

УДК 519.872.5

О СЕТЕВОЙ СТОХАСТИЧЕСКОЙ МОДЕЛИ КОЛЛ-ЦЕНТРА

Т. В. Русилко

*Гродненский государственный университет имени Янки Купалы, ул. Ожешко, 22,
230023, г. Гродно, Беларусь, tatiana.rusilko@gmail.com*

Описана математическая модель процесса обработки звонков клиентов в колл-центре. В качестве модели используется замкнутая экспоненциальная сеть массового обслуживания с нетерпеливыми заявками. Целью исследования является расчет характеристик эффективности функционирования колл-центра и решение задачи подбора оптимального штатного расписания агентов с целью максимизации производительности колл-центра.

Ключевые слова: колл-центр; сеть массового обслуживания; математическое моделирование; метод асимптотического анализа.

ABOUT THE NETWORK STOCHASTIC CALL CENTER MODEL

T. V. Rusilko

*Yanka Kupala State University of Grodno, Ozheshko St., 22,
220083, Grodno, Belarus, tatiana.rusilko@gmail.com*

A mathematical model of the process of processing customer calls in a call center is described. A closed exponential queueing network with impatient customers is used as a model. The purpose of the study is to calculate the performance characteristics of a call center. The problem of finding the optimal staffing list of agents in order to maximize call center performance is investigated.

Keywords: call center; queueing network; mathematical modeling; asymptotic analysis method.

Обзор проблемы. Колл-центр – это специализированная организация или внутреннее штатное подразделение компании, в котором телефонные звонки клиентов обрабатываются группой консультантов, также известных как агенты. Центры обработки вызовов могут обрабатывать входящие и исходящие вызовы с помощью голосовых каналов связи, хотя некоторые из них могут специализироваться лишь на обработке входящих или исходящих звонков [1].

Чтобы удовлетворять потребности клиентов, качество и операционная эффективность телефонных услуг должны быть исключительными [2]. Агентам крупных колл-центров приходится обслуживать тысячи телефонных звонков в час. При этом часто требуется чтобы время ожида-

ния звонков в очереди не превышало нескольких секунд. Для достижения высокого уровня качества и эффективности обслуживания необходимо понимать научные принципы, лежащие в основе работы колл-центров. Несомненно, необходимы сложные подходы для точного описания процесса их функционирования, математическое моделирование этого процесса может значительно улучшить производительность колл-центров [1, 3, 4]. Возможности колл-центров включают также накопление очереди вызовов в режиме ожидания, информирование о времени ожидания, распределение и маршрутизацию вызовов между агентами, классификацию вызовов и определение стратегии взаимодействия в зависимости от статуса клиента.

Качество оперативного обслуживания вызовов агентами колл-центра обычно количественно определяется с помощью некоторых показателей эффективности или показателей производительности. Среди них можно отметить следующие: рабочая нагрузка, размещение вызовов между агентами, ожидаемое время обработки вызовов, процент успешно обработанных вызовов, продолжительность цикла обслуживания одного вызова, доля клиентов, покидающих очередь не дождавшись обслуживания, длина очереди, время ожидания.

Проанализировав организацию колл-центров, приходим к естественному заключению, что для анализа эффективности их работы, прогнозирования загрузки и оптимизации структуры можно использовать теорию сетей массового обслуживания, история возникновения которой связана с решением задач упорядочения и оптимизации работы телефонных станций. Перечисленные выше показатели эффективности также подчеркивают естественное соответствие между моделями массового обслуживания и колл-центрами [5]. Системы массового обслуживания, в частности марковские, широко используются для моделирования и анализа производительности колл-центров [6–10]. Системы массового обслуживания с нетерпеливыми заявками [7], [10], системы массового обслуживания с повторными вызовами [6], [9] и многие другие возникли из-за необходимости математического моделирования поведения клиентов при получении телефонных услуг.

Колл-центры часто имеют разветвленную или многоуровневую структуру. В функции агентов колл-центра первого уровня входит регистрация звонков и предоставление общей информации. Когда тема запроса клиента выходит за рамки специализации агентов текущего уровня, клиент перенаправляется на более высокий уровень, специалисты которого компетентны в решении проблемы. Таким образом, вызовы маршрутизируются в соответствии с их спецификой между разными уровнями агентов. Следовательно, организационная структура центра

обработки вызовов может быть графически изображена в виде сетевой диаграммы, показывающей иерархию агентов центра обработки вызовов, маршрутизацию и переадресацию вызовов в соответствии с их особенностями между различными уровнями консультантов. Таким образом, сети массового обслуживания подходят в качестве математических моделей многоуровневых колл-центров.

Целью данной статьи является математическое моделирование колл-центра, оптимизация и анализ эффективности процесса обработки звонков в колл-центре с использованием замкнутых экспоненциальных сетей массового обслуживания с нетерпеливыми заявками. Модель исследуется в асимптотическом предположении большого числа клиентов колл-центра [11–12], следовательно точность расчетов увеличивается с ростом числа клиентов.

