ИНФОРМАТИКА

2015 **июль-сентябрь** **№ 3**

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

УДК 519.872

С.А. Дудин, О.С. Дудина

ИССЛЕДОВАНИЕ МНОГОЛИНЕЙНОЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С АБСОЛЮТНЫМ ПРИОРИТЕТОМ И ПОВТОРНЫМИ ВЫЗОВАМИ

Исследуется многолинейная система массового обслуживания с повторными вызовами без буфера и двумя типами запросов как модель системы когнитивного радио. Запросы первого типа имеют абсолютный приоритет. Процессы поступления запросов первого и второго типов моделируются марковскими входными потоками. Время обслуживания запроса имеет экспоненциальное распределение с зависимым от типа запроса параметром. Запросы второго типа принимаются на обслуживание, если число занятых каналов меньше заданного порогового значения. Запросы второго типа могут совершать повторные попытки попасть на обслуживание. Находится условие существования стационарного режима. Приводятся формулы для вычисления основных характеристик производительности системы.

Введение

В настоящее время ресурсы большинства проводных телекоммуникационных систем не являются дефицитными, в то время как ресурсы беспроводных сетей связи строго дефицитные. Это обусловлено тем, что доступный спектр радиочастот ограничен физической природой. В связи с быстрым ростом числа смартфонов и других мобильных устройств передача данных в беспроводных сетях может столкнуться с катастрофической нехваткой частот. Мобильные операторы и научно-исследовательские лаборатории прилагают значительные усилия для решения задачи оптимального использования спектра радиочастот. Одним из возможных решений является применение технологии когнитивного радио. В системах когнитивного радио предлагается давать возможность нелицензированным пользователям сети временно «захватывать» неиспользуемый спектр частот, не создавая при этом помех для лицензированных пользователей (см., например, [1]). Проблема оптимизации совместного доступа первичных и вторичных пользователей может быть эффективно решена с помощью теории систем массового обслуживания.

На сегодняшний день существует достаточно большое количество работ, посвященных исследованию систем когнитивного радио. Как правило, в литературе рассматриваются модели с двумя типами запросов, где запросы первого типа имеют абсолютный приоритет над запросами второго типа. В случае если запрос первого типа поступает в систему в момент, когда все приборы заняты, но на обслуживании есть запросы второго типа, обслуживание одного запроса второго типа прекращается (запрос выбивается с обслуживания) и запрос первого типа занимает освободившийся прибор. Таким образом, запрос первого типа получает отказ только тогда, когда в момент его прихода все приборы заняты запросами первого типа, т. е. запросы второго типа не оказывают влияние на обслуживание запросов первого типа. Обслуживание запросов второго типа может прерываться, что может приводить к бесполезной трате пропускной способности системы и низкому качеству обслуживания запросов второго типа. Чтобы минимизировать вероятность наступления таких прерываний, предлагается осуществлять управление доступом в систему запросов второго типа, т. е. остановить прием вторичных запросов, когда число занятых приборов превышает некоторое пороговое значение и риск выбивания вторичных запросов является высоким. Таким образом, разумно резервировать некоторое число каналов только для обслуживания первичных пользователей. Сведения о состоянии дел в данной предметной области можно получить, например, из статей [1–4].

Практически все модели систем когнитивного радио, рассматриваемые в литературе, имеют следующие существенные недостатки. В частности, предполагается, что потоки запросов лицензированных и нелицензированных пользователей описываются стационарным пуассоновским входным потоком, в то время как входные потоки могут быть коррелированными. Также предполагается, что нелицензированные (вторичные) пользователи, которые получают отказ в обслуживании в момент прихода, покидают систему навсегда, в то время как на самом деле вторичный пользователь может повторять попытки получить обслуживание.

