



Math-Net.Ru

Общероссийский математический портал

Н. И. Головки, В. О. Каретник, В. Е. Танин, И. И. Сафонюк, Исследование моделей систем массового обслуживания в информационных сетях, *Сиб. журн. индустр. матем.*, 2008, том 11, номер 2, 50–64

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.139.83.210

20 ноября 2024 г., 02:09:31



ИССЛЕДОВАНИЕ МОДЕЛЕЙ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ В ИНФОРМАЦИОННЫХ СЕТЯХ

**Н. И. Головки, В. О. Каретник,
В. Е. Танин, И. И. Сафонюк**

Исследуются информационные сети (ИС) с точки зрения информационных потоков. Построены модели систем массового обслуживания (СМО) в ИС, показаны емкость накопителя и количество обслуживающих приборов. С применением статистических методов исследованы модели СМО и классифицированы типы возможных входных потоков и законы распределения интервалов времени обслуживания. Приводятся рекомендации для применения моделей СМО в ИС.

Ключевые слова: информационные сети, статистический анализ, система массового обслуживания, функция распределения, экспоненциальное обслуживание, дважды стохастический пуассоновский поток.

1. ПОСТРОЕНИЕ ПРИКЛАДНЫХ МОДЕЛЕЙ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ В ИНФОРМАЦИОННЫХ СЕТЯХ

1.1. Состояние проблемы. При моделировании систем массового обслуживания (СМО) большое внимание уделяется изучению дважды стохастических потоков, как правило, пуассоновских [1]. Достаточно хорошо изучены СМО с входными так называемыми МС-потоками [1–3]. В таких потоках интенсивность входного потока представляет собой марковскую цепь или полумарковский процесс с дискретным пространством состояний. В работе [4] исследуются дважды стохастические потоки, интенсивность которых является процессом с независимыми приращениями или гауссовским процессом. В работе [5] исследовались СМО с параметрами, изменяющимися во времени случайным образом, т. е. функционирующие в случайной среде. В настоящей работе такие СМО для краткости будем называть *дважды стохастическими* (ДС) по аналогии с дважды стохастическими потоками.

Как правило, рассматривались ДС СМО, параметры которых постоянны в течение некоторого случайного времени, а затем мгновенно изменяются [6]. Набор значений параметров конечен, а процесс их переключения есть либо марковский [7], либо полумарковский [8]. В [9] для системы с бесконечной очередью получено необходимое и достаточное условие эргодичности процесса, а для СМО с конечным накопителем получены выражения для некоторых показателей производительности СМО. Также изучались скачкообразные процессы и ДС СМО со скачкообразной интенсивностью входного потока (см. список литературы в [10]). В указанных работах интенсивность входного потока имела дискретное пространство состояний и расчет характеристик ДС СМО, как правило, сводился к решению матричных уравнений.

В отличие от указанных работ в последнее время большое внимание уделяется исследованию СМО с входным дважды стохастическим пуассоновским потоком заявок, пространство состояний интенсивности которого представляет собой интервал или совокупность интервалов. В работе авторов [11] исследуются ДС СМО с диффузионной интенсивностью входного потока и конечным

накопителем. В [12, 13] приводятся результаты исследования ДС СМО с диффузионной и скачкообразной интенсивностью входного потока.

Исследовались информационные потоки на канальном и транспортном сетевых уровнях семиуровневой модели сетевого обмена. Например, в работе [14] рассмотрено применение модели СМО М/М/1 в информационных сетях (ИС) на указанных уровнях. В данной работе исследуется применение моделей СМО на прикладном седьмом сетевом уровне. Работа имеет прикладной характер на границе трех направлений: информационные сети, теория вероятностей, теория массового обслуживания. В п. 1.2 работы описана предметная область, в которой обозначена прикладная проблема — разработка новых средств мониторинга информационных сетей. В последующих разделах работы строятся математические модели СМО и проверяются соответствующие статистические гипотезы. В разд. 3 работы показано применение предложенных новых средств мониторинга ИС.

1.2. Общие закономерности построения локальных сетей. Рассмотрим общие закономерности построения локальных сетей, важные с точки зрения анализа информационных потоков в ИС.

Независимо от топологии локальной сети (ЛС) — кольцевой, звезда или общая шина — пользователи ЛС в процессе информационного обмена общаются между собой в ЛС и с внешней глобальной сетью с помощью специализированных серверов. Чаще всего в качестве внешней глобальной сети выступает Интернет. Концентраторы и коммутаторы используются в таком информационном обмене в качестве вспомогательных инструментов. Рассматривая сложившуюся в российской практике классификацию локальных сетей по количеству рабочих станций: малых (100–300), средних (300–600) и больших размеров (более 600) — можно отметить общую закономерность в построении таких ЛС. Пользователи ЛС через коммутаторы этажей и корпусов, затем через корневой коммутатор подключаются к специализированным серверам. К ним относятся: сервер локальных ресурсов, библиотечные серверы, серверы баз данных, Прoxy-сервер, web-сервер, сервер мониторинга. На сервере локальных ресурсов размещается программное обеспечение, файлы общего использования ЛС. На библиотечных серверах, как правило, размещаются базы данных электронных библиотек, представляющие пользователям Интернет-услуги с применением web-технологии.

Прoxy-сервер получает заявки от пользователей ЛС на доступ к Интернет-ресурсам, выполняет запрос доступности данных ресурсов и, получив положительный ответ, реализует заявки. Информационный обмен между Прoxy-сервером и Интернет-ресурсами реализуется с помощью пакетной коммутации, при которой информация передается пакетами с использованием установленных протоколов связи в течение определенных сеансов. При этом каждая заявка порождает совокупность дополнительных сеансов, которые называют рожденными сеансами. Если пользователь ЛС обратился, например, к сайту, то каждый дополнительный сеанс служит обычно для передачи в виде отдельных файлов объектов, размещенных на сайте (картинок, отдельных фреймов, банеров и т. д.). На web-сервере размещаются информационные ресурсы для пользователей Интернет: сайты, базы данных, архивы. Пользователи ЛС также имеют доступ к информационным ресурсам web-сервера ЛС. Сервер мониторинга выполняет функции наблюдения и защиты информационных ресурсов ЛС от внешнего и внутреннего несанкционированного доступа.

