t+\Delta} = i_t, x_{t+\Delta} = x_t + \Delta \mid j_t = 1\} = \\ \quad = \mu_2(X_t)\Delta + o(\Delta). \end{array} \right. \quad (10)$$

Соотношения (10) задают переходную функцию процесса  $X_t$ , т. е. процесс  $X_t$  является марковским процессом с пространством состояний  $\mathcal{X}$  и с переходными вероятностями (10).

Описанная здесь модель включает в себя рассмотренный ранее экспоненциальный случай, а также ситуации, когда в случае большой продолжительности работы основного элемента резервный элемент заранее начинает работать в более интенсивном режиме, чтобы быстрее включиться в работу в случае отказа основного элемента. Также работа основного элемента может быть менее интенсивной, если неисправен резервный элемент, чтобы исключить возможность одновременной неисправности обоих элементов, и пр.

**Предположения.** Положим, что

$$\begin{aligned} 0 < \lambda_m^- \leq \lambda_m(X_t) \leq \lambda_m^+ < \infty, \\ 0 < \mu_m^- \leq \mu_m(X_t) \leq \mu_m^+ < \infty. \end{aligned} \quad (11)$$

Введём дополнительные случайные величины  $\xi_m^+$ ,  $\xi_m^-$ ,  $\zeta_m^+$ ,  $\zeta_m^-$ :

$$\begin{aligned} \mathbb{P}\{\xi_m^+ \leq s\} &= 1 - \exp\left(-\int_0^s \lambda_m^+(u) du\right), & \mathbb{P}\{\xi_m^- \leq s\} &= 1 - \exp\left(-\int_0^s \lambda_m^-(u) du\right), \\ \mathbb{P}\{\zeta_m^+ \leq s\} &= 1 - \exp\left(-\int_0^s \mu_m^+(u) du\right), & \mathbb{P}\{\zeta_m^- \leq s\} &= 1 - \exp\left(-\int_0^s \mu_m^-(u) du\right). \end{aligned} \quad (12)$$

В дальнейшем мы будем рассматривать наборы таких случайных величин  $\xi_m^+(k)$ ,  $\xi_m^-(k)$ ,  $\zeta_m^+(k)$ ,  $\zeta_m^-(k)$ ,  $k \in \mathbb{N}$ , полагая эти случайные величины независимыми в совокупности.

**Замечание 4.** Обозначим через  $\xi_m(k)$   $k$ -й период работы  $m$ -го элемента ( $m = 1, 2$ ), а через  $\zeta_m(k)$  —  $k$ -й период восстановления  $m$ -го элемента. В соответствии с замечанием 3

$$\zeta_m(k) \prec \zeta_m^-, \quad \xi_m(k) \prec \xi_m^-.$$

### 3.3. Эргодичность процесса $X_t$

Обозначим через  $\mathcal{P}_t$  распределение процесса  $X_t$  в момент времени  $t$ , т. е.  $\mathcal{P}_t(A) = \mathbb{P}\{X_t \in A\}$ ,  $A \in \mathcal{B}(\mathcal{X})$ . Очевидно,  $\mathcal{P}_t$  зависит от  $X_0$ . При этом процесс  $X_t$  не является *регенерирующим*, и стандартные методы доказательства его эргодичности как регенерирующего процесса неприменимы (см. [6, 7]).

**Лемма 1.** В условиях (11) процесс  $X_t$  является эргодическим, т. е. существует инвариантная вероятностная мера  $\mathcal{P}$  на пространстве  $\mathcal{X}$ , такая что  $\mathcal{P}_t \Rightarrow \mathcal{P}$  для всех начальных состояний  $X_0$  процесса  $X_t$ .

