

УДК 519.2

О РЕГЕНЕРАТИВНОМ МОДЕЛИРОВАНИИ ОДНОСЕРВЕРНЫХ СИСТЕМ ОБСЛУЖИВАНИЯ С УСТАРЕВАНИЕМ ИНФОРМАЦИИ

И. В. Пешкова¹⁾, А. С. Алексеев²⁾

¹⁾*Петрозаводский государственный университет, просп. Ленина, 33, 185910, г. Петрозаводск, Россия, Институт прикладных математических исследований Карельского научного центра Российской академии наук, email: iaminova@petrsu.ru*

²⁾*Петрозаводский государственный университет, просп. Ленина, 33, 185910, г. Петрозаводск, Россия*

Работа посвящена изучению односерверных систем без буфера, в которых при поступлении с некоторой вероятностью принимается решение о допуске заявки на обслуживание, а также два типа блокировки старых сообщений: снятие читаемого сообщения в момент прихода нового или блокировка всех новых сообщений, пока не прочитается текущее. Представлены результаты численного моделирования для системы с гиперэкспоненциальным входным потоком и экспоненциальным обслуживанием.

Ключевые слова: система с устареванием информации; регенеративное моделирование.

ON REGENERATIVE MODELING OF SINGLE-SERVER QUEUEING SYSTEMS WITH AGING INFORMATION

I. V. Peshkova¹⁾, A. S. Alekseev²⁾

¹⁾*Petrozavodsk State University, Lenin Str., 33, 185910, Petrozavodsk, Russia, Institute of Applied Mathematical Research, Karelian Research Center of RAS, iaminova@petrsu.ru*

²⁾*Petrozavodsk State University, Lenin Str., 33, 185910, Petrozavodsk, Russia*

The work is devoted to the study of single-server systems without a buffer, in which, upon arrival, a decision is made with some probability to admit a service request, as well as two types of blocking old messages: removing a readable message at the time of arrival of a new one or blocking all new messages until the current one is read. The results of numerical simulation for a system with hyperexponential interarrivals and exponential maintenance are presented

Keywords: system with aging information.

В настоящее время актуальной является проблема устаревания (возраста) передаваемой информации, являющейся частью процесса управления или принятия решений в реальном времени [1, 2]. Например, си-

системы мониторинга или сбора киберфизических значений (температура, влажность и т. д.) включают в себя несколько датчиков, которые периодически передают показания на базовую станцию, при этом актуальность такой информации чувствительна к задержке сообщения. Любое тревожное значение параметра делает устаревшим любое сообщение о нормальной работе системы.

Как уже говорилось выше, такие системы зависят от возраста последнего полученного сообщения с удаленного датчика. При поступлении нового информационного пакета необходимо принять решение: считать его или нет. Выбор имеет решающее значение и зависит от различных параметров, таких как время задержки, длина пакета и частота поступления информационных пакетов [3].

В нашем исследовании мы рассматриваем односерверные системы обслуживания без буфера, в которых при поступлении с некоторой вероятностью принимается решение о допуске заявки на обслуживание, а также два типа блокировки старых сообщений: снятие читаемого сообщения в момент прихода нового или блокировка всех новых сообщений пока не прочтется текущее. Входной поток имеет гиперэкспоненциальное распределение, а обслуживания – показательное. В работе [4] рассматриваются подобные системы с пуассоновским входным потоком без учета вероятности допуска.

Рассмотрим односерверную систему массового обслуживания типа $M_2/M/1$ [5], в которой отсутствует буфер. Обозначим через T_n – момент прихода n -го сообщения, $n \geq 1, T_0 = 0$. Пусть $\tau_n := T_{n+1} - T_n$ – интервалы между поступлениями сообщений на сервер. Введем для каждого $n \geq 1$ индекс допуска сообщения χ_n следующим образом:

$$\chi_n = \begin{cases} 1, & \text{если сообщение, прибывшее в момент } T_n, \text{ допущено к прочтению} \\ 0, & \text{иначе.} \end{cases}$$

Пусть $P(\chi_n = 1) = p$. Обозначим через T'_n время ухода сообщения из системы, s_n – время чтения сообщения, поступившего в момент T_n .

