

УДК 519.872.1

СРАВНЕНИЕ ДВУХ ОЦЕНОК ЭФФЕКТИВНОЙ ПРОПУСКНОЙ СПОСОБНОСТИ СИСТЕМЫ ОБСЛУЖИВАНИЯ

А. В. Бородина, Е. В. Морозов

*Институт прикладных математических исследований
Карельского научного центра РАН
Петрозаводский государственный университет*

В статье рассматриваются методы оценивания эффективной пропускной способности системы с постоянной скоростью обслуживания. Проводится сравнение оценок эффективной пропускной способности, построенных с помощью метода группового среднего и регенеративного метода. Регенеративный подход предлагается в качестве альтернативы традиционному методу группового среднего, который не учитывает структуру входного потока. С помощью имитационного моделирования проводится анализ свойств обеих оценок, а также их сравнение с известными аналитическими результатами и с оценкой по методу Монте-Карло. Численные эксперименты показывают определенные преимущества регенеративной оценки над оценкой по методу группового среднего.

Ключевые слова: система с постоянной скоростью обслуживания, эффективная пропускная способность, теория больших уклонений, регенеративная оценка, метод группового среднего.

A. V. Borodina, E. V. Morozov. COMPARISON OF TWO QUEUING SYSTEM EFFECTIVE BANDWIDTH ESTIMATES

This article discusses methods for estimating effective bandwidth in a system with constant service rate. The effective bandwidth estimates constructed using the group average method and the regenerative method are compared. The regenerative approach is suggested as an alternative to the traditional batch means method, which does not take into account the structure of the input stream. Simulation helps analyze the properties of both estimates and compare the experimental values with known theoretical results and Monte-Carlo evaluation. These experiments confirm the conservatism of regenerative estimates, and good consistency with known theoretical values.

Key words: system with constant service rate, effective bandwidth, large deviation theory, regenerative estimator, batch means method.

ВВЕДЕНИЕ

Быстрое развитие телекоммуникационных технологий неразрывно связано с проблемой улучшения технической оснащенности и

производительности устройств связи, которые гарантировали бы требуемое качество сетевого обслуживания (QoS гарантии). Для изучения современных IP-ориентированных мультисервисных сетей с высокоагрегированным

трафиком приходится привлекать методы теории очередей, теории больших уклонений, ускоренные методы моделирования редких событий. В зависимости от назначения, сетевые устройства могут предъявлять различные требования к определенному уровню качества обслуживания. Например, передача данных требует определенного уровня надежности, связанного с вероятностью перегрузки и потери данных; мультимедийные приложения предъявляют повышенные требования к пропускной способности канала, а для VoIP-сетей более важны показатели задержки и джиттера (т. е. вариации задержки передачи пакета в сети). Другим примером является управление энергопотреблением мобильных устройств, которое связано с оцениванием вероятности превышения нагрузкой буфера передатчика. В беспроводных каналах связи с затуханием не только коэффициент затухания и интерференция могут существенно снижать пропускную способность канала, но также параметры средней задержки и вероятность переполнения буфера, см. [2]. Кроме того, на уровень надежности критически влияет *эффективная пропускная способность* (ЭПС) устройства, которая гарантирует заданную (малую) вероятность превышения стационарной нагрузкой заданного уровня (или объема буфера).

В данной статье с помощью имитационного моделирования сравниваются оценки ЭПС обслуживающего устройства, принимающего регенеративный входной поток. Рассматривается оценка на основе метода группового среднего и оценка, учитывающая регенеративную структуру входного процесса.

ЭФФЕКТИВНАЯ ПРОПУСКНАЯ СПОСОБНОСТЬ

Одним из важнейших показателей качества обслуживания является вероятность превышения $P_b := P(W > b)$ стационарным процессом нагрузки W (т. е. незавершенной работы) некоторого достаточно большого уровня b . Такая вероятность переполнения тесно связана с ЭПС. Если мощность обслуживающего устройства C (величина работы, которую прибор может сделать за единицу времени) можно выбирать, то естественная задача QoS состоит в выборе такого значения C , которое гарантирует, что вероятность P_b не превысит заданной (малой) величины Γ , т. е.

$$P_b := P(W > b) \leq \Gamma. \quad (1)$$

Упомянутая задача называется *выбором ЭПС*. Пусть $W(n)$ есть незавершенная работа в мо-

мент (дискретного времени) n , v_i – величина работы, поступившая в интервале $(i-1, i]$, тогда $V(n) = \sum_{i=1}^n v_i$ есть суммарная работа, поступившая в интервале $(0, n]$. Имеет место рекурсия Линдли

$$W(n+1) = [W(n) + v_n - C]^+, \quad n = 0, 1, \dots \quad (2)$$

Если буфер для ожидания неограничен, то предполагается также выполненным следующее условие стационарности

$$EX = Ev - C := \lambda - C < 0. \quad (3)$$

(Здесь и далее типичный элемент последовательности н. о. р. с. в. обозначается без соответствующего индекса.) Процесс W для широкого класса систем обслуживания удовлетворяет *принципу больших уклонений* [4, 6], именно:

$$\lim_{b \rightarrow \infty} \frac{1}{b} \ln P(W > b) = -\theta^*, \quad (4)$$

где $\theta^* > 0$ – некоторый (искомый) параметр. Это влечет экспоненциальную аппроксимацию вероятности большого уклонения стационарного процесса нагрузки, т. е.

$$P(W > b) = e^{-\theta^* b + o(b)} \asymp e^{-\theta^* b}, \quad (5)$$

(\asymp означает логарифмическую асимптотику). Соотношения (1), (5) дают такое значение параметра θ^* :

$$\theta^* = -\ln \Gamma / b > 0. \quad (6)$$

Далее предполагается, что существует конечный (для всех θ в некоторой положительной окрестности $(0, \theta_0)$) предел

$$\Lambda_V(n) := \frac{1}{n} \ln E e^{\theta V(n)} \rightarrow \Lambda_V(\theta), \quad n \rightarrow \infty, \quad (7)$$