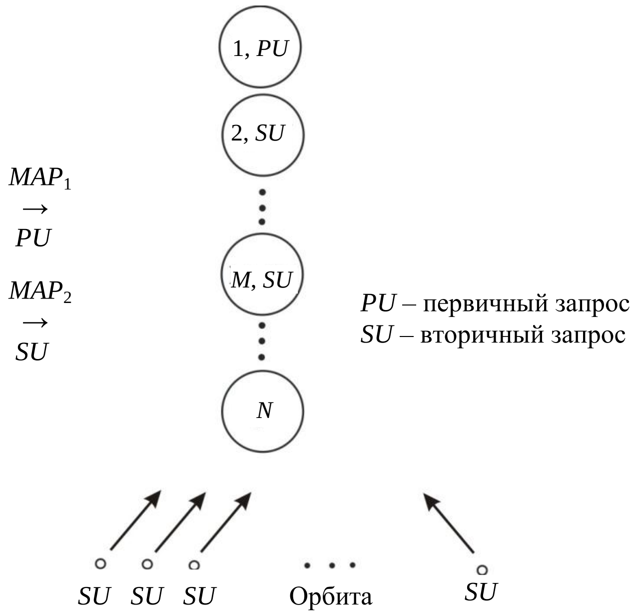
Преимущества модели, исследуемой в данной статье, по сравнению с моделями, посвященными анализу систем когнитивного радио посредством теории массового обслуживания, заключаются в следующем:

1. Предполагается, что входные потоки первичных и вторичных запросов описываются марковскими входными потоками (MAP, от англ. Markovian Arrival Process). MAP-поток коррелированный, поэтому он идеально подходит для моделирования коррелированного или пульсирующего трафика в современных телекоммуникационных сетях. Стационарный пуассоновский процесс является простейшим случаем MAP-потока. Если попытаться описать некоторый реальный поток с помощью стационарного пуассоновского входного потока, то модель будет отражать только среднюю скорость поступления, но не дисперсию или более высокие моменты распределения времени между моментами поступления запросов и возможную корреляцию между этими временами.

2. Предполагается, что вторичный запрос, который не получает доступ в систему в момент поступления или обслуживание которого прервано, имеет возможность покинуть систему навсегда или пойти на так называемую в литературе орбиту (некоторое виртуальное место) и повторить попытку получить доступ к ресурсам системы через случайный временной промежуток. Считается, что вторичные запросы могут быть ненастойчивыми (могут покидать систему после неудачной попытки) и (или) нетерпеливыми (могут покидать систему через некоторое время пребывания на орбите).

1. Математическая модель

Рассмотрим систему массового обслуживания без буфера, с двумя типами запросов и повторными вызовами, состоящую из N приборов (рисунок).



Структура системы

В систему поступают два MAP-потока запросов. Запросы первого типа (первичных пользователей) поступают в соответствии с потоком, обозначаемым как MAP_1 и заданным неприводимой цепью Маркова $w_t,\ t \geq 0,$ с непрерывным временем и конечным пространством состояний $\{0,1,...,W\}$. Время пребывания цепи в состоянии w экспоненциально распределено с положительным параметром λ_w. Когда время пребывания в состоянии w истекло, с вероятностью $p^{(k)}_{w,w'}$ процесс w_t переходит в состояние w' и при этом генерируется k запросов, $k = 0,1;\ w, w' = \overline{0,W}$. Поведение MAP-потока полностью характеризуется матрицами $D_k,\ k = 0, 1,$ элементы которых определяются следующим образом:

$$(D_k)_{w,w'} = \lambda_w p^{(k)}_{w,w'},\ k = 1,\quad k = 0,\ w \neq w';$$
$$(D_0)_{w,w} = -\lambda_w,\ w = \overline{0,W}.$$

Матрица $D(1) = D_0 + D_1$ представляет собой инфинитезимальный генератор цепи $w_t,\ t \geq 0$.

Средняя интенсивность поступления запросов первого типа λ_1 имеет вид

$$\lambda_1 = \boldsymbol{\theta} D_1 \boldsymbol{e},$$

где $\boldsymbol{\theta}$ – вектор стационарного распределения цепи Маркова $w_t,\ t \geq 0$. Вектор $\boldsymbol{\theta}$ является единственным решением системы линейных алгебраических уравнений

$$\boldsymbol{\theta} D(1) = \mathbf{0},\ \boldsymbol{\theta}\boldsymbol{e} = 1.$$

Здесь и далее $\boldsymbol{e}$ – вектор-столбец, состоящий из единиц; $\mathbf{0}$ – вектор-строка, состоящая из нулей.

