

УДК 519.872.1

РЕГЕНЕРАТИВНОЕ МОДЕЛИРОВАНИЕ ВЕРОЯТНОСТИ ПОТЕРИ В СИСТЕМАХ ОБСЛУЖИВАНИЯ С КОНЕЧНЫМ БУФЕРОМ

Р. С. Горичева, Е. В. Морозов

*Институт прикладных математических исследований
Карельского научного центра РАН*

В статье рассматривается проблема моделирования и статистического оценивания вероятности переполнения (потери) в системах обслуживания с конечным буфером. Приведены известные аналитические результаты, а также новая асимптотика вероятности потери при растущем размере буфера. Основное внимание уделено построению точечных и интервальных оценок вероятности потери с использованием регенеративного моделирования. Приведены численные примеры. Работа выполнена при поддержке РФФИ, проект 10-07-00017.

К л ю ч е в ы е с л о в а : система обслуживания с потерями, конечный буфер, регенерация, формула Эрланга, вероятность потери, распределение Парето, доверительное оценивание.

R. S. Goricheva, E. V. Morozov. REGENERATIVE SIMULATION OF LOSS PROBABILITY IN QUEUEING SYSTEMS WITH FINITE BUFFER

In the paper we consider the problem of modeling and statistical estimation of the overflow (loss) probability in queueing systems with finite buffer. Both known analytical results and a new asymptotic for loss probability under increasing buffer size are given. The main focus is on construction of point and interval estimates of the loss probability using regenerative simulation. Some numerical examples are given. The research is supported by RFBR, project 10-07-00017.

Key words: queueing loss systems, finite buffer, regeneration, Erlang formula, loss probability, Pareto distribution, confidence estimation.

Введение

При изучении современных телекоммуникационных систем важное место занимает анализ потока отказов или потерь. К потерям относится та часть поступающей нагрузки, которая не обслуживается из-за занятости обслуживающих устройств или переполнения буферов для ожидания в очереди.

В данной статье рассматриваются основные модели систем с потерями, обсуждаются известные аналитические результаты для таких систем, в частности, формулы Эрланга. Изучаются свойства вероятности потери, включая свойства монотонности, а также ее асимптотика при растущем размере буфера в условиях, когда коэффициент загрузки больше 1. Приведены результаты

о свойствах среднего числа потерь на цикле занятости, которые затем используются при оценивании стационарной вероятности потери.

Регенерирующие процессы и тесно связанные с ними процессы восстановления играют существенную роль при моделировании трафиков и оценивании характеристик широкого класса телекоммуникационных сетей. Важно также заметить, что регенеративный подход часто может быть использован для доверительного оценивания сетевых характеристик. В статье описана регенеративная структура системы с конечным буфером, а затем регенеративный метод использован для построения точечных и доверительных оценок как среднего числа потерь на цикле регенерации, так и вероятности потери. Подробно представлены результаты имитационного моделирования некоторых систем с потерями, в том числе систем с распределением Парето времени обслуживания. Данная статья не может рассматриваться как сколько-нибудь полный обзор проблематики систем с потерями. (В этой связи укажем недавнюю работу [Abramov, 2008].) Основной интерес для авторов представляет возможность применения регенеративного метода для надежного оценивания вероятности потери с перспективой использования этих результатов при оценке вероятности потери данных (оценки качества обслуживания (QoS) в телекоммуникационной сети.

Авторы посвящают эту работу 100-летию теории А. К. Эрланга, которое отмечалось в 2009 г.

Системы с пуассоновским входным потоком

Рассмотрим систему с буфером размера $n < \infty$, в которой заявка, поступающая при величине очереди, равной n , получает отказ – теряется. Пусть T_i – интервал между приходом i -й и $i+1$ -й заявок, а S_i – время обслуживания i -й заявки $i=1, 2, \dots$. Предполагается, что последовательности $\{T_i\}$ и $\{S_i\}$ независимы и состоят из независимых, одинаково распределенных случайных величин (н.о.р.с.в.). Всюду далее, чтобы обозначать типичный элемент такой последовательности, опускается соответствующий индекс. Коэффициент загрузки (или трафика) системы задается следующим образом: $\rho = ES / ET$, где ES , ET – среднее время обслуживания и средняя длина интервала между приходами заявок, соответственно. Стационарная вероятность по-

тери обозначается через P_l . В широких предположениях, выполненных во всех моделях, рассматриваемых в данной статье, эта вероятность существует и равна пределу отношения числа потерянных заявок к общему числу приходов, когда число приходов неограниченно растет [Abramov, 1997].

Система, в которой входной поток является пуассоновским и отсутствует буфер для ожидания в очереди, называется *системой Эрланга* $M/G/m/0$, m – число обслуживающих каналов. (Подробнее можно ознакомиться в [Greiner et al., 1999].) В такой системе вероятность P_l определяется *формулой Эрланга*

$$P_l = \frac{\rho^m / m!}{\sum_{k=1}^m \rho^k / k!}. \quad (1)$$

Подчеркнем, что формулы Эрланга (включающие также вероятности других состояний) верны при любом распределении времени обслуживания. Если время обслуживания имеет экспоненциальное распределение и система обладает буфером размера $0 \leq n < \infty$, то стационарную вероятность потери можно рассчитать по известной формуле

$$P_l = \frac{\rho^{m+n}}{m! m^n} P_0, \quad (2)$$

где вероятность простоя системы P_0 определяется следующим образом :

$$P_0 = \left(\sum_{k=0}^m \rho^k / k! + \rho^m / m! \sum_{k=1}^n (\rho / m)^k \right)^{-1}. \quad (3)$$

