

ОПТИМИЗАЦИЯ РАЗМЕЩЕНИЯ И ПЕРЕДИСЛОКАЦИИ БРИГАД СКОРОЙ МЕДИЦИНСКОЙ ПОМОЩИ

Ю. А. Кочетов^{1, a}, Н. Б. Шамрай^{2, b}

¹ Институт математики им. С. Л. Соболева,
пр. Акад. Коптюга, 4, 630090 Новосибирск, Россия

² Институт автоматики и процессов управления,
ул. Радио, 5, 690041 Владивосток, Россия

E-mail: ^a jkochet@math.nsc.ru, ^b shamray@dvo.ru

Аннотация. Рассматривается задача размещения врачебных бригад по станциям скорой медицинской помощи (СМП) с целью минимизации среднего времени ожидания их прибытия к пациентам. Построена имитационная модель работы службы СМП, учитывающая стохастический характер проблемы и меняющуюся загруженность транспортной сети города. Для решения задачи разработан алгоритм генетического локального поиска с четырьмя типами окрестностей. Имитационная модель в этом алгоритме используется для вычисления целевой функции. Исследовано влияние окрестностей на точность получаемого решения. Проведено компьютерное моделирование работы службы СМП на примере г. Владивостока. Показана возможность снижения среднего времени ожидания в 1,5 раза. Получены оценки влияния загруженности дорог на среднее время ожидания. Табл. 4, ил. 11, библиогр. 28.

Ключевые слова: оптимизация, служба скорой медицинской помощи, имитационное моделирование, генетический алгоритм, локальный поиск.

Введение

Служба скорой медицинской помощи (СМП) — важный компонент системы здравоохранения. Её главной задачей является оказание первой медицинской помощи и, при необходимости, транспортировка больного в стационар. Оперативность работы службы критическим образом влияет на выживаемость и дальнейший процесс выздоровления пациентов.

Работа первого автора выполнена в рамках гос. задания ИМ СО РАН (проект № 0314–2019–0014), второго — при поддержке Российского фонда фундаментальных исследований (проект № 18–29–03071).

Время ожидания прибытия бригады на вызов — один из ключевых показателей эффективности её работы. В данной статье рассматривается задача минимизации среднего времени ожидания за счёт выбора наилучшего размещения бригад на станциях СМП. Предполагается, что в период наблюдения меняются интенсивность поступления вызовов и уровень загруженности улично-дорожной сети (УДС). Количество и месторасположение станций СМП фиксировано. Задано общее число работающих бригад в смене, к каждой станции приписаны свои бригады и закреплена приоритетная территория обслуживания. Свободная бригада ждёт назначения очередного вызова на станции приписки, промежуточные пункты ожидания не рассматриваются.

Эффективная организация работы службы СМП — комплексная задача, включающая решение проблем на стратегическом, тактическом и оперативном уровнях. В частности, на стратегическом уровне определяется наилучшее расположение станций СМП, на тактическом уровне — оптимальное распределение бригад по станциям, на оперативном уровне выбирается сценарий диспетчеризации карет СМП. В работе [1] представлен обзор логистических задач, с которыми сталкиваются службы СМП разных стран. В [2] описана эволюция детерминированных, вероятностных и динамических задач размещения парка карет СМП. Хронологию развития моделей и систем принятия решений для служб СМП можно проследить по обзору [3]. В [4] акцент сделан на задачи экстремального покрытия и оптимизационных методах их решения. В работе [5] основное внимание уделяется проблемам диспетчеризации, маршрутизации, прогнозирования спроса и планирования рабочей силы служб СМП. Обзор [6] сконцентрирован на задачах принятия решений о допустимом местоположении и перемещении карет СМП, описываются разные сценарии диспетчеризации и их влияние на целевые показатели. Проблема поиска наилучшего плана передислокации бригад между станциями СМП исследуется в [7, 8].

Процесс принятия решений в обозначенной проблеме происходит в динамической среде, где пространственное распределение запросов на медпомощь и дорожная ситуация меняются с течением времени. Эта же среда является стохастической — вызовы поступают случайным образом из разных районов в разном количестве, время обслуживания каждого вызова заранее не определено. Учёт всех внешних и внутренних факторов, влияющих на работу службы СМП, при расчёте среднего времени ожидания требует создания имитационной модели, которая используется в оптимизационных алгоритмах. Достоинства такого подхода отмечены в [9] при определении наилучших локаций станций СМП в г. Остин, США. В работе [10] данный подход используется для анализа эффективности работы СМП при оптимизации разных целевых показателей. В [11]

описан опыт реализации имитационной модели в программном комплексе поддержки принятия решений, используемом одной из служб СМП в Новой Зеландии. Размещение станций СМП и распределение по ним бригад службы СМП Лондона оптимизировались в [12] путём синтеза имитационной модели и генетического алгоритма. Аналогичный подход применялся в [14] для решения задачи распределения бригад по станциям СМП Шанхая. Авторы выделяют несколько следующих друг за другом независимых периодов наблюдения. Для каждого периода рассчитывается свой вариант размещения бригад. Зная варианты для двух соседних периодов, строится обобщённая задача о назначениях. По результатам её решения определяется вариант передислокации бригад по станциям. Время движения по УДС предопределено для каждого периода и представляет собой усреднённые данные наблюдений. Практика использования имитационных моделей для компьютерной симуляции работы служб СМП приведена в обзоре [13].

В настоящей работе имитационная модель имеет больше возможностей для описания условий работы СМП по сравнению с аналогичной моделью из [14]. В частности, работа службы моделируется в непрерывном режиме на протяжении всего периода наблюдений. Для передислокации бригад определяются моменты времени, когда диспетчерская принимает решение о возможном изменении станции приписки для каждой бригады. Если такое изменение было принято, то после окончания обслуживания текущего вызова бригада должна будет вернуться на новую назначенную ей станцию. Преимуществом такой реализации, по сравнению с разбиением суток на независимые периоды моделирования, является большая реалистичность описываемых процессов и, как следствие, более точная оценка целевого показателя. При этом в моменты принятия решений о передислокации бригад сохраняются 1) их текущая занятость и местонахождение; 2) очередь вызовов, ожидающих обслуживания; 3) количество вызовов, уже обслуженных каждой бригадой (этот показатель отвечает за равномерное распределение рабочей нагрузки на врачей) и другие наблюдаемые параметры. Кроме того, загрузка УДС определяется для разных моментов времени путём решения реализованной внутри имитационной модели задачи транспортного равновесия с эластичным спросом [15, 16]. Такой подход позволяет оценить влияние активности использования личного автотранспорта населением на работу службы СМП и строить разные сценарии размещения бригад по станциям в зависимости от транспортной ситуации в городе.

Для решения поставленной задачи предложен алгоритм генетического локального поиска. Построены четыре типа окрестностей и проведены численные эксперименты по оценке вычислительной эффективности работы алгоритма при их использовании. Компьютерное моделирование

проведено на статистических и тестовых данных для службы СМП г. Владивостока. Для разных уровней загруженности УДС определялось два решения: вариант размещения бригад на двенадцатичасовой период работы смены и варианты с возможной передислокацией между утренним, дневным и вечерним периодами. Результаты расчётов показали возможность сократить в 1,5 раза среднее время прибытия на вызовы по сравнению с существующим вариантом распределения бригад по станциям. Передислокация бригад в утренние, дневные и вечерние периоды радикально не меняет ситуацию по сравнению с расчётным вариантом без передислокации. Наблюдается лишь незначительное снижение времени прибытия, в среднем на 4%. Основной причиной малой эффективности введения передислокации видится высокая загруженность бригад и, как следствие, малое число возвратов на станции приписки. Транспортная ситуация сильно влияет на решение задачи, но наблюдаемый эффект от передислокации остаётся при любом уровне загруженности дорог. Если же более половины экономически активного населения будет использовать личный автотранспорт, то среднее время ожидания бригад становится неприемлемо большим. Бригады не возвращаются на свои станции, сразу едут с одного вызова на другой, и передислокация между утренним, дневным и вечерним периодами эффекта не даёт. В этом случае требуется больше бригад для обслуживания пациентов.

Статья организована следующим образом. В разд. 1 приводится содержательная постановка задачи. В разд. 2 вводятся обозначения и описание математической модели. Разд. 3 посвящён имитационной модели и расчётам загрузки транспортной сети. Здесь вместо сбора эмпирических данных о загрузке УДС решается задача потокового равновесия с эластичным спросом и балансовыми ограничениями на объёмы въезда и выезда. В разд. 4 даётся описание алгоритма генетического локального поиска. В разд. 5 определяются окрестности локального поиска, а разд. 6 содержит результаты численных экспериментов и выводы.

1. Содержательная постановка задачи

Задана территория, население которой обслуживается службой СМП. Служба включает множество станций СМП. За каждой станцией закреплены свои врачебные бригады и приоритетная зона обслуживания. Бригады выезжают на вызовы и возвращаются на станцию приписки для ожидания следующего назначения. Запросы на оказание помощи поступают в единую диспетчерскую службу. Диспетчер оценивает повод для вызова и текущую занятость бригад. Он либо помещает вызов в очередь ожидания, либо передаёт вызов бригаде на исполнение. По приезде бригада оказывает пациенту первую помощь и, при необходимости, везёт больного в стационар. На этом бригада заканчивает обслуживание

вызова, сообщает об этом в диспетчерскую и далее либо возвращается на станцию приписки, либо выезжает на следующий вызов. На рис. 1 приведена блок-схема организации работы службы СМП.