Описание модели. Постановка задачи. Исследуется колл-центр обработки входящих звонков в крупной компании с ограниченным числом клиентов. Обозначим это количество клиентов через K , $K \gg 1$. Такие центры обрабатывают входящие вызовы, инициированные внешними абонентами, звонящими в колл-центр. Обычно центры такого типа предоставляют поддержку клиентам, поддержку бронирования и продаж для авиакомпаний и отелей, а также осуществляют функции приема заказов для интернет-магазинов.

В качестве модели предлагается замкнутая сеть массового обслуживания. Сеть состоит из $n + 1$ марковских m_i –линейных систем массового обслуживания S_i , $i = \overline{1, n+1}$, и K -линейной гипотетической системы S_0 , которая играет роль внешней среды и рассматривается как зависимый источник поступающих заявок. Каждый звонок в колл-центр соответствует поступлению заявки в сеть массового обслуживания. Пребывание заявки в определенном сетевом узле и ее маршрутизация между узлами соответствуют статусу звонка клиента в колл-центре и процессу его маршрутизации между телефонными агентами разных уровней.

Рассмотрим двухуровневый колл-центр компании, который обрабатывает звонки по n темам. Каждый клиент может находиться в одном из следующих состояний:

- S_0 – не нужно обращаться в колл-центр,
- S_{n+1} – требуется обслуживание звонка одним из m_{n+1} агентов первого уровня, которые регистрируют и распределяют звонки в зависимости от их тематики,
- S_i – требуется обслуживание звонка одним из m_i агентов второго уровня, консультирующих по i -й тематике, $i = \overline{1, n}$.

Соответственно, каждая заявка сети массового обслуживания может находиться в одной из СМО S_i , $i = \overline{0, n+1}$.

Переход клиента из состояния S_0 в состояние S_{n+1} соответствует звонку клиента в колл-центр, который происходит в случайное время, независимо от состояния других клиентов, но зависит от времени суток. Вероятность этого перехода на временном интервале $[t, t + \Delta t]$ равна $\lambda(t)\Delta t + o(\Delta t)$, где $\lambda(t)$ – интенсивность поступления вызовов. Таким образом, поток входящих вызовов в общем случае не стационарен.

Из состояния S_{n+1} вызов направляется в состояние S_i с вероятностью $p_{(n+1)i}$, $i = \overline{1, n}$, или с вероятностью $p_{(n+1)0}$ на этап S_0 в случае ошибочного обращения клиента. После завершения обслуживания вызова в состоянии S_i , $i = \overline{1, n}$, клиент переходит в S_0 с вероятностью p_{i0} , если его запрос полностью удовлетворен или перенаправляется специалистам по другой тематике на этап S_j , $j = \overline{1, n}$, с вероятностью p_{ij} , $i \neq j$, если агенты состояния S_i не компетентны ответить на вопросы клиента. Матрица перехода обслуженных заявок представляет собой стохастическую матрицу $(p_{ij})_{(n+2) \times (n+2)}$ со следующими ненулевыми элементами:

$$p_{0(n+1)} = 1; p_{(n+1)i}, i = \overline{0, n}, \sum_{i=0}^n p_{(n+1)i} = 1; p_{ij}, i = \overline{1, n}, j = \overline{0, n}, i \neq j, \sum_{j=0}^n p_{ij} = 1.$$

Маршрутизация обслуженных клиентов в сетевой модели колл-центра изображена сплошной линией на рис. 1.

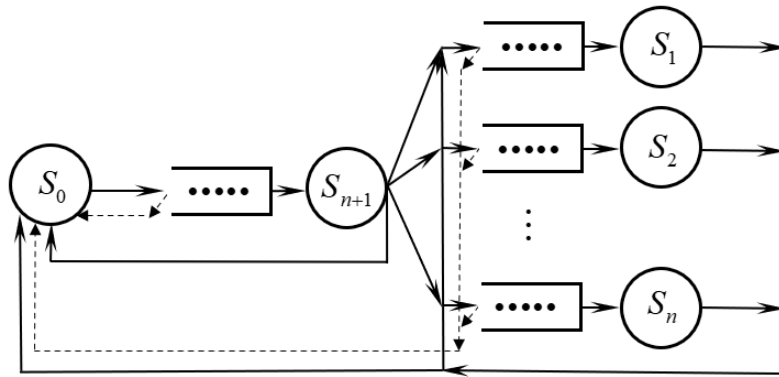


Рис. 1. Граф перехода между СМО обслуженных и нетерпеливых заявок

Если хотя бы одна из линий обслуживания системы S_i бездействует, то входящий звонок сразу начинает обслуживаться в момент перехода в состояние S_i , $i = \overline{1, n+1}$. Звонки обслуживаются в порядке поступления.

Для сети массового обслуживания это означает, что запросы на обслуживание выбираются в соответствии с дисциплиной FIFO. Все m_i агентов (линий) в состоянии (узле) S_i идентичны и имеют экспоненциально распределенное время обслуживания с интенсивностью μ_i , т. е. μ_i^{-1} – среднее время обслуживания, $i = \overline{1, n+1}$.