Исследуем вероятность потери P_l в системе Эрланга на монотонность как функцию коэффициента загрузки ρ . Продифференцировав по ρ выражение (1), получаем:

$$P_l' = \frac{\rho^{m-1}}{(m-1)!} \cdot \frac{\left(\sum_{k=0}^{m-1} \rho^k / k! (1 - \rho / m) + \rho^m / m! \right)}{\left(\sum_{k=0}^m \rho^k / k! \right)^2}.$$

Нетрудно увидеть, что числитель второго множителя этого выражения может быть записан как

$$\begin{aligned} & \sum_{k=0}^{m-1} \rho^k / k! (1 - \rho / m) + \rho^m / m! = \\ & = 1 + \sum_{k=1}^{m-1} (1 - k / m) \rho^k / k! > 0. \end{aligned}$$

Это доказывает (интуитивно ожидаемый результат), что вероятность потери возрастает с увеличением загрузки системы.

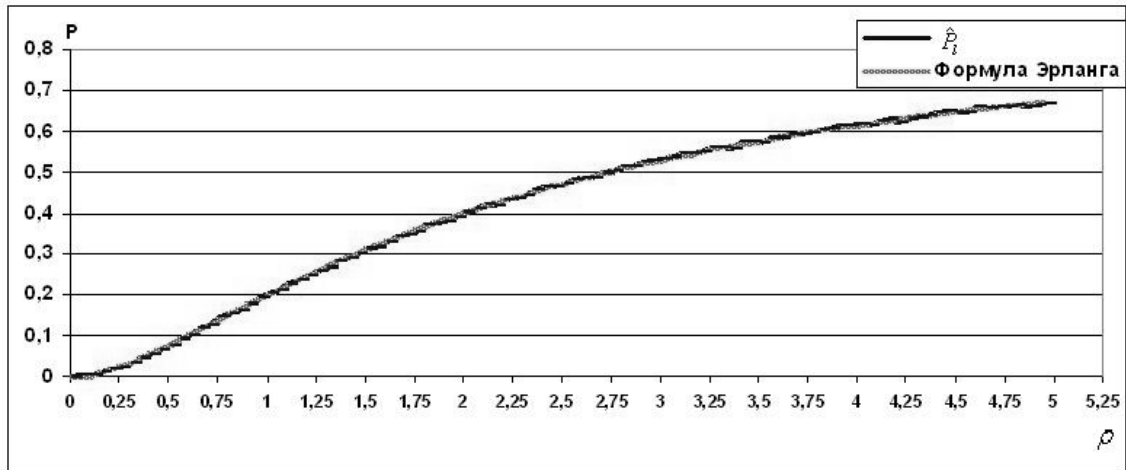


Рис. 1. Зависимость вероятности потери в системе $M/Pareto/2/0$ (с параметром распределения Парето $\alpha = 2,6$) от коэффициента загрузки ρ

Дальнейший анализ в случае $m=1$, дает $P_1'' = -2/(1+\rho)^3 < 0$, т. е. вероятность P_1 является выпуклой вверх функцией. А для $m=2$ нетрудно получить, что вероятность потери будет менять выпуклость вниз на выпуклость вверх в точке перегиба $\rho \approx 0,73$. Рис. 1 хорошо иллюстрирует установленные выше свойства выпуклости и монотонности на примере двухканальной системы Эрланга с распределением Парето времени обслуживания, т. е. $P(S > x) = x^{-\alpha}$, $x \geq 1$, (и $P(S > x) = 1$ при $x < 1$) с параметром $\alpha = 2,6$. Отметим также очень хорошее совпадение между оценкой вероятности потери \hat{P}_1 , вычисленной как отношение числа потерянных заявок к общему числу заявок, и ее точным стационарным значением, вычисленным по формуле Эрланга (1). (Этот результат может также служить тестом корректности программы имитационного моделирования.)

Расчет вероятности в системе $M/M/m/n$

Рассмотрим систему $M/M/m/n$ с пуассоновским входным потоком с параметром $\lambda > 0$ и экспоненциальным обслуживанием с параметром $\mu > 0$ и изучим зависимость вероятности потери P_1 от значения коэффициента трафика ρ . При выполнении условия $\rho/m < 1$ соответствующая система с неограниченным буфером ($n = \infty$) обладает стационарным режимом. В этом случае существует такой уровень, выше которого загрузка системы практически не под-

нимается. Если достижение этого уровня интерпретировать как вероятность потери (переполнения буфера в аналогичной системе с конечным буфером), то вероятность такого события оказывается очень малой. Заметим, что эта область исследований относится к теории оценивания вероятностей редких событий. Поэтому очевидно, что вероятность потери в системе с конечным буфером при фиксированном значении коэффициента трафика $\rho/m < 1$ должна быть весьма чувствительна к величине буфера. В частности, можно ожидать, что она убывает с ростом размера буфера. Действительно, это легко следует из формул (2), (3). Более того, с ростом величины буфера n , $P_1 = O((\rho/m)^n)$.

Таким образом, скорость убывания вероятности потери является экспоненциальной, что хорошо иллюстрирует принцип больших уклонов для данной простейшей системы, у которой распределение времени обслуживания (экспоненциальное) имеет так называемый легкий хвост [Glynn, Whitt, 1994].