В данной работе приводится анализ информационных потоков в информационной сети на примере ЛС Тихоокеанского государственного экономического университета (ТГЭУ, г. Владивосток). Аналогичные результаты получены для информационных сетей ряда других организаций. К серверам ЛС ТГЭУ подключено более 600 рабочих станций. На библиотечном сервере размещается автоматизированная библиотечная информационная система (АБИС) «Руслан», представляющая пользователям Интернет-услуги каталога и полных текстов

библиотеки ТГЭУ с применением web-технологии. Связь Проху-сервера, web-сервера и сервера мониторинга с Интернет осуществляется через концентратор внешней сети, многополосный быстродействующий модем и выделенный канал связи — асимметричную цифровую абонентскую линию (Asymmetric Digital Subscriber Line, ADSL). Максимальная пропускная способность рассматриваемого ADSL-модема составляет 10 Мбит/с, из которой администрацией ЛС заказана провайдеру Интернет-услуг выделенная пропускная способность ADSL-модема 512 Кбит/с.

Согласно стандартам «Семиуровневой модели сетевого обмена» (OSI Ref. Model) на седьмом прикладном уровне работа каждого сервера регистрируется специальными протоколами, в которых, как правило, отмечается для каждого сеанса внутренний адрес заявки ЛС (в виде IP-адреса и/или ЛС-идентификатора), дата и время заявки, web-адрес запрашиваемых ресурсов, код результата запроса выполнения заявки, объем передаваемой информации (в байтах) — длина пакета. Протоколы размещаются в специальных файлах, в результате статистического анализа которых можно показать характеристики работы сервера.

Из рассмотрения общих закономерностей построения ЛС можно сделать вывод, что каждый сервер можно рассматривать в качестве системы массового обслуживания. В такой СМО заявки размещаются в накопительном буфере, емкость которого N_0 зависит от типа сервера. Например, для Проху-сервера емкость накопителя составляет 350000 заявок. Для больших сетей такая емкость может считаться конечной, для малых — практически бесконечной. Совокупность заявок, поступающих на сервер, образует входной поток заявок СМО.

Закон распределения интервалов времени $\tau_{\lambda,i}$ между моментами поступления заявок характеризует тип входного потока заявок. Совокупность порожденных сеансов одного пользователя ЛС в принципе можно считать тоже одной общей (интегральной) заявкой. Закон распределения интервалов между моментами поступления интегральных заявок характеризует тип входного потока интегральных заявок.

Обслуживание каждой заявки заключается в передаче файла определенного объема V_i со скоростью v , равной максимальной пропускной способности сервера, например, для Проху-сервера выделенная пропускная способность равна $v = 512$ Кбит/с. Таким образом, совокупность случайных величин $\eta_i = V_i/v$ образует совокупность длительностей обслуживания, закон распределения которых характеризует тип обслуживания.

В рассмотренных серверах заявки обслуживаются в порядке поступления, что и будет рассматриваться в дальнейшем в данной работе, хотя в некоторых случаях в сети возможно установление приоритетов обслуживания заявок.

Таким образом, тип входного потока и закон обслуживания сервера можно выяснить с помощью статистического анализа законов распределения величин V_i и η_i в протоколах сервера.

2. КЛАССИФИКАЦИЯ ПРИКЛАДНЫХ МОДЕЛЕЙ СМО В ИНФОРМАЦИОННЫХ СЕТЯХ

Рассмотрим статистический анализ входного потока и обслуживания на примере статистического анализа протокола Проху-сервера. Для поступивших заявок проводились наблюдения в моменты времени $9^{30}, 10^{30}, \dots, 18^{30}, 19^{30}$ с интервалом в один час, которые для удобства будем называть в дальнейшем контрольными моментами наблюдения или для краткости — контрольными моментами.

2.1. Статистический анализ распределения интервалов времени обслуживания. В контрольные моменты наблюдались выборки длин пакетов V_1, \dots, V_n объема n , $n = 450$. На основе данной выборки строился вариацион-

ный ряд

$$\begin{array}{c|cccc} V_i & V_1 & V_2 & \dots & V_k \\ \hline n_i & n_1 & n_2 & \dots & n_k \end{array} \quad (1)$$

объема $n = n_1 + \dots + n_k$, $V_1 < V_2 < \dots < V_k$, с выборочными частотами n_1, \dots, n_k . Выборка (1) использовалась для построения выборки времен обслуживания $\eta_i = V_i/v$:

$$\begin{array}{c|cccc} \eta_i & \eta_1 & \eta_2 & \dots & \eta_k \\ \hline n_i & n_1 & n_2 & \dots & n_k \end{array}. \quad (2)$$

На основе выборки (2) строилась эмпирическая функция распределения $F_e(t)$ согласно определению по формуле

$$F_e(t) = \sum_i \frac{n_i}{n}, \quad (3)$$

где суммирование проводится по всем значениям индекса i , для которых выполняется условие $\eta_i < t$. Для краткости формула (3) записывается в виде

$$F_e(t) = \sum_{\eta_i < t} \frac{n_i}{n}.$$

На рис. 1 показаны графики эмпирической функции распределения $F_e(t)$ в контрольный момент времени 10^{30} . Аналогичный вид имеют графики эмпирической функции распределения в остальных контрольных точках наблюдения. По графикам виден экспоненциальный характер распределения времен обслуживания. Поэтому в качестве теоретического уравнения регрессии принята функция экспоненциального распределения

$$F_t(t) = 1 - e^{-\mu t},$$

где μ — параметр экспоненциального распределения. Функцию $F_t(t)$ будем называть *теоретической функцией экспоненциального распределения*. На рис. 1 график эмпирической функции показан слева и справа, тогда как график теоретической функции распределения показан только слева. Такое построение графиков выполнено в силу достаточной близости данных функций, из-за которой графики функций сливаются.