Коэффициент вариации c_{var} длин интервалов между моментами поступления запросов первого типа определяется формулой

$$c^2_{var} = 2\lambda_1 \boldsymbol{\theta}(-D_0)^{-1}\boldsymbol{e} - 1.$$

Коэффициент корреляции c_{cor} длин двух соседних интервалов между поступлением запросов первого типа вычисляется следующим образом:

$$c_{cor} = (\lambda_1 \boldsymbol{\theta}(-D_0)^{-1}(D(1) - D_0)(-D_0)^{-1}\boldsymbol{e} - 1) / c^2_{var}.$$

Запросы второго типа (вторичных пользователей) поступают в систему в соответствии с марковским входным потоком MAP_2, который управляется случайным процессом $v_t, t \geq 0,$ с конечным пространством состояний $\{0,1,\ldots,V\}$ и описывается квадратными матрицами H_0 и H_1 размером $\bar{V} = V + 1$. Средняя интенсивность λ_2 поступления запросов второго типа вычисляется как $\lambda_2 = \boldsymbol{\chi} H_1 \boldsymbol{e},$ где $\boldsymbol{\chi}$ – вектор стационарного распределения процесса $v_t,\ t \geq 0,$ – определяется как единственное решение системы линейных алгебраических уравнений

$$\boldsymbol{\chi}(H_0 + H_1) = \mathbf{0},\ \boldsymbol{\chi}\boldsymbol{e} = 1.$$

Запросы первого типа считаются приоритетными, а запросы второго типа – неприоритетными. Полагаем, что запросы второго типа принимаются на обслуживание, если число занятых каналов меньше чем $M, N \geq M > 0$. Если в момент прихода запроса второго типа число занятых приборов больше либо равно M, то запрос идет на орбиту с вероятностью q_1, а с дополнительной вероятностью покидает систему.

Если в момент прибытия запроса первого типа есть свободный прибор, то этот запрос принимается в систему. Если в момент прибытия запроса первого типа все приборы заняты, но на обслуживании есть запрос второго типа, то этот приоритетный запрос «выбивает» неприоритетный запрос с обслуживания, т. е. приоритетный запрос начинает обслуживание, а выбитый неприоритетный запрос идет на орбиту с вероятностью q_2 или покидает систему с дополнительной вероятностью. Если в момент прибытия запроса первого типа все приборы заняты и на обслуживании нет запросов второго типа, то пришедший запрос покидает систему.

Запросы, находящиеся на орбите, совершают повторные попытки попасть на обслуживание через экспоненциально распределенное с параметром α, $\alpha > 0$, время. В случае если попытка попасть на обслуживание осуществляется, когда число занятых приборов больше либо равно M, то запрос возвращается на орбиту с вероятностью q_1, а с дополнительной вероятностью покидает систему, т. е. запросы с орбиты могут проявлять ненастойчивость.

Запросы с орбиты также могут проявлять нетерпеливость. Считаем, что запрос покидает орбиту через экспоненциально распределенное время с параметром γ, $\gamma > 0$, с момента попадания на нее.

Время обслуживания запросов первого и второго типов имеет экспоненциальное распределение с параметрами μ_1 и μ_2 соответственно.

Условие, что запросы второго типа принимаются в систему, когда число занятых каналов меньше чем $M, M > 0$, по сути, означает, что для приоритетных запросов зарезервировано $N - M$ приборов. Это может показаться бессмысленным, так как приоритетные запросы могут выбивать неприоритетные с обслуживания и не нуждаются в резервировании каналов. На самом деле резервирование имеет смысл для улучшения качества обслуживания запросов второго типа. Резервирование может помочь избежать частых выбиваний запросов второго типа с обслуживания и тем самым увеличить пропускную способность системы. В случае частых выбиваний ресурсы системы на обслуживание выбитых запросов второго типа теряются впустую, а резервирование приборов может помочь использовать эти ресурсы оптимальным образом.

