

УСКОРЕННОЕ МОДЕЛИРОВАНИЕ ВЕРОЯТНОСТИ БЛОКИРОВКИ ТРЕБОВАНИЙ В СЕТЯХ ОБСЛУЖИВАНИЯ С МНОЖЕСТВЕННЫМ ДОСТУПОМ

Аннотация. Рассматривается модель сети обслуживания, в которую поступают разнотипные пуассоновские потоки требований, требующие соединения между заданными терминалами. Путь соединения зависит от типа поступающего требования, от запрашиваемого ресурса, от занятых в данный момент путей и от загруженности всех каналов связи сети. Предложен метод ускоренного моделирования вероятности блокировки требований определенного потока с запрашиваемым ресурсом не ниже заданного.

Ключевые слова: сеть обслуживания, множественный доступ, вероятность блокировки, метод Монте-Карло, метод расслоенной выборки, ускоренное моделирование, дисперсия оценки.

ВВЕДЕНИЕ

В настоящее время наблюдается стремительное развитие телекоммуникационных сетей и связанных с ними информационных технологий. Создание эффективной инфраструктуры и принятие обоснованных проектных решений невозможно без создания адекватного математического аппарата оценки вероятностных характеристик сети, а также алгоритмов и соответствующего программного обеспечения [1, 2]. Одним из важных показателей, характеризующих эффективность работы сети, является вероятность блокировки [3, 4], т.е. вероятность такого состояния сети, при котором поступающие требования получают отказ и покидают сеть необслуженными.

Рассмотрим некоторые особенности телекоммуникационных сетей, существенно усложняющих их анализ известными методами. К ним относится разнотипность поступающих потоков требований, которые запрашивают существенно различные ресурсы (аудио- и видеопотоки). Линии связи (ребра сети) имеют ограниченную пропускную способность и составляют общую ресурсную базу всех потоков (множественный доступ). Поэтому выбор пути обслуживания поступающего требования зависит от уже выбранных путей обслуживания всех поступивших до него требований потоков (даже тех, обслуживание которых уже закончилось). Длительность обслуживания имеет функцию распределения общего вида, зависящую как от типа потока, так и от запрашиваемого ресурса. Существенная загрузка сети делает малоэффективными методы теории надежности, рассчитанные в основном на системы с низкой загрузкой [5–7].

Предлагаемая в настоящей статье модель сети обслуживания учитывает все перечисленные особенности. Зависимость текущего состояния сети не только от моментов поступления требований, но и от порядка их поступления, а также от запрашиваемого ресурса практически исключает использование явных аналитических, рекуррентных или асимптотических формул. В то же время малые значения вероятности блокировки делает применение стандартного метода Монте-Карло малоэффективным ввиду чрезмерных затрат вычислительного ресурса.

Наиболее перспективным направлением исследований, по мнению авторов, является создание метода направленного моделирования, позволяющего достичь форсированного попадания сети в состояние блокировки (почти в каждой реализа-

ции алгоритма). Несмещенность оценки достигается за счет соответствующего нормирующего множителя. Развитию подобных техник моделирования (называемых также ускоренным моделированием), направленных на уменьшение дисперсии оценки, уделяется особое внимание исследователей в последние несколько десятилетий. Среди многих методов уменьшения дисперсии отметим наиболее часто используемые: метод существенной выборки [8–13], аналитико-статистический метод [5–7, 14], метод расслоенной выборки [15, 16], метод многоуровневого расщепления [17, 18] и ряд других. Основные идеи, положенные в основу уменьшения дисперсии оценок, изложены в обзорах [6, 19–21].

Настоящая статья включает шесть разделов. В разд. 1 рассмотрена модель сети обслуживания с множественным доступом. В разд. 2, 3 вводится марковский процесс, описывающий текущее состояние сети и ее временное поведение, а также приводится алгоритм моделирования изменений состояний данного процесса. Алгоритм ускоренного моделирования вероятности блокировки и общая схема моделирования, основанная на методе расслоенной выборки, приведены в разд. 4, 5. Разд. 6 содержит численный пример, иллюстрирующий точность оценок, получаемых предложенным методом.

1. СЕТЬ ОБСЛУЖИВАНИЯ С МНОЖЕСТВЕННЫМ ДОСТУПОМ

Рассмотрим сеть обслуживания, структуру которой задает граф $G = [V, E]$, где V — конечное множество вершин, E — множество неориентированных ребер: $e = \langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle \in E$, $v_1, v_2 \in V$. Последовательность ребер $\langle v_0, v_1 \rangle, \langle v_1, v_2 \rangle, \dots, \langle v_{n-1}, v_n \rangle$ такая, что $\langle v_{i-1}, v_i \rangle \in E$, $i = 1, \dots, n$, называется путем, ведущим из v_0 в v_n . При этом предполагается отсутствие циклов, т.е. $v_i \neq v_j$, если $i \neq j$. Для каждого ребра $e \in E$ задана его максимальная пропускная способность c_e .

Сеть предназначена для обслуживания потока поступающих требований. Задано множество пар $W \subset V \times V$. Если $w = (v_1, v_2) \in W$, то это значит, что между вершинами (терминалами) v_1 и v_2 может быть установлена связь. Интенсивность входящего потока требований на установление такой связи равна λ_w (далее требования w -го типа).

Каждое поступающее требование w -го типа для своего обслуживания запрашивает определенный ресурс — это дискретная случайная величина θ_w , принимающая значения из конечного множества $\Theta_w = \{r_w^{(1)}, \dots, r_w^{(k_w)}\}$, а именно $\theta_w = r_w^{(m)}$ с вероятностью $q_w^{(m)}$ ($m = 1, \dots, k_w$). Для обслуживания каждого требования w -го типа задано множество S_w путей, упорядоченных в порядке возрастания длины (длина пути — это количество ребер, из которых он состоит). Для обслуживания выбирается первый доступный путь. Путь $s \in S_w$ называется доступным для требования w -го типа с запрашиваемым ресурсом $r = r_w^{(m)} \in \Theta_w$, если текущая пропускная способность каждого ребра пути s не ниже, чем r . Если требование занимает путь s , то пропускная способность каждого ребра пути уменьшается на r . Если в момент поступления требования w -го типа с запрашиваемым ресурсом r множество доступных путей является пустым, то требование получает отказ и покидает сеть необслуженным. Такое состояние сети называется состоянием блокировки требований w -го типа с запрашиваемым ресурсом не ниже, чем r . Поскольку разные пути могут содержать одни и те же ребра, то поступление требований одного типа может существенно влиять на блокировку требований другого типа (множественный доступ). Если требование w -го типа с запрашиваемым ресурсом $r_w^{(m)}$ принято на обслуживание, то его длительность определяется случайной величиной $\eta_w^{(m)}$ с функцией

распределения $G_w^{(m)}(x)$. При этом предполагаем, что существует плотность распределения $g_w^{(m)}(x)$ и $\mathbf{M}\eta_w^{(m)} < \infty$. По окончании обслуживания пропускная способность каждого ребра пути s повышается на $r_w^{(m)}$.

