

**О ПИКОВОМ ВОЗРАСТЕ ИНФОРМАЦИИ В ГРУППЕ ПЕРЕДАЧИ,
МОДЕЛИРУЕМОЙ СИСТЕМОЙ $PH/PN/1/R/LCFS$ С ЗАХВАТОМ
ПРИБОРА И ПЕРЕОБСЛУЖИВАНИЕМ ЗАЯВОК**

Гайдамака Ю.В.^{*,}, Матюшенко С.И.^{*}, Самуйлов К.Е.^{*,**}**

^{*}Российский университет дружбы народов, г. Москва

^{**}Институт проблем информатики Федерального исследовательского центра
«Информатика и управление» РАН, г. Москва

Поступила в редакцию 09.01.2024, после переработки 29.01.2024.

Данная статья продолжает цикл работ авторов, посвященных проблеме возраста информации (Age of Information, AoI) – метрики, используемой в информационных системах для количественной оценки «свежести» информации, доставляемой в центр управления от периферийных источников. В работе рассматривается двухузловая группа передачи информации, состоящая из узла-отправителя, узла-получателя и канала связи между ними. Процесс передачи информации моделируется посредством однолинейной системы обслуживания конечной емкости с распределениями фазового типа, которая в обозначениях Кендалла кодируется как $PH/PN/1/r$. При этом учитываются особые условия передачи, состоящие в том, что поступающий в систему пакет, минуя очередь, сразу отправляется на передачу, захватывая канал у предыдущего пакета, если тот не завершил передачу. Пакет, передача которого была прервана, перемещается на первое место в очереди и после освобождения канала заново осуществляет попытку передачи. Для данной системы получено выражение для преобразования Лапласа–Стилтьеса стационарной функции распределения пикового возраста информации и его среднего значения. Проведено численное исследование зависимости пикового возраста информации от загрузки системы. Корректность аналитических результатов проверена путем их сравнения с результатами имитационного моделирования.

Ключевые слова: возраст информации, пиковый возраст информации, система массового обслуживания, распределение фазового типа.

Вестник ТвГУ. Серия: Прикладная математика. 2024. № 1. С. 37–52.
<https://doi.org/10.26456/vtpmk701>

Введение

Понятие «возраст информации» было введено для количественной оценки «свежести» знаний об удаленном объекте. В специальной литературе для обозначения этого понятия стали использовать аббревиатуру AoI от английского «Age of

Information». Со временем это понятие превратилось в метрику, отражающую качество предоставляемых услуг, для которых фактор времени играет наиважнейшую роль. Как правило, в системах подобного рода обмен информацией между оконечными устройствами и центром управления обеспечивает контроль за удаленными процессами и повышает их безопасность за счет эффективного использования информации о них [1–3].

Наиболее удобным средством для исследования проблемы возраста информации является аппарат теории массового обслуживания. Перечень работ, использующих этот аппарат, можно найти, например, в [4]. Однако следует отметить, что большинство исследователей ограничиваются простыми моделями, например, с экспоненциальным распределением времени между моментами генерации пакетов в узле-отправителе и экспоненциальным [5] либо детерминированным [6] распределением длительности обработки информации в узле-получателе. Такие модели позволяют получить лишь грубую оценку возраста информации, поскольку однопараметрические распределения не дают возможности учесть все особенности протоколов работы современных систем диспетчерского управления. В настоящей работе процесс передачи информации от узла-отправителя к узлу-получателю моделируется посредством системы массового обслуживания с распределениями фазового типа, выбор параметров фаз которого позволяет гибко моделировать сложные зависимости, возникающие в современных системах передачи данных.

Однако заметим, что авторы данной работы не являются первыми, кто решил использовать распределения фазового типа для решения проблемы «возраста информации». Так в работе [7] передача информации моделировалась с помощью системы $M/PH/1/1$ с k источниками, $1 \leq k < \infty$, и с возможностью замены заявки в очереди, поступившей от источника n , на более «свежую» заявку, поступившую от источника m с наперед заданной вероятностью $p(n, m)$, $n, m = 1, \dots, k$.

В работе [8] передача информации моделировалась двумя способами: с помощью системы $PH/PH/1/0$ с вероятностным захватом прибора вновь поступающей заявкой и системой $M/PH/1/1$ с вероятностной заменой заявки в накопителе более «свежей» заявкой.

В данной работе мы также рассматриваем систему с захватом прибора, но с более сложным алгоритмом функционирования системы. Усложнение состоит в том, что в нашей системе заявка, обслуживание которой было прервано, не теряется, а становится на первое место в очереди и затем переобслуживается.

1. Системная модель

Рассмотрим группу передачи (ГП) информации, состоящую из узла-отправителя (УО), узла-получателя (УП) и канала связи между ними (Рис. 1).

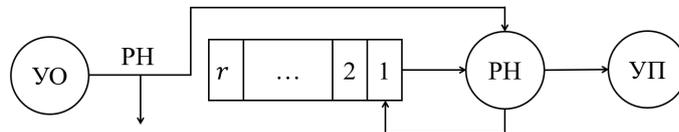


Рис. 1: Двухузловая ГП

Процесс передачи информации из УО в УП будем моделировать с помощью однолинейной системы массового обслуживания (СМО) с накопителем конечной емкости r . При этом будем полагать, что поток заявок является рекуррентным с функцией распределения (ФР) $A(t)$ фазового типа:

$$A(t) = 1 - \alpha^T e^{\Lambda t} \mathbf{1}, \quad t > 0, \quad \alpha^T \mathbf{1} = 1,$$

допускающей неприводимое PH -представление (α, Λ) порядка l .

Длительности обслуживания заявок имеют ФР фазового типа $B(t)$:

$$B(t) = 1 - \beta^T e^{\mathbf{M}t} \mathbf{1}, \quad t > 0, \quad \beta^T \mathbf{1} = 1,$$

допускающую неприводимое PH -представление (β, \mathbf{M}) порядка m и не зависящую от длительности генерации заявок.

