

ИНФОРМАТИКА

Известия Саратовского университета. Новая серия. Серия: Математика. Механика. Информатика. 2026. Т. 26, вып. 2. С. 280–295

Izvestiya of Saratov University. Mathematics. Mechanics. Informatics, 2026, vol. 26, iss. 2, pp. 280–295

<https://mmi.sgu.ru>

DOI: <https://doi.org/10.18500/1816-9791-2026-26-2-280-295>

EDN: <https://elibrary.ru/RDHJQE>

Научная статья

УДК 519.2

Модели зависимостей между приборами в системах массового обслуживания типа split-merge. Случай экспоненциального обслуживания

А. В. Горбунова¹✉, А. В. Лебедев²

¹Институт проблем управления им. В. А. Трапезникова Российской академии наук, Россия, 117997, г. Москва, ул. Профсоюзная, д. 65, стр. 2

²Московский государственный университет имени М. В. Ломоносова, Россия, 119991, г. Москва, ГСП-1, Ленинские горы, д. 1

Горбунова Анастасия Владимировна, кандидат физико-математических наук, старший научный сотрудник, avgorbunova@list.ru, <https://orcid.org/0000-0002-9183-0426>, SPIN: 5354-0834, AuthorID: 851308

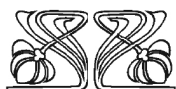
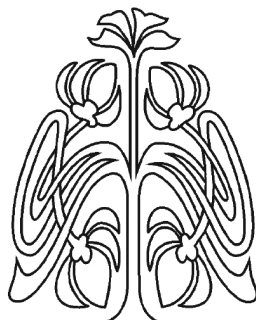
Лебедев Алексей Викторович, доктор физико-математических наук, доцент кафедры теории вероятностей, avlebed@yandex.ru, <https://orcid.org/0000-0002-9258-0588>, AuthorID: 16748

Аннотация. Рассматривается система с разделением заявок типа split-merge с двумя приборами, времена обслуживания на которых одинаково распределены и зависят в соответствии с различными копулами. Изучается влияние зависимости времен обслуживания на среднее время пребывания заявки в системе: получены выражения для математического ожидания времени отклика системы, анализируется его поведение для различных значений коэффициентов корреляции Кендалла и Бломквиста, а также получены формулы для границ среднего времени отклика в зависимости от значения коэффициента Бломквиста.

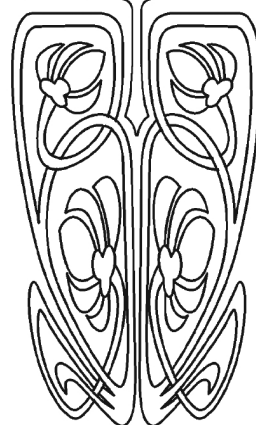
Ключевые слова: система массового обслуживания, система типа split-merge, зависимые времена обслуживания, копула, время пребывания

Для цитирования: Горбунова А. В., Лебедев А. В. Модели зависимостей между приборами в системах массового обслуживания типа split-merge. Случай экспоненциального обслуживания // Известия Саратовского университета. Новая серия. Серия: Математика. Механика. Информатика. 2026. Т. 26, вып. 2. С. 280–295. DOI: <https://doi.org/10.18500/1816-9791-2026-26-2-280-295>, EDN: RDHJQE

Статья опубликована на условиях лицензии Creative Commons Attribution 4.0 International (CC-BY 4.0)



Научный
отдел





Article

Servers dependency models in split-merge queueing systems. The case of exponential service

A. V. Gorbunova¹✉, A. V. Lebedev²

¹V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, 65 Profsoyuznaya St., Moscow 117997, Russia

²Lomonosov Moscow State University, GSP-1, Leninskie Gory, Moscow 119991, Russia

Anastasia V. Gorbunova, avgorbunova@list.ru, <https://orcid.org/0000-0002-9183-0426>, SPIN: 5354-0834, AuthorID: 851308

Alexey V. Lebedev, avlebed@yandex.ru, <https://orcid.org/0000-0002-9258-0588>, AuthorID: 16748

Abstract. A split-merge queueing system with two servers, the service times of which are identically distributed and dependent according to different copulas, is considered. The influence of service time dependence on the average sojourn time of a request in the system is studied. Expressions are derived for the expected value of the system response time, its behavior is analyzed for various values of the Kendall and Blomquist correlation coefficients, and formulas are derived for the bounds of the average response time depending on the value of the Blomqvist coefficient.

Keywords: queueing system, split-merge system, dependent service times, copula, sojourn time

For citation: Gorbunova A. V., Lebedev A. V. Servers dependency models in split-merge queueing systems. The case of exponential service. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*, 2026, vol. 26, iss. 2, pp. 280–295 (in Russian). DOI: <https://doi.org/10.18500/1816-9791-2026-26-2-280-295>, EDN: RDHJQE

This is an open access article distributed under the terms of Creative Commons Attribution 4.0 International License (CC-BY 4.0)

Введение

Рассмотрим систему массового обслуживания (СМО) с пуассоновским входным потоком постоянной интенсивности, двумя обслуживающими приборами с одинаковым распределением времени обслуживания и накопителем неограниченной емкости. Поступающие заявки обслуживаются в порядке поступления. Перед началом обслуживания каждая заявка разделяется на две подзаявки, которые обслуживаются на двух приборах параллельно. Когда обслуживание обеих подзаявок заканчивается, заявка считается обслуженной.

Заметим, что в подобных системах (с разделением заявок) возможны различные дисциплины обращения с подзаявками. Основными из них являются следующие: либо подзаявки от разных заявок обслуживаются независимо друг от друга и образуются две параллельные очереди, каждая к своему прибору (fork-join), либо обслуживание подзаявок следующей заявки не начинается, пока не будут обслужены все подзаявки текущей (split-merge). Возможны и другие варианты.

Так, например, в [1] исследуется открытая сеть массового обслуживания с отдельной обработкой фрагментов заявок. Мы будем рассматривать системы типа split-merge. Такие системы фактически сводятся к системам $M|G|1$, в которых времена обслуживания заявок формируются как максимумы времен обслуживания их подзаявок на соответствующих приборах.

Системы типа split-merge и их обобщения исследуются, в частности, в работах [2–7].

Системы с разделением заявок могут иметь приложения в медицине, делопроизводстве, торговле, информатике и др. В общем случае возможно разделение заявок на произвольное фиксированное число $N \geq 2$ или даже случайное число подзаявок, а времена обслуживания на различных приборах могут иметь разное распределение, однако здесь мы ограничимся простейшим случаем $N = 2$ и однородными приборами.

Обычно времена обслуживания на приборах считаются независимыми. Настоящая же работа мотивирована статьей [6], где времена обслуживания полагаются зависимыми. В отличие от нашей работы, в [6] при этом рассматривается дисциплина cancel-on-finish, согласно которой заявка считается обслуженной, если обслужена хотя бы одна из ее подзаявок, т. е. времена обслуживания заявок формируются как минимумы, а не максимумы времен обслуживания подзаявок. Однако это различие не принципиально. По поводу механизма возникновения зависимости в [6] говорится, что во многих ситуациях, когда один прибор (фактически, возможно, человек) работает быстрее (или медленнее), это мотивирует второй прибор также работать быстрее (или медленнее). Это положительная зависимость между приборами.

