ИНФОРМАТИКА

2016 апрель-июнь № 2

УДК 519.872

В.И. Клименок

СТАЦИОНАРНЫЕ ХАРАКТЕРИСТИКИ НЕНАДЕЖНОЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ГРУППОВЫМ МАРКОВСКИМ ПОТОКОМ И РЕЗЕРВНЫМ ПРИБОРОМ

Исследуется система массового обслуживания с ненадежным прибором и холодным резервированием, состоящая из бесконечного буфера, основного ненадежного прибора и резервного надежного прибора. Термин «холодный резерв» подразумевает, что резервный прибор выключен, пока исправен основной, и включается только при выходе основного прибора из строя. В систему поступает групповой марковский поток запросов. Поломки на основной прибор поступают в марковском потоке. Времена обслуживания на приборах и времена ремонтов распределены по фазовому закону. Выводится условие существования стационарного режима. Приводятся формулы для вычисления основных характеристик производительности системы. Находится преобразование Лапласа – Стилтьеса распределения времени пребывания запроса в системе.

Введение

В последние годы технологии атмосферных оптических линий связи, с помощью которых информация передается по оптическому лазерному каналу, стали широко распространенными благодаря их несомненным достоинствам. Основные преимущества оптического лазерного канала заключаются в том, что он является высокоскоростным, не создает радиопомех и обеспечивает высокую конфиденциальность передачи информации. Однако этот канал очень чувствителен к состоянию атмосферы. Он не может передавать данные в условиях плохой видимости – тумана или пасмурной погоды. Как отмечено в [1], одним из важных направлений создания высокоскоростных надежных сетей связи является развитие гибридных сетей связи, основанных на лазерных и радиотехнологиях.

В гибридной системе оптический лазерный канал резервируется радиоканалом, функционирующим под управлением протокола IEEE 802.11n, и (или) радиоканалом миллиметрового диапазона радиоволн. Радиоканал миллиметрового диапазона является высокоскоростным, но он также чувствителен к состоянию атмосферы и не может осуществлять передачу во время осадков (дождь, снег и т. д.). Беспроводной широкополосный радиоканал, функционирующий под управлением протокола IEEE 802.11n, мало чувствителен к состоянию атмосферы, но имеет существенно меньшую скорость передачи. Поэтому и возникает проблема построения гибридных систем, использующих совместно или поочередно различные технологии передачи.

Вследствие высокой практической важности гибридных сетей связи в последние годы в литературе появился ряд публикаций, посвященных исследованию таких систем. Некоторые результаты этих исследований представлены в [1–4]. Работа [2] в основном посвящена анализу характеристик надежности, методам и алгоритмам выбора оптимальных стратегий переключения каналов в гибридных системах путем имитационного моделирования. В [1, 3–4] используется аналитическое моделирование гибридных систем связи системами массового обслуживания. В статье [3] рассматривается гибридная система связи с так называемым горячим резервированием, при котором резервный канал функционирует параллельно с основным и при выходе основного прибора из строя продолжает обслуживание. Достоинства горячего резервирования – быстрая реакция на выход основного прибора из строя и более высокая скорость обслуживания. В то же время следует отметить и его основной недостаток – неэкономное использование ресурсов, например электроэнергии, который отсутствует при холодном резервировании. В работе [4] рассматривается гибридная система с холодным резервированием, в которой резервный надежный радиоканал подключается к обслуживанию только тогда, когда основной лазерный канал выходит из строя вследствие неблагоприятных погодных условий. Статья [1] посвящена гибридной системе связи, в которой миллиметровый канал используется как резервный по отношению к оптическому лазерному каналу. В качестве математической модели авторы рассматривают систему массового обслуживания с двумя ненадежными приборами.

В настоящей работе исследуется система массового обслуживания, предназначенная для моделирования гибридной сети связи с холодным резервированием, при гораздо более общих по сравнению с [4] предположениях относительно характера входных потоков и потоков поломок и распределений времен обслуживания и восстановления.

1. Математическая модель

Рассматривается система массового обслуживания с ожиданием, состоящая из двух приборов, один из которых (основной, прибор 1) является ненадежным, а другой (резервный, прибор 2) – абсолютно надежным. Последний находится в холодном резерве. Ненадежный прибор – это лазерный канал, а надежный – беспроводной широкополосный радиоканал, функционирующий под управлением протокола IEEE 802.11n. Под влиянием смога, тумана, снега лазерный канал может выйти из строя и сразу начать восстанавливаться. Одновременно начнется период переключения на резервный прибор с целью дальнейшей передачи информации по этому каналу. Во время восстановления (ремонта) информация будет передаваться по резервному каналу. После восстановления основного канала резервный канал отключится до следующей поломки основного канала.

Опишем более подробно сценарий взаимодействия основного и резервного приборов. Обслуживание заявок производится основным прибором. Если произошла поломка основного прибора, то одновременно запускаются механизмы ремонта этого прибора и переключения на резервный прибор. В случае когда ремонт закончится раньше, чем само переключение, переключение прекращается и основной прибор возобновляет обслуживание, прерванное поломкой, или берет первую поступившую заявку из очереди за время ремонта. Если прибор был пуст как в момент поломки, так и в момент окончания ремонта, то он ждет поступления новой заявки. В случае когда переключение закончилось раньше, чем ремонт, резервный прибор еще раз заново обслуживает заявку, а после – только заявки из буфера и вновь поступающие заявки до тех пор, пока не закончится ремонт основного прибора. С момента окончания ремонта обслуживание заявок немедленно переносится на основной прибор. Обслуживание заявки, которая находилась на резервном приборе (если таковая имелась), начинается на основном приборе заново.