Цель исследования — оценка стационарной вероятности $Q_w(r)$ того, что сеть находится в состоянии блокировки требований w -го типа с запрашиваемым ресурсом, не меньшим, чем r , $r \in \Theta_w$.

2. ПОСТРОЕНИЕ МАРКОВСКОГО ПРОЦЕССА

Состояние сети обслуживания и ее временное поведение описываются непрерывным справа марковским процессом следующего вида:

$$\zeta(t) = (\sigma(t); w^{(m)}(t); s^{(m)}(t); \theta^{(m)}(t); \gamma^{(m)}(t), m = 1, \dots, \sigma(t)), t \geq 0,$$

где $\sigma(t)$ — общее количество требований, обслуживаемых в сети в момент t ; $w^{(m)}(t)$ — тип m -го требования, задающий номера вершин, которые соединяет путь $s^{(m)}(t)$; $\theta^{(m)}(t)$ — ресурс, необходимый для обслуживания m -го требования; $\gamma^{(m)}(t)$ — время, которое остается до окончания его обслуживания.

Заметим, что все введенные величины являются одномерными, за исключением пути $s^{(m)}(t)$, который имеет следующий вид: если $w^{(m)}(t) = (v_1, v_2)$, то

$$s^{(m)}(t) = (\langle v_0^{(m)}, v_1^{(m)} \rangle, \langle v_1^{(m)}, v_2^{(m)} \rangle, \dots, \langle v_{n^{(m)}-1}^{(m)}, v_{n^{(m)}}^{(m)} \rangle), v_0^{(m)} = v_1, v_{n^{(m)}}^{(m)} = v_2,$$

где $n^{(m)}$ — количество ребер в пути. Имеющаяся в момент t пропускная способность $\alpha_e(t)$ каждого ребра $e = \langle z_1, z_2 \rangle \in E$ определяется по формуле

$$\alpha_e(t) = c_e - \sum_{m=1}^{\sigma(t)} \theta^{(m)}(t) \sum_{l=1}^{n^{(m)}} I(v_{l-1}^{(m)} = z_1, v_l^{(m)} = z_2), \quad (1)$$

где $I(\cdot)$ — индикатор соответствующего события. Тогда ресурс $\Delta_w(t)$ для обслуживания требований w -го типа ($w \in W$) определяется по формуле

$$\Delta_w(t) = \max_{s \in S_w} \min_{e \in s} \alpha_e(t). \quad (2)$$

Обозначим $Q_w(t; r) = \mathbf{P}\{\Delta_w(t) < r\}$, где $r \in \Theta_w$. Цель исследования — оценка стационарной вероятности

$$Q_w(r) = \lim_{t \rightarrow \infty} Q_w(t; r)$$

при фиксированных $w \in W$ и $r \in \Theta_w$.

3. МОДЕЛИРОВАНИЕ МЕТОДОМ МОНТЕ-КАРЛО

Как известно, метод Монте-Карло ввиду своей универсальности является наиболее используемым в инженерной практике методом оценки вероятностных характеристик систем и сетей, структурная сложность и функциональные особенности которых делают невозможным использование аналитических (в частности, асимптотических) методов. Если оцениваются вероятности маловероятных событий (например, попадание в состояние блокировки), то единственной альтернативой является применение специальных алгоритмов моделирования (ускоренное моделирование) для понижения дисперсии оценок за счет использования встроенных в алгоритм аналитических формул вычисления малых вероятностей. Тем не менее алгоритм метода Монте-Карло является обязатель-

ным элементом ускоренного моделирования; этим объясняется детальное описание данного алгоритма в настоящем разделе.

Предположим, что моделированием необходимо определить состояние процесса $\zeta(u)$, $u \geq 0$, в момент t при условии отсутствия требований в начальный момент времени, т.е. $\sigma(0) = 0$. Один из наиболее простых способов моделирования состоит в следующем.

Обозначим $\Lambda = \sum_{\mu \in W} \lambda_{\mu}$ суммарную интенсивность всех потоков требований, поступающих в сеть; $p_{\mu} = \lambda_{\mu} / \Lambda$ — вероятность того, что поступившее требование является μ -го типа. Алгоритм построения оценки в одной реализации для вероятности блокировки в момент t требований w -го типа с запрашиваемым ресурсом, не меньшим, чем $r \in \Theta_w$, состоит из семи шагов.

1. Задаем состояние сети в начальный момент времени: $\sigma(0) = 0$.

2. Моделируем количество K требований, поступивших в промежутке $[0, t]$. Это случайная величина, имеющая распределение Пуассона с параметром Λt , т.е.

$$\mathbf{P}\{K = n\} = \frac{(\Lambda t)^n}{n!} e^{-\Lambda t}, \quad n = 0, 1, \dots$$

3. Моделируем K независимых равномерно распределенных в $[0, t]$ случайных величин $\omega_1, \dots, \omega_K$ и упорядочиваем их в порядке возрастания. В результате получим последовательность $0 < t_1 < t_2 < \dots < t_K < t$ моментов поступления требований. Положим $t_0 = 0$.

4. Для каждого поступающего требования ($n = 1, \dots, K$) моделируем:

— тип требования: $\mu_n = \mu$ с вероятностью $p_{\mu} = \lambda_{\mu} / \Lambda$, $\mu \in W$;

— требуемый ресурс: $\theta_n = r_{\mu_n}^{(m_n)}$, где $m_n = m$ с вероятностью $q_{\mu_n}^{(m)}$,

$m = 1, \dots, k_{w_n}$;

— длительность обслуживания η_n — случайную величину с функцией распределения $G_{\mu_n}^{(m_n)}(x)$.

5. Для каждого $n = 1, \dots, K$ при известном состоянии $\zeta(t_{n-1})$ находим:

— состояние $\zeta(t_n - 0)$, если известно, что в интервале (t_{n-1}, t_n) требования не поступали;

— состояние $\zeta(t_n)$, если известно, что в момент t_n поступило требование μ_n -го типа с запрашиваемым ресурсом θ_n и длительностью обслуживания η_n .

6. Находим состояние $\zeta(t)$, если известно, что в интервале (t_K, t) требования не поступали.

7. Согласно формулам (1) и (2) находим ресурс $\Delta_w(t)$, который имеется в момент t для обслуживания требований w -го типа. Если $\Delta_w(t) < r$, то оценкой в одной реализации для вероятности $Q_w(t; r)$ выбираем $\hat{Q}_w^{(1)}(t; r) = 1$. В противном случае полагаем $\hat{Q}_w^{(1)}(t; r) = 0$.