**Доказательство.** Обозначим (для  $m = 1, 2, \dots$ )

$$\begin{aligned} \theta_1 &\stackrel{\text{def}}{=} \inf\{t > 0 : i_t = 0\}, & \theta'_1 &\stackrel{\text{def}}{=} \inf\{t > \theta_1 : i_t = 1\}, \\ \theta_{m+1} &\stackrel{\text{def}}{=} \inf\{t > \theta'_m : i_t = 0\}, & \theta'_{m+1} &\stackrel{\text{def}}{=} \inf\{t > \theta'_{m+1} : i_t = 0\}. \end{aligned}$$

$\theta_k$  — это момент  $k$ -го восстановления основного элемента.

Из замечаний 3 и 4 получаем, что

$$\mathbb{E}\theta_1 \leq \frac{1}{\lambda_1} + \frac{1}{\mu_1}, \quad \mathbb{E}(\theta_{k+1} - \theta_k) \leq \frac{1}{\lambda_1} + \frac{1}{\mu_1}. \quad (13)$$

Возьмём достаточно малое  $\varepsilon > 0$  и обозначим

$$\mathcal{S}^\varepsilon \stackrel{\text{def}}{=} \{(1, x, 1, y) : \max\{x, y\} < \varepsilon\}, \quad \mathcal{S}^\varepsilon \subset \mathcal{X}.$$

В момент времени  $\theta_k$  ( $k > 1$ ) первый (основной) восстанавливаемый элемент переходит в рабочее состояние;  $i_{\theta_k} = 0$ ,  $x_{\theta_k} = 0$ .

Относительно второго (резервного) элемента возможны два случая.

1. Если  $j_{\theta_k} = 1$ , т. е. второй элемент неисправен, то остаточное время его пребывания в неисправном состоянии по распределению меньше случайной величины  $\zeta_2^-$  и с вероятностью, большей чем

$$\varpi_1 \stackrel{\text{def}}{=} 1 - e^{-\varepsilon\mu_2^-},$$

второй элемент перейдёт в исправное состояние за время, меньшее  $\varepsilon$ , т. е. до момента времени  $\theta_k + \varepsilon$  процесс  $X_t$  окажется в множестве  $\mathcal{S}^\varepsilon$ .

2. Если  $j_{\theta_k} = 0$ , т. е. второй элемент исправен, то его остаточное время пребывания в исправном состоянии по распределению меньше случайной величины  $\xi^-$ , а последующее время восстановления по распределению меньше случайной величины  $\zeta_2^-$ . Поэтому с вероятностью, большей чем

$$\varpi_2 \stackrel{\text{def}}{=} (1 - e^{-(\varepsilon/2)\mu_2^-})(1 - e^{-(\varepsilon/2)\lambda_2^-}),$$

второй элемент откажет быстрее, чем через время  $\varepsilon/2$ , и затем восстановится быстрее, чем за время  $\varepsilon/2$ , т. е. с вероятностью, большей чем  $\varpi_2$ , к моменту времени  $\theta_k + \varepsilon$  процесс  $X_t$  окажется в множестве  $\mathcal{S}^\varepsilon$ .

Таким образом, с вероятностью, большей чем

$$\varpi \stackrel{\text{def}}{=} \min\{\varpi_1, \varpi_2\} > 0,$$

при некотором  $\vartheta_k \in [\theta_k, \theta_k + \varepsilon]$  процесс  $X_{\vartheta_k}$  окажется в маленьком подмножестве  $\mathcal{S}^\varepsilon$  пространства состояний  $\mathcal{X}$ . Значит, время между последовательными попаданиями процесса  $X_t$  в множество  $\mathcal{S}^\varepsilon$  является геометрической суммой случайных величин с конечными средними (см. (13)) и имеет конечное математическое ожидание. Поэтому процесс  $X_t$  удовлетворяет условию Дёблина—Дуба [3, гл. 6, § 2], и по теореме Дуба [3, теорема 2.1] процесс  $X_t$  эргодичен, т. е. существует единственная вероятностная мера  $\mathcal{P}$ , такая что распределение процесса  $X_t$  сходится к нему в метрике полной вариации с экспоненциальной скоростью.  $\square$