Дисциплина обслуживания заявок предполагает обратный порядок обслуживания с захватом прибора и повторным обслуживанием прерванной заявки с тем же распределением $B(t)$. Другими словами, каждая новая заявка обладает абсолютным приоритетом, а в очередь размещаются только прерванные заявки, причем последняя прерванная заявка помещается на первое место в очереди.

Заявка, заставшая в момент поступления накопитель занятым, покидает систему и снова в нее не возвращается.

В соответствии с обозначениями Кенделла будем кодировать нашу систему как $PH/PH/1/r/LCFS/PRR$, где PRR -аббревиатура от Preemptive Resume Repetition.

2. Построение математической модели

Цель нашего исследования – найти распределение пикового возраста информации, передаваемой из УО в УП. Как известно [9] под пиковым возрастом Z^{n-1} ($n-1$)-ой заявки, поступающей в систему, понимают сумму:

$$Z^{n-1} = G^n + T^n,$$

где G^n – интервал времени между поступлениями $(n-1)$ -ой и n -ой заявки, а T^n – время пребывания n -ой заявки в системе. Либо используют другой вариант:

$$Z^{n-1} = T^{n-1} + D^n,$$

где D^n – интервал времени между выходами из системы n -ой и $(n-1)$ -ой заявки. Оба варианта, естественно, приводят к одинаковому результату. Однако в нашем случае порядок входа заявок в систему и порядок выхода из нее различаются. Сам принцип функционирования направлен на то, чтобы наиболее «свежая» информация передавалась в первую очередь. В результате часть заявок передается без прерывания, а другая часть попадает на переобслуживание. Далее заявки, переданные без прерывания, будут называться «успешными».

Не смотря на то, что «неуспешные» заявки приходят в УП с «устаревшей» информацией, их передача имеет смысл с точки зрения «полноты картины». Ведь не исключена ситуация, когда в реальной технической системе из-за «увлечения свежей информацией» могут быть пропущены важные сведения, содержащиеся в

«неуспешных» заявках, что приведет к какому-либо сбою или катастрофе. И для того, чтобы разобраться в произошедшем, потребуется весь архив информации.

Однако, учитывая, что главная роль отводится все же «успешным» заявкам, мы в дальнейшем будем исследовать распределение пикового возраста информации, передаваемой именно «успешным» заявками и для расчетов используется второй вариант, т.е. под пиковым возрастом $(n-1)$ -ой «успешной» заявки $Z^{n-1,+}$ будем понимать сумму:

$$Z^{n-1,+} = T^{n-1,+} + D^{n,+},$$

где знак «+» здесь и далее является признаком «успешности», в то время как знак «-» – «неуспешности».

Учитывая вероятностную интерпретацию распределений фазового типа, функционирование рассматриваемой системы опишем однородным марковским процессом (МП) $X(t)$, $t \geq 0$, над множеством состояний:

$$\begin{aligned} \mathcal{X} &= \mathcal{X}_0 + \bigcup_{k=1}^{r+1} \mathcal{X}_k^+ + \bigcup_{k=1}^r \mathcal{X}_k^-, \\ \mathcal{X}_0 &= \{(i, 0), \quad i = \overline{1, l}\}, \\ \mathcal{X}_k^+ &= \{(i, k, j)^+, \quad i = \overline{1, l}, \quad j = \overline{1, m}\}, \quad k = \overline{1, r+1}, \\ \mathcal{X}_k^- &= \{(i, k, j)^-, \quad i = \overline{1, l}, \quad j = \overline{1, m}\}, \quad k = \overline{1, r}. \end{aligned}$$

Здесь для произвольного момента времени t состояние $(i, 0)$ означает, что в момент времени t система пуста, а генерация новой заявки происходит по фазе i , $i = \overline{1, l}$. Состояние $(i, k, j)^+$ означает, что в системе k заявок, генерация новой заявки происходит на фазе i , а обслуживание «успешной» заявки – на фазе j . Состояние $(i, k, j)^-$ отличается от предыдущего тем, что на приборе на фазе j происходит обслуживание «неуспешной» заявки.

Все состояния МП $X(t)$ сообщаются между собой. Поэтому предельные вероятности:

$$p_x = \lim_{t \rightarrow \infty} P\{X(t) = x\}, \quad x \in \mathcal{X},$$

существуют, строго положительны, не зависят от начального распределения и совпадают со стационарными вероятностями.

Введем векторы:

$$\begin{aligned} \mathbf{p}_0^T &= (p(1, 0), \dots, p(l, 0)), \\ \mathbf{p}_{k,+}^T &= (p(1, k, 1)^+, \dots, p(l, k, 1)^+, p(1, k, 2)^+, \dots, p(l, k, m)^+), \quad k = \overline{1, r+1}, \\ \mathbf{p}_{k,-}^T &= (p(1, k, 1)^-, \dots, p(l, k, 1)^-, p(1, k, 2)^-, \dots, p(l, k, m)^-), \quad k = \overline{1, r}. \end{aligned}$$