Рис. 1. Общая схема диспетчеризации бригад и ключевые временные периоды в работе службы скорой медицинской помощи

При назначении бригады на вызов диспетчер руководствуется следующим набором инструкций:

- 1) определить свободные бригады станции, закреплённой за зоной вызова;
- 2) если такие бригады есть, то назначить бригаду с ближайшим местонахождением; если таких несколько, то вызов примет бригада с меньшим количеством уже отработанных вызовов;
- 3) если таковых нет, то назначить свободную бригаду, ближайшую к месту вызова;
- 4) если нет свободных бригад, то вызов будет передан первой освободившейся бригаде любой станции.

На вызов назначается свободная бригада по профилю. Если все профильные бригады заняты, то диспетчер действует по инструкциям о взаимозаменяемости бригад.

Эффективность работы службы СМП будем оценивать по среднему времени ожидания прибытия бригад на вызовы в наблюдаемый период.

При этом по каждой станции будем считать своё среднее время ожидания и в качестве значения целевого показателя брать максимальное время среди всех станций. Определим *время ожидания* как время, прошедшее с момента принятия вызова диспетчером до прибытия бригады к месту вызова. Оно складывается из следующих временных затрат (см. рис. 1):

времени диспетчеризации — время, необходимое диспетчеру для принятия решения;

транспортных задержек — время на проезд к месту вызова.

Время обслуживания начинается с момента прибытия бригады к пациенту и длится до окончания его полного обслуживания. Оно складывается из следующих временных затрат (см. рис. 1):

времени первой помощи — время, проведённое бригадой на месте вызова;

транспортных задержек — время на проезд от места вызова до стационара, если требуется госпитализация;

времени госпитализации — время на передачу больного сотрудникам стационара.

Времена диспетчеризации, первой помощи и госпитализации заранее неизвестны и являются случайными величинами. Рассмотрим два ключевых динамично меняющихся внешних фактора, оказывающих влияние на эффективность работы службы СМП — это интенсивность поступления вызовов и загруженность УДС автотранспортом. За длительный период наблюдения отмечаются некоторые закономерности. В утренние часы интенсивность поступления вызовов непрерывно растёт, в обеденный период остаётся на одном уровне и продолжает расти в вечерние часы. Для УДС характерны утренние и вечерние пики транспортной загруженности, когда осуществляется так называемая трудовая миграция: люди едут на работу утром и возвращаются домой вечером. Будем предполагать, что для каждой зоны обслуживания известна интенсивность поступления вызовов в течение суток. Вызовы делятся на три категории: экстренные, срочные и неотложные. Пациенты по экстренным вызовам госпитализируются в обязательном порядке, остальные с заданной вероятностью. Время диспетчеризации и время оказания помощи не зависят от категории вызова и являются случайными величинами с заданным распределением. Время госпитализации также является случайной величиной с заданным распределением.

Транспортные задержки карет СМП будем определять по результатам решения задачи потокового равновесия с эластичным спросом и балансовыми ограничениями [15, 16] на каждый час наблюдаемого периода. Принцип потокового равновесия широко используется для прогнозирования загруженности УДС в практике современного транспортно-го моделирования [17–19]. УДС опишем в виде ориентированного графа

с заданными пропускными способностями дуг и временем свободного проезда по ним. Изменяя параметры трудовой миграции в течение периода наблюдения, будем рассчитывать загруженность УДС частным автотранспортом и транспортные задержки бригад СМП.

Таким образом, содержательная постановка задачи состоит в следующем. При сделанных выше предположениях о правилах работы службы СМП определить вариант размещения врачебных бригад по станциям с минимальным средним временем ожидания. Будем искать два типа решения: размещение на весь период наблюдения и вариант размещения с возможной передислокацией бригад по станциям в заданные моменты времени. Ключевым являлся вопрос: можно ли за счёт передислокации сократить время прибытия на вызовы?

2. Математическая модель

Пусть n — количество станций СМП с известным расположением, Q — общее число работающих бригад на станциях. Интервал наблюдения разделим на T периодов $t = 1, 2, \dots, T$, для каждого из которых будем определять свой вариант размещения бригад на станциях. При $T = 1$ получаем задачу поиска оптимального размещения бригад без передислокации. Определим целочисленные неотрицательные переменные x_i^t — количество бригад, приписанных к станции i в период времени t . Множество допустимых решений X задаётся равенствами:

$$\sum_{i=1}^n x_i^t = Q, \quad t = 1, 2, \dots, T. \quad (1)$$

Построим векторы

$$x^t = (x_i^t \in \mathbb{Z}_+ \mid i = 1, 2, \dots, n), \\ x = (x^t \mid t = 1, 2, \dots, T).$$

К началу периода t состояние службы СМП, которое обозначим через S_t , задаётся следующими параметрами:

- 1) местонахождение и статус (занята/свободна) каждой бригады;
- 2) очередь вызовов, ожидающих обслуживания (сформированная за предыдущий период).

Начальное состояние каждого следующего периода формируется по результатам работы службы за предыдущий период, т. е. $S_t = S_t(S_{t-1})$.

Обозначим через $\varphi_i(x^t, S_t)$ среднее время ожидания бригад, приписанных к станции i , в течение периода t . Эффективность работы службы СМП будем оценивать по среднему времени ожидания всех бригад за всё время их работы, выбирая в каждый период t станцию с максимальным значением показателя $\varphi_i(x^t, S_t)$ и вычисляя среднее по этим величинам.

Другими словами, в качестве минимизируемой целевой функции $\Phi(x)$ рассматривается

$$\Phi(x) = \frac{1}{T} \sum_{t=1}^T \max_{i=1, \dots, n} \varphi_i(x^t, S_t). \quad (2)$$

Стохастический характер рассматриваемой проблемы и динамично изменяющиеся внешние условия затрудняют построение аналитической формы записи функций $\varphi_i(x^t, S_t)$ и, как следствие, целевого показателя $\Phi(x)$. Поэтому в следующем разделе приводится описание имитационной модели для получения статистических оценок наблюдаемых показателей.

3. Имитационная модель

Для заданного допустимого вектора x будем имитировать оперативную работу службы СМП на разных возможных сценариях поступления вызовов в рассматриваемый период. Сценарии генерируются случайным образом и задаются следующими величинами:

- 1) вызовами на обслуживание с указанием типа (экстренный, срочный, неотложный), времени поступления в диспетчерскую СМП, местонахождения на рассматриваемой территории;
- 2) временем диспетчеризации и временем оказания первой помощи для каждого вызова;
- 3) принимающим больному стационаром и временем госпитализации (если требуется).

Основные этапы работы имитатора состоят из следующих шагов.

1. Загрузить сценарий поступления вызовов на медпомощь.
2. В процессе обработки вызовов выполнять процедуры:
 - 2.1) рассчитать прогнозную загрузку УДС на текущий период;
 - 2.2) в определённые моменты времени принять решение о перемещении бригад;
 - 2.3) реализовать диспетчеризацию бригад.
3. Оценить среднее время ожидания по каждой станции СМП.

На этапе 2.1 рассчитывается загрузка УДС транспортными потоками и оценивается среднее время проезда карет СМП по каждой дуге сети. Прогноз строится по результатам решения задачи потокового равновесия с эластичным спросом и дополнительными балансовыми ограничениями на объёмы выезда и въезда [15, 16]. УДС описывается следующими входными параметрами: A и V — множества дуг и вершин (территориальных зон) сети; $\delta_{va} = \{0, \pm 1\}$ — признак инцидентности вершин $v \in V$ и дуги $a \in A$; c_a и τ_a — пропускная способность и время свободного проезда

по дуге $a \in A$; $O \subseteq V$ и $D \subseteq V$ — вершины, порождающие и притягивающие транспортные потоки; o_i — объём выезда из вершины $i \in O$, d_j — объём въезда в вершину $j \in D$ такие, что $\sum_{i \in O} o_i = \sum_{j \in D} d_j$. Введём неотрицательные переменные: z_{aij} — величина потока, идущего по дуге $a \in A$ из $i \in O$ в $j \in D$; ρ_{ij} — общий объём потока, направляющегося из $i \in O$ в $j \in D$ (транспортная корреспонденция). Допустимая область задаётся условиями сохранения потоков

$$\sum_{a \in A} \delta_{va} z_{aij} = \{0, \pm \rho_{ij}\}, \quad v \in V, i \in O, j \in D, \quad (3)$$

и балансовыми ограничениями на объёмы выезда и въезда:

$$\sum_{j \in D} \rho_{ij} = o_i, \quad \sum_{i \in O} \rho_{ij} = d_j, \quad i \in O, j \in D. \quad (4)$$

Предполагаем, что транспортные потоки формирует трудоспособное население, использующее личный автомобиль для передвижения по сети. Водители независимо друг от друга выбирают пути следования, соответствующие их минимальным транспортным задержкам (первый принцип Вардропа [20]), т. е. потоковое равновесие соответствует конкурентному бескоалиционному равновесию по Нэшу. Пусть транспортные задержки по каждой дуге $a \in A$ зависят только от общего потока по этой дуге $z_a = \sum_{(i,j) \in O \times D} z_{aij}$ и описываются широко используемой в транспортном моделировании так называемой BPR-функцией [21]. Минимальные транспортные задержки на проезд из вершины $i \in O$ в вершину $j \in D$ определяются из гравитационной модели построения матрицы корреспонденций [22, 23]. Тогда потоковое равновесие в сети — это решение задачи минимизации целевой функции (см., например, [24]):

$$\sum_{a \in A} \tau_a z_a \left(1 + 0,2 \left(\frac{z_a}{c_a} \right)^4 \right) - \sum_{(i,j) \in O \times D} \rho_{ij} \left(1 + \ln \left(\frac{o_i d_j}{\rho_{ij}} \right) \right) \quad (5)$$

на допустимой области (3), (4). Изменяя объёмы выезда o_i и въезда d_j , можно моделировать разные варианты загруженности УДС, что, в свою очередь, повышает практическую ценность имитатора для принятия эффективных решений по диспетчеризации бригад СМП в соответствии с транспортной ситуацией на дорогах.