## 4. Основной результат

**Теорема 1.** *Если выполнены условия (11), то для любого начального состояния  $X_0$  процесса  $X_t$  можно вычислить числа  $\alpha > 0$  и  $\mathfrak{K} = \mathfrak{K}(\alpha, \lambda_m^-, \lambda_m^+, \mu_m^-, \mu_m^+)$ , такие что  $\|\mathcal{P}_t - \mathcal{P}\|_{TV} \leq \mathfrak{K}e^{-\alpha t}$  для всех  $t \geq 0$ . Здесь  $\|\cdot\|_{TV}$  обозначает метрику полной вариации.*

**Следствие 1.** *В условиях теоремы 1 верно неравенство*

$$|K_{\Gamma_2}(t) - K_{\Gamma_2}(\infty)| \leq \mathfrak{K}e^{-\alpha t},$$

поскольку коэффициент готовности рассматриваемой восстанавливаемой системы равен

$$K_{\Gamma_2}(t) = \mathbb{P}\left\{X_t \notin \{(1, \cdot), (1, \cdot)\}\right\} = 1 - \mathcal{P}_t((1, \cdot), (1, \cdot)).$$

Доказательство теоремы 1 — это построение алгоритма вычисления чисел  $\alpha > 0$  и  $\mathfrak{K}$ , и этот алгоритм основан на методе склеивания марковских процессов (см., например, [8, 13]). В этой работе не будет полного описания этого алгоритма, поскольку оно занимает много места и требует большого количества вычислений. Мы изложим только основные идеи доказательства.

Напомним, что «классический» метод склеивания неприменим для исследуемой модели, поскольку её поведение описывается марковским процессом

в непрерывном времени. Поэтому будет использовано понятие успешной склейки, предложенное в [9]. Также будет использовано понятие «общей части распределений», т. е.

$$\kappa \stackrel{\text{def}}{=} \int_{\mathbb{R}} \min\{f_1(u), f_2(u)\} du,$$

где  $f_m(\cdot)$  являются плотностями распределений. В доказательстве применяется основная лемма склеивания в её простейшем виде (см., например, [12, 14]). О применении метода склеивания в теории массового обслуживания см., например, [15, 16].

#### 4.1. Успешная склейка

Пусть  $X_t$  и  $\hat{X}_t$  — два независимых марковских процесса с одной и той же переходной функцией (10), но с разными начальными состояниями в момент времени  $t = 0$ .

Пусть на некотором вероятностном пространстве построены (зависимые) процессы  $Y_t = ((i_t, x_t), (j_t, y_t))$  и  $\hat{Y}_t = ((\hat{i}_t, \hat{x}_t), (\hat{j}_t, \hat{y}_t))$  таким образом, что

- 1)  $Y_t \stackrel{D}{=} X_t$  и  $\hat{Y}_t \stackrel{D}{=} \hat{X}_t$  для всех фиксированных  $t$ ;
- 2)  $\mathbb{P}\{\tau(X_0, \hat{X}_0) < \infty\} = 1$ , где  $\tau(X_0, \hat{X}_0) = \tau(Y_0, \hat{Y}_0) = \inf\{t > 0: Y_t = \hat{Y}_t\}$ .

Такая пара процессов  $Y_t = ((i_t, x_t), (j_t, y_t))$  и  $\hat{Y}_t = ((\hat{i}_t, \hat{x}_t), (\hat{j}_t, \hat{y}_t))$  называется *успешной склейкой* процессов  $X_t$  и  $\hat{X}_t$ , а  $\tau(X_0, \hat{X}_0)$  называется *моментом склеивания*.

Нашей задачей будет конструирование успешной склейки и оценка экспоненциальных моментов случайной величины  $\tau(X_0, \hat{X}_0)$ .