Более интересная проблема – изучение поведения величины P_1 при большой загрузке системы, т. е. при $\rho/m > 1$. Это типичная ситуация в системах с потерями, где не возникает проблема нестационарности, и поэтому ограничения на коэффициент трафика не требуются. В этом случае предполагаем, что вероятность потери должна быть скорее чувствительна к коэффициенту трафика, чем к величине буфера. Для пояснения отметим, что при $\rho/m > 1$ система (в стационарном режиме) почти все время находится в состоянии максимальной загрузки независимо

от величины буфера. (Величина буфера, однако, влияет на время выхода системы на стационарный режим.) Для проверки этого предположения исследуем поведение вероятности P_l в системе $M/M/m/n$ с загрузкой $\rho/m > 1$ при неограниченном увеличении буфера. Из формулы (2) имеем:

$$P_l = \frac{\rho^{m+n} / (m!m^n)}{\sum_{k=0}^m \rho^k / k! + \rho^m / m! \sum_{k=0}^n (\rho/m)^k} \quad (4)$$

Выполнив замены $A := (\sum_{k=0}^m \rho^k / k!) / (\rho^m / m!)$,

$\delta := \rho/m$, перейдем к пределу при $n \rightarrow \infty$ и, учитывая, что $\delta > 1$:

$$\begin{aligned} \lim_{n \rightarrow \infty} P_l &= \lim_{n \rightarrow \infty} \frac{\delta^n}{A + \sum_{k=1}^n \delta^k} = \\ &= \lim_{n \rightarrow \infty} \frac{\delta^n}{A + \delta \frac{\delta^n - 1}{\delta - 1}} = \\ &= \lim_{n \rightarrow \infty} \frac{\delta^n (\delta - 1)}{A(\delta - 1) + \delta^{n+1} - \delta} = \\ &= \lim_{n \rightarrow \infty} \frac{1 - 1/\delta}{A(\delta - 1) / \delta^{n+1} + 1 - 1/\delta^n} = \\ &= 1 - 1/\delta = 1 - m/\rho. \end{aligned} \quad (5)$$

Таким образом, действительно, вероятность потери в системе $M/M/m/n$ в случае большой загрузки сходится при растущем размере буфера n к фиксированному значению, зависящему лишь от коэффициента ρ/m . Важно подчеркнуть, что переход к пределу при $n \rightarrow \infty$ в самой формуле (4) (при $\rho/m > 1$) не имеет смысла, так в этом случае система с бесконечным буфером оказывается нестационарной и поэтому, в частности, $P_0 = 0$.

В табл. 1 представлены результаты численного эксперимента (по 20 000 наблюдений), где сравниваются значения оценки вероятности потери \hat{P}_l , вычисленной как доля отказов и вероят-

ности потери P_l , вычисленной по формуле (4).

Полученная близость \hat{P}_l и P_l совершенно ожидаема и указывает на состоятельность оценки \hat{P}_l . Заметим, что строгое обоснование (сильной) состоятельности следует из регенеративной теории, что будет обсуждаться далее. Проблема, связанная с обоснованием состоятельности, состоит в том, что потери, точнее соответствующие индикаторы потерь, не являются н.о.р.с.в.

Таблица 1. Результаты имитационного моделирования системы $M/M/m/n$

ρ	m	n	\hat{P}_l	P_l
4,06	2	0	0,6162	0,6193
4,06	2	5	0,5072	0,5096
4,06	2	10	0,5093	0,5069
4,06	2	20	0,5116	0,5069
4,06	2	50	0,5101	0,5069
4,06	2	100	0,5035	0,5069
4,06	4	0	0,3146	0,3160
4,06	4	10	0,0852	0,0822
4,06	4	20	0,0564	0,0500
1,56	2	0	0,3228	0,3220
1,56	2	5	0,0521	0,0512
1,56	2	7	0,0011	0,0009
8	4	10	0,5034	0,5000
8	4	100	0,5054	0,5000
8	2	10	0,7493	0,7500
8	2	100	0,7418	0,7500
8	2	500	0,7275	0,7500

Кроме того, результаты экспериментов подтверждают, что при $\rho/m < 1$, оценка \hat{P}_l убывает с увеличением буфера, а при $\rho/m > 1$, оценка \hat{P}_l перестает зависеть от его размера и стремится к фиксированному значению, определяемому формулой (5). Рис. 2 иллюстрирует последний результат на примере системы $M/M/2/n$ при $n = 50$.

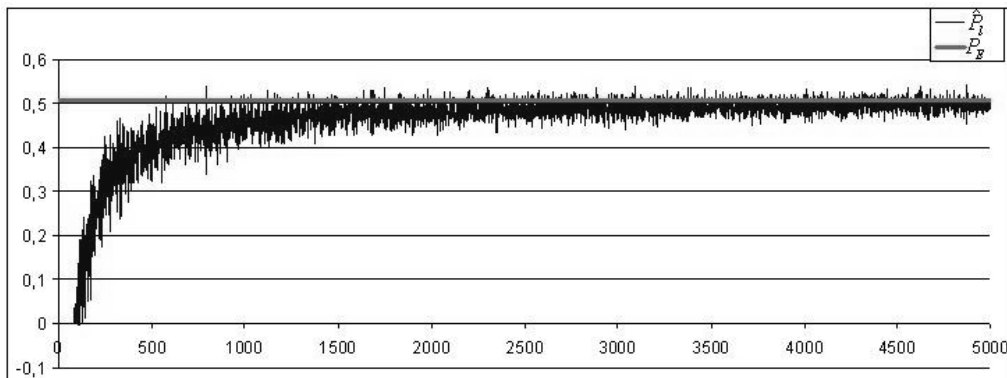


Рис. 2. Зависимость вероятности потери в системе $M/M/2/50$ от числа поступивших заявок при значении коэффициента загрузки $\rho = 4,06$