Для решения задачи (3)–(5) адаптируются методы гладкой выпуклой оптимизации, например, метод Франка — Вульфа, методы проекции градиента, двойственные методы. Обзор таких методов приведён в [24]. Современный рынок программных продуктов позволяет решить задачу (3)–(5) при помощи специализированных оптимизационных пакетов. Например, в рамках данной работы задача решалась пакетом MINOS.

На этапе 2.2 решается задача передислокации бригад между станциями в зависимости от транспортной ситуации. В качестве критерия оптимизации рассматривается суммарное расстояние для передислокации при переходе от вектора x^t к вектору x^{t+1} . По сути решается следующая задача о назначениях:

$$\begin{aligned} \sum_{j=1}^Q \sum_{i=1}^n c_{ji}(t) y_{ji} &\rightarrow \min, \\ \sum_{j=1}^Q y_{ji} &= x_i^{t+1}, \quad i = 1, 2, \dots, n, \\ \sum_{i=1}^n y_{ji} &= 1, \quad j = 1, 2, \dots, Q, \\ y_{ji} &\in \{0, 1\}, \quad j = 1, 2, \dots, Q, i = 1, 2, \dots, n, \end{aligned} \tag{6}$$

где $c_{ji}(t)$ — время проезда от текущего местонахождения бригады j до станции i с учётом загрузки сети к концу периода t , переменные задачи y_{ji} принимают значение 1, если бригада j получает приписку к станции i , и 0 в противном случае. Оптимальные значения y_{ji} определялись с помощью специализированного программного модуля PuLP для решения задач линейной и целочисленной оптимизации.

4. Алгоритм генетического локального поиска

Идеи генетических алгоритмов впервые были предложены в работах Фогеля [25] и Холланда [26] по эволюционному моделированию. Почти в то же время независимо от них А. Г. Ивахненко и Л. А. Растргиным были предложены методы группового учёта аргументов [27] и случайного поиска [28], созвучные идеям эволюции и случайных мутаций. При решении задач дискретной оптимизации основная схема такого подхода состоит в построении начального множества допустимых решений задачи (начальной популяции) и её развития за счёт отбора родительских решений и применения к ним рандомизированных процедур, имитирующих мутацию генов и скрещивание хромосом. В алгоритмах генетического локального поиска к каждому новому решению, полученному таким способом, применяются алгоритмы локального улучшения. Полученный локальный оптимум проверяется на эксклюзивность для исключения клонов и добавляется в популяцию. Гибкость и удивительная эффективность такой схемы показана в многочисленных приложениях и является источником вдохновения для огромного числа разработчиков (см., например, труды конференции EvoCOP <http://www.evostar.org/2021/evocop/>).

Основной трудностью для применения данного подхода к решению задачи размещения и передислокации бригад СМП является стохастический характер вызовов, быстро меняющаяся нагрузка транспортной сети и непредсказуемость во времени обслуживания пациентов. Предложенная в предыдущем разделе имитационная модель будет применяться в генетическом алгоритме для оценки функции $\Phi(x)$ (определённой в (2)) — среднего времени ожидания при размещении бригад по варианту x . Усреднение будет проводиться по результатам многократной работы имитатора на разных возможных сценариях поступления вызовов (в данной работе рассмотрено 500 сценариев).

Пусть набор допустимых решений задачи $\{x^1, x^2, \dots, x^l\}$ задаёт начальную популяцию из l элементов. Тогда общая схема алгоритма генетического локального поиска может быть представлена следующим образом.

1. Построить начальную популяцию и вычислить значения целевой функции для каждого её элемента.
2. Повторять, пока не будет выполнен критерий остановки:
 - 2.1) выбрать родительские решения x', x'' ;
 - 2.2) построить по ним новое решение x ;
 - 2.3) модифицировать решение x и получить из него локальный оптимум \bar{x} ;
 - 2.4) Если решения \bar{x} нет в популяции, то добавить его и удалить решение с наибольшим значением целевой функции.
3. Предъявить наилучшее решение в популяции.

Начальная популяция строится случайным образом: каждая бригада получает приписку к одной из станций, выбираемой с равномерным распределением независимо от других бригад. Критерием остановки является либо достижение максимального числа итераций, либо стагнация процесса, т. е. отсутствие смены рекорда за фиксированное число итераций. На шаге 2.1 всегда выбираются два разных родительских решения. Один из них выбирается по турнирной селекции: формируется случайное подмножество элементов популяции и берётся лучший из них. Второе решение выбирается по одному из следующих трёх правил:

- 1) наилучший элемент в популяции;
- 2) по турнирной селекции;
- 3) случайный элемент, не совпадающий с первым родителем.

Численные эксперименты показали, что случайный порядок применения этих правил приводит к лучшим результатам, чем использование каждого из них в отдельности.

На шаге 2.2 скрещивание родительских решений также проводится по одному из трёх правил, применяемых в случайном порядке:

1) *одноточечный кроссовер*: случайным образом с равномерным распределением среди целых чисел от 1 до n выбирается координата k , и потомок получает первые k координат от первого родительского решения, а остальные — от второго;

2) *равномерный кроссовер*: потомок получает в каждой координате среднее арифметическое от родительских решений;

3) *эвристический кроссовер* [12]: случайным образом выбираются два непересекающихся подмножества координат I, J и параметр ν на отрезке $[0, 1]$; потомок формируется следующим образом: $x(I) = x''(I)$, $x(J) = \nu x'(J) + (1 - \nu)x''(J)$, а на остальных координатах потомок совпадает с первым родительским решением.

Применение разных вариантов кроссоверов приводит к большему разнообразию в популяции и уменьшает вероятность появления «клонов». Заметим, что в ходе скрещивания могут получаться недопустимые решения с дробными компонентами, поэтому к потомку применяется процедура нормализации и округления для получения допустимого решения.

На шаге 2.3 с априори заданной вероятностью значение каждой координаты потомка меняется на случайное число от 0 до Q . Полученное решение приводится к допустимому и используется как стартовое решение для получения локального оптимума по одной из четырёх окрестностей, описание которых приводится в следующем разделе.

5. Окрестности локального поиска

При вычислении целевой функции $\Phi(x)$ имитационная модель выдаёт подробную информацию о работе каждой бригады. В частности, для каждого периода времени t известно среднее число поступивших вызовов α_{il}^t , территориально принадлежащих станции l , но обслуженных бригадами станции i . По сути, это матрица взаимопомощи между станциями. Она будет использоваться при построении окрестностей. Для станции i определим вектор

$$\Delta_i^t = \sum_{l=1}^n \alpha_{li}^t - \sum_{l=1}^n \alpha_{il}^t, \quad t = 1, 2, \dots, T,$$

полагая $\alpha_{ii}^t = 0$. Будем использовать этот вектор как признак избытка или дефицита бригад в каждый период времени. Обозначим через γ среднее число вызовов одной бригады.

Окрестность $N_1(x)$ состоит из одного элемента $\bar{x}_i^t = x_i^t + [\frac{1}{\gamma}\Delta_i^t]$, $i = 1, \dots, n$, $t = 1, 2, \dots, T$. Бригады перераспределяются согласно матрице $\{\Delta_i^t\}$.

Выделим для периода t множество станций с дефицитом бригад $S_1^t = \{i \mid \Delta_i^t > 0\}$, и пусть $S_2^t = \{i \mid \Delta_i^t \leq 0\}$ — все остальные станции.

Окрестность $N_2(x)$ состоит из всех допустимых решений, получаемых из решения x изъятием одной бригады из станции множества S_2^t и добавлением её к одной из станций множества S_1^t для каждого периода времени t . Мощность этой окрестности оценивается величиной $O(Tn^2)$.

Положим $\beta_{ij}^t = \alpha_{ij}^t - \alpha_{ji}^t$ для всех пар i, j , $i \neq j$, и всех периодов t . Это показатель помощи станции i для станции j .

Окрестность $N_3(x)$ состоит из всех решений, получаемых из решения x выбором одной станции, скажем, i , и перераспределением бригад между ней и другими станциями пропорционально величинам β_{ij}^t для каждого периода времени, т. е.

$$\bar{x}_j^t = \left\{ x_j^t + \left[\frac{1}{\gamma} \beta_{ij}^t \right] \mid j = 1, 2, \dots, n, t = 1, 2, \dots, T \right\}$$

в предположении, что $\beta_{ii}^t = - \sum_{j \neq i} \beta_{ij}^t$, $i = 1, 2, \dots, n$, $t = 1, 2, \dots, T$. Эта окрестность состоит ровно из n элементов.

Окрестность $N_4(x)$ состоит из всех решений, получаемых из решения x выбором двух станций и передачей одной бригады с одной станции на другую, если соответствующий коэффициент α_{ij}^t положителен. Мощность этой окрестности оценивается величиной $O(Tn^2)$. Как показали вычислительные эксперименты, коэффициент α_{ij}^t почти всегда оказывался положительным, т. е. окрестность $N_4(x)$ является фактически самой большой из рассматриваемых окрестностей.