#### 4.2. Конструирование процессов $Y_t$ и $\hat{Y}_t$

Для конструирования процессов  $Y_t$  и  $\hat{Y}_t$  возьмём пространство

$$\prod_{k=0}^{\infty} (\Omega_k, \mathbb{P}_k, \mathcal{B}_k),$$

где  $(\Omega_k, \mathbb{P}_k, \mathcal{B}_k)$  — некоторое вероятностное пространство, на котором в дальнейшем будут строиться нужные для конструкции успешной склейки случайные величины. Конструирование будет проводиться пошагово, в те моменты времени  $t_1, t_2, \dots$ , когда меняется одна из компонент  $i, j, \hat{i}, \hat{j}$  пары  $(Y_t, \hat{Y}_t)$ .

Пусть в момент  $t_k$  известно состояние обоих процессов и, следовательно, известны интенсивности окончания пребывания обоих приборов в том состоянии, в котором они находятся, т. е. известны совместные распределения этих остаточных времён пребывания приборов в их состояниях. Значит, в соответствии с определением состояния процесса (7) в момент  $t_k$  известны величины

$i_{t_k}, j_{t_k}, \hat{i}_{t_k}, \hat{j}_{t_k}$  и совместное распределение остаточных времён случайных величин  $x_{t_k}, y_{t_k}, \hat{x}_{t_k}, \hat{y}_{t_k}$ . Это четыре случайные величины, и известно их совместное распределение.

На этом шаге используется одно из пространств  $(\Omega_k, \mathbb{P}_k, \mathcal{B}_k)$ , и на нём строятся эти четыре случайные величины — таким образом, что их совместное распределение совпадает с заданным. После этого выбирается минимальная из этих четырёх случайных величин, пусть её значение есть  $\varsigma_k$ .

В момент времени

$$t_{k+1} \stackrel{\text{def}}{=} t_k + \varsigma_k$$

происходит следующее изменение одной из компонент  $i, j, \hat{i}, \hat{j}$ .

В момент  $t_{k+1}$  процедура повторяется.

Понятно, что предложенная здесь процедура конструирования случайных процессов  $Y_t$  и  $\hat{Y}_t$  не гарантирует совпадения в какой-либо момент времени этих процессов.

### 4.3. Основная лемма склеивания

Здесь мы приводим основную лемму склеивания в наипростейшей формулировке (см., например, [4]).

**Лемма 2.** Если случайные величины  $\vartheta_1$  и  $\vartheta_2$  имеют функции распределения  $\Phi_1(s)$  и  $\Phi_2(s)$  соответственно и их общая часть

$$\kappa \stackrel{\text{def}}{=} \int_{\mathbb{R}} \min\{\Phi_1'(s), \Phi_2'(s)\} ds$$

больше 0, то на некотором вероятностном пространстве можно построить случайные величины  $\hat{\vartheta}_1$  и  $\hat{\vartheta}_2$  таким образом, что

- 1)  $\hat{\vartheta}_1 \stackrel{\mathcal{D}}{=} \vartheta_1, \hat{\vartheta}_2 \stackrel{\mathcal{D}}{=} \vartheta_2$ ;
- 2)  $\mathbb{P}\{\hat{\vartheta}_1 = \hat{\vartheta}_2\} = \kappa$ .

**Замечание 5.** Утверждение леммы 2 естественным образом переносится на любое конечное количество случайных величин. Пусть  $\vartheta_1, \vartheta_2, \dots, \vartheta_n$  — случайные величины с плотностями распределения  $\varphi_1(s), \varphi_2(s), \dots, \varphi_n(s)$  соответственно и

$$\kappa \stackrel{\text{def}}{=} \int_{\mathbb{R}} \min\{\varphi_1(s), \varphi_2(s), \dots, \varphi_n(s)\} ds > 0.$$

Тогда на некотором вероятностном пространстве можно построить случайные величины  $\hat{\vartheta}_1(s), \hat{\vartheta}_2(s), \dots, \hat{\vartheta}_n(s)$ , такие что

- 1)  $\hat{\vartheta}_i \stackrel{\mathcal{D}}{=} \vartheta_i, i = 1, 2, \dots, n$ ;
- 2)  $\mathbb{P}\{\hat{\vartheta}_1 = \hat{\vartheta}_2 = \dots = \hat{\vartheta}_n\} = \kappa$ .