Запросы поступают в систему в групповом марковском потоке (Batch Markovian Arrival Process – BMAP). Поступление групп запросов в этом потоке возможно только в моменты скачков некоторой неприводимой цепи Маркова $\nu_t,\, t\ge 0,$ с непрерывным временем и конечным пространством состояний $\{0,1,\ldots,W\}$, которая называется управляющим процессом BMAP-потока. Интенсивности переходов процесса $\nu_t,\, t\ge 0$, сопровождающиеся поступлением групп запросов размером $k\ge 1$ (не сопровождающиеся поступлением запросов), задаются элементами матриц D_k (недиагональными элементами матрицы D_0) или производящей функцией $D(z)=\sum_{k=0}^{\infty} D_k z^k,\ |z|\le 1.$ При этом матрица $D(1)$ является неприводимым генератором цепи $\nu_t,\ t\ge 0$. Вектор-строка $\mathbf{q}$ стационарного распределения данного процесса является решением системы уравнений $\mathbf{q}D(1)=0,\ \mathbf{qe}=1$, где $\mathbf{e}$ – вектор-столбец, состоящий из единиц. Интенсивность λ поступления запросов в BMAP-потоке имеет вид $\lambda=\mathbf{q}D'(z)|_{z=1}\mathbf{e}$, интенсивность λ_b поступления групп запросов определяется как $\lambda_b=\mathbf{q}(-D_0)\mathbf{e}$. Коэффициент вариации c_{var} длин интервалов между моментами поступления последовательных групп определяется формулой $c_{var}^2=2\lambda_b\mathbf{q}(-D_0)^{-1}\mathbf{e}-1$. Коэффициент корреляции c_{cor} длин двух соседних интервалов вычисляется следующим образом: $c_{cor}=[\lambda_b\mathbf{q}(-D_0)^{-1}(D(1)-D_0(-D_0)^{-1}\mathbf{e}-1)]/c_{var}^2$. Более подробное описание BMAP-потока можно найти, например, в [5].

Поломки поступают на основной прибор в любом его состоянии. Если прибор находится на ремонте, то поступившая поломка игнорируется. В противном случае поломка вызывает начало периода переключения на резервный прибор. Поломки поступают в MAP-потоке, характеризующимся матрицами H_0 и H_1 порядка $(V+1)\times(V+1)$. Стационарный вектор MAP определяется как решение системы $\mathbf{g}(H_0+H_1)=0,\ \mathbf{ge}=1.$ Интенсивность поломок $h=\mathbf{g}H_1\mathbf{e}$.

Времена обслуживания на приборах, времена ремонтов и переключений имеют распределения фазового типа. Краткое описание этого типа распределений следующее. Время, имеющее фазовое распределение (Phase Type distribution – PH) с неприводимым представлением $(\mathbf{b},S)$, определяется как время, за которое цепь Маркова с непрерывным временем $m_t,\ t\ge 0,$ имеющая несущественные состояния $(1,\dots,M)$ и поглощающее состояние $M+1$, достигнет поглощающего состояния. Начальное состояние цепи выбирается в соответствии с распределением, задаваемым вектор-строкой $(\mathbf{b},0)$, где $\mathbf{b}$ – стохастический вектор размерности M. Переходы цепи Маркова $m_t,\ t\ge 0,$ описываются генератором $\begin{pmatrix} S & \mathbf{S}_0 \\ \mathbf{0} & 0 \end{pmatrix}$, где матрица S – субгенератор, а вектор-столбец $\mathbf{S}_0$ определяется как $\mathbf{S}_0=-S\mathbf{e}$. Матрица S является квадратной матрицей порядка M с отрицательными диагональными элементами и неотрицательными недиагональными элементами. Вектор $\mathbf{S}_0$ является вектором с неотрицательными элементами, среди которых есть хотя бы один положительный. Математическое ожидание рассматриваемого интервала времени вычисляется по формуле $b_1=\mathbf{b}(-S)^{-1}\mathbf{e},$ интенсивность обслуживания – $\mu=b_1^{-1}.$ Более подробную информацию о распределениях фазового типа можно найти, например, в [6, 7].

В нашем случае предполагаем, что время обслуживания запроса k-м прибором имеет PH-распределение с неприводимым представлением $(\mathbf{b}^{(k)},\ S^{(k)}),\ k=1,2.$ Процесс обслуживания на k-м приборе происходит под управлением цепи Маркова $m_t^{(k)},\ \ t\ge 0,$ с пространством состояний $\{1,\dots,M^{(k)},\ M^{(k)}+1\},$ где $M^{(k)}+1$ есть поглощающее состояние. Интенсивности переходов в поглощающее состояние задаются вектором $\mathbf{S}_0^{(k)}=-S^{(k)}\mathbf{e}$. Интенсивности обслуживания вычисляются как $\mu^{(k)}=-[\mathbf{b}^{(k)}(S^{(k)})^{-1}\mathbf{e}]^{-1},\ k=1,2.$

Время ремонта имеет PH-распределение с неприводимым представлением $(\mathbf{t},T)$. Процесс ремонта происходит под управлением цепи Маркова $\vartheta_t,\ \ t\ge 0,$ с пространством состояний $\{1,\dots,R,\ R+1\},$ где $R+1$ есть поглощающее состояние. Интенсивности переходов в поглощающее состояние задаются вектором $\mathbf{T}_0=-T\mathbf{e}$. Интенсивность ремонта вычисляется как $\phi=-(\mathbf{t}T^{-1}\mathbf{e})^{-1}.$

Время переключения на резервный прибор имеет PH-распределение с неприводимым представлением $(\mathbf{a},A)$. Процесс переключения происходит под управлением цепи Маркова $l_t,t\ge 0,$ с пространством состояний $\{1,\dots,L,L+1\},$ где $L+1$ есть поглощающее состояние. Интенсивности переходов в поглощающее состояние задаются вектором $\mathbf{A}_0=-A\mathbf{e}$. Интенсивность переключения вычисляется как $a=-[\mathbf{a}A^{-1}\mathbf{e}]^{-1}.$

