This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

The Multi-Depot Cumulative Vehicle Routing Problem with Mandatory Visit Times and Minimum Delayed Latency

ALAN OSORIO-MORA[1], MAURICIO SOTO-BUSTOS[2], GUSTAVO GATICA[3], PEDRO PALOMINOS[4], RODRIGO LINFATI[5]

[1]Department of Electric, Electronic and Information Engineering, University of Bologna, Viale Risorgimento 2, Bologna I-40136, Italy (email: alan.osorio2@unibo.it)
[2]School of Industrial Engineering, Universidad del Bío-Bío, Concepción 4030000, Chile (email: mauricio.soto1401@alumnos.ubiobio.cl)
[3]Faculty of Engineering, Universidad Andres Bello, Santiago 7500000, Chile (email: ggatica@unab.cl)
[4]Department of Industrial Engineering, Universidad de Santiago de Chile, Santiago 9170020, Chile (email: pedro.palominos@usach.cl)
[5]Department of Industrial Engineering, Universidad del Bío-Bío, Concepción 4030000, Chile (email: rlinfati@ubiobio.cl)

Corresponding author: Rodrigo Linfati (e-mail: rlinfati@ubiobio.cl).

This research was partially funded by Universidad del Bío-Bío, through the project "ALGORITMOS EXACTOS Y METAHEURÍSTICOS PARA NUEVAS VARIANTES DEL PROBLEMA DE RUTEO DE VEHÍCULOS", code 2060222 IF/R.

ABSTRACT This paper introduces a novel variant of the cumulative vehicle routing problem (CCVRP) that deals with home health care (HHC) logistics. It includes multiple nonfixed depots and emergency trips from patients to the closest depot. The aim is to minimize the system's delayed latency by satisfying mandatory visit times. Delayed latency corresponds to caregivers' total overtime hours worked while visiting patients. A new mixed-integer linear programming model is proposed to address this problem. Computational experiments, with more than 165 new benchmark instances, are carried out using the CPLEX and Gurobi MIP solvers. The results indicate that patients' geographical distribution directly impacts the complexity of the problem. An analysis of the model parameters proves that instances with more depots/vehicles or longer workdays are significantly easier to solve than are original cases. The results show that Gurobi outperforms CPLEX in 55% of the instances analyzed, while CPLEX performs better in only 16% of them. To the best of our knowledge, this is the first VRP that minimizes delayed latency and the first HHC routing study to use a cumulative objective function.

INDEX TERMS CCVRP, Home Health Care, HHC, MDCCVRP, Mixed Integer Programming.

I. INTRODUCTION

In recent years, home health care (HHC) services have grown considerably in many countries. This growth can be explained by the increase in the number of patients with chronic diseases and physical disabilities [1]. These types of services help prevent queues and congestion at hospitals and allow patients to receive timely attention. HHC logistics management can be addressed at three different decision levels. Strategic planning includes districting, setting the location of infrastructure, and fleet sizing. Tactical decisions include those related to medical staff allocation (to hospitals or other facilities) and inventory policies. Operative planning mainly involves routing decisions but also includes scheduling and other short-term decisions [2]. A limited team of caregivers tending to patients at their scheduled appointment times can be viewed as an operative planned logistic activity [3]. From the point of view of service providers, such a situation can therefore be addressed as a vehicle routing problem.

Vehicle routing problems (VRPs) are one of the most studied types of combinatorial optimization problems due to their varied applications. A VRP aims to find the best sequence of client visits for a fleet, generally using cost-minimization criteria. Vehicles may or may not have a limited capacity; if they do, then the problem can be classified as a capacitated vehicle routing problem (CVRP) [4]. This type of problem was introduced by [5], who proposed the first formulation and algorithm to solve a real problem of gasoline distribution. Since then, several authors have conducted research in this field. Some authors have worked on exact solution methods, [6], [7], while others have focused on heuristics-metaheuristics methods [8], [9]. Still others have developed new VRP applications including HHC [10], drone routing [11], [12], blood transportation processes [13], and several others [14]. The depot is the node at which vehicles start and end routes. A classical VRP has only one depot, but the multi-depot vehicle routing problem (MDVRP) has also been identified. In MDVRPs, vehicles are forced to start and end a route at the same depot [15]. However, it is possible to relax this condition and allow vehicles to close their routes at other depots [16]. This concept is known as a nonfixed destination or simply as a nonfixed MDVRP [17]. A literature review of MDVRPs can be found in [18]. One of the most studied variants of VRPs is a vehicle routing problem with time windows, VRPTW, which incorporates an interval of time in which vehicles can visit each customer. There are two kinds of time windows (TWs) in VRPs. The most studied type is hard TWs, in which the interval must be respected. In this case, if a vehicle arrives early, then service must be provided at the lower bound of the TW [19]. Conversely, a soft TW allows vehicles to violate the TW but penalizes them for doing so [20], [21].

A relatively new branch of VRPs is the use of latency as an objective function. Latency can be defined as the total distance traveled, or the time required, to reach each node. In latency problems, costs are influenced not only by the location of the nodes but also by their position in the route [22]. These problems are known as cumulative vehicle routing problems.

This paper is structured as follows. Section 1 presents the problem to be addressed and describes the relevant recent research. Section 2 describes the materials and methods used in this work; then, a new mathematical model used to address the problem is introduced. Section 3 discusses the main research results. Finally, Section 4 presents the conclusions and directions for future research.

A. PROBLEM DESCRIPTION

Traditionally, HHC routing has been addressed mainly from a business point of view; that is, it has focused primarily on minimizing costs. This can be a problem if these routes do not guarantee service quality. Low-quality services not only imply economic consequences due to a loss of clients but can also lead to legal repercussions. Several authors agree that problems with cumulative objective functions (latency) are well suited for addressing real-life problems in which the focus is on client satisfaction [23] [24] [25]. The above aspect is due to the nature of the objective function, which implies that patients need to be attended to as soon as possible. Figure 1 compares solutions under criteria for minimum total cost/time and minimum latency. In the example, two clients must be visited by two available vehicles. The travel time between nodes is indicated close to the edges. The optimal total cost-based route is D1-C2-C1-D1, using only one vehicle. The optimal latency-based routes are D1-C1-D1 and D1-C2-D1, using two vehicles (each line type represents a different vehicle). The total cost for the first case is 30 units, while its total latency is 37 units because C2 is visited at 12 and C1 is visited at 25. The total cost for the second case is 34 units, while its total latency is 17 units (C1 is visited at 5 and C2 is visited at 12).

As presented in Section 1.2, some authors have included clients' preferences in their HHC models. However, none of them have looked for routes defined by the desirable maximum workload per client. Due to the limited nature of resources such as vehicles, caregivers, and hospitals, it is not always possible to tend to each patient without exceeding the aforementioned limit. Nevertheless, it is desirable to minimize such situations. It has been proven that a heavy workload for caregivers implies burnout, which leads to low-quality services [26]. According to these authors, work overload causes emotional exhaustion and depersonalization in caregivers, leading to bad work performance.

