Расчет многоканальных СМО с отрицательными заявками

Р. С. Хабаров, к.т.н.
В. А. Лохвицкий, д.т.н.

Аннотация

Предложен численный метод расчёта многоканальных систем массового обслуживания с произвольным распределением времени обслуживания (аппроксимируемым гиперэкспоненциальным H_2 с комплексными параметрами) и двумя дисциплинами обработки отрицательных заявок: **RCS** (удаление заявки на обслуживании) и **катастрофы** (мгновенная очистка системы). Метод расширяет алгоритм Такахаси–Таками модификацией матриц переходов микросостояний и позволяет получать не только стационарные вероятности состояний, но и временные характеристики — начальные моменты времени ожидания и пребывания. Получены результаты для двух вариантов поведения положительных заявок при поступлении отрицательной: при уходе из системы и при направлении на повторное обслуживание.

Ключевые слова: теория массового обслуживания, отрицательные заявки, катастрофы, RCS, метод Такахаси–Таками, фазовые аппроксимации, H_2, ПЛС (преобразование Лапласа–Стилтьеса).

Введение

Метод Такахаси–Таками [1] изначально предложен для общего класса систем $GI/G/c$; в монографии Рыжикова [2] подробно изложены многоканальные фазовые системы и итерационное решение векторно-матричных уравнений баланса с блоками A_j, B_j, C_j, D_j. Метод применим к достаточно широкому классу систем (в т. ч. $M/H_k/n$, $H_k/M/n$ и др.); в данной работе рассматриваются системы с *пуассоновским входом* интенсивности λ и *распределением времени обслуживания* B, аппроксимируемым гиперэкспоненциальным H_2 с комплексными параметрами, позволяющими описывать системы с произвольным коэффициентом вариации при n каналах. Предположения о пуассоновском потоке заявок и произвольном обслуживании являются достаточно общими для многих реальных систем.

В последние десятилетия значительный интерес вызывает исследование систем с отрицательными заявками (negative customers), введёнными в работах Геленбе [3, 4]. Отрицательные заявки при поступлении могут удалять положительные заявки согласно различным дисциплинам. В работе рассматриваются две дисциплины обслуживания: **RCS (Remove Customer in Service)** — воздействие на одну заявку в обслуживании (удаление); **катастрофы** — мгновенная очистка системы от всех положительных заявок. Далее через δ обозначается интенсивность потока отрицательных событий, W и V — времена ожидания в очереди и пребывания в системе (случайные величины), w_1 и v_1 — их первые начальные моменты (математические ожидания).

Исследования СМО с отрицательными заявками обычно связывают с работами Э. Геленбе (G-сети с «сигналами»); ключевой результат — существование стационарных распределений в product-form при марковских предпосылках и ограничениях на механизмы удаления [3, 4]. G-сети обобщались на случаи зависящих от состояния интенсивностей, пакетных перемещений и многоклассовых заявок, сохраняя product-form стационарных распределений при экспоненциальных предпосылках. Обзор подходов к сигналам и стратегиям удаления приведён в [5]. Для одноканальных систем $M/G/1$ сформировалась аналитическая ветка (дополнительная переменная, интегральные уравнения), позволяющая описывать стационарные распределения и временные характеристики [6]; развивались также факторизация Винера–Хопфа [7] и интерпретации удаления работы [8]. Отдельная линия

— *катастрофы*: для $M/G/1$ получены аналоги формулы Поллачека–Хинчина для стационарного распределения числа заявок в системе и времени пребывания, подчёркивающие роль преобразования Лапласа–Стилтьеса (ПЛС), $T^*(s) = \mathbb{E}[e^{-sT}]$, $s \geq 0$, и корректной трактовки «обрезания» траектории катастрофой [9]. Модели с катастрофами обобщались на BMAP/SM/1 [10], дисциплину с разделением процессора [11], дискретное время [12], повторные попытки и отказы прибора [13, 14]. Для многоканальных моделей встречаются марковские частные случаи [15]; интерес к системам с отрицательными заявками и катастрофами сохраняется в последние годы (дискретное время, параллельные очереди, гетерогенные приборы), однако расчёт при произвольном обслуживании в многоканальном случае остаётся нетривиальным.

В данной работе предложено расширение метода Такахаси–Таками для многоканальных СМО с отрицательными заявками: вводятся правила формирования матриц переходов для дисциплин RCS и катастроф, а также алгоритмы вычисления первых начальных моментов времени ожидания и пребывания w_1 и v_1, согласованные с имитационным моделированием. В результате метод позволяет получать стационарные вероятности состояний и временные характеристики в рамках единого алгоритма; обслуживание аппроксимируется H_2 при любом коэффициенте вариации, при необходимости с комплексными параметрами.

Рассматриваемая модель: пуассоновские потоки положительных (λ) и отрицательных (δ) заявок, время обслуживания B аппроксимируется H_2 с заданным средним и коэффициентом вариации, число каналов $n \geq 1$. Коэффициент загрузки $\rho = n^{-1}\lambda\,\mathbb{E}[B]$ предполагается меньше 1. Пространство состояний усекается по уровню $i \leq N-1$ (N — максимальный уровень); сходимость итераций контролируется по изменению $x[i]$ или $t[i]$ с заданным допуском.

Метод Такахаси–Таками: базовые принципы

2.1 Структура микросостояний

Метод основан на представлении системы $M/H_2/n$ в виде марковского процесса с дискретным пространством состояний. Состояние описывается:

- i — число заявок в системе (уровень, level);
- j — число заявок, находящихся в фазе 2 H_2-аппроксимации (с интенсивностью μ_2; микросостояние, microstate).