Если все m_i агентов в S_i заняты, то заказчик ожидает в неограниченной по числу мест очереди. Однако время ожидания клиента в очереди системы S_i ограничено экспоненциальной случайной величиной τ со средним значением η_i^{-1} , $i = \overline{1, n+1}$. В этом случае заявки в сети массового обслуживания называются нетерпеливыми. Если в течение времени τ на звонок не ответили, то он теряется, т. е. с вероятностью, равной единице, заявка по истечении этого времени переходит во внешнюю среду S_0 . Матрица перехода нетерпеливых (неудовлетворенных) заявок имеет вид $(q_{ij})_{(n+1) \times (n+2)}$ со следующими ненулевыми элементами: $q_{i0} = 1$, $i = \overline{1, n+1}$.

Пусть $k_i(t)$ обозначает количество заявок в состоянии S_i в момент времени t , $i = \overline{1, n+1}$. Очевидно, что $\sum_{i=0}^{n+1} k_i(t) = K$. Распределение клиентов по возможным состояниям в момент времени t полностью описывает ситуацию в колл-центре на этот момент. Соответственно, распределение заявок по узлам сети массового обслуживания полностью определяет состояние сетевой модели. Таким образом, состояние исследуемой сетевой модели в момент времени t представлено вектором $k(t) = (k_1(t), k_2(t), \dots, k_{n+1}(t))$.

Принимая во внимание вышеизложенное, процесс $k(t)$ изменения количества заявок в узлах сети представляет собой цепь Маркова с непрерывным временем и конечным числом состояний. Понятно, что количество клиентов во внешней среде S_0 равно $k_0(t) = K - \sum_{i=1}^{n+1} k_i(t)$.

Представляет интерес следующая задача: учитывая прибыль от звонков клиентов и учитывая фонд оплаты труда агентов, необходимо найти такое штатное расписание (набор серверов в узлах сети), которое обеспечило бы максимальную эффективность колл-центра, при условии, что K , $\lambda(t)$, μ_i , η_i , p_{ij} фиксированы $i = \overline{1, n+1}$, $j = \overline{0, n+1}$.

Введем следующие обозначения:

- F_i – доход (денежный эффект или некоторый вид эффективности), генерируемый обслуживанием одного обращения в состоянии S_i за единицу времени, $i = \overline{1, n+1}$;
- L_i – заработная плата одного агента в системе S_i за единицу времени, $i = \overline{1, n+1}$;
- F_0 – доход, создаваемый одним клиентом, когда ему не нужно обращаться в колл-центр (может быть равна нулю).

Тогда общий доход или денежная эффективность колл-центра в момент времени t определяется следующим выражением

$$F(t) = F_0 \left(K - \sum_{i=1}^{n+1} \min(m_i, k_i(t)) \right) + \sum_{i=1}^{n+1} F_i \min(m_i, k_i(t)) - \sum_{i=1}^{n+1} L_i m_i. \quad (1)$$

Поскольку $k_i(t)$ являются компонентами случайного процесса, доход $F(t)$ также является случайным процессом. Применяя оператор математического ожидания M к обеим частям выражения (1), мы получаем следующее выражение для среднего ожидаемого дохода колл-центра:

$$M_F(t) = F_0 K + \sum_{i=1}^{n+1} (F_i - F_0) \min(m_i, M_{k_i}(t)) - \sum_{i=1}^{n+1} L_i m_i, \quad (2)$$

Предполагается, что работа колл-центр имеет смысл, если $M_F(t) > 0$, и его эффективность увеличивается с увеличением $M_F(t)$. Интересно рассмотреть задачу максимизации функционала (2) по количеству серверов m_i в СМО S_i , $i = \overline{1, n+1}$, при условии, ограничений на число заявок в каждом из возможных состояний. В простейшем случае предположения об отсутствии очередей математическая задача максимизации среднего дохода колл-центра в момент времени t задается так:

$$\begin{cases} M_F(t) = F_0 K + \sum_{i=1}^{n+1} (F_i - F_0) M_{k_i}(t) - \sum_{i=1}^{n+1} L_i m_i \rightarrow \max_{m_i, i=\overline{1, n+1}}, \\ M_{k_i}(t) \leq m_i, i = \overline{1, n+1} \end{cases} \quad (3)$$

Понятно, что для решения этой задачи, прежде всего, необходимо определить распределение вероятностей процесса $k(t)$. Только после этого можно определить ожидаемые значения $M_{k_i}(t)$, $i = \overline{1, n+1}$. Распределение $k(t)$ параметрически зависит от количества серверов в системах массового обслуживания.

Асимптотический анализ сетевой модели. Асимптотический анализ подразумевает приближенный метод исследования сети массового обслуживания в предположении большого количества заявок. Математический подход, используемый в данной статье, основан на дискретной модели непрерывного марковского процесса, описанной во многих работах по теории диффузионных марковских процессов [13–15].