который называется (нормированной) предельной логарифмической производящей функцией моментов входного процесса. Из теории больших уклонений [4] следует, что ЭПС определяется из условия

$$\Lambda_V(\theta^*) = \theta^* C. \quad (8)$$

В данной статье не обсуждаются подробно условия существования предела $\Lambda_V(\theta^*)$ (называемого функционалом действия). В простейшем случае н. о. р. с. в. $\{v_i\}$, в частности, требуется $E e^{\theta^* v} < \infty$ при $\theta^* \in (0, \theta_0)$. (Условие $E e^{\theta^* v} < \infty$ означает, что с. в. v имеет так называемое распределение с *легким хвостом*.) В более общем случае требуется, чтобы последовательность $\{v_i\}$ была стационарной с перемешиванием. Ввиду (8) ЭПС определяется (разумеется, приближенно) из условия

$P(W > b) = \Gamma = e^{-\theta^* b}$ как

$$C := \frac{\Lambda_V(\theta^*)}{\theta^*} = \frac{\Lambda_V(-\ln \Gamma/b)b}{-\ln \Gamma}. \quad (9)$$

Этот результат верен также и для системы с конечным буфером (большого) размера b , т. е. для системы с потерями.

ОЦЕНИВАНИЕ ЭПС МЕТОДОМ ГРУППОВОГО СРЕДНЕГО

Когда с. в. $\{v_i\}$ — н. о. р., то функция $\Lambda_V(\theta^*) = \ln \mathbb{E} e^{\theta^* v}$. (Напомним, что $\mathbb{E} e^{\theta^* v} < \infty$.) Несмещенной и сильно состоятельной оценкой функции Λ_V является выборочное среднее, т. е. с вероятностью 1 (с в. 1)

$$\ln \frac{1}{k} \sum_{i=1}^k e^{\theta^* v_i} \rightarrow \Lambda_V(\theta^*), \quad k \rightarrow \infty. \quad (10)$$

Гораздо труднее оценить Λ_V для зависимых $\{v_i\}$. Наиболее распространенным методом оценивания функции Λ_V является метод группового среднего (*batch-means method*), когда последовательность $\{v_i\}$ разбивается на блоки данных одинаковой длины B . Существуют разные подходы к выбору величины B и способу формирования блоков [10, 11, 14]. (В частности, могут рассматриваться блоки с перекрытием, когда часть данных попадает в два соседних блока ([8, 12].) Далее рассматривается стандартный вариант разбиения с блоками фиксированной длины B , определяемыми следующим образом:

$$\hat{X}_j = \sum_{i=(j-1)B+1}^{jB} v_i, \quad j \geq 1.$$

Основное предположение состоит в том, что при большом B эти блоки являются (приближенно) н. о. р. Считая, что общее число наблюдений $n = kB$ ратно величине блока, получим

$$\frac{1}{n} \ln \mathbb{E} e^{\theta \sum_{i=1}^n v_i} = \frac{\ln \mathbb{E} e^{\theta \hat{X}}}{B} := \Lambda_V(\theta, B), \quad (11)$$

где \hat{X} означает типичный блок (и предполагается, что $\ln \mathbb{E} e^{\theta \hat{X}} < \infty$, $\theta \in (0, \theta_0)$). Равенство в (11) надо понимать приближенно, так как возможная зависимость между блоками игнорируется. Очевидно, при фиксированном размере блока B , выборочная оценка

$$\hat{\Lambda}_k(\theta, B) := \frac{1}{B} \ln \frac{1}{k} \sum_{i=1}^k e^{\theta \hat{X}_i} \quad (12)$$

является сильно состоятельной оценкой функции $\Lambda_V(\theta, B)$, т. е. при $k \rightarrow \infty$

$$\hat{\Lambda}_k(\theta, B) \rightarrow \Lambda_V(\theta, B) \text{ с в. 1.}$$

Соответствующая оценка ЭПС определяется тогда следующим образом (см. (9)):

$$\hat{C}_k(\theta^*, B) = \frac{\hat{\Lambda}_k(\theta^*, B)}{\theta^*}, \quad \theta^* = -\ln \Gamma/b. \quad (13)$$

Оценка $\hat{C}_k(\theta^*, B)$ является смещенной (см. [9]). Действительно, поскольку функция \ln является вогнутой, то неравенство Йенсена дает

$$\begin{aligned} \mathbb{E} \left[\hat{C}_k(\theta^*, B) \right] &= \mathbb{E} \left[\frac{1}{\theta^* B} \ln \frac{1}{k} \sum_{i=1}^k e^{\theta^* \hat{X}_i} \right] \\ &= \frac{1}{\theta^* B} \mathbb{E} \left[\ln \frac{1}{k} \sum_{i=1}^k e^{\theta^* \hat{X}_i} \right] \\ &\leq \frac{1}{\theta^* B} \ln \mathbb{E} \left[\frac{1}{k} \sum_{i=1}^k e^{\theta^* \hat{X}_i} \right] \\ &= \frac{1}{\theta^* B} \ln \frac{1}{k} \sum_{i=1}^k \mathbb{E} \left[e^{\theta^* \hat{X}_i} \right]. \end{aligned}$$

В силу предположения о независимости блоков $\{\hat{X}_i\}$ отсюда следует

$$\mathbb{E} \left[\hat{C}_k(\theta^*, B) \right] \leq \frac{1}{\theta^* B} \ln \mathbb{E} \left[e^{\theta^* \hat{X}} \right] = C(\theta^*, B). \quad (14)$$

Ниже это смещение оценки $\hat{C}_k(\theta^*, B)$ иллюстрируется результатами численного моделирования. Отметим, что при выборе числа наблюдений кратным числу блоков, т. е. при $n = kB$, оценки по шкале блоков и по шкале наблюдений очевидным образом согласованы:

$$\hat{\Lambda}_k(\theta, B) = \frac{1}{B} \ln \frac{B}{n} \sum_{i=1}^{n/B} e^{\theta \hat{X}_i}. \quad (15)$$

Замечание 1. Смещенность оценки (13) в сторону меньших значений говорит о том, что она не очень пригодна для оценки ЭПС в системах высокой надежности. Кроме того, основной проблемой при данном подходе является выбор размера блока B : небольшой размер блока может не соответствовать предположению о независимости, а слишком большая величина неэффективна с точки зрения требуемого времени моделирования.