The problem addressed is as follows. Consider an incomplete graph, $G = (N, A)$, with $N = C \cup T$ as the set of nodes, C as the set of patients, T as the set of hospitals, and A as the set of feasible arcs. Additionally, let B be the set of vehicles/caregivers. Each patient, i, has a scheduled visit time, h_i, that needs to be fulfilled. Each hospital (depot), k, has an available fleet at the beginning of the day, QI_t, and a demand for vehicles at the end of the day, QF_t. The supply and demand may be different. After visiting all their assigned clients, vehicles $b \in B$ can finish their route at the same or different depot than the one at which they started. An arc, (i, j), with $i \in N$ and $j \in C$, is feasible if the sum of h_i (if $i \in T, h_i = 0$) and travel time d_{ij} is less than or equal to h_j. As visit time is strict, caregivers can arrive early and wait or arrive just in time, but they can never arrive late. For each patient, an emergency travel time to the closest hospital must be considered in case the patient suffers an unexpected emergency or if the medical team needs medical supplies that are not available in their vehicle. Such a situation is an important consideration since most of the patients on these routes have chronic conditions. The challenge is to find the best route for each vehicle, considering a maximum desirable quota, l_b, for client visits. This quota can be represented by the standard workday. If necessary, services with overtime are allowed. Nevertheless, it is an undesirable condition due to the risk of low-quality services that it implies. The aim is to minimize the sum of the overtime hours with which clients are visited (if any). The more patients who are served after the regular workday, the greater the penalization.

B. LITERATURE REVIEW

The structure of this literature review is presented in Figure 2. It focuses mainly on two kinds of problems: i) VRPs in the HHC industry and ii) cumulative VRPs. For an in-depth review of VRP variants and their respective practical applications, the reader may refer to [4].

1) Vehicle Routing Problems in HHC

Two of the main operational decisions in HHC are routing and scheduling. These two decisions can be made individually or simultaneously [27]. The literature contains a number of studies that focus only on routing in HHC. The authors of [1] solved an HHC routing problem by addressing it as a CVRP. They used the well-known savings algorithm to solve a situation involving 20 patients. The work of [28] addressed a simultaneous pickup and delivery VRPTW. Drugs are delivered to patients, and simultaneously, unused drugs or biosamples are collected. To deal with this problem, the authors proposed two mixed-integer programming (MIP) models. Two heuristic algorithms were also used: a genetic algorithm (GA) and a tabu search (TS). An HHC routing problem in the context of natural disasters was addressed by [29], who proposed a model that included soft time windows, mandatory rest for staff, and workday constraints. Some interesting considerations, such as nurses' qualifications and language proficiency, were also incorporated. The objective was to minimize the time spent and both client and staff dissatisfaction. They found that the Xpress v7 solver was not capable of solving large instances, so they proposed a variable neighborhood search algorithm (VNS) to handle them. The proposed model and algorithm were tested with real data from Austria. In [30], mixed-integer linear programming (MILP) models for a VRPTW with precedence and synchronization constraints were proposed. Patients needed to receive different kinds of services at specific intervals of time, following a precedence pattern. The objective was to minimize travel time, clients' dissatisfaction with their assigned caregivers, and wait times. It also considered a maximum route length. Models were solved using OPL v12.5, and they handled situations involving 45 clients. A similar case was addressed in [31], in which the aim was to minimize costs, and the proposed method was an iterative local search algorithm (ILS). In turn, [32] studied an HHC application of a VRP with the synchronization constraints proposed by [33]. The objective was to minimize caregivers' travel time. The authors proposed a metaheuristic based on an ILS combined with the random variable neighborhood descent method (RVND) to solve the problem. A multiobjective approach to this kind of problem was addressed by [34], which sought to optimize both cost and client preferences. The authors proposed different variants of the nondominated sorting genetic algorithm (NSGA-II) to solve the problem with up to 73 clients.

The home health care routing and scheduling problem (HHCRSP) is a variant of the VRP that includes simultaneous routing and scheduling decisions [35]. In [36], different kinds of services had to be provided to clients, and both client and staff availability had to be considered in the time windows. The aim was to minimize the staff required to satisfy client demand. The authors proposed an integer linear programming (ILP) model. Small and medium-sized instances were solved by CPLEX. For large instances, a matheuristic that decomposed the model into two subproblems (staff rostering and vehicle routing) was proposed. A multiobjective MILP model was proposed by [37], which optimized four objectives by using the weighted linear aggregation method: to minimize the total travel time and the arrival times of each caregiver and to maximize caregiver operability and patient satisfaction. This model also considered patients' relative priority, which was related to the health condition of each patient. The authors used CPLEX to test the proposed model and optimally solve instances with up to 40 nodes. Another multiobjective model for this kind of problem was addressed by [38]. The objective of the proposed model was to minimize cost and maximize service level, and it also incorporated time windows and client preferences regarding caregivers and visit times. The proposed solution corresponded to a metaheuristic combining a large neighborhood search with a multidirectional LS. In the work of [39], a matheuristic algorithm was proposed to address a periodic scheduling and routing problem applied to medication delivery. This algorithm consisted of two phases. In the first phase, a mathematical optimization model was used to solve the scheduling problem; the model was solved with CPLEX. The second phase incorporated a metaheuristic that combined simulated annealing with a record-to-record algorithm to solve the routing problem. The proposed algorithm was tested with real-life pilot cases in Chile, which included up to 800 clients. Another two-phase matheuristic algorithm that sought to minimize transport and labor costs was proposed to solve an HHCRSP in [40]. This approach was based on MIP modeling and separated decisions by type of caregiver. In the first phase, only nurses were scheduled. The solution determined by the first phase was added as a constraint to the second phase, in which the other caregivers and synchronization constraints were considered. The authors used Gurobi to perform their experiments. They found that the proposed approach was computationally more efficient than was solving the complete model. Two recent literature reviews of HHCRSP can be found in [35], [41]. TW has been a common feature observed and analyzed in most HHC papers. Mandatory visit times are a particular case of hard TW; here, intervals' lower bounds are open [21].

2) Cumulative Vehicle Routing Problems

Research on cumulative routing problems began with the minimum latency problem (MLP), which is a variant of the traveling salesman problem (TSP) [42]. The objective of the MLP is to minimize the sum of the arrival times at each client. According to [43], the MLP has also been referred to as the traveling repairman problem (TRP) [44] and delivery man problem (DMP) [45]. Despite the fact that cumulative TSP variants were introduced in the 1990s, cumulative VRP variants were not introduced until much more recently, within approximately the last ten years. Seminal work on this problem was published by [23]. The authors called this problem The Cumulative Capacitated Vehicle Routing Problem (CCVRP). In the same paper, the authors also presented upper and lower bounding procedures, which are based on memetic algorithms (MAs) and the properties of the CCVRP, respectively. The CCVRP seeks to minimize system latency and is a generalization of the MLP, adding capacity constraints and homogeneous fleets.