Введем $(n+1)$ -вектор вида $I_i = (0, 0, \dots, 0, \overset{i}{1}, 0, \dots, 0)$ и функцию Хевисайда $\theta(x)$. Как упоминалось выше, процесс $k(t)$ представляет собой цепь Маркова с непрерывным временем и конечным числом состояний. Предположения, сделанные при постановке задачи, определяют, что за короткое время Δt марковский процесс $k(t) = (k, t)$ может совершить один из следующих переходов в состояние $(k, t + \Delta t)$:

- из состояния $(k - I_{n+1}, t)$ процесс переходит в $(k, t + \Delta t)$ с вероятностью $\lambda(t)(K - \sum_{i=1}^{n+1} k_i(t) + 1)\Delta t + o(\Delta t)$, что соответствует переходу клиента из состояния S_0 в состояние S_{n+1} ;

- из состояния $(k - I_i, t)$ процесс переходит в $(k, t + \Delta t)$ с вероятностью $\eta_i(k_i(t) + 1 - m_i)\theta(k_i(t) + 1 - m_i)\Delta t + \mu_i p_{i0} \min(m_i, k_i(t) + 1)\Delta t + o(\Delta t)$, что соответствует переходу нетерпеливой клиента из состояния S_i в состояние S_0 , $i = \overline{1, n+1}$;

- из состояния $(k + I_i - I_j, t)$ процесс переходит в $(k, t + \Delta t)$ с вероятностью $\mu_i p_{ij} \min(m_i, k_i(t) + 1)\Delta t + o(\Delta t)$, что соответствует переходу обслуженного запроса из состояния S_i в состояние S_j , $i = \overline{1, n+1}$, $j = \overline{1, n}$;

- из состояния (k, t) в $(k, t + \Delta t)$ с вероятностью

$$1 - (\lambda(t)(K - \sum_{i=1}^{n+1} k_i(t)) + \sum_{i,j=1}^{n+1} \mu_i p_{ij} \min(m_i, k_i(t)) + \sum_{i=1}^{n+1} (\mu_i p_{i0} \min(m_i, k_i(t) + \eta_i(k_i(t) - m_i)\theta(k_i(t) - m_i)))\Delta t + o(\Delta t),$$

- что соответствует отсутствию изменения состояния клиента клиентов;

- из других состояний – в $(k, t + \Delta t)$ с вероятностью $o(t)$.

Вероятности p_{ij} являются элементами матрицы переходных вероятностей $(p_{ij})_{(n+2) \times (n+2)}$, которая была указана в разделе описания модели. Например, $p_{i(n+1)} = 0$, $i = \overline{1, n+1}$.

Принимая во внимание перечисленные выше переходы за короткое время Δt , используя закон полной вероятности, для вероятности $P(k, t) = P(k(t) = k)$ справедлива следующая система уравнений:

$$\begin{aligned} P(k, t + \Delta t) = & P(k - I_{n+1}, t) \lambda(t) \left(K - \sum_{i=1}^{n+1} k_i(t) + 1 \right) \Delta t + \\ & + \sum_{i=1}^{n+1} P(k + I_i, t) (\mu_i p_{i0} \min(m_i, k_i(t) + 1) + \eta_i (k_i(t) + 1 - m_i) \theta(k_i(t) + 1 - m_i)) \Delta t + \\ & + \sum_{i,j=1}^{n+1} P(k + I_i - I_j, t) \mu_i p_{ij} \min(m_i, k_i(t) + 1) \Delta t + P(k, t) \times \\ & \times (1 - (\lambda(t)(K - \sum_{i=1}^{n+1} k_i(t)) + \sum_{i,j=1}^{n+1} \mu_i p_{ij} \min(m_i, k_i(t)) + \\ & + \sum_{i=1}^{n+1} (\mu_i p_{i0} \min(m_i, k_i(t)) + \eta_i (k_i(t) - m_i) \theta(k_i(t) - m_i))) \Delta t) + o(\Delta t). \end{aligned}$$

Обозначим

$$\begin{aligned} \Delta_{i0} P(k, t) &= P(k + I_i, t) - P(k, t), \\ \Delta_{0(n+1)} P(k, t) &= P(k - I_{n+1}, t) - P(k, t), \\ \Delta_{ij} P(k, t) &= P(k + I_i - I_j, t) - P(k, t), i = \overline{1, n+1}, j = \overline{1, n}. \end{aligned}$$