РЕГЕНЕРАТИВНАЯ ОЦЕНКА ЭПС

В работах [3, 13] для входной последовательности $\{v_n\}$ с моментами регенерации β_k и периодами регенерации $\alpha_k = \beta_{k+1} - \beta_k$, было предложено строить оценку ЭПС путем группировки данных по циклам регенерации, что приводит к *регенеративным блокам* вида

$$\hat{X}_k := \sum_{i=\beta_k}^{\beta_{k+1}-1} v_i, \quad k \geq 0, \quad \beta_0 = 0, \quad (16)$$

которые, действительно, являются н. о. р. В основе этого подхода лежит предположение о том, что более точный учет структуры зависимости между данными $\{v_i\}$ может улучшить качество оценки (уменьшить дисперсию) по сравнению с оценкой в методе группового среднего. В предположении

$$\begin{aligned} E\alpha < \infty, \quad \ln E e^{\theta^* \hat{X}} < \infty, \quad \theta^* \in (0, \theta_0), \\ E(\alpha - E\alpha)^2 := \sigma^2 \in (0, \infty), \end{aligned} \quad (17)$$

рассмотрим функцию $\Lambda_V(n)$ из (7) и определим число регенераций для n наблюдений как $k(n) = \max(k \geq 0 : \beta_k \leq n)$. Рассмотрим, как отразится регенеративная структура входного процесса на свойствах выборочной оценки. Учитывая неотрицательность всех рассматриваемых величин, имеем

$$\begin{aligned} \frac{1}{n} \ln E \left[e^{\theta^* \sum_{i=1}^{k(n)} \hat{X}_i} \right] &\leq \frac{1}{n} \ln E \left[e^{\theta^* \sum_{i=1}^n v_i} \right] \\ &\leq \frac{1}{n} \ln E \left[e^{\theta^* \sum_{i=1}^{k(n)+1} \hat{X}_i} \right]. \end{aligned}$$

В силу моментных условий (17) из теории восстановления следует, что величина работы $\sum_{i=k(n)}^n v_i$, поступающей на текущем (в момент n) цикле регенерации, с в. 1 есть величина $o(n)$, $n \rightarrow \infty$. Далее, по элементарной теореме восстановления, $E k(n)/n \rightarrow 1/E\alpha$. Поэтому в силу независимости блоков можно ожидать, что при $n \rightarrow \infty$

$$\begin{aligned} \frac{1}{n} \ln E e^{\theta^* \sum_{i=1}^n v_i} &\approx \\ \frac{E k(n)}{n} \frac{1}{E k(n)} \ln E e^{\theta^* \sum_{i=1}^{k(n)} \hat{X}_i} &\rightarrow \frac{1}{E\alpha} \ln E e^{\theta^* \hat{X}}. \end{aligned}$$

Обозначим

$$\Lambda_{REG}(\theta^*) := \frac{1}{E\alpha} \ln E e^{\theta^* \hat{X}}. \quad (18)$$

Приведенные выше аргументы приводят к такой *регенеративной оценке*, которая строится по k регенеративным блокам,

$$\hat{\Lambda}_k(\theta^*) := \frac{k}{\beta_k} \ln \frac{1}{k} \sum_{i=1}^k e^{\theta^* \hat{X}_i}. \quad (19)$$

Поскольку $\beta_k/k \rightarrow E\alpha$, то с в. 1

$$\hat{\Lambda}_k(\theta^*) \rightarrow \Lambda_{REG}(\theta^*), \quad k \rightarrow \infty,$$

и в предельной функции присутствует *средний размер блока* $E\alpha$. Таким образом, регенеративную оценку ЭПС естественно определить как

$$\hat{C}_k(\theta^*) = \frac{\hat{\Lambda}_k(\theta^*)}{\theta^*}. \quad (20)$$

(В обзорной работе [7] также рассматривается возможность использовать блоки случайной длины, в частности, отражающие структуру зависимости между величинами $\{v_i\}$.) В работе [1] было показано, что

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln E e^{\theta^* \sum_{i=1}^n v_i} \geq \Lambda_{REG}(\theta^*), \quad (21)$$

т. е., что $\Lambda_{REG}(\theta^*)$ является нижней границей для функции $\Lambda_V(\theta^*)$. Однако обоснование (той же самой) верхней границы для (18) пока остается открытой проблемой. Тем не менее результаты численных экспериментов позволяют сделать предположение о строгой состоятельности и *консервативности* регенеративной оценки $\hat{C}_k(\theta^*)$. Консервативность означает, что оценка несколько завышает требования к ЭПС, т. е. всегда гарантирует выполнение условия (1) (однако, вообще говоря, с вероятностью, меньшей чем Γ). Это смещение очень важно для систем *высокой надежности*.