Доказательство этого факта повторяет доказательство леммы 2 из [4].

#### 4.4. Склеивание процессов $Y_t$ и $\hat{Y}_t$

Сначала рассмотрим моменты  $\theta_k$  — моменты восстановления основного элемента процесса  $Y_t$ ;

$$\theta_k < \sum_{i=1}^k (\xi_1^-(i) + \zeta_1^-(i))$$

(см. лемму 1).

Аналогично тому, как это было показано в лемме 1, для некоторого фиксированного  $\varepsilon > 0$  можно утверждать следующее:

1) с вероятностью, большей чем

$$\pi_1 \stackrel{\text{def}}{=} (1 - e^{-(\varepsilon/2)\lambda_2^-})(1 - e^{-(\varepsilon/2)\mu_2^-}),$$

на интервале  $[\theta_k, \theta_k + \varepsilon]$  процесс  $Y_t$  окажется в множестве  $\mathcal{S}^\varepsilon$ ;

2) с вероятностью, большей чем

$$\begin{aligned} \pi_2 \stackrel{\text{def}}{=} & (1 - e^{-(\varepsilon/2)\lambda_1^-})(1 - e^{-(\varepsilon/2)\mu_1^-}) \times \\ & \times (1 - e^{-(\varepsilon/2)\Lambda^-})(1 - e^{-(\varepsilon/2)\lambda_2^-})(1 - e^{-(\varepsilon/2)\mu_2^-}), \end{aligned}$$

на интервале  $[\theta_k, \theta_k + \varepsilon]$  процесс  $\hat{Y}_t$  окажется в множестве  $\mathcal{S}^\varepsilon$ .

Таким образом, с вероятностью, большей чем  $\pi_1\pi_2$ , на интервале  $[\theta_k, \theta_k + \varepsilon]$  оба процесса  $Y_t$  и  $\hat{Y}_t$  окажутся в множестве  $\mathcal{S}^\varepsilon$ . В этом множестве  $\mathcal{S}^\varepsilon$  все приборы исправны и остаточное время пребывания всех приборов в исправном состоянии имеет общую часть

$$\begin{aligned} \kappa_1 \stackrel{\text{def}}{=} & \int_0^\infty \min\{\lambda_1(X_{\theta_k+\varepsilon+s})e^{-\lambda_1(X_{\theta_k+\varepsilon+s})}, \lambda_2(X_{\theta_k+\varepsilon+s})e^{-\lambda_2(X_{\theta_k+\varepsilon+s})}\} ds \geq \\ & \geq \frac{\min\{\lambda_1^-, \lambda_2^-\}}{\max\{\lambda_1^+, \lambda_2^+\}} > 0. \end{aligned}$$

Поэтому на одном вероятностном пространстве можно построить четыре случайные величины — остаточные времена пребывания всех элементов в исправном состоянии — таким образом, что они совпадут с вероятностью, не меньшей чем  $\kappa_1$ . Таким образом, оба процесса окажутся одновременно в состоянии  $(0, 0, 0, 0)$  через время  $\eta$ , которое по распределению меньше, чем случайная величина с функцией распределения  $\Phi(s) = 1 - e^{-\min\{\lambda_1^-, \lambda_2^-, \mu_1^-, \mu_2^-\}s}$ .

На самом деле можно сначала «склеить» состояния первых (основных) элементов обоих процессов, а затем — состояния резервных элементов, но описание этой процедуры существенно более сложное. Также можно рассмотреть моменты  $\theta'_k$  — моменты отказов первого (основного) элемента. Опять повторяя рассуждения леммы 1, можно оценить вероятность того, что на интервале  $[\theta'_k, \theta'_k + \varepsilon]$  все элементы обеих систем окажутся неисправны, и снова можно применить модифицированную основную лемму склеивания (лемму 2) и показать, что с вероятностью, большей чем некоторое  $\kappa_2 > 0$ , процессы совпадут.