В предположении, что $\Delta t \rightarrow 0$, мы переходим к дифференциальному уравнению Колмогорова для этих вероятностей:

$$\begin{aligned} \frac{\partial p(k, t)}{\partial t} = & \lambda(t) (K - \sum_{i=1}^{n+1} k_i(t)) \Delta_{0(n+1)} P(k, t) + \lambda(t) P(k - I_{n+1}, t) + \\ & + \sum_{i=1}^{n+1} (\mu_i p_{i0} \min(m_i, k_i(t)) + \eta_i (k_i(t) - m_i) \theta(k_i(t) - m_i)) \Delta_{i0} P(k, t) + \\ & + \sum_{i,j=1}^{n+1} \mu_i p_{ij} (\min(m_i, k_i(t) + 1) - \min(m_i, k_i(t))) P(k + I_i, t) + \end{aligned} \quad (4)$$

$$\begin{aligned}
& + \sum_{i=1}^{n+1} \eta_i p_{i0} ((k_i(t) + 1 - m_i) \theta(k_i(t) + 1 - m_i) - (k_i(t) - m_i) \theta(k_i(t) - m_i)) P(k + I_i, t) + \\
& + \sum_{i,j=1}^{n+1} \mu_i p_{ij} \min(m_i, k_i(t)) \Delta_{ij} P(k, t) + \\
& + \sum_{i,j=1}^{n+1} \mu_i p_{ij} (\min(m_i, k_i(t) + 1) - \min(m_i, k_i(t))) P(k + I_i - I_j, t).
\end{aligned}$$

Последнее уравнение не поддается аналитическому решению при больших n . Колл-центры, несомненно, обрабатывают большое количество звонков клиентов. В связи с этим рассмотрим важный асимптотический случай большого количества клиентов $K \gg 1$ и найдем распределение вероятностей вектора состояния модели колл-центра с точностью $O(\varepsilon^2)$, где $\varepsilon = K^{-1}$.

Предположим, что нас интересуют свойства процесса $\xi(t) = \frac{k(t)}{K}$, когда K становится очень большим. Вектор $\xi(t)$ указывает, какая часть клиентов компании связалась с колл-центром и как звонки распределялись по состояниям модели в момент времени t . За время $\Delta t \rightarrow 0$ возможные изменения вектора $\xi(t)$ составляют e_i , где $e_i = I_i \cdot \varepsilon = \frac{I_i}{K}$. В случае, когда $K \rightarrow \infty$ имеем $e_i \rightarrow 0$ и вектор $\xi(t) = \frac{k(t)}{K} = \left(\frac{k_1(t)}{K}, \frac{k_2(t)}{K}, \dots, \frac{k_{n+1}(t)}{K} \right)$ будет марковским процессом с непрерывным временем и непрерывным множеством состояний, функция его плотности распределения вероятностей следующая:

$$\begin{aligned}
p(x, t) &= \lim_{\varepsilon \rightarrow 0} \frac{P(x_1 \leq \xi_1(t) < x_1 + \varepsilon, x_2 \leq \xi_2(t) < x_2 + \varepsilon, \dots, x_{n+1} \leq \xi_{n+1}(t) < x_{n+1} + \varepsilon)}{\varepsilon^{n+1}} = \\
&= \lim_{\varepsilon \rightarrow 0} \frac{P(Kx_1 \leq k_1(t) < Kx_1 + 1, Kx_2 \leq k_2(t) < Kx_2 + 1, \dots, Kx_{n+1} \leq k_{n+1}(t) < Kx_{n+1} + 1)}{\varepsilon^{n+1}},
\end{aligned}$$

т. е.

$$\begin{aligned}
& K^{n+1} P(k, t) \xrightarrow{K \rightarrow \infty} p(x, t), \text{ где } x \in X, \\
& X = \left\{ x = (x_1, x_2, \dots, x_{n+1}) : x_i \geq 0, i = \overline{1, n+1}, \sum_{i=1}^{n+1} x_i \leq 1 \right\}.
\end{aligned} \tag{5}$$

Реализуя асимптотический переход (5) для уравнения (4), получаем следующее уравнение в частных производных:

$$\begin{aligned}
\frac{\partial p(x,t)}{\partial t} = & \lambda(t)K(1 - \sum_{i=1}^{n+1} x_i(t))\Delta_{0(n+1)}p(x,t) + \lambda(t)p(x - e_{n+1},t) + \\
& + K \sum_{i=1}^{n+1} (\mu_i p_{i0} \min(l_i, x_i(t)) + \eta_i(x_i(t) - l_i)\theta(x_i(t) - l_i))\Delta_{i0}p(x,t) + \\
& + \sum_{i=1}^{n+1} \mu_i p_{i0} \frac{\partial \min(l_i, x_i)}{\partial x_i} p(x + e_i, t) + \sum_{i=1}^{n+1} \eta_i p_{i0} \frac{\partial (x_i(t) - l_i)\theta(x_i(t) - l_i)}{\partial x_i} p(x + e_i, t) + \\
& + K \sum_{i,j=1}^{n+1} \mu_i p_{ij} \min(l_i, x_i(t))\Delta_{ij}p(x,t) + \sum_{i,j=1}^{n+1} \mu_i p_{ij} \frac{\partial \min(l_i, x_i)}{\partial x_i} p(x + e_i - e_j, t).
\end{aligned} \tag{6}$$