Since the publication of these seminal works, several

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

solution methods have been proposed. The authors of [46] proposed an adaptive large neighborhood search heuristic that outperforms the MAs of [23]. An ILS was proposed by [47], and it outperformed the algorithm of [23] but was less effective than that proposed by [46]. The two-phase metaheuristic presented by [48] also outperformed the MAs. In some instances, this algorithm was capable of finding better solutions than those found by [46], but it took more time to do so. In [49], the authors proposed and compared three methods: a GA, an evolutionary algorithm using particle swarm optimization, and a TS. Their results indicated that the TS provides the best solutions in most of the instances tested. However, these algorithms were not compared with others from the literature. In [25], a two-stage adaptive VNS was proposed. The authors compared their proposed method with the algorithms proposed by [23], [46], and [48]. New best-known solutions were found for a number of instances in the literature, proving the competitiveness of the algorithm in terms of solution quality and CPU time. The authors also tested the effectiveness of the algorithm with a CCVRP that minimizes the maximum arrival time. A brainstorm optimization algorithm was presented in [50], and the authors compared it with those proposed by [46], [48], and [25]. This algorithm was capable of finding new best-known solutions for several instances, and it took less time on average. Two new mathematical formulations were proposed by [51]. One of them could optimally solve instances of up to 44 nodes in less than two hours using CPLEX. The authors also proposed two versions of the greedy algorithm for the CCVRP. The average GAP between their obtained results and the best known solutions was less than 1%. Their algorithm was capable of finding new and improved solutions for a few instances and also arrived at the best known solution for several instances in less CPU time than that required by the algorithms in [23], [46], [48], and [25]. Recently, [52] proposed a skewed VNS heuristic. The authors compared it with the algorithms from [23] and [46], and it was proven to be more efficient in terms of CPU time and solution quality. The only exact algorithm found in the literature was proposed by [53]. They developed a branch and cut and price algorithm, which was capable of optimally solving instances with more than 100 clients.

Below, the variants of the CCVRP studied in the literature are presented. In [54], the authors proposed a hybrid ant colony algorithm to solve a multi-depot cumulative VRP (MDCCVRP) as applied to postdisaster route planning. More recently, an LS-based algorithm was presented in [55] to solve the same theoretical problem, which was also studied in [24], in which the authors proposed a matheuristic that decomposed the problem into subproblems that could be easily solved to optimality. Their approach was called partial optimization metaheuristic under special intensification conditions.

The multitrip cumulative capacitated vehicle routing problem (MTCCVRP) is another variant of the problem that has been studied in the literature. Unlike in the classical problem, the MTCCVRP allows vehicles to perform several trips. It was addressed for the first time by [56], who proposed two MIP formulations and a GRASP-based metaheuristic for the problem; they considered a single-vehicle scenario. The same question was addressed in [57], where the authors proposed two MILP models and an exact algorithm to solve the problem. This method was based on a resource-constrained shortest path formulation and was capable of optimally solving instances with up to 40 clients. The generalized version of the problem, that is, the non-single-vehicle variant, was introduced by [58], who presented an MILP model and a hybrid metaheuristic for the MTCCVRP. The proposed solution combined a multistart ILS with variable neighborhood descent. The authors tested it with instances with more than 400 clients. More recently, the authors in [59] studied another variant of the problem, a CCVRP with priority indexes. This variant incorporated precedence constraints to ensure that certain clients were visited before others. The problem was addressed with biobjective optimization, and the objectives were to minimize the system's total latency and total tardiness. An MIP model was proposed but was capable of solving instances with only up to 15 customers. Therefore, the authors proposed MAs with random keys to solve larger instances. Regarding CCVRPs with time windows, the only work found was that of [60]. These authors introduced the problem and proposed a metaheuristic based on the hybridization of a large neighborhood search algorithm with a GA. They asserted that this problem was suitable for several applications, especially humanitarian logistics in postdisaster contexts.

Cumulative routing problems have many possible applications. The work of [61] presented a latency VRP applied to postdisaster management; specifically, it addressed the routing of assessment teams in disaster areas. The authors proposed a continuous approximation approach to solve this problem. Another latency routing problem was addressed in [62], which also dealt with postdisaster routing. This method considered minimum service level constraints, and the authors developed a VNS-based heuristic. Unnamed aerial vehicle surveillance services were also addressed with latency routing problems in [63]. Here, the authors used linear programming (LP) to minimize the delivery latency in this context, proposing a model that allowed vehicles to perform multiple trips. The cumulative VRP was also a natural fit for modeling trucks' fuel consumption [64]. In [65] and [66], column generation-based algorithms were used to address CCVRP in a fuel-consumption context. Cumulative objective functions may also be suitable for modeling energy management systems for electric/hybrid vehicles [67], [68]. To the best of our knowledge, VRPs minimizing latency have not yet been used in the HHC logistics context.

C. MAIN CONTRIBUTION OF THIS WORK

The literature review revealed that studies on cumulative routing problems have focused on solution methods for the classic CCVRP (mainly heuristics). Few variants have been addressed until now, including the CCVRP with priority index, the MTCCVRP, the CCVRPTW and the MDCCVRP. The main contribution of the present article is that it introduces a novel variant of the CCVRP to address HHC logistics. To propose a suitable method for supporting routing planning in HHC, our problem combines the features of CCVRP variants that have been studied only individually. The problem includes multiple nonfixed depots, emergency trips to the closest depot, and mandatory visit times. To the best of our knowledge, this is the first VRP that seeks to minimize delayed latency. It is also the first HHC routing problem addressed through a cumulative objective function. The concept of a system's delayed latency corresponds to the total overtime hours with which each patient is visited. A novel MILP model is proposed to handle this problem. The proposed model is implemented with two of the most-used commercial solvers, CPLEX and Gurobi, to benchmark their performance. This method is designed for planning the attention of patients with chronic diseases and physical disabilities. Nevertheless, it is suitable for several other applications, such as humanitarian logistics, forest fire control, and food transportation. The common feature in all of these scenarios is that customer satisfaction is the key factor.

II. MATERIALS AND METHODS

To address the problem described in Section 1.1, we propose an MILP model. The problem is a novel variant of the CCVRP. We have called it the nonfixed Multi-depot Cumulative Vehicle Routing Problem with mandatory visit times (MDCCVRmvt).

The proposed model has the following sets:

Sets

- C: Clients
- T: Depots
- N: Nodes, $N = C \cup T$
- B: Vehicles
- A: Feasible arcs that comply with scheduled visit times

In the model, clients represent patients, while depots represent hospitals or other medical facilities. In this article, "clients" are equivalent to "customers" in the classical VRP. We use the terms "patient" and "client" because the person is receiving a service. The term "customer" is more suitable for goods transportation.