2. Процесс изменения состояний системы и стационарное распределение числа запросов в системе

Введем в рассмотрение следующие процессы:

i_t, $i_t \geq 0$, – число запросов на орбите;

$n_t, n_t = \overline{0, N}$, – число занятых приборов;

l_t, $l_t = \overline{0, \min\{n_t, M\}}$, – число запросов второго типа на обслуживании;

w_t, $w_t = \overline{0, W}$, – состояние управляющего процесса MAP_1;

v_t, $v_t = \overline{0, V}$, – состояние управляющего процесса MAP_2

в момент времени $t, t \geq 0$.

Процесс $\xi_t = \{i_t, n_t, l_t, w_t, v_t\}$, $t \geq 0$, является неприводимой цепью Маркова с непрерывным временем.

Примем следующие обозначения:

I – единичная матрица, O – нулевая матрица соответствующего размера;

$\oplus$ и $\otimes$ – символы кронекеровой суммы и кронекерова произведения матриц (см., например, [5]);

$\bar{W} = W + 1$;

$C_l = \mathrm{diag}\{0, 1, \ldots, l\}$, $l = \overline{0, M}$;

$\bar{C}_l = \mathrm{diag}\{l,\ l-1, \ldots, 0\}, l = \overline{0, M}$;

$\tilde{C}_l = \mathrm{diag}\{l,\ l-1,\dots,l-M+1,l-M\},\ l=\overline{M,N};$

$\mathrm{diag}\{A_1,\dots,A_l\}$ – блочно-диагональная матрица с диагональными блоками $A_1,\dots,A_l$;

$E_l^+,\ \tilde{E}_l^+,\ l=\overline{0,M-1},$ – матрицы размером $(l+1)\times(l+2),$ которые определяются как

$$E_l^+ = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix},\ \tilde{E}_l^+ = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix};$$

$E_l^-,\tilde{E}_l^-,\ l=\overline{1,M},$ – матрицы размером $(l+1)\times l,$ которые определяются как

$$E_l^- = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix},\ \tilde{E}_l^- = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix};$$

$\tilde{I}$ – диагональная матрица размером $(M+1)(N-M/2+1)$ с диагональными элементами $\{\underbrace{0,\dots,0}_{M(M+1)/2},1,\dots,1\};$

$E^-,\ \tilde{I}$ – квадратные матрицы размером $M+1,$ которые имеют вид

$$E^- = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},\ \tilde{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Перенумеруем состояния цепи ξ_t в лексикографическом порядке компонент $(i,n,\ l,w,v)$. Множество состояний, имеющих значение $(i,\ n)$ двух первых компонент цепи, будем называть макросостоянием $(i,\ n)$.

Пусть A – генератор цепи Маркова $\xi_t,\ t\ge 0,$ который сформирован из блоков $A_{i,j}$, состоящих из матриц $(A_{i,j})_{n,n'}$ интенсивностей переходов цепи $\xi_t,\ t\ge 0,$ из макросостояния $(i,\ n)$ в макросостояние $(j,\ n'),\ n,\ n'=\overline{0,\min\{i,N\}}.$ Диагональные элементы матрицы $A_{i,i}$ отрицательны и их модули определяют интенсивность выхода из соответствующего состояния цепи Маркова.

Теорема 1. *Генератор A имеет следующую блочно-трехдиагональную структуру:*

$$A = \begin{pmatrix} A_{0,0} & A_{0,1} & O & O & \dots \\ A_{1,0} & A_{1,1} & A_{1,2} & O & \dots \\ O & A_{2,1} & A_{2,2} & A_{2,3} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Ненулевые блоки $A_{i,j},\ i,\ j \geq 0,$ *имеют вид*

$$A_{i,i} = \begin{pmatrix} Q_i^{(0)} & B^{(0)} & O & \dots & O & O \\ F^{(1)} & Q_i^{(1)} & B^{(1)} & \dots & O & O \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ O & O & O & \ddots & Q_i^{(N-1)} & B^{(N-1)} \\ O & O & O & \dots & F^{(N)} & Q_i^{(N)} \end{pmatrix} +$$

$$+(1-q_1)\tilde{I} \otimes I_{\bar{W}} \otimes H_1 + I_{(M+1)(N-M/2+1)} \otimes (D_0 \oplus H_0),\ i \geq 0,$$