Если $p(x,t)$ – дважды непрерывно дифференцируемая функция относительно x , то используем разложения в ряд Тейлора функций $p(x \pm e_i, t)$, $p(x + e_i - e_j, t)$, $i, j = 1, n+1$, ограничиваясь членами порядка малости ε^2 . Подставляя эти разложения в уравнение (6) и группируя члены при одинаковых порядках производной $p(x,t)$, получаем, что $p(x,t)$ удовлетворяет многомерному уравнению Фоккера – Планка – Колмогорова (7):

$$\frac{\partial p(x,t)}{\partial t} = - \sum_{i=1}^{n+1} \frac{\partial}{\partial x_i} (A_i(x,t)p(x,t)) + \frac{\varepsilon}{2} \sum_{i,j=1}^{n+1} \frac{\partial^2}{\partial x_i \partial x_j} (B_{ij}(x,t)p(x,t)), \tag{7}$$

с коэффициентами сноса

$$\begin{aligned}
A_i(x,t) = & \sum_{j=1}^{n+1} \mu_j (p_{ji} - \delta_{ji}) \min(l_j, x_j(t)) - \eta_i(x_i(t) - l_i)\theta(x_i(t) - l_i) + \\
& + \lambda(t)p_{0i} \left(1 - \sum_{j=1}^{n+1} x_j(t) \right),
\end{aligned} \tag{8}$$

и коэффициентами диффузии

$$\begin{aligned}
B_{ii}(x,t) = & \sum_{j=1}^{n+1} \mu_j (p_{ji} + \delta_{ji}) \min(l_j, x_j(t)) + \eta_i(x_i(t) - l_i)\theta(x_i(t) - l_i) + \\
& + \lambda(t)p_{0i} \left(1 - \sum_{j=1}^{n+1} x_j(t) \right),
\end{aligned}$$

$$B_{ij}(x, t) = -\mu_i p_{ij}(l_i, x_i(t)), i \neq j;$$

δ_{ij} – дельта Кронекера, $i, j = \overline{1, n+1}$. Погрешность этого приближения не более $O(\varepsilon^2)$.

Известно [16–17], что математическое ожидание $M_{\xi_i}(t)$ определяется с точностью $O(\varepsilon^2)$ из системы обыкновенных дифференциальных уравнений:

$$\frac{dM_{\xi_i}(t)}{dt} = A_i(M_{\xi_i}(t), t), i = \overline{1, n+1},$$

где $A_i(x, t)$ – коэффициенты сноса, определяемы формулой (8). Система дифференциальных уравнений для компонент вектора $k(t) = K\xi(t)$ имеет вид

$$\frac{dM_{k_i}(t)}{dt} = KA_i\left(\frac{1}{K}M_{k_i}(t), t\right), i = \overline{1, n+1}, \quad (9)$$

Принимая во внимание (9), (8) и матрицу переходов $(p_{ij})_{(n+2) \times (n+2)}$ сетевой модели, ожидаемое количество заявок $M_{k_i}(t)$ в состоянии S_i , $i = \overline{1, n+1}$ можно найти, решив следующую систему:

$$\begin{aligned} \frac{dM_{k_i}(t)}{dt} &= \sum_{j=1}^{n+1} \mu_j p_{ji} \min(m_j, M_{k_j}(t)) - \mu_i \min(m_i, M_{k_i}(t)) - \\ &\quad - \eta_i (M_{k_i}(t) - m_i) \theta(M_{k_i}(t) - m_i), i = \overline{1, n}; \\ \frac{dM_{k_{n+1}}(t)}{dt} &= \lambda(t)(K - \sum_{i=1}^{n+1} M_{k_i}(t)) - \mu_{n+1} \min(m_{n+1}, M_{k_{n+1}}(t)) - \\ &\quad - \eta_{n+1} (M_{k_{n+1}}(t) - m_{n+1}) \theta(M_{k_{n+1}}(t) - m_{n+1}). \end{aligned} \quad (10)$$

Ожидаемое количество клиентов во внешней среде составляет $M_{k_0}(t) = K - \sum_{i=1}^{n+1} M_{k_i}(t)$.

Решение системы (10) с определенным начальным условием, во-первых, позволяет прогнозировать среднее количество звонков в каждом состоянии модели, а, во-вторых, позволяет решить задачу максимизации дохода колл-центра (3). Для решения (10) можно использовать численные методы или решить систему аналитически в областях линейности правой части. Решив (10) для разных наборов количества агентов

$m = (m_1, m_2, \dots, m_{n+1})$, подставив это решение в (2) и сравнив значения ожидаемой производительности $M_F(t)$ для каждого $m = (m_1, m_2, \dots, m_{n+1})$, находим оптимальный набор m , для которого $M_F(t)$ принимает наибольшее значение в фиксированный момент t .

Хорошо известно, что при определенных условиях сеть достигает «установившегося состояния», стационарное значение ожидаемого дохода равно $M_F = \lim_{t \rightarrow \infty} M_F(t)$. Стационарные средние $M_{k_i} = \lim_{t \rightarrow \infty} M_{k_i}(t)$, где $i = \overline{1, n+1}$, можно найти из системы алгебраических уравнений баланса вида (10), где левые части равны нулю. В этом случае интенсивность поступления заявок должна быть постоянной: $\lambda(t) = \lambda$. Относительное число

ло $\frac{M_{k_i}}{K}$ численно равно вероятности того, что клиент находится в состоянии (узле) S_i , $i = \overline{1, n+1}$, в стационарном режиме. В стационарном режиме задача (3) превращается в задачу линейного программирования.