Для получения регенеративной оценки ЭПС можно использовать два основных сценария. Первый предполагает искусственно заданный входной поток, с известным распределением длины цикла регенерации α . В этом случае можно заранее потребовать выполнение необходимых моментных условий как для α , так и для объема \hat{X} поступающей на цикле работы. (Заметим, что эти условия в любом случае исключают распределения \hat{X} с *тяжелым хвостом*.) Во втором сценарии можно рассматривать двухузловую тандемную сеть. На вход узла 1 поступает процесс восстановления (к примеру, пуассоновский) с интенсивностью λ и н. о. р. временами обслуживания $\{S_n\}$ со средним $ES = 1/\mu$ при условии стационарности $\rho_1 := \lambda/\mu < 1$. В этом случае, как хорошо известно, выходной процесс из первого узла является положительно возвратным регенерирующим, т. е. $E\alpha < \infty$. Поэтому на вход узла 2 поступает положительно возвратный регенерирующий процесс, моменты регенерации β_n которого определяются как моменты прихода заявок в пустой узел 1. Однако в

качестве моментов регенерации мы будем рассматривать *моменты ухода из узла 1 тех заявок, которые пришли в пустой узел 1*. Именно, пусть t_n – момент прихода заявки с номером n в узел 1, t_n^* – момент ее ухода из узла $n \geq 1$. Тогда упомянутые моменты регенерации определяются следующим образом:

$$\beta_0 = 0, \beta_{n+1} = \min_k (t_k^* > \beta_n : \nu(t_k^-) = 0), n \geq 0, \quad (22)$$

где $\{\nu(t), t \geq 0\}$ – непрерывный справа процесс очереди в узле 1. Хотя в моменты $\{\beta_n\}$ узел 1, вообще говоря, не пустой, легко проверить, что распределение накопленной в этот момент нагрузки является одним и тем же и не зависит от n . (Это вариант классической, но *невыврожденной* регенерации. Такие моменты $\{\beta_n\}$ связаны с событиями выходного потока и более удобны для процесса имитационного моделирования.) Хотя длины циклов $\alpha_n = \beta_{n+1} - \beta_n$ выражаются в числе поступивших заявок, в методе группового среднего поступающая нагрузка рассматривается на единичных интервалах времени (слотах). Чтобы иметь возможность сравнивать оценки, получаемые обоими методами, необходимо перейти от шкалы, считающей заявки, поступающие в узел 2, к шкале слотов. Будем далее считать, что заявки в узел 2 поступают по одной в слоте. На практике это предположение может означать, что заявки из узла 1 поступают в узел 2 через некоторый синхронизирующий буфер достаточно большого объема. Такие буфера широко применяются в коммуникационных сетях. Например, пакеты от одного отправителя к одному и тому же получателю могут иметь в сети различные маршруты. Это ведет к разбросу значений задержки пакетов (джиттеру) и нарушает синхронизацию исходного трафика. Величина джиттера является важным показателем качества обслуживания. В частности, в широко используемых в настоящее время VoIP сетях, неконтролируемый джиттер может существенно влиять на качество передаваемого звука. Одним из вариантов сокращения джиттера является введение синхронизирующего буфера, который позволяет восстановить равные промежутки времени между пакетами при получении (см. [5]). Таким образом, для сравнения двух оценок достаточно моделировать процесс незавершенной нагрузки согласно соотношению (1).

При проведении экспериментов рассматривалась некоторая комбинация двух описанных выше сценариев. Именно, в узел 2 поступает последовательность данных v_i (по одному на единичных слотах), однако регенеративные

блоки формируются в соответствии с циклами регенерации выходного потока из (гипотетического) узла 1. Кроме того, в большинстве экспериментов предполагалось, что с. в. $\{v_i\}$ являются н. о. р. Даже при таких упрощающих предположениях, представленные ниже результаты численного моделирования демонстрируют важные свойства описанных выше оценок.

РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ ЭКСПЕРИМЕНТОВ

Целью численных экспериментов является выяснение свойств регенеративной оценки ЭПС и их сравнение со свойствами оценки по методу группового среднего. (Эти эксперименты существенно расширяют результаты, полученные в работе [1].) Опираясь на сделанные выше предположения, рассмотрим модель с интервалами единичной длины между поступлениями заявок в узел 1 и входной последовательностью н. о. р. экспоненциальных с. в. $\{v_n\}$ с параметром μ_2 . (Случай зависимых с. в. $\{v_n\}$ рассматривается в эксперименте 7.)

Хотя для независимых с. в. $\{v_i\}$ группировка не требуется, тем не менее, интересно сравнить обе оценки (13), (20) с известным аналитическим значением ЭПС и с оценкой, полученной методом Монте-Карло. Это трудно (или невозможно) сделать в случае зависимых $\{v_i\}$. Более того, такой анализ может быть полезен, если, например, независимость $\{v_i\}$ априори неизвестна. Поскольку в рассматриваемом случае

$$E e^{\theta^* v} = \int_0^\infty e^{\theta^* x} \mu_2 e^{-\mu_2 x} dx = \frac{\mu_2}{\mu_2 - \theta^*},$$

то ЭПС (9) принимает вид

$$C = \frac{\ln\left(\frac{\mu_2}{\mu_2 - \theta^*}\right)}{\theta^*}, \quad (23)$$

где необходимо выбрать параметр $\mu_2 > \theta^* = -\ln \Gamma/b$. Для удобства будем использовать сокращения *метод Р* для регенеративного метода и *метод ГС* для метода группового среднего, соответственно. Также назовем величину $e^{\theta^* \hat{X}}$ *экспоненциальной нагрузкой* на блоке. Введем следующие обозначения:

- B – размер блока в методе ГС;
- k – количество блоков или циклов регенерации;
- \hat{C}_{REG} – оценка ЭПС с помощью метода Р, вычисляется по формуле (20);

- \hat{C}_{BM} — оценка ЭПС методом ГС, вычисляется по формуле (13);
- \hat{C}_{MC} — оценка ЭПС методом Монте-Карло в случае н. о. р. $\{v_i\}$;
- n_{REG} — число наблюдений, проведенных при моделировании методом Р;
- n_{BM} — число наблюдений, проведенных при моделировании методом ГС, $n_{BM} = Bk$;
- $V_{REG}[\alpha]$ — выборочная оценка дисперсии длины цикла регенерации;
- $V_{REG}[e^{\theta^* \hat{X}}]$ — выборочная оценка дисперсии экспоненциальной нагрузки на цикле-блоке в методе Р;
- $V_{BM}[e^{\theta^* \hat{X}}]$ — выборочная оценка дисперсии экспоненциальной нагрузки на блоке в методе ГС;
- $V[\hat{C}_{REG}]$ — выборочная дисперсия регенеративной оценки;
- $V[\hat{C}_{BM}]$ — выборочная дисперсия оценки по методу ГС;
- $V_{REG}[C]$ — среднеквадратическое отклонение регенеративной оценки от теоретического значения ЭПС C ;
- $V_{BM}[C]$ — среднеквадратическое отклонение оценки по методу ГС от теоретического значения ЭПС C .