В некоторых других ситуациях, если один прибор работает в спокойном темпе, другому приходится работать быстрее, чтобы сбалансировать нагрузку, это является примером отрицательной зависимости между приборами. Добавим от себя, что зависимость может быть также связана со случайной неравномерностью разделения работы (когда одна подзаявка оказывается больше, а другая меньше) или неоднородностью заявок по объему работ (когда обе подзаявки оказываются одновременно больше или меньше).

Для описания зависимости используется современная теория копул. В качестве учебника по ней можно указать [8]. Любое совместное распределение случайных величин можно разложить на его копулу и частные распределения. При этом выбор известных копул весьма широк.

Далее мы рассмотрим нашу систему при показательном распределении времен обслуживания для трех различных копул. Во всех случаях найдены средние времена пребывания заявок (времена отклика), изучено поведение коэффициентов корреляции Кендалла и Бломквиста, а также получены выражения для границ среднего времени отклика системы.

1. Математическая модель системы типа split-merge

В данном разделе более детально опишем исследуемую систему (рис. 1).

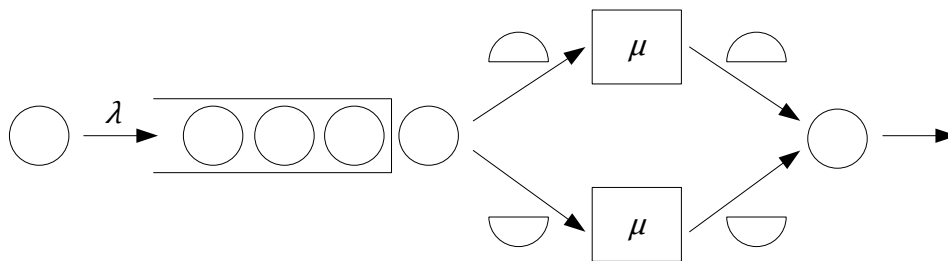


Рис. 1. Модель split-merge системы массового обслуживания с двумя подсистемами типа $M|M|1$

Fig. 1. Split-merge model of a queueing system with two subsystems of type $M|M|1$

Итак, рассматривается система массового обслуживания с пуассоновским входящим потоком с параметром λ , двумя зависимыми между собой обслуживающими приборами и накопителем неограниченной емкости. Заявка перед поступлением на обслуживание разделяется на две подзаявки, каждая из которых обслуживается на своем приборе. При этом подзаявка может покинуть свой прибор только одновременно со второй заявкой, т. е. только после того, как обслужится вторая подзаявка, поступившая одновременно с ней на обслуживание. Таким образом, время обслуживания всей заявки определяется выражением

$$X_{2:2} = \max\{X_1, X_2\},$$

где X_i — это случайное время обслуживания подзаявки на i -м приборе, $i = 1, 2$. Среднее время ожидания начала обслуживания и среднее время пребывания заявки в данной системе



фактически определяется выражениями для среднего времени ожидания начала обслуживания и среднего времени пребывания заявки в системе $M|G|1$, а именно

$$E[W] = \frac{\lambda E[X_{2:2}^2]}{2(1 - \lambda E[X_{2:2}])}, \tag{1}$$

$$E[V] = E[W] + E[X_{2:2}], \tag{2}$$

где (1) представляет собой известную формулу Поллачека – Хинчина.

Предположим, что времена обслуживания подзаявок имеют экспоненциальную функцию распределения:

$$F_{Exp}(x) = 1 - e^{-\mu x}, \quad x \geq 0.$$

Приборы, а соответственно, и времена обслуживания на них являются зависимыми. Далее рассмотрим несколько вариантов моделей зависимостей между приборами, которые опишем с помощью трех типов копула-функций.

Согласно определению с помощью, например, двумерной копулы можно представить функцию распределения двумерной случайной величины

$$F(x_1, x_2) = P(X_1 < x_1, X_2 < x_2) = C(F_1(x_1), F_2(x_2)),$$

где $F_i(x)$ — это частные функции распределения одномерных случайных величин, составляющих исходную двумерную. В том случае, когда частные распределения совпадают $F_1(x) = F_2(x) = F(x)$, имеем

$$F(x_1, x_2) = C(F(x_1), F(x_2)) = C(u_1, u_2), \quad 0 \leq u_1, u_2 \leq 1.$$

Обозначим через

$$F_{2:2}(x) = F_{X_{2:2}}(x) = P(X_{2:2} < x) = P(\max\{X_1, X_2\} < x)$$

функцию распределения для случайной величины времени обслуживания заявки на приборах, тогда, учитывая, что времена обслуживания являются одинаково распределенными случайными величинами, можем записать

$$F_{2:2}(x) = P(\max\{X_1, X_2\} < x) = P(X_1 < x, X_2 < x) = C(F(x), F(x)) = \delta(F(x)),$$

где $\delta(u)$ называется диагональным сечением копулы.

Для моделирования зависимостей между временами обслуживания на приборах используются следующие виды копул: копула Гумбеля, копула Фарли – Гумбеля – Моргенштерна и копула Тапа – Чжао, предложенная в работе [6].

2. Копулы и среднее время пребывания в СМО

Для экспоненциального распределения времени обслуживания на приборах рассмотрим три модели зависимостей между приборами и, соответственно, временами обслуживания на них.

2.1. Копула Гумбеля

Двумерная копула Гумбеля имеет следующий вид:

$$C(u_1, u_2) = \exp\{-((-\ln u_1)^\theta + (-\ln u_2)^\theta)^{\frac{1}{\theta}}\}, \quad 0 \leq u_1, u_2 \leq 1, \quad \theta \geq 1.$$

Данная копула, в частности, использовалась авторами в [9] для приближения зависимости времен пребывания в системе с разделением заявок типа fork-join.

Случай $\theta = 1$ соответствует независимости случайных величин, а в пределе при $\theta \rightarrow \infty$ получаем совершенную положительную зависимость (комонотонность).

Для копулы Гумбеля диагональное сечение ($u_1 = u_2 = u$) определяется как

$$\delta(u) = C(u, u) = u^{2^{1/\theta}}, \quad 0 \leq u \leq 1, \quad \theta \geq 1.$$

Тогда функция распределения времени обслуживания заявки примет вид

$$F_{2:2}(x) = \delta(F_{Exp}(x)) = (1 - e^{-\mu x})^{2^{1/\theta}}, \quad x \geq 0,$$

а ее плотность распределения, соответственно,

$$f_{2:2}(x) = F'_{2:2}(x) = 2^{1/\theta} \mu e^{-\mu x} (1 - e^{-\mu x})^{2^{1/\theta}-1}, \quad x \geq 0.$$

Далее можно рассчитать первый и второй моменты времени обслуживания:

$$E[X_{2:2}] = \int_0^\infty 2^{1/\theta} \mu x e^{-\mu x} (1 - e^{-\mu x})^{2^{1/\theta}-1} dx,$$

$$E[X_{2:2}^2] = \int_0^\infty 2^{1/\theta} \mu x^2 e^{-\mu x} (1 - e^{-\mu x})^{2^{1/\theta}-1} dx,$$

вычислив представленные интегралы.