Приведенный достаточно простой алгоритм непосредственного моделирования изменений состояний сети позволяет оценить вероятность блокировки в фиксированный момент t . Увеличивая t , можно проследить изменение вероятности $Q_w(t; r)$ и найти момент t_0 , начиная с которого колебания $Q_w(t; r)$ будут несущественными с практической точки зрения (данное утверждение не является математически четко сформулированным; в то же время, построив доверительные интервалы оценок, легко определить момент, начиная с которого все интервалы будут иметь общие точки).

Существенным недостатком метода Монте-Карло является его низкая эффективность при оценке малых вероятностей, т.е. необходимо найти способ моделирования, позволяющий достичь более частого попадания в состояние блокировки, сохраняя при этом свойство несмещенности оценки.

4. АЛГОРИТМ УСКОРЕННОГО МОДЕЛИРОВАНИЯ ВЕРОЯТНОСТИ БЛОКИРОВКИ

Прежде чем формулировать непосредственно алгоритм, приведем некоторые замечания, раскрывающие его основную суть. Пусть t — некоторый момент времени, а T — длительность промежутка времени, который начинается в момент t (выбор T будет описан ниже). До момента t моделируем сеть обслуживания методом Монте-Карло (см. разд. 3), а в промежутке $[t, t+T]$ используем ускоренное моделирование, в основе которого метод расслоенной выборки. Далее момент t будем считать нулевым и сеть обслуживания моделировать в промежутке $[0, T]$, считая начальным состоянием процесса $\zeta(0) = \zeta_0$ то состояние, в которое сеть перейдет за время t . Цель ускоренного моделирования — оценка вероятности $Q_w(T; r)$ при фиксированных $w \in W$ и $r \in \Theta_w$.

Предположим, что в промежутке $[0, T]$ поступило n требований (распределение Пуассона). Моменты $\omega_1, \dots, \omega_n$ поступления каждого требования имеют равномерное распределение в $[0, T]$. В отличие от метода Монте-Карло эти моменты не упорядочиваем в порядке возрастания. Параметром метода, задаваемого среди начальных данных, выберем вероятность $p \in (0, 1)$ того, что обслуживание поступившего требования не окончится к моменту T :

$$p = \varphi(T) = \mathbf{P} \{ \omega_i + \eta_i > T \} = \frac{1}{T} \int_0^T [1 - G(T - x)] dx = \frac{1}{T} \int_0^T [1 - G(x)] dx. \quad (3)$$

Здесь $\{ \eta_i \}$ — независимые одинаково распределенные случайные величины (длительности обслуживания), имеющие функцию распределения

$$G(x) = \mathbf{P} \{ \eta_i < x \} = \sum_{\mu \in W} p_\mu \sum_{m=1}^{k_w} q_\mu^{(m)} G_\mu^{(m)}(x), \quad (4)$$

где $p_\mu = \lambda_\mu / \Lambda$, $\mu \in W$. Поскольку $\mathbf{M} \eta_\mu^{(m)} < \infty$ для всех $\mu \in W$ и m , то функция $\varphi(x)$ является непрерывной и монотонно убывающей, причем $\varphi(0) = 1$ и $\varphi(\infty) = 0$. Поэтому существует единственное решение уравнения (3). Именно $T = T(p)$ используется в дальнейшем.

Обозначим L число требований, поступивших в промежутке $[0, T]$. Эта случайная величина имеет распределение Пуассона с параметром ΛT . Если $L = n$, то число требований ν_T , обслуживание которых не закончится к моменту T , имеет биномиальное распределение с параметрами n и p . Обозначим ν_0 число требований, находящихся в сети при начальном состоянии $\zeta(0) = \zeta_0$, обслуживание которых не закончится к моменту T . Справедливы следующие соотношения:

$$Q_w(T; r) = \sum_{z=1}^{\infty} Q_w^{(z)}(T; r), \quad (5)$$

$$Q_w^{(z)}(T; r) = \mathbf{P} \{ \nu_0 + \nu_T = z, L \geq \nu_T, \Delta_w(T) < r \}. \quad (6)$$

Формулы (5), (6) реализуют принцип расслоенной выборки. Параметром z расслоения является число требований, которые могли бы находиться на обслуживании в момент T , если бы ни одно из них не получило отказ (блокировка) в момент поступления. Обозначим $z_0(w, r)$ минимальное количество требований, нахождение которых в сети может привести к попаданию в состояние блокиров-

ки требований w -го типа с запрашиваемым ресурсом не ниже, чем r . Это значит, что $Q_w^{(z)}(T; r) = 0$ при $z < z_0(w, r)$. Пусть $z_{\max}(w, r)$ — некоторое фиксированное число такое, что $z_{\max}(w, r) \geq z_0(w, r)$ ($z_{\max}(w, r)$ также является параметром метода; рекомендации по его выбору указаны ниже). Тогда формулу (5) можно записать в виде

$$Q_w(T; r) = \sum_{z=z_0(w, r)}^{z_{\max}(w, r)-1} Q_w^{(z)}(T; r) + Q_w^*(T; r), \quad (7)$$

где

$$Q_w^*(T; r) = \mathbf{P} \{ \nu_0 + \nu_T \geq z_{\max}(w, r), L \geq \nu_T, \Delta_w(T) < r \}. \quad (8)$$

Алгоритмы моделирования вероятностей $Q_w^{(z)}(T; r)$ и $Q_w^*(T; r)$ являются подобными, отличаясь незначительно. Пусть $z \in \{z_0(w, r), \dots, z_{\max}(w, r) - 1\}$ является фиксированным числом. Тогда

$$Q_w^{(z)}(T; r) = \mathbf{M}_{\xi(0)} \left\{ \sum_{n=z-\nu_0}^{\infty} \frac{(\Lambda T)^n}{n!} e^{-\Lambda T} C_n^{z-\nu_0} p^{z-\nu_0} (1-p)^{n-z+\nu_0} \times \right. \\ \left. \times \mathbf{P} \{ \Delta_w(T) < r \mid L = n, \nu_0 + \nu_T = z \} \right\}, \quad (9)$$

$$Q_w^*(T; r) = \mathbf{M}_{\xi(0)} \left\{ \sum_{j=z_{\max}(w, r)-\nu_0}^{\infty} \sum_{n=j}^{\infty} \frac{(\Lambda T)^n}{n!} e^{-\Lambda T} C_n^j p^j (1-p)^{n-j} \times \right. \\ \left. \times \mathbf{P} \{ \Delta_w(T) < r \mid L = n, \nu_0 + \nu_T = j \} \right\}, \quad (10)$$

где символ $\mathbf{M}_{\xi(0)}$ означает усреднение по состояниям сети обслуживания в начальный момент времени. Формулы (9) и (10) описывают случаи, когда в промежутке $[0, T]$ поступит ровно $z - \nu_0$ требований, обслуживание которых не закончится к моменту T , и соответственно не менее $z_{\max}(w, r) - \nu_0$ таких требований.