В заключение отметим, что вероятность потери в системе $M/M/1/0$ $P_1 = \rho/(1+\rho)$ действительно приближается с $1-1/\rho$, т. е. с правой частью формулы (5) при $m=1$ и при большом ρ . Интересно отметить, что аналогичный результат (однако доказываемый гораздо более сложно) верен и для более общей системы $GI/M/m/n$, см. [Абрамов, 2008]. Можно ожидать, что результат верен и для системы с потерями $M/GI/m/n$.

Потери на цикле занятости в системе $GI/G/1/n$

В данном разделе, который опирается на работу [Wolff, 2002], приведены результаты для среднего числа потерь EL_n на цикле занятости системы (который является также циклом регенерации всех типичных процессов системы, см. [Perkoz et al., 2003]) в зависимости от размера буфера n . Этот результат полезен для вычисления вероятности потери. Пусть A – количество поступивших заявок, K – число обслуженных заявок, и $L_n = A - K$ – число потерь, все – за один период занятости. Обозначим через β – длительность цикла, включающего период занятости B и следующий за ним период простоя I , т. е. $\beta = B + I$. Таким образом, имеем (стохастические) равенства

$$\beta = \sum_{i=1}^A T_i, B = \sum_{i=1}^K S_i. \quad (6)$$

Естественно предположить, что в системе с потерями $EA < \infty$. Это условие выполняется в случае пуассоновского входного потока и дисциплины обслуживания FIFS (заявки обслуживаются в порядке поступления в систему). Для других распределений и при $\rho > 1$ необходимы дополнительные условия, иначе возможно возникновение ситуации, при которой система никогда не станет свободной. Применяя тождество Вальда к равенствам (6), получим: $\lambda E\beta = EA$, $\mu EB = EK$, что в итоге дает

$$EL_n = \lambda E\beta - \mu EB. \quad (7)$$

Рассмотрим два важных случая распределения интервала входного потока. Говорят, что интервал между приходами заявок T обладает свойством НСЛИ (новое в среднем лучше использованного), если среднее незавершенное время между приходами меньше среднего времени между приходами, т. е. $E(T-t|T>t) \leq ET$, $t \geq 0$. (Принятый англоязычный термин – NBUE: *new better than used*

in expectation. Русскоязычная терминология (НСЛИ) дается в [Байхельт, Франкен, 1988].) Противоположное свойство НСХИ (NWUE – *new worse than used in expectation*) означает, что среднее незавершенное время между приходами больше среднего времени между приходами, т. е., $E(T-t|T>t) \geq ET$, $t \geq 0$. Отметим, что в первом случае $EI \leq ET$, а во втором $EI \geq ET$. Отсюда и из равенства (7) легко получается следующий результат [Wolff, 2002].

Теорема 1. В системе $GI/G/1/n$

1) если $\lambda \leq \mu$ и T удовлетворяет свойству НСЛИ, то

$$EL_n \leq 1, \quad n \geq 0,$$

2) если $\lambda \geq \mu$ и T удовлетворяет свойству НСХИ, то

$$EL_n \geq 1, \quad n \geq 0.$$

В обоих случаях при $\lambda = \mu$ выполняется равенство $EL_n = 1$ для любого $n \geq 0$.

Заметим, что равенство $EL_n = 1$ (при $\lambda = \mu$) легко следует из (7) в системе $M/GI/1/n$, поскольку для пуассоновского входного потока $EI = 1/\lambda$.

Можно рассмотреть более общую одноканальную систему с бесконечным буфером $GI/G/1$, в которой, тем не менее, заявки могут быть потеряны по различным причинам. К примеру, если в системе находится j заявок, то новая заявка встает в очередь с вероятностью p_j .

Получается система с конечным буфером n при $p_j = 1$, если $j \leq n$ и $p_j = 0$, если $j > n$. В качестве еще одного примера можно предложить правило, где поступившая заявка, встречающаяся в системе работу больше некоторой константы v , теряется. В общем случае заявка, встречающаяся в системе работу, равную v , встает в очередь с вероятностью q_v , $v \geq 0$. Перечисленные правила удовлетворяют следующему условию J : для любого $i \geq 1$ событие $\{A \geq i \text{ и } i\text{-я заявка встает в очередь}\}$ не зависит от последовательности $\{S_i, S_{i+1}, \dots, T_i, T_{i+1}, \dots\}$. При последнем условии времена обслуживания заявок, вставших в очередь, образуют последовательность н.о.р.с.в. (Заявки могут обслуживаться в произвольном порядке.)

Теорема 2. Пусть в системе с потерями $GI/G/1$ выполняется условие J . Тогда

1) если $\lambda \leq \mu$ и T удовлетворяет свойству НСЛИ, то среднее число потерь на цикле

$$EL \leq 1;$$

2) если $\lambda \geq \mu$ и T удовлетворяет свойству НСХИ, то

$$EL \geq 1.$$

В обоих случаях при $\lambda = \mu$ выполняется равенство $EL = 1$.