The parameters used in the model are as follows:

Parameters

- QI_t: Number of vehicles at depot t at the beginning of the day.
- QF_t: Number of vehicles at depot t at the end of the day.
- l_b: Maximum desirable quota with which vehicle b can visit each client before being penalized (workday).
- h_i: Scheduled visit time of client i.
- d_{ij}: Travel time from node i to node j.
- $BigM$: Sum of worst arcs from each node, mathematically, $\sum_{i \in N} \max_{j \in N}\{d_{ij}\}$

Finally, the following decision variables are introduced:

Decision variables

- X_{ij}^b: 1 if the arc going from i to j is considered part of the route of vehicle b and 0 otherwise
- U_i^b: Accumulated time spent traveling to client i by vehicle b
- CU_i: Overtime hours with which client i is visited

The proposed model's formulation corresponds to (1) – (12):

$$min\ Z = \sum_{i \in C} CU_i \tag{1}$$

subject to:

$$\sum_{b \in B, j \in N | (i,j) \in A} X_{ij}^b = 1 \qquad \forall i \in C \tag{2}$$

$$\sum_{j \in N | (j,i) \in A} X_{ji}^b = \sum_{j \in N | (i,j) \in A} X_{ij}^b \qquad \forall i \in C, b \in B \tag{3}$$

$$\sum_{b \in B, j \in C | (t,j) \in A} X_{tj}^b \leq QI_t \qquad \forall t \in T \tag{4}$$

$$\sum_{b \in B, j \in C | (j,t) \in A} X_{jt}^b \leq QF_t \qquad \forall t \in T \tag{5}$$

$$\sum_{t \in T, j \in N | (t,j) \in A} X_{tj}^b \leq 1 \qquad \forall b \in B \tag{6}$$

$$d_{tj} \leq U_j^b + BigM(1 - X_{tj}^b) \qquad \forall b \in B, t \in T, j \in C | (t,j) \in A \tag{7}$$

$$U_i^b + d_{ij} \leq U_j^b + BigM(1 - X_{ij}^b) \qquad \forall b \in B, i \in C, j \in C | (i,j) \in A \tag{8}$$

$$CU_i \geq U_i^b + \min_{t \in T}\{d_{it}\} - l_b \qquad \forall i \in C, b \in B \tag{9}$$

$$X_{ij}^b \in \{0,1\} \qquad \forall b \in B, i \in N, j \in N | (i,j) \in A \tag{10}$$

$$U_i^b \geq 0 \qquad \forall i \in C, b \in B \tag{11}$$

$$CU_i \geq 0 \qquad \forall i \in C \tag{12}$$

The objective function (1) minimizes the delayed latency of the system. It corresponds to the sum of the overtime hours with which each patient is visited. The greater the number of patients visited by the vehicle, the more greater the penalization. To the best of our knowledge, our model is the first VRP that intends to optimize delayed latency. The set of constraints (2) ensures that all clients are visited once. In (3), we present flow balance constraints, which establish consistency in the use of vehicles. The set of constraints (4) indicates the maximum number of vehicles that can start a route from each depot. Similarly, (5) provides the maximum number of vehicles that can end a route at depot t. It is important to note that the model does not force vehicles to return to their starting depot; it is a nonfixed MDVRP. The constraints in (6) prevent vehicles from performing more than one route. The constraint groups (7) and (8) calculate the accumulated time for each vehicle at each client; they also correspond to the subtour elimination constraints. While set (7) considers arcs from depots, (8) accounts for between-client arcs. The set of constraints (9) calculates the overtime hours spent with each client. Overtime hours correspond to the difference between the workday and the cumulative time plus an emergency trip to the closest depot. Equations (10) to (12) present the decision variable domains.

Finally, set A, which represents all feasible arcs, must be defined. This definition is given by 13-15.

$$A1 : (t, i) \mid h_t + d_{ti} \leq h_i \qquad \forall t \in T, i \in C \qquad (13)$$
$$A2 : (i, j) \mid h_i + d_{ij} \leq h_j \qquad \forall i \in C, j \in C | i \neq j \qquad (14)$$
$$A3 : (i, t) \qquad \forall i \in C, t \in T \qquad (15)$$

$A1$ includes all the arcs from depots that are close enough to patients for vehicles to arrive by the established appointment time. Note that h_t represents departure times from depots; we assume that this value is equal to zero. $A2$ represents arcs between patients. Finally, as depots have no visit times, $A3$ states that all arcs from clients to depots are feasible.

Therefore, the set of feasible arcs is represented as follows:

$$A = A1 \cup A2 \cup A3 \qquad (16)$$

It is important to note that the condition sets (13) – (16) have a similar role to time window constraints. In this MDC-CVRmvt, independent of the arrival time, service is provided at the scheduled time. In real-life HHC problems, service must be provided at the scheduled time, instead of having an interval of possible visit times. Set A is computed prior to the solving process.

Figure 3 illustrates an example of the proposed model with a small instance. It consists of two depots (A and B), four clients, and one vehicle. To force the vehicle to travel a route from depot A to depot B, we define parameters $QI_A = 1$, $QF_A = 0$, $QI_B = 0$, and $QF_B = 1$. The final depot does not affect the objective function value. To simplify this example, we assume that visit time h_i is equal to accumulated travel time U_i. Parameter l_b is set equal to 150 time units. Travel times measured in time units are indicated at the edges of the diagram. Dashed lines represent emergency trips to the closest depot.

Dashed lines around patients indicate that overtime hours were used on those visits. The objective function value is equal to 25 time units. The route performed by the vehicle is A-C1-C2-C3-C4-B, and the normal workday time limit is exceeded during the visit to client C3. It is important to note that CU_i is equal to l_b, less the sum of the accumulated time traveled, U_i, and the emergency trip to the closest depot. Thus, if the emergency trip was not considered, then none of the visits would incur overtime.

III. RESULTS AND DISCUSSION

The proposed MILP model was implemented on AMPL and solved using the commercial solvers CPLEX 12.9.0 and Gurobi 8.1.0. The computer used had the following features: Intel Core TM i7-7700K, 4.2 GHz processor, 32 GB RAM, and a 64-bit RedHat Enterprice 8.0 operating system. Both solvers were used with default settings and with the time limit parameter set to $3600s$.

A. DATA GENERATION

To test the computational performance of the model, we created structured instances. These are combinations of variations in sets' cardinality and parameters. These features are described as follows:

- Clients (C) : $\{10, 20, 30, 40, 50\}$.
- Depots (T) : $\lceil \sqrt{C}/2 \rceil$.
- Vehicles(B) : $3T$, where $QI_t = QF_t = 3$
- Quota (l_b) : $\{100, 125, 150, 175, rand[125, 175], 200\}$.
- Geographic distribution of clients: random (R), clusters (CL), and geometric (G). An example of each can be seen in Figure [X]. Squares represent depots, and circles represent clients.
- Cluster density parameter (D) : $\{3, 6, 9, 12, 15\}$. This parameter affects only cluster-type instances. The smaller the value of D is, the denser the cluster. For a more comprehensive explanation of the cluster creation methodology, the reader may refer to [69].

Coordinates x and y for all nodes are in the range $[-100, 100]$. While the type of geographical distribution of the clients is variable, depots are fixed and independent of client distribution. The travel time parameter d_{ij} is assumed to be equal to the value of the Euclidean distance between points i and j. The visit time parameter h_i may define instance feasibility. To generate these values with a low risk of infeasibility, the methodology considers a number of parameters. First, for each client and depot, we define A_{ij} as the approximation of the next multiple of 15 of the travel time between client i and depot j. For example, if $d_{ij} = 40$, then A_{ij} is set equal to 45. Then, for each client, one of the $|T|$ parameters, A_{ij}, is randomly selected. The next step is to sort clients by their selected A_{ij} values and create G groups with m members. The number of groups created for each instance is $G = \{2, 3, 4, 5, 6\}$ for $C = \{10, 20, 30, 40, 50\}$. When the remainder of $C/G = 0$, we have G groups with $m = C/G$ members. Otherwise, we will have $G - 1$ groups with $m = \lceil C/G \rceil$ members and one group with $m = C - (G - 1)\lceil C/G \rceil$ members. Each member of each group is subject to variation parameter V_i, which affects the selected A_{ij}. The values of V_i are given by probabilities, as shown in Table 1. It should be noted that ϵ corresponds to an additional variation parameter, which is included to add variability to h_i values. In our experiments, we set $\epsilon = 15$.