Стационарные вероятности $\{\mathbf{p}_0; \mathbf{p}_{k,+}, k = \overline{1, r+1}; \mathbf{p}_{k,-}, k = \overline{1, r}\}$ являются единственным решением системы уравнений равновесия (СУР):

$$\mathbf{0}^T = \mathbf{p}_0^T \mathbf{\Lambda} + [\mathbf{p}_{1,-}^T + \mathbf{p}_{1,+}^T](\mathbf{I} \otimes \boldsymbol{\mu}), \quad (1)$$

$$\mathbf{0}^T = \mathbf{p}_0^T (\lambda \boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) + \mathbf{p}^T (\mathbf{\Lambda} \oplus \mathbf{M}), \quad (2)$$

$$\mathbf{0}^T = [\mathbf{p}_{k-1,-}^T + \mathbf{p}_{k-1,+}^T](\lambda \boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) + \mathbf{p}_{k,+}^T (\mathbf{\Lambda} \oplus \mathbf{M}), \quad k = \overline{1, r}, \quad (3)$$

$$\mathbf{0}^T = [\mathbf{p}_{k+1,-}^T + \mathbf{p}_{k+1,+}^T](\mathbf{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T) + \mathbf{p}_{k,-}^T(\boldsymbol{\Lambda} \oplus \mathbf{M}), \quad k = \overline{1, r-1}, \quad (4)$$

$$\mathbf{0}^T = \mathbf{p}_{r+1,+}^T(\mathbf{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T) + \mathbf{p}_{r,-}^T(\boldsymbol{\Lambda} \oplus \mathbf{M}), \quad (5)$$

$$\mathbf{0}^T = [\mathbf{p}_{k,-}^T + \mathbf{p}_{k,+}^T](\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) + \mathbf{p}_{r+1,+}^T(\boldsymbol{\Lambda} \oplus \mathbf{M} + \boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \mathbf{I}) \quad (6)$$

с условием нормировки

$$\mathbf{p}_0^T \mathbf{1} + \sum_{k=1}^{r+1} \mathbf{p}_{k,+}^T \mathbf{1} + \sum_{k=1}^r \mathbf{p}_{k,-}^T \mathbf{1} = \mathbf{1}. \quad (7)$$

Здесь и далее $\boldsymbol{\lambda} = -\boldsymbol{\Lambda}\mathbf{1}$, $\boldsymbol{\mu} = -\mathbf{M}\mathbf{1}$. Символы \otimes и \oplus означают кронекерово произведение и кронекерову сумму матриц соответственно.

Заметим, что состояния $(i, r+1, j)^-$, $i = \overline{1, l}$, $j = \overline{1, m}$ не существует, т.к. в системе может появиться $r+1$ заявка только за счет прихода новой заявки, когда в системе уже было k заявок. Эта новая заявка «успешная», т.к. из-за отсутствия места в накопителе, следующая заявка войти не может, поэтому мы попадем в одно из состояний $(i, r+1, j)^+$, $i = \overline{1, l}$, $j = \overline{1, m}$

Далее введем макросостояния:

$$(i, k, j) = \begin{cases} (i, k, j)^+ \cup (i, k, j)^-, & k = \overline{1, r}, \\ (i, k, j)^+, & k = r+1. \end{cases}$$

Состояние (i, k, j) означает, что в системе в момент времени t имеется k заявок, генерация новой заявки происходит на фазе i , а обслуживание – на фазе j , причем не важно, какая заявка («успешная» или «неуспешная») обслуживается на приборе.

Обозначим через $p(i, k, j)$ стационарные вероятности введенных макросостояний и введем векторы:

$$\mathbf{p}_k^T = (p(1, k, 1); \dots; p(l, k, 1); p(1, k, 2); \dots; p(l, k, m)). \quad (8)$$

Из (4) и (5) с учетом того, что $\mathbf{p}_{r+1,+}^T = \mathbf{p}_{r+1}^T$ получим

$$\mathbf{p}_{k,-}^T = \mathbf{p}_{k+1}^T(\mathbf{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T)(\boldsymbol{\Lambda} \oplus \mathbf{M})^{-1}, \quad k = \overline{1, r}. \quad (9)$$

При этом невырожденность матрицы $(\boldsymbol{\Lambda} \oplus \mathbf{M})$ следует из неприводимости PH -представления [10].

Очевидно, что:

$$\mathbf{p}_{k,+}^T = \mathbf{p}_k^T - \mathbf{p}_{k,-}^T. \quad (10)$$

Следовательно, согласно (9) и (10) для определения стационарного распределения $\{p_x, x \in \mathcal{X}\}$ нет необходимости знать \mathbf{p}_k , $k = \overline{1, r+1}$.

Введем обозначения:

$$\widetilde{\boldsymbol{\Lambda}} = -(\boldsymbol{\Lambda} \oplus \mathbf{M}) - \mathbf{1}\boldsymbol{\alpha}^T \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T,$$

$$\widetilde{\mathbf{M}} = -(\boldsymbol{\Lambda} \oplus \mathbf{M}) + \boldsymbol{\Lambda} \otimes \mathbf{1}\boldsymbol{\beta}^T,$$

$$\widetilde{\widetilde{\mathbf{M}}} = \boldsymbol{\Lambda} \oplus \mathbf{M} + \boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \mathbf{I},$$

$$\begin{aligned}
\mathbf{W}_0 &= -(\mathbf{A} \otimes \boldsymbol{\beta}^T) \widetilde{\mathbf{M}}^{-1}, \\
\mathbf{W} &= \widetilde{\mathbf{A}} \widetilde{\mathbf{M}}^{-1}, \\
\mathbf{W}_r &= -(\boldsymbol{\lambda} \alpha^T \otimes \mathbf{1} \boldsymbol{\beta}^T) \widetilde{\mathbf{M}}^{-1}, \\
\mathbf{Z} &= \mathbf{W}_0 \mathbf{W}^{r-1} (\mathbf{W}_r \widetilde{\mathbf{M}} - \widetilde{\mathbf{A}}) (\mathbf{I} \otimes \mathbf{1}), \\
\mathbf{v} &= (\mathbf{I} + \sum_{k=1}^r \mathbf{W}_0 \mathbf{W}^{k-1} + \mathbf{W}_0 \mathbf{W}^{r-1} \mathbf{W}_r) \mathbf{1}.
\end{aligned}$$

Заметим, что доказательство неприводимости матриц $\widetilde{\mathbf{M}}$ и $\widetilde{\mathbf{M}}$ можно найти, например, в [10], глава 5.