Для произвольной случайной величины η , имеющей обобщенное экспоненциальное распределение с функцией распределения вида [10]

$$F(x) = (1 - e^{-\frac{x}{\beta}})^\alpha, \quad x \geq 0,$$

математическое ожидание и дисперсия определяются с помощью дигамма-функции $\psi(x)$:

$$E[\eta] = \beta[\psi(\alpha + 1) - \psi(1)],$$

$$Var[\eta] = \beta^2[\psi'(1) - \psi'(\alpha + 1)],$$

которая, в свою очередь, является производной от натурального логарифма гамма-функции

$$\psi(x) = \frac{d}{dx} \ln \Gamma(x),$$

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt.$$

Таким образом, с помощью дигамма-функции можем вычислить полученные интегралы, а именно для математического ожидания имеем

$$E[X_{2:2}] = \frac{1}{\mu} [\psi(2^{\frac{1}{\theta}} + 1) - \psi(1)],$$

а для второго момента будет справедливо

$$E[X_{2:2}^2] = \frac{1}{\mu^2} [\psi'(1) - \psi'(2^{\frac{1}{\theta}} + 1) + (\psi(2^{\frac{1}{\theta}} + 1) - \psi(1))^2].$$

Далее по (1),(2) вычисляем $E[V]$ (формула не приводится из-за ее громоздкости).

На рис. 2 приведены графики среднего времени пребывания заявки в системе при $\mu = 1$ и разных θ . Видно, что чем больше θ , тем ниже проходит график (меньше среднее).

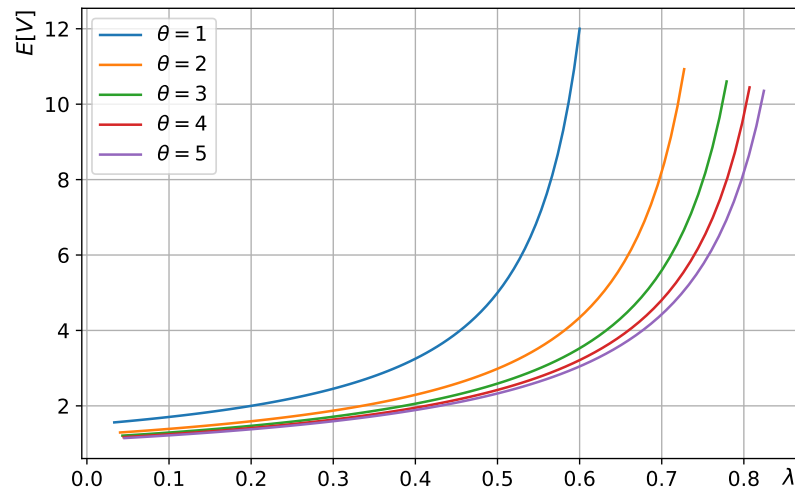


Рис. 2. Среднее время пребывания заявки в СМО для случая копулы Гумбеля (цвет онлайн)
 Fig. 2. Average sojourn time of a request in the system for the case with Gumbel copula (color online)

2.2. Копула Фарли – Гумбеля – Morgenштерна

Копула-функция имеет в этом случае следующий вид:

$$C(u_1, u_2) = u_1 u_2 (1 + \theta(1 - u_1)(1 - u_2)), \quad 0 \leq u_1, u_2 \leq 1, \quad -1 \leq \theta \leq 1.$$

Случай $\theta = 0$ соответствует независимости, $\theta < 0$ — отрицательной зависимости, а $\theta > 0$ — положительной зависимости (но совершенная зависимость не достигается при $\theta = \pm 1$).

Диагональное сечение определяется выражением

$$\delta(u) = C(u, u) = u^2(1 + \theta(1 - u^2)).$$

Далее определим выражение для функции и плотности распределения времени обслуживания

$$F_{2:2}(x) = \delta(F_{Exp}(x)) = 1 - 2e^{-\mu x} + (1 + \theta)e^{-2\mu x} - 2\theta e^{-3\mu x} + \theta e^{-4\mu x},$$

$$f_{2:2}(x) = F'_{2:2}(x) = 2\mu e^{-\mu x} - 2\mu(1 + \theta)e^{-2\mu x} + 6\mu\theta e^{-3\mu x} - 4\mu\theta e^{-4\mu x}.$$

Тогда

$$E[X_{2:2}] = \int_0^\infty x f_{2:2}(x) dx = \frac{18 - \theta}{12\mu},$$

$$E[X_{2:2}^2] = \int_0^\infty x^2 f_{2:2}(x) dx = \frac{252 - 13\theta}{72\mu^2}.$$

Следовательно,

$$Var[X_{2:2}] = \frac{\theta(10 - \theta) + 180}{144\mu^2} = \frac{10\theta - \theta^2 + 180}{144\mu^2}.$$

Теперь можем вычислить среднее время ожидания начала обслуживания и среднее время пребывания заявки в системе, подставив соответствующие выражения в общие формулы:

$$E[W] = \frac{\lambda(252 - 13\theta)}{12\mu(12\mu - \lambda(18 - \theta))},$$

$$E[V] = \frac{12\mu(18 - \theta) - \lambda(\theta^2 - 23\theta + 72)}{12\mu(12\mu - \lambda(18 - \theta))}.$$

На рис. 3 представлены графики среднего времени пребывания заявки в системе при $\mu = 1$ и разных θ . Видно, что чем больше θ , тем ниже проходит график (меньше среднее).

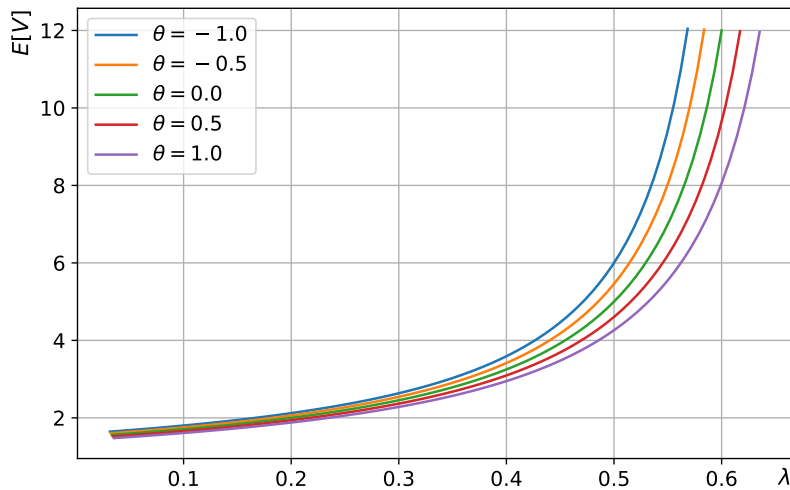


Рис. 3. Среднее время пребывания заявки в СМО для случая копулы Фарли–Гумбеля–Моргенштерна (цвет онлайн)
 Fig. 3. Average sojourn time of a request in the system for the case with Farlie–Gumbel–Morgenstern copula (color online)

2.3. Копула Тапа – Чжао

Далее рассмотрим двумерную копулу, предложенную в работе [6]. Она имеет вид

$$C(u_1, u_2) = (1 - \theta^2)u_1u_2 + \frac{\theta}{4} \left[(1 + \theta)^2 \min(u_1, u_2) - (1 - \theta)^2 \max(u_1 + u_2 - 1, 0) \right], \quad -1 \leq \theta \leq 1.$$

Случаю $\theta = 0$ соответствует независимость, $\theta = -1$ – совершенная отрицательная зависимость (контрамонотонность), $\theta = 1$ – совершенная положительная зависимость (комонотонность).