Эксперимент 1. В табл. 1 приведены результаты моделирования описанной выше системы при $\Gamma = 10^{-8}$, $b = 60$ (т. е. $\theta^* = 0,307011$) и н. о. р. показательных $\{v_i\}$ с параметром μ_2 . Напомним, что в этом случае можно сравнить результаты моделирования с явным значением C из (23).

Таблица 1. Оценивание ЭПС при умеренной нагрузке узлов

ρ_1	μ_2	\hat{C}_{BM}	C	\hat{C}_{REG}	n_{REG}
0,5	0,6	1,88705	2,33475	2,65482	1007
0,5	0,9	1,18603	1,35897	1,39296	1003
0,5	1,3	0,79928	0,87749	1,03916	10047
0,3	2,5	0,414878	0,42677	0,42745	19815

Число наблюдений n_{BM} , зависящее от заданной величины блока B и заданного числом блоков k , принималось равным 1000, 10000 и

20000, соответственно. Для метода Р число наблюдений n_{REG} определялось в процессе моделирования на (целом) числе циклов, исходя из условия $n_{REG} \geq n_{BM}$. Значение коэффициента ρ_1 и параметра $\mu_2 > \theta^*$ соответствуют умеренной нагрузке на оба узла. При этом заметные изменения коэффициента μ_2 с 0,6 до 2,5 почти не влияют на оценки \hat{C}_{REG} и \hat{C}_{BM} , которые оказываются близки к теоретическому значению C . В частности, из табл. 1 видно, что наиболее близкое к теоретическому значению C метод Р дает при $\rho_1 = 0,3$ и $\mu_2 = 2,5$. Данный результат обусловлен тем, что при выбранной нагрузке узла 1 длины циклов регенерации (размер блоков) варьируются слабо. (Вариация длин циклов в зависимости от коэффициента загрузки ρ_1 обсуждается в последующих экспериментах.) Заметим также, что $\hat{C}_{BM} < C < \hat{C}_{REG}$, и в частности, эксперименты подтверждают консервативность оценки \hat{C}_{REG} .

Эксперимент 2. Сохранив значения Γ , b , θ^* такими же как в эксперименте 1, возьмем $B = 1500$, $n_{BM} = 30000$ и, уменьшая значение параметра $\mu_2 (> \theta^* = 0,3070011)$, увеличим тем самым нагрузку на узел 2.

Таблица 2. Оценивание ЭПС при увеличении нагрузки на узел 2

ρ_1	μ_2	\hat{C}_{BM}	C	\hat{C}_{REG}	n_{REG}
0,4	0,5	2,09545	3,10079	4,32887	29680
0,3	0,4	2,61066	4,75223	5,04857	30236
0,4	0,35	2,97821	6,83036	11,9739	29700
0,6	0,32	3,27348	10,4369	32,6247	26611

Из табл. 2 видно, что при этом растет различие между C и обеими оценками ЭПС. Такое поведение оценки \hat{C}_{REG} можно объяснить тем, что даже незначительное увеличение ρ_1 при большой нагрузке на узел 2 ведет к значительному росту дисперсии суммарной нагрузки, поступающей на цикле. Тем не менее, оценка \hat{C}_{REG} остается консервативной. Напротив, оценка \hat{C}_{BM} нечувствительна к увеличению ρ_1 , а большие значения B нивелируют дисперсию величины нагрузки, поступающей на блоке. Однако значение оценки \hat{C}_{BM} существенно меньше, чем C .

Таким образом, решение, основанное на \hat{C}_{BM} , может привести к драматической ошибке в выборе ЭПС и, как следствие, к серьезной перегрузке системы обслуживания. В то

же время, решение на основе (консервативной) оценки \hat{C}_{REG} может быть полезно при расчете высоконадежных систем.

Эксперимент 3. Как было отмечено ранее, для н. о. р. $\{v_n\}$ можно либо найти явное значение функции $\Lambda(\theta^*) = \ln E e^{\theta^* v}$, либо использовать для ее оценки стандартное выборочное среднее \hat{C}_{MC} на основе метода Монте-Карло. Рассмотрение этого простейшего случая позволяет экспериментально оценить качество оценок \hat{C}_{BM} и \hat{C}_{REG} путем их сравнения либо с известным значением, либо с оценкой \hat{C}_{MC} . В табл. 3 даны результаты вычисления оценок ЭПС при различных значениях параметра θ^* и при $\rho_1 = 0, 4, \mu_2 = 0, 9$.

Таблица 3. Сравнение оценок ЭПС

θ^*	\hat{C}_{REG}	\hat{C}_{MC}	C	\hat{C}_{BM}
0,8008	3,37599	2,29366	2,75472	1,4685
0,7084	3,3423	2,13291	2,18415	1,44997
0,6140	2,48644	1,82857	1,86717	1,42216
0,4605	1,87695	1,55257	1,5565	1,38025
0,3684	1,58691	1,42744	1,42918	1,35425
0,3349	1,51458	1,38869	1,38967	1,3329

Варьирование параметра θ^* происходит за счет изменения Γ и b . Для всех методов оценки строились по 10000 наблюдений. Отметим, что оценка \hat{C}_{BM} во всех случаях дает заниженное значение ЭПС.