6. Численные эксперименты

6.1. Подготовка входных данных. Для проведения численных экспериментов использовались статистические данные работы службы СМП г. Владивостока. Территория города была поделена на зоны размером $3,2 \times 3,2$ км², в выделенных зонах локализованы станции СМП и стационары, включая родильные дома города (рис. 2). За каждой станцией закреплены приоритетные зоны обслуживания. Согласно отчётной статистике ГБУЗ «Станции скорой медицинской помощи г. Владивостока» новая смена приступает к работе в 8:00 и состоит из 26 медицинских бригад. Время работы бригад — 12 часов. Среднегодовое количество вызовов по часам показано на рис. 3.

Согласно плотности жилой застройки каждой зоны можно оценить число её жителей. Будем предполагать, что вызовы поступают независимо друг от друга, их количество пропорционально числу жителей, а интенсивность поступления звонков подчинена пуассоновскому закону

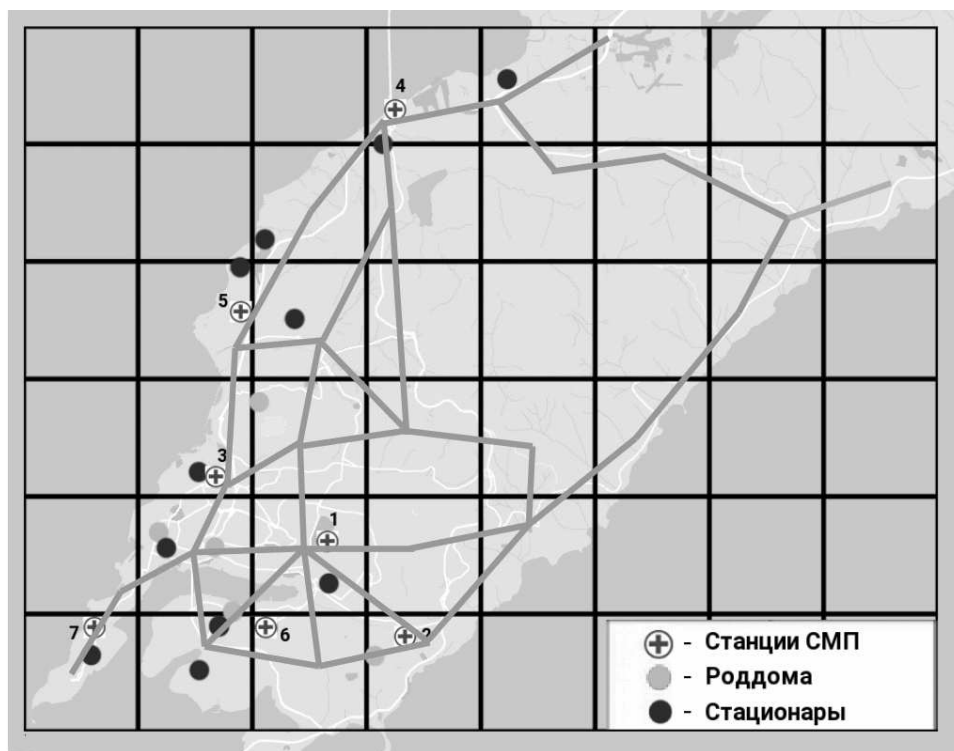


Рис. 2. Станции скорой медицинской помощи и стационары на агрегированном графе улично-дорожной сети г. Владивостока

распределения. С использованием генератора случайных величин для пуассоновского распределения было сформировано 500 сценариев поступления вызовов в службу СМП. Каждому вызову присвоен тип срочности: экстренный, срочный или неотложный. На основе данных отчётов службы СМП Владивостока доля каждого типа в общем количестве вызовов составляет: 17% — экстренные, 24% — срочные, 59% — неотложные. Для каждого сценария проводились расчёты на имитационной модели и подсчитывались средние значения целевой функции и параметров работы каждой станции. Расчёты проводились на процессоре Intel(R) Xeon(R) CPU E5-2680 0 @ 2,70 ГГц. Процессорное время на вычисление среднего времени ожидания, т. е. на реализацию имитационной модели для 500 сценариев поступления вызовов, в среднем составило 20 секунд.

В численных экспериментах использовались следующие параметры имитационной модели:

1) время диспетчеризации (мин) — случайная величина с равномерным распределением на отрезке $[1, 3]$;

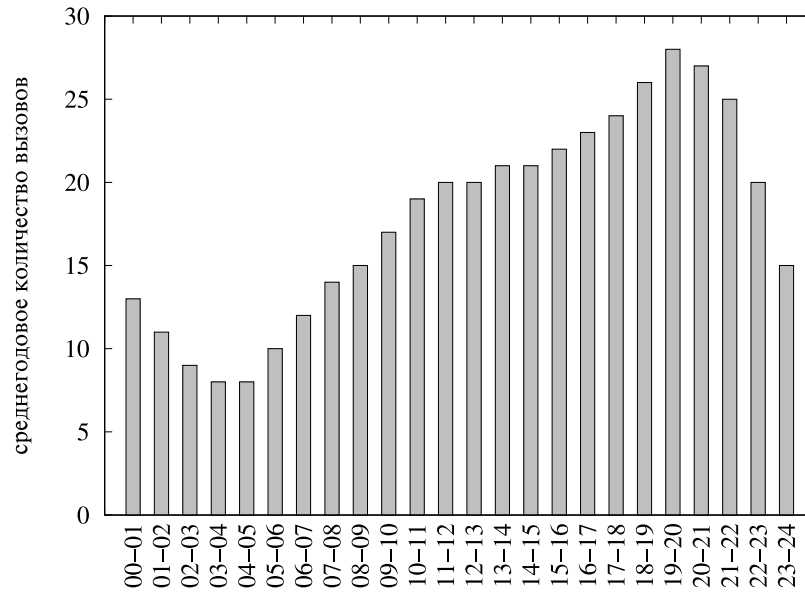


Рис. 3. Среднегодовое количество вызовов в течение суток

- 2) время первой помощи (мин) — случайная величина с равномерным распределением на отрезке $[15, 50]$;
- 3) время госпитализации (мин) — случайная величина с равномерным распределением на отрезке $[7, 15]$;
- 4) экстренные вызовы госпитализируются в обязательном порядке, остальные — с вероятностью 0,5.

Загрузка УДС прогнозировалась по результатам решения задачи (3)–(5) на каждый час рассматриваемого периода. Графовая модель реальной УДС Владивостока приводит к задаче большой размерности, что делает её трудно решаемой на практике. Поэтому с целью улучшения вычислительных свойств задачи (3)–(5) УДС Владивостока была агрегирована графом с 26 вершинами ($V = 26$) и 72 дугами ($A = 72$), общая топология которого приведена на рис. 2. Интерактивные сервисы маршрутизации позволили подсчитать время свободного передвижения по УДС, т. е. для агрегированной сети определить величины τ_a с учётом движения по реальным дорогам УДС Владивостока.

С использованием доступной информации (интерактивные карты города, статистические отчёты и пр.) для каждой зоны города (см. рис. 2) проведена оценка числа рабочих мест. Предполагается, что загрузку УДС создают потоки трудовой миграции: в утренние часы люди едут из мест проживания к местам приложения труда, в вечерние часы возвращаются обратно. Согласно статистике уровень занятости населения в экономике

Таблица 1

Доля экономически активного населения
в трудовой миграции по часам

8–9	9–10	10–11	11–12	16–17	17–18	18–19	19–20
0,4	0,3	0,2	0,1	0,1	0,35	0,35	0,2

Владивостока составляет примерно 60% от общего числа жителей и именно это количество владивостокцев заложено в модель в качестве потенциальных пользователей городских дорог.

Утренняя трудовая миграция с 8:00 до 12:00 загружает УДС и препятствует свободному проезду бригад. В данный период времени в качестве зон, порождающих потоки, рассматривались зоны проживания, в соответствии с этим определялись величины o_i ; в качестве зон, притягивающих потоки, рассматривались зоны мест приложения труда, в соответствии с этим определялись величины d_j задачи (3)–(5). Вечерняя трудовая миграция занимает период 16:00–20:00, для которого характерно изменение направления трудовой миграции: зоны приложения труда порождают потоки, зоны проживания их притягивают. Предполагалось, что в период 12:00–16:00 транспортные потоки не препятствуют движению бригад, поэтому в этот период задача (3)–(5) не решалась, за время проезда между зонами принято время свободного движения.