Формулы (9) и (10) служат основой для алгоритмов построения несмещенных оценок для вероятностей $Q_w^{(z)}(T; r)$ и $Q_w^*(T; r)$. Алгоритм построения оценки $\hat{Q}_{w1}^{(z)}(T; r)$ в одной реализации формулируется следующим образом (изменения к этому алгоритму для построения оценки $\hat{Q}_{w1}^*(T; r)$ указаны ниже в замечаниях к алгоритму).

1. Методом Монте-Карло моделируем состояние сети обслуживания в начальный момент времени: $\xi(0) = \xi_0$ (см. алгоритм, описанный в разд. 3). При этом определяем как число требований ν_0 , обслуживание которых не закончится к моменту T , так и задействованный этими требованиями ресурс.

2. При условии отсутствия поступления требований в промежутке $[0, T]$ проверяем, находится ли сеть в состоянии блокировки по отношению к требованиям w -го типа с запрашиваемым ресурсом, не меньшим, чем r . В случае блокировки полагаем $\hat{Q}_{w1}^{(z)}(T; r) = 1$; оценка построена. В противном случае переходим на следующий шаг алгоритма.

3. Вычисляем (см. (9)) $A(z - \nu_0)$, где

$$A(k) = \sum_{n=k}^{\infty} a_n(k), \quad k \geq 0, \quad (11)$$

$$a_n(k) = \frac{(\Lambda T)^n}{n!} e^{-\Lambda T} C_n^k p^k (1-p)^{n-k}, \quad n \geq k. \quad (12)$$

4. Полагаем $\hat{Q}_{wl}^{(z)}(T; r) = A(z - \nu_0)$ (начальное значение оценки).

5. Моделируем число L требований, поступивших в $[0, T]$, а именно $L = n$ с вероятностью $a_n(z - \nu_0) / A(z - \nu_0)$. Пусть $L = n$. Это значит, что в промежутке $[0, T]$ поступило n требований, причем у $z - \nu_0$ требований из них обслуживание не закончится к моменту T . В то же время у $n - z + \nu_0$ требований обслуживание закончится до момента T .

6. Моделируем моменты поступления требований и соответствующие длительности обслуживания. Для этого моделируем равномерно распределенные в промежутке $[0, T]$ случайные величины $\{\omega_i\}$ и длительности обслуживания $\{\eta_i\}$ с функцией распределения (4). Если $\omega_i + \eta_i > T$, то ω_i — момент поступления требования, обслуживание которого не закончится к моменту T (общее количество таких моментов должно составить $z - \nu_0$). Если $\omega_i + \eta_i < T$, то обслуживание данного требования закончится до момента T (общее количество таких моментов $\{\omega_i\}$ должно составить $n - z + \nu_0$).

7. Упорядочиваем моменты поступления в порядке возрастания: $0 = t_0 < t_1 < \dots < t_n$. Соответствующие длительности обслуживания обозначим y_1, \dots, y_n .

8. В каждый момент t_i при известной длительности обслуживания y_i моделируем: а) как тип μ_i поступившего требования, так и запрашиваемый им ресурс θ_i (если $t_i + y_i < T$); б) только тип μ_i поступившего требования (если $t_i + y_i > T$); запрашиваемые ресурсы требований, которые находятся в сети на обслуживании в момент T , должны быть выбраны таким образом, чтобы сеть в момент T находилась в состоянии блокировки требований определенного типа.

Если $t_i + y_i < T$, то двухмерную случайную величину (μ_i, θ_i) моделируем согласно распределению

$$\mathbf{P}\{\mu_i = \mu, \theta_i = r_\mu^{(m)}\} = \frac{p_\mu q_\mu^{(m)} g_\mu^{(m)}(y_i)}{\sum_{\pi \in W} p_\pi \sum_{j=1}^{k_\pi} q_\pi^{(j)} g_\pi^{(j)}(y_i)}, \quad \mu \in W, \quad m = 1, \dots, k_\mu,$$

где $\{g_\mu^{(m)}(y_i)\}$ — плотности распределения соответствующих функций распределения при известной длительности обслуживания y_i .

Если $t_i + y_i > T$, то случайную величину μ_i моделируем согласно распределению

$$\mathbf{P}\{\mu_i = \mu\} = \frac{p_\mu \sum_{j=1}^{k_\mu} q_\mu^{(j)} g_\mu^{(j)}(y_i)}{\sum_{\pi \in W} p_\pi \sum_{j=1}^{k_\pi} q_\pi^{(j)} g_\pi^{(j)}(y_i)}, \quad \mu \in W.$$

9. Находим $i_1 = \min\{i: t_i + y_i > T\}$ и моделируем состояние сети в момент $t_{i_1} - 0$ (известны все моменты поступления требований, их длительности обслуживания, типы и запрашиваемые ресурсы).

10. Запоминаем состояние сети $\zeta^* = \zeta(t_{i_1} - 0)$.

11. Требование типа μ_{i_1} , поступающее в момент t_{i_1} , может запросить один из $k_{\mu_{i_1}}$ ресурсов множества $\Theta_{\mu_{i_1}}$. При начальном состоянии ζ^* , предполагая, что запрашиваемым ресурсом в момент t_{i_1} является $r_{\mu_{i_1}}^{(m)} \in \Theta_{\mu_{i_1}}$, моделируем состояние сети в момент T , считая, что в каждый момент $t_i > t_{i_1}$, $t_i + y_i > T$ запрошен максимально допустимый ресурс, не приводящий к потере требования в этот момент времени. Если при этом сеть попадает в состояние блокировки требований

w -го типа с запрашиваемым ресурсом не ниже, чем r , то полагаем $J(r_{\mu_{i_1}}^{(m)}) = 1$, и $J(r_{\mu_{i_1}}^{(m)}) = 0$ — в противном случае.

12. Вычисляем вероятности ресурсов в момент t_{i_1} :

$$b_{\mu_{i_1}}^{(m)}(r_{\mu_{i_1}}^{(m)}) = \mathbf{P}\{\theta_{i_1} = r_{\mu_{i_1}}^{(m)}\} = \frac{q_{\mu_{i_1}}^{(m)} g_{\mu_{i_1}}^{(m)}(y_{i_1})}{\sum_{j=1}^{k_{\mu_{i_1}}} q_{\mu_{i_1}}^{(j)} g_{\mu_{i_1}}^{(j)}(y_{i_1})}, \quad r_{\mu_{i_1}}^{(m)} \in \Theta_{\mu_{i_1}}.$$

13. Вычисляем

$$B = \sum_{r_{\mu_{i_1}}^{(m)} \in \Theta_{\mu_{i_1}}} J(r_{\mu_{i_1}}^{(m)}) b_{\mu_{i_1}}^{(m)}(r_{\mu_{i_1}}^{(m)})$$

и полагаем $\hat{Q}_{wl}^{(z)}(T; r) := \hat{Q}_{wl}^{(z)}(T; r) \cdot B$, где символ $:=$ означает, что новому значению оценки присваивается функция от ее старого значения.