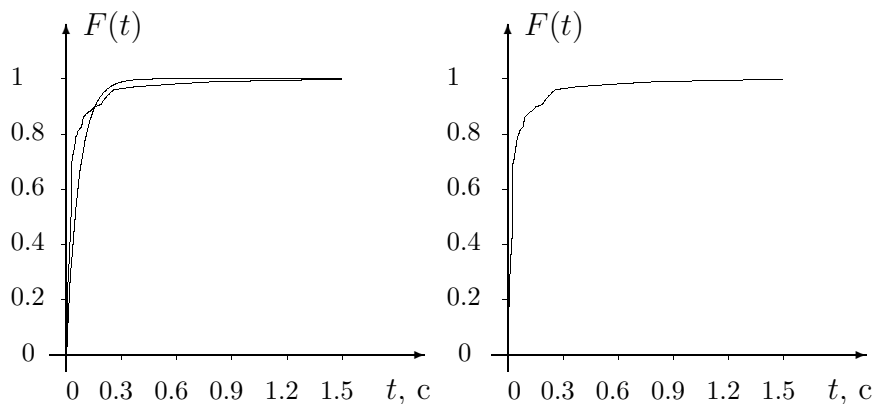


Рис. 1. Распределение интервалов времени обслуживания

Параметр μ находился двумя различными способами: методом наименьших квадратов (для экспоненциальной регрессии) и с применением основного свойства экспоненциального распределения. Метод наименьших квадратов хорошо известен в литературе, поэтому здесь не описывается. По основному свойству экспоненциального распределения для экспоненциально распределенной случайной величины η с параметром μ выполняется

$$\mu = 1/\mathbf{M}\eta. \quad (4)$$

Согласно (4) оценка параметра μ найдена по формуле

$$\mu = 1/\bar{\eta}, \quad (5)$$

где среднее время обслуживания $\bar{\eta}$ по определению равно

$$\bar{\eta} = \sum_{i=1}^k \eta_i n_i / n.$$

Наименьшая погрешность получена при значении параметра μ , найденном вторым методом по формуле (5).

Введем обозначения плотностей распределения: $f_e(t)$ — эмпирической, $f_t(t)$ — теоретической. Для вычисления $f_e(t)$ задавалось число интервалов $s = 50$ и интервалы группирования строились равномерным разбиением интервала выборочных значений на s подынтервалов шириной Δ . На каждом i -м интервале находились интервальные частоты v_i и плотность $f_e(t)$ приравнивалась отношению $v_i/(n\Delta)$, где n — объем выборки. Рассматривались погрешности отклонения эмпирического и теоретического распределений, т. е. отклонения эмпирической $F_e(t)$ от теоретической $F_t(t)$ функции распределения и отклонения эмпирической $f_e(t)$ от теоретической $f_t(t)$ плотности распределения.

Первая погрешность — это относительная погрешность отклонения эмпирической $F_e(t)$ от теоретической $F_t(t)$ функции распределения. Эта погрешность ϵ_{L_1} рассматривалась по норме пространства L_1 интегрируемых функций

$$\epsilon_{L_1} = \frac{\|F_e(t) - F_t(t)\|_{L_1}}{\|F_t(t)\|_{L_1}}, \quad (6)$$

где

$$\|F_e(t) - F_t(t)\|_{L_1} = \int_0^r |F_e(t) - F_t(t)| dt, \quad \|F_t(t)\|_{L_1} = \int_0^r |F_t(t)| dt,$$

значение r задавалось из условия поточечного отклонения $F_e(t)$, $F_t(t)$ не более чем на малое значение $\varepsilon = 10^{-6}$ для $t > r$. Для рис. 1 значения r равны максимальным значениям абсцисс. По сути, относительная погрешность ϵ_{L_1} — это относительное отклонение площадей под функциями распределениями на интервале $(0, r)$. Как показали расчеты, погрешность ϵ_{L_1} оказалась весьма незначительной. Для рис. 1 $\epsilon_{L_1} = 0,14\%$. Среднее значение ϵ_{L_1} по всем контрольным точкам равно $\mathbf{M}\epsilon_{L_1} = 0,14\%$.

Однако погрешность ϵ_{L_1} не учитывает особенностей поведения кривых плотностей распределения $f_e(t)$, $f_t(t)$. Поэтому рассматривалась также относительная погрешность отклонения эмпирической $f_e(t)$ от теоретической $f_t(t)$ плотности распределения. Эта погрешность ϵ_{L_2} рассматривалась по норме пространства L_2 интегрируемых с квадратом функций

$$\epsilon_{L_2} = \frac{\|f_e(t) - f_t(t)\|_{L_2}}{\|f_t(t)\|_{L_2}}, \quad (7)$$

где

$$\|f_e(t) - f_t(t)\|_{L_2} = \int_0^{\infty} (f_e(t) - f_t(t))^2 dt, \quad \|f_t(t)\|_{L_2} = \int_0^{\infty} f_t^2(t) dt.$$

Согласно расчетам для рис. 1 погрешность $\epsilon_{L_2} = 64\%$. Среднее значение ϵ_{L_2} по всем контрольным точкам равно $M\epsilon_{L_2} = 66\%$. Для рис. 2 вычислены погрешности ϵ_{L_1} , ϵ_{L_2} во всех контрольных точках, которые характеризуют достаточно хорошую близость плотностей распределения $f_e(t)$, $f_t(t)$.

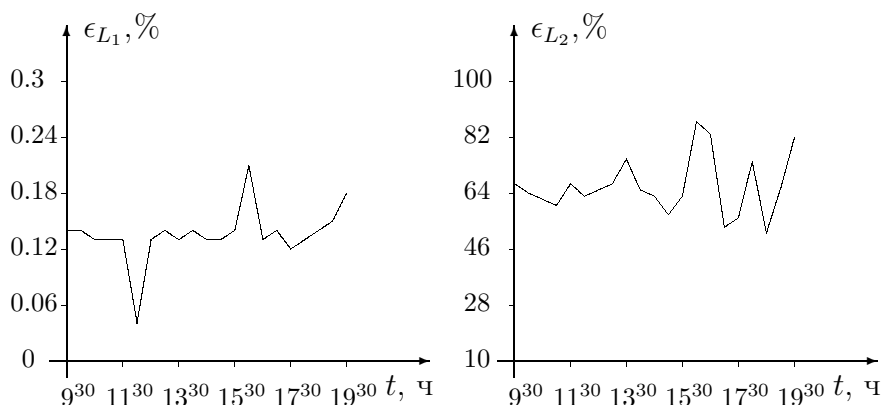


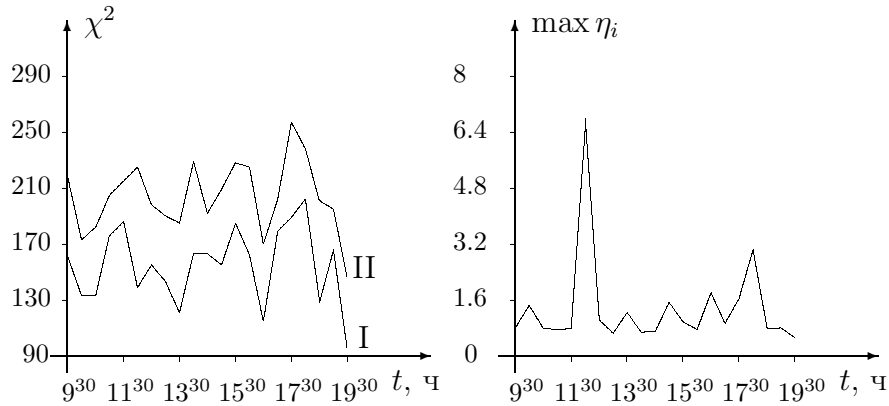
Рис. 2. Проверка гипотезы о законе обслуживания