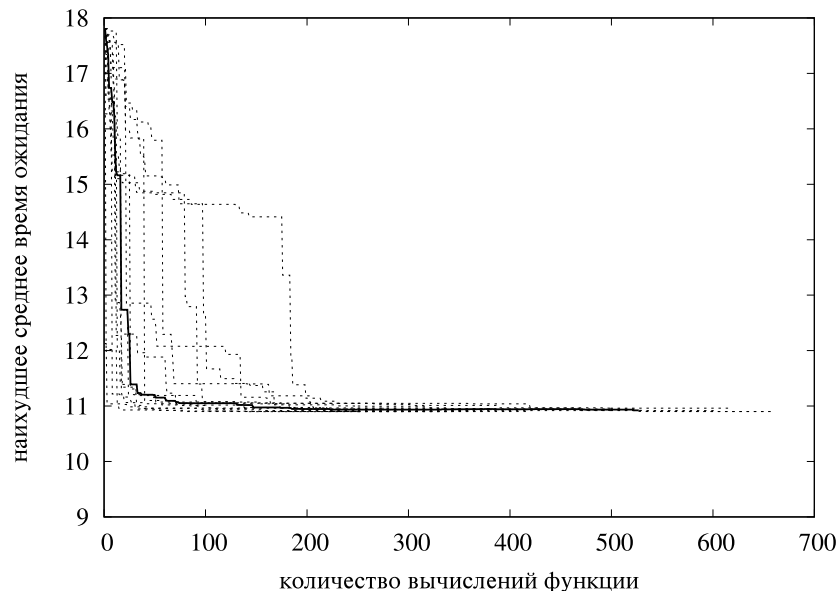


Рис. 4. Алгоритм ГА0

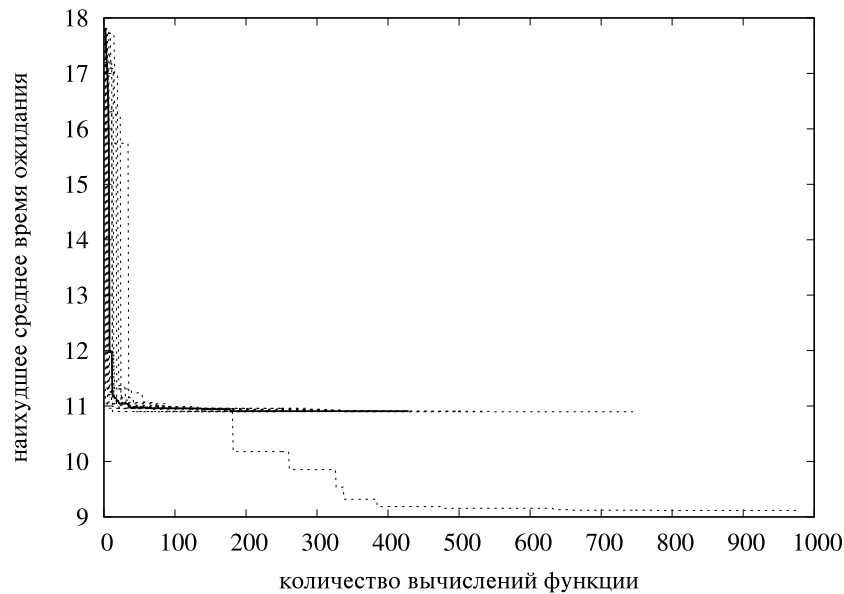


Рис. 5. Алгоритм ГА1

Доля экономически активного населения, выезжающего утром на работу и возвращающегося домой вечером, приведена в табл. 1.

Основную загрузку УДС создаёт трудовое население, передвигающееся на личном транспорте. Владивосток занимает первое место в России по уровню автомобилизации населения: каждый второй житель города имеет личный автомобиль. Если рассматривать только экономически активное население, то, скорее всего, доля автомобилизации получится выше. Варьируя процент автомобилистов (от 0% до 90%) и тем самым изменяя входные параметры o_i и d_j в задаче (3)–(5), мы рассчитывали разную возможную загрузку УДС Владивостока. Далее при помощи имитационно-оптимизационной модели определялись наилучшее размещение бригад на станциях в зависимости от транспортной ситуации.

6.2. Выбор окрестности для локального поиска. Первый эксперимент был направлен на сравнение генетических алгоритмов с разными окрестностями для локального поиска. Для сокращения временных затрат все вычисления проводились на четырёхчасовом периоде 8:00–12:00. Рассматривались пять версий генетического алгоритма: без локального поиска (ГА0) и с процедурой локального поиска по каждой из четырёх окрестностей $N_1(x), \dots, N_4(x)$ (алгоритмы ГА1, ..., ГА4 соответственно). На рис. 4–8 пунктирная линия отображает работу алгоритма при каждом из 15 испытаний. Сплошная линия показывает среднее значение

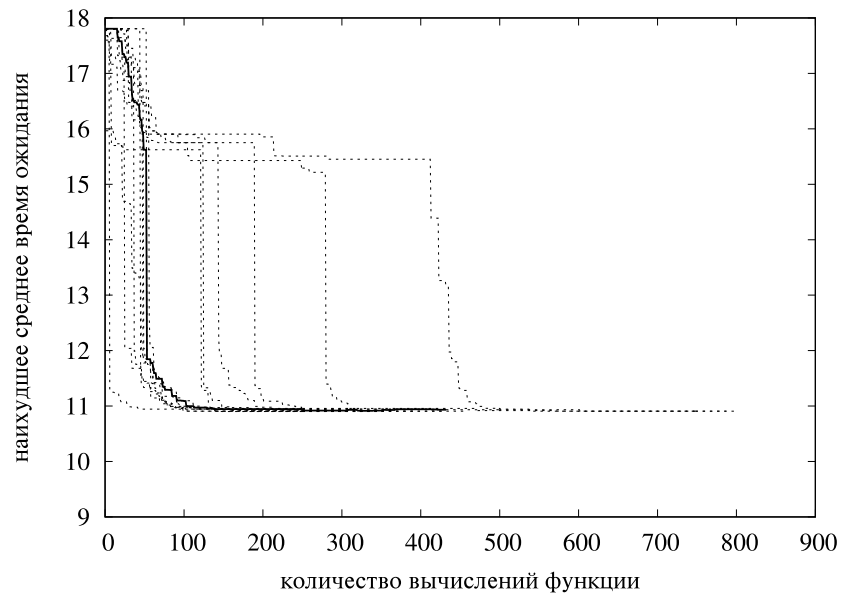


Рис. 6. Алгоритм ГА2

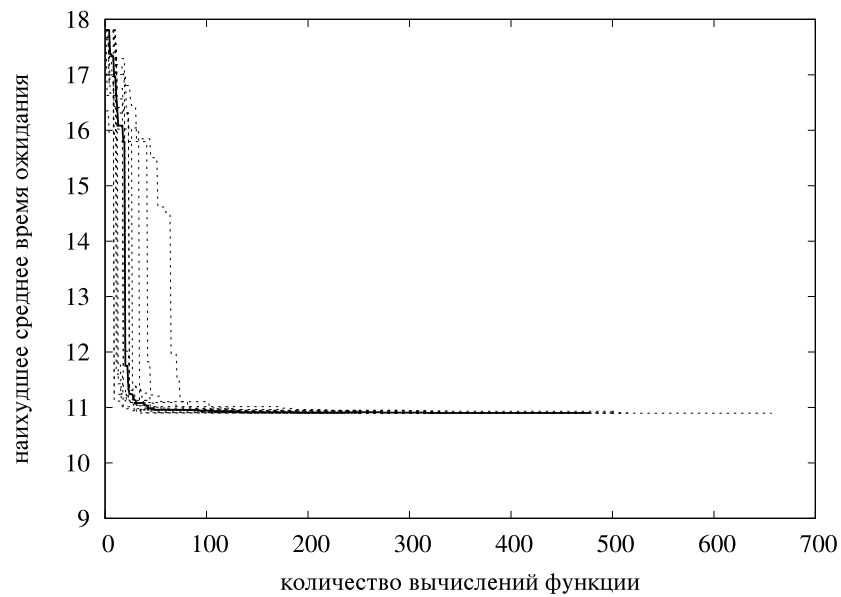


Рис. 7. Алгоритм ГА3

рекорда $\bar{\Phi}(x)$. Начальная популяция и значения всех параметров алгоритма для всех испытаний одинаковы. На рис. 9 приведён сводный график изменения рекорда. На всех графиках по оси абсцисс показано количество вычислений среднего значения целевой функции $\Phi(x)$ на имитационной модели по 500 сценариям поступления вызовов.

