РАСШИРЕНИЕ МОДЕЛЬНЫХ ПРЕДПОЛОЖЕНИЙ И СООТВЕТСТВУЮЩИХ РЕЗУЛЬТАТОВ ДЛЯ СЕТЕЙ ГЕЛЕНБЕ

Ю. В. Малинковский

Гомельский государственный университет им. Ф. Скорины
Гомель, Беларусь
E-mail: malinkovsky@gsu.by

Рассматривается экспоненциальная сеть массового обслуживания, отличающаяся от сети Геленбе (с обычными положительными и так называемыми отрицательными заявками) тем, что существенно ослаблены условия на возможные значения параметров входящих в сеть пуассоновских потоков положительных и отрицательных заявок в теореме Геленбе. Это позволяет вложить класс сетей Джексона с однолинейными узлами в класс сетей Геленбе.

Ключевые слова: сети Джексона, сети Геленбе, уравнения трафика, стационарное распределение, эргодичность, форма произведения.

ВВЕДЕНИЕ

В известной работе [1] впервые рассматривались однолинейные сети с положительными и отрицательными заявками, для которых было найдено в мультипликативной форме стационарное распределение, выраженное через решение нелинейных уравнений трафика. В [2] с помощью теоремы Брауэра о неподвижной точке доказано существование положительного решения нелинейных уравнений трафика. Существенным недостатком работ [1, 2] и более позднего обзора Арталехо [3] является требование наличия входящих потоков положительных и отрицательных заявок во все без исключения узлы сети. Цель настоящей работы состоит в перенесении результатов [1–3] на сети Геленбе, в которых будет ослаблено требование наличия всех входящих потоков в узлы сети извне. Вместо него потребуется только наличие хотя бы одного входящего потока положительных заявок и неприводимости вводимой ниже матрицы маршрутизации, что, в частности, позволяет вложить модель Джексона с однолинейными узлами в модифицированную таким образом модель Геленбе.

1. МОДИФИЦИРОВАННЫЕ G-СЕТИ

В сеть массового обслуживания, состоящую из N однолинейных экспоненциальных узлов (систем) с интенсивностью обслуживания μ_i для i-го узла, поступают $2N$ независимых стационарных пуассоновских потоков заявок, причем в i-й узел поступают два потока – поток положительных (обычных заявок, требующих обслуживания) с интенсивностью Λ_i и поток отрицательных заявок с интенсивностью λ_i $(i=\overline{1,N})$. В работе [1] предполагалось, что для всех индексов i выполняется $\Lambda_i > 0$, $\lambda_i > 0$ (предположение I), хотя, как будет показано, это не принципиально и необходимо только для неприводимости описывающей поведение сети цепи Маркова с непрерывным временем. В дальнейшем это предположение будет заменено более слабым. При поступлении в i-й узел, в котором имеется хотя бы одна положительная заявка, отрицательная заявка вычеркивает, удаляет из сети ровно одну положительную заявку и мгновенно покидает сеть, не оказывая в дальнейшем на нее никакого влияния. Если отрицательная заявка поступает в узел, свободный от положительных заявок, то она вообще не оказывает влияния на сеть и мгновенно пропадает. Процессы поступления и обслуживания предполагаются независимыми. Число мест для ожидания заявок в узле бесконечное. Время обслуживания положительной заявки единственным прибором i-го узла имеет показательное распределение с параметром μ_i $(i=\overline{1,N})$. Если положительная заявка поступает в узел, свободный от положительных заявок, она сразу начинает обслуживаться. Для определенности будем предполагать, что заявки обслуживаются в порядке поступления в узлы (дисциплина FCFS). Положительная заявка, обслуженная в i-м узле, мгновенно и независимо от других заявок с вероятностью p_{ij}^{+} направляется в j-й узел как положительная заявка, с вероятностью p_{ij}^{-} как отрицательная заявка, а с вероятностью p_{i0} покидает сеть $(i,j=\overline{1,N},\ \sum_{j=0}^{N} p_{ij}=1)$, где $p_{ij}=p_{ij}^{+}+p_{ij}^{-}$ для $j\neq 0$.