2. Цепь Маркова, описывающая функционирование системы

Пусть в момент t:

i_t – число запросов в системе, $i_t\ge 0$;

$$n_t=\begin{cases}0, & \text{если основной прибор (прибор 1) исправен;}\\ 1, & \text{если основной прибор на ремонте;}\end{cases}$$

$$r_t=\begin{cases}0, & \text{если момент } t \text{ не принадлежит периоду переключения;}\\ 1, & \text{если в момент } t \text{ идет переключение;}\end{cases}$$

$m_t^{(j)}$ – состояние управляющего процесса PH-времени обслуживания на j-м занятом приборе, $j=1,2, m_t^{(j)}=\overline{1,M^{(j)}}$;

l_t – состояние управляющего процесса PH-времени переключения с прибора 1 на прибор 2, $l_t = \overline{1,L}$;

ϑ_t – состояние управляющего процесса PH-времени ремонта, $\vartheta_t = \overline{1,R}$;

ν_t и η_t – состояния управляющих процессов входящего BMAP-потока и MAP-потока поломок соответственно, $\nu_t = \overline{0,W}$, $\eta_t = \overline{0,V}$.

Процесс функционирования системы описывается неприводимой цепью Маркова ξ_t, $t \geq 0$, с пространством состояний

$$X = \{(i,n,\nu,\eta),\ i=0,\ \ n=0,\ \ \nu=\overline{0,W},\ \eta=\overline{0,V}\}\bigcup$$

$$\bigcup\{(i,n,\nu,\eta,\vartheta),\ \ i=0,\ \ n=1,\ \ \nu=\overline{0,W},\ \ \eta=\overline{0,V},\ \vartheta=\overline{0,R}\}\bigcup$$

$$\bigcup\{(i,n,r,\nu,\eta,m^{(1)}),\ i>0,\ \ n=0,\ \ r=0,\ \ \nu=\overline{0,W},\ \eta=\overline{0,V},\ \ m^{(1)}=\overline{1,M^{(1)}}\}\bigcup$$

$$\bigcup\{(i,n,r,\nu,\eta,m^{(2)},\vartheta),\ i>0,\ \ n=1,\ \ r=0,\ \ \nu=\overline{0,W},\ \eta=\overline{0,V},\ \ \vartheta=\overline{1,R},\ \ m^{(2)}=\overline{1,M^{(2)}}\}\bigcup$$

$$\bigcup\{(i,n,r,\nu,\eta,\vartheta,l),\ \ i>0,\ \ n=1,\ \ r=1,\ \ \nu=\overline{0,W},\ \ \eta=\overline{0,V},\ \ \vartheta=\overline{1,R},\ \ l^{(2)}=\overline{1,L}\}.$$

Далее будем предполагать, что состояния цепи ξ_t внутри каждого из приведенных подмножеств упорядочены в лексикографическом порядке и таким образом упорядоченные подмножества упорядочены в том порядке, в котором они перечислены выше.

Теорема 1. *Инфинитезимальный генератор Q цепи Маркова ξ_t, $t \geq 0$, имеет следующую блочную структуру:*

$$Q = \begin{pmatrix} \tilde{Q}_0 & \tilde{Q}_1 & \tilde{Q}_2 & \tilde{Q}_3 & \cdots \\ \widehat{Q}_0 & Q_1 & Q_2 & Q_3 & \cdots \\ O & Q_0 & Q_1 & Q_2 & \cdots \\ O & O & Q_0 & Q_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

где ненулевые блоки имеют вид

$$\tilde{Q}_0 = \begin{pmatrix} D_0 \oplus H_0 & I_{\bar{W}} \otimes H_1 \otimes \mathbf{t} \\ I_a \otimes \mathbf{T}_0 & D_0 \oplus H \oplus T \end{pmatrix};$$

$$\tilde{Q}_k = \begin{pmatrix} 0 & D_k \otimes I_{\bar{V}} \otimes \mathbf{b}^{(1)} & O & O \\ 1 & O & D_k \otimes I_{\bar{V}} \otimes I_R \otimes \mathbf{b}^{(2)} & O \end{pmatrix},\quad k \geq 1;$$

$$\widehat{Q}_0 = \begin{pmatrix} I_a \otimes \mathbf{S}_0^{(1)} & O \\ O & I_{aR} \otimes \mathbf{S}_0^{(2)} \\ O & O \end{pmatrix},\quad Q_0 = \begin{pmatrix} I_a \otimes \mathbf{S}_0^{(1)}\mathbf{b}^{(1)} & O & O \\ O & I_{aR} \otimes \mathbf{S}_0^{(2)}\mathbf{b}^{(2)} & O \\ O & O & O \end{pmatrix};$$

$$Q_1 = \begin{pmatrix} D_0 \oplus H_0 \oplus S^{(1)} & O & I_{\bar{W}} \otimes H_1 \otimes \mathbf{e}_{M^{(1)}} \otimes \mathbf{t} \otimes \mathbf{a} \\ I_a \otimes \mathbf{T}_0 \otimes \mathbf{e}_{M^{(2)}} \otimes \mathbf{b}^{(1)} & D_0 \oplus H \oplus T \oplus S^{(2)} & O \\ I_a \otimes \mathbf{T}_0 \otimes \mathbf{e}_L \otimes \mathbf{b}^{(1)} & I_a \otimes I_R \otimes \mathbf{A}_0 \otimes \mathbf{b}^{(2)} & D_0 \oplus H \oplus T \oplus A \end{pmatrix},$$

$$Q_k = \mathrm{diag}\{D_{k-1} \otimes I_{\bar{V}} \otimes I_{M^{(1)}},\ D_{k-1} \otimes I_{\bar{V}} \otimes I_R \otimes I_{M^{(2)}},\ D_{k-1} \otimes I_{\bar{V}} \otimes I_R \otimes I_L\},\quad k \geq 2.$$

Здесь $H = H_0 + H_1$; $\otimes, \oplus$ *– символы кронекерова произведения и суммы матриц соответственно,* $\bar{W} = W + 1, \bar{V} = V + 1, a = \bar{W}\bar{V}$.