Кроме того, во всех контрольных точках проверялась гипотеза об экспоненциальном законе распределения времени обслуживания с применением критерия Пирсона [15] при уровне значимости $\gamma = 0,05$. В процедуре использования критерия Пирсона задавалось число интервалов $s = 50$ и области группирования строились равномерным разбиением интервала выборочных значений на s подынтервалов. Количество степеней свободы критерия бралось равным $k = s - 2$. В частности, согласно расчетам для рис. 1 наблюдаемое значение критерия $\chi_{\text{набл}}^2$ равно 133; критическая точка $\chi_{\text{крит}}^2$ равна 173. Согласно расчетам во всех контрольных точках наблюдаемое значение критерия меньше критической точки. Это означает, что результаты проверки гипотезы не противоречат гипотезе об экспоненциальном характере распределения времени обслуживания.

На рис. 3 справа показаны максимальные интервалы $\max \eta_i$ обслуживания во всех контрольных точках, где среднее время обслуживания одного порожденного сеанса $M(\max \eta_i) = 1,35$ с, а слева два графика, характеризующих изменение во времени значений критериев: нижний график $\chi_{\text{набл}}^2$ помечен меткой I, верхний график $\chi_{\text{крит}}^2$ помечен меткой II.

Проведенное сравнение эмпирического и теоретического распределений показывает, что время обслуживания заявок на Ргоху-сервере имеет экспоненциальный закон распределения. Аналогичные результаты получены для других дней наблюдения и остальных серверов.

2.2. Статистический анализ функции распределения входящего потока. Статистический анализ функции распределения входящего потока проводился аналогично статистическому анализу распределения интервалов времени обслуживания. Сначала осуществлялся статистический анализ потока согласно временному описанию потока. В контрольные моменты наблюдения брались выборки интервалов времени между соседними моментами появления заявок τ_1, \dots, τ_n объема $n = 450$. Так как протоколы работы серверов регистрируют время прихода заявок с точностью до секунд, то если в какую-то

Рис. 3. Погрешности ϵ_{L_1} и ϵ_{L_2}

секунду появилось несколько заявок, тогда при проведении статистического анализа секунда делилась на интервалы времени между соседними моментами появления заявок пропорционально объемам пакетов заявок. Это объясняется тем, что перед появлением заявки создаются вспомогательные промежуточные сеансы связи и чем больше объем заявки, тем больше времени уходит на предварительную обработку информации. На основе выборки τ_1, \dots, τ_n строился вариационный ряд

$$\begin{array}{c|cccc} \tau_i & \tau_1 & \tau_2 & \dots & \tau_k \\ \hline n_i & n_1 & n_2 & \dots & n_k \end{array} \quad (8)$$

объема $n = n_1 + \dots + n_k$, $\tau_1 < \tau_2 < \dots < \tau_k$, с выборочными частотами n_1, \dots, n_k . На основе выборки (8) строилась эмпирическая функция распределения $F_e(t)$ согласно определению по формуле

$$F_e(t) = \sum_{\tau_i < t} \frac{n_i}{n}.$$

Анализ графиков эмпирической функций распределения $F_e(t)$, имеющих аналогичный графикам рис. 1 вид, в контрольные моменты времени показал, что они имеют экспоненциальный характер распределения интервалов времени между соседними моментами появления заявок. Поэтому в качестве теоретического уравнения регрессии принята функция экспоненциального распределения

$$F_t(t) = 1 - e^{-\lambda t},$$

где λ — параметр экспоненциального распределения. Параметр λ находился двумя различными способами: методом наименьших квадратов (для экспоненциальной регрессии) и с применением основного свойства экспоненциального распределения. По основному свойству экспоненциального распределения для экспоненциально распределенной случайной величины τ с параметром λ выполняется

$$\lambda = 1/(M\tau). \quad (9)$$

Согласно (9) оценка параметра λ находилась по формуле

$$\lambda = 1/\bar{\tau}, \quad (10)$$

где среднее значение интервалов времени между соседними моментами появления заявок $\bar{\tau}$ по определению равно

$$\bar{\tau} = \sum_{i=1}^k \tau_i n_i / n.$$

Наименьшая погрешность получена при значении параметра λ , найденном вторым методом по формуле (10). В качестве критериев погрешности отклонения эмпирической $F_e(t)$ от теоретической $F_t(t)$ функции распределения использовались определенные ранее погрешности ϵ_{L_1} , ϵ_{L_2} . Первая погрешность — это относительная погрешность ϵ_{L_1} по норме пространства L_1 интегрируемых функций, задаваемая формулой (6). Например, в расчетах в контрольные моменты 9^{30} , 18^{00} значения погрешностей равны $\epsilon_{L_1} = 0,27\%$; $\epsilon_{L_1} = 0,29\%$ соответственно. Среднее значение ϵ_{L_1} по всем контрольным точкам равно $M\epsilon_{L_1} = 0,32\%$.

Однако, как ранее указывалось, погрешность ϵ_{L_1} не учитывает особенностей поведения соответствующих плотностей распределения $f_e(t)$, $f_t(t)$. Поэтому рассматривалась также относительная погрешность ϵ_{L_2} по норме пространства L_2 интегрируемых с квадратом функций, задаваемая формулой (7). Например, в расчетах в контрольные моменты 9^{30} , 18^{00} значения погрешностей равны $\epsilon_{L_2} = 55\%$, $\epsilon_{L_2} = 53\%$ соответственно. Среднее значение ϵ_{L_2} по всем контрольным точкам равно $M\epsilon_{L_2} = 57\%$. Расчеты показали погрешности ϵ_{L_1} , ϵ_{L_2} во всех контрольных точках, которые характеризуют достаточно хорошую близость плотностей распределения $f_e(t)$, $f_t(t)$.