Пример. Рассмотрим колл-центр компании, который обрабатывает звонки по трем темам, $n=3$. Пусть число клиентов компании равно $K = 20000$, и работа колл-центра определяется следующими параметрами:

- матрица перехода $P = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0.8 & 0 & 0.1 & 0.1 & 0 \\ 0.9 & 0.03 & 0 & 0.07 & 0 \\ 0.75 & 0.1 & 0.15 & 0 & 0 \\ 0.1 & 0.45 & 0.3 & 0.15 & 0 \end{pmatrix},$

- интенсивность поступления заявок

$$\lambda(t) = 0.002 \cdot \sin\left(\frac{2\pi t}{24} + 2\right) + 0.0083 \text{ или } \lambda = 0.01,$$

- количество серверов в сетевых узлах –

$$m_1 = 3, m_2 = 4, m_3 = 3, m_4 = 2;$$

- интенсивности обслуживания заявок –

$$\mu_1 = 40, \mu_2 = 20, \mu_3 = 15, \mu_4 = 135;$$

- среднее время ожидания в очереди –

$$\eta_1^{-1} = 2, \eta_2^{-1} = 2, \eta_3^{-1} = 2.5, \eta_4^{-1} = 10;$$

- начальное размещение клиентов – $E_{k_0}(t) = 20000, E_{k_i}(t) = 0, i = \overline{1, 4}$.

Решим систему (10) численными методами и представим ее решение графически на рис. 2. На рис. 2 – динамика ожидаемого количества заявок в узлах сети S_i , $i = \overline{1, 4}$, когда интенсивность их поступления в сеть является

функцией времени: $\lambda(t) = 0.002 \cdot \sin\left(\frac{2\pi t}{24} + 2\right) + 0.0083$. В этой ситуации рекомендуется варьировать количество агентов в зависимости от интенсивности поступления звонков. В пиковых ситуациях в узле S_3 есть очередь.

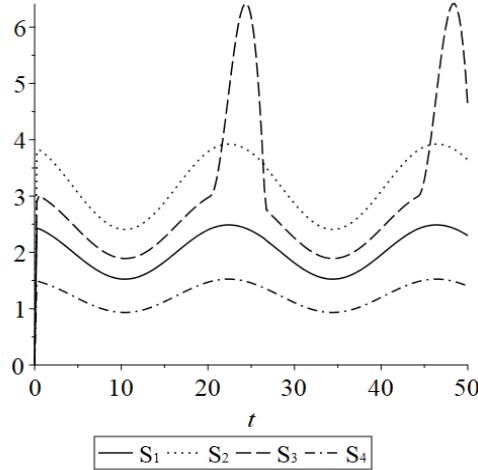


Рис. 2. Динамика $M_{k_i}(t)$, $i = \overline{1, 4}$, интенсивность поступления заявок не постоянна

На рис. 3 показана динамика ожидаемого количества заявок в системах массового обслуживания S_i , $i = \overline{1, 4}$, при постоянной интенсивности поступления заявок: $\lambda = 0.01$. Рис. 3 демонстрирует, что процесс быстро достигает стационарного состояния. Наибольшее количество заявок накапливается в узле S_2 . Очевидно, что серверы в узлах сети в среднем простаивают, и в этой ситуации количество агентов, обрабатывающих вызовы, достаточно при условии отсутствия очередей.

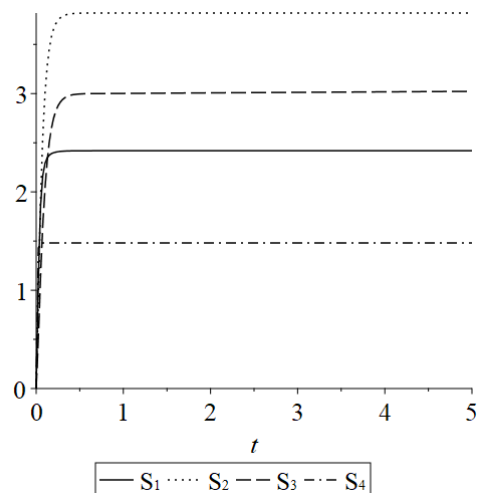


Рис. 3. Динамика $M_{k_i}(t)$, $i = \overline{1, 4}$, интенсивность поступления заявок постоянна

Рассмотрим ожидаемый доход колл-центра $M_F(t)$, который рассчитывается по формуле (2). Рассчитаем значение (2) в предположении, что

$F_0 = 0, F_1 = F_2 = 20, F_3 = 40, F_4 = 10$ $L_1 = L_2 = 0$ $5L_3 = 0$ $L_4 = 7$ в денежных единицах. На рис. 4 показана динамика $M_F(t)$ в случае нестационарного приходящего потока.