Доказательство теоремы 1 проводится путем анализа поведения цепи на бесконечно малом интервале времени.

Следствие 1. *Цепь Маркова* ξ_t, $t \geq 0$, *принадлежит классу квазитеплицевых цепей Маркова с непрерывным временем [8].*

Доказательство следует из вида генератора, заданного леммой 1, и определения квазитеплицевых цепей Маркова, данного в [8].

В дальнейшем будем использовать выражения для производящих функций $\tilde{Q}(z) = \sum_{k=1}^{\infty} \tilde{Q}_k z^k$ и $Q(z) = \sum_{k=0}^{\infty} Q_k z^k$, $|z| \leq 1$.

Следствие 2. *Матричные производящие функции* $\tilde{Q}(z)$, $Q(z)$ *имеют следующий вид:*

$$\tilde{Q}(z) = \begin{pmatrix} (D(z) - D_0) \otimes I_{\bar{V}} \otimes \mathbf{b}^{(1)} & O & O \\ O & (D(z) - D_0) \otimes I_{\bar{V}} \otimes I_R \otimes \mathbf{b}^{(2)} & O \end{pmatrix}; \tag{1}$$

$$Q(z) = Q_0 + Qz + z\mathrm{diag}\{D(z) \otimes I_{\bar{V}} \otimes I_{M^{(1)}},\ D(z) \otimes I_{\bar{V}} \otimes I_R \otimes I_{M^{(2)}},\ D(z) \otimes I_{\bar{V}} \otimes I_R \otimes I_L\}, \tag{2}$$

где

$$Q = \begin{pmatrix} I_{\bar{W}} \otimes H_0 \oplus S^{(1)} & O & I_{\bar{W}} \otimes H_1 \otimes \mathbf{e}_{M^{(1)}} \otimes \mathbf{t} \otimes \mathbf{a} \\ I_a \otimes \mathbf{T}_0 \otimes \mathbf{e}_{M^{(2)}} \otimes \mathbf{b}^{(1)} & I_{\bar{W}} \otimes H \oplus T \oplus S^{(2)} & O \\ I_a \otimes \mathbf{T}_0 \otimes \mathbf{e}_L \otimes \mathbf{b}^{(1)} & I_a \otimes I_R \otimes \mathbf{A}_0 \otimes \mathbf{b}^{(2)} & I_{\bar{W}} \otimes H \oplus T \oplus A \end{pmatrix}.$$

3. Условие эргодичности

Теорема 2. *Цепь Маркова* ξ_t, $t \geq 0$, *эргодична тогда и только тогда, когда выполняется неравенство*

$$\lambda < \mathbf{p}_1 \mathbf{S}_0^{(1)} + \mathbf{p}_2 \mathbf{S}_0^{(2)}, \tag{3}$$

где

$$\mathbf{p}_1 = \mathbf{x}_1(\mathbf{e}_{\bar{V}} \otimes I_{M^{(1)}}),\ \mathbf{p}_2 = \mathbf{x}_2(\mathbf{e}_{\bar{V}} \otimes \mathbf{e}_R \otimes I_{M^{(2)}}), \tag{4}$$

а вектор $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ *является единственным решением системы линейных алгебраических уравнений*

$$\mathbf{x}\Gamma = \mathbf{0},\ \mathbf{x}\mathbf{e} = 1. \tag{5}$$

Здесь

$$\Gamma = \begin{pmatrix} H_0 \oplus S^{(1)} + I_{\bar{V}} \otimes \mathbf{S}_0^{(1)} \mathbf{b}^{(1)} & O & H_1 \otimes \mathbf{e}_{M^{(1)}} \otimes \mathbf{t} \otimes \mathbf{a} \\ I_{\bar{V}} \otimes \mathbf{T}_0 \otimes \mathbf{e}_{M^{(2)}} \otimes \mathbf{b}^{(1)} & H \oplus T \oplus S^{(2)} + I_{\bar{V}R} \otimes \mathbf{S}_0^{(2)} \mathbf{b}^{(2)} & O \\ I_{\bar{V}} \otimes \mathbf{T}_0 \otimes \mathbf{e}_L \otimes \mathbf{b}^{(1)} & I_{\bar{V}} \otimes I_R \otimes \mathbf{A}_0 \otimes \mathbf{b}^{(2)} & H \oplus T \oplus A \end{pmatrix}.$$

Доказательство. Согласно [8] необходимое и достаточное условие эргодичности рассматриваемой квазитеплицевой цепи Маркова $\xi_t,\ t \ge 0,$ состоит в выполнении неравенства

$$\mathbf{y}Q'(1)\mathbf{e} < 0, \tag{6}$$

где вектор $\mathbf{y}$ является единственным решением системы

$$\mathbf{y}Q(1) = \mathbf{0}; \tag{7}$$

$$\mathbf{y}\mathbf{e} = 1. \tag{8}$$

Представим вектор $\mathbf{y}$ в виде

$$\mathbf{y} = (\mathbf{q}\otimes\mathbf{x}_1, \mathbf{q}\otimes\mathbf{x}_2, \mathbf{q}\otimes\mathbf{x}_3), \tag{9}$$

где $\mathbf{x} \overset{def}{=} (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ – стохастический вектор.

Тогда, учитывая, что $\mathbf{q}\sum\limits_{k=0}^{\infty} D_k = \mathbf{0}$, система (7), (8) сводится к виду (5).