Обозначим через λ_i^{+} и λ_i^{-} соответственно интенсивности потоков положительных и отрицательных заявок, поступающих в i-й узел (извне и из других узлов, $i=\overline{1,N}$) в стационарном режиме. Изолируем i-й узел от сети и направим на него два независимых простейших потока заявок – потока положительных заявок с интенсивностью λ_i^{+} и потока отрицательных заявок с интенсивностью λ_i^{-}. В остальном, касающемся процесса обслуживания и ограничения на время пребывания, поведение изолированного узла точно такое, как поведение в сети. Очевидно, число заявок в такой системе массового обслуживания $\tilde{n}_i(t)$ является процессом размножения и гибели с постоянными интенсивностями рождения λ_i^{+} и постоянными интенсивностями гибели $\mu_i+\lambda_i^{-}$. Следовательно, для эргодичности марковского процесса $\tilde{n}_i(t)$ необходимо и достаточно, чтобы $\rho_i<1$, где $\rho_i=\dfrac{\lambda_i^{+}}{\mu_i+\lambda_i^{-}}$ – загрузка системы, а стационарное распределение процесса $\tilde{n}_i(t)$

$$p_i(n_i)=\rho_i^{n_i}(1-\rho_i),\quad n_i=0,1,\ldots . \tag{1}$$

В силу (1) ρ_i можно трактовать как вероятность занятости прибора в изолированном узле в стационарном режиме. По формуле полного математического ожидания интенсивность выходящего из изолированного i-го узла потока положительных заявок $\mu_i\rho_i$. Поэтому в стационарном режиме выполняется следующий закон сохранения:

$$\lambda_i^{+}=\Lambda_i+\sum_{j=1}^{N}\mu_j\rho_j p_{ji}^{+},\quad \lambda_i^{-}=\lambda_i+\sum_{j=1}^{N}\mu_j\rho_j p_{ji}^{-},\quad i=\overline{1,N}. \tag{2}$$

Решением нелинейных уравнений трафика (2) назовем $(\lambda_1^{+},\ldots,\lambda_N^{+},\lambda_1^{-},\ldots,\lambda_N^{-})$. В [2] с помощью теоремы Брауэра о неподвижной точке доказано, что эти уравнения имеют решение. При выполнении предположения I это решение строго положительно, так как из вида (2) следует, что для каждого i выполняется $\lambda_i^{+}\geq\Lambda_i$, $\lambda_i^{-}\geq\lambda_i$.

Состояния сети в момент t будем описывать цепью Маркова с непрерывным временем $\vec{n}(t)=(n_1(t),\dots,n_N(t))$, где $n_i(t)$ – число положительных заявок в i-м узле в момент времени t. Пространство состояний этой цепи $X=\mathrm{Z}_+^N$, где $\mathrm{Z}_+^N=\{0,1,\dots\}$. В силу предположения I цепь Маркова $\vec{n}(t)$ неприводима. Интуитивно понятно, что если загрузки узлов меньше 1,

$$\rho_i<1,\quad i=\overline{1,N}, \tag{3}$$

то цепь Маркова $\vec{n}(t)$ является эргодической, что доказано в [1] с помощью эргодической теоремы Фостера [4]. Пусть $\{p(\vec{n}),\vec{n}\in X\}$ – ее предельное эргодическое распределение, которое в этом случае будет единственным решением глобальных уравнений равновесия

$$\begin{aligned} p(\vec{n})\sum_{i=1}^{N}\left[\Lambda_i+(\mu_i+\lambda_i)I_{\{n_i\neq 0\}}\right]=\sum_{i=1}^{N}\Big\{p(\vec{n}-\vec{e}_i)\Lambda_i I_{\{n_i\neq 0\}}+p(\vec{n}+\vec{e}_i)(\mu_i p_{i0}+\lambda_i)+\\ +\sum_{j=1}^{N}\left[p(\vec{n}+\vec{e}_j-\vec{e}_i)\mu_j p_{ji}^{+}I_{\{n_i\neq 0\}}+p(\vec{n}+\vec{e}_j+\vec{e}_i)\mu_j p_{ji}^{-}+p(\vec{n}+\vec{e}_j)\mu_j p_{ji}^{-}I_{\{n_i=0\}}\right]\Big\}, \end{aligned} \tag{4}$$

удовлетворяющим условию нормировки $\sum_{\vec{n}\in X}p(\vec{n})=1$. Здесь $\vec{e}_i$ – единичный вектор i-го направления, I_A – индикатор события A, равный 1, если A происходит, и 0 в противном случае.