$$A_{i,i+1} = A^{+} = \mathrm{diag}\{H^{(0)},\dots,H^{(N)}\},\ i \geq 0,$$

$$A_{i,i-1} = \begin{pmatrix} L_i^{(0)} & \tilde{B}_i^{(0)} & O & \dots & O & O \\ O & L_i^{(1)} & \tilde{B}_i^{(1)} & \dots & O & O \\ O & O & L_i^{(2)} & \dots & O & O \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ O & O & O & \ddots & L_i^{(N-1)} & \tilde{B}_i^{(N-1)} \\ O & O & O & \dots & O & L_i^{(N)} \end{pmatrix},\ i \geq 1,$$

где

$$Q_i^{(n)} = \begin{cases} -(\mu_2 C_n + \mu_1 \bar{C}_n + i(\alpha+\gamma) I_{n+1}) \otimes I_{\bar{W}\bar{V}},\ n < M, \\ -(\mu_2 C_M + \mu_1 \tilde{C}_n + i((1-q_1)\alpha+\gamma) I_{M+1}) \otimes I_{\bar{W}\bar{V}},\ M \leq n < N, \\ -(\mu_2 C_M + \mu_1 \tilde{C}_n + i((1-q_1)\alpha+\gamma) I_{M+1}) \otimes I_{\bar{W}\bar{V}} + \\ +((1-q_2)E^{-} + \tilde{I}) \otimes D_1 \otimes I_{\bar{V}},\ n = N, \end{cases} \qquad i \geq 0;$$

$$B^{(n)} = \begin{cases} E_n^{+} \otimes I_{\bar{W}} \otimes H_1 + \tilde{E}_n^{+} \otimes D_1 \otimes I_{\bar{V}},\ n < M, \\ I_{M+1} \otimes D_1 \otimes I_{\bar{V}},\ M \leq n < N; \end{cases}$$

$$F^{(n)} = \begin{cases} (\mu_1 \bar{C}_n E_n^{-} + \mu_2 C_n \tilde{E}_n^{-}) \otimes I_{\bar{W}\bar{V}},\ n \leq M, \\ (\mu_1 \tilde{C}_n + \mu_2 C_M E^{-}) \otimes I_{\bar{W}\bar{V}},\ M < n \leq N; \end{cases}$$

$$H^{(n)} = \begin{cases} O,\ n < M, \\ q_1 I_{(M+1)\bar{W}} \otimes H_1,\ M \leq n < N, \\ q_2 E^{-} \otimes D_1 \otimes I_{\bar{V}} + q_1 I_{M+1} \otimes I_{\bar{W}} \otimes H_1,\ n = N; \end{cases}$$

$$\tilde{B}_i^{(n)} = \begin{cases} i\alpha E_n^{+} \otimes I_{\bar{W}\bar{V}},\ n < M, \\ O,\ n \geq M, \end{cases} \qquad i \geq 1;$$

$$L_i^{(n)} = \begin{cases} i\gamma I_{(n+1)\bar{W}\bar{V}},\ n < M, \\ i(\gamma + (1-q_1)\alpha) I_{(M+1)\bar{W}\bar{V}},\ M \leq n \leq N, \end{cases} \qquad i \geq 1.$$

Замечание. Цепь Маркова $\xi_t,\ t \geq 0,$ принадлежит классу асимптотически квазитеплицевых цепей Маркова (АКТЦМ) с непрерывным временем [6].

Как следует из [6], необходимым условием существования стационарного распределения АКТЦМ $\xi_t,\ t \geq 0,$ является условие

$$\boldsymbol{y}Y_0\boldsymbol{e} > \boldsymbol{y}Y_2\boldsymbol{e}, \qquad (1)$$

где вектор-строка $\boldsymbol{y}$ – единственное решение системы линейных алгебраических уравнений

$$\boldsymbol{y}(Y_0 + Y_1 + Y_2) = \boldsymbol{y},\ \boldsymbol{ye} = 1, \qquad (2)$$