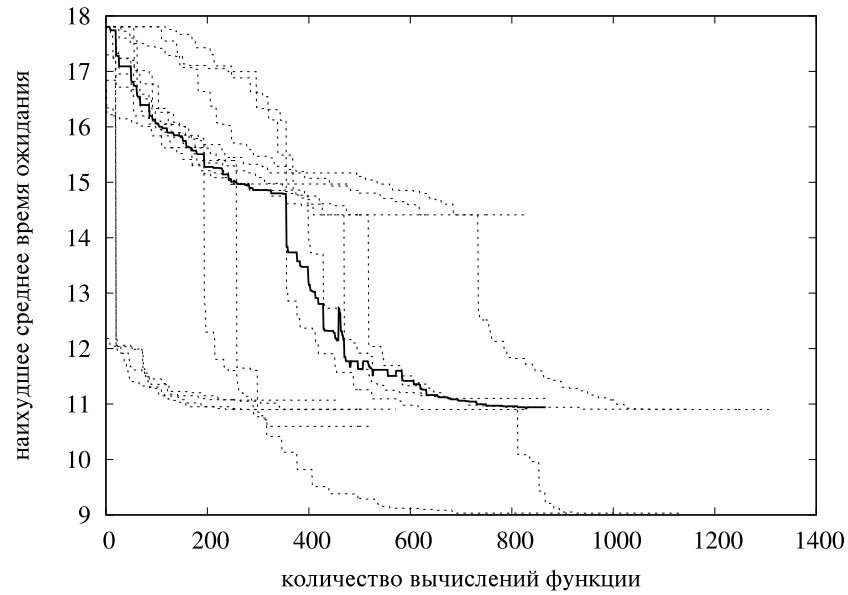


Рис. 8. Алгоритм ГА4

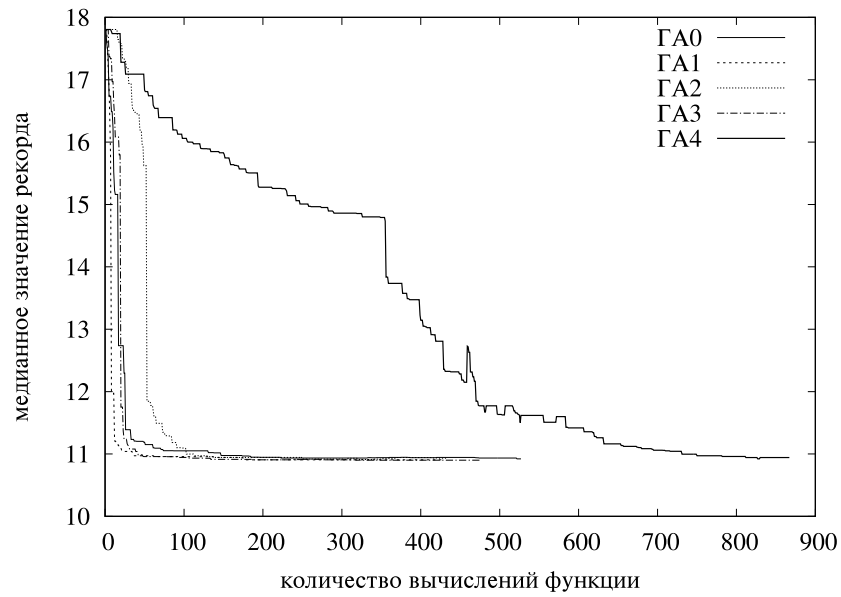


Рис. 9. Сводный график медианного значения рекорда $\bar{\Phi}(x)$

Как следует из представленных графиков, алгоритмы ГА0, ГА1 и ГА3 в среднем демонстрируют идентичный характер сходимости. Чуть медленнее сходится алгоритм ГА2, наиболее медленная скорость падения рекорда наблюдается у алгоритма ГА4. В проведённых экспериментах

все алгоритмы получили один и тот же результат. Тем не менее, для алгоритмов ГА0 (рис. 4), ГА2 (рис. 6) и ГА4 (рис. 8) характерны сильные отклонения текущего значения рекорда от среднего, что говорит о нестабильной работе данных алгоритмов. Противоположная картина наблюдается для ГА1 (рис. 5) и ГА3 (рис. 7) — при каждом запуске текущее наилучшее значение $\overline{\Phi}(x)$ близко к медианному рекорду по всем экспериментам.

Таким образом, с учётом результатов проведённых численных экспериментов, для минимизации функции $\Phi(x)$ целесообразно применять алгоритмы ГА1 и ГА3. Отметим, что в некоторых экспериментах ГА1 удалось построить решение с меньшим, чем медианный рекорд, значением функции $\Phi(x)$ (см. рис. 5). Для алгоритма ГА3 такого результата не наблюдалось. Далее в численных экспериментах применялся алгоритм ГА1.

6.3. Расчёты по размещению бригад. В следующем эксперименте рассчитываются и сравниваются между собой два варианта размещения бригад:

- 1) единый вариант размещения бригад на станциях на протяжении всего периода 8:00–20:00;
- 2) вариант с возможностью передислокации бригад между утренним (8:00–12:00), дневным (12:00–16:00) и вечерним (16:00–20:00) периодами.

Эксперимент преследует две следующие цели:

- 1) оценить влияние передислокации бригад на среднее время ожидания;
- 2) исследовать влияние загруженности дорог личным автотранспортом на среднее время ожидания бригад.

Моделирование проведено для разных уровней активности использования населением личного автомобиля для проезда к месту работы и обратно домой. Среди трёх независимых реализаций модели в качестве решения принимался вариант, соответствующий наименьшему значению целевого показателя $\Phi(x)$

На первом этапе была рассмотрена идеальная ситуация, когда транспортные потоки в УДС не препятствуют движению бригад на протяжении всего периода наблюдения. Это возможно, например, когда население предпочитает общественный транспорт личному и тем самым не нагружает УДС или когда автомобилисты дисциплинированно и заранее освобождают полосу движения для проезда бригад СМП. При этих предположениях полученный единый вариант размещения бригад по станциям представлен в табл. 2.

Бригады распределены по станциям № 1, 3, 5, 6. Силами этих четырёх станций обслуживается вся территория. Такое решение вполне

Таблица 2

**Единый вариант размещения бригад на период
8:00–20:00 при беспрепятственном движении по УДС**

№ станции	1	2	3	4	5	6	7
кол-во бригад (8:00–20:00)	8	0	8	0	7	3	0

объяснимо: в окрестности этих станций находится множество плотно населённых зон и имеет место хорошая транспортная связность с зонами, закреплёнными за станциями № 2, 4 и 7 (см. рис. 2).

Максимальное среди всех станций среднее время ожидания бригад на вызовы составило 13,43 минут (станция № 6), средняя нагрузка на бригаду получилась равной 9,8 вызовов, минимальное и максимальное количества обслуженных вызовов бригадами в среднем равно 7,9 и 11,2 соответственно. За весь период каждой из бригад удалось вернуться на станцию приписки 3 или 4 раза.

При тех же исходных данных получены наилучшие варианты размещения бригад для утреннего, дневного и вечернего периодов (табл. 3). Значение целевого показателя $\Phi(x)$ при таком решении незначительно снизилось до 13,12 минут (станция № 6).

Таблица 3

**Вариант размещения бригад с передислокацией
при беспрепятственном движении по УДС**

№ станции	1	2	3	4	5	6	7
кол-во бригад (8:00–12:00)	4	1	3	0	11	6	1
кол-во бригад (12:00–16:00)	7	2	6	0	6	4	1
кол-во бригад (16:00–20:00)	6	1	7	0	5	5	2

Видим, что модель реагирует на повышение интенсивности поступления звонков, стягивая «силы» на станции № 1 и 3 и «ослабляя» станцию № 5. В утренний период бригады в среднем выполняли от 1,7 до 4,1 вызовов, большей части бригад удалось только раз вернуться на станцию приписки. В дневной период нагрузка на бригады выровнялась: минимальная — 2,8, максимальная — 3,4, в среднем 3,2 вызова на бригаду; каждая бригада гарантированно один раз возвращалась на станцию приписки. В вечерний период нагрузка на бригады была равномерной: 3,8 вызова на бригаду, единожды каждая бригада возвращалась на станцию приписки. Из полученных результатов моделирования можно предположить, что количество бригад в утренний период может быть сокращено без снижения эффективности работы службы СМП,

но с выравниванием нагрузки на бригады, а в вечерний период увеличение количества бригад может позволить снизить нагрузку на врачей.

Отсутствие штата медиков на станции № 4 можно объяснить качеством исходных данных: в окрестных зонах, учтённых в модели, проживает малое число населения. Но в реалиях данная станция также обслуживает районы, не вошедшие в рассмотренную территорию. Отсюда делаем вывод, что работа бригад станции № 4 может быть ориентирована только на окрестные, но неучтенные в модели районы. Небольшая численность населения в зонах, закреплённых за станцией № 7, вполне может быть обслужена соседними станциями, поэтому здесь нет необходимости держать штат медиков. В реалиях Владивостока станция № 7 практически не работает.

Будем наращивать количество автомобилистов в УДС, увеличивая их долю на 10% при каждом очередном эксперименте. Наша цель — оценить влияние числа автомобилистов на работу службы СМП. Для каждого уровня загрузки, как и ранее, рассматривалось два варианта управленческих решений: единый и с возможностью передислокации бригад.

Полученные результаты приведены в табл. 4. Нижняя строка таблицы содержит данные о размещении бригад согласно отчётам службы СМП Владивостока. Из таблицы видим, что при загрузке сети не более 20% экономически активных автомобилистов расчёты приводят к весьма схожим вариантам размещения бригад по станциям. Аналогичная ситуация наблюдается, если доля автомобилистов превышает 50%: здесь основной штат бригад сосредоточен на станции № 1, отличающейся хорошей транспортной связностью со всеми районами города. Ключевыми станциями

Таблица 4

**Единый вариант размещения бригад по станциям
на период 8:00–20:00 при разной загрузке УДС**

% авто	№ 1	№ 2	№ 3	№ 4	№ 5	№ 6	№ 7
0	8	0	8	0	7	3	0
10	8	1	8	0	5	4	0
20	8	0	8	0	7	3	0
30	9	1	3	0	9	4	0
40	12	0	8	0	4	2	0
50	11	2	6	0	4	3	0
60	10	1	7	0	4	3	0
70	10	1	7	0	5	3	0
80	9	1	9	0	4	3	0
90	13	1	7	0	4	1	0
по факту	4	3	9	4	3	3	0

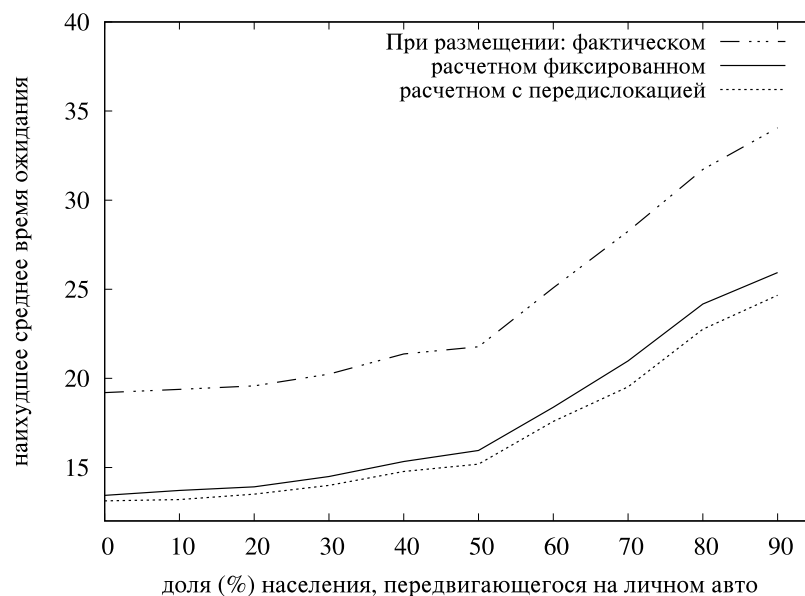


Рис. 10. Зависимость среднего времени ожидания от загруженности УДС

при любой загруженности УДС остаются станции № 1, 3, 5. Фактическое размещение бригад службы СМП Владивостока кардинально отличается от каждого из расчётных вариантов. В реалиях основной штат располагается на станции № 3 (центральная станция города), по остальным станциям, за исключением станции № 7, бригады распределены примерно в равном количестве.