Введем функцию состояния сервера в момент времени t :

$$q(t) := \sum_{n \geq 1} \chi_n I_{T_n \leq t \leq T'_n}, \quad (1)$$

где I – индикатор случайного события. Функция $q(t) = 0$, если сервер простаивает, или $q(t) = 1$, если занят.

Будем называть сообщение *n* *успешным*, если оно покидает систему после полного прочтения.

Обозначим

$$\psi_n := I_{T'_n = T_n + s_n} \quad (2)$$

– индекс успешного прочтения *n*-го сообщения. Заметим, что по определению для всех *n* выполнено условие $\psi_n \leq \chi_n$.

Обозначим D_n – момент ближайшего к *n*-му ”успешному” ($\psi_n = 1$) сообщению поступления нового сообщения. Тогда

$$\gamma_k = \psi_n (T_{D_n} - T'_n)(1 - q(T'_{D_n})), k \geq 0$$

– интервалы простоя сервера. Пусть $X(t)$ – процесс, описывающий состояние системы в момент времени *t*, а именно ”время жизни” сообщения, находящегося на обслуживании в момент времени *t*, обозначим $\beta_0 = 0, \beta_{k+1} = \inf(T_n : q(T_n^-) = 0)$ – моменты прихода сообщений, в которые система простаивает, $k \geq 1$. Тогда процесс X является регенерирующим, а моменты β_k являются его моментами регенерации с $\alpha_k = \beta_{k+1} - \beta_k \geq 1$ – длинами циклов регенерации.

Далее рассмотрим два типа систем, в которых по-разному определяется, попадёт ли сообщение на обслуживание и будет ли оно полностью прочитано [4].

Система с обновлением старых заявок. Рассмотрим систему, в которой при поступлении нового сообщения находящееся на прочтении сообщение немедленно снимается и покидает сервер. Время ухода T'_n сообщения из такой системы определяется соотношением

$$T'_n = \begin{cases} T_n, & \text{если } \chi_n = 0; \\ \min(T_n + s_n, \inf\{T_r : r > n, \chi_r = 1\}), & \text{если } \chi_n = 1. \end{cases} \quad (3)$$

Таким образом, прибывшее в систему сообщение может быть либо немедленно отклонено с вероятностью $1 - p$ и покинет систему без прочтения, либо оно поступит на обслуживание. Причем принятое сообщение может быть либо полностью прочитано, либо вытеснено другим новым принятым во время его обслуживания сообщением.

На рис. 1 представлена схема работы системы с обновлением старых заявок.

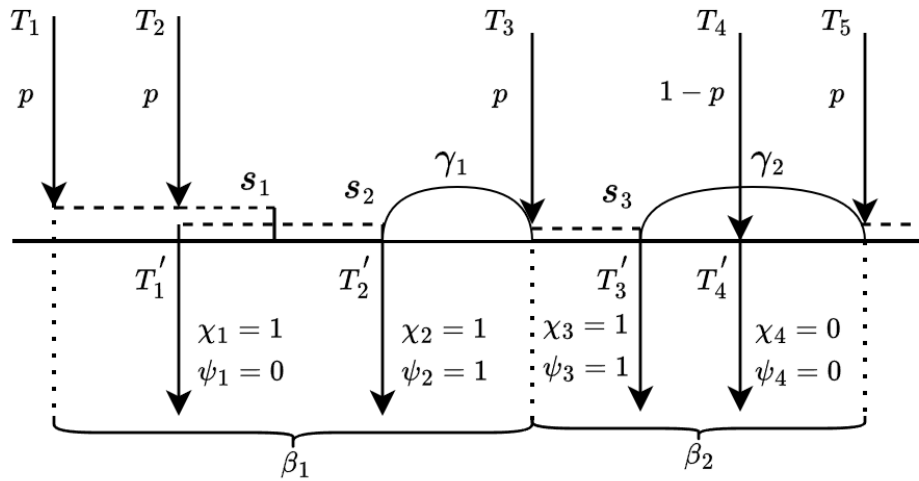


Рис. 1. Система с обновлением старых заявок

Из (3) и (1) следует, что индекс успешного прочтения n -го сообщения равен

$$\psi_n = I_{T_n + s_n \leq T_{n+1}} = I_{\tau_n \geq s_n}, n \geq 1.$$

Система с блокированием новых заявок. Рассмотрим систему, в которой каждое новое сообщение после проверки на допуск ($\chi_n = 1$) блокируется, если в момент его поступления сервер занят. Таким образом, "успешными" являются все сообщения, которые пришли в пустую систему. В этом случае $\psi_n = \chi_n$ для всех n , т. е. каждое принятое на чтение сообщение является успешным.