В работе [11] был разработан рекуррентный матричный алгоритм для расчета стационарного распределения $\{\mathbf{p}_k, k = \overline{0}, r+1\}$ и доказана следующая

Теорема 1. *Стационарное распределение $\{\mathbf{p}_k, k = \overline{0}, r+1\}$ представляется в виде:*

$$\mathbf{p}_k^T = \begin{cases} \mathbf{p}_0^T \mathbf{W}_0 \mathbf{W}^{k-1}, & k = \overline{1}, r \\ \mathbf{p}_0^T \mathbf{W}_0 \mathbf{W}^{r-1} \mathbf{W}_r, & k = r+1, \end{cases} \quad (11)$$

где \mathbf{p}_0 определяется как единственное решение системы уравнений

$$\mathbf{p}_0^T \mathbf{Z} = \mathbf{0}^T, \quad (12)$$

$$\mathbf{p}_0^T \mathbf{v} = 1. \quad (13)$$

Далее, пусть τ_k – с.в. – интервал между выходом $(n-1)$ -ой «успешной» заявки и началом обслуживания n -ой «успешной» заявки при условии, что в момент $t_0 + 0$ выхода $(n-1)$ -ой заявки в системе находилось k заявок, $k = \overline{0}, r$.

Пусть $\eta_k(t)$, $t \in [t_0 + 0, t_0 + \tau_k)$ – случайный процесс (СП), определенный на множестве состояний

$$\mathcal{Y}_k = \bigcup_{n=1}^{r+1} \mathcal{Y}_{(k,n)^+} + \bigcup_{n=1}^r \mathcal{Y}_{(k,n)^-} + \mathcal{Y}_{(k,0)},$$

где

$$\begin{aligned}
\mathcal{Y}_{(k,n)^+} &= \{(i, j)_{(k,n)^+}, \quad i = \overline{1}, \overline{l}, \quad j = \overline{1}, \overline{m}\}, \quad n = \overline{1}, r+1, \\
\mathcal{Y}_{(k,n)^-} &= \{(i, j)_{(k,n)^-}, \quad i = \overline{1}, \overline{l}, \quad j = \overline{1}, \overline{m}\}, \quad n = \overline{1}, r, \\
\mathcal{Y}_{(k,0)} &= \{(i)_{(k,0)}, \quad i = \overline{1}, \overline{l}\}.
\end{aligned}$$

Кроме этого, введем макросостояния:

$$\begin{aligned}
(i)_{(k,n)^+} &= \bigcup_{j=1}^m (i, j)_{(k,n)^+}, \quad i = \overline{1}, \overline{l}, \quad n = \overline{1}, k+1, \\
(i)_{(k,n)^-} &= \bigcup_{j=1}^m (i, j)_{(k,n)^-}, \quad i = \overline{1}, \overline{l}, \quad n = \overline{1}, r,
\end{aligned}$$

$$(i)_{(k,0)-} = (i)_{(k,0)}; \quad (\cdot)_{(k,0)} = \bigcup_{i=1}^l (i)_{(k,0)}.$$

Обозначим через \mathbf{D}_k – матрицу интенсивностей переходов СП $\eta_k(t)$, а через $\mathbf{Q}_k(t) = (q_{y,z}^k)_{y,z \in \mathcal{Y}_k}$ – матрицу переходных вероятностей:

$$q_{y,z}^k(t) = P\{\tau_k > t, \eta_k(t) = z | \eta_k(0) = y\}_{y,z \in \mathcal{Y}_k}$$

и вектор начального распределения

$$\mathbf{h}_k^T = (h_k(y), y \in \mathcal{Y}_k),$$

$$h_k(y) = P\{\eta_k(0) = y\}.$$

Далее положим:

$$\mathbf{q}_k^T(t) = \mathbf{h}_k^T \mathbf{Q}_k(t),$$

$$q_k^T(t) = \mathbf{q}_k^T(t) \mathbf{1}.$$

Заметим, что q_k^T есть дополнение к ФР с.в. τ_k .

Учитывая вероятностную интерпретацию PH -распределения, можно утверждать, что СП $\eta_k(t)$ является поглощающим марковским процессом, для которого выражение $\eta_k(t) = (i, j)_{(k,n)\pm}$ означает, что в системе в момент времени t находится n заявок при условии, что в момент времени $t_0 + 0$ их было k , на приборе на фазе j обслуживается «успешная» (состояние со знаком «+») либо «неуспешная» (состояние со знаком «-») заявка, а генерация новой заявки происходит на фазе i .

Равенство $\eta_k(t) = (i)_{(k,0)}$ означает, что в момент времени t система пуста, а генерация новой заявки происходит на фазе i .

Далее введем обозначения

$$q(i, j)_{(k,n)+}(t) = P\{\eta_k(t) = (i, j)_{(k,n)+}\}, \quad n = \overline{1, r+1},$$

$$q(i, j)_{(k,n)-}(t) = P\{\eta_k(t) = (i, j)_{(k,n)-}\}, \quad n = \overline{1, r},$$

$$q(i)_{(k,0)}(t) = P\{\eta_k(t) = (i)_{(k,0)}\}$$

и введем векторы

$$\mathbf{q}_{(k,n)+}^T(t) = (q(i, j)_{(k,n)+}(t), i = \overline{1, l}, j = \overline{1, m}), \quad n = \overline{1, r+1},$$

$$\mathbf{q}_{(k,n)-}^T(t) = (q(i, j)_{(k,n)-}(t), i = \overline{1, l}, j = \overline{1, m}), \quad n = \overline{1, r},$$

$$\mathbf{q}_{(k,0)}^T(t) = (q(i)_{(k,0)}(t), i = \overline{1, l}),$$

$$\mathbf{q}_k^T(t) = \left(\mathbf{q}_{(k,0)}^T(t), \mathbf{q}_{(k,1)-}^T(t), \mathbf{q}_{(k,1)+}^T(t), \dots, \mathbf{q}_{(k,r)-}^T(t), \mathbf{q}_{(k,r)+}^T(t), \mathbf{q}_{(k,r+1)+}^T(t) \right).$$