Регенеративный метод оценивания

В данном разделе кратко описан регенеративный метод имитационного моделирования. Кроме того, этот метод применен для моделирования работы систем с конечным буфером и доверительного оценивания вероятности потери в таких системах.

Процесс $X = \{X_t, t \in T\}$, где $T = [0, \infty)$ или $T = \{0, 1, \dots\}$, называется *регенерирующим процессом*, если существует бесконечная последовательность *моментов регенерации* $0 = \beta_0 < \beta_1 < \beta_2 < \dots$ таких, что сегменты (*циклы регенерации*) $G_p = (X_t, \beta_{k-1} \leq t < \beta_k)$ являются н.о.р. [Serfozo, 2009]. Разумеется, интервалы $\alpha_k = \beta_k - \beta_{k-1}$, при $k \geq 1$, также н.о.р.с.в. Обозначим через β типичную длину цикла регенерации, считая, что первая заявка поступает в пустую систему в момент времени $t = 0$ (процесс без задержки). Будем предполагать, что процесс X является положительно возвратным, т. е. $E\beta < \infty$, и длина цикла регенерации β – непериодическая, т. е. $P(\beta = 1) > 0$. (Условие положительной возвратности обеспечивается конечностью буфера, исключающей неограниченный рост значений процесса.) Тогда существует (стационарный) предел по распределению $X_k \Rightarrow X$ при $k \rightarrow \infty$ такой, что $P(X < \infty) = 1$. Более того, если f – измеримая функция, то (для процесса с дискретным временем) верно следующее соотношение:

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^k f(X_i) = \frac{E \sum_{i=1}^{\beta} f(X_i)}{E\beta} = Ef(X) = r. \quad (8)$$

(Аналог для случая непрерывного времени очевиден.) Предполагается также, что

$$E \sum_{i=0}^{\beta} |f(X_i)| < \infty.$$

Для оценки неизвестного параметра r , являющегося искомой стационарной характеристикой регенерирующего процесса, объединим данные в пределах одного цикла регенерации и получим н.о.р. с.в.

$$Y_p := \sum_{k=\beta_{p-1}}^{\beta_p-1} f(X_k), \quad p \geq 0.$$

Если выполнено (наиболее слабое) условие $0 < E(Y - r\beta)^2 < \infty$,

$$(9)$$

то можно применить *регенеративную Центральную Предельную Теорему* [Glynn, Iglehart, 1993]:

$$k^{1/2} \bar{\alpha}_k |r - r_k| \Rightarrow \sigma N(0, 1), \quad k \rightarrow \infty,$$

где

$$\bar{\alpha}_k = \frac{1}{k} \sum_{i=1}^k \alpha_i, \quad r_k := \frac{\bar{Y}_k}{\bar{\alpha}_k}, \quad \bar{Y}_k = \frac{1}{k} \sum_{i=1}^k Y_i$$

– соответствующие выборочные средние, $\sigma^2 = E(Y - r\beta)^2$, $N(0, 1)$ – стандартная нормальная с. в., а k – число завершённых циклов регенерации в процессе моделирования.

Таким образом, последовательность $k^{1/2} \bar{\alpha}_k [r_k - r] \sigma$ сходится (по распределению) к стандартному нормальному закону с ростом числа циклов регенерации. Следовательно, $100(1 - \gamma)\%$ доверительный интервал для неизвестного оцениваемого параметра r имеет вид

$$\left[r_k - \frac{z_\gamma s(k)}{\bar{\alpha}_k \sqrt{k}}, r_k + \frac{z_\gamma s(k)}{\bar{\alpha}_k \sqrt{k}} \right],$$

где квантиль z_γ находится из условия $2\Phi(z_\gamma) = 1 - \gamma$, $\Phi(x)$ – функция Лапласа, а $s^2(k)$ – эмпирическая дисперсия, причем $s^2(k) \rightarrow \sigma^2$, $k \rightarrow \infty$ с вероятностью 1. (Подробнее описание регенеративного метода имитационного моделирования см. в [Bratley, Fox, 1987], [Law, Kelton, 1991], [Glynn, Iglehart, 1993].)

Важным аспектом изучения систем с потерями является анализ потока отказов, возникающих из-за переполнения буфера. Такой поток может являться входным потоком для других телекоммуникационных устройств. Для того чтобы проиллюстрировать регенеративный метод оценивания, опишем регенеративную структуру системы $M/G/m/n$ и построим доверительный интервал для оценки \hat{P}_l вероятности потери заявки, получаемой как отношение числа отказов к общему числу поступивших заявок.

Пусть $\{t_i\}$ – моменты прихода заявок в систему и v_i – число заявок в системе в момент прихода заявки $i \geq 1$. Легко понять, что регенерации в процессе $\{v_i\}$ возникают при приходе заявки в пустую систему. Определим рекурсивно

$$\beta_{k+1} = \min\{l > \beta_k : v_l = 0\}, \quad k \geq 0,$$

где $\beta_0 = 0$ (первый приход в пустую систему в момент времени $t = 0$). Тогда $\{\beta_k\}$ – есть моменты

регенерации процесса $\{v_i\}$ (и других процессов) в данной системе обслуживания. (Данная конструкция легко может быть расширена на случай ненулевых начальных условий.)

Для того чтобы проиллюстрировать применение регенеративного метода для построения точечных и интервальных оценок вероятности потери, рассмотрим m – серверную систему $M/G/m/n$ с конечным буфером.