Диагональное сечение определяется следующим выражением:

$$\delta(u) = C(u, u) = (1 - \theta^2)u^2 + \frac{\theta}{4} \left[(1 + \theta)^2 u - (1 - \theta)^2 \max(2u - 1, 0) \right].$$

Следовательно, функция распределения времени обслуживания равна

$$F_{2:2}(x) = \delta(F_{Exp}(x)) = (1 - \theta^2)(1 - 2e^{-\mu x} + e^{-2\mu x}) + \frac{\theta}{4} \left[(1 + \theta)^2(1 - e^{-\mu x}) - (1 - \theta)^2 \max(1 - 2e^{-\mu x}, 0) \right], \quad x \geq 0,$$

а ее плотность распределения равна

$$f_{2:2}(x) = (1 - \theta^2)(2\mu e^{-\mu x} - 2\mu e^{-2\mu x}) + \frac{\theta}{4} \left[(1 + \theta)^2 \mu e^{-\mu x} - (1 - \theta)^2 2\mu e^{-\mu x} \mathbf{I}(x \geq \frac{1}{\mu} \ln 2) \right], \quad x \geq 0.$$

Для удобства запишем плотность распределения как

$$f_{2:2}(x) = \begin{cases} g(x), & 0 \leq x < \frac{1}{\mu} \ln 2, \\ g(x) - \frac{\theta}{4}(1 - \theta)^2 2\mu e^{-\mu x}, & x \geq \frac{1}{\mu} \ln 2, \end{cases}$$

где

$$g(x) = (1 - \theta^2)(2\mu e^{-\mu x} - 2\mu e^{-2\mu x}) + \frac{\theta}{4}(1 + \theta)^2 \mu e^{-\mu x}.$$



Теперь вычислим первый и второй моменты времени обслуживания:

$$\begin{aligned}
 E[X_{2:2}] &= \int_0^\infty x f_{2:2}(x) dx = \int_0^{\frac{1}{\mu} \ln 2} x g(x) dx + \int_{\frac{1}{\mu} \ln 2}^\infty x [g(x) - \frac{\theta}{4}(1-\theta)^2 2\mu e^{-\mu x}] dx = \\
 &= \int_0^\infty x g(x) dx - \frac{\theta}{4}(1-\theta)^2 \int_{\frac{1}{\mu} \ln 2}^\infty 2x\mu e^{-\mu x} dx = \\
 &= (1-\theta^2) \frac{3}{2\mu} + \frac{\theta}{4} \cdot \frac{1}{\mu} \left[(1+\theta)^2 - (1-\theta)^2(1+\ln 2) \right], \\
 E[X_{2:2}^2] &= \int_0^\infty x^2 f_{2:2}(x) dx = \int_0^\infty x^2 g(x) dx - \frac{\theta}{4}(1-\theta)^2 \int_{\frac{1}{\mu} \ln 2}^\infty 2x^2\mu e^{-\mu x} dx = \\
 &= (1-\theta^2) \frac{7}{2\mu^2} + \frac{\theta(1+\theta)^2}{2\mu^2} - \frac{\theta(1-\theta)^2}{4\mu^2} \left[\ln^2 2 + 2 \ln 2 + 2 \right].
 \end{aligned}$$

Далее уже можем определить среднее время ожидания начала обслуживания и среднее время пребывания заявки в системе, подставляя полученные выражения в соответствующие формулы. В результате для среднего времени ожидания начала обслуживания получим

$$E[W] = \frac{\lambda(14 - 6\theta^2 - \theta(1-\theta)^2(\ln^2 2 + 2 \ln 2))}{2\mu(4\mu - 6\lambda + \theta\lambda(2\theta + \ln 2(1-\theta)^2))}.$$

На рис. 4 приведены графики среднего времени пребывания заявки в системе при $\mu = 1$ и разных θ . Видно, что чем больше θ , тем ниже проходит график (меньше среднее).

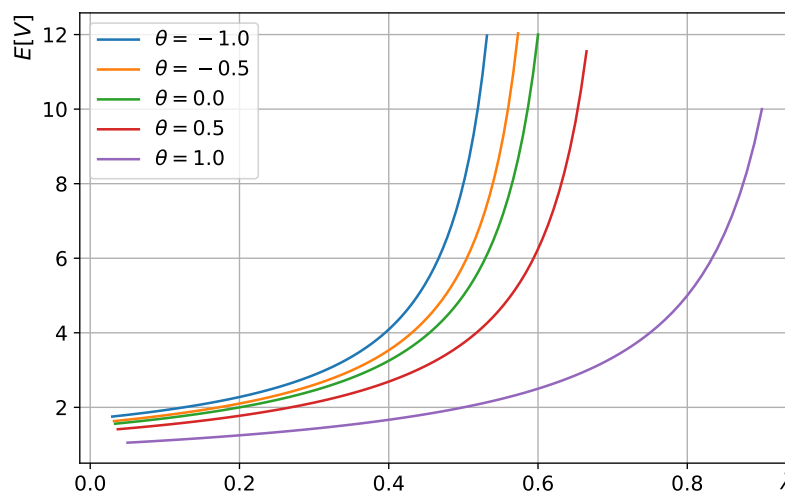


Рис. 4. Среднее время пребывания заявки в СМО для случая копулы Тапа–Чжао (цвет онлайн)

Fig. 4. Average sojourn time of a request in the system for the case with Thapa–Zhao copula (color online)

3. Сравнительный анализ с учетом коэффициентов корреляции

У каждой из рассмотренных копул свой параметр θ , имеющий разный смысл, поэтому результаты для них нельзя сравнивать непосредственно. Чтобы провести сравнение, необходимо параметризовать копулы единым образом. Это можно сделать с помощью различных мер зависимости, однозначно определяемых копулами, такими как коэффициенты корреляции Спирмена, Кендалла и Бломквиста [8, гл. 5].

К сожалению, для копулы Гумбеля коэффициент корреляции Спирмена не выражается в явном виде, поэтому ограничимся коэффициентами Кендалла и Бломквиста.

3.1. Коэффициент корреляции Кендалла

Коэффициент корреляции Кендалла случайных величин X_1 и X_2 определяется формулой

$$\tau = E[\text{sign}(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2)],$$

где $(\tilde{X}_1, \tilde{X}_2)$ — независимая вероятностная копия (X_1, X_2) .

Коэффициент корреляции Кендалла для произвольной копула-функции $C(u_1, u_2)$ рассчитывается по формуле [8, теорема 5.1.3]

$$\tau = 4E[C(u_1, u_2)] - 1 = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1$$

и принимает значения от -1 до 1 .

Соответственно, для копулы Гумбеля он находится как [8, пример 5.4.]

$$\tau_G = 1 - 1/\theta, \quad \theta \geq 1,$$

для копулы Фарме–Гумбеля–Моргенштерна [8, пример 5.2] —

$$\tau_{FGM} = \frac{2}{9}\theta, \quad -1 \leq \theta \leq 1,$$

а для копулы Тапа–Чжао [6, теорема 2.2] —

$$\tau_{TCh} = \frac{\theta(8 - 3\theta + 12\theta^2 + 6\theta^3 + 4\theta^4 - 3\theta^5)}{24}, \quad -1 \leq \theta \leq 1.$$

На рис. 5 представлена зависимость среднего времени отклика системы от значения коэффициента корреляции Кендалла для всех рассматриваемых трех типов копул при $\mu = 1$ и $\lambda = 1/2$ (при этих значениях параметров все средние времена пребывания конечны).