Далее заметим, что начальными состояниями МП $\eta_k(t)$ являются состояния $(i)_{(0,0)}$ и $(i, j)_{(k,k)-}$, т.к. процесс начинает движение с того момента, когда сразу после окончания обслуживания «успешной» заявки в системе оказывается k заявок, а сам факт завершения обслуживания приводит к тому, что из накопителя на обслуживание поступает «неуспешная» заявка, фаза обслуживания которой

выбирается в соответствии с β , за исключением случая, когда система оказывается свободной. В свою очередь, стационарное распределение вероятностей состояний $(i)_{(0,0)}$ и $(i, j)_{(k,k)-}$ для моментов времени $t_0 + 0$ окончания обслуживания «успешных» заявок можно определить через стационарные вероятности $\mathbf{p}_D(i)_{(0,0)}$ и $\mathbf{p}_D(i)_{(k,k)-}$ соответствующих состояний ЦМ, вложенной по этим моментам в МП $X(t)$. Для подсчета этих вероятностей воспользуемся результатами [12], согласно которым:

$$\mathbf{p}_D^T(i)_{(k,k)-} = \left[\sum_{n=1}^{r+1} \mathbf{p}_{n,+}^T \mathbf{1} \right]^{-1} \mathbf{p}_{k+1,+}^T (\mathbf{I} \otimes \mathbf{1}), \quad k = \overline{0, r}. \quad (14)$$

Таким образом, для вектора начального распределения имеем:

$$\mathbf{h}_k^T = \begin{cases} (\mathbf{p}_D^T(i)_{(0,0)}; 0^{d_0}), & d_0 = lm + 2lmr, k = 0, \\ (0^{d_1}; \mathbf{p}_D^T(i)_{(k,k)-} \otimes \beta^T; 0^{d_2}), & d_1 = l + 2lm(k-1), \\ & d_2 = 2lm(r+1-k), k = \overline{1, r}. \end{cases}$$

Известно [10], глава 1, что для $\mathbf{q}_k^T(t)$ применима система дифференциальных уравнений Колмогорова:

$$\frac{d}{dt} \mathbf{q}_k^T(t) = \mathbf{q}_k^T(t) \mathbf{D}_k \quad (15)$$

с начальным условием

$$\mathbf{q}_k(0) = \mathbf{h}_k. \quad (16)$$

Умножая (15) на e^{-st} и интегрируя по t от 0 до ∞ , получим:

$$\begin{aligned} \int_0^\infty (\mathbf{q}_k^T)'_t e^{-st} dt &= \int_0^\infty \mathbf{q}_k^T(t) e^{-st} \mathbf{D}_k dt - \\ &- \mathbf{q}_k^T(0) + s \int_0^\infty \mathbf{q}_k^T(t) e^{-st} dt = \int_0^\infty \mathbf{q}_k^T(t) e^{-st} \mathbf{D}_k dt. \end{aligned}$$

Замечая, что $\int_0^\infty \mathbf{q}_k^T(t) e^{-st} dt \triangleq \hat{\mathbf{q}}_k^T(s)$ – преобразование Лапласа (ПЛ) $\mathbf{q}_k^T(t)$ и учитывая (16), приходим к следующему равенству:

$$\mathbf{h}_k^T = \hat{\mathbf{q}}_k^T(s) (s\mathbf{I} - \mathbf{D}_k). \quad (17)$$

Поскольку матрица \mathbf{D}_k неразложима, то существует $(s\mathbf{I} - \mathbf{D}_k)^{-1}$. Кроме того, элементы \mathbf{D}_k не зависят от k , поэтому индекс k можно опустить.

Итак, нами доказана

Теорема 2. *Преобразование Лапласа стационарного дополнения ФР интервалов времени между выходами из системы «успешной» заявки при условии, что в момент $t = 0$ ее выхода в системе имелось k заявок, и началом обслуживания следующей «успешной» заявки определяется выражением:*

$$\hat{\mathbf{q}}_k^T(s) = \mathbf{h}_k^T (s\mathbf{I} - \mathbf{D}_k)^{-1}. \quad (18)$$

Обозначим через $\tilde{q}_k(s)$ – ПЛС плотности распределения с.в. τ_k (ПЛС с.в. τ_k) и положим $\hat{q}_k(s) = \hat{\mathbf{q}}_k^T(s) \mathbf{1}$. Учитывая наличие связи между $\tilde{q}_k(s)$ и $\hat{q}_k(s)$, выражаемой формулой [10]:

$$\tilde{q}_k(s) = 1 - s\hat{q}_k(s),$$

получим очевидное

Следствие 6. ПЛС интервала времени между выходом из системы «успешной» заявки при условии, что в момент $t + 0$ ее выхода в системе находилось k заявок, и началом обслуживания следующей «успешной» заявки определяется выражением:

$$\tilde{q}_k(s) = 1 - s\mathbf{h}_k^T (s\mathbf{I} - \mathbf{D})^{-1}\mathbf{1}. \quad (19)$$

Известно [13], что начальные моменты $M\tau_k^n$ порядка n неотрицательной с.в. (если они существуют) можно определять через ПЛ дополнения ее ФР по формуле:

$$M\tau_k^n = (-1)^{n-1} \left. \frac{d^{n-1}}{ds^{n-1}} \hat{q}_k(s) \right|_{s=0}.$$

Отсюда получаем еще одно очевидное

Следствие 7. Среднее значение интервала времени между выходом из системы «успешной» заявки при условии, что в момент $t + 0$ ее выхода в системе находилось k заявок, и началом обслуживания следующей «успешной» заявки определяется выражением:

$$M\tau_k = -\mathbf{h}_k^T \mathbf{D}^{-1}\mathbf{1}. \quad (20)$$

Очевидно, что пиковый возраст «успешных» заявок помимо с.в. τ_k включает в себя время пребывания «успешной» заявки в системе и время обслуживания следующей «успешной» заявки. В свою очередь, время пребывания равно времени обслуживания, т.к. «успешные» заявки, как и остальные, сразу направляются на прибор, но их обслуживание не прерывается. Учитывая независимость интервалов и применяя формулу полной вероятности, приходим к следующему результату:

Теорема 3. ПЛС $z(s)$ пикового возраста информации, передаваемой «успешными» заявками системы $PH/PH/1/r/LCFS/PRR$, определяется выражением:

$$z(s) = \sum_{k=0}^r [\mathbf{p}_D^T(i)_{(k,k)} - \mathbf{1}] \tilde{q}_k(s) \cdot \beta^2(s), \quad (21)$$

где $\beta(s)$ – ПЛС ФР $B(t)$, а $\mathbf{p}_D(i)_{(k,k)}$ и $\tilde{q}_k(s)$ определяются в соответствии с (14) и (19).