Кроме того, во всех контрольных точках проверялась гипотеза об экспоненциальном законе распределения интервалов времени между соседними моментами появления заявок с применением критерия Пирсона при уровне значимости $\gamma = 0,05$. В процедуре использования критерия Пирсона задавалось число интервалов $s = 50$ и области группирования строились равномерным разбиением интервала выборочных значений на s подынтервалов. В частности, согласно расчетам в контрольный момент 9^{30} наблюдаемое значение критерия $\chi_{\text{набл}}^2$ равно 228; критическая точка $\chi_{\text{крит}}^2$ равна 276. Согласно расчетам во всех контрольных точках наблюдаемое значение критерия меньше критической точки. Это означает, что результаты проверки гипотезы не противоречат гипотезе об экспоненциальном характере распределения интервалов времени между соседними моментами появления заявок.

Проведенное сравнение эмпирического и теоретического распределений показывает, что интервалы времени между соседними моментами появления заявок на Proху-сервере имеют экспоненциальный закон распределения. Аналогичные результаты получены для других дней наблюдения и остальных серверов.

Таким образом, выборочные данные не противоречат гипотезе о том, что входной поток имеет экспоненциальное распределение с постоянной интенсивностью λ^* . Однако в более точной модели интенсивность входного потока может иметь переменный характер. В этом случае λ^* равно усредненному значению выборочных значений переменной интенсивности, как показано выше. Кроме того, показано, что погрешности эмпирической и теоретической функций распределения для интенсивности входного потока больше, чем соответствующие погрешности распределения для времени обслуживания. Более подробное исследование интенсивности входного потока приведено в следующем пункте.

2.3. Статистический анализ интенсивности потока. Ранее были показаны результаты статистического анализа потока согласно временному описанию потока. Покажем теперь результаты статистического анализа потока согласно его количественному описанию. При количественном описании потока временной интервал наблюдения \hat{T} разбивается на детерминированные интервалы z_1, \dots, z_n . Пусть в результате наблюдения в этих детерминированных интервалах появилось количество заявок ν_1, \dots, ν_n . При таком подходе поток задается совместным распределением случайных величин ν_1, \dots, ν_n для любого $n \in \mathbb{N}$. Если в качестве детерминированных интервалов z_1, \dots, z_n взять малые единичные интервалы, то величины $M\nu_1, \dots, M\nu_n$ будут представлять

собой последовательность значений интенсивности входного потока $\lambda(t)$. В работе вычисление величин $M\nu_1, \dots, M\nu_n$ проводилось усреднением количества появившихся заявок на подынтервалах интервалов z_1, \dots, z_n . Такой подход использовался при статистическом анализе интенсивности входного потока заявок порожденных сеансов $\lambda(t)$, который проводился в указанные выше контрольные моменты времени.

2.4. Статистический анализ потока заявок порожденных сеансов. Рассмотрим статистический анализ потока порожденных сеансов. На рис. 4 показаны графики интенсивности входного потока заявок порожденных сеансов $\lambda(t)$ в контрольные моменты времени 9^{30} (слева), 10^{00} (справа). Внешний вид интенсивности $\lambda(t)$ показывает, что интенсивность представляет собой диффузионный процесс, на рис. 4 — с упругими границами. В другие контрольные моменты времени наблюдались также поглощающие границы. Согласно определению диффузионным процессом называется марковский случайный процесс второго порядка с независимыми приращениями, диффузионные операторные моменты первого и второго порядков которого равны коэффициенту сноса $a(t, x)$ и коэффициенту диффузии $b(t, x)$ соответственно (см. [16, 17]). Коэффициент сноса a диффузионного процесса $\lambda(t)$ по определению равен

$$a(t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|y-x|<\varepsilon} (y-x)f(t, x; t+\Delta t, y) dy$$

для произвольной сколь угодно малой величины ε .

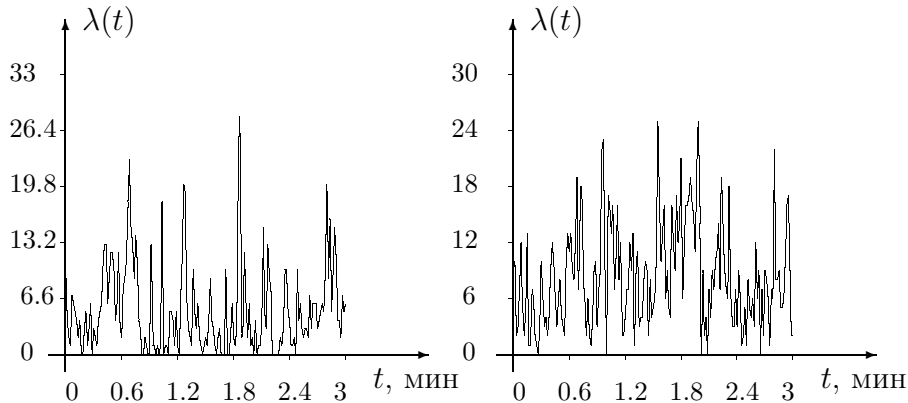


Рис. 4. Интенсивность входного потока заявок порожденных сеансов

Коэффициент диффузии $b(t, x)$ диффузионного процесса $\lambda(t)$ по определению равен

$$b(t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|y-x|<\varepsilon} (y-x)^2 f(t, x; t+\Delta t, y) dy \neq 0, \quad (11)$$

где $f(t, x; s, y)$ — переходная плотность распределения диффузионного процесса $\lambda(t)$:

$$f(t, x; s, y) dy = \mathbf{P}\{y < \lambda(s) < y + dy \mid \lambda(t) = x\}.$$

Диффузионные операторные моменты k -го порядка равны нулю для всех $k > 2$:

$$c_k(t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|y-x|<\varepsilon} (y-x)^k f(t, x; t+\Delta t, y) dy = 0, \quad k > 2. \quad (12)$$

Условия (11), (12) означают, что диффузионный процесс имеет второй порядок [14].

Расчеты показали малость приращений диффузионного процесса $\lambda(t)$. Проверка гипотез о независимости каждой пары приращений X, Y диффузионного процесса $\lambda(t)$ выполнена с применением критерия, построенного на основе корреляционного отношения для парной выборки наблюдений X, Y и описана в приложении.

Статистический анализ показал, что в некоторые контрольные моменты времени коэффициенты сноса и диффузии не зависят от t, x .

Для проверки независимости коэффициента сноса a от t, x последовательно проверялись две гипотезы: независимость коэффициента сноса a от x и независимость коэффициента сноса a от t . Для проверки гипотезы независимости a от t для каждого фиксированного x_i рассматривались всевозможные пары величин (a, t) , а для проверки гипотезы независимости a от x для каждого фиксированного t_i рассматривались всевозможные пары величин (a, x) и с использованием техники, изложенной в приложении, проверялась независимость указанных величин с применением критерия Фишера — Снедекора при уровне значимости $\gamma = 0,05$. Для проверки независимости коэффициента диффузии b от x и t применялась та же методика. В расчетах коэффициенты сноса a и диффузии b находились приближенно согласно приведенным определениям с учетом выражений под знаками пределов.