Подставляя в неравенство (6) вектор $\mathbf{y}$ в виде (9), выражение для $Q'(1)$, вычисленное с помощью формулы (2), и учитывая, что $\mathbf{q}D'(1)\mathbf{e} = \lambda,$ сводим это неравенство к выражению

$$\lambda + xQ^{-}\mathbf{e} < 0, \tag{10}$$

где

$$Q^{-} = \begin{pmatrix} H_0 \oplus S^{(1)} & O & H_1 \otimes \mathbf{e}_{M^{(1)}} \otimes \mathbf{t} \otimes \mathbf{a} \\ I_{\bar V} \otimes \mathbf{T}_0 \otimes \mathbf{b}^{(1)} & H \oplus T \oplus S^{(2)} & O \\ I_{\bar V} \otimes \mathbf{T}_0 \otimes \mathbf{e}_L \otimes \mathbf{b}^{(1)} & I_{\bar V} \otimes I_R \otimes \mathbf{A}_0 \otimes \mathbf{b}^{(2)} & H \oplus T \oplus A \end{pmatrix}.$$

Используя соотношения $H\mathbf{e} = (H_0 + H_1)\mathbf{e} = \mathbf{0},\ T\mathbf{e} + \mathbf{T}_0 = \mathbf{0},\ A\mathbf{e} + \mathbf{A}_0 = \mathbf{0},$ сведем неравенство (10) к виду

$$\lambda < x_1(\mathbf{e}_{\bar V} \otimes I_{M^{(1)}})\mathbf{S}_0^{(1)} + x_2(\mathbf{e}_{\bar V} \otimes \mathbf{e}_R \otimes I_{M^{(2)}})\mathbf{S}_0^{(2)}.$$

После использования обозначений (4) последнее приобретает вид (3).

Замечание. При физической интерпретации условия эргодичности (3) учитываем, что данное условие отражает процесс обслуживания в системе в условиях перегрузки и векторы $\mathbf{p}_n, n = 1, 2,$ имеют следующий смысл: компонента $\mathbf{p}_1(m^{(1)})$ вектора $\mathbf{p}_1$ есть вероятность того, что прибор 1 исправен и обслуживает запрос на фазе $m^{(1)}$, $m^{(1)} = \overline{1, M^{(1)}},$ компонента $\mathbf{p}_2(m^{(2)})$ вектора $\mathbf{p}_2$ есть вероятность того, что прибор 1 на ремонте, переключение закончилось и прибор 2 обслуживает запрос на фазе $m^{(2)}$, $m^{(2)} = \overline{1, M^{(2)}}.$ Тогда правая часть неравенства (3) выражает суммарную интенсивность обслуживания запросов в условиях перегрузки. Очевидно, что для существования стационарного режима в системе необходимо и достаточно, чтобы интенсивность входного потока λ была меньше суммарной интенсивности обслуживания в условиях перегрузки.

Следствие 3. *В случае стационарного пуассоновского потока поломок и экспоненциальных распределений времен обслуживания, переключения и ремонтов условие эргодичности (3) сводится к следующему неравенству:*

$$\lambda < \pi_1\mu_1 + \pi_2\mu_2,$$

где

$$\pi_1 = \frac{a+\phi}{h}\Big[1 + \frac{a}{\phi} + \frac{a+\phi}{h}\Big]^{-1},\ \pi_2 = \frac{a}{\phi}\Big[1 + \frac{a}{\phi} + \frac{a+\phi}{h}\Big]^{-1}.$$

4. Стационарное распределение

Предполагаем, что условие эргодичности, заданное теоремой 2, выполняется. Введем обозначения для предельных вероятностей состояний цепи Маркова ξ_t, которые в нашем случае являются и стационарными вероятностями состояний рассматриваемой системы массового обслуживания:

$$p_0^{(0)}(\nu,\eta)=\lim_{t\to\infty}P\{i_t=0,\ n_t=0,\ \nu_t=\nu,\ \eta_t=\eta\},\quad \nu=\overline{0,W},\quad \eta=\overline{0,V};$$

$$p_0^{(1)}(\nu,\eta)=\lim_{t\to\infty}P\{i_t=0,\ n_t=1,\ \nu_t=\nu,\ \eta_t=\eta,\ \vartheta_t=\vartheta\},\ \nu=\overline{0,W},\ \ \eta=\overline{0,V},\ \ \vartheta=\overline{1,R};$$

$$p_i^{(0,0)}\{(\nu,\eta,m^{(1)})\}=\lim_{t\to\infty}P\{i_t=i,\ \ n_t=0,\ \ r_t=0,\ \nu_t=\nu,$$

$$\eta_t=\eta,\ \ m_t^{(1)}=m^{(1)}\},\ \ i>0,\ \nu=\overline{0,W},\ \ \eta=\overline{0,V},\ \ m^{(1)}=\overline{1,M^{(1)}};$$

$$p_i^{(1,0)}\{(\nu,\eta,m^{(2)},\vartheta)\}=\lim_{t\to\infty}P\{i_t=i,\ \ n_t=1,\ r_t=0,\ \nu_t=\nu,\ \eta_t=\eta,\ \vartheta_t=\vartheta,\ \ m_t^{(2)}=m^{(2)}\},$$

$$i>0,\ \nu=\overline{0,W},\ \ \eta=\overline{0,V},\ \ \vartheta=\overline{1,R},\quad m^{(2)}=\overline{1,M^{(2)}};$$

$$p_i^{(1,1)}\{(\nu,\eta,\vartheta,l)\}=\lim_{t\to\infty}P\{i_t=i,\ n_t=1,\ r_t=1,\ \nu_t=\nu,\ \eta_t=\eta,\ \vartheta_t=\vartheta,\ l_t=l\},$$

$$i>0,\quad \nu=\overline{0,W},\ \eta=\overline{0,V},\quad \vartheta=\overline{1,R},\quad l=\overline{1,L}.$$