матрицы $Y_0,\ Y_1$ и Y_2 определяются выражениями

$$Y_0 = \lim_{t\to\infty} R_i^{-1} A_{i,i-1},\ Y_1 = \lim_{t\to\infty} R_i^{-1} A_{i,i} + I,\ Y_2 = \lim_{t\to\infty} R_i^{-1} A_{i,i+1},$$

а R_i – диагональная матрица, диагональные элементы которой определяются как модули соответствующих диагональных элементов матрицы $A_{i,i},\ i \geq 0.$

Легко убедиться, что матрица R_i является блочно-диагональной матрицей с диагональными блоками $T_i^{(n)},\ n = \overline{0,N},\ i \geq 0,$ которые определяются следующим образом:

$$T_i^{(n)} = \begin{cases} -A_i^{(n)} + I_{n+1} \otimes (\Sigma_1 \oplus \Lambda_1),\ n = \overline{0, M-1}; \\ -A_i^{(n)} - (1-q_1) I_{M+1} \otimes I_{\bar{W}} \otimes \Lambda_2 + I_{M+1} \otimes (\Sigma_1 \oplus \Lambda_1),\ n = \overline{M, N-1}; \\ (\mu_1 \tilde{C}_N + \mu_2 C_M + i((1-q_1)\alpha + \gamma)) \otimes I_{\bar{W}\bar{V}} - \\ -\tilde{I} \otimes \Sigma_2 \otimes I_{\bar{V}} - (1-q_1) I_{(M+1)\bar{W}} \otimes \Lambda_2 + I_{M+1} \otimes (\Sigma_1 \oplus \Lambda_1),\ n = N, \end{cases} \qquad i \geq 0,$$

где $\Sigma_1,\ \Sigma_2,\ \Lambda_1$ и Λ_2 – диагональные матрицы, диагональные элементы которых являются диагональными элементами матриц $-D_0,\ D_1,\ -H_0$ и H_1 соответственно.

В рассматриваемом случае матрицы $Y_0,\ Y_1$ и Y_2 определяются как

$$Y_0 = \begin{pmatrix} \frac{\gamma}{\gamma+\alpha} & \frac{\alpha}{\gamma+\alpha}E_0^+ & \cdots & O & O & O & \cdots & O \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \frac{\gamma}{\gamma+\alpha}I_M & \frac{\alpha}{\gamma+\alpha}E_M^+ & O & \cdots & O \\ O & O & \cdots & O & I_{M+1} & O & \cdots & O \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & O & O & O & \cdots & O \\ O & O & \cdots & O & O & O & \cdots & I_{M+1} \end{pmatrix} \otimes I_{\bar{W}\bar{V}},\quad Y_1 = O,\ Y_2 = O.$$

Очевидно, что при этом условие (1) может быть переписано в виде $\boldsymbol{y}Y_0\boldsymbol{e} > 0,$ а система (2) – в виде $\boldsymbol{y}Y_0 = \boldsymbol{y},\ \boldsymbol{ye} = 1.$ Исходя из этого, получаем неравенство

$$\boldsymbol{y}Y_0\boldsymbol{e} = \boldsymbol{ye} = 1 > 0,$$

т. е. условие (1) выполняется для любых параметров системы.

Поскольку условие эргодичности системы выполнено для любых параметров системы, существуют следующие пределы (стационарные вероятности):

$$p(i, n, l, w, v) = \lim_{t\to\infty} P\{i_t = i,\ n_t = n,\ l_t = l,\ w_t = w,\ v_t = v\},\ i \geq 0,\ n = \overline{0, \min\{i, N\}},$$

$$l = \overline{0, \min\{n, M\}},\ \ w = \overline{0, W},\ \ v = \overline{0, V}.$$

Сформируем вектор-строки $\boldsymbol{p}_i$:

$$\boldsymbol{p}(i,n,l,w) = (p(i,n,l,w,0), p(i,n,l,w,1),\ldots, p(i,n,l,w,V)),\quad w = \overline{0,W};$$

$$\boldsymbol{p}(i,n,l) = (\boldsymbol{p}(i,n,l,0), \boldsymbol{p}(i,n,l,1),\ldots, \boldsymbol{p}(i,n,l,W)),\quad l = \overline{0,\min\{n,M\}};$$

$$\boldsymbol{p}(i,n) = (\boldsymbol{p}(i,n,0), \boldsymbol{p}(i,n,1),\ldots, \boldsymbol{p}(i,n,\min\{n,M\})),\quad n = \overline{0,\min\{i,N\}};$$

$$\boldsymbol{p}_i = (\boldsymbol{p}(i,0), \boldsymbol{p}(i,1),\ldots, \boldsymbol{p}(i,\min\{i,N\})),\quad i \geq 0.$$