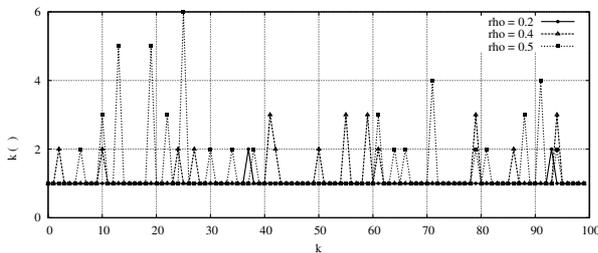


Рис. 1. Длины циклов выходного потока узла 1 при умеренной нагрузке

Эксперимент 4. Чтобы проанализировать, каким образом дисперсия длины цикла регенерации влияет на качество оценки \hat{C}_{REG} , снова исследуем случаи умеренной и сильной загрузки узла 1. Выберем $\mu_2 = 1, 5$, $\Gamma = 10^{-8}$, $b = 60$ и будем варьировать коэффициент загрузки ρ_1 .

На рис. 1 и 2 показан характер изменения длины цикла регенерации выходного потока из узла 1 при заданном ρ_1 для 100 циклов. На рис. 2 (при большом ρ_1) видны большие колебания длины цикла, что существенно влияет на суммарную величину нагрузки в регенеративном блоке.

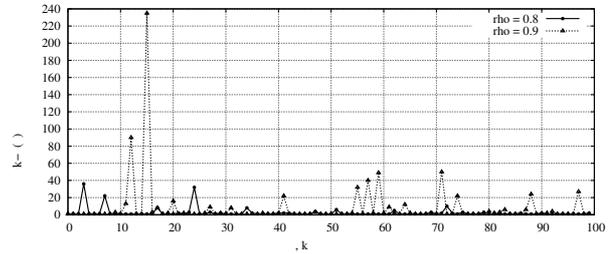


Рис. 2. Длины циклов выходного потока узла 1 при большой нагрузке

Соответствующее данным рис. 1 и 2 изменение величины $e^{\theta^* \hat{X}}$ на циклах регенерации показано на рис. 3 и 4. Рис. 3 отражает изменение величины $e^{\theta^* \hat{X}}$ при умеренном ρ_1 , а рис. 4 — при большом значении ρ_1 .

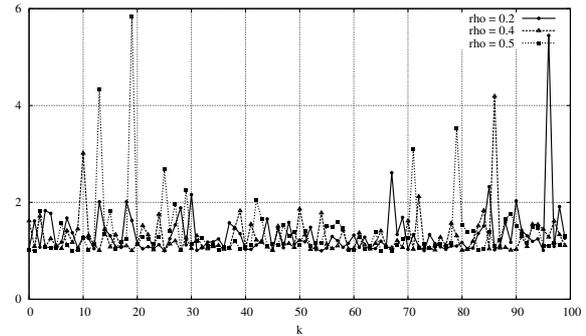


Рис. 3. Экспоненциальная нагрузка узла 2 на блоках при умеренной нагрузке узла 1

Видно, что небольшие изменения длины цикла приводят к значительному изменению экспоненциальной нагрузки. Таким образом, колебания длины цикла при увеличении ρ_1 влекут существенный рост дисперсии регенеративной оценки \hat{C}_{REG} .

Заметим, что для $\rho_1 = 0, 8$ максимум экспоненциальной нагрузки $e^{\theta^* \hat{X}_3} = 4709, 35$ (см. рис. 4), что соответствует максимальной длине цикла $\alpha_3 = 36$ (см. рис. 2). Для $\rho_1 = 0, 9$ максимум равен $e^{\theta^* \hat{X}_{15}} = 7, 63318e + 22$ при $\alpha_{15} = 235$.

Таблица 4. Зависимость дисперсии оценки ЭПС от ρ_1

ρ_1	0,1	0,3	0,4	0,5	0,7	0,8
$V_{REG}[\alpha]$	0,0001	0,063635	0,249424	0,753058	7,91327	34,38
$V_{REG}[e^{\theta^* \hat{X}}]$	0,744988	1,01771	4,32138	51,2856	3,82251e+16	3,90876e+35
$V_{BM}[e^{\theta^* \hat{X}}]$	2,78e+163	3,94e+156	5,30e+159	8,69e+156	4,23e+155	2,36e+159
$V[\hat{C}_{REG}]$	4,394e-4	4,39e-4	1,13e-3	1,022	59,10	566,80
$V[\hat{C}_{BM}]$	6,64e-4	6,13e-4	6,81e-4	8,85e-4	6,29e-4	7,15e-4

Следует отметить, что изменение ρ_1 не влияет на оценку \hat{C}_{BM} , так как блоки имеют фиксированный размер B . Это также подтверждается результатами следующего эксперимента.

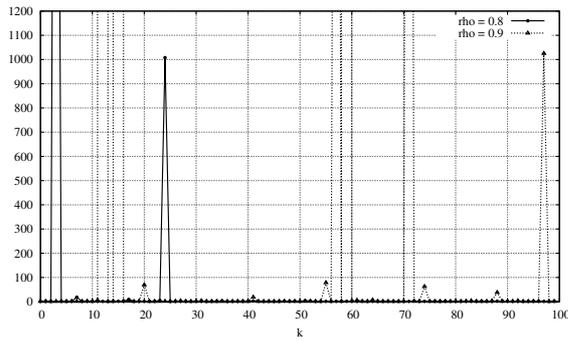


Рис. 4. Экспоненциальная нагрузка узла 2 на блоках при сильной загрузке узла 1

Эксперимент 5. Выберем $B = 500$, $\mu_2 = 0,9$, $\theta^* = 0,307011$, $b = 60$, число наблюдений $n_{BM} = 10000$, размер выборки для вычисления оценок дисперсии равен 50.

В табл. 4 приведены выборочные дисперсии оценок экспоненциальной нагрузки в блоке для метода ГС и метода Р, а также оценки ЭПС для обоих методов при различных значениях ρ_1 .

Изменение ρ_1 не влияет на вариацию экспоненциальной нагрузки в блоке для метода ГС, однако, как видно из табл. 4, значение $V_{BM}[e^{\theta^* \hat{X}}]$ стабильно велико (имеет порядок 10^{156}). Тем не менее, эти колебания в блоке нивелируются за счет выбора большого значения B , и поэтому выборочная дисперсия оценки \hat{C}_{BM} оказывается невелика. Эта оценка по-прежнему смещена в сторону меньших значений.