Внутри каждой из выделенных групп упорядочим вероятности в лексикографическом порядке компонент и сформируем векторы этих вероятностей

$$\mathbf{p}_0^{(0)},\ \mathbf{p}_0^{(1)},\ \mathbf{p}_i^{(0,0)},\ \mathbf{p}_i^{(1,0)},\ \mathbf{p}_i^{(1,1)},\ \ i\ge 1.$$

Далее сформируем векторы стационарных вероятностей, соответствующих значениям счетной компоненты, как

$$\mathbf{p}_0=(\mathbf{p}_0^{(0)},\ \mathbf{p}_0^{(1)}),\ \ \mathbf{p}_i=(\mathbf{p}_i^{(0,0)},\ \mathbf{p}_i^{(1,0)},\ \mathbf{p}_i^{(1,1)}),\ i\ge 1.$$

Порядок вектора $\mathbf{p}_0$ равен $a(R+1)$, порядок каждого из векторов $\mathbf{p}_i, i\ge 1$, равен $a(M^{(1)}+RM^{(2)}+RL)$.

Для вычисления векторов $\mathbf{p}_i, i>0$, используется алгоритм, разработанный в [8] для вычисления стационарного распределения многомерных квазитеплицевых цепей Маркова. Идея этого алгоритма состоит в выводе другой системы линейных алгебраических уравнений для искомых векторов на основе использования понятия сенсорной цепи Маркова (см., например, [9]).

Алгоритм. Векторы стационарных вероятностей $\mathbf{p}_i,\ i\ge 0$, находятся следующим образом:

1. Вычисляется матрица G как стохастическое решение нелинейного матричного уравнения

$$\sum_{k=0}^{\infty}Q_kG^k=\mathbf{0}.$$

2. Вычисляется матрица G_0 из уравнения

$$\widehat{Q}_0+(Q_1+\sum_{n=2}^{\infty}Q_nG^{n-1})G_0=O,$$

откуда

$$G_0=-(Q_1+\sum_{n=2}^{\infty}Q_nG^{n-1})^{-1}\widehat{Q}_0.$$

3. Вычисляются матрицы $\bar{Q}_{i,l},\ l\ge i,\ \ i\ge 0$, по формулам

$$\bar{Q}_{i,l} = \begin{cases} \tilde{Q}_l + \sum_{n=l+1}^{\infty} \tilde{Q}_n G_{n-1} G_{n-2} \ldots G_l, \; i = 0, \; l \ge 0, \\ Q_{l-i} + \sum_{n=l+1}^{\infty} Q_{n-i} G_{n-1} G_{n-2} \ldots G_l, \; i \ge 1, \; l \ge i, \end{cases}$$

где $G_i = G, \; i \ge 1$.

4. Вычисляются матрицы $\Phi_l, \; l \ge 0$, по рекуррентным формулам

$$\Phi_0 = I, \; \Phi_l = \sum_{i=0}^{l-1} \Phi_i \bar{Q}_{i,l} (-\bar{Q}_{l,l})^{-1}, \quad l \ge 1.$$

5. Вектор $\mathbf{p}_0$ вычисляется как единственное решение системы линейных алгебраических уравнений

$$\mathbf{p}_0(-\bar{Q}_{0,0}) = \mathbf{0}, \; \mathbf{p}_0(\mathbf{e}_{a(R+1)} + \sum_{l=1}^{\infty} \Phi_l \mathbf{e}) = 1.$$

6. Векторы стационарных вероятностей $\mathbf{p}_l, \; l \ge 1$, вычисляются как

$$\mathbf{p}_l = \mathbf{p}_0 \Phi_l, \; l \ge 1.$$

Преимуществом этого алгоритма выступает отсутствие операции вычитания в рекурсиях. Все матрицы, фигурирующие в алгоритме, являются неотрицательными, что обеспечивает устойчивость вычислений при компьютерной реализации алгоритма.

5. Векторная производящая функция стационарного распределения. Характеристики производительности системы

Вычислив векторы стационарных вероятностей $\mathbf{p}_i, \; i \ge 0$, можно найти также различные характеристики производительности системы. При этом будет полезным следующий результат.

Теорема 3. *Векторная производящая функция* $\mathbf{P}(z) = \sum_{i=1}^{\infty} \mathbf{p}_i z^i, \; |z| \le 1$, *удовлетворяет уравнению*

$$\mathbf{P}(z) Q(z) = z[\mathbf{p}_1 Q_0 - \mathbf{p}_0 \tilde{Q}(z)]. \tag{11}$$

Формула (11), в частности, может быть использована для вычисления значений функции $\mathbf{P}(z)$ и ее производных в точке $z = 1$ без вычисления бесконечных сумм. Полученные значения позволят найти моменты числа заявок в системе и ряд других ее характеристик. Заметим, что непосредственно вычислить величину $\mathbf{P}(z)$ и ее производных в точке $z = 1$ из уравнения (11) не удается, так как матрица $Q(1)$ вырожденная. Преодолеть эту трудность можно путем использования рекуррентных формул, приведенных в следствии 4.

Обозначим через $f^{(m)}(z)$ m-ю производную функции $f(z), \; m \ge 1$, и $f^{(0)}(z) = f(z)$.