Известно, что векторы $\boldsymbol{p}_i$, $i \geq 0$, удовлетворяют следующей системе линейных алгебраических уравнений:

$$(\boldsymbol{p}_0, \boldsymbol{p}_1,\ldots)A = \boldsymbol{0},\ (\boldsymbol{p}_0, \boldsymbol{p}_1,\ldots)\boldsymbol{e} = 1, \tag{3}$$

где A – инфинитезимальный генератор цепи Маркова ξ_t, $t \geq 0$.

С учетом того что генератор A в данной модели имеет блочно-трехдиагональную форму, для решения системы (3) может быть использован численно устойчивый алгоритм, приведенный в работе [7].

3. Характеристики производительности

Найдя векторы стационарных вероятностей $\boldsymbol{p}_i$, $i \geq 0$, можно вычислить различные характеристики производительности системы:

– распределение вероятностей числа запросов на орбите

$$\lim_{t\to\infty} P\{i_t = i\} = \boldsymbol{p}_i\boldsymbol{e},\ i \geq 0;$$

– среднее число запросов в системе

$$L = \sum_{i=0}^{\infty}\sum_{n=0}^{N}(i+n)\boldsymbol{p}_i\boldsymbol{e};$$

– среднее число запросов на орбите

$$L_{orbit} = \sum_{i=1}^{\infty} i\boldsymbol{p}_i\boldsymbol{e};$$

– среднее число занятых приборов

$$N_{server} = \sum_{i=0}^{\infty}\sum_{n=1}^{N} n\boldsymbol{p}(i,n)\boldsymbol{e};$$

– среднее число занятых приборов, обслуживающих запросы первого типа,

$$N_{server-1} = \sum_{i=0}^{\infty}\sum_{n=1}^{N}\sum_{l=0}^{\min\{n,M\}}(n-l)\,\boldsymbol{p}(i,n,l)\boldsymbol{e};$$

– среднее число занятых приборов, обслуживающих запросы второго типа,

$$N_{server-2} = \sum_{i=0}^{\infty}\sum_{n=1}^{N}\sum_{l=1}^{\min\{n,M\}} l\boldsymbol{p}(i,n,l)\boldsymbol{e} = N_{server} - N_{server-1};$$

– среднюю интенсивность выходящего потока обслуженных запросов первого типа

$$\lambda_{out}^{(1)} = \mu_1 N_{server-1};$$

– среднюю интенсивность выходящего потока обслуженных запросов второго типа

$$\lambda_{out}^{(2)} = \mu_2 N_{server-2};$$

– среднюю интенсивность выходящего потока обслуженных запросов

$$\lambda_{out} = \lambda_{out}^{(1)} + \lambda_{out}^{(2)}.\frac{1}{n};$$

– вероятность потери запроса первого типа

$$P_1^{(loss)} = \lambda_1^{-1}\sum_{i=0}^{\infty}\boldsymbol{p}(i,N,0)(D_1 \otimes I_{\bar{V}})\boldsymbol{e} = 1 - \frac{\lambda_{out}^{(1)}}{\lambda_1};$$

– вероятность потери запроса второго типа

$$P_2^{(loss)} = 1 - \frac{\lambda_{out}^{(2)}}{\lambda_2};$$

– вероятность потери произвольного запроса

$$P^{(loss)} = 1 - \frac{\lambda_{out}}{\lambda_1 + \lambda_2};$$

– вероятность потери запроса второго типа на входе в систему из-за занятости более $M-1$ приборов

$$P^{(ent-loss)} = (1-q_1)\lambda_2^{-1}\sum_{i=0}^{\infty}\sum_{n=M}^{N}\boldsymbol{p}(i,n)(I_{(M+1)\bar{W}} \otimes H_1)\boldsymbol{e};$$