Эксперимент подтверждает, что рост дисперсии оценки \hat{C}_{REG} прямо пропорционален росту дисперсии длины цикла, которая, в свою очередь, связана с увеличением коэффициента

ρ_1 . При этом отметим незначительную разницу в дисперсии обеих оценок \hat{C}_{REG} и \hat{C}_{BM} при малых и умеренных значениях ρ_1 .

Замечание 2. Важно отметить, что в системах с высокими требованиями QoS наиболее интересным является случай средней и малой загрузки. И именно в этих случаях эксперименты подтверждают преимущество оценки \hat{C}_{REG} над оценкой \hat{C}_{BM} .

Таблица 5. Отклонение оценок ЭПС от теоретического значения C

ρ_1	0,1	0,3	0,5	0,7
$V[\hat{C}_{REG}]$	0,344	0,482	64,37	295,13
$V[\hat{C}_{BM}]$	5,69e-3	6,4e-3	5,56e-3	9,13e-3
$V_{REG}[C]$	54,48e-3	3,20e-3	1,89	34,97
$V_{BM}[C]$	36,86	36,91	37,05	36,70

Эксперимент 6. Теперь сравним оба метода, увеличив нагрузку на узел 2, когда объемы поступающей нагрузки распределены показательно с параметром $\mu_2 = 0,35$. Будем изменять ρ_1 на первом узле, а остальные параметры оставим без изменения (см. эксперимент 5). Поскольку для рассматриваемой модели известно точное значение ЭПС, то можно найти среднееквадратическое отклонение обеих оценок \hat{C}_{REG} и \hat{C}_{BM} от теоретического значения C . (Выше мы обозначили эти отклонения $V_{REG}[C]$ и $V_{BM}[C]$, соответственно.)

Из табл. 5 видно, что отклонение $V[\hat{C}_{BM}]$ оценки \hat{C}_{BM} от выборочного среднего остается небольшим, однако, отклонение $V_{BM}[C]$ от известного значения C стабильно велико. Отклонение $V[\hat{C}_{REG}]$ оценки \hat{C}_{REG} от выборочного среднего больше, чем у \hat{C}_{BM} , но она существенно ближе к C (за исключением больших значений ρ_1). Более того, эксперименты снова подтверждают установленный ранее результат $\hat{C}_{BM} < C < \hat{C}_{REG}$.

Таблица 6. Сравнение оценок ЭПС для зависимых с. в. v ($\eta \sim exp$) на циклах

n	50	100	500	1000	5000	10000	20000
\hat{C}_{REG}	0,846919	0,829186	0,883955	0,87963	0,877652	0,879206	0,878917
\hat{C}_{BM}	0,801576	0,801067	0,847855	0,864714	0,899761	0,846297	0,857191
$V[\hat{C}_{REG}]$	0,0135	0,0141	0,0072	0,0017	0,000361	0,0001497	7,84896e-05
$V[\hat{C}_{BM}]$	0,0177	0,0116	0,005	0,0029	0,0023	0,00114	0,001046

Таблица 7. Сравнение оценок ЭПС для зависимых v ($\eta \sim weib$) на циклах

n	100	500	1000	5000	10000	20000
\hat{C}_{REG}	0,576674	0,575671	0,579387	0,576897	0,576467	0,57678
\hat{C}_{BM}	0,570715	0,568081	0,567508	0,569535	0,569611	0,56978
$V[\hat{C}_{REG}]$	0,000649	8,89196e-05	4,73991e-05	1,12482e-05	7,6528e-06	4,28579e-06
$V[\hat{C}_{BM}]$	0,000747277	0,00010108	7,78538e-05	1,37676e-05	9,87332e-06	3,90953e-06

Таким образом, в случае н. о. р. $\{v_i\}$ при малой и средней нагрузке первого узла, оценка \hat{C}_{REG} имеет сравнительно небольшую дисперсию и при этом ее отклонение от известного значения C существенно меньше, чем у оценки \hat{C}_{BM} . Таким образом, оценка \hat{C}_{REG} имеет явные преимущества перед оценкой \hat{C}_{BM} .

Эксперимент 7. Однако на практике, как правило, имеет место зависимость между с. в. $\{v_i\}$, и тогда получить теоретическое значение ЭПС (9) очень сложно (или невозможно), в силу трудности вычисления функции моментов (7). Для построения оценок ЭПС в этом случае рассмотрим пример зависимости между v_i на цикле регенерации следующего вида:

$$v_j = \frac{\sum_{k=1}^j \eta_k}{j}, \quad (24)$$

где с. в. $\{\eta_k\}$ — н. о. р., j — номера заявок на цикле регенерации, α_i — длина i -го цикла, $1 \leq j \leq \alpha_i$, $i \in [1, k]$. Такую зависимость можно считать сильной в том смысле, что все с. в. на одном цикле зависимы (причем, функционально). На самом деле зависимость может быть произвольной и определяться, к примеру, протоколом передачи данных. В то время как \hat{C}_{REG} строится по действительно независимым блокам, при построении оценки \hat{C}_{BM} не исключена ситуация, когда зависимые данные попадают в разные блоки. Это обстоятельство может негативно отразиться на дисперсии оценки.

Пусть $b = 80$, $\Gamma = 10^{-8}$, $\theta^* = 0,230259$, $\rho_1 = 0,4$. Дисперсия оценок вычисляется по

выборке размера 100. Будем варьировать общее количество наблюдений n . В табл. 6 приведены результаты для экспоненциальной с. в. η с параметром $\mu_2 = 1,3$. Как и в предыдущих экспериментах, $\hat{C}_{BM} < \hat{C}_{REG}$, и кроме того, оценка \hat{C}_{REG} имеет меньшую дисперсию. (Напомним, что в случае независимых данных дисперсия \hat{C}_{REG} была, как правило, больше.) В табл. 7 представлены оценки при $\rho_1 = 0,4$ для с. в. η с распределением Вейбулла (*weib*) с легким хвостом,

$$F_\eta(x) = 1 - e^{-\gamma x^c}, \quad \gamma > 0, \quad c \geq 1, \quad (25)$$

где $\gamma = 3$, $c = 4$, а зависимость на цикле определяется по формуле (24).