В [1] доказано, что в предположении I при выполнении условия эргодичности (3) процесс $\vec{n}(t)$ эргодичен, а вероятности, задаваемые с помощью равенства $p(\vec{n})=p_1(n_1)\dots p_N(n_N)$, где $p_i(n_i)$ определены равенством (1), удовлетворяют уравнениям глобального равновесия (4), т.е. образуют эргодическое распределение $\vec{n}(t)$.

Предположение I, достаточное для неприводимости $\vec{n}(t)$, в постановке задачи можно заменить на более слабое условие. Прежде всего, предположим, что все $\Lambda_i\geq 0,\lambda_i\geq 0$ и хотя бы одно $\Lambda_i>0$, т.е. могут отсутствовать все или некоторые потоки отрицательных заявок в узлы сети извне и обязательно должен присутствовать хотя бы один поток положительных заявок в один из узлов извне. Заметим, что рассматриваемая сеть эквивалентна следующей сети без отрицательных заявок (эквивалентность понимается в том смысле, что эта сеть описывается теми же уравнениями Колмогорова и равновесия, что и первоначальная сеть с отрицательными заявками). В сеть, состоящую из N однолинейных узлов с показательным распределением времени обслуживания прибором j-го узла с параметром $\mu_j+\lambda_j$, поступает N независимых простейших потоков с интенсивностью Λ_j в j-й узел. Заявка, завершившая обслуживание в j-м узле, мгновенно и независимо от других заявок покидает сеть без других заявок с вероятностью $\dfrac{\mu_j p_{j0}+\lambda_j}{\mu_j+\lambda_j}$, либо покидает сеть вместе с заявкой из i-го узла ($i=\overline{1,N}$) с вероятностью $\dfrac{\mu_j p_{ji}^{-}}{\mu_j+\lambda_j}$ (включая фиктивный случай, когда в i-м узле отсутствуют заявки), либо присоединяется к i-му узлу с вероятностью $\dfrac{\mu_j p_{ji}^{+}}{\mu_j+\lambda_j}$. Легко видеть, что сумма этих вероятностей

$$\frac{\mu_j p_{j0}+\lambda_j}{\mu_j+\lambda_j}+\sum_{i=1}^{N}\left(\frac{\mu_j p_{ji}^-}{\mu_j+\lambda_j}+\frac{\mu_j p_{ji}^+}{\mu_j+\lambda_j}\right)$$

действительно равняется единице. Идея состоит в том, чтобы построить подходящим способом матрицу маршрутизации для этой эквивалентной сети, а затем найти условия на эту матрицу, которые гарантируют неприводимость процесса, описывающего эквивалентную сеть, а значит, и цепи $\vec{n}(t)$.

Добавим к узлам $1,2,\ldots,N$ узел 0, трактуемый как внешность сети, представляющая собой одновременно источник и сток. Введем квадратную матрицу $(N+1)$-го порядка $S=(s_{ji},\ j,i=\overline{1,N})$, где $s_{00}=0$, $s_{0i}=\frac{\Lambda_i}{\Lambda}$ для $i=\overline{1,N}$, $s_{j0}=\frac{\mu_j p_{j0}+\lambda_j}{\mu_j+\lambda_j}+\sum_{i=1}^{N}\frac{\mu_j p_{ji}^-}{\mu_j+\lambda_j}$ для $j=\overline{1,N}$, $s_{ji}=\frac{\mu_j p_{ji}^+}{\mu_j+\lambda_j}$ для $j,i=\overline{1,N}$. Назовем эту стохастическую матрицу матрицей маршрутизации. Для краткости будем называть описывающую сеть цепь Маркова $\vec{n}(t)$ с непрерывным временем цепью Маркова, а цепь Маркова с дискретным временем с матрицей вероятностей перехода S за один шаг S-цепью. Отметим, что, в отличие от матрицы маршрутизации сети Джексона, матрица маршрутизации рассматриваемой сети неоднозначно определяет будущее значение процесса $\vec{n}(t)$ при переходе S-цепи в состояние 0 сразу после момента скачка процесса по текущему его значению.