Finally, clients' appointment time h_i is given by the sum of the selected approximation parameter and the variation that occurred. This parameter is expressed as follows:

$$h_i = A_{ij} + V_i \qquad (17)$$

TABLE 1. Probabilities of possible values of V_i according to problem size and assigned group.

C	Group	V_i: 0+ϵ	90+ϵ	180+ϵ	270+ϵ	360+ϵ	450+ϵ
10	1	70%	20%	10%			
	2	20%	50%	30%			
20	1	70%	20%	10%			
	2	20%	50%	30%			
	3		20%	50%	30%		
30	1	70%	20%	10%			
	2	20%	50%	30%			
	3		20%	50%	30%		
	4			20%	50%	30%	
40	1	70%	20%	10%			
	2	20%	50%	30%			
	3		20%	50%	30%		
	4			20%	50%	30%	
	5			10%	30%	60%	
50	1	70%	20%	10%			
	2	20%	50%	30%			
	3		20%	50%	30%		
	4			20%	50%	30%	
	5				20%	50%	30%
	6				10%	30%	60%

TABLE 2. Average CPU time (s), objective function value and GAP by instance size for each solver.

C	CPLEX CPU time	CPLEX Z	CPLEX GAP	Gurobi CPU time	Gurobi Z	Gurobi GAP
10	0.32	172.19	0.00%	0.40	172.19	0.00%
20	0.44	0.00	0.00%	0.23	0.00	0.00%
30	1582.00	69.36	30.18%	791.74	70.31	3.92%
40	1590.22	0.71	20.00%	751.79	0.36	20.00%

B. EXPERIMENT A: PROBLEM SIZE VS. CPU TIME

The first experiment consists of an exploratory analysis. We tested the model's performance by instance size and compared the results of both commercial solvers. Parameter l_b was set to 200 time units. This experiment does not contain $C = 50$. For each number of clients C, five instances with different random seeds are created (in total 20 instances). All instances are distributed randomly (R) in the plane.

The results summarized in Table 2 indicate that the problem's complexity grows with the number of clients. Instances with 10 and 20 patients are solved in a few seconds on average. However, for instances with $C = 30$ and 40, the CPU time increases considerably. We found that in terms of CPU time and GAP, Gurobi is more efficient than CPLEX for solving the analyzed instances. Another result worth noting is the objective function value in instances with 20 patients. A small Z value indicates that in these instances, little or no overtime hours are required for client visits. This can be explained by the combination of the quota's value and the number of vehicles. For instances with $C = 20$, $B = 9$, while for those with $C = 10$, $B = 6$, which is approximately 2 patients per vehicle. This implies a low risk of overtime.

C. EXPERIMENT B: VARIATIONS IN QUOTA PARAMETER

The objective of Experiment B is to determine the influence of l_b on model performance. This analysis was applied only to instances that presented an objective function value equal to 0 in Experiment A (with $l_b = 200$). Furthermore, we only used Gurobi. A total of 35 instances were analyzed, 7 for each value of $l_b = \{100, 125, 150, 175, rand[125, 175]\}$. Three of these 7 instances correspond to $C = 20$, two to $C = 30$ and two others to $C = 40$.

As the value of l_b increases, the objective function's value for each instance size C improves. This situation can be explained by the fact that penalization starts later when the value of l_b is larger. CPU time is demonstrated to be indirectly proportional to the value of l_b. As the quota parameter increases, the problem becomes less complex. Indeed, it is possible to observe a significant diminution in computational time between $l_b = 150$ and 175 and from 175 to 200. This phenomenon is visible in all instance sizes, but especially in $C = 30$ and 40. The results of this experiment are summarized in Figure 5. The first horizontal axis contains the number of clients, while the second contains l_b values. CPU time is represented by blue bars (left vertical axis), and Z is represented by green bars (right).

D. EXPERIMENT C: ANALYSIS OF GEOGRAPHICAL DISTRIBUTION TYPE

Here, we analyze whether the geographical distribution of clients influences model behavior. Additionally, both commercial solvers are compared. Instances include variations in the quota and the number of clients. While $l_b = \{150, rand[150, 200], 200\}$, $C = \{10, 20, 30, 40, 50\}$. For the random distribution case, five seeds were generated for each value of C. These seeds were used for the three quota values. In the cluster distribution, for each combination of number of patients C and quota l_b, five instances were created, each with different cluster density D. In the geometric case, only one seed for each number of clients was created and used for the different values of l_b. A total of 165 instances were created in this experiment: 75 for random distribution, 15 for geometric distribution, and 75 for clustering.

By analyzing the results in Table 3, we conclude that the geometric distribution is the most complex, while the cluster distribution is the easiest to solve. All small instances, that is, $C = 10$ and $C = 20$, were solved to optimality for random and cluster distributions but not for geometric distribution. By analyzing instances with 30 or more clients, we found that all cluster types were solved to optimality. By comparing random and geometric distributions, we found that in most cases, random distribution presents better average CPU times. For geometric type, none of the instances with $C = 30$ could be optimally solved. Regarding cluster distribution, the objective function was found to be equal to 0 for all instances analyzed. The above situation means that no overtime has occurred. Regarding cluster density, we found that the denser the cluster is, the less CPU time

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

required to solve the problem. The effect of density on model performance is easily visible in large instances. The behavior of cluster instances (and its difference compared to other distributions) can be explained by the level of closeness between clients and depots. The results of instances with 50 clients are summarized in Figure 6. Regarding solvers' benchmark, for instances with $C = 10$ and $C = 20$, we found that CPLEX performs better than Gurobi in geometric distribution and tight quotas. For $l_b = 200$, a significant difference in favor of Gurobi was observed. For instances with 30 and 40 clients, Gurobi requires considerably less computational time to solve most of the random and cluster instances. In geometric instances of these sizes, both solvers reach the time limit, but Gurobi presents lower GAPs than CPLEX. Furthermore, in three instances with $C = 30$, CPLEX runs out of memory. In these instances, the branch and bound algorithm explores a large number of nodes that are not able to be pruned, leading to a large number of nodes left to explore and high memory usage. In cluster instances with $C = 50$, Gurobi outperforms CPLEX. For other instances, no notable differences are found.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

TABLE 3. Benchmarking solvers: Average results by distribution type, problem size, and allowed quota.