Предложенные здесь оценки вероятностей попадания пары процессов в некоторое множество не оптимальны, они могут быть улучшены с помощью рассмотрения различных комбинаций состояний элементов обоих процессов и различных способов выбора величины  $\varepsilon$  и долей этой величины вместо  $\varepsilon/2$ .

#### 4.5. Выводы

Итак, с некоторой вероятностью  $\tilde{\kappa}$ , оценка которой была дана здесь достаточно грубо, можно утверждать, что на любом интервале  $[\theta_k, \theta_{k+1}]$  процессы  $Y_t$  и  $\hat{Y}_t$  совпадут. Момент совпадения  $\tau = \tau(X_0, \hat{X}_0)$ . Поэтому можно применить основное неравенство склеивания для построенной успешной склейки:

$$\begin{aligned} |\mathbb{P}\{X_t \in S\} - \mathbb{P}\{\hat{X}_t \in S\}| &= |\mathbb{P}\{Y_t \in S\} - \mathbb{P}\{\hat{Y}_t \in S\}| = \\ &= |\mathbb{P}\{Y_t \in S \& \tau > t\} - \mathbb{P}\{\hat{Y}_t \in S \& \tau > t\} + \\ &+ |\mathbb{P}\{Y_t \in S \& \tau < t\} - \mathbb{P}\{\hat{Y}_t \in S \& \tau < t\}| = \\ &= |\mathbb{P}\{Y_t \in S \& \tau > t\} - \mathbb{P}\{\hat{Y}_t \in S \& \tau > t\}| \leq \\ &\leq \mathbb{P}\{\tau > t\} \leq \mathbb{P}\{e^{\alpha\tau} > e^{\alpha t}\} \leq \frac{\mathbb{E} e^{\alpha\tau}}{e^{\alpha t}}, \end{aligned} \quad (14)$$

поскольку при  $\tau < t$  «склеивание» процессов  $Y_t$  и  $\hat{Y}_t$  уже произошло, их распределения в момент  $t$  одинаковы и

$$\mathbb{P}\{Y_t \in S \& \tau < t\} = \mathbb{P}\{\hat{Y}_t \in S \& \tau < t\};$$

здесь  $S \in \mathcal{B}(\mathcal{X})$ . Таким образом, осталось оценить  $\mathbb{E} e^{\beta\tau}$  и выяснить, при каких  $\beta > 0$  это математическое ожидание конечно.

Учитывая, что имеется равномерная (не зависящая от начальных условий) оценка

$$\tau(X_0, \hat{X}_0) < \theta_1 + \sum_{k=1}^{\nu} (\theta_{k+1} - \theta_k) + \eta < \eta + \sum_{k=1}^{\nu} (\xi_k^- + \zeta_k^-),$$

где

$$\theta_{k+1} - \theta_k < \xi_1^- + \zeta_1^- \stackrel{D}{=} \xi_k^- + \zeta_k^-,$$

а также  $\mathbb{P}\{\nu > n\} \leq (1 - \tilde{\kappa})^n$ , можно оценить величину  $\alpha$  и  $\mathfrak{K} = \mathbb{E} e^{\alpha\tau}$ .

Поскольку оценка (14) равномерна по начальным условиям процессов  $X_t$  и  $\hat{X}_t$ , она верна и в том случае, когда процесс  $\hat{X}_t$  имеет в качестве начального распределения стационарное распределение  $\mathcal{P}$ , откуда следует, что

$$\sup_{S \in \mathcal{B}(\mathcal{X})} |\mathbb{P}\{X_t \in S\} - \mathcal{P}(S)| = \|\mathcal{P}_t(S) - \mathcal{P}(S)\|_{TV} \leq \frac{\mathbb{E} e^{\alpha\tau}}{e^{\alpha t}} = \mathfrak{K} e^{-\alpha t}.$$

## 5. Заключение

Представленная здесь конструкция успешной склейки существенно упрощена, и оценки, которые можно получить по описанной схеме, достаточно

грубые. Однако использование предложенного подхода с анализом всех возможных ситуаций, подходящих для склеивания, а также возможность учесть некоторые специфические характеристики изучаемой восстанавливаемой системы и начальное состояние процесса  $X_t$  позволяют существенно улучшить предложенную здесь оценку.