Рис. 10 показывает зависимость среднего времени ожидания от доли экономически активного населения, использующего личный транспорт для поездок на работу и обратно. Вычисления проведены для фактического (штрих-пунктирная линия), единого (сплошная линия) и варианта с возможностью передислокации бригад (пунктирная линия). Если трудовая миграция включает в себя до 20% автомобилистов, то они фактически не препятствуют свободному передвижению бригад. Среднее время ожидания близко к значению, полученному в условиях беспрепятственного движения по сети. Дальнейший рост этого показателя в пределах от 30% до 50% приводит к увеличению значения $\Phi(x)$ в 1,18 раз от среднего времени ожидания при свободной сети. Резкое увеличение времени ожидания наблюдается, если более половины трудоспособного населения использует автомобиль для проезда на работу и обратно. Если 70% трудоспособного населения едет на работу на личном транспорте (для Владивостока это реальный сценарий), время ожидания увеличивается в 1,5 раза.

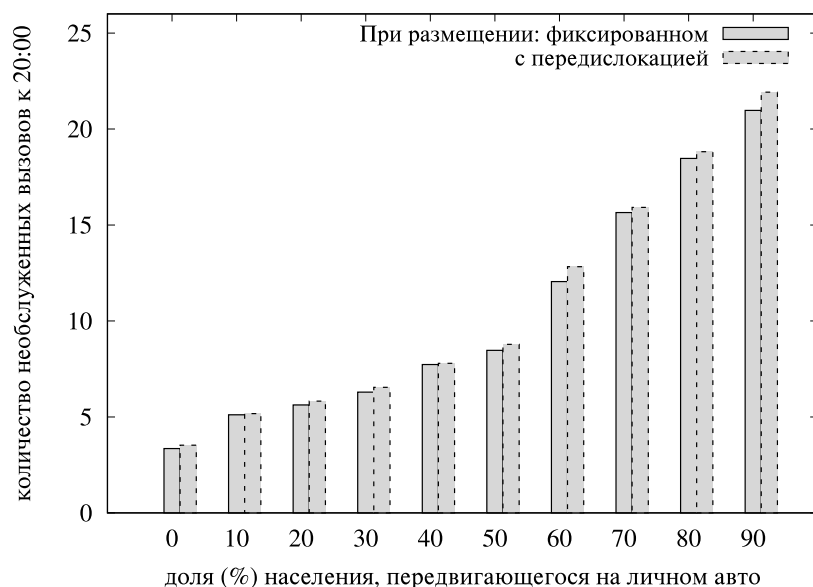


Рис. 11. Количество необслуженных вызовов к 20:00

Обращает на себя внимание размер очереди необслуженных вызовов, образующийся к концу периода наблюдения, т. е. к 20:00. На рис. 11 столбец со сплошным контуром соответствует расчётам при едином варианте размещения бригад, столбец с пунктирным контуром — при варианте размещения с возможностью передислокации бригад. Необслуженные вызовы накапливаются по мере увеличения загруженности сети. Если больше половины работающего населения использует личный транспорт для трудовых поездок, то очередь стремительно растёт и увеличивается в разы. Этот эффект делает целесообразным на следующий двенадцатичасовой период 20:00–8:00 привлечь к работе новую смену, бригады которой начнут обслуживать вызовы, поступающие после 20:00. Текущая смена закончит работу, когда она обслужит все вызовы, поступившие до 20:00. Такое управленческое решение позволит поддержать эффективную работу службы СМП и не допустить увеличения времени ожидания в следующем периоде из-за накопившейся очереди необслуженных вызовов предыдущего периода. Отметим, что после 20:00 вызовы продолжают поступать с высокой интенсивностью (см. рис. 3).

Заключение

Для решения задачи размещения и передислокации бригад по станциям СМП построена имитационная модель, учитывающая изменение

интенсивности поступления вызовов, загруженность УДС города, стохастический характер времени обслуживания пациентов. Для решения задачи разработан алгоритм генетического локального поиска, с помощью которого получены следующие результаты:

- 1) оптимизация размещения бригад по станциям может сократить время ожидания в 1,5 раза по сравнению с фактической вариантом;
- 2) передислокация бригад по станциям между утренним, дневным и вечерним периодами радикально не меняет ситуации по сравнению с единым расчётным вариантом размещения, наблюдается лишь незначительное уменьшение среднего времени ожидания в пределах 4%;
- 3) если более 50% экономически активного населения использует личный автотранспорт для трудовых поездок, то время ожидания стремительно растёт;
- 4) высокая активность использования личного автотранспорта негативно влияет на оперативную работу службы СМП, а управленческие решения о передислокации бригад между станциями не позволяют заметно сократить время ожидания.

Причиной, снижающей целесообразность передислокации бригад, может служить малое количество возвратов бригад на станции приписки (не более четырёх раз). В большинстве случаев, заканчивая обслуживание вызов, медики сразу отправляются на помощь следующему пациенту. По мере увеличения загруженности сети минимальная нагрузка на бригаду возрастает, а средняя и максимальная нагрузки остаются неизменными. При такой ситуации актуальность передислокации бригад в промежуточные периоды сомнительна.

В будущем возможны следующие пути развития проведённого исследования. Во-первых, необходимо улучшить вычислительные свойства генетического алгоритма за счёт применения параллельных вычислений. Генетический алгоритм ГА1 демонстрирует хорошие свойства сходимости, но его работа, в том числе реализация имитационной модели, занимает много машинного времени. Время расчётов для двенадцатичасового периода без передислокации бригад варьируется от 3 до 10 часов, с передислокацией — от 2 до 8 часов.

Во-вторых, интересно исследовать эффективность других управленческих решений, например, влияние разных сценариев диспетчеризации, ситуационное наращивание/сокращение работающих бригад, нормирование рабочих часов врачей, планирование перерывов для приёма пищи, контроль за равномерной нагрузкой бригад и т. п. Имитационная модель может быть дополнена новыми деталями работы скорой помощи, что даёт шанс повышения качества работы всей службы СМП.

ЛИТЕРАТУРА

1. **Reuter-Oppermann M., van den Berg P. L., Vile J. L.** Logistics for emergency medical service systems // *Health Syst.* 2017. Vol. 6. P. 187–208.
2. **Brotcorne L., Laporte G., Semet F.** Ambulance location and relocation models // *Eur. J. Oper. Res.* 2003. Vol. 147. P. 451–463.
3. **Goldberg J.** Operations research models for the deployment of emergency services vehicle // *EMS Manage. J.* 2004. Vol. 1. P. 20–39.
4. **Li X., Zhao Z., Zhu X., Wyatt T.** Covering models and optimization techniques for emergency response facility location and planning: a review // *Math. Methods Oper. Res.* 2011. Vol. 74, No. 3. P. 281–310.
5. **Aringhieri R., Bruni M. E., Khodaparasti S., van Essen J. T.** Emergency medical services and beyond: Addressing new challenges through a wide literature review // *Comput. Oper. Res.* 2017. Vol. 78. P. 349–368.
6. **Bélanger V., Ruiz A., Soriano P.** Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles // *Eur. J. Oper. Res.* 2019. Vol. 272. P. 1–23.
7. **Andersson N., Värbrand P.** Decision support tools for ambulance dispatch and relocation // *J. Oper. Res. Soc.* 2007. Vol. 58, No. 2. P. 195–201.
8. **Schmid V.** Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming // *Eur. J. Oper. Res.* 2012. Vol. 219, No. 3. P. 611–621.
9. **Fitzsimmons J. A., Srikar B. N.** Emergency ambulance location using the contiguous zone search routine // *J. Oper. Manage.* 1982. Vol. 2, No. 4. P. 225–237.
10. **Zaffar M. A., Rajagopalan H. K., Saydam C., Mayorga M., Sharer E.** Coverage, survivability or response time: a comparative study of performance statistics used in ambulance location models via simulation-optimization // *Oper. Res. Health Care.* 2016. Vol. 11. P. 1–12.
11. **Henderson S. G., Mason A. J.** Ambulance service planning: simulation and data visualization // *Oper. Res. Health Care.* 2004. Vol. 70. P. 77–102.
12. **McCormack R., Coates G.** A simulation model to enable the optimization of ambulance fleet allocation and base station location for increased patient survival // *Eur. J. Oper. Res.* 2015. Vol. 247. P. 294–309.
13. **Aboueljinane L., Sahin E., Jemai Z.** A review of simulation models applied to emergency medical service operations // *Comput. Ind. Eng.* 2013. Vol. 66. P. 734–750.
14. **Zhen L., Wang K., Hu H., Chang D.** A simulation optimization framework for ambulance deployment and relocation problems // *Comput. Ind. Eng.* 2014. Vol. 72. P. 12–23.
15. **Garcia R., Marin A.** Network equilibrium models with combined modes: models and solution algorithms // *Transp. Res. B.* 2005. Vol. 39. P. 223–254.