l_b	C	Random						Geometric						Cluster					
		CPLEX			Gurobi			CPLEX			Gurobi			CPLEX			Gurobi		
		CPU (s)	Z	GAP %	CPU (s)	Z	GAP %	CPU (s)	Z	GAP %	CPU (s)	Z	GAP %	CPU (s)	Z	GAP %	CPU (s)	Z	GAP %
200	10	0.1	108.2	0.0	0.1	108.2	0.0	0.1	364.5	0.0	0.1	364.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	20	0.4	0.0	0.0	0.2	0.0	0.0	1,611.1	135.3	0.0	43.7	135.3	0.0	0.2	0.0	0.0	0.1	0.0	0.0
	30	2,271.0	38.1	41.8	939.1	38.1	20.0	Out of memory			3,600.0	619.0	17.0	1.5	0.0	0.0	0.2	0.0	0.0
	40	1,378.5	0.3	20.0	125.5	0.0	0.0	3,600.0	166.7	100.0	3,600.0	166.3	100.0	8.7	0.0	0.0	1.0	0.0	0.0
	50	3,600.0	169.7	100.0	3,600.0	117.7	100.0	3,600.0	755.1	99.2	3,600.0	834.6	92.1	140.4	0.0	0.0	6.0	0.0	0.0
150	10	0.1	262.4	0.0	0.2	262.4	0.0	0.1	700.6	0.0	0.1	700.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	20	40.8	98.6	0.0	308.2	98.6	0.0	14.9	587.9	0.0	296.0	587.9	0.0	0.2	0.0	0.0	0.2	0.0	0.0
	30	3,600.0 *	434.6*	82.4*	3,149.5	398.1	43.5	3,600.0	1,270.1	71.4	3,600.0	1,270.1	24.0	15.7	0.0	0.0	1.3	0.0	0.0
	40	3,600.0	212.7	97.5	3,600.0	210.8	94.4	3,600.0	893.2	100.0	3,600.0	986.8	86.7	131.4	0.0	0.0	9.1	0.0	0.0
	50	3,600.0	748.2	99.4	3,600.0	768.0	99.4	3,600.0	1,887.8	91.2	3,600.0	2,089.8	87.0	1,752.8	0.0	0.0	77.7	0.0	0.0
Rand[150,200]	10	0.3	130.6	0.0	0.4	130.6	0.0	0.3	450.5	0.0	0.3	450.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	20	2.4	0.3	0.0	2.9	0.3	0.0	1,393.1	231.4	0.0	3,600.0	231.4	20.9	0.2	0.0	0.0	0.1	0.0	0.0
	30	3,295.8*	46.6*	50.0*	2,968.0	84.8	40.0	3,600.0	816.3	90.3	3,600.0	813.3	86.9	1.9	0.0	0.0	0.6	0.0	0.0
	40	3,015.2	11.0	80.0	2,539.9	25.5	60.0	3,600.0	357.8	100.0	3,600.0	406.8	100.0	47.8	0.0	0.0	2.9	0.0	0.0
	50	3,600.0	375.7	100.0	3,600.0	483.2	100.0	3,600.0	1,179.1	100.0	3,600.0	1,515.2	95.1	806.8	0.0	0.0	114.8	0.0	0.0

*CPLEX runs out of memory with one of the five analyzed instances of this type.

E. EXPERIMENT D: INSTANCE SMOOTHING – INCREASE IN THE NUMBER OF DEPOTS AND SIZE OF THE FLEET

The final experiment involves tripling the number of depots ($smoothed(D)$) and vehicles ($smoothed(V)$). We consider both cases separately and together ($smoothed(V + D)$). Only instances with random and geometric distributions were analyzed because these reached the time limit with high GAP values in the previous experiment. The same instances as those in experiment C were smoothed: 75 instances of random type and 15 instances of geometric distribution. Furthermore, we used only Gurobi.

The reader can observe the results of this experiment in Table 4. In most of the cases studied, compared to the original situation, the proposed variations decrease the problem's complexity in terms of CPU time, GAP, and objective function value for both distribution types. By comparing the effect of depots and vehicles (geometric instances), we found that larger fleets lead to lower GAPS and CPU times in most cases. These results are consistent and can be validated by property 1 of the MDCCVRP described in [24]. Furthermore, there is no pattern that indicates the parameter with the greatest impact on the objective function value. The improvement produced by smoothing both parameters is greater than that by doing so separately. As the number of available vehicles and depots increases, the possibility of a caregiver using overtime hours to visit a client decreases. In the new scenario, each vehicle can perform shorter routes, which reduces the risk of incurring overtime. All instances were solved to optimality in considerably low computing times for the $smoothed(V + D)$ case.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

TABLE 4. Comparison between original and smoothed instances.

l_b	C	Geometric											
		Original			$Smoothed(V+D)$			$Smoothed(V)$			$Smoothed(D)$		
		CPU (s)	Z	GAP	CPU (s)	Z	GAP	CPU (s)	Z	GAP	CPU (s)	Z	GAP
200	10	0.1	364.5	0%	0.0	5.8	0%	0.2	280.0	0%	0.1	153.5	0%
	20	43.7	135.3	0%	0.2	0.0	0%	1.7	45.6	0%	1.2	45.3	0%
	30	3,600.0	619.0	17%	0.6	0.0	0%	3,587.8	45.7	60%	3,600.0	338.2	31%
	40	3,600.0	166.3	100%	1.2	0.0	0%	4.2	5.8	0%	859.4	0.0	0%
	50	3,600.0	834.6	92%	5.1	0.0	0%	9.4	5.8	0%	3,600.0	266.7	100%
150	10	0.1	700.6	0%	0.6	143.2	0%	0.2	566.2	0%	0.1	398.0	0%
	20	296.0	587.9	0%	1.7	60.5	0%	2,702.3	293.2	0%	67.8	242.1	0%
	30	3,600.0	1,270.1	24%	13.6	70.8	0%	3,600.0	332.4	26%	3,600.0	861.5	41%
	40	3,600.0	986.8	87%	11.9	17.7	0%	3,600.0	202.5	76%	3,600.0	319.5	100%
	50	3,600.0	2,089.8	87%	34.5	17.7	0%	3,600.0	235.9	76%	3,600.0	1,013.6	97%
Rand[150,200]	10	0.3	450.5	0%	0.0	9.8	0%	5.5	323.2	0%	0.4	192.3	0%
	20	3,600.0	231.4	21%	0.2	0.0	0%	5.1	62.6	0%	19.7	55.5	0%
	30	3,600.0	813.3	87%	0.5	0.0	0%	3,600.0	62.7	15%	3,600.0	507.8	91%
	40	3,600.0	406.8	100%	2.9	0.0	0%	3.9	7.8	0%	3,600.0	77.0	100%
	50	3,600.0	1,515.2	95%	6.7	0.0	0%	10.8	7.8	0%	3,600.0	716.8	100%

l_b	C	Random†								
		Original			$Smoothed(V+D)$			$Smoothed(V)$		
		CPU (s)	Z	GAP	CPU (s)	Z	GAP	CPU (s)	Z	GAP
200	10	0.1	108.2	0%	0.0	0.0	0%	0.1	81.9	0%
	20	0.2	0.0	0%	0.2	0.0	0%	0.0	0.0	0%
	30	939.1	38.1	20%	0.6	0.0	0%	0.4	2.6	0%
	40	125.5	0.0	0%	1.5	0.0	0%	1.1	0.0	0%
	50	3,600.0	117.7	100%	4.4	0.0	0%	2.8	0.0	0%
150	10	0.2	262.4	0%	0.1	10.5	0%	0.3	201.9	0%
	20	308.2	98.6	0%	0.2	0.0	0%	2.6	41.6	0%
	30	3,149.5	398.1	44%	173.6	7.7	0%	1,508.5	100.5	12%
	40	3,600.0	210.8	94%	2.4	0.3	0%	38.9	29.6	0%
	50	3,600.0	768.0	99%	5.7	0.0	0%	2,864.5	71.3	63%
Rand[150,200]	10	0.4	130.6	0%	0.0	0.0	0%	0.2	90.9	0%
	20	2.9	0.3	0%	0.2	0.0	0%	0.1	0.0	0%
	30	2,968.0	84.8	40%	0.6	0.0	0%	0.7	4.2	0%
	40	2,539.9	25.5	60%	1.1	0.0	0%	0.7	0.0	0%
	50	3,600.0	483.2	100%	3.9	0.0	0%	3.5	0.0	0%