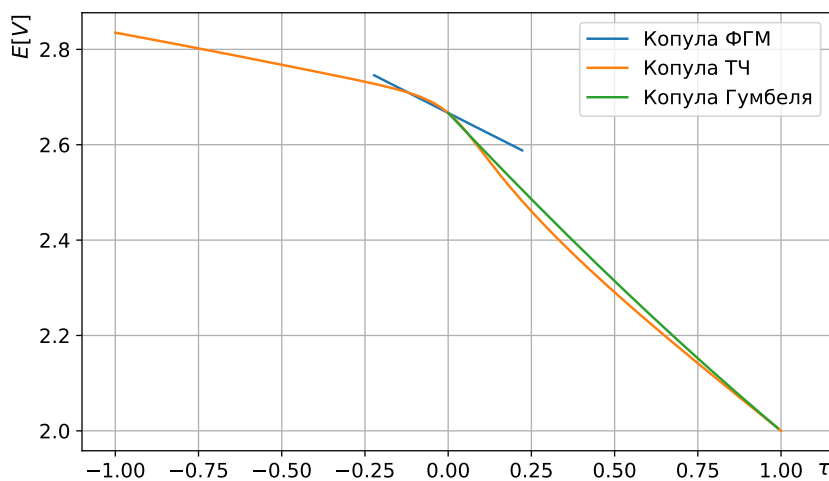


Рис. 5. Среднее время пребывания заявки в СМО в зависимости от коэффициента корреляции Кендалла, рассчитанного для трех типов копул, $\mu = 1$, $\lambda = 1/2$ (цвет онлайн)

Fig. 5. Average sojourn time of a request in the system depending on the Kendall correlation coefficient, calculated for three types of copulas, $\mu = 1$, $\lambda = 1/2$ (color online)

Как ясно из аналитических формул и представленных на рис. 5 графиков, наиболее широкий диапазон значений коэффициента корреляции Кендалла позволяет получить модель копулы Тапа–Чжао, т. е. весь спектр возможных значений от -1 до 1 , на втором месте стоит копула Гумбеля с положительными значениями от 0 до 1 и, наконец, копула Фарли–Гумбеля–Моргенштерна со значениями коэффициента корреляции в диапазоне от $-2/9$ до $2/9$.



Видно, что все три графика пересекаются при $\tau = 0$ (случай независимости), при $\tau < 0$ копулы Фарли–Гумбеля–Моргенштерна и Тапа–Чжао дают близкие значения средних времен пребывания (и их графики имеют еще одну точку пересечения), при $\tau > 0$ копулы Гумбеля и Тапа–Чжао дают близкие значения и сходятся в точке $\tau = 1$ (случай совершенной положительной зависимости), а копула Фарли–Гумбеля–Моргенштерна дает сильное расхождение с остальными.

3.2. Коэффициент корреляции Бломквиста

Коэффициент корреляции Бломквиста случайных величин X_1 и X_2 определяется формулой

$$\beta = E[(X_1 - \text{med}X_1)(X_2 - \text{med}X_2)],$$

где $\text{med}X$ — медиана X .

Коэффициент корреляции Бломквиста для копула-функции $C(u_1, u_2)$ рассчитывается по формуле [8, (5.1.27)]

$$\beta = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1$$

и принимает значения от -1 до 1 .

Таким образом, для копулы Гумбеля имеем

$$\beta_G = 2^{2-2^{1/\theta}} - 1, \quad \theta \geq 1,$$

для копулы Фарли–Гумбеля–Моргенштерна —

$$\beta_{FGM} = \frac{3}{4}\theta, \quad -1 \leq \theta \leq 1,$$

а для копулы Тапа–Чжао [6, теорема 2.2] —

$$\beta_{TCh} = \frac{\theta(\theta^2 + 1)}{2}, \quad -1 \leq \theta \leq 1.$$

На рис. 6 представлена зависимость среднего времени отклика системы от значения коэффициента корреляции Бломквиста также для всех рассматриваемых трех типов копул.

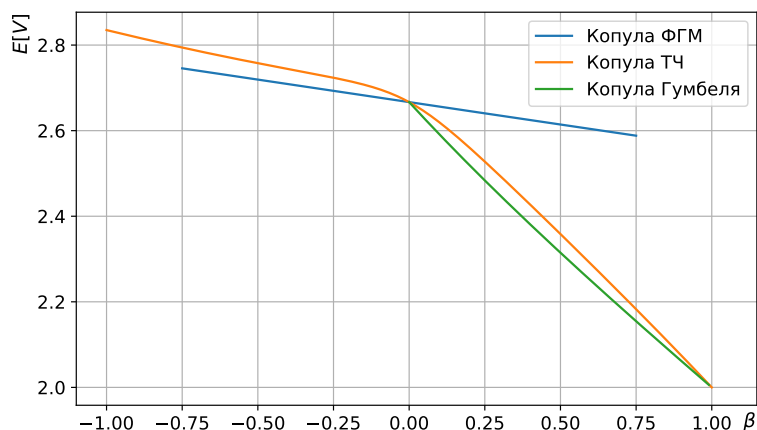


Рис. 6. Среднее время пребывания заявки в СМО в зависимости от коэффициента корреляции Бломквиста, рассчитанного для трех типов копул, $\mu = 1, \lambda = 1/2$ (цвет онлайн)

Fig. 6. Average sojourn time of a request in the system depending on the Blomqvist correlation coefficient, calculated for three types of copulas, $\mu = 1, \lambda = 1/2$ (color online)

Так же как и в предыдущем случае, исходя из аналитических формул и представленных на рис. 6 графиков, наиболее широкий диапазон значений коэффициента корреляции Бломквиста от -1 до 1 позволяет получить именно модель копулы Тапа – Чжао, далее следует копула Гумбеля с положительными значениями от 0 до 1 и затем копула Фарли – Гумбеля – Моргенштерна со значениями коэффициента корреляции в диапазоне от $-3/4$ до $3/4$ (что значительно шире, чем у коэффициента Кендалла).

Видно, что все три графика пересекаются при $\beta = 0$ (случай независимости), при $\beta < 0$ копулы Фарли – Гумбеля – Моргенштерна и Тапа – Чжао дают близкие значения средних времен пребывания (но больше точек пересечения нет), при $\beta > 0$ копулы Гумбеля и Тапа – Чжао дают близкие значения и сходятся в точке $\beta = 1$ (случай совершенной положительной зависимости), а копула Фарли – Гумбеля – Моргенштерна дает сильное расхождение с остальными.

4. Верхняя и нижняя границы для среднего времени отклика

Мы рассмотрели три класса известных копул, однако этот выбор в значительной степени произволен. Можно задаться вопросом, каковы вообще верхние и нижние границы для среднего времени отклика для любых копул, но с известными коэффициентами корреляции?

В [8] приведена теорема 5.1.16, дающая верхние и нижние (поточечные) границы для копул с известными коэффициентами корреляции Спирмена, Кендалла и Бломквиста. При этом сами границы также являются копулами, что наводит на мысль рассчитать соответствующие средние времена отклика. К сожалению, в явном виде это получается только в случае коэффициента Бломквиста.