14. Если $B = 0$, то алгоритм окончен, в данной реализации сеть не может попасть в состояние блокировки требований w -го типа с запрашиваемым ресурсом не ниже, чем r , т.е. $\hat{Q}_{wl}^{(z)}(T; r) = 0$. Если $B > 0$, то согласно распределению $\frac{1}{B} J(r_{\mu_{i_1}}^{(m)}) b_{\mu_{i_1}}^{(m)}(r_{\mu_{i_1}}^{(m)})$, $r_{\mu_{i_1}}^{(m)} \in \Theta_{\mu_{i_1}}$, моделируем запрашиваемый ресурс θ_{i_1} в момент t_{i_1} .

15. Моделируем состояние сети в момент t_{i_1} , если известно, что $\xi(t_{i_1} - 0) = \xi^*$ и в этот момент поступает требование μ_{i_1} -го типа с длительностью обслуживания y_{i_1} и запрашиваемым ресурсом θ_{i_1} .

16. Находим $i_2 = \min\{i > i_1 : t_i + y_i > T\}$ и моделируем состояние сети в момент $t_{i_2} - 0$ (известны все моменты поступления требований, их длительности обслуживания, типы и запрашиваемые ресурсы).

Затем, начиная с шага 10, алгоритм повторяется с заменой i_1 на i_2 . Данный циклический алгоритм продолжается до тех пор, пока ни будут промоделированы запрашиваемые ресурсы всех $z - \nu_0$ требований, обслуживание которых не заканчивается к моменту T . В результате будет построено состояние сети в момент T , которое является состоянием блокировки требований w -го типа с запрашиваемым ресурсом не ниже, чем r . В качестве оценки выбираем $\hat{Q}_{wl}^{(z)}(T; r)$.

Замечания. 1. Алгоритм ускоренного моделирования основан на аналитическом вычислении вероятностей некоторых событий, наступление которых является необходимым условием попадания сети в состояние блокировки, и на моделировании состояния сети при условии наступления этих событий, что гарантирует несмещенность оценки.

2. Если на первом цикле алгоритма, относящегося к моменту t_{i_1} , имеем $B > 0$, то сеть гарантированно попадает в состояние блокировки в момент T .

3. Вероятность p легко выбрать таким образом, что $\nu_0 = 0$ с вероятностью, близкой к единице, что уменьшает дисперсию оценки.

4. Алгоритм построения оценки $\hat{Q}_{wl}^*(T; r)$ в одной реализации для $Q_w^*(T; r)$ формулируется аналогично описанному выше. Единственное изменение: шаг 3 и шаг 4 алгоритма заменяем на следующие действия:

— вычисляем (см. (10))

$$D(\nu_0) = \sum_{k=z_{\max}(w,r)-\nu_0}^{\infty} A(k), \quad k \geq 0,$$

где $\{A(k)\}$ определяются согласно (11), (12);

— моделируем число требований κ , обслуживание которых не закончится к моменту T : $\kappa = k$ с вероятностью $A(k)/D(\nu_0)$;

— полагаем $\hat{Q}_{w1}^*(T; r) = D(\nu_0)$ (начальное значение оценки);

— начиная с шага 5, алгоритм полностью повторяет описанный выше (с заменой $z - \nu_0$ на κ).

5. ПРИМЕНЕНИЕ МЕТОДА РАССЛОЕННОЙ ВЫБОРКИ ДЛЯ ОЦЕНКИ ВЕРОЯТНОСТИ $Q_w(T; r)$

Формулы (7), (8) удобны для использования метода расслоенной выборки. Предположим, что параметры p и $z_{\max}(w, r)$ фиксированы (заметим, что $T = T(p)$ является решением уравнения (3)). Тогда оценка $\hat{Q}_w(T; r)$ вероятности $Q_w(T; r)$ строится по формуле

$$\hat{Q}_w(T; r) = \sum_{z=z_0(w,r)}^{z_{\max}(w,r)-1} \frac{1}{N_z} \sum_{i=1}^{N_z} \hat{Q}_{wi}^{(z)}(T; r) + \frac{1}{N^*} \sum_{i=1}^{N^*} \hat{Q}_{wi}^*(T; r), \quad (13)$$

где N_z и N^* — количества реализаций алгоритма, использованных для построения оценок $\{\hat{Q}_{wi}^{(z)}(T; r)\}$ и $\{\hat{Q}_{wi}^*(T; r)\}$ соответствующих вероятностей. Поскольку все оценки являются независимыми случайными величинами, то

$$\mathbf{D}\hat{Q}_w(T; r) = \sum_{z=z_0(w,r)}^{z_{\max}(w,r)-1} \frac{\mathbf{D}\hat{Q}_{w1}^{(z)}(T; r)}{N_z} + \frac{\mathbf{D}\hat{Q}_{w1}^*(T; r)}{N^*}. \quad (14)$$

Предположим, что N — общее количество реализаций для построения оценки $\hat{Q}_w(T; r)$. Как известно, минимальное значение дисперсии достигается при

$$N_z = N \cdot \frac{\sqrt{\mathbf{D}\hat{Q}_{w1}^{(z)}(T; r)}}{D_w(T; r)}, \quad N^* = N \cdot \frac{\sqrt{\mathbf{D}\hat{Q}_{w1}^*(T; r)}}{D_w(T; r)}, \quad (15)$$

где

$$D_w(T; r) = \sum_{z=z_0(w,r)}^{z_{\max}(w,r)-1} \sqrt{\mathbf{D}\hat{Q}_{w1}^{(z)}(T; r)} + \sqrt{\mathbf{D}\hat{Q}_{w1}^*(T; r)}. \quad (16)$$

Подставив (15) в (14), получим

$$\mathbf{D}\hat{Q}_w(T; r) = \frac{1}{N} [D_w(T; r)]^2. \quad (17)$$

Согласно центральной предельной теореме оценка $\hat{Q}_w(T; r)$ имеет асимптотически нормальное распределение. Поэтому относительная погрешность оценки, построенной за N реализаций алгоритма с достоверностью $1 - \gamma$, вычисляется по формуле

$$R_w(T; r) = \frac{\alpha(\gamma) D_w(T; r)}{\sqrt{N} Q_w(T; r)} \cdot 100 \%, \quad (18)$$

где $\alpha(\gamma)$ определяется из уравнения $\frac{2}{\sqrt{2\pi}} \int_0^{\alpha(\gamma)} e^{-u^2/2} du = 1 - \gamma$.