– вероятность того, что запрос второго типа поступит в систему, когда число занятых приборов больше $M-1$, и пойдет на орбиту,

$$P^{(ent-to-orbit)} = q_1\lambda_2^{-1}\sum_{i=0}^{\infty}\sum_{n=M}^{N}\boldsymbol{p}(i,n)(I_{(M+1)\bar{W}} \otimes H_1)\boldsymbol{e};$$

– вероятность того, что запрос второго типа будет выбит запросом первого типа и покинет систему,

$$P^{(knock-out-loss)} = (1-q_2)\lambda_2^{-1}\sum_{i=0}^{\infty}\sum_{l=1}^{M}\boldsymbol{p}(i,N,l)(D_1 \otimes I_{\bar{V}})\boldsymbol{e};$$

– вероятность потери запроса с орбиты

$$P^{(loss-from-orbit)} = P_2^{(loss)} - P^{(ent-loss)} - P^{(knock-out-loss)};$$

– вероятность того, что запрос с орбиты совершит повторную попытку, когда число занятых приборов больше $M-1$, и вернется на орбиту,

$$P^{(return-to-orbit)} = \frac{q_1}{\alpha L_{orbit}}\sum_{i=1}^{\infty}\sum_{n=M}^{N} i\alpha\boldsymbol{p}(i,n)\boldsymbol{e}.$$

Заключение

В статье изучена многолинейная система массового обслуживания с повторными вызовами и запросами разных приоритетов. Построен процесс изменения состояний системы, найдено условие его эргодичности, приведены формулы для нахождения основных характеристик производительности системы. Результаты исследования могут применяться для оценивания производительности и оптимизации функционирования систем когнитивного радио.

Данная работа выполнена при частичной поддержке Белорусского республиканского фонда фундаментальных исследований Ф15М-005.

Список литературы

1. Feasibility analysis of vehicular dynamic spectrum access via queueing theory model / S. Chen [et al.] // IEEE Communications Magazine. – 2011. – Vol. 49. – P. 156–163.

2. Next generation dynamic spectrum access cognitive radio wireless networks: A survey / I.F. Akyildiz [et al.] // Computer Networks. – 2006. – Vol. 50(13). – P. 2127–2159.

3. Performance analysis of dynamic spectrum handoff scheme with variable bandwidth demand on secondary users for cognitive radio networks / Y. Konishi [et al.] // Wireless Networks. – 2013. – Vol. 19. – P. 607–617.

4. Zahed, S. Analytical modeling for spectrum handoff decision in cognitive radio networks / S. Zahed, I. Awan, A. Cullen // Simulation modelling practice and theory. – 2013. – Vol. 38. – P. 98–114.

5. Graham, A. Kronecker Products and Matrix Calculus with Applications / A. Graham. – Cichester : Ellis Horwood, 1981. – 130 p.

6. Klimenok, V.I. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory / V.I. Klimenok, A.N. Dudin // Queueing Systems. – 2006. – Vol. 54. – P. 245–259.

7. Dudina, O. Retrial queueing system with Markovian arrival flow and phase-type service time distribution / O. Dudina, Ch. Kim, S. Dudin // Computers and Industrial Engineering. – 2013. – Vol. 66. – P. 360–373.

Поступила 12.05.2015

Белорусский государственный университет,
Минск, пр. Независимости, 4
e-mail: dudin85@mail.ru,
dudina_olga@email.com

S.A. Dudin, O.S. Dudina

ANALYSIS OF MULTI-SERVER QUEUEING SYSTEM WITH PREEMPTIVE PRIORITY AND REPEATED CALLS

Multi-server retrial queueing system with no buffer and two types of customers is analyzed as the model of cognitive radio system. Customers of type 1 have a preemptive priority. Customers of both types arrive according to Markovian Arrival Processes. Service times have exponential distribution with parameter depending on the customer type. Type 2 customers are admitted for service only if the number of busy servers is less than the predefined threshold. The rejected type 2 customers retry for the service. Existence condition of the stationary mode of system operation is derived. Formulas for computing key performance measures of the system are presented.