Следствие 4. *m-я производная векторной производящей функции* $\mathbf{P}(z), \; |z| \le 1$, *в точке* $z = 1$ *(факториальный момент m-го порядка) вычисляется рекуррентно как решение следующей системы линейных алгебраических уравнений:*

$$\begin{cases} \mathbf{P}^{(m)}(1) Q(1) = B^{(m)}(1) - \sum_{l=0}^{m-1} C_m^l \mathbf{P}^{(l)}(1) Q^{(m-l)}(1); \\ \mathbf{P}^{(m)}(1) Q'(1) \mathbf{e} = \frac{1}{m+1} [B^{(m+1)}(1) - \sum_{l=0}^{m-1} C_{m+1}^l \mathbf{P}^{(l)}(1) Q^{(m+1-l)}(1)] \mathbf{e}, \end{cases}$$

где

$$B^{(m)}(1)=\begin{cases}\mathbf{p}_1Q_0-\mathbf{p}_0\tilde{Q}(1),\ m=0;\\ \mathbf{p}_1Q_0-\mathbf{p}_0\tilde{Q}(1)-\mathbf{p}_0\tilde{Q}'(1),\ m=1;\\ -\mathbf{p}_0[m\tilde{Q}^{(m-1)}(1)+Q^{(m)}(1)],\ m>1,\end{cases}$$

а производные $Q^{(m)}(1)$ и $\tilde{Q}^{(m)}(1)$ находятся с помощью формул (1), (2).

Доказательство следствия аналогично доказательству из работы [10].

Приведем формулы для некоторых важных характеристик производительности системы:

$\rho=\mathbf{p}_1\mathbf{S}_0^{(1)}+\mathbf{p}_2\mathbf{S}_0^{(2)}$ – пропускная способность системы (максимальное значение интенсивности потока, который может быть пропущен через систему);

$L=\mathbf{P}^{(1)}(1)\mathbf{e}$ – среднее число запросов в системе;

$V=\mathbf{P}^{(2)}(1)\mathbf{e}+L-L^2$ – дисперсия числа запросов в системе;

$P_0^{(0)}=\mathbf{p}_0^{(0)}\mathbf{e}$ – вероятность того, что система пуста и прибор 1 в исправном состоянии;

$P_0^{(1)}=\mathbf{p}_0^{(1)}\mathbf{e}$ – вероятность того, что система пуста и прибор 1 на ремонте;

$P_0^{(0,0)}=\mathbf{P}(1)\mathrm{diag}\{I_{aM^{(1)}},\mathbf{0}_{aR(M^{(2)}+L)}\}\mathbf{e}$ – вероятность того, что прибор 1 обслуживает запрос;

$P_0^{(1,0)}=\mathbf{P}(1)\mathrm{diag}\{\mathbf{0}_{aM^{(1)}},I_{aRM^{(2)}},\mathbf{0}_{aRL}\}\mathbf{e}$ – вероятность того, что прибор 1 на ремонте, а прибор 2 обслуживает запрос;

$P_0^{(1,1)}=\mathbf{P}(1)\mathrm{diag}\{\mathbf{0}_{a(M^{(1)}+RM^{(2)})},I_{aRL}\}\mathbf{e}$ – вероятность того, что прибор 1 на ремонте и идет переключение с этого прибора на прибор 2;

$P_0=\mathbf{p}_0^{(0)}\mathbf{e}+\mathbf{P}(1)\mathrm{diag}\{I_{aM^{(1)}},\mathbf{0}_{aR(M^{(2)}+L)}\}\mathbf{e}$ – вероятность того, что прибор 1 в исправном состоянии;

$P_1=\mathbf{p}_0^{(1)}\mathbf{e}+\mathbf{P}(1)\mathrm{diag}\{\mathbf{0}_{aM^{(1)}},I_{aR(M^{(2)}+L)}\}\mathbf{e}$ – вероятность того, что прибор 1 на ремонте.

6. Стационарное распределение времени пребывания

Пусть $V(x)$ – стационарная функция распределения времени пребывания произвольного запроса в системе, $v(s)=\int\limits_0^\infty e^{-sx}dV(x),\ Re(s)\geq 0,$ – преобразование Лапласа – Стилтьеса этой функции.

Теорема 4. *Преобразование Лапласа – Стилтьеса стационарного распределения времени пребывания произвольного запроса в системе имеет вид*

$$v(s)=\lambda^{-1}\{\mathbf{p}_0\sum_{k=1}^{\infty}\tilde{Q}_k\sum_{l=1}^{k}\Phi^l(s)+\sum_{i=1}^{\infty}\mathbf{p}_i\sum_{k=2}^{\infty}Q_k\sum_{l=1}^{k-1}\Phi^{i+l}(s)\}\mathbf{e},\tag{12}$$

где

$$\Phi(s)=(sI-\bar{Q})^{-1}Q_0,\quad \bar{Q}=Q(1)-Q_0.$$

Доказательство основано на вероятностном смысле преобразования Лапласа – Стилтьеса. Предполагается, что независимо от функционирования системы поступает стационарный пуассоновский поток так называемых катастроф с интенсивностью s, $s>0$. Тогда преобразование Лапласа – Стилтьеса $v(s)$ можно интерпретировать как вероятность того, что за время пребывания запроса в системе не произойдет катастрофа. Такая интерпретация позволяет вывести выражение для $v(s)$ путем вероятностных рассуждений.

Предполагаем, что в начале обслуживания запроса первоначальная фаза времени обслуживания уже установлена. Тогда матрица вероятностей того, что за время обслуживания не произойдет катастрофа и произойдут соответствующие переходы конечных компонент цепи

Маркова ξ_t, $t \geq 0$, вычисляется как $\widehat{\Phi}(s) = \int_0^\infty e^{(-sI+\bar{Q})t}\widehat{Q}_0 dt = (sI - \bar{Q})^{-1}\widehat{Q}_0$, если в момент окончания обслуживания нет очереди, и как $\Phi(s) = \int_0^\infty e^{(-sI+\bar{Q})t}Q_0 dt = (sI - \bar{Q})^{-1}Q_0$, если в момент окончания обслуживания есть очередь. Заметим, что

$$\widehat{\Phi}(s)\mathbf{e} = \Phi(s)\mathbf{e}, \tag{13}$$

так как $\widehat{Q}_0\mathbf{e} = Q_0\mathbf{e}$.