16. **Shamray N. B.** The general multimodal network equilibrium problem with elastic balanced demand // Discrete Optimization and Operations Research. Suppl. Proc. 9th Int. Conf. DOOR (Vladivostok, Russia, Sept. 19–23, 2016). Aachen: RWTH Aachen Univ., 2017. P. 404–414. (CEUR Workshop Proc.; Vol. 1623). Available at <http://ceur-ws.org/Vol-1623> (accessed Jan. 20, 2021).
17. PTV Group. Karlsruhe: PTV Planung Transport Verkehr, 2021. Available at <http://www.ptvgroup.com> (accessed Jan. 20, 2021).
18. Aimsun. Barcelona: TSS-Transport Simulation Systems, 2021. Available at <http://www.aimsun.com> (accessed Jan. 20, 2021).
19. TRANSIMS Open-Source. Available at <http://code.google.com/archive/p/transims> (accessed Jan. 20, 2021).
20. **Beckmann M., McGuire C. B., Winsten C. B.** Studies in the economics of transportation. New Haven: Yale Univ. Press, 1956.
21. Traffic Assignment Manual. Washington D.C.: Bureau of Public Roads, 1964.
22. **Вильсон А. Дж.** Энтропийные методы моделирования сложных систем. М.: Наука, 1978.
23. **Fang S.-C., Rajasekera J. R., Tsao H.-S. J.** Entropy optimization and mathematical programming. Boston: Kluwer Acad. Publ., 1997.
24. **Patriksson M.** The traffic assignment problem — Models and methods. Utrecht: VSP, 1994.
25. **Fogel L. J., Owens A. J., Walsh M. J.** Artificial intelligence through simulated evolution. Chichester, UK: Wiley, 1966.
26. **Holland J. H.** Adaptation in natural and artificial systems. Ann Arbor: Univ. Michigan Press, 1975.
27. **Ивахненко А. Г.** Системы эвристической самоорганизации в технической кибернетике. Киев: Техника, 1971.
28. **Растринин Л. А.** Статистические методы поиска. М.: Наука, 1968.

*Кочетов Юрий Андреевич
Шамрай Наталья Борисовна*

Статья поступила
9 ноября 2020 г.
После доработки —
13 января 2021 г.
Принята к публикации
15 января 2021 г.

OPTIMIZATION OF THE AMBULANCE FLEET
LOCATION AND RELOCATIONYu. A. Kochetov^{1, a} and N. B. Shamray^{2, b}¹ Sobolev Institute of Mathematics,
4 Acad. Koptuyug Avenue, 630090 Novosibirsk, Russia² Institute of Automation and Control Processes,
5 Radio Street, 690041 Vladivostok, RussiaE-mail: ^a jkochet@math.nsc.ru, ^b shamray@dvo.ru

Abstract. We consider the problem of optimal location of an ambulance fleet at the base stations. The objective is to minimize the average waiting time for ambulance arrival. We elaborate a simulation model that describes a day of work of the Emergency Medical Service (EMS). This model takes into account the stochastic nature of the problem and the changing busyness of the city's transport network. To solve the problem, a genetic local search algorithm was developed with 4 types of neighborhoods. The simulation model in this algorithm is used to compute the value of the objective function. We investigate the influence of neighborhoods on the accuracy of the obtained solutions and conduct computer simulations on the example of the Vladivostok EMS. We show that it is possible to reduce the average waiting time by 1.5 times. Some estimates on the impact of traffic congestion on the average waiting time are obtained. Tab. 4, illustr. 11, bibliogr. 28.

Keywords: optimization, emergency medical service, simulation model, genetic algorithm, local search.

REFERENCES

1. M. Reuter-Oppermann, P. L. van den Berg, and J. L. Vile, Logistics for emergency medical service systems, *Health Syst.* **6**, 187–208 (2017).

The study of the first author is carried out under the state contract of the Sobolev Institute of Mathematics (Project 0314–2019–0014). The study of the second author is supported by the Russian Foundation for Basic Research (Project 18–29–03071).

English version: Journal of Applied and Industrial Mathematics **15** (2), 234–252 (2021), DOI 10.1134/S1990478921020058.

2. **L. Brotcorne, G. Laporte, and F. Semet**, Ambulance location and relocation models, *Eur. J. Oper. Res.* **147**, 451–463 (2003).
3. **J. Goldberg**, Operations research models for the deployment of emergency services vehicle, *EMS Manag. J.* **1**, 20–39 (2004).
4. **X. Li, Z. Zhao, X. Zhu, and T. Wyatt**, Covering models and optimization techniques for emergency response facility location and planning: A review, *Math. Methods Oper. Res.* **74** (3), 281–310 (2011).
5. **R. Aringhieri, M. E. Bruni, S. Khodaparasti, and J. T. van Essen**, Emergency medical services and beyond: Addressing new challenges through a wide literature review, *Comput. Oper. Res.* **78**, 349–368 (2017).
6. **V. Bélanger, A. Ruiz, and P. Sorianoa**, Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles, *Eur. J. Oper. Res.* **272**, 1–23 (2019).
7. **N. Andersson and P. Värbrand**, Decision support tools for ambulance dispatch and relocation, *J. Oper. Res. Soc.* **58** (2), 195–201 (2007).
8. **V. Schmid**, Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming, *Eur. J. Oper. Res.* **219** (3), 611–621 (2012).
9. **J. A. Fitzsimmons and B. N. Srikar**, Emergency ambulance location using the contiguous zone search routine, *J. Oper. Manag.* **2** (4), 225–237 (1982).
10. **M. A. Zaffar, H. K. Rajagopalan, C. Saydam, M. Mayorga, and E. Sharer**, Coverage, survivability or response time: A comparative study of performance statistics used in ambulance location models via simulation-optimization, *Oper. Res. Health Care* **11**, 1–12 (2016).
11. **S. G. Henderson and A. J. Mason**, Ambulance service planning: simulation and data visualization, *Handb. Oper. Res. Health Care Methods Appl.* **70**, 77–102 (2004).
12. **R. McCormack and G. Coates**, A simulation model to enable the optimization of ambulance fleet allocation and base station location for increased patient survival, *Eur. J. Oper. Res.* **247**, 294–309 (2015).
13. **L. Aboueljinane, E. Sahin, and Z. Jemai**, A review of simulation models applied to emergency medical service operations, *Comput. Ind. Eng.* **66**, 734–750 (2013).
14. **L. Zhen, K. Wang, H. Hu, and D. Chang**, A simulation optimization framework for ambulance deployment and relocation problems, *Comput. Ind. Eng.* **72**, 12–23 (2014).
15. **R. Garcia and A. Marin**, Network equilibrium models with combined modes: Models and solution algorithms, *Transp. Res. Part B*, **39**, 223–254 (2005).
16. **N. B. Shamray**, The general multimodal network equilibrium problem with elastic balanced demand, in *Discrete Optimization and Operations Research, Suppl.* (Proc. 9th Int. Conf. DOOR, Vladivostok, Russia, Sept. 19–23, 2016) (RWTH Aachen Univ., Aachen, 2017), pp. 404–414 (CEUR Workshop Proc., Vol. 1623). Available at <http://ceur-ws.org/Vol-1623> (accessed Jan. 20, 2021).

17. *PTV Group* (PTV Planung Transport Verkehr, Karlsruhe, 2021). Available at <http://www.ptvgroup.com> (accessed Jan. 20, 2021).
18. *Aimsun* (TSS-Transport Simulation Systems, Barcelona, 2021). Available at <http://www.aimsun.com> (accessed Jan. 20, 2021).
19. *TRANSIMS Open-Source*. Available at <http://code.google.com/archive/p/transims> (accessed Jan. 20, 2021).
20. **M. Beckmann, C. B. McGuire, and C. B. Winsten**, *Studies in the Economics of Transportation* (Yale Univ. Press, New Haven, 1956).
21. *Traffic Assignment Manual* (Bureau of Public Roads, Washington D.C., 1964).
22. **A. G. Wilson**, *Entropy in Urban and Regional Modelling* (Pion, London, 1970; Nauka, Moscow, 1978 [Russian]).
23. **S. C. Fang, J. R. Rajasekera, and H.-S. J. Tsao**, *Entropy Optimization and Mathematical Programming* (Kluwer Acad. Publ., Boston, 1997).
24. **M. Patriksson**, *The Traffic Assignment Problem – Models and Methods* (VSP, Utrecht, 1994).
25. **L. J. Fogel, A. J. Owens, and M. J. Walsh**, *Artificial Intelligence Through Simulated Evolution* (Wiley, Chichester, UK, 1966).
26. **J. H. Holland**, *Adaptation in Natural and Artificial Systems* (Univ. Michigan Press, Ann Arbor, 1975).
27. **A. G. Ivakhnenko**, *Heuristic Self-Organization Systems in Engineering Cybernetics* (Tekhnika, Kyiv, 1971) [Russian].
28. **L. A. Rastrigin**, *Statistical Search Methods* (Nauka, Moscow, 1968) [Russian].

Yury A. Kochetov
Natalia B. Shamray

Received November 9, 2020
Revised January 13, 2021
Accepted January 15, 2021