Усредненные оценки постоянных коэффициентов сноса a и b находились следующим образом. При фиксированном x по определению коэффициент сноса зависит от t и представляет собой для конкретной реализации скорость изменения сглаженного тренда $S(t)$ диффузионного процесса λ , другими словами, график коэффициента сноса — это график производной $S'(t)$ сглаженного тренда диффузионного процесса $S(t)$. При фиксированном x для графика $z = S(t)$ по методу наименьших квадратов можно найти уравнение линейной регрессии z от t , аппроксимирующей график $S(t)$, вида $z = K_x t + d$. Согласно определению при фиксированном x линейная оценка коэффициента сноса равна K_x . Затем для нахождения оценки постоянного выборочного коэффициента сноса a проводилось усреднение полученных значений K_x по x .

Выборочная оценка постоянного коэффициента диффузии b находилась следующим образом. Наблюдалась выборка значений процесса λ

$$\begin{array}{c|cccc} L_i & L_0 & L_1 & \dots & L_k \\ \hline m_i & m_0 & m_1 & \dots & m_k \end{array}$$

объема $m = m_0 + \dots + m_k$, $L_0 < L_1 < \dots < L_k$, с выборочными частотами m_0, \dots, m_k . Для вычисления выборочной усредненной оценки постоянного коэффициента диффузии b согласно определению использовалась формула

$$b = \sum_{i=1}^m \frac{(L_i - L_{i-1})^2}{\Delta t} h_i \Delta x / m, \quad h_i = m_i / m, \quad \Delta x = (\beta - \alpha) / m,$$

где $[\alpha, \beta]$ — границы изменения диффузионного процесса λ , Δt — малый единичный отрезок времени наблюдения. В статистическом анализе использовалось значение $\Delta t = 1$ с.

На рис. 4 для левого графика $[\alpha, \beta] = [0, 28]$, $a = -0,09$, $b = 2,44$, а для правого графика $[\alpha, \beta] = [0, 25]$, $a = 0,36$, $b = 2,64$.

Статистический анализ потока входных заявок показал также наличие диффузионной интенсивности входного потока на других серверах, как правило, на таких, где заявки во входном потоке появляются достаточно часто, например на web-сервере.

2.5. Статистический анализ потока заявок интегральных сеансов. Рассмотрим статистический анализ потока интегральных сеансов Ргоху-сервера. На рис. 5 показаны графики интенсивности входного потока заявок порожденных сеансов $\lambda(t)$ в контрольные моменты времени 10^{30} (слева), 11^{00} (справа). Внешний вид интенсивности $\lambda(t)$ показывает, что интенсивность представляет собой скачкообразный процесс. Как показал статистический анализ, аналогичный приведенному выше, интервалы постоянства T представляют собой экспоненциальные величины с параметром θ . Кроме того, с применением статистического анализа показано, что значения процесса $\lambda(t)$ в точках разрыва слева и справа представляют собой независимые числа, причем значения справа имеют равномерное распределение на конечном интервале $[a, b]$.

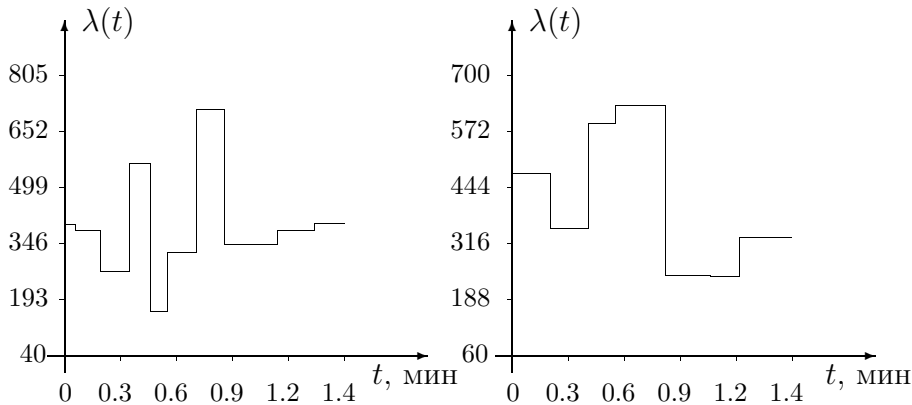


Рис. 5. Интенсивность входного потока интегральных заявок

Статистический анализ потока входных заявок показал также наличие скачкообразной интенсивности на других серверах, как правило, на таких, где заявки во входном потоке появляются достаточно часто, например на библиотечном сервере и сервере баз данных.

3. РЕКОМЕНДАЦИИ К ПРИМЕНЕНИЮ МОДЕЛЕЙ СМО

Создание и эксплуатация ИС предполагает последовательное выполнение ряда задач: определение круга пользователей, проектирование и построение топологии информационных сетей, выбор типа операционной системы, выбор и приобретение соответствующего оборудования, выбор подрядчика для реализации информационной сети, реализация сети, администрирование ИС и т. д.

Основные проблемы администрирования сетей с использованием web-технологий отражены в [18]. В рамках данной работы основное внимание уделяется вопросам разработки теоретических положений и научно обоснованных технических решений, обеспечивающих эффективность и повышение качества администрирования информационных сетей на основе исследования информационных потоков в информационных сетях.

Мониторинг информационных потоков ИС и прогнозирование их состояния с применением моделей СМО позволяет выбирать соответствующие параметры оборудования, программного обеспечения, настроек ИС в процессе проектирования, а затем эксплуатации сетей. Таким образом, моделирование информационных сетей является актуальной технической и научной проблемой на стадии проектирования, а затем эксплуатации ИС.

В настоящее время наряду с модернизацией оборудования ИС актуальной задачей является необходимость оптимизации затрат, выделяемых на поддержку ресурсов ИС. Одним из факторов оптимизации является применение моделей информационных потоков ИС, в том числе моделей СМО ИС.

Одной из важных задач администрирования локальных сетей является предоставление пользователям ресурсов Интернет. Это обстоятельство существенным образом влияет на администрирование ИС с применением Интернет-технологий.