$^\dagger smoothed(D)$ is not included. Half of the cases became infeasible, which can lead to skewed results

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

F. SUMMARY OF THE BENCHMARKING OF COMMERCIAL SOLVERS

A total of 165 instances were used to compare the commercial solvers CPLEX v12.9 and Gurobi v8.1 in experiments A and C. While Gurobi exhibits better computational performance in 90 instances (55%), CPLEX does better in $27 (16\%)$. In 48 instances, (29%), the solvers are found to have a similar performance, which is mainly due to the fact that they both reach the time limit in the more complex instances. Nevertheless, Gurobi presents lower GAPs than CPLEX in most of these instances. Another important conclusion from the comparison is illustrated in Figure 7, which presents a graphical analysis of the benchmarking by plotting $CPLEX_{CPUtime} - Gurobi_{CPUtime}$ for each instance. Thus, instances with negative values indicate that CPLEX is more efficient than Gurobi, while positive values indicate the opposite. The reader can observe not only that Gurobi has better behavior in more instances but also that when this solver outperforms CPLEX, the magnitude of the difference in CPU time is considerable. Figure 7 presents only those instances in which differences in CPU time are greater than 50 s. For further details of all instances, the reader may check the supplementary materials. Instances' names are constructed as follows: distribution type, number of patients, and quota. For cluster instances, density D is also included at the end. For example, instance CL50-200-D9 is a type-one cluster distribution with $C = 50$, $l_b = 200$, and $D = 9$.

IV. CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

In this paper, we introduced a new variant of the CCVRP. It corresponds to the nonfixed MDCCVRPmvt with minimum delayed latency. Contrary to similar CCVRPs found in the literature, our problem includes features present in real-life HHC routing problems. These characteristics include fulfilling mandatory visit times, multiple depots, and emergency trips in case patients need special attention at a hospital. The novel objective function used helps define routes by the minimizing overtimes hours with which clients are visited, leading to better service quality. A novel MILP model to address this problem is proposed, the performance of which is analyzed with new structured instances. The main idea is to simulate different potential real-life cases. We find that instance size is directly proportional to computation time. Most instances with 10 and 20 clients are solved with relative ease by both solvers. In our analysis of patients' geographical distribution, we conclude that geometric is the most difficult instance type to address, while cluster distribution is the easiest. The density parameter also influences model performance. The denser the cluster is, the better the model's behavior. It is important to note that instances of up to 50 clients are solved optimally for cluster cases. For the other two distribution types, most instances with $C = 40$ and 50 reach the time limit without finding the optimal solution. In some of these cases, the GAP is 100%. Regarding the effect of parameters on the problem's complexity, we find that tighter quotas make the instance more difficult to solve. By increasing the number of depots and the fleet size, the complexity can be reduced. This situation can be explained by the improved workload distribution that additional vehicles/depots provide. The benchmarking between CPLEX and Gurobi indicates that Gurobi requires less computational time in most of the instances analyzed. Remarkable differences in CPU time and GAP are found in favor of Gurobi for several instances.

Although we present this novel model for HHC applications, it can easily be extended to other problems as follows:

- In forest fire control (considering planes with more than one water discharge mechanism). In this case, hot spots are clients. Depending on when a fire starts, hot spots have a defined maximum time to be reached by plane. The parameter l_b could correspond to the maximum flight time, which could be limited by gas availability or other technical constraints. Planes could stay on route even when the quota is reached. Nevertheless, it is a dangerous situation, the risk of which increases with each new hot spot visited. Emergency trips to the closest depot must be performed when the plane runs out of water or gas or when it faces other technical emergencies.
- Humanitarian logistics in a postdisaster context. Natural disasters such as earthquakes have the potential to affect wide swaths of the population. Rescue vehicles must take certain routes to pick up these individuals and take them to shelters. In this context, a good response time is essential, and the goal is to reach all the people before a maximum time, l_b, after the event occurrence. If a person needs emergency medical attention, then an emergency trip to the closest facility must be made.
- In the delivery of frozen/fresh products or warm bread. Note that in many Latin American countries, it is common to eat fresh, warm bread at breakfast and dinner. In both cases, products have a certain time available before the cold chain is broken or they get cold. The desirable outcome is to deliver products that have not had their quality negatively impacted by transport. Standard time l_b violations are expected to be minimized. The emergency trip could be related to clients who reject the delivery if, for example, the order is incomplete or the products are in poor condition. This would imply a trip to the closest depot to correct the order.
- Truck/bus companies that provide scheduled delivery or pickup services. In this case, l_b corresponds to the maximum number of hours that drivers should be at the wheel. They can drive more hours, but it increases the risk of accidents. In the same industry, l_b can also correspond to the number of km at which each vehicle should undergo routine maintenance. Vehicles can be on the road without maintenance, but it increases the risk of operational failures, which can, in turn, lead to accidents or scheduling problems. In this case, since l_b, U_i^b and

CU_i are distances, a simple conversion using a speed parameter is needed.

For future research, we propose adding new features and components to the model to improve its representation of the real world. Such components can include service time, different types of services, or precedence constraints. The proposed model can also be extended to a cumulative HHCRSP. A multiobjective approach that allows for the study of the tradeoff between delayed latency and total cost/time-based objectives may be adequate. Furthermore, new solution methods to handle more complex instances are also an interesting possibility.

REFERENCES

[1] R. Çolak, A. H. Işik, and T. Yiğit, "A NEW METHOD FOR ROUTING IN HOME HEALTH CARE SERVICES," BRAIN. BROAD RESEARCH IN ARTIFICIAL INTELLIGENCE AND NEUROSCIENCE, vol. 10, no. 4, pp. 14–27, dec 2019. [Online]. Available: https://lumenpublishing.com/journals/index.php/brain/article/view/2072/pdf

[2] E. V. Gutiérrez and C. J. Vidal, "Home health care logistics management problems: a critical review of models and methods," Revista Facultad de Ingeniería Universidad de Antioquia, vol. 68, pp. 160–175, 2013. [Online]. Available: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-62302013000300016

[3] M. Á. Suarez Cuba, "ATENCIÓN INTEGRAL A DOMICILIO," Rev. Méd. La Paz [online], vol. 18, no. 2, pp. 52–58, 2012. [Online]. Available: http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S1726-89582012000200010&lng=es&nrm=iso

[4] P. Toth and D. Vigo, Eds., Vehicle Routing: Problems, Methods, and Applications, 2nd ed. Philadelpia: Society for Industrial and Applied Mathematics, 2014.