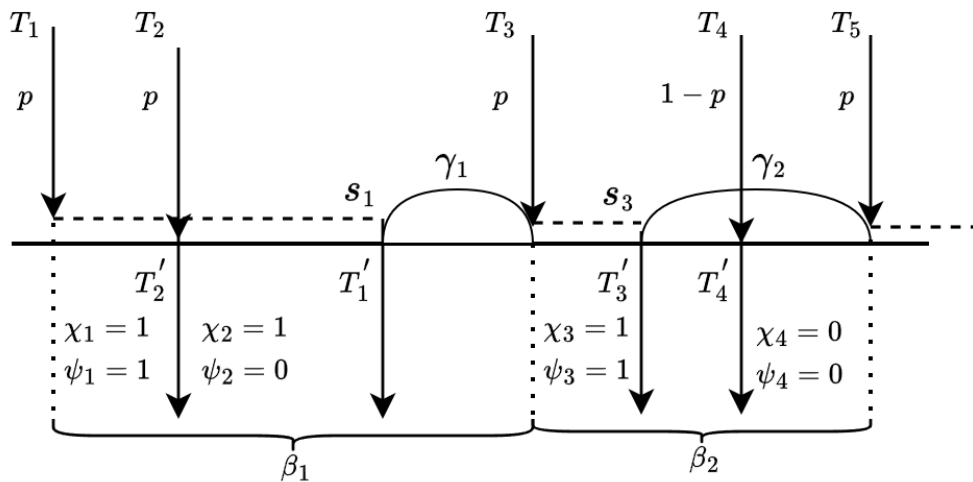


Рис. 2. Система с блокированием новых заявок.

На рис. 2 представлена схема работы системы с блокированием старых заявок.

Время ухода T_n' сообщения из такой системы определяется соотношением

$$T_n' = \begin{cases} T_n, & \text{если } \chi_n = 0 \text{ или } q(T_n^-) = 1; \\ \min(T_n + s_n, \inf\{T_r : r > n, \chi_r = 1\}), & \text{если } \chi_n = 1 \text{ или } q(T_n^-) = 0. \end{cases} \quad (4)$$

Для проведения численных экспериментов рассматривалась система $H_2/M/1$, в которой поток входящих сообщений τ описывается гиперэкспоненциальным распределением [6]:

$$A(t) = 1 - qe^{-\lambda_1 t} - (1 - q)e^{-\lambda_2 t}, \quad t \geq 0,$$

где $\lambda_1, \lambda_2 \geq 0$ и $0 \leq q \leq 1$. Для такой системы были получены эмпирические оценки среднего времени простоя и числа успешных заявок при следующих параметрах: интенсивность обслуживания $\mu = 3; 3,5; \dots; 10$; $\lambda_1 = 6$; $\lambda_2 = 2$; $q = 0,6$. Количество входящих сообщений $N = 4 \cdot 10^4$, вероятность допуска $p = 0,7$. Результаты моделирования представлены на рис. 3 и рис. 4.