Вначале рассмотрим систему $M/Pareto/m/n$ с распределением Парето времени обслуживания с параметром $\alpha > 2$, т. е. $P(S \geq x) = x^{-\alpha}$, $x \geq 1$. В настоящее время общепризнано, что распределение Парето (распределение с тяжелым хвостом) достаточно хорошо описывает трафики широкого класса современных телекоммуникационных систем [Leland et al., 1994], [Taqqi et al., 1997], [Willinger et al., 1997], в то время как экспоненциальное распределение (распределение с легким хвостом) также традиционно продолжает играть важную роль в таком анализе. В случае конечного буфера средняя длина цикла регенерации конечна, $E\beta < \infty$, и кроме того, величина β является непериодической ввиду свойства входного потока. Поскольку $\alpha > 2$, то время обслуживания имеет конечный второй момент, $ES^2 < \infty$. В свою очередь это влечет конечность второго момента длины цикла регенерации, $E\beta^2 < \infty$ [Wolff, 1989].

Более того, при оценивании вероятности потери из этого факта также следует $EL_n^2 < \infty$, поскольку величина L_n (число потерь на цикле регенерации) в данном случае составлена из индикаторов потерь на цикле и поэтому $L_n \leq \beta$ [Mogozov, 2004]. Другими словами, в данном случае можно применить доверительное оценивание, основанное на регенеративном методе. Очевидно, это же верно и в случае экспоненциального времени обслуживания, поскольку тогда $ES^2 = 1/\mu^2 < \infty$.

Замечание. Напомним, что система регенерирует каждый раз, когда заявка поступает в пустую систему. Это означает, что момент регенерации не является событием в потоке отказов, поскольку отказ происходит только в момент прихода при переполненном буфере. Другими словами, нельзя обнаружить момент регенерации, наблюдая лишь за потоком потерь. С этой точки зрения моменты регенерации можно назвать *скрытыми*.

Число потерь на интервале дискретного времени $[0, i]$ (считающего приходы заявок) можно задать в следующем виде:

$$L_n(i) = \sum_{j=1}^i I_j,$$

где индикатор $I_j = 1$, если j -я заявка потерялась.

Так как циклы регенерации

$$G_k = \{I_i, \beta_{k-1} \leq i < \beta_k\} \text{ н.о.р.,}$$

то и фрагменты потока отказов, состоящие из заявок, потерянных на циклах $[\beta_{k-1}, \beta_k)$, $k \geq 1$, также будут н. о. р. с. в.

Таким образом, последовательность $I = \{I_i, i \geq 1\}$ образует регенерирующий процесс, и выполняется следующее предельное соотношение:

$$\lim_{k \rightarrow \infty} \frac{1}{k} L_n(k) = \frac{E \sum_{i=1}^{\beta} I_i}{E\beta} = \frac{EL}{E\beta} \equiv P_l := r.$$

Ниже приведены результаты численного моделирования для построения $(1-\gamma)\%$ доверительного интервала для оценки вероятности потери в системе $M/G/m/n$ при $\gamma = 0,95$, т. е. $z_\gamma = 1,96$. Численные значения результатов моделирования представлены в табл. 2, где

$$\Delta(k) = \frac{z_\gamma s(k)}{\bar{\alpha}_k \sqrt{k}} - \text{ширина доверительного интервала}$$

для оценки вероятности потери, построенного по k завершенным циклам регенерации.

В частности, из табл. 2 видно, что оценка вероятности потери, полученная на основе регенеративного метода, близка к значению, полученному по формуле (4) для системы с экспоненциальным обслуживанием. Заметим, что для систем с распределением Парето времени обслуживания, данная формула неприменима. Полученные данные также показывают, что повысить число циклов можно, уменьшая размер буфера или увеличивая число каналов. Поведение интервальной оценки, в зависимости от числа циклов регенерации k , представлено на рис. 3. (Фактически, это иллюстрирует регенеративную Центральную Предельную Теорему.)

Далее проиллюстрируем утверждения Теоремы 1 на примере системы $M/M/1/n$, в которой верны соотношения $EL_n \leq 1$, $EL_n = 1$, $EL_n \geq 1$, при $\rho \leq 1$, $\rho = 1$, $\rho \geq 1$, соответственно. Заметим, что интервал между приходами заявок T обладает как свойством НСЛИ, так и свойством НСХИ, и поэтому Теорема 1 применима. В частности, при $\rho = 1$ получаем

$$P_l = \frac{EL_n}{E\beta} = \frac{1}{E\beta}, \quad (10)$$

Таблица 2. Результаты регенеративного моделирования системы $M / G / m / n$

Service	ρ	m	n	k	\hat{L}_n	$\hat{P}_l(k)$	$\Delta(k)$	$\hat{P}_l(k) - \Delta(k)$	$\hat{P}_l(k) + \Delta(k)$	P_l
Pareto	1,28	1	0	8 789	1,27	0,5604	0,0072	0,5532	0,5675	–
Pareto	1,28	1	5	759	6,56	0,2494	0,0211	0,2283	0,2706	–
M	1	1	0	10 042	0,99	0,4977	0,007	0,4907	0,5048	0,5
M	1	3	0	18 646	0,07	0,0676	0,0043	0,0633	0,0719	0,0625
M	4,06	2	0	7 570	1,64	0,6213	0,0077	0,6136	0,629	0,6193
M	4,06	2	10	13	472,23	0,5203	0,0104	0,5099	0,5307	0,5069
M	4,06	4	0	13 678	0,46	0,316	0,0084	0,3076	0,3244	0,316
M	4,06	4	10	3 198	0,47	0,0756	0,0106	0,065	0,0862	0,0822
M	4,06	4	20	1 718	0,7	0,0604	0,0131	0,0473	0,0734	0,05