Предположим, что произвольный запрос, поступающий в группе размером k, размещается на j-м месте в группе с вероятностью $1/k$. Тогда, используя формулу полной вероятности и свойство (13), получим следующее выражение для $v(s)$:

$$v(s) = \mathbf{p}_0 \sum_{k=1}^{\infty} \frac{k}{\lambda} \tilde{Q}_k \sum_{l=1}^{k} \frac{1}{k} \Phi^l(s)\mathbf{e} + \sum_{i=1}^{\infty} \mathbf{p}_i \sum_{k=2}^{\infty} \frac{k-1}{\lambda} Q_k \sum_{l=1}^{k-1} \frac{1}{k-1} \Phi^{i+l}(s)\mathbf{e}. \tag{14}$$

Формула (12) следует непосредственно из (14).

Следствие 5. *Среднее время* $\bar{v}$ *пребывания произвольного запроса в системе вычисляется как*

$$\bar{v} = -\lambda^{-1}[\mathbf{p}_0 \sum_{k=1}^{\infty} \tilde{Q}_k \sum_{l=1}^{k} \sum_{m=0}^{l-1} \Phi^m(0) + \sum_{i=1}^{\infty} \mathbf{p}_i \sum_{k=2}^{\infty} Q_k \sum_{l=1}^{k-1} \sum_{m=0}^{i+l-1} \Phi^m(0)]\Phi'(0)\mathbf{e}, \tag{15}$$

где $\Phi'(0) = -(\bar{Q})^{-2}Q_0$.

Доказательство. Чтобы получить формулу (15), используется соотношение $\bar{v} = -v'(0)$ и тот факт, что матрица $\Phi(0)$ является стохастической.

Заключение

В настоящей работе исследована система массового обслуживания с бесконечным буфером и двумя неоднородными приборами, один из которых является основным, а другой – резервным. Основной прибор является высокоскоростным, но ненадежным. Если этот прибор выходит из строя, обслуживание текущей заявки переходит на резервный прибор, который является низкоскоростным, но надежным. При достаточно общих предположениях о характере потоков запросов и поломок и о распределениях времен обслуживания, переключений, ремонтов поведение системы описывается в терминах многомерной цепи Маркова с непрерывным временем. Найдено конструктивное условие существования стационарного режима в системе, вычислено ее стационарное распределение и факториальные моменты этого распределения, получены формулы для основных характеристик производительности системы, а также формулы для преобразования Лапласа – Стилтьеса времени пребывания запроса в системе и для математического ожидания этого времени.

Результаты исследования могут быть использованы для вычисления производительности, проектирования и модернизации гибридных систем связи, состоящих из основного лазерного и резервного беспроводного широкополосного радиоканала, функционирующего под управлением протокола IEEE 802.11n.

Список литературы

1. Vishnevsky, V. Redundant queueing system with unreliable servers / V. Vishnevsky, D. Kozyrev, O. Semenova // Proc. of the 6th Intern. Congress on Ultra Modern Telecommunications and Control Systems and Workshops, Moscow, 2014. – Moscow, 2014. – P. 383–386.

2. Advanced Optical Wireless Communication Systems / S. Arnon [et al.]. – Cambridge University Press, 2012. – 404 p.

3. Vishnevsky, V.M. Modeling and analysis of a hybrid communication channel based on free-space optical and radio-frequency technologies / V.M. Vishnevsky, O.V. Semenova, S.Yu. Sharov // Automation and Remote Control. – 2013. – Vol. 72. – P. 345–352.

4. Sharov, S.Yu. Simulation model of wireless channel based on FSO and RF technologies / S.Yu. Sharov, O.V. Semenova // Distributed Computer and Communication Networks. Theory and Applications (DCCN–2010). – Moscow, 2010. – P. 368–374.

5. Lucantoni, D.M. New results on the single server queue with a batch Markovian arrival process / D.M. Lucantoni // Communications in Statistics-Stochastic Models. – 1991. – Vol. 7. – P. 1–46.

6. Neuts, M.F. Matrix-geometric solutions in stochastic models / M.F. Neuts. – Baltimore : The Johns Hopkins University Press, 1981. – 332 p.

7. Бочаров, П.П. Теория массового обслуживания / П.П. Бочаров, А.В. Печинкин. – М. : Изд-во РУДН, 1999. – 528 с.

8. 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.

9. Kemeni, J.G. Denumerable Markov Chains / J.G. Kemeni, J.L. Snell, A.W. Knapp. – N. Y. : Van Nostrand, 1966. – 348 p.

10. Dudin, A. Recursive formulas for the moments of queue length in the $BMAP/G/1$ queue / A. Dudin, V. Klimenok, M. Ho Lee // IEEE Communication Letters. – 2009. – Vol. 13 – P. 351–353.

Поступила 20.01.2016

Белорусский государственный университет,
Минск, пр. Независимости, 4
e-mail: klimenok@bsu.by

V.I. Klimenok

STATIONARY CHARACTERISTICS OF UNRELIABLE QUEUEING SYSTEM WITH BATCH MARKOVIAN ARRIVAL PROCESS AND RESERVE SERVER

In the paper, a queueing system with unreliable server and so called «cold» redundancy is analyzed. The system consists of an infinite buffer, the main unreliable server and the reliable reserve server. The term «cold reserve» means that the reserve server is off until the main server is in good order and is activated when the main server is under repair. The input flow to the system is a BMAP (Batch Markovian Arrival Process). Breakdowns arrive to the main server according to a MAP (Markovian Arrival Process). Service times as well as repair time have PH (Phase type) distribution. A condition for the stable operation of the system is shown, its stationary distribution and the main characteristics are calculated and the expression for the Laplace – Stieltjes transform of the sojourn time distribution is derrived.