Лемма. Пусть хоть одно из $\Lambda_i>0$. Цепь Маркова $\vec{n}(t)$ неприводима тогда и только тогда, когда неприводима S-цепь.

Сначала докажем достаточность. Пусть S неприводима. Любому пути из состояния $\vec{n}$ в состояние $\vec{m}$ цепи, связанному с перемещениями заявок в моменты скачков процесса $\vec{n}(t)$, соответствует некоторое перемещение заявок между узлами сети $0,1,\ldots,N$, т.е. некоторый путь в графе переходов S-цепи. Так как интенсивности выхода из любых состояний цепи строго положительны, а интенсивность перехода по некоторой дуге в графе переходов цепи равна произведению интенсивности выхода из начала дуги на вероятность перехода в соответствующей S-цепи, то произведение интенсивностей перехода цепи по любому конечному пути из состояния $\vec{n}$ в состояние $\vec{m}$ будет положительным тогда и только тогда, когда будет положительным произведение вероятностей перехода по соответствующему пути в графе переходов S-цепи.

Докажем, что любое состояние $\vec{n}$ цепи достижимо из состояния $\vec{0}$. В силу неприводимости S для любого $i=\overline{1,N}$ найдется конечная цепочка переходов S-цепи $0\to i_1\to i_2\to\ldots\to i_l\to i$ такая, что $s_{0i_1}s_{i_1i_2}\ldots s_{i_{l-1}i_l}s_{i_li}>0$, причем, не нарушая общности, можно считать, что состояние 0 встречается только в ее начале. Поскольку пути $\vec{n}\to\vec{n}+\vec{e_{i_1}}\to\vec{n}+\vec{e_{i_2}}\to\ldots\to\vec{n}+\vec{e_{i_l}}\to\vec{n}+\vec{e_i}$ цепи из $\vec{n}$ в $\vec{n}+\vec{e_i}$ соответствует именно приведенный выше путь в S-цепи, то произведение интенсивностей перехода цепи по пути $\vec{n}\to\vec{n}+\vec{e_{i_1}}\to\vec{n}+\vec{e_{i_2}}\to\ldots\to\vec{n}+\vec{e_{i_l}}\to\vec{n}+\vec{e_i}$ будет положительным. Итак, для любого $i=\overline{1,N}$ состояние $\vec{n}+\vec{e_i}$ достижимо из состояния $\vec{n}$. Обозначим отношение достижимости значком $\Rightarrow$, т.е. запись $\vec{n}\Rightarrow\vec{m}$ или $\vec{m}\Leftarrow\vec{n}$ означает, что $\vec{m}$ достижимо из $\vec{n}$. Следовательно,

$$\vec{0} \Rightarrow \vec{e_1} \Rightarrow 2\vec{e_1} \Rightarrow \ldots \Rightarrow n_1\vec{e_1} \Rightarrow n_1\vec{e_1} + \vec{e_2} \Rightarrow n_1\vec{e_1} + 2\vec{e_2} \Rightarrow \ldots \Rightarrow n_1\vec{e_1} + n_2\vec{e_2} \Rightarrow \ldots \Rightarrow$$
$$\Rightarrow n_1\vec{e_1} + n_2\vec{e_2} + \ldots + n_N\vec{e_N} = \vec{n}.$$

В силу транзитивности отношения достижимости отсюда следует, что любое состояние $\vec{n}$ цепи достижимо из состояния $\vec{0}$.