Поскольку точные значения $Q_w(T; r)$ и $D_w(T; r)$ неизвестны, то на практике используют их статистические оценки. Иначе говоря, на предварительном этапе при относительно небольшом количестве реализаций строят оценки для $\mathbf{D}\hat{Q}_{w1}^{(z)}(T; r)$, $\mathbf{D}\hat{Q}_{w1}^*(T; r)$ и согласно (15), (16) находят оценки \hat{N}_z и \hat{N}^* . Затем на основном этапе моделирования согласно (13) строят оценку для $\hat{Q}_w(T; r)$, заменив N_z и N^* на \hat{N}_z и \hat{N}^* соответственно. При этом моделирование на основном этапе можно организовать без потери информации, полученной на предварительном этапе (т.е. с учетом всех полученных ранее оценок).

Рассмотрим более подробно рекомендации по выбору p и $z_{\max}(w, r)$. Если выбрать $p=1$, то получим классический метод Монте-Карло. Если $p \rightarrow 0$, то $T = T(p) \rightarrow \infty$ (см. (3)) и время моделирования сильно возрастет. Реально же условием выбора T может служить соотношение

$$T = \inf \{t: \mathbf{P} \{ \eta > t \} = 1 - G(t) < \varepsilon \}, \quad (19)$$

где функция распределения длительности обслуживания определяется согласно (4), а ε — малая величина (например, $\varepsilon = 10^{-5}$). В этом случае с вероятностью, близкой к единице, $\nu_0 = 0$, т.е. максимально уменьшается влияние непосредственного моделирования на состояние сети в момент T . Выбрав T из условия (19), находим $p = \varphi(T)$ (см. (3)). Величина $z_{\max}(w, r)$ также влияет на дисперсию оценки. Если выбрать $z_{\max}(w, r) = z_0(w, r)$, то расслоения не произойдет и значение дисперсии будет существенно отличаться от оптимального. С увеличением $z_{\max}(w, r)$ дисперсия уменьшится, но при этом незначительно увеличится затрачиваемое время (за счет предварительного этапа). Выбор параметра $z_{\max}(w, r)$ можно осуществлять согласно следующему соотношению (см. (16)):

$$\sqrt{\mathbf{D}\hat{Q}_{w1}^*(T; r)} \ll \sum_{z=z_0(w, r)}^{z_{\max}(w, r)-1} \sqrt{\mathbf{D}\hat{Q}_{w1}^{(z)}(T; r)}. \quad (20)$$

Выбор величин ε , p , T и $z_{\max}(w, r)$ рассмотрим более подробно на численном примере.

6. ЧИСЛЕННЫЙ ПРИМЕР

Рассмотрим сеть обслуживания, изображенную на рис. 1. Она включает 13 вершин и 26 ребер. Характеристики этой сети зададим в соответствии с описанной выше моделью.

Пусть $W = \{(1, 13), (2, 13), (3, 13)\}$, т.е. имеются три пары терминалов, а соответственно и три входящих потока требований. Положим $\lambda_{i,13} = i$, $i = 1, 2, 3$. Для каждой указанной пары зададим множество доступных путей обслуживания по-

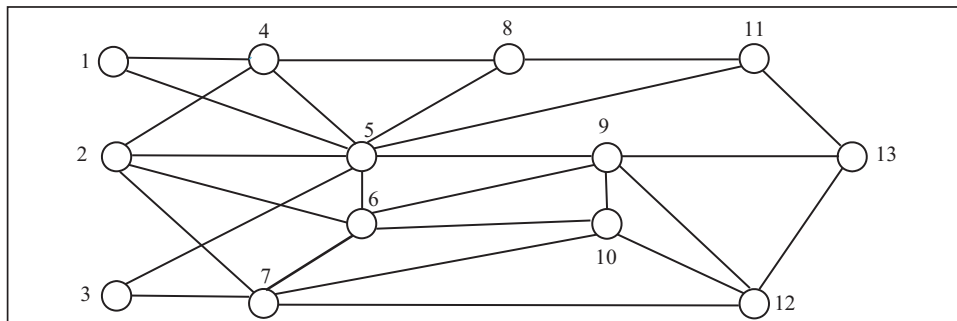


Рис. 1. Сеть обслуживания

ступающих требований (в виде последовательности вершин, а не последовательности ребер):

$$S_{1,13} = \{(1, 5, 9, 13), (1, 5, 11, 13), (1, 4, 5, 9, 13), (1, 4, 5, 11, 13), \\ (1, 4, 8, 11, 13), (1, 5, 6, 9, 13), \dots \}, \\ S_{2,13} = \{(2, 5, 9, 13), (2, 5, 11, 13), (2, 6, 9, 13), (2, 7, 12, 13), \\ (2, 4, 5, 9, 13), (2, 4, 5, 11, 13), \dots \}, \\ S_{3,13} = \{(3, 5, 9, 13), (3, 5, 11, 13), (3, 7, 12, 13), (3, 5, 6, 9, 13), \\ (3, 5, 8, 11, 13), (3, 5, 9, 12, 13), \dots \}.$$

Предположим, что все ребра имеют одну и ту же пропускную способность $c_e = c$, $e \in E$ (значение c определено ниже). Каждое поступающее требование w -го типа для своего обслуживания требует определенный ресурс — дискретную случайную величину θ_w , принимающую значения из множества $\Theta_w = \{2, 3, 5, 7\}$ с одной и той же вероятностью 0.25. Длительность обслуживания требования $w = (i, j) \in W$ имеет распределение Вейбулла:

$$\mathbf{P} \{ \eta_{ij} < x \} = G_{ij}(x) = 1 - e^{-(\beta_{ij}x)^2}, \quad x \geq 0,$$

где $\beta_{1,13} = 0.5$, $\beta_{2,13} = 1$, $\beta_{3,13} = 1.5$. В рассматриваемом примере исследуется точность оценок стационарной вероятности $Q_w(r)$ для случая $w=2$ и $r=3$, т.е. вероятности того, что сеть находится в состоянии блокировки требований потока (2, 13) с запрашиваемым ресурсом, не меньшим, чем 3.

Верхней оценкой вероятности того, что длительность обслуживания превосходит T , выберем $\varepsilon = 10^{-7}$. Согласно (19) определяем $T = 7.57$. При этом $p = \varphi(T) = 0.117$ (см. (3)).

Исследуем, как выбор параметра $z_{\max}(w, r)$ (количество слоев) влияет на дисперсию оценки. Пусть $c = 30$. В этом случае попадание в состояние блокировки возможно лишь в случае, когда по крайней мере $z_0(w, r) = 12$ требований одновременно обслуживаются в сети. Поэтому минимальным значением $z_{\max}(w, r)$ является 12.

Для оптимального выбора $z_{\max}(w, r)$ оценим величину каждого слагаемого в формуле (16). Пусть $z_{\max}(w, r) = 22$. Обозначим оценки дисперсий $\mathbf{D} \hat{Q}_{wl}^{(z)}(T; r)$ и $\mathbf{D} \hat{Q}_{wl}^{(z)}(T; r)$ через $[\hat{\sigma}_w^{(z)}(T; r)]^2$ и $[\hat{\sigma}_w^*(T; r)]^2$ соответственно. Для построения каждой оценки, приведенной в табл. 1, использовано 20000 реализаций алгоритма.