Автор благодарит Э. Ю. Калимулину, В. В. Козлова и А. Ю. Веретенникова за ценные рекомендации при подготовке статьи. Работа выполнена при финансовой поддержке РФФИ (проект № 17-01-00633 А).

## Литература

- [1] Гнеденко Б. В., Беляев Ю. К., Соловьёв А. Д. Математические методы в теории надёжности. — М.: Наука, 1965.
- [2] Гнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания. — М.: Наука, 1966.
- [3] Дуб Дж. Л. Вероятностные процессы. — М.: Изд. иностр. лит., 1956.
- [4] Зверкина Г. А. Об одной обобщённой системе Эрланга—Севастьянова и её скорости сходимости // *Фундамент. и прикл. матем.* — 2018. — Т. 22, вып. 3. — С. 57–82.
- [5] Штойян Д. Качественные свойства и оценки стохастических моделей. — М.: Мир, 1979.
- [6] Afanasyeva L. G., Tkachenko A. V. On the convergence rate for queueing and reliability models described by regenerative processes // *J. Math. Sci.* — 2016. — Vol. 218, no. 2. — P. 119–136.
- [7] Asmussen S. *Applied Probability и Queues.* — New York: Springer, 2003.
- [8] Doeblin W. Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états // *Rev. Math. Union Interbalkan.* — 1938. — Vol. 2. — P. 77–105.
- [9] Griffeath D. A maximal coupling for Markov chains // *Z. Wahrsch. Verw. Gebiete.* — 1975 — Vol. 31, no. 2. — P. 95–106.
- [10] Kalimulina E. Rate of convergence to stationary distribution for unreliable Jackson-type queueing network with dynamic routing // *Distributed Computer and Communication Networks. 19th International Conference, DCCN 2016 / Vishnevskiy V. M., Samouylov K. E., Kozyrev D. K., eds.* — Cham: Springer, 2016. — (Commun. Comput. Inform. Sci.; Vol. 678). — P. 253–265.
- [11] Kalimulina E. Analysis of unreliable open queueing network with dynamic routing // *Distributed Computer and Communication Networks. DCCN 2017 / Vishnevskiy V. M., Samouylov K. E., Kozyrev D. K., eds.* — Cham: Springer, 2017. — (Commun. Comput. Inform. Sci.; Vol. 700). — P. 355–367.
- [12] Kato K. Coupling lemma and its application to the security analysis of quantum key distribution // *Tamagawa Univ. Quantum ICT Research Inst. Bull.* — 2014. — Vol. 4, no. 1. — P. 23–30.
- [13] Thorisson H. *Coupling, Stationarity, and Regeneration.* — Berlin: Springer, 2000.
- [14] Veretennikov A., Butkovsky O. A. On asymptotics for Vaserstein coupling of Markov chains // *Stoch. Process. Appl.* — 2013. — Vol. 123, no. 9. — P. 3518–3541.

- [15] Veretennikov A. Yu., Zverkina G. A. Simple proof of Dynkin's formula for single-server systems and polynomial convergence rates // *Markov Process. Relat. Fields.* — 2014. — Vol. 20. — P. 479–504.
- [16] Zverkina G. On strong bounds of rate of convergence for regenerative processes // *Distributed Computer and Communication Networks. 19th International Conference, DCCN 2016 / Vishnevskiy V. M., Samouylov K. E., Kozyrev D. K., eds.* — (Commun. Comput. Inform. Sci.; Vol. 678). — Cham: Springer, 2016. — P. 381–393.