Запросы внешних пользователей из Интернет к ресурсам ИС серверов — библиотечного, Проху-сервера, Web-сервера, а также запросы внутренних пользователей к ресурсам Интернет создают общую нагрузку на выделенный канал связи в Интернет, который обладает заданной пропускной способностью, в настоящее время равной K Кбит/с. Пропускная способность K представляет собой максимальное количество килобит в секунду, которое может пропустить выделенный канал, и определяется по договору администрации ИС с провайдером, обеспечивающим доступ в Интернет. Ежемесячная оплата S_{net} администрации ИС услуг провайдера зависит от объема трафика информации V_{inf} , проходящей через выделенный канал связи в Интернет, и стоимости единицы трафика $C(K)$, которая является функцией от пропускной способности K :

$$S_{\text{net}} = V_{\text{inf}}C(K).$$

Общий объем трафика информации V_{inf} равен сумме объемов информации

$$V_{\text{inf}} = \sum_{j=1}^{m_{\text{ser}}} V_{\text{inf},j},$$

где $V_{\text{inf},j}$ — объем информации, проходящей через j -й сервер, $1 \leq j \leq m_{\text{ser}}$, m_{ser} — общее количество серверов. Объем информации $V_{\text{inf},j}$, проходящей через j -й сервер, равен

$$V_{\text{inf},j} = \mathbf{M}V_{i,j}\mathbf{M}\nu_j,$$

где $V_{i,j}$ — объем в байтах i -го сообщения на j -м сервере, ν_j — число сообщений, проходящих через j -й сервер. Таким образом, построена функция зависимости S_{net} от K :

$$S_{\text{net}} \equiv S(K) = C(K) \sum_{j=1}^{m_{\text{ser}}} \mathbf{M}V_{i,j}\mathbf{M}\nu_j(K). \quad (13)$$

Величины $V_{i,j}$ наблюдаются в результате мониторинга ИС, величина $C(K)$ определяется провайдером сети, методы и примеры вычисления среднего числа сообщений $\mathbf{M}\nu_j$ приводятся в данной работе. Отметим, что среднее число сообщений $\mathbf{M}\nu_j$ является функцией от K , так как $\mathbf{M}\nu_j$ зависит от интенсивности обслуживания μ , а интенсивность обслуживания μ зависит от K : $\mu = \mathbf{M}V_{i,j}/K$.

Завершая данные рассуждения, заключаем, что если заказана высокая пропускная способность K , а сеть простаивает, т. е. велика вероятность простоя сети p_0 , то администрация ИС несет неоправданно высокие затраты S_{net} и, следовательно, затраты на ресурсы ИС необходимо снижать. Если заказана невысокая пропускная способность K , а сеть перегружена, т. е. время ожидания пользователями начала обслуживания U велико (обслуживание пользователей тормозится), то пользователи ИС несут неоправданно высокие потери при низких затратах на ресурсы ИС S_{net} и, следовательно, затраты на ресурсы ИС необходимо повышать.

Таким образом, мониторинг ИС и прогноз состояния ИС является важным инструментом регуляции многокритериальной оптимизации затрат по показателям: расходы на ресурсы ИС S_{net} и потери пользователей ИС по времени ожидания обслуживания U . Функционал расходов на ресурсы ИС S_{net} задан в формуле (13), методы вычисления среднего и дисперсии числа сообщений на серверах и времени ожидания обслуживания U для моделей СМО при скачкообразной и диффузионной интенсивности входного потока приведены в монографиях [12, 13].

ПРИЛОЖЕНИЕ

Проверка независимости приращений диффузионного процесса $\lambda(t)$ выполнена с помощью следующего статистического анализа. Вводилось понятие наблюдения, т. е. набора последовательных замеров $\lambda(t)$ в моменты времени $t_1, \dots, t_{n_\lambda}$ с интервалом 1 с, начиная с контрольного момента времени t_1 , где n_λ — количество замеров. В статистическом анализе использовалось значение $n_\lambda = 181$. Вводилось обозначение для приращений процесса $\Delta\lambda(t_i) = \lambda(t_{i+1}) - \lambda(t_i)$. Для проверки независимости случайных величин $X = \Delta\lambda(t_{i_0})$ и $Y = \Delta\lambda(t_{j_0})$ для всех моментов времени $1 \leq t_{i_0}, t_{j_0} \leq n_\lambda$, $t_{i_0} \neq t_{j_0}$, проводилась серия $m_\lambda = 100$ наблюдений. В результате наблюдений над случайными величинами X, Y получалась выборка $(x_1, y_1), \dots, (x_{m_\lambda}, y_{m_\lambda})$ объема m_λ . Затем диапазон наблюдаемых значений x_i разбивался на k подынтервалов с серединами u_1, \dots, u_k . Диапазон наблюдаемых значений y_j разбивался на $n = k$ подынтервалов с серединами v_1, \dots, v_n . Строилась корреляционная матрица интервальных частот u_i, v_j , в которой отмечались различные значения: $u_1, \dots, u_k, v_1, \dots, v_n$ и частоты $n_{u_i v_j}$ пары значений u_i, v_j . Частота $n_{u_i v_j}$ равна количеству парных значений x, y таких, что x принадлежит i -му подынтервалу по u , а y принадлежит j -му подынтервалу по v . На основе корреляционной таблицы строилось корреляционное отношение

$$\eta = \frac{\sum_{i=1}^k n_{u_i} (\bar{v}_{u_i} - \bar{v})^2}{\sum_{j=1}^n n_{v_j} (v_j - \bar{v})^2},$$

где \bar{v}_{u_i} — соответствующие условные средние. Независимость значений X, Y соответствует малым значениям корреляционного отношения η . Поэтому на основе корреляционного отношения η строился критерий F проверки независимости X, Y следующим образом. Вводились обозначения $\tilde{\xi}_i = \bar{v}_{u_i} - \bar{v}$, $\tilde{\eta}_j = v_j - \bar{v}$, $\xi_i = \tilde{\xi}_i / \sigma(\tilde{\xi}_i)$, $\eta_j = \tilde{\eta}_j / \sigma(\tilde{\eta}_j)$, где $\sigma(\tilde{\xi}_i)$, $\sigma(\tilde{\eta}_j)$ — среднеквадратические отклонения выборочных значений $\tilde{\xi}_i, \tilde{\eta}_j$, полученных в результате наблюдений.