[5] G. B. Dantzig and J. H. Ramser, "The Truck Dispatching Problem," Management Science, vol. 6, no. 1, pp. 80–91, oct 1959.

[6] F. Semet, P. Toth, and D. Vigo, "Classical Exact Algorithms for the Capacitated Vehicle Routing Problem," in Vehicle Routing: Problems, Methods, and Applications, 2nd ed., P. Toth and D. Vigo, Eds. Society for Industrial and Applied Mathematics, 2014, ch. 2, pp. 37–57.

[7] M. Poggi and Uchoa. Eduardo, "New Exact Algorithms for the Capacitated Vehicle Routing Problem," in Vehicle Routing: Problems, Methods, and Applications, 2nd ed., P. Toth and D. Vigo, Eds. Society for Industrial and Applied Mathematics, 2014, ch. 3, pp. 59–86.

[8] G. Laporte, S. Ropke, and T. Vidal, "Heuristics for the Vehicle Routing Problem," in Vehicle Routing: Problems, Methods, and Applications, 2nd ed., P. Toth and D. Vigo, Eds. Society for Industrial and Applied Mathematics, 2014, ch. 4, pp. 87–116.

[9] A. A. R. Hosseinabadi, A. Slowik, M. Sadeghilalimi, M. Farokhzad, A. K. Sangaiah et al., "An ameliorative hybrid algorithm for solving the capacitated vehicle routing problem," IEEE Access, vol. 7, pp. 175 454–175 465, 2019.

[10] S. H. Taieb, T. Loukil, and A. E. Mhamedi, "Home (Health)-Care Routing and Scheduling Problem," in International Colloquium on Logistics and Supply Chain Management, LOGISTIQUA 2019. Institute of Electrical and Electronics Engineers Inc., jun 2019.

[11] K. Sundar and S. Rathinam, "Algorithms for routing an unmanned aerial vehicle in the presence of refueling depots," IEEE Transactions on Automation Science and Engineering, vol. 11, no. 1, pp. 287–294, jan 2014.

[12] W. P. Coutinho, M. Battarra, and J. Fliege, "The unmanned aerial vehicle routing and trajectory optimisation problem, a taxonomic review," Computers & Industrial Engineering, vol. 120, pp. 116 – 128, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360835218301815

[13] K. F. Doerner, M. Gronalt, R. F. Hartl, G. Kiechle, and M. Reimann, "Exact and heuristic algorithms for the vehicle routing problem with multiple interdependent time windows," Computers and Operations Research, vol. 35, no. 9, pp. 3034–3048, sep 2008.

[14] J. Widuch, "Current and emerging formulations and models of real-life rich vehicle routing problems," in Smart Delivery Systems. Elsevier, 2020, pp. 1–35.

[15] J. Renaud, G. Laporte, and F. F. Boctor, "A tabu search heuristic for the multi-depot vehicle routing problem," Computers and Operations Research, vol. 23, no. 3, pp. 229–235, 1996.

[16] J. Li, Y. Li, and P. M. Pardalos, "Multi-depot vehicle routing problem with time windows under shared depot resources," Journal of Combinatorial Optimization, vol. 31, no. 2, pp. 515–532, feb 2016.

[17] N. Sharma and M. Monika, "A literature survey on multi-depot vehicle routing problem," Int. J. Res. Dev, vol. 3, pp. 1752–1757, 2015.

[18] J. R. Montoya-Torres, J. L. Franco, S. N. Isaza, H. F. Jiménez, and N. Herazo-Padilla, "A literature review on the vehicle routing problem with multiple depots," Computers & Industrial Engineering, vol. 79, pp. 115–129, 2015.

[19] J. Zhang, F. Yang, and X. Weng, "An evolutionary scatter search particle swarm optimization algorithm for the vehicle routing problem with time windows," IEEE Access, vol. 6, pp. 63 468–63 485, 2018.

[20] G. Li and J. Li, "An improved tabu search algorithm for the stochastic vehicle routing problem with soft time windows," IEEE Access, vol. 8, pp. 158 115–158 124, 2020.

[21] G. Desaulniers, O. B. Madsen, and S. Ropke, "The Vehicle Routing Problem with Time Windows," in Vehicle Routing: Problems, Methods, and Applications, 2nd ed., P. Toth and D. Vigo, Eds. Society for Industrial and Applied Mathematics, 2014, ch. 5, pp. 119–159.

[22] I. Kara, B. Y. Kara, and M. K. Yetiş, "Cumulative vehicle routing problems," in Vehicle routing problem, T. Caric and H. Gold, Eds. In-Teh, 2008, ch. 6, pp. 85–98.

[23] S. U. Ngueveu, C. Prins, and R. Wolfler Calvo, "An effective memetic algorithm for the cumulative capacitated vehicle routing problem," Computers and Operations Research, vol. 37, no. 11, pp. 1877–1885, nov 2010.

[24] E. Lalla-Ruiz and S. Voß, "A POPMUSIC approach for the Multi-Depot Cumulative Capacitated Vehicle Routing Problem," Optimization Letters, vol. 14, no. 3, pp. 671–691, apr 2020. [Online]. Available: https://doi.org/10.1007/s11590-018-1376-1

[25] J. F. Sze, S. Salhi, and N. Wassan, "The cumulative capacitated vehicle routing problem with min-sum and min-max objectives: An effective hybridisation of adaptive variable neighbourhood search and large neighbourhood search," Transportation Research Part B: Methodological, vol. 101, pp. 162–184, jul 2017.

[26] E. Rusminingsih, S. Istiqomah, T. Romadhani et al., "What is the impact of burnout on nurses work behavior in nursing care?" in 1st International Conference on Science, Health, Economics, Education and Technology (ICoSHEET 2019). Atlantis Press, 2020, pp. 448–452.

[27] W. Emiliano, J. Telhada, and M. d. S. Carvalho, "Home health care logistics planning: a review and framework," Procedia Manufacturing, vol. 13, pp. 948–955, jan 2017.

[28] R. Liu, X. Xie, V. Augusto, and C. Rodriguez, "Heuristic algorithms for a vehicle routing problem with simultaneous delivery and pickup and time windows in home health care," European Journal of Operational Research, vol. 230, no. 3, pp. 475–486, nov 2013.

[29] A. Trautsamwieser, M. Gronalt, and P. Hirsch, "Securing home health care in times of natural disasters," OR Spectrum, vol. 33, no. 3, pp. 787–813, jul 2011.

[30] S. R. Ait Haddadene, N. Labadie, and C. Prodhon, "The VRP With Time Windows, Synchronization and Precedence Constraints: Application In Home Health Care Sector," in MOSIM 2014, 10ème Conférence Francophone de Modélisation, Optimisation et Simulation, Nancy, France, 2014. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01166650

[31] N. Labadie, C. Prins, and Y. Yang, "Iterated Local Search for a Vehicle Routing Problem with Synchronization Constraints," in ICORES 2014, Angers, France, 2014, pp. 257–263.

[32] L. En-nahli, S. Afifi, H. Allaoui