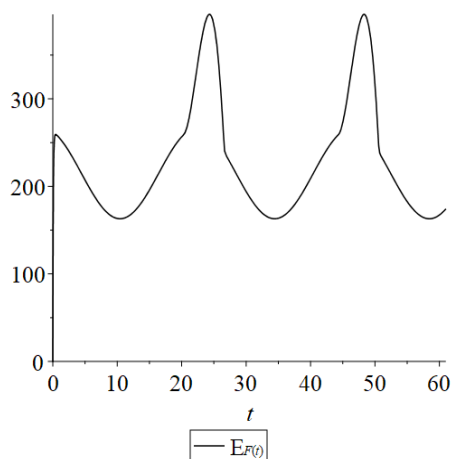


Рис. 4. Динамика $M_F(t)$, когда скорость прибытия зависит от времени

Результаты расчета ожидаемого денежного эффекта колл-центра для разных наборов серверов $m = (m_1, m_2, m_3, m_4)$ при $t = 25$ и стационарном потоке поступления звонков $\lambda = 0.01$ представлены в табл. 1. Как видно из результатов расчета, наиболее оптимальным является набор $m = (3, 4, 3, 2)$, в котором нет очередей и денежная эффективность максимальна.

Таблица 1

Средний денежный эффект колл-центра в момент времени $t=25$

No.	m_1	m_2	m_3	m_4	$M_F(25)$
1	3	4	3	2	251.59
2	4	4	3	2	251.54
3	3	3	3	2	229.97
4	2	4	3	2	234.67
5	3	4	2	2	207.09

Таким образом, результаты расчетов могут быть полезны при анализе и принятии решений относительно работы колл-центра при разных входных параметрах.

Библиографические ссылки

1. *Gans N., Koole G., Mandelbaum A.* Telephone call centers: tutorial, review, and research prospects // *Manufacturing & service operations management*. 2003. Vol. 5. Iss. 2. P. 79–141.
11. *Bennington L., Cummane J., Conn P.* Customer satisfaction and call centers: an Australian study // *International journal of service industry management*. 2000. Vol. 11. Iss. 2. P. 162–173.
12. *Stolletz R.* Performance Analysis and Optimization of Inbound Call Centers. Berlin, Heidelberg: Springer, 2003.
13. *Srinivasan R., Talim J., Wang J.* Performance analysis of a call center with interactive voice re-sponse units // *Top*. 2004. Vol. 12. P. 91–110.
14. *Koole G., Mandelbaum A.* Queueing models of call centers: an introduction // *Annals of operations research*. 2002. Vol. 113. P. 41–59.
15. *Aguir S., Karaesmen F., Akşin O., Chauvet F.* The impact of retrials on call center performance // *OR Spectrum*. 2004. Vol. 26. P. 353–376.
16. *Zeltyn S., Mandelbaum A.* Call centers with impatient customers: many-server asymptotics of the M/M/n + G queue // *Queueing Systems*. 2005. Vol. 51. P. 361–402.
17. *Kim C., Dudina O., Dudin A., Dudin S.* Queueing system MAP/M/N as a model of call center with call-back option // *Proceedings of 19th international conference on Analytical and Stochastic Modeling Techniques and Applications, ASMTA 2012, Lecture Notes in Computer Science, LNCS*. Berlin, Heidelberg: Springer, 2012, P. 1–15.
18. *Pustova S. V.* Investigation of call centers as retrial queueing systems // *Cybernetics and System Analysis*. 2010. Vol. 46. P. 494–499.
19. *Iravani F., Balcioglu B.* On priority queues with impatient customers // *Queueing Systems*. 2008. Vol. 58. P. 239.
20. *Rusilko T. V.* Network Stochastic Call Center Model // *CEUR workshop proceedings*. Vol. 3057 : selected papers of the 6th International scientific and practical conference "Distance learning technologies". Yalta : [s. n.], 2021. P. 91–101.
21. *Rusilko T. V., Pankov A. V.* Queueing network model of a call center with customer retrials and impatient customers // *Известия Саратовского университета. Новая серия. Сер. Математика. Механика. Информатика*. 2024. Т. 24. № 2. С. 287–297.
22. *Kobayashi H.* Application of the diffusion approximation to queueing networks // *Journal of ACM*. 1974. V. 21. P. 316–328.
23. *Медведев Г. А.* Замкнутые системы массового обслуживания и их оптимизация // *Известия АН СССР. Техническая кибернетика*. 1978. № 6. С. 199–203.
24. *Матальцкий М. А., Романюк Т. В.* Приближенные методы анализа сетей с центральной системой обслуживания и их применения. Гродно : ГрГУ, 2003.
25. *Тихонов А. Н., Миронов М. А.* Марковские процессы. М.: Советское радио, 1977.
26. *Русилко Т. В.* Метод определения моментов первых двух порядков для вектора состояния сети массового обслуживания в асимптотическом случае // *Веснік ГрДУ імя Янкі Купалы. Сер. 2, Матэматыка. Фізіка. Інфарматыка, вылічальная тэхніка і кіраванне*. 2021. Т. 11. № 2. С. 152–161.