Значению $z_{\max}(w, r) = 22$ соответствует $\hat{\sigma}_w^*(T; r) = 2.02 \cdot 10^{-8}$. Нетрудно видеть, что $\hat{\sigma}_w^{(20)}(T; r) + \hat{\sigma}_w^{(21)}(T; r) + \hat{\sigma}_w^*(T; r) = 2.63 \cdot 10^{-7}$, что составляет всего 1.9% суммы $\sum_{z=12}^{19} \hat{\sigma}_w^{(z)}(T; r) = 1.40 \cdot 10^{-5}$. Поэтому целесообразно выбирать

Таблица 1. Оценки дисперсий, $c = 30$, $z = 12, \dots, 21$

$\hat{\sigma}_w^{(12)}(T; r)$	$1.44 \cdot 10^{-10}$	$\hat{\sigma}_w^{(17)}(T; r)$	$2.09 \cdot 10^{-6}$
$\hat{\sigma}_w^{(13)}(T; r)$	$9.23 \cdot 10^{-8}$	$\hat{\sigma}_w^{(18)}(T; r)$	$1.08 \cdot 10^{-6}$
$\hat{\sigma}_w^{(14)}(T; r)$	$4.83 \cdot 10^{-6}$	$\hat{\sigma}_w^{(19)}(T; r)$	$4.97 \cdot 10^{-7}$
$\hat{\sigma}_w^{(15)}(T; r)$	$2.45 \cdot 10^{-6}$	$\hat{\sigma}_w^{(20)}(T; r)$	$1.86 \cdot 10^{-7}$
$\hat{\sigma}_w^{(16)}(T; r)$	$2.96 \cdot 10^{-6}$	$\hat{\sigma}_w^{(21)}(T; r)$	$5.67 \cdot 10^{-8}$

$z_{\max}(w, r) = 20$, т.е. количество слоев равно $z_{\max}(w, r) - z_0(w, r) + 1 = 9$. Дальнейшее увеличение количества слоев не приводит к заметному увеличению точности. Именно девять слоев используется во всех приведенных ниже оценках.

Оценим эффект от расслоения. Если $z_{\max}(w, r) = 12$ (всего один слой), то для $Q_w(T; r)$ и $D\hat{Q}_w(T; r)$ были построены следующие оценки: $\hat{Q}_w(T; r) = 1.63 \cdot 10^{-6}$, $\hat{\sigma}_w^2(T; r) = 4.66 \cdot 10^{-14}$ (использовано $N = 100000$ реализаций). Если $z_{\max}(w, r) = 20$, то при том же количестве реализаций получены оценки $\hat{Q}_w(T; r) = 1.70 \cdot 10^{-6}$, $\hat{\sigma}_w^2(T; r) = 1.46 \cdot 10^{-15}$. Таким образом, применение расслоения позволило уменьшить дисперсию в 32 раза.

Далее исследуем, насколько заметным является влияние начального момента t на вероятность блокировки $Q_w(T; r)$. Отметим, что в методе ускоренного моделирования в промежутке $[0, t]$ сеть моделируется методом Монте-Карло, а момент t для ускоренного моделирования рассматривается как нулевой. Проследим, насколько быстро процесс $\zeta(u)$, $u \geq 0$, входит в стационарный режим. Соответствующие оценки для вероятности блокировки $Q_w(T; r)$, дисперсии $D\hat{Q}_w(T; r)$ и относительной погрешности $R_w(T; r)$ (достоверность $1 - \gamma = 0.95$), построенные при различных значениях t , приведены в табл. 2 (для построения каждой оценки использовано $N = 1000000$ реализаций).

Приведенные оценки демонстрируют практически полное отсутствие зависимости от начального момента t , что свидетельствует о достаточно быстром вхождении процесса $\zeta(u)$, $u \geq 0$, в стационарный режим. При этом относительная погрешность оценок, определяемая дисперсией, изменяется от 1.22 % до 1.61 %.

Исследуем изменение относительной погрешности с возрастанием пропускной способности дуг. Соответствующие оценки для вероятности блокировки $Q_w(T; r)$, дисперсии $D\hat{Q}_w(T; r)$ и относительной погрешности $R_w(T; r)$ (достоверность $1 - \gamma = 0.95$), построенные при различных значениях c , приведены в табл. 3 (для построения каждой оценки использовано $N = 1000000$ реализаций).

Данная таблица наглядно иллюстрирует эффективность предложенного метода ускоренного моделирования: при значительном убывании вероятности блокировки (c $8.86 \cdot 10^{-4}$ до $5.15 \cdot 10^{-10}$) наблюдается сравнительно небольшой рост относительной погрешности (c 0.51% до 2.60 %).

Таблица 2

Обозначение оценки	Оценки для $Q_w(T; r)$, $D\hat{Q}_w(T; r)$ и $R_w(T; r)$ при различных t				
	$t = 0$	$t = 5$	$t = 10$	$t = 15$	$t = 20$
$\hat{Q}_w(T; r)$	$1.68 \cdot 10^{-6}$	$1.68 \cdot 10^{-6}$	$1.69 \cdot 10^{-6}$	$1.67 \cdot 10^{-6}$	$1.68 \cdot 10^{-6}$
$\hat{\sigma}_w^2(T; r)$	$1.91 \cdot 10^{-16}$	$1.08 \cdot 10^{-16}$	$1.24 \cdot 10^{-16}$	$1.16 \cdot 10^{-16}$	$1.40 \cdot 10^{-16}$
$\hat{R}_w(T; r)$, %	1.61	1.22	1.29	1.26	1.38

Таблица 3

Обозначение оценки	Оценки для $Q_w(T; r)$, $D\hat{Q}_w(T; r)$ и $R_w(T; r)$ при различных c				
	$c = 20$	$c = 25$	$c = 30$	$c = 35$	$c = 40$
$\hat{Q}_w(T; r)$	$8.86 \cdot 10^{-4}$	$4.22 \cdot 10^{-5}$	$1.69 \cdot 10^{-6}$	$3.76 \cdot 10^{-8}$	$5.15 \cdot 10^{-10}$
$\hat{\sigma}_w^2(T; r)$	$5.36 \cdot 10^{-12}$	$2.31 \cdot 10^{-14}$	$1.24 \cdot 10^{-16}$	$1.47 \cdot 10^{-19}$	$4.67 \cdot 10^{-23}$
$\hat{R}_w(T; r)$, %	0.51	0.71	1.29	2.00	2.60

ЗАКЛЮЧЕНИЕ

Известно, что протоколы множественного доступа определяют порядок использования общего ресурса каналов связи. Сложность исследования сетей обслуживания с множественным доступом состоит в том, что путь соединения (обслуживания) существенно зависит от типа поступающего требования, запрашиваемого ресурса, от занятых к данному моменту путей и от загруженности всех каналов связи сети. В статье построен метод моделирования, основанный на сочетании метода расслоенной выборки и направленного моделирования, позволяющего значительно увеличить вероятность попадания в состояние блокировки на каждой реализации алгоритма. Аналогичный подход можно использовать и для оценки других характеристик сети обслуживания, в частности для оптимизации пропускной способности линий связи в целях достижения заданной вероятности блокировки для каждого типа требования и каждого запрашиваемого ресурса.