Докажем теперь, что $\vec{0}$ достижимо из любого состояния $\vec{n}$. Идея доказательства похожая, но оно осложняется тем, что заявка может покидать сеть без других заявок, а может покидать сеть вместе с еще одной заявкой любого узла, в связи с чем знание s_{j0} не позволяет установить, какая из этих возможностей реализуется при уходе заявки j-го узла из сети. В силу неприводимости S-цепи для любого $i = \overline{1,N}$ найдется конечная цепочка переходов S-цепи $i \to i_1 \to i_2 \to \ldots \to i_q \to j \to 0$ такая, что $s_{ii_1}s_{i_1i_2}\ldots s_{i_{q-1}i_q}s_{i_qj}s_{j0} > 0$, причем $i_1 \neq 0, i_2 \neq 0, \ldots, i_q \neq 0, j \neq 0$. Так как $s_{j0} > 0$, то из определения s_{j0} вытекает, что с положительной вероятностью произойдет хотя бы одно из следующих N + 1 событий: либо заявка j-го узла покинет сеть без других заявок $\left(\text{ вероятность } \dfrac{\mu_j p_{j0} + \lambda_j}{\mu_j + \lambda_j}\right)$, либо одновременно две заявки j-го узла покинут сеть $\left(\text{ вероятность } \dfrac{\mu_j p_{jj}^-}{\mu_j + \lambda_j}\right)$, либо заявка j-го узла покинет сеть вместе с заявкой другого, k-го узла ($k = \overline{1,N}, k \neq j$; вероятность $\dfrac{\mu_j p_{jk}^-}{\mu_j + \lambda_j}$). Но тогда хотя бы для одной из $N+1$ цепочек $\vec{n} \to \vec{n} - \vec{e_i} + \vec{e_{i_1}} \to \vec{n} - \vec{e_i} + \vec{e_{i_2}} \to \ldots \to \vec{n} - \vec{e_i} + \vec{e_{i_q}} \to \vec{n} - \vec{e_i} + \vec{e_j} \to \vec{n} - \vec{e_i}$ (эта цепочка заканчивается уходом заявки j-го узла из сети без других заявок), $\vec{n} \to \vec{n} - \vec{e_i} + \vec{e_{i_1}} \to \vec{n} - \vec{e_i} + \vec{e_{i_2}} \to \ldots \to \vec{n} - \vec{e_i} + \vec{e_{i_q}} \to \vec{n} - \vec{e_i} + \vec{e_j} \to \vec{n} - \overrightarrow{2e_i}$ (эта цепочка заканчивается уходом заявки j-го узла из сети вместе с заявкой i-го узла), $\vec{n} \to \vec{n} - \vec{e_i} + \vec{e_{i_1}} \to \vec{n} - \vec{e_i} + \vec{e_{i_2}} \to \ldots \to \vec{n} - \vec{e_i} + \vec{e_{i_q}} \to \vec{n} - \vec{e_i} + \vec{e_j} \to \vec{n} - \vec{e_i} + \vec{e_k}$ ($k = \overline{1,N}, k \neq i$; эти цепочки заканчиваются уходом заявки j-го узла из сети вместе с заявкой k-го узла) произведение интенсивностей перехода цепи $\vec{n}(t)$ вдоль цепочки положительное. Другими словами, для любого $i = \overline{1,N}$ из состояния $\vec{n}$ достижимо хотя бы одно из $N+1$ состояний $\vec{n} - \vec{e_i}$, $\vec{n} - \overrightarrow{2e_i}$, $\vec{n} - \vec{e_i} - \vec{e_k}$ $(k = \overline{1,N}, k \neq i)$.

Используя транзитивность отношения достижимости, можно, уменьшая первую координату то на 1, то на 2, довести ее значение до 0, остальные координаты могут не измениться или уменьшиться. Таким образом, существуют $m_2 \le n_2, \ldots, m_N \le n_N$ такие, что $\vec{n} \Rightarrow (0, m_2, \ldots, m_N)$. Применяя аналогичные рассуждения ко второй координате вектора $(0, m_2, \ldots, m_N)$, найдем такие $l_3 \le m_3, \ldots, l_N \le m_N$, что $(0, m_2, \ldots, m_N) \Rightarrow (0, 0, l_3, \ldots, l_N)$. По индукции получим $\vec{n} \Rightarrow (0, m_2, \ldots, m_N) \Rightarrow (0, 0, l_3, \ldots, l_N) \Rightarrow \ldots \Rightarrow \vec{0}$. Итак, любое состояние $\vec{n}$ цепи $\vec{n}(t)$ сообщается с состоянием $\vec{0}$, т.е. эта цепь неприводима.