Из анализа данных табл. 7 можно сделать вывод о близости оценок \hat{C}_{REG} и \hat{C}_{BM} . При этом (см. табл. 6) для экспоненциальной с. в. η дисперсия \hat{C}_{REG} существенно меньше чем у \hat{C}_{BM} . Таким образом, в рассмотренном примере выбор длины цикла регенерации в качестве длины блока позволяет более эффективно оценить ЭПС.

ЗАКЛЮЧЕНИЕ

В статье обсуждаются свойства регенеративной оценки и оценки по методу группового среднего для нахождения ЭПС 2-го узла в tandemной сети. Хотя случай зависимых данных во входном потоке узла 2 также рассмотрен, основное внимание уделено свойству оценок для н. о. р. данных. При упрощающих предположениях показано, что широко используемая

оценка по методу группового среднего занижает требования к ЭПС, в то время как регенеративная оценка является консервативной. Последнее свойство может рассматриваться как преимущество при анализе и проектировании высоконадежных коммуникационных систем.

Работа выполнена в рамках Программы стратегического развития на 2012–2016 гг. «Университетский комплекс ПетрГУ в научно-образовательном пространстве Европейского Севера: стратегия инновационного развития». Работа поддержана РФФИ, проект 10-07-00017.

ЛИТЕРАТУРА

1. *Бородина А., Дюденко И., Морозов Е.* Ускоренное оценивание вероятности переполнения регенеративных систем обслуживания. ОППМ. 2009. Т. 16, вып. 4. С. 577–593.
2. *De Turck K.* Performance Evaluation of Buffers in Wireless Networks, PhD thesis, University of Ghent, 2010–2011.
3. *Dyudenko I., Morozov E., Pagano M.* Regenerative estimator for effective bandwidth, Proceedings of the International Conference "Mathematical methods for analysis and optimization of information telecommunication networks". Minsk: Belarusian State University, 2009. С. 58–60.
4. *Ganesh A., O'Connell N., Wischik D.* Big Queues. Springer-Verlag, Berlin, 2004. 260 p.
5. *Ganz A., Ganz Z., Wongthavarawat K.* Multimedia Wireless Networks: Technologies, Standards, and QoS. Prentice Hall PTR. 2003. 352 p.
6. *Glynn P. W., Whitt W.* Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. Journal of Applied Probability 31A. 1994. P. 131–156.
7. *Lewis J. T., Russell R.* An introduction to large deviation for teletraffic engineers, DIAS Report, 1996.
8. *Meketon M. S., Schmeiser B. W.* Overlapping batch means: Something for nothing? // Proceedings of the Winter Simulation Conference. 1984. P. 227–230.
9. *Rabinovitch P.* Statistical estimation of effective bandwidth, M.Sc.thesis, University of Cambridge, 2000. 75 p.
10. *Schmeiser B.* Batch size effects in the analysis of simulation output. Oper. Res. 30. 1982. P. 556–568.
11. *Song W.-M. T.* On the estimation of optimal batch sizes in the analysis of simulation output. European Journal of Operational Research. 1996. Vol. 88, N 2. P. 304–319, January 20.
12. *Song W. T., Mingchang Chih.* Implementable mse-optimal dynamic partial-overlapping batch means estimators for steady-state simulations // Proceedings of the 2008 Winter Simulation Conference. 2008. P. 426–435.
13. *Vorobieva I., Morozov E., Pagano M., Procissi G.* A New Regenerative Estimator for Effective Bandwidth Prediction // Proceedings of AMICT 2007. Petrozavodsk, Russia, 2008. P. 175–186.
14. *Welch P. D.* On the relationship between batch means, overlapping batch means and spectral estimation // Proceedings of the Winter Simulation Conference, 1987. P. 320–323.

СВЕДЕНИЯ ОБ АВТОРАХ:

Бородина Александра Валентиновна
младший научный сотрудник, к. ф.-м. н.
Институт прикладных математических исследований
Карельского научного центра РАН
ул. Пушкинская, 11, Петрозаводск, Республика Карелия, Россия, 185610
старший преподаватель кафедры ИМО
Петрозаводский государственный университет
пр. Ленина, 33, Петрозаводск, Республика Карелия, Россия, 185910
эл. почта: borodina@krc.karelia.ru
тел.: (8142) 763370

Морозов Евсей Викторович
ведущий научный сотрудник, д. ф.-м. н.
Институт прикладных математических исследований
Карельского научного центра РАН
ул. Пушкинская, 11, Петрозаводск, Республика Карелия, Россия, 185910
профессор кафедры ПМиК
Петрозаводский государственный университет
пр. Ленина, 33, Петрозаводск, Республика Карелия, Россия, 185910
эл. почта: emorozov@karelia.ru
тел.: (8142) 763370

Borodina, Alexandra
Institute of Applied Mathematical Research, Karelian
Research Centre, Russian Academy of Sciences
11 Pushkinskaya St., 185610 Petrozavodsk, Karelia,
Russia
Department of Computer Science
Petrozavodsk State University
33 Lenina St., 185910, Petrozavodsk, Karelia, Russia
e-mail: borodina@krc.karelia.ru
tel.: (8142) 763370

Morozov, Evsey
Institute of Applied Mathematical Research, Karelian
Research Centre, Russian Academy of Sciences
11 Pushkinskaya St., 185610 Petrozavodsk, Karelia,
Russia
Professor at the Department of Applied Mathematics and
Cybernetics
Petrozavodsk State University
33 Lenina St., 185910, Petrozavodsk, Karelia, Russia
e-mail: emorozov@karelia.ru
tel.: (8142) 763370