3. Пример расчета среднего значения пикового возраста информации

В качестве частного случая рассмотрим систему $M/M/1/1/LCFS/PRR$ с пуассоновским потоком интенсивности λ и экспоненциальным обслуживанием μ . При этом с точки зрения основной системы ее параметры определяются следующими выражениями:

$$\begin{aligned} \boldsymbol{\alpha} &= (1); & \boldsymbol{\Lambda} &= (-\lambda); & \boldsymbol{\lambda} &= (\lambda), \\ \boldsymbol{\beta} &= (1); & \mathbf{M} &= (-\mu); & \boldsymbol{\mu} &= (\mu). \end{aligned} \quad (22)$$

Наша задача определить среднее значение пикового возраста информации.

Функционирование рассматриваемой системы можно описать однородным МП $X(t)$, $t \geq 0$, над множеством состояний $\{(k), k = \overline{0, 2}\}$, где $X(t) = (k)$ означает, что в момент времени t в системе имеется k заявок. Для расчета стационарных вероятностей

$$p_k = \lim_{t \rightarrow \infty} P\{X(t) = k\}$$

можно применить теорему 1, либо воспользоваться готовыми формулами, например из главы 3 работы [10]:

$$p_0 = \frac{1 - \rho}{1 - \rho^3}, \quad (23)$$

$$p_1 = \rho p_0, \quad (24)$$

$$p_2 = \rho p_1, \quad \text{где } \rho = \frac{\lambda}{\mu} - \text{загрузка системы.} \quad (25)$$

Далее, используя (9) и (10), находим стационарные вероятности состояний, учитывающих «успешность» обслуживаемой заявки:

$$p_{1,-} = \frac{\mu}{\lambda + \mu} p_2 = \frac{\mu}{\lambda + \mu} \rho p_1 = \frac{\lambda}{\lambda + \mu} p_1, \quad (26)$$

$$p_{1,+} = p_1 - \frac{\lambda}{\lambda + \mu} p_1 = \frac{\mu}{\lambda + \mu} p_1, \quad (27)$$

$$p_{2,+} = p_2. \quad (28)$$

В соответствии с (14) определим стационарное распределение вероятностей состояний ЦМ, вложенной по моментам $t + 0$ окончания обслуживания:

$$p_D(1)_{(0;0)} = \frac{p_{1,+}}{p_{1,+} + p_{2,+}}, \quad (29)$$

$$p_D(1)_{(1;1)-} = \frac{p_{2,+}}{p_{1,+} + p_{2,+}}. \quad (30)$$

Далее согласно (17) определим \mathbf{h}_k , задающее начальное распределение состояний МП $\eta_k(t)$:

$$\mathbf{h}_0^T = (p_D(1)_{(0;0)}; 0; 0; 0),$$

$$\mathbf{h}_1 = (0; p_D(1)_{(1;1)-}; 0; 0)$$

и с учетом (22) выписываем матрицу:

$$D = \begin{bmatrix} -\lambda & \lambda & 0 & 0 \\ 0 & -(\lambda + \mu) & 0 & \lambda \\ \mu & 0 & -(\lambda + \mu) & \lambda \\ 0 & 0 & 0 & -\mu \end{bmatrix}.$$

Далее находим

$$D^{-1} = \begin{bmatrix} -\frac{1}{\lambda} & -\frac{1}{\lambda + \mu} & 0 & -\frac{\lambda}{\mu(\lambda + \mu)} \\ 0 & -\frac{1}{\lambda + \mu} & 0 & -\frac{\mu(\lambda + \mu)}{\lambda} \\ -\frac{\mu}{\lambda(\lambda + \mu)} & -\frac{\mu}{(\lambda + \mu)^2} & -\frac{1}{\lambda + \mu} & -\frac{\lambda^2 + 2\lambda\mu}{\mu(\lambda + \mu)^2} \\ 0 & 0 & 0 & -\frac{1}{\mu} \end{bmatrix}$$

и в соответствии с (20) получаем:

$$E(\tau_0) = p_D(1)_{(0;0)} \left[\frac{1}{\lambda} + \frac{1}{\lambda + \mu} + \frac{\lambda}{\mu(\lambda + \mu)} \right], \quad (31)$$

$$E(\tau_1) = p_D(1)_{(1;1)} \left[\frac{1}{\lambda + \mu} + \frac{\lambda}{\mu(\lambda + \mu)} \right]. \quad (32)$$

Подставляя в (31)–(32) правые части (29)–(30) и используя (23)–(28) приходим к следующему результату:

$$E(\tau) = \frac{\mu^2}{\lambda^2 + \lambda\mu + \mu^2} \cdot \frac{\lambda + \mu}{\lambda\mu} + \frac{\lambda^2 + \mu\lambda}{\lambda^2 + \lambda\mu + \mu^2} \frac{1}{\mu} = \frac{\lambda^3 + \lambda\mu^2 + \lambda^2\mu + \mu^3}{\lambda\mu(\lambda^2 + \lambda\mu + \mu^2)}. \quad (33)$$