Среднеквадратические отклонения выборочных значений $\tilde{\xi}_i, \tilde{\eta}_j$ вычислялись следующим образом. Наблюдение над случайными величинами $X = \Delta\lambda(t_{i_0})$ и $Y = \Delta\lambda(t_{j_0})$ для фиксированных моментов времени $t_{i_0} \neq t_{j_0}$ называлось испытанием. Количество испытаний равно числу сочетаний из n_λ по 2: $n_0 = C_{n_\lambda}^2$. В каждом r -м испытании ($1 \leq r \leq n_0$) наблюдались величины $\tilde{\xi}_i^r, \tilde{\eta}_j^r$, где индекс r обозначает номер испытания. Из усреднения и вычисления дисперсии по верхнему индексу r следуют выборочные моменты для величин $\tilde{\xi}_i^r, \tilde{\eta}_j^r$ при фиксированных i, j , $1 \leq i \leq k, 1 \leq j \leq n$:

$$\begin{aligned} \mathbf{M}_{\text{выб}} \tilde{\xi}_i^l &= \sum_{r=1}^{n_0} \tilde{\xi}_i^r / n_0, & \mathbf{M}_{\text{выб}} \tilde{\eta}_j^l &= \sum_{r=1}^{n_0} \tilde{\eta}_j^r / n_0, \\ \mathbf{D}_{\text{выб}} \tilde{\xi}_i^l &= \sum_{r=1}^{n_0} (\tilde{\xi}_i^r - \mathbf{M}_{\text{выб}} \tilde{\xi}_i^l)^2 / n_0, & \mathbf{D}_{\text{выб}} \tilde{\eta}_j^l &= \sum_{r=1}^{n_0} (\tilde{\eta}_j^r - \mathbf{M}_{\text{выб}} \tilde{\eta}_j^l)^2 / n_0. \end{aligned}$$

Следует отметить, что разбиение диапазона значений x_i, y_j на одинаковое значение подынтервалов $k = n$ позволяет вычислить указанные моменты. В дальнейшем во избежание громоздкости обозначений верхний индекс r подразумевается для соответствующих величин, но не указывается. В рамках

введенных обозначений корреляционное отношение равно

$$\eta = \frac{\sum_{i=1}^k n_{u_i} \xi_i^2}{\sum_{j=1}^n n_{v_j} \eta_j^2}.$$

В случае независимости X, Y малым значениям $\tilde{\xi}_i, \tilde{\eta}_j$ соответствуют малые значения отнормированных величин ξ_i, η_j . Поэтому в наблюдениях малым значениям корреляционного отношения η соответствуют малые значения величины

$$F = \frac{\sum_{i=1}^k n_{u_i} \xi_i^2 / k_1}{\sum_{j=1}^n n_{v_j} \eta_j^2 / k_2}.$$

Очевидно, что $M\tilde{\xi}_i = 0, M\tilde{\eta}_j = 0$. С применением критерия Пирсона показано, что случайные величины $\tilde{\xi}_i, \tilde{\eta}_j$ распределены по нормальному закону распределения. Следовательно, случайные величины ξ_i, η_j распределены по нормальному закону распределения с параметрами $(0, 1)$. По построению случайная величина F имеет распределение Фишера — Снедекора с количеством степеней свободы $k_1 = k, k_2 = n$. С применением критерия Фишера — Снедекора F проверялись, и подтверждены гипотезы о независимости всех приращений диффузионного процесса X, Y с уровнем значимости $\gamma = 0, 05$.

ЛИТЕРАТУРА

1. Башарин Г. П., Кокотушкин В. А., Наумов В. А. Метод эквивалентных замен в теории трафика. Т. 2. М.: Электросвязь, 1980. С. 82–122.
2. Горцев А. М., Назаров А. А., Терпугов А. Ф. Управление и адаптация в системах массового обслуживания. Томск: Изд-во Томского ун-та, 1978.
3. Дудин А. Н., Клименок В. И. Расчет характеристик однолинейной системы обслуживания, функционирующей в синхронной случайной среде // Автоматика и телемеханика. 1997. № 1. С. 74–84.
4. Alvarez-Andrade S. Strong approximation of doubly stochastic Poisson processes // C. R. Acad. Paris Ser. I Math. 1993. V. 316, N 8. P. 869–872.
5. Таташев А. Г. Система массового обслуживания с переменной интенсивностью входного потока // Автоматика и телемеханика. 1995. № 12. С. 78–84.
6. Afanas'ev L. G., Kibkalo A. A. Uniform bounds for the periodic queue in the M(t)/G/1/system // Soviet Math. 1988. V. 40, № 4. P. 454–457.
7. Портенко Н. И., Скороход А. В., Шуренков В. М. Марковские процессы // Современные проблемы математики. Фундаментальные направления. Т. 46. М.: ВИНТИ, 1989. С. 5–245. (Итоги науки и техники).
8. Анисимов В. В., Алиев А. О. Предельные теоремы для рекуррентных процессов полумарковского типа // Теория вероятностей и мат. статистика. 1989. Т. 41. С. 9–15.
9. Rolski T. Approximation of periodic queues // Adv. Appl. Probab. 1987. V. 17, N 3. P. 691–707.
10. Клейнрок Л. Теория массового обслуживания. М.: Машиностроение, 1979.
11. Головкин Н. И., Катрахов В. В., Писаренко Т. А. Краевые задачи в стационарных системах массового обслуживания с диффузионной интенсивностью входного потока // Дифференц. уравнения. 2002. Т. 38, № 3. С. 305–312.
12. Головкин Н. И., Катрахов В. В. Анализ систем массового обслуживания, функционирующих в случайной среде. Владивосток: изд. ДВГАЭУ, 2000.

13. Катрахов В. В., Головки Н. И., Рыжков Д. Е. Введение в теорию марковских дважды стохастических систем массового обслуживания. Владивосток: изд. ДВГУ, 2005.
14. Клейнрок Л. Коммуникационные сети. Стохастические потоки и задержки сообщений. М.: Наука, 1970.
15. Гмурман В. Е. Теория вероятностей и математическая статистика. М.: Высш. школа, 2001.
16. Гнеденко Б. В. Курс теории вероятностей. М.: Физматгиз, 1961.
17. Баруча-Рид А. Т. Элементы теории марковских процессов и их приложения. М.: Наука, 1969.
18. Камер Д. Сети TCP/IP. Т. 1. Принципы, протоколы и структура. М.: Вильямс, 2003.

*Головки Николай Иванович
Каретник Владимир Олегович
Танин Виктор Евгеньевич
Сафонюк Иван Иванович
Тихоокеанский государственный
экономический университет
Океанский пр., 19, 690091 г. Владивосток
E-mail: cdo@psue.ru*

*Статья поступила 21 марта 2007 г.
Окончательный вариант 13 марта 2008 г.*