Для этого обратимся к выражениям для верхних и нижних границ копул с известным значением коэффициента корреляции Бломквиста β , который, напомним, принимает значения в диапазоне от -1 до 1 включительно:

$$\begin{aligned} \underline{B}(u_1, u_2) &= \max \left(W(u_1, u_2), \frac{\beta + 1}{4} - \left(u_1 - \frac{1}{2}\right)^+ - \left(u_2 - \frac{1}{2}\right)^+ \right), \\ \overline{B}(u_1, u_2) &= \min \left(M(u_1, u_2), \frac{\beta + 1}{4} + \left(u_1 - \frac{1}{2}\right)^+ + \left(u_2 - \frac{1}{2}\right)^+ \right), \end{aligned}$$

где $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$, $M(u_1, u_2) = \min(u_1, u_2)$, $x^+ = \begin{cases} 0, & x < 0, \\ x, & x \geq 0. \end{cases}$

Теперь выразим границы для диагонального сечения $\delta_\beta(u)$ копулы $C(u_1, u_2)$ в зависимости от значения β , т. е. когда $u_1 = u_2 = u$, тогда нижняя граница для диагонального сечения примет вид

$$\delta_\beta(u) = \begin{cases} 0, & 0 \leq u \leq \frac{3-\beta}{8}, \\ 2u - 1 + \frac{\beta+1}{4}, & \frac{3-\beta}{8} \leq u \leq \frac{1}{2}, \\ \frac{\beta+1}{4}, & \frac{1}{2} \leq u \leq \frac{5+\beta}{8}, \\ 2u - 1, & \frac{5+\beta}{8} \leq u \leq 1, \end{cases}$$

а верхняя, соответственно,

$$\overline{\delta}_\beta(u) = \begin{cases} u, & 0 \leq u \leq \frac{\beta+1}{4}, \\ \frac{\beta+1}{4}, & \frac{\beta+1}{4} \leq u \leq \frac{1}{2}, \\ 2u - 1 + \frac{\beta+1}{4}, & \frac{1}{2} \leq u \leq \frac{3-\beta}{4}, \\ u, & \frac{3-\beta}{4} \leq u \leq 1. \end{cases}$$

Далее воспользуемся полученными выражениями для определения верхней и нижней границы для среднего времени пребывания заявки в системе. Для этого выразим соответствующие функции и плотности распределения времени обслуживания заявки (частей заявки) на обоих приборах.



При этом необходимо отметить, что нижняя граница копул дает верхнюю границу среднего времени пребывания, а верхняя граница копул — нижнюю границу среднего времени пребывания. Действительно, большие значения копулы ведут к большим значениям функции распределения (как большей вероятности, что случайная величина меньше заданного уровня), т. е. к стохастически меньшим временам обслуживания, имеющим меньшие средние и средние квадраты (в силу неотрицательности случайных величин), что приводит к меньшим средним временам ожидания и пребывания по формулам (1) и (2).

Через \bar{F} и \underline{F} обозначим функции распределения для нахождения верхней и нижней границ среднего времени пребывания, тогда

$$\underline{F}_{\beta,2:2}(x) = \bar{\delta}_{\beta}(F(x)), \quad \bar{F}_{\beta,2:2}(x) = \underline{\delta}_{\beta}(F(x)), \quad F(x) = 1 - e^{-\mu x}, \quad x \geq 0,$$

в результате получаем следующие формулы для искомых функций распределения:

$$\underline{F}_{\beta,2:2}(x) = \begin{cases} 1 - e^{-\mu x}, & 0 \leq x \leq -\frac{1}{\mu} \ln \frac{3-\beta}{4}, \quad x \geq -\frac{1}{\mu} \ln \frac{\beta+1}{4}, \quad \beta \neq -1, \\ \frac{\beta+1}{4}, & -\frac{1}{\mu} \ln \frac{3-\beta}{4} \leq x \leq -\frac{1}{\mu} \ln \frac{1}{2}, \\ 1 - 2e^{-\mu x} + \frac{\beta+1}{4}, & -\frac{1}{\mu} \ln \frac{1}{2} \leq x \leq -\frac{1}{\mu} \ln \frac{\beta+1}{4}, \quad \beta \neq -1, \end{cases}$$

причем для частного случая $\beta = -1$ имеем

$$\underline{F}_{-1,2:2}(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{\mu} \ln 2, \\ 1 - 2e^{-\mu x}, & x \geq \frac{1}{\mu} \ln 2, \end{cases}$$

$$\bar{F}_{\beta,2:2}(x) = \begin{cases} 0, & 0 \leq x \leq -\frac{1}{\mu} \ln \frac{5+\beta}{8}, \\ 1 - 2e^{-\mu x} + \frac{\beta+1}{4}, & -\frac{1}{\mu} \ln \frac{5+\beta}{8} \leq x \leq -\frac{1}{\mu} \ln \frac{1}{2}, \\ \frac{\beta+1}{4}, & -\frac{1}{\mu} \ln \frac{1}{2} \leq x \leq -\frac{1}{\mu} \ln \frac{3-\beta}{8}, \\ 1 - 2e^{-\mu x}, & x \geq -\frac{1}{\mu} \ln \frac{3-\beta}{8}, \end{cases}$$

и для плотностей распределения справедливо

$$\underline{f}_{\beta,2:2}(x) = \begin{cases} 0, & -\frac{1}{\mu} \ln \frac{3-\beta}{4} \leq x \leq -\frac{1}{\mu} \ln \frac{1}{2}, \\ 2\mu e^{-\mu x}, & -\frac{1}{\mu} \ln \frac{1}{2} \leq x \leq -\frac{1}{\mu} \ln \frac{\beta+1}{4}, \\ \mu e^{-\mu x}, & 0 \leq x \leq -\frac{1}{\mu} \ln \frac{3-\beta}{4}, \quad x \geq -\frac{1}{\mu} \ln \frac{\beta+1}{4}, \end{cases}$$

причем для $\beta = -1$ получаем $\underline{f}_{-1,2:2}(x) = 2\mu e^{-\mu x}, \quad x \geq \frac{1}{\mu} \ln 2,$

$$\bar{f}_{\beta,2:2}(x) = \begin{cases} 0, & 0 \leq x \leq -\frac{1}{\mu} \ln \frac{5+\beta}{8}, \quad -\frac{1}{\mu} \ln \frac{1}{2} \leq x \leq -\frac{1}{\mu} \ln \frac{3-\beta}{8}, \\ 2\mu e^{-\mu x}, & -\frac{1}{\mu} \ln \frac{5+\beta}{8} \leq x \leq -\frac{1}{\mu} \ln \frac{1}{2}, \quad x \geq -\frac{1}{\mu} \ln \frac{3-\beta}{8}. \end{cases}$$

Далее, используя функции плотностей, находим нижние и верхние границы для первого и второго момента времени обслуживания соответственно.