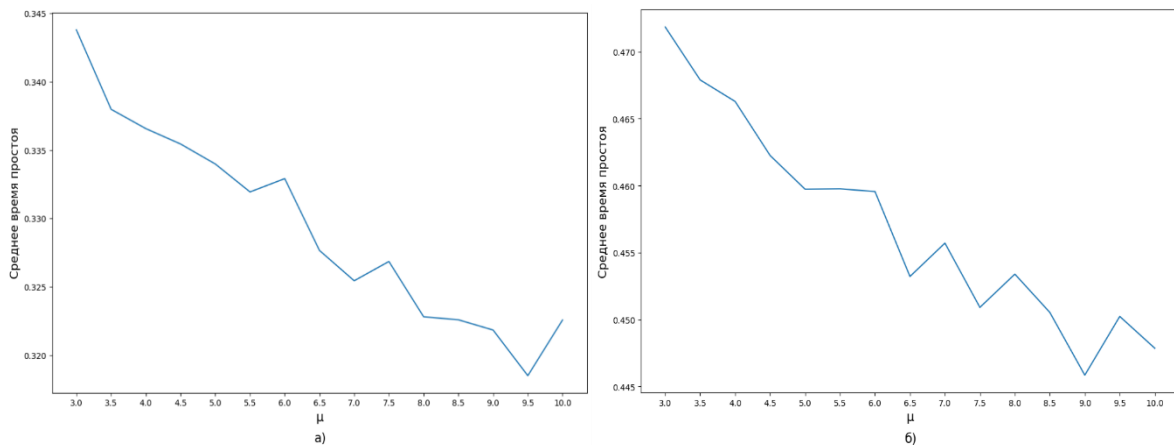


Рис. 3. Зависимость среднего времени простоя γ от интенсивности обслуживания μ :

- а) в системе с обновлением старых заявок;
- б) в системе с блокированием новых заявок.

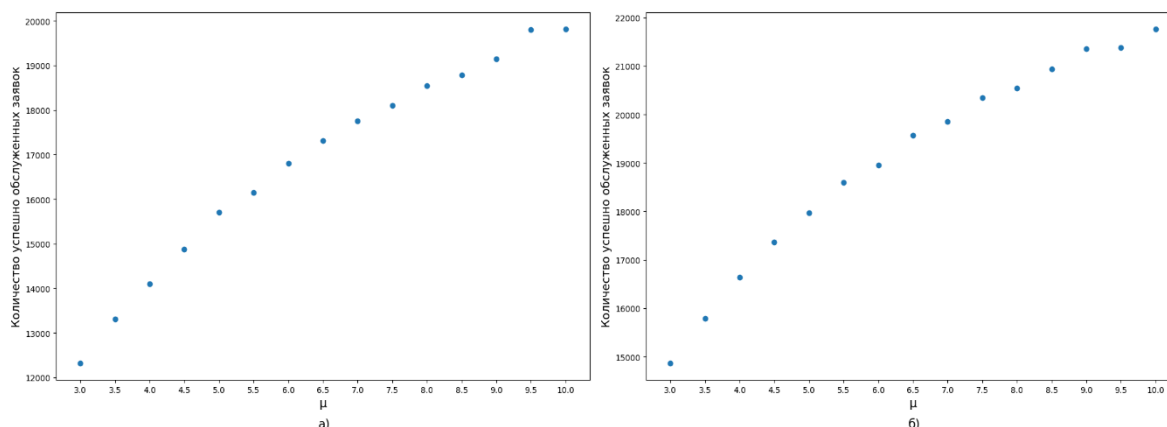


Рис. 4. Зависимость количества ”успешных” заявок от интенсивности обслуживания μ :
а) в системе с обновлением старых заявок;
б) в системе с блокированием новых заявок.

В статье рассмотрены два типа систем обслуживания: с обновлением старых заявок и с блокировкой новых. В обоих случаях при приходе заявки принимается решение о ее допуске. Результаты численного моделирования показали, что время простоя и число успешно обслуженных заявок растет с ростом интенсивности обслуживания. При этом вторая дисциплина обслуживания оказывается оптимальнее с точки зрения роста числа ”успешных” заявок.

Работа выполнена при финансовой поддержке РФФИ (проект 21-71-10135).

Библиографические ссылки

1. He Q., Yuan D., Ephremide A. Optimizing freshness of information: On minimum age link scheduling in wireless systems // Proceedings of the 14th International Symposium on Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks (WiOpt). 2016. P. 1–8.
2. Kaul S., Yates R., Gruteser M. Real-time status: How often should one update? // Proceedings – IEEE INFOCOM. 2012. P. 2731–2735.
3. Kosta A., Pappas N., Angelakis V. Age of information: A new concept, metric, and tool // Foundations and Trends in Networking. 2017. Vol. 12. P. 162–259.
4. Konstantopoulos T., Kesidis G., Zazanis M. A. The distribution of age-of-information performance measures for message processing systems // Queueing Systems. 2020. Vol. 95(3). P. 1–48.
5. Kleinrock L. The theory of queuing. Mashinostroenie, M., 1979. 432 p.
6. Sudhofen H.-D., Pawlita P.F. Modeling of Compound Traffic Streams in Computer Communication Network // Proc ITC–11. Kyoto. 1985. P. 3.1–3.7.