СПИСОК ЛИТЕРАТУРЫ

1. Ross K.W. Multiservice loss models for broadband telecommunication networks. London: Springer-Verlag, 1995. 343 p.
2. Frenkel I.B., Karagrigoriou A., Lisnianski A., Kleyner A.V. Applied reliability engineering and risk analysis: probabilistic models and statistical inference. New York: Wiley, 2013. 448 p.
3. Nyberg E., Virtamo J., Aalto S. An exact algorithm for calculating blocking probabilities in multicast networks. Pujolle G., Perros H., Fdida S., Körner U., Stavrakakis I. (Eds.). *Networking*. Paris, 2000. P. 275–286.
4. Karvo J. Efficient simulation of blocking probabilities for multi-layer multicast streams. Gregori E., Conti M., Campbell A., Omidyar C., Zukerman M. (Eds.). *Networking*. Berlin: Springer-Verlag, 2002. P. 1020–1031.
5. Коваленко И.Н., Кузнецов Н.Ю. Методы расчета высоконадежных систем. Москва: Радио и связь, 1988. 176 с.
6. Kovalenko I.N., Kuznetsov N.Yu., Pegg Ph.A. Mathematical theory of reliability of time dependent systems with practical applications. Chichester: Wiley, 1997. 303 p.
7. Kuznetsov N.Yu. Fast simulation technique in reliability evaluation of Markovian and non-Markovian systems. *Simulation and Optimization Methods in Risk and Reliability Theory*. New York: Nova Science Publishers, 2009. P. 69–112.
8. Heidelberger P. Fast simulation of rare events in queueing and reliability models. *ACM Transactions on Modeling and Computer Simulation*. 1995. Vol. 5, Iss. 1. P. 43–85.
9. Glasserman P. Monte Carlo methods in financial engineering. New York: Springer, 2004. 575 p.
10. Ермаков С.М. Метод существенной выборки для моделирования вероятностей умеренных и больших уклонений оценок и критериев. *Теория вероятности и ее применен.* 2006. Т. 51, № 2. С. 319–332.
11. Li J., Mosleh A., Kang R. Likelihood ratio gradient estimation for dynamic reliability applications. *Reliab. Engin. and System Safety*. 2011. Vol. 96, N 12. P. 1667–1679.
12. Kouikoglou V.S., Yannis A.Ph. Review of a fast simulation method for the analysis of queueing networks. *Applied Stochastic Models and Data Analysis*. 1998. Vol. 13, Iss. 2. P. 73–83.
13. Falkner M., Devetsikiotis M., Lambadaris I. Fast simulation of networks of queues with effective and decoupling bandwidths. *ACM Transactions on Modeling and Computer Simulation*. 2021. Vol. 31, Iss. 1.
14. Кузнецов Н.Ю., Шумская А.А. Оценка опасности отказа резервированной системы методом ускоренного моделирования. *Проблемы управления и информатики*. 2013. № 3. С. 50–62.
15. Fox B.L., Glynn P.W. Discrete-time conversion for simulating finite-horizon Markov processes. *SIAM J. Appl. Math.* 1990. Vol. 50, N 5. P. 1457–1473.

16. Шумская А.А. Ускоренное моделирование коэффициента неготовности восстанавливаемой системы с ограниченной относительной погрешностью оценки. *Кибернетика и системный анализ*. 2003. № 3. С. 45–58.
17. Glasserman P., Heidelberger Ph., Shahabuddin P., Zajic T. Multilevel splitting for estimating rare event probabilities. *Oper. Research*. 1999. Vol. 47, N 4. P. 585–600.
18. Juneja S., Shahabuddin P., Zajic T. Splitting-based importance-sampling algorithm for fast simulation of Markov reliability models with general repair-policies. *IEEE Transactions on Reliab.* 2001. Vol. 50, N 3. P. 235–245.
19. Lagnoux A. Rare event simulation. *Probab. Eng. and Inf. Sci.* 2006. Vol. 20, N 1. P. 45–66.
20. Gertsbakh I.B., Shpungin Y. Models of network reliability: Analysis, combinatorics, and Monte Carlo. Boca Raton: CRC Press, 2009. 203 p.
21. Blanchet J., Lam H. Rare event simulation techniques. *Proc. of the 2011 Winter Simulation Conference*. 2011. P. 217–231.

Надійшла до редакції 27.01.2021

М.Ю. Кузнєцов, І.М. Кузнєцов

ПРИСКОРЕНЕ МОДЕЛЮВАННЯ ЙМОВІРНІСТІ БЛОКУВАННЯ ВИМОГ У МЕРЕЖАХ ОБСЛУГОВУВАННЯ ІЗ МНОЖИННИМ ДОСТУПОМ

Анотація. Розглянуто модель мережі обслуговування, в яку надходять різноманітні пуассонівські потоки вимог, що потребують з'єднання між заданими терміналами. Шлях з'єднання залежить від типу вимоги, від ресурсу, що вимагається для її обслуговування, від зайнятих у цей момент шляхів та від завантаженості всіх каналів зв'язку мережі. Запропоновано метод прискореного моделювання ймовірності блокування вимог певного потоку із необхідним для обслуговування ресурсом, рівень якого не нижче заданого.

Ключові слова: мережа обслуговування, множинний доступ, ймовірність блокування, метод Монте-Карло, метод розширеної вибірки, прискорене моделювання, дисперсія оцінки.

N.Yu. Kuznetsov, I.N. Kuznetsov

FAST SIMULATION OF THE CUSTOMER BLOCKING PROBABILITY IN QUEUEING NETWORKS WITH MULTICAST ACCESS

Abstract. A model of a queueing network with several input Poisson flows is considered. These flows require connections between given terminals. The connection path depends on the type of the customer, on the requested resource, on the paths currently occupied and on the load on all communication channels of the network. A fast simulation method for the evaluation of the blocking probability for customers of the certain flow with a required resource not lower than a given one is proposed.

Keywords: queueing network, multicast access, blocking probability, Monte Carlo method, stratified sampling, fast simulation method, variance of estimate.

Кузнєцов Микола Юрійович,

чл.-кор. НАН України, доктор техн. наук, завідувач відділу Інституту кібернетики імені В.М. Глушкова НАН України; професор Фізико-технічного інституту Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, e-mail: kuznetsov2016@icloud.com.

Кузнєцов Ігор Миколайович,

кандидат фіз.-мат. наук, старший викладач Фізико-технічного інституту Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, e-mail: sea_hawk@icloud.com.