Докажем теперь необходимость. Так как цепь неприводима, то для любого $i = \overline{1,N}$ ее состояния $\vec{0}$ и $\vec{e_i}$ сообщаются, т.е. $\vec{0} \Rightarrow \vec{e_i}$ и $\vec{e_i} \Rightarrow \vec{0}$. Из $\vec{0} \Rightarrow \vec{e_i}$ следует, что существует конечная цепочка переходов между состояниями $\vec{0} \to \vec{f_1} \to \vec{f_2} \to \ldots \to \vec{f_l} \to \vec{e_i}$, вдоль которой произведение интенсивностей перехода положительно. Пусть этой цепочке соответствует цепочка $0 \to i_1 \to i_2 \to \ldots \to i_l \to i$ переходов между состояниями *S*-цепи. Очевидно, произведение вероятностей перехода *S*-цепи по этой цепочке будет положительным, т.е. в *S*-цепи $0 \Rightarrow i$ для любого $i = \overline{1, N}$. Аналогично, из $\vec{e_i} \Rightarrow \vec{0}$ следует, что для каждого $i = \overline{1, N}$ в *S*-цепи $i \Rightarrow 0$. Итак, любое состояние *S*-цепи сообщается с состоянием 0. В силу транзитивности отношения сообщаемости отсюда вытекает, что все состояния *S*-цепи сообщаются, т.е. *S*-цепь неприводима. Доказательство завершено.

Отметим, что для неприводимости матрицы *S* достаточно неприводимости хотя бы одной из *N* + 1 частичных матриц маршрутизации $S_0, S_1, \ldots, S_N$, элементы которых за исключением элементов нулевого столбца совпадают с элементами матрицы *S*, нулевой элемент нулевого столбца равен 0, *j*-й элемент нулевого столбца матрицы S_0 равен $\frac{\mu_j p_{j0} + \lambda_j}{\mu_j + \lambda_j}$, *j*-й элемент нулевого столбца матрицы S_i равен $\frac{\mu_j p^-_{ji}}{\mu_j + \lambda_j}$ ($j, i = \overline{1, N}$).

Назовем предположением II требование, что все $\Lambda_i \geq 0, \lambda_i \geq 0$ и хотя бы одно из $\Lambda_i > 0$, а матрица маршрутизации *S* неприводима.

В предположении II уравнения трафика (2) имеют неотрицательное решение такое, что все $\lambda^+_i > 0$. Действительно, если предположить, что какое-то $\lambda^+_i = 0$, то из вида уравнений трафика (2) вытекает, что $\Lambda_i = 0,\ \mu_j \rho_j p^+_{ji} = 0$ для всех $j \neq 0$, т.е. $s_{ji} = 0$ для всех $j = \overline{0, N}$, значит, в матрице *S* найдется столбец, со стоящий из нулей. Но это противоречит неприводимости *S*.

Доказательство эргодичности процесса $\vec{n}(t)$ и проверка выполнения уравнений локального равновесия при выполнении условия II проводится точно так же, как для классической модели Геленбе. Окончательный результат имеет следующий вид.

Теорема. Если в предположении II выполнено условие (3), то цепь Маркова $\vec{n}(t)$ эргодична, а ее единственное стационарное распределение имеет форму произведения $p(\vec{n}) = p_1(n_1) \ldots p_N(n_N)$ с множителями (1), где $\{\lambda^+_i, \lambda^-_i,\ i = \overline{1, N}\}$ – решение уравнения трафика (2).

Отметим, что в отличие от классической сети Геленбе, модифицированная здесь сеть содержит сеть Джексона с однолинейными узлами как частный случай, если положить все $\lambda_i = 0,\ p^-_{ij} = 0$. Построенная модель позволяет охватить гораздо большее число практически важных случаев. В ней нелинейные уравнения трафика в определенных случаях могут вырождаться, превращаясь в линейные.

ЛИТЕРАТУРА

1. *Gelenbe, E.* Product-form Queueing Networks with Negative and Positive Customers / E. Gelenbe // J. Appl. Prob. 1991. V. 28. P. 656–663.
2. *Gelenbe, E.* Stability of G-networks / E. Gelenbe, R. Schassberger // Probability in the Engineering and Informational Sciences. 1992. V. 6. P. 271–276.
3. *Artalejo, J. R.* Invited Review. G-networks: A Versatile Approach for Work Removal in Queueing Networks / J. R. Artalejo // European Journal of Operational Research. 2000. V. 126. P. 233–249.
4. *Бочаров, П. П.* Теория массового обслуживания / П. П. Бочаров, А. В. Печинкин. М. : РУДН, 1995.