В итоге для нижней границы имеем следующие выражения:

$$E[\underline{X}_{\beta,2:2}] = \frac{1}{\mu} \left[\frac{3-\beta}{4} \left(\ln \frac{3-\beta}{4} - 1 \right) + \frac{1+\beta}{4} \left(\ln \frac{1+\beta}{4} - 1 \right) + \ln 2 + 2 \right], \quad \beta \neq -1,$$

$$E[\underline{X}_{-1,2:2}] = \frac{1 + \ln 2}{\mu},$$

$$E[\underline{X}_{\beta,2:2}^2] = \frac{1}{\mu^2} \left[\ln^2 2 + 2 \ln 2 + 4 - \frac{3-\beta}{4} \left(\ln^2 \frac{3-\beta}{4} - 2 \ln \frac{3-\beta}{4} + 2 \right) - \frac{1+\beta}{4} \left(\ln^2 \frac{1+\beta}{4} - 2 \ln \frac{1+\beta}{4} + 2 \right) \right], \quad \beta \neq -1,$$

$$E[X_{-1,2:2}^2] = \frac{\ln^2 2 + 2 \ln 2 + 2}{\mu^2},$$

которые при подстановке в общую формулу для среднего времени ожидания начала обслуживания

$$E[W_\beta] = \frac{\lambda E[X_{\beta,2:2}^2]}{2(1 - \lambda E[X_{\beta,2:2}])}$$

и среднего времени пребывания заявки в системе

$$E[V_1] = E[V_\beta] = E[W_\beta] + E[X_{\beta,2:2}]$$

позволят определить нижние границы этих величин в зависимости от значения коэффициента корреляции Бломквиста.

Соответственно, для верхних границ среднего времени ожидания начала обслуживания и среднего времени пребывания заявки в системе справедливо

$$E[\bar{W}_\beta] = \frac{\lambda E[\bar{X}_{\beta,2:2}^2]}{2(1 - \lambda E[\bar{X}_{\beta,2:2}])}, \quad E[V_2] = E[\bar{V}_\beta] = E[\bar{W}_\beta] + E[\bar{X}_{\beta,2:2}],$$

где

$$E[\bar{X}_{\beta,2:2}] = \frac{1}{\mu} \left[\frac{3-\beta}{4} \left(1 - \ln \frac{3-\beta}{8} \right) + \frac{5+\beta}{4} \left(1 - \ln \frac{5+\beta}{8} \right) - \ln 2 - 1 \right],$$

$$E[\bar{X}_{\beta,2:2}^2] = \frac{1}{\mu^2} \left[\frac{3-\beta}{4} \left(\ln^2 \frac{3-\beta}{8} - 2 \ln \frac{3-\beta}{8} + 2 \right) + \right.$$

$$\left. + \frac{5+\beta}{4} \left(\ln^2 \frac{5+\beta}{8} - 2 \ln \frac{5+\beta}{8} + 2 \right) - \ln^2 2 - 2 \ln 2 - 2 \right].$$

На рис. 7 представлены графики зависимостей для нижних и верхних границ среднего времени пребывания заявки в системе в зависимости от значения коэффициента корреляции Бломквиста β при $\mu = 1$.

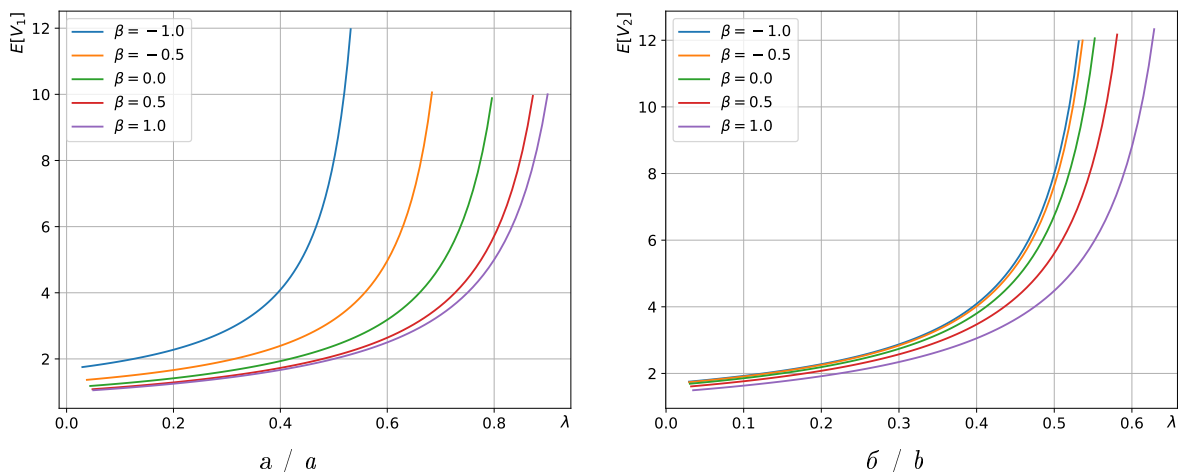


Рис. 7. Зависимость границы среднего времени пребывания заявки в системе от интенсивности входящего потока λ : а — нижней; б — верхней (цвет онлайн)

Fig. 7. Dependence of the limit of the average sojourn time of a request in the system on the arrival rate λ : a is lower; b is upper (color online)

На рис. 8 можно увидеть, как расположены значения среднего времени пребывания заявки в системе, моделируемые тремя типами копул для различных значений коэффициента корреляции Бломквиста при $\mu = 1, \lambda = 1/2$. Полученные верхняя и нижняя границы для



среднего времени пребывания очень широкие, причем они сходятся при $\beta = -1$ (случай совершенной отрицательной зависимости), но не сходятся при $\beta = 1$ (поскольку это значение коэффициента Бломквиста не задает однозначно совершенную положительную зависимость по нижней границе копул). Таким образом, для получения более узких границ нужно либо опираться на другие меры зависимости, либо накладывать дополнительные ограничения на возможные копулы.

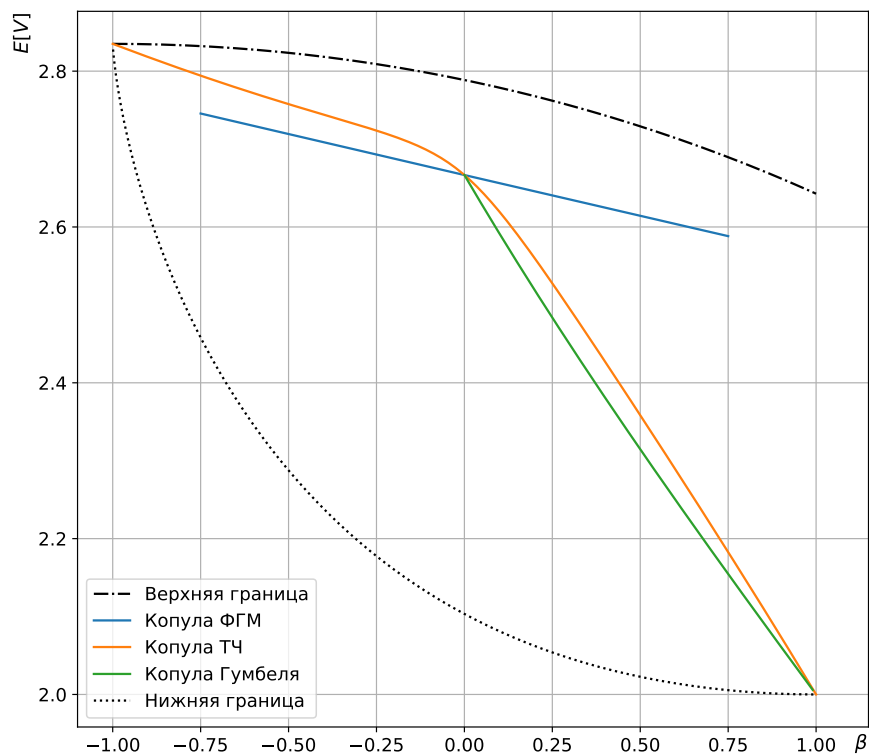


Рис. 8. Расположение среднего времени пребывания заявки в СМО в зависимости от коэффициента корреляции Бломквиста, рассчитанного для трех типов копул, относительно верхней и нижней границ (цвет онлайн)