Для уровня $i < n$ число микросостояний равно $i+1$, для $i \geq n$ постоянно и равно $n+1$.

2.2 Матрицы переходов

Процесс описывается четырьмя типами матриц переходов между микросостояниями соседних уровней: A (поступление), B (уход вниз), C (горизонтальные переходы), D (диагональная матрица интенсивностей выхода).

Пусть $y_1 + y_2 = 1$, где y_1 и y_2 — вероятности выбора фаз 1 и 2 в H_2-аппроксимации, а μ_1, μ_2 — интенсивности экспоненциального обслуживания в фазах. На уровне $i < n$ все заявки находятся в обслуживании, а на уровне $i \geq n$ в обслуживании ровно n заявок, остальные ждут в очереди.

Матрица A (поступление положительных заявок)

Для $i < n$ при приходе новая заявка немедленно занимает свободный канал и выбирает фазу:

$$(i,j) \to (i+1,j) \text{ с интенсивностью } \lambda y_1, \qquad (i,j) \to (i+1,j+1) \text{ с интенсивностью } \lambda y_2.$$

Для $i \geq n$ пришедшая заявка становится в очередь, и микросостояние j (число заявок в фазе 2 среди обслуживаемых) не меняется:

$$(i,j) \to (i+1,j) \text{ с интенсивностью } \lambda.$$

Матрица B (уход на уровень вниз)

Для $i \le n$ уход вниз обусловлен завершением обслуживания одной из i заявок:

$$(i,j) \to (i-1,j) \text{ с интенсивностью } (i-j)\mu_1, \qquad (i,j) \to (i-1,j-1) \text{ с интенсивностью } j\mu_2.$$

Для $i > n$ уход вниз сопровождается тем, что освободившийся канал немедленно занимает заявка из очереди, выбирая фазу 1/2 с вероятностями y_1/y_2. Поэтому для $j \in \{0,1,\dots,n\}$:

$$\begin{aligned}
(i,j) &\to (i-1,j) \text{ с интенсивностью } (n-j)\mu_1 y_1 + j\mu_2 y_2,\\
(i,j) &\to (i-1,j+1) \text{ с интенсивностью } (n-j)\mu_1 y_2, \qquad (j<n),\\
(i,j) &\to (i-1,j-1) \text{ с интенсивностью } j\mu_2 y_1, \qquad (j>0).
\end{aligned}$$

Матрица D (интенсивности выхода)

Диагональные элементы задаются суммарной интенсивностью выхода из микросостояния. Для $i < n$:

$$D(i,j) = \lambda + (i-j)\mu_1 + j\mu_2,$$

а для $i \ge n$:

$$D(i,j) = \lambda + (n-j)\mu_1 + j\mu_2.$$

Примеры для $n = 3$, уровень $i = 2$: матрица A_1 (переход 1→2) — строка $A_1 = \begin{pmatrix}\lambda y_1 & \lambda y_2 & 0\end{pmatrix}$; матрица B_2 (переход 2→1) — блок 3×2:

$$B_2 = \begin{pmatrix} 2\mu_1 & 0 \\ \mu_2 & \mu_1 \\ 0 & 2\mu_2 \end{pmatrix};$$

матрица D_2 (диагональ выхода):

$$D_2 = \begin{pmatrix} \lambda + 2\mu_1 & 0 & 0 \\ 0 & \lambda + \mu_1 + \mu_2 & 0 \\ 0 & 0 & \lambda + 2\mu_2 \end{pmatrix}.$$

2.3 Итерационный алгоритм

Алгоритм Такахаси–Таками основан на итерационном вычислении вспомогательных величин для уровней $i = 0,1,\dots,N-1$ и последующем восстановлении вероятностей уровней $p[i]$. В типичной реализации используются:

- вспомогательные матрицы $G_i = (D_i - C_i)^{-1}$, а также произведения $A_{i-1}G_i$ и $B_{i+1}G_i$;
- итерации по уровням для обновления векторов $t[i]$ и скаляров $x[i], z[i]$ до сходимости;
- восстановление вероятностей уровней по рекуррентной формуле $p[i] = p[i-1]x[i-1]$ при нормировке.

Построение матриц: для каждого уровня i по микросостоянию j заполняются блоки: (1) A — переходы вверх по правилам для $i < n$ и $i \ge n$; (2) B — переходы вниз из завершений обслуживания и выбора фазы при $i > n$; (3) C — горизонтальные переходы; (4) диагональ D так, чтобы сумма по строкам в $A + B + C - D$ равнялась нулю. Размеры блоков определяются числом микросостояний на соседних уровнях.

Выбор N, сходимость, сложность: N выбирают так, чтобы $p[N-1]x[N-1] < \varepsilon$. Итерации прекращают при $\max_i |x[i]^{\text{new}} - x[i]^{\text{old}}| < \varepsilon$. Время одного шага итерации — $O(N(n+1)^3)$, память — $O(N(n+1)^2)$.

2.4 Временные характеристики в базовой модели

Для базовой системы без отрицательных заявок моменты ожидания $w^{(k)}$ могут вычисляться по вероятностям уровней:

$$w^{(k)} = \frac{1}{\lambda^{k+1}} \sum_{j=k+1}^{N-n} j(j-1)\cdots(j-k)\, p[n+j].$$

Моменты пребывания при отсутствии отрицательных заявок получаются стандартной свёрткой моментов ожидания и обслуживания. Для систем с отрицательными воздействиями моменты вычисляются либо по вероятностям уровней (в редуцированных аппроксимациях к $M/G/n$), либо восстанавливаются из ПЛС численным дифференцированием: $(-1)^k T^{*(k)}(0)$ аппроксимируется конечными разностями по s в окрестности нуля (шаг $h \sim 10^{-5}$–10^{-4}); устойчиво ограничены первые 2–4 момента.

Расширение метода для систем с отрицательными заявками

3.1 Общая идея

Расширение основано на модификации матриц переходов базового метода без изменения общей структуры итерационного процесса. В зависимости от дисциплины отрицательного воздействия переопределяются правила построения матриц B и D, а для катастроф — также структура микросостояний.

3.2 Дисциплина RCS (Remove Customer in Service)

При дисциплине RCS отрицательная заявка воздействует на одну заявку в обслуживании, выбирая занятый канал равновероятно. На уровнях $i \le n$ (очереди нет) отрицательные события добавляют к уходам вниз интенсивности

$$(i,j) \to (i-1,j) \text{ с интенсивностью } \delta\frac{i-j}{i}, \qquad (i,j) \to (i-1,j-1) \text{ с интенсивностью } \delta\frac{j}{i},$$

то есть эквивалентны дополнительной «опасности удаления» δ/i для каждой обслуживаемой заявки. Диагональные элементы дополняются интенсивностью отрицательных приходов:

$$D_{\mathrm{RCS}}(i,j) = D_{\mathrm{base}}(i,j) + \delta.$$

На уровнях $i > n$ отрицательное событие также удаляет одну заявку из обслуживания, после чего канал немедленно занимает заявка из очереди и выбирает фазу. Для практического построения матрицы B удобно использовать приближение $\mu_1' = \mu_1 + \delta/n$, $\mu_2' = \mu_2 + \delta/n$ (дополнительная опасность удаления δ/n для каждого канала) и формулы базового случая для $i > n$, подставляя μ_1', μ_2' вместо μ_1, μ_2.

3.3 Дисциплина катастроф

При базовой дисциплине обслуживания катастрофа очищает систему (удаляет все положительные заявки). Для согласования с матричной структурой метода вводится искусственное микросостояние D («состояние катастрофы») на каждом уровне $i > 0$. При поступлении отрицательной заявки система переходит в состояние D, а затем с большой искусственной интенсивностью $\gamma \gg \mu_1, \mu_2, \lambda$ — в нулевой уровень, что моделирует мгновенную очистку.

Структура микросостояний модифицируется так:

- уровень 0: одно состояние $(0,0)$;

- уровни $1 \le i < n+1$: $1 + (i+1)$ микросостояний (первое — D, затем $(i,0),\dots,(i,i)$);
- уровни $i \ge n+1$: $1 + (n+1)$ микросостояний (первое — D, затем $(i,0),\dots,(i,n)$).

Поступление положительных заявок (A) «пропускает» состояние D. Для уровня 0:

$$(0,0) \to (1,0) \text{ с интенсивностью } \lambda y_1, \qquad (0,0) \to (1,1) \text{ с интенсивностью } \lambda y_2.$$

Для уровня $i > 0$ положительный приход переводит (i,j) в соответствующее обычное состояние уровня $i+1$, сохраняя описанную выше структуру (без попадания в D).

Переход в катастрофическое состояние происходит из любого обычного микросостояния уровня $i > 0$ с интенсивностью δ. Состояние D имеет уход в $(0,0)$ с интенсивностью γ, поэтому диагональный элемент для этого состояния (суммарная интенсивность выхода) равен γ. Для уровня $i = 1$ при $n = 3$ микросостояния суть D, $(1,0)$, $(1,1)$; матрица B_1 имеет размер 3×1 (переход в уровень 0).

3.4 Примеры матриц переходов

Переходы по завершению обслуживания (интенсивности μ_1, μ_2) в базе $\mathrm{M}/\mathrm{H}_2/n$ стандартны, поэтому ниже для каждой дисциплины приведены только переходы, возникающие из отрицательных воздействий (интенсивность δ и производные; для катастроф также γ).

RCS (удаление из системы)

При построении матриц B_i к «стандартным» интенсивностям обслуживания добавляются интенсивности отрицательных воздействий. В качестве наглядного примера приведём матрицу B_2 для дисциплины RCS при $n \ge 2$ (уровень $i = 2$, микросостояния $j = 0,1,2$ на уровне 2 и $j = 0,1$ на уровне 1). Тогда

$$B_2 = \begin{pmatrix} 2\mu_1 + \delta & 0 \\ \mu_2 + \delta/2 & \mu_1 + \delta/2 \\ 0 & 2\mu_2 + \delta \end{pmatrix},$$

где, например, элемент $(2,1) \to (1,1)$ равен $\mu_1 + \delta/2$: он складывается из обслуживания фазы 1 и удаления заявки в фазе 1 отрицательным прибытием (равновероятный выбор канала среди двух занятых). Матрица C_i совпадает с базовой (горизонтальные переходы отсутствуют при удалении из системы).

RCS (повторное обслуживание)

При повторном обслуживании отрицательное воздействие не меняет числа заявок в системе, а лишь прерывает обслуживание одной случайной заявки и возвращает её в очередь с рестартом. Это моделируется горизонтальными переходами в матрице C_i: при $i \le n$ заявок каждое микросостояние j переходит в другие микросостояния j' с интенсивностью, определяемой вероятностью выбора нового распределения фаз после рестарта одного из i каналов. Матрица B_i совпадает с базовой (без дополнительных уходов вниз по уровню).

Пример матрицы C_2 при уровне $i = 2$ (микросостояния $j = 0,1,2$ — число заявок в фазе 2): прерывание случайного канала даёт переход $j \to j+1$ с интенсивностью $\delta\,\frac{i-j}{i}\,y_2$ (рестарт в фазе 2) и $j \to j-1$ с интенсивностью $\delta\,\frac{j}{i}\,y_1$ (рестарт в фазе 1). Тогда

$$C_2 = \begin{pmatrix} 0 & \delta y_2 & 0 \\ \delta y_1/2 & 0 & \delta y_2/2 \\ 0 & \delta y_1 & 0 \end{pmatrix} \quad \text{(только внедиагональные переходы; диагональ дополняется в } D_i\text{).}$$

Катастрофы (очищение системы)

Для катастроф с очищением вводится дополнительное микросостояние D на каждом уровне $i > 0$. Приведём пример матрицы B_1 при $n \geq 1$ (уровень $i = 1$, микросостояния: D, $(1, 0)$, $(1, 1)$; переход в уровень 0):

$$B_1 = \begin{pmatrix} \gamma \end{pmatrix}.$$

Здесь первая строка соответствует состоянию D, из которого система мгновенно переходит в $(0, 0)$ с интенсивностью $\gamma \gg \mu_1, \mu_2, \lambda$. Обычные состояния $(1, 0)$ и $(1, 1)$ уходят в уровень 0 по стандартным интенсивностям μ_1 и μ_2 (не показаны, так как не зависят от катастроф).

Матрица C_1 моделирует горизонтальные переходы из обычных состояний в катастрофическое состояние D:

$$C_1 = \begin{pmatrix} 0 & 0 & 0 \\ \delta & 0 & 0 \\ \delta & 0 & 0 \end{pmatrix},$$

где элементы $(1, 0) \to D$ и $(1, 1) \to D$ равны δ (поступление катастрофы мгновенно переводит систему в состояние D).

Катастрофы (повторное обслуживание)

При повторном обслуживании катастрофа не очищает систему, а прерывает все обслуживания и возвращает заявки в очередь с рестартом. Состояние D не вводится; вместо этого используются горизонтальные переходы в матрице C_i между микросостояниями с интенсивностью δ, моделирующие перераспределение фаз всех $m = \min(i, n)$ занятых каналов после катастрофы. Каждый из m каналов независимо выбирает фазу 1 или 2 с вероятностями y_1 и y_2, поэтому из любого микросостояния (i, j) система переходит в (i, k) с интенсивностью $\delta \cdot \binom{m}{k} y_2^k y_1^{m-k}$ (независимо от j). Матрицы B_i совпадают с базовыми.

Пример матрицы C_2 при уровне $i = 2$ (микросостояния $j = 0, 1, 2$; $m = 2$):

$$C_2 = \delta \begin{pmatrix} y_1^2 & 2y_1y_2 & y_2^2 \\ y_1^2 & 2y_1y_2 & y_2^2 \\ y_1^2 & 2y_1y_2 & y_2^2 \end{pmatrix}.$$

В общем случае для уровня i при $m = \min(i, n)$ матрица C_i имеет размер $(m + 1) \times (m + 1)$, и все строки совпадают: $C_i(j, k) = \delta \binom{m}{k} y_2^k y_1^{m-k}$ при $j, k \in \{0, \dots, m\}$. Диагональ уровня в блоке D_i дополняется интенсивностью δ (суммарный выход из микросостояния из-за катастрофы).

Расчёт временных характеристик

4.1 Дисциплина RCS

4.1.1 RCS с удалением из системы

Для типичной заявки, находящейся в обслуживании при m занятых каналах, интенсивность удаления аппроксимируется как $r(m) = \delta/m$. При дисциплине «удаление из обслуживания» время в обслуживании:

$$S = \min(B, Y), \qquad Y \sim \mathrm{Exp}(r),$$

и ПЛС времени обслуживания $S^*(s) = \frac{r}{s+r} + \frac{s}{s+r}\beta(s + r)$, где $\beta(s) = \mathbb{E}[e^{-sB}]$. Дальнейшие моменты W, V вычисляются через ПЛС и численное дифференцирование в нуле.

4.1.2 RCS с повторным обслуживанием

Если отрицательное событие интерпретируется как прерывание обслуживания с возвратом заявки в очередь (рестарт обслуживания и повторная генерация длительности), расчёт сводится к *эффективному времени обслуживания* B_{eff} при пуассоновских рестартах. Пусть B имеет ПЛС $\beta(s)$, рестарты

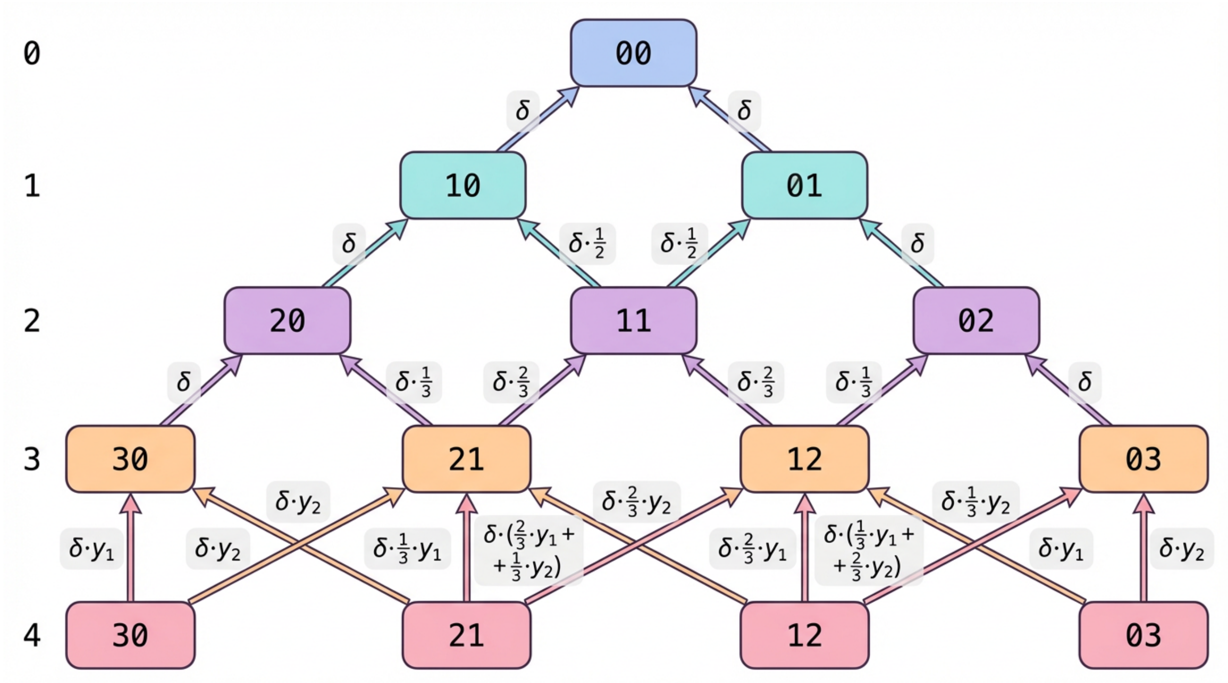


Рис. 1: Переходы, вызванные отрицательными воздействиями, для дисциплины RCS (удаление заявки из обслуживания), $n = 3$, уровни $0, \dots, n$. Для $i > n$ в расчётах используется приближение через μ_1', μ_2' (см. текст).

— с интенсивностью r. Для модели «прерывание–повтор» ПЛС времени завершения обслуживания:

$$B^*_{\text{eff}}(s) = \beta(s+r) \cdot \frac{s+r}{s + r\,\beta(s+r)}.$$

Для RCS типичная заявка при m занятых каналах имеет интенсивность рестарта $r(m) = \delta/m$; в расчёте r оценивается самосогласованно (итерационно), начиная с приближения $r \approx \delta/n$. Система аппроксимируется стандартной $M/G/n$ без отрицательных выходов с распределением B_{eff}; тогда $V = W + B_{\text{eff}}$, $q = 1$, $V_{\text{served}} \equiv V$, $V_{\text{broken}} \equiv 0$. Моменты W и V вычисляются базовым алгоритмом Такахаси–Таками для $M/G/n$ с распределением B_{eff}.

4.2 Дисциплина катастроф

4.2.1 Катастрофы с очищением

Пусть $Y \sim \text{Exp}(\delta)$ — время до ближайшей катастрофы после прихода типичной заявки. Для согласования с имитационным моделированием вводятся гипотетические величины W_0 и $Z_0 = W_0 + B$, соответствующие динамике, в которой катастрофы *после прихода данной заявки* отключены (но состояние при приходе берётся из стационара исходной системы). Тогда

$$W = \min(W_0, Y), \qquad V = \min(Z_0, Y),$$

и для любой независимой пары T и $Y \sim \text{Exp}(\delta)$ справедливо:

$$\mathbb{E}[e^{-s\min(T,Y)}] = \frac{\delta}{s+\delta} + \frac{s}{s+\delta} T^*(s+\delta).$$

Отсюда:

$$W^*(s) = \frac{\delta}{s+\delta} + \frac{s}{s+\delta} W_0^*(s+\delta), \qquad V^*(s) = \frac{\delta}{s+\delta} + \frac{s}{s+\delta} Z_0^*(s+\delta).$$

Вероятность успешного обслуживания:

$$p_{\text{served}} = \mathbb{P}(Z_0 < Y) = \mathbb{E}[e^{-\delta Z_0}] = Z_0^*(\delta),$$

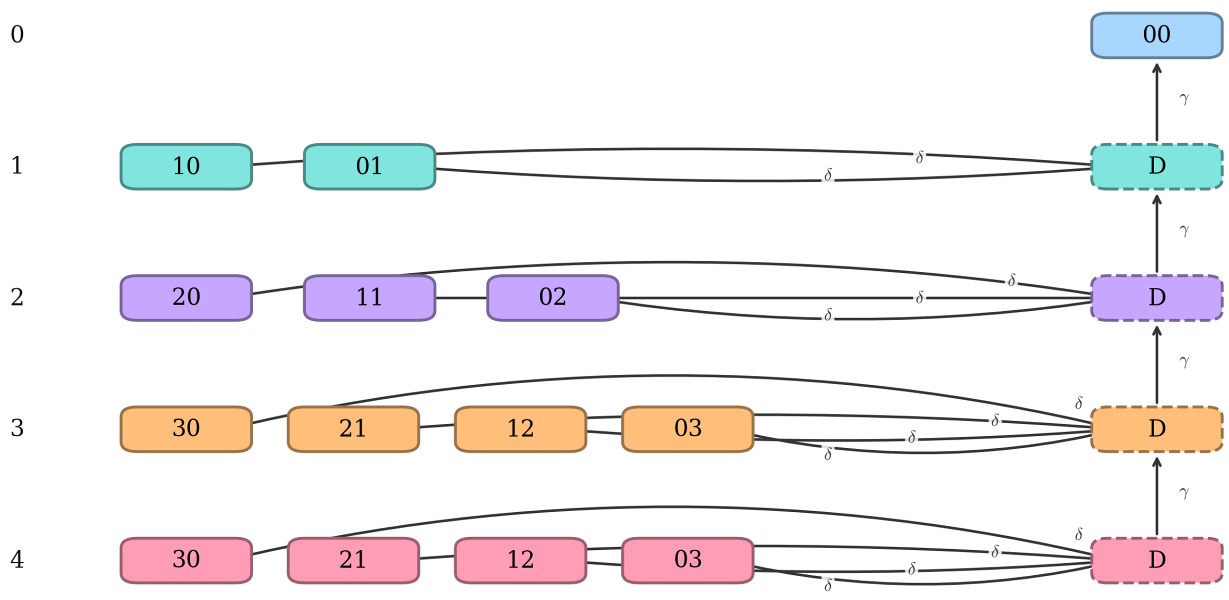


Рис. 2: Переходы, вызванные катастрофическими воздействиями (интенсивности δ и γ), $n = 3$, уровни $0, \dots, n+1$.

а ПЛС условных распределений:

$$\mathbb{E}[e^{-sV} \mid \text{served}] = \frac{Z_0^*(s+\delta)}{Z_0^*(\delta)}, \qquad \mathbb{E}[e^{-sV} \mid \text{broken}] = \frac{\delta}{s+\delta} \cdot \frac{1 - Z_0^*(s+\delta)}{1 - Z_0^*(\delta)}.$$

4.2.2 Катастрофы с повторным обслуживанием

При дисциплине «катастрофа прерывает обслуживание всех заявок в каналах и возвращает их в очередь с рестартом» прерванных в смысле удалённых заявок нет: $q = 1$, $V_{\text{served}} \equiv V$, $V_{\text{broken}} \equiv 0$. Эффективное время обслуживания B_{eff} при пуассоновских рестартах с интенсивностью δ (каждая катастрофа прерывает любую заявку в канале):

$$B_{\text{eff}}^*(s) = \beta(s+\delta) \cdot \frac{s+\delta}{s + \delta\, \beta(s+\delta)}.$$

Система в приближении сводится к стандартной $M/G/n$ без отрицательных выходов с распределением B_{eff}; $V = W + B_{\text{eff}}$, моменты получаются численным дифференцированием ПЛС (с гамма-аппроксимацией $\beta(s)$ по первым двум моментам B).

Численные результаты и сравнение с имитационным моделированием

5.1 Параметры тестов

В качестве базового примера использовались параметры: $n = 3$, $\lambda_{\text{pos}} = 1.0$, $\delta = 0.3$, $\rho = 0.7$, коэффициент вариации времени обслуживания $CV = 1{,}2$, размер выборки в имитационном моделировании 1,000,000. Дополнительно, чтобы показать применимость метода при различных коэффициентах вариации времени обслуживания, в таблице 1 приведены численные значения для $CV = 0{,}8$ и $CV = 1{,}2$.

5.2 Сравнение численного метода и ИМ по w_1 и v_1

В таблице 1 приведены первые начальные моменты времени ожидания (w_1) и пребывания (v_1) для четырёх сценариев при двух значениях коэффициента вариации времени обслуживания; для каждого значения CV приведено сравнение с имитационным моделированием (ИМ).

	w_1			v_1		
Сценарий	Числ	ИМ	Отн.%	Числ	ИМ	Отн.%
Коэффициент вариации $CV = 0{,}8$						
RCS (удаление)	0.4205	0.4111	2.29	2.0952	2.0849	0.50
RCS (повторное обслуж.)	1.2067	1.1650	3.58	3.4030	3.4062	−0.10
Катастрофы (очищение)	0.1964	0.1918	2.42	1.4853	1.4829	0.16
Катастрофы (повторное обслуж.)	1.8268	1.8513	−1.32	4.1527	4.1725	−0.47
Коэффициент вариации $CV = 1{,}2$ (базовое значение)						
RCS (удаление)	0.3881	0.3968	−2.18	1.9127	1.9190	−0.33
RCS (повторное обслуж.)	1.0627	1.0273	3.44	3.0547	3.0674	−0.41
Катастрофы (очищение)	0.1613	0.1660	−2.82	1.3096	1.3160	−0.49
Катастрофы (повторное обслуж.)	0.7583	0.7781	−2.55	2.6441	2.6732	−1.09

Таблица 1: Первые начальные моменты времени ожидания w_1 и пребывания v_1 при фиксированных $n = 3$, $\lambda_{\text{pos}} = 1.0$, $\delta = 0.3$, $\rho = 0.7$ и различных коэффициентах вариации времени обслуживания. Числ — численный расчёт, ИМ — имитационное моделирование, Отн.% — относительная погрешность.

В таблице 2 приведены среднеквадратичные отклонения σ_W и σ_V для тех же сценариев и значений коэффициента вариации.

	σ_W			σ_V		
Сценарий	Числ	ИМ	Отн.%	Числ	ИМ	Отн.%
Коэффициент вариации $CV = 0{,}8$						
RCS (удаление)	0.9477	0.9233	2.57	1.789	1.767	1.23
RCS (повторное обслуж.)	1.976	1.984	0.42	2.743	2.782	1.41
Катастрофы (очищение)	0.5639	0.5603	0.65	1.305	1.302	0.19
Катастрофы (повторное обслуж.)	2.448	2.412	1.50	3.185	3.149	1.11
Коэффициент вариации $CV = 1{,}2$ (базовое значение)						
RCS (удаление)	1.003	1.008	0.50	2.053	2.05	0.14
RCS (повторное обслуж.)	2.033	2.017	0.81	3.074	3.074	0.01
Катастрофы (очищение)	0.5621	0.5743	2.17	1.418	1.422	0.25
Катастрофы (повторное обслуж.)	1.643	1.636	0.40	2.702	2.697	0.18

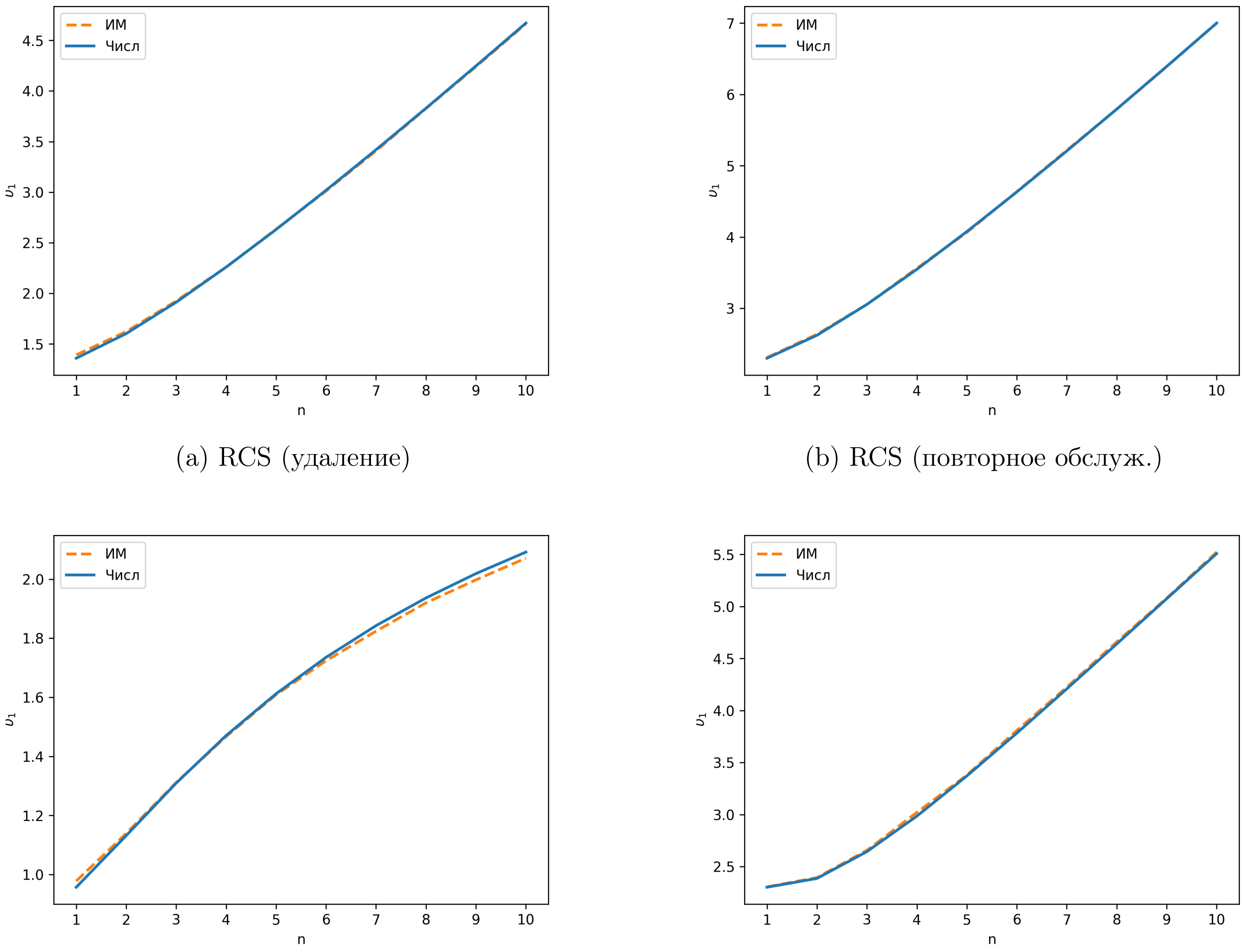
Таблица 2: Среднеквадратичные отклонения σ_W и σ_V при фиксированных $n = 3$, $\lambda_{\text{pos}} = 1.0$, $\delta = 0.3$, $\rho = 0.7$ и различных коэффициентах вариации времени обслуживания.

Заметны две устойчивые закономерности. Во-первых, сценарии с *удалением* (RCS) и *очищением* (катастрофы) дают наименьшие значения w_1 и v_1: отрицательные воздействия в этих постановках уменьшают «эффективную загрузку» системы за счёт преждевременного завершения обслуживания и/или мгновенного сброса заявок. Во-вторых, при *повторном обслуживании* (рестартах) значения w_1 и v_1 выше при малом CV, так как рестарт приводит к потере уже выполненной части обслуживания. При этом с ростом CV значения w_1 и v_1 снижаются во всех четырёх сценариях: в модели с пуассоновскими прерываниями характеристики определяются Лапласовым образом $\beta(\cdot)$ распределения обслуживания (в частности, через $\beta(\delta/n)$ и/или $\beta(s + \delta)$), и при увеличении вариативности гамма-распределения возрастает вероятность «быстрых» реализаций обслуживания после рестарта, что уменьшает *эффективное* время обслуживания и, следовательно, снижает средние времена ожидания и пребывания.

Согласование численного метода с имитационным моделированием в таблицах 1–2 в целом хорошее. В таблице 1 относительная погрешность для v_1, как правило, не превышает 1–2%, а для w_1 — нескольких процентов (в приведённых тестах максимум порядка 3,6%). Для среднеквадратичных отклонений (таблица 2) погрешности также остаются на уровне нескольких процентов. Наибольшая чувствительность наблюдается для w_1 и σ_W при больших CV, что связано с влиянием хвостов распределений и приближениями, используемыми в расчёте (в частности, аппроксимацией интенсивностей рестартов/удалений и моментной аппроксимацией распределений).

5.3 Графики параметрических зависимостей

На рисунке 3 приведены зависимости первого начального момента времени пребывания v_1 от числа каналов n для четырёх сценариев. Видно, что увеличение числа каналов приводит к росту v_1 при фиксированной загрузке ρ (при задании среднего времени обслуживания $\mathbb{E}[B] = n\rho/\lambda$). Влияние коэффициента вариации времени обслуживания показано в таблице 1.



(a) RCS (удаление)

(b) RCS (повторное обслуж.)

(c) Катастрофы (очищение)

(d) Катастрофы (повторное обслуж.)

Рис. 3: Зависимость v_1 от числа каналов n для четырёх дисциплин обслуживания отрицательных заявок (пунктир — ИМ, сплошная — численный расчёт).

Заключение

Предложено расширение метода Такахаси–Таками для расчёта многоканальных систем с отрицательными заявками. Показано, как модифицировать матрицы переходов для дисциплин RCS и катастроф, сохраняя структуру итерационного алгоритма. Временные характеристики получают через ПЛС и численное дифференцирование; для катастроф корректный расчёт времени пребывания использует операцию минимума с экспоненциальным временем до ближайшей катастрофы после прихода.

Численные эксперименты показывают хорошее согласование с имитационным моделированием по вероятностям состояний и по временным характеристикам (первые моменты и среднеквадратичные отклонения), а вычислительное время существенно меньше по сравнению с прямым имитационным моделированием. Это делает метод пригодным для параметрического анализа и оптимизации. Реализация доступна в [16].

Список литературы

[1] Y. Takahashi и Y. Takami. «A Numerical Method for the Steady-State Probabilities of a GI/G/c Queuing System in a General Class». В: *Journal of the Operations Research Society of Japan* 19.2 (1976), с. 147—157.

[2] Ю. И. Рыжиков. *Численные методы теории очередей.* Санкт-Петербург: Лань, 2018.

[3] Erol Gelenbe. «Product-form queueing networks with negative and positive customers». В: *Journal of Applied Probability* 28.3 (1991), с. 656—663.

[4] Erol Gelenbe и Rudolf Schassberger. «Stability of product form queueing networks with negative customers». В: *Journal of Applied Probability* 29.4 (1992), с. 890—901.

[5] J. R. Artalejo. «G-networks: A versatile approach for work removal in queueing networks». В: *European Journal of Operational Research* 126.2 (2000), с. 233—249.

[6] Peter G. Harrison и E. Pitel. «The M/G/1 queue with negative customers». В: *Advances in Applied Probability* 28.2 (1996), с. 540—566.

[7] N. Bayer и O. J. Boxma. «Wiener–Hopf analysis of an M/G/1 queue with negative customers and of a related class of random walks». В: *Queueing Systems* 23.1-4 (1996), с. 301—316.

[8] Richard J. Boucherie и Onno J. Boxma. «The Workload in the M/G/1 Queue with Work Removal». В: *Probability in the Engineering and Informational Sciences* 10.2 (1996), с. 261—277.

[9] Gautam Jain и Karl Sigman. «A Pollaczek–Khintchine formula for M/G/1 queues with disasters». В: *Journal of Applied Probability* 33.4 (1996), с. 1191—1200.

[10] Alexander Dudin и S. Nishimura. «A BMAP/SM/1 queueing system with Markovian arrival input of disasters». В: *Journal of Applied Probability* 36.3 (1999), с. 868—881.

[11] Q.-L. Li и C. Lin. «The M/G/1 processor-sharing queue with disasters». В: *Computers & Mathematics with Applications* 51.6-7 (2006), с. 987—998.

[12] Ignacio Atencia и Pedro Moreno. «The discrete-time Geo/Geo/1 queue with negative customers and disasters». В: *Computers & Operations Research* 31.9 (2004), с. 1537—1550.

[13] J. R. Artalejo и A. Gómez-Corral. «On a single server queue with negative arrivals and request repeated». В: *Journal of Applied Probability* 35.3 (1998), с. 907—918.

[14] J. Wang, B. Liu и J. Li. «Transient analysis of an M/G/1 retrial queue subject to disasters and server failures». В: *European Journal of Operational Research* 189.3 (2008), с. 1338—1353.

[15] S. Dharmaraja и R. Kumar. «Transient solution of a Markovian queueing model with heterogeneous servers and catastrophes». В: *Opsearch* 52.4 (2015), с. 810—824.

[16] Р. С. Хабаров и В. А. Лохвицкий. *Most-Queue: Библиотека методов расчёта СМО с отрицательными заявками.* `https://github.com/xabarov/most-queue/tree/main/works/negative_queues`. Дата обращения: 14.02.2026. 2026.