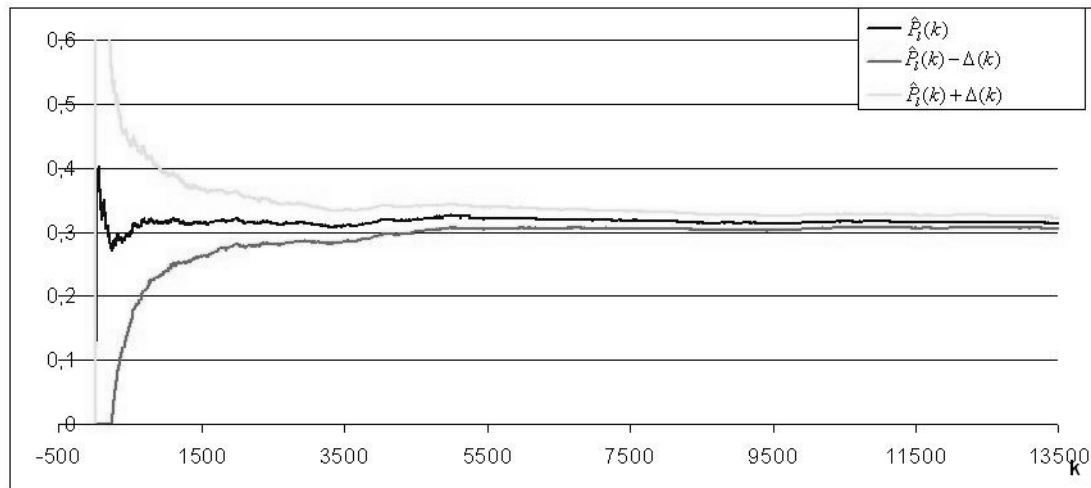


Рис. 3. Зависимость оценки и границ доверительного интервала в системе $M / M / 4 / 0$ от числа регенераций при значении коэффициента загрузки $\rho = 4,06$

Таблица 3. Результаты регенеративного моделирования системы $M / M / 1 / n$

λ	μ	ρ	n	k	\hat{L}_n	$\hat{P}_l(k)$	$\Delta(k)$	$\hat{P}_l(k) - \Delta(k)$	$\hat{P}_l(k) + \Delta(k)$	P_l
7	1,5	4,67	5	4	3456,750	0,7823	0,0076	0,7748	0,7899	0,7857
5	3	1,67	5	340	24,230	0,4123	0,0122	0,4001	0,4244	0,4115
4	4	1	3	3 991	0,980	0,1965	0,0091	0,1873	0,2056	0,2000
3	5	0,6	3	8 699	0,140	0,0621	0,0167	0,0454	0,0788	0,0562
1,5	8	0,19	3	16 096	0,001	0,0013	0,0007	0,0006	0,0019	0,0010

что позволяет использовать выборочное среднее длины цикла регенерации при оценке P_l . Численные результаты доверительного оценивания представлены в табл. 3, где P_l – вероятность потери заявки, вычисленная по формуле (4), $\hat{P}_l(k)$ – оценка, построенная по k завершённым циклам регенерации, $\Delta(k)$ – ширина доверительного интервала для оценки вероятности потери. Эксперименты проведены для систем с размером буфера $n \leq 5$, что даёт достаточное число циклов регенерации за приемлемое время моделирования. Из табл. 3 видно, что оценка \hat{P}_l вероятности потери, полученная с использованием регенеративного метода, близка к значению P_l , вычисленной по формуле (4). Рис. 4 де-

монстрирует зависимость ширины доверительного интервала для \hat{P}_l от числа циклов регенерации k . Из табл. 3 видно, что эксперименты подтверждают зависимость оценки \hat{L}_n среднего EL_n от ρ , установленную Теоремой 1. В частности, видно, что при $\rho = 1$, значение \hat{L}_n близко к 1. В целом, результаты, полученные в ходе моделирования, дают хорошее согласие с утверждениями Теоремы 1.

Заключение

В статье рассматриваются различные вопросы, связанные с вероятностью потери в системах обслуживания с конечным буфером. Рассмотрены некоторые аналитические свойства