Fig. 8. The location of the average sojourn time relative to the upper and lower bounds, depending on the Blomqvist correlation coefficient calculated for three types of copulas (color online)

Заключение

В результате исследования получены выражения для среднего времени пребывания в системах с разделением заявок типа split-merge с двумя приборами и зависимыми временами обслуживания на них, с показательным распределением. Разобраны три типа копул, описывающих зависимость: Гумбеля, Фарли – Гумбеля – Моргенштерна и Тапа – Чжао. Во всех случаях выведены формулы для первых двух моментов времени обслуживания заявки и среднего времени пребывания заявки (времени отклика).

Проведен сравнительный анализ с учетом коэффициентов корреляции Кендалла и Бломквиста. В обоих случаях при отрицательной корреляции копулы Фарли – Гумбеля – Моргенштерна и Тапа – Чжао дают близкие значения средних времен пребывания, при положительной корреляции копулы Гумбеля и Тапа – Чжао дают близкие значения, а копула Фарли – Гумбеля – Моргенштерна дает сильное расхождение с остальными.

Получены верхние и нижние границы для среднего времени пребывания для любых



копул при известном коэффициенте корреляции Бломквиста. Эти границы оказываются очень широкими. Для получения более узких границ нужно либо опираться на другие меры зависимости, либо накладывать дополнительные ограничения на возможные копулы.

Представленный подход может быть распространен на другие распределения времен обслуживания подзаявок (в том числе с тяжелыми хвостами) и на системы с разделением на $N > 2$ подзаявок с использованием многомерных копул.

Список литературы

1. Постнова О. С., Тананко И. Е., Рогачко Е. С. Приближенный анализ длительности пребывания требований в сети массового обслуживания с делением и слиянием требований // Управление большими системами. 2025. Вып. 115. С. 33–51.
2. Fiorini P. M., Lipsky L. Exact analysis of some split-merge queues // ACM SIGMETRICS Performance Evaluation Review. 2015. Vol. 43, iss. 2. P. 51–53. DOI: <https://doi.org/10.1145/2825236.2825257>
3. Harrison P., Zertal S. Queueing models with maxima of service times // Computer Performance Evaluation. Modelling Techniques and Tools / eds. P. Kemper, W. H. Sanders. Berlin ; Heidelberg : Springer, 2003. P. 152–168. (TOOLS 2003. Lecture Notes in Computer Science. Vol. 2794). DOI: https://doi.org/10.1007/978-3-540-45232-4_10
4. Joshi G., Soljanin E., Wornell G. Efficient redundancy techniques for latency reduction in cloud systems // ACM Transactions on Modeling and Performance Evaluation of Computing Systems (TOMPECS). 2017. Vol. 2, iss. 2. P. 1–30. DOI: <https://doi.org/10.1145/3055281>
5. Румянцев А. С., Долгалева Д. С., Головин А. С. Исследование стационарных характеристик многосерверных моделей с избыточностью // Программные системы: теория и приложения. 2023. Т. 14, № 1. С. 55–94. DOI: <https://doi.org/10.25209/2079-3316-2023-14-1-55-94>
6. Thapa S., Zhao Y. Q. Construction of new copulas with queueing application. 2021. DOI: <https://doi.org/10.48550/arXiv.2101.12401>
7. Tsimashenka I., Knottenbelt W. J., Harrison P. G. Controlling variability in split-merge systems and its impact on performance // Annals of Operations Research. 2016. Vol. 239. P. 569–588. DOI: <https://doi.org/10.1007/s10479-014-1560-3>
8. Nelsen R. B. An introduction to copulas. New York : Springer, 2006. 276 p. DOI: <https://doi.org/10.1007/0-387-28678-0>
9. Gorbunova A. V., Lebedev A. V. Copulas and quantiles in fork-join queueing systems // Advances in Systems Science and Applications. 2024. Vol. 24, iss. 1. P. 1–19. DOI: <https://doi.org/10.25728/assa.2024.24.1.1585>
10. Gupta R. D., Kundu D. Theory & methods: Generalized exponential distributions // Australian and New Zealand Journal of Statistics. 1999. Vol. 41. P. 173–188. DOI: <https://doi.org/10.1111/1467-842X.00072>

References

1. Postnova O. S., Tananko I. E., Rogachko E. S. An approximation for the response time in fork-join queueing networks. *Large-Scale Systems Control*, 2025, iss. 115, pp. 33–51 (in Russian).
2. Fiorini P. M., Lipsky L. Exact analysis of some split-merge queues. *ACM SIGMETRICS Performance Evaluation Review*, 2015, vol. 43, iss. 2, pp. 51–53. DOI: <https://doi.org/10.1145/2825236.2825257>
3. Harrison P., Zertal S. Queueing models with maxima of service times. In: Kemper P., Sanders W. H. (eds.) *Computer Performance Evaluation. Modelling Techniques and Tools*. TOOLS 2003. Lecture Notes in Computer Science, vol. 2794. Berlin, Heidelberg, Springer, 2003, pp. 152–168. DOI: https://doi.org/10.1007/978-3-540-45232-4_10
4. Joshi G., Soljanin E., Wornell G. Efficient redundancy techniques for latency reduction in cloud systems. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems (TOMPECS)*, 2017, vol. 2, iss. 2, pp. 1–30. DOI: <https://doi.org/10.1145/3055281>



5. Rumyantsev A. S., Dolgaleva D. S., Golovin A. S. Steady-state performance analysis of multiserver queueing models with redundancy. *Program Systems: Theory and Applications*, 2023, vol. 14, iss. 1, pp. 55–94 (in Russian). DOI: <https://doi.org/10.25209/2079-3316-2023-14-1-55-94>
6. Thapa S., Zhao Y. Q. *Construction of new copulas with queueing application*. 2021. DOI: <https://doi.org/10.48550/arXiv.2101.12401>
7. Tsimashenka I., Knottenbelt W. J., Harrison P. G. Controlling variability in split-merge systems and its impact on performance. *Annals of Operations Research*, 2016, vol. 239, pp. 569–588. DOI: <https://doi.org/10.1007/s10479-014-1560-3>
8. Nelsen R. B. *An introduction to copulas*. New York, Springer, 2006. 276 p. DOI: <https://doi.org/10.1007/0-387-28678-0>
9. Gorbunova A. V., Lebedev, A. V. Copulas and quantiles in fork-join queueing systems. *Advances in Systems Science and Applications*, 2024, vol. 24, iss. 1, pp. 1–19. DOI: <https://doi.org/10.25728/assa.2024.24.1.1585>
10. Gupta R. D., Kundu D. Theory & methods: Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, 1999, vol. 41, pp. 173–188. DOI: <https://doi.org/10.1111/1467-842X.00072>

Поступила в редакцию / Received 20.01.2026

Принята к публикации / Accepted 20.02.2026

Опубликована / Published 01.06.2026