Учитывая, что пиковый возраст включает в себя помимо $E(\tau)$ две средних длительности обслуживания, окончательно получаем:

$$E(z) = E(\tau) + \frac{2}{\mu} = \frac{3\lambda^3 + \mu^3 + 3\lambda\mu(\lambda + \mu)}{\lambda\mu(\lambda^2 + \lambda\mu + \mu^2)}.$$

Заключение

В результате проведенного исследования нам удалось получить выражение для преобразования Лапласа-Стилтьеса стационарной функции распределения пикового возраста информации, передаваемой из периферийного источника в центр управления, моделируя процесс передачи с помощью системы массового обслуживания с распределениями фазового типа. При этом были учтены требования протокола передачи, в соответствии с которыми рассматривалась система с инверсионной дисциплиной обслуживания, абсолютным приоритетом и переобслуживанием прерванных заявок. Данное исследование обобщает результаты работ [6] и [14] и дает возможность получать оценки возраста информации, максимально близкие к реальным, благодаря наличию большого числа параметров у распределений фазового типа.

Список литературы

- [1] Kaul S., Gruteser M., Rai V., Kenney J. Minimizing age of information in vehicular networks // 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks. 2011. Pp. 350–358.
- [2] Горбунова А.В., Зарядов И.С., Матюшенко С.И., Самуйлов К.Е., Шоргин С.Я. Аппроксимация времени отклика системы облачных вычислений // Информатика и ее применения. 2015. Т. 9, № 3. С. 32–38. <https://doi.org/10.14357/19922264150304>
- [3] Meykhanadzhyan L., Matyushenko S., Pyatkina D., Razumchik R. Revisiting joint stationary distribution in two finite capacity queues operating in parallel // Информатика и ее применения. 2017. Т. 11, № 3. С. 106–112. <https://doi.org/10.14357/19922264170312>

- [4] The Ongoing History of the Age of Information [Electronic resource]. 2023. URL: <https://webhome.auburn.edu/yzs0078/AoI.html>.
- [5] Costa M., Codreanu M., Ephremides A. On the age of information in status update systems with packet management // *IEEE Transactions on Information Theory*. 2016. Vol. 62, № 4. Pp. 1897–1910.
- [6] Kaul S.K., Yates R.D., Gruteser M. Status updates through queues // *46th Annual Conference on Information Sciences and Systems (CISS)*. 2012. Pp. 1–6.
- [7] Dogen O., Akar N. The Multi-Source Preemptive $M/PH/1/1$ Queue with Packet Error: Exact Distribution of the Age of Information and Its Peak // *IEEE Transactions on Communications*. 2021. Vol. 69, № 11. Pp. 7297–7308. <https://doi.org/10.1109/TCOMM.2021.3106347>
- [8] Akar N., Dogan O., Atay E.U. Finding the Exact Distribution of (Peak) Age of Information for Queues of $PH/PH/1/1$ and $M/PH/1/2$ Type // *IEEE Transactions on Communications*. 2020. Vol. 68, № 9. <https://doi.org/10.48550/arXiv.1911.07274>
- [9] Kosta A., Pappas N., Angelakis V. Age of Information: A New Concept, Metric, and Tool // *Foundations and Trends in Networking*. 2017. Vol. 12, № 3. Pp. 162–259. <https://doi.org/10.1561/13000000060>
- [10] Бочаров П.П., Печинкин А.В. Теория массового обслуживания. М.: Изд-во РУДН, 1995. 529 с.
- [11] Бочаров П.П., Павлова О.И. Матрично-геометрическое распределение очереди при дисциплине LCFS с прерываниями и распределениями фазового типа // *Автоматика и телемеханика*. 1991. № 9. С. 112–122.
- [12] Башарин Г.П., Бочаров П.П., Коган Я.Ф. Анализ очередей в вычислительных сетях. Теория и методы анализа. М.: Наука, 1989.
- [13] Наумов В.А. О предельных вероятностях полумарковского процесса // *Современные задачи в точных науках*. М.: Университет дружбы народов, 1975. С. 35–39.
- [14] Najm E., Yates R., Soljanin E. Status updates through $M/G/1/1$ queues with HARQ // *Proceedings IEEE International Symposium on Information Theory (ISIT)*. 2017. Pp. 131–135.

Образец цитирования

Гайдамака Ю.В., Матюшенко С.И., Самуйлов К.Е. О пиковом возрасте информации в группе передачи, моделируемой системой $PH/PH/1/r/LCFS$ с захватом прибора и переобслуживанием заявок // *Вестник ТвГУ. Серия: Прикладная математика*. 2024. № 1. С. 37–52. <https://doi.org/10.26456/vtppmk701>

Сведения об авторах**1. Гайдамака Юлия Васильевна**

профессор кафедры теории вероятностей и кибербезопасности Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» РАН.

Россия, 117198, Москва, ул. Миклухо-Маклая, д. 6, РУДН.

E-mail: gaydamaka-yuv@rudn.ru

2. Матюшенко Сергей Иванович

доцент кафедры теории вероятностей и кибербезопасности Российского университета дружбы народов.

Россия, 117198, г. Москва, ул. Миклухо-Маклая, д. 6, РУДН.

E-mail: matyushenko-si@rudn.ru

3. Самуйлов Константин Евгеньевич

заведующий кафедрой теории вероятностей и кибербезопасности Российского университета дружбы народов; ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» РАН.

Россия, 117198, Москва, ул. Миклухо-Маклая, д. 6, РУДН.

E-mail: samuylov-ke@rudn.ru

**ABOUT THE PEAK AGE OF INFORMATION
IN A SINGLE-CHANNEL TRANSMISSION GROUP MODELED
BY THE $PH/PH/1/R/LCFS$ SYSTEM WITH INVERSION
DISCIPLINE, INTERRUPTION MAINTENANCE AND RE-SERVICE
OF APPLICATIONS**

Gaidamaka Yu.V.^{*,}, Matyushenko S.I.^{*}, Samouylov K.E.^{*,**}**

^{*}RUDN University, Moscow

^{**}Federal Research Center “Computer Science and Control”,
Russian Academy of Sciences, Moscow

Received 09.01.2024, revised 29.01.2024.