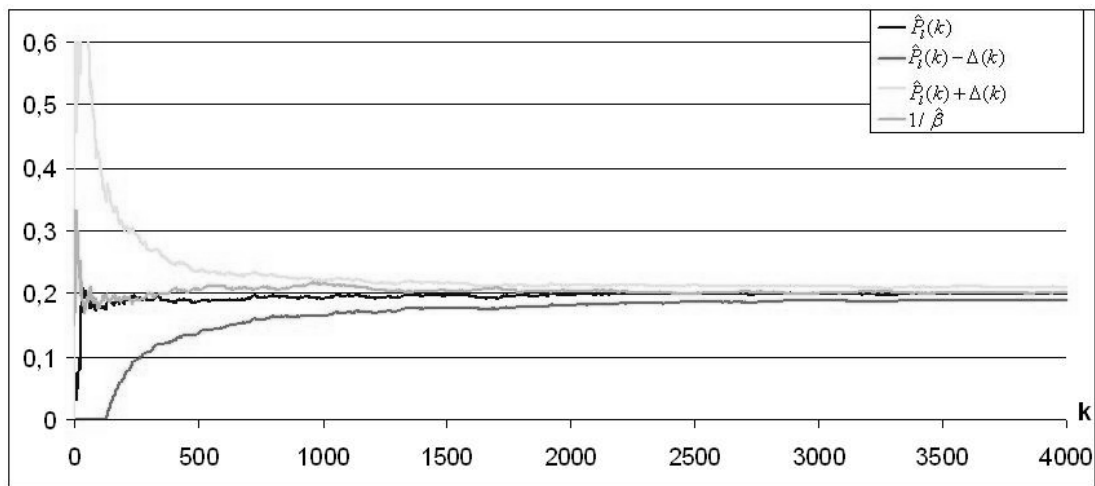


Рис. 4. Зависимость оценки и границ доверительного интервала в системе $M/M/3/0$ и величины, обратной оценке длины цикла от числа регенераций при значении коэффициента загрузки $\rho = 1$

вероятности потери, в частности, свойство монотонности. Изучена зависимость вероятности потери от величины буфера. В частности, показано, что если коэффициент трафика больше 1, то эта вероятность асимптотически перестает зависеть от величины буфера, когда он растет. Рассмотрены также свойства среднего числа потерь на цикле занятости системы, являющегося также циклом регенерации, в зависимости от типа распределения входного потока. Другим важным аспектом данной работы является использование регенеративного имитационного моделирования для точечного и доверительного оценивания вероятности потери. В статье также приведены численные результаты моделирования.

Литература

Байхельт Ф., Франкен П. Надежность и техническое обслуживание. Математический подход // Радио и связь. М., 1988. С. 392.

Abramov V. M. One property of a refusals stream // J. Appl. Prob. 1997. Vol. 34, N. 3. P. 800–805.

Abramov V. M. Takács' Asymptotic Theorem and Its Applications: A Survey // Acta Applicandae Mathematica / Springer Science+Business Media B.V. 2008. Vol. 109. P. 609–651.

Asmussen Applied Probability and Queues / New York: Springer, 2003. 2nd ed.

Bratley P., Fox B. L. A guide to simulation / New York: Springer-Verlag, 1987. 2nd ed.

Glynn P. W., Iglehart D. L. Simulation methods for queues: an overview // Queueing systems. 1988. Vol. 3. P. 221–237.

Glynn P. W., Iglehart D. L. Conditions for the applicability of the regenerative method // Management Science. 1993. 39. P. 1108–1111.

Glynn P. W., Whitt W. Logarithmic asymptotics for steady-state tail probabilities in a single-server queue // Adv. Appl. Prob. 1994. P. 131–156.

Greiner M., Jobmann M., Kluppelberg C. Telecommunication traffic, queuing models and subexponential distributions // Queueing Systems. 1999. Vol. 33. P. 125–152.

Law A. M., Kelton W. D. Simulation modeling and analysis/ New York: McGraw- Hill, 1991. 2nd ed.

Leland W., Taqqu M., Willinger W., Wilson D. On the Self-Similar Nature of Ethernet Traffic // IEEE/ACM Transactions on Networking, 1994. Vol. 2. N. 1. P. 1–15.

Morozov E. Weak regeneration in modeling of queueing processes // Queueing Systems. 2004. 46. P. 295–315.

Morozov E. Communications systems: rare event simulation and effective bandwidths/ Pamplona: Universidad Publica de Navarra, 2004.

Perkoz E. A., Richter R., Xia C. H. Characterizing Losses during busy periods in finite buffer systems // J. Appl. Prob. 2003. 40. P. 242–249.

Resnick S. Heavy tail modeling and teletraffic data // The Annals of Statistics. 1997. Vol. 25. N. 5. P. 1805–1869.

Serfozo R. Basics of applied stochastic processes / Berlin: Springer-Verlag, 2009.

Taqqu M. S., Willinger W., Sherman R. Proof of a fundamental result in self-similar traffic modeling // Computer Communication Review. 1997. Vol. 27. P. 5–23.

Willinger W., Taqqu M., Sherman R., Wilson D. Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level // IEEE/ACM Transactions on Networking. 1997. Vol. 5. N. 1. P. 71–86.

Wolff R. W. Stochastic Modeling and the Theory of Queues/ Prentice-Hall, 1989.

Wolff R. W. Loses per cycle in a single-server queue // J. Appl. Prob. 2002. P. 905–909.

СВЕДЕНИЯ ОБ АВТОРАХ:

Горичева Руслана Сергеевна

аспирантка

Институт прикладных математических исследований
КарНЦ РАН

ул. Пушкинская, 11, Петрозаводск, Республика Карелия,
Россия, 185910

эл. почта: lana--@mail.ru

тел.: (8142) 763370

Goricheva, Ruslana

Institute of Applied Mathematical Research, Karelian Research
Centre, Russian Academy of Science

11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia

e-mail: lana--@mail.ru

tel.: (8142) 763370

Морозов Евсей Викторович

ведущий научный сотрудник, д. ф.-м. н.

Институт прикладных математических исследований
КарНЦ РАН

ул. Пушкинская, 11, Петрозаводск, Республика Карелия,
Россия, 185910

эл. почта: emorozov@karelia.ru

тел.: (8142) 763370

Morozov, Evsey

Institute of Applied Mathematical Research, Karelian Research
Centre, Russian Academy of Science

11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia

e-mail: emorozov@karelia.ru

tel.: (8142) 763370