This article continues the author’s cycle of works devoted to the problem of the Age of Information (AoI), a metric used in information systems to quantify the “freshness” of information delivered to the control center from peripheral sources. The paper considers a two-node information transmission group consisting of a sender node, a recipient node and a communication channel between them. The information transfer process is modeled by means of a single-line finite capacity queuing system with phase-type distributions, which in Kendall’s notation is encoded as $PH/PH/1/r$. This takes into account the special requirements of the transmission protocol, which consist in the fact that a packet entering the system, bypassing the queue, is immediately sent for transmission, capturing the channel from the previous packet if it has not completed transmission. The packet whose transmission was interrupted is moved to the first place in the queue and, after the channel is released, attempts transmission again. For this system, an expression is obtained for the Laplace–Stieltjes transformation of the stationary distribution function of the peak age of information and its average value. A numerical study of the dependence of the peak age of information on the system load was carried out. The correctness of the analytical results was verified by comparing them with the results of simulation modeling.

Keywords: age of information, peak age of information, queuing system, phase type distribution.

Citation

Gaidamaka Yu.V., Matyushenko S.I., Samouylov K.E., “About the Peak Age of Information in a Single-Channel Transmission Group Modeled by the $PH/PH/1/r/LCFS$ System with Inversion Discipline, Interruption Maintenance and Re-service of Applications”, *Vestnik TvgU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2024, № 1, 37–52 (in Russian). <https://doi.org/10.26456/vtprm701>

References

- [1] Kaul S., Gruteser M., Rai V., Kenney J., “Minimizing age of information in vehicular networks”, *8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 2011, 350–358.
- [2] Gorbunova A.V., Zaryadov I.S., Matyushenko S.I., Samujlov K.E., Shorgin S.Ya., “Approximation of the response time of a cloud computing system”, *Informatika i ee primeneniya [Computer science and its applications]*, **9**:3 (2015), 32–38 (in Russian), <https://doi.org/10.14357/19922264150304>.
- [3] Meykhanadzhyan L., Matyushenko S., Pyatkina D., Razumchik R., “Revisiting joint stationary distribution in two finite capacity queues operating in parallel”, *Informatika i ee primeneniya [Computer science and its applications]*, **11**:3 (2017), 106–112 (in Russian), <https://doi.org/10.14357/19922264170312>.
- [4] *The Ongoing History of the Age of Information*, 2023, <https://webhome.auburn.edu/yzs0078/AoI.html>.
- [5] Costa M., Codreanu M., Ephremides A., “On the age of information in status update systems with packet management”, *IEEE Transactions on Information Theory*, **62**:4 (2016), 1897–1910.
- [6] Kaul S.K., Yates R.D., Gruteser M., “Status updates through queues”, *46th Annual Conference on Information Sciences and Systems (CISS)*, 2012, 1–6, <https://ieeexplore.ieee.org/document/6310931>.
- [7] Dogen O., Akar N., “The Multi-Source Preemptive $M/PH/1/1$ Queue with Packet Error: Exact Distribution of the Age of Information and Its Peak”, *IEEE Transactions on Communications*, **69**:11 (2021), 7297–7308, <https://doi.org/10.1109/TCOMM.2021.3106347>.
- [8] Akar N., Dogan O., Atay E.U., “Finding the Exact Distribution of (Peak) Age of Information for Queues of $PH/PH/1/1$ and $M/PH/1/2$ Type”, *IEEE Transactions on Communications*, **68**:9 (2020), <https://doi.org/10.48550/arXiv.1911.07274>.
- [9] Kosta A., Pappas N., Angelakis V., “Age of Information: A New Concept, Metric, and Tool”, *Foundations and Trends in Networking*, **12**:3 (2017), 162–259, <https://doi.org/10.1561/13000000060>.
- [10] Bocharov P.P., Pechinkin A.V., *Teoriya massovogo obsluzhivaniya [Queueing theory]*, Publishing House of RUDN, Moscow, 1995 (in Russian), 529 pp.
- [11] Bocharov P.P., Pavlova O.I., “Matrix-geometric queue distribution in LCFS discipline with interruptions and phase-type distributions”, *Avtomatika i telemekhanika [Automation and telemechanics]*, 1991, № 9, 112–122 (in Russian).
- [12] Basharin G.P., Bocharov P.P., Kogan Ya.F., *Queue analysis in computer networks. Theory and methods of analysis*, Nauka Publ., Moscow, 1989 (in Russian).

- [13] Naumov V.A., “On the limiting probabilities of a semi-Markov process”, *Modern tasks in the exact sciences*, Peoples’ Friendship University, Moscow, 1975, 35–39 (in Russian).
- [14] Najm E., Yates R., Soljanin E., “Status updates through $M/G/1/1$ queues with HARQ”, *Proceedings IEEE International Symposium on Information Theory (ISIT)*, 2017, 131–135.

Author Info

1. **Gaidamaka Yuliya Vasilyevna**

Professor of Department of Applied Probability and Informatics, RUDN University; Senior Researcher at Federal Research Center “Computer Science and Control”, Russian Academy of Sciences.

Russia, 119992, Moscow, 6 Miklukho-Maklaya St., RUDN University.

E-mail: gaydamaka-yuv@rudn.ru

2. **Matyushenko Sergey Ivanovich**

Assistant Professor of Department of Probability Theory and Cyber Security, RUDN University.

Russia, 119992, Moscow, 6 Miklukho-Maklaya St., RUDN University.

E-mail: matyushenko-si@rudn.ru

3. **Samouylov Konstantin Evgenyevich**

Head of Department of Probability Theory and Cyber Security, RUDN University; Lead scientist at Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences.

Russia, 119992, Moscow, 6 Miklukho-Maklaya St., RUDN University.

E-mail: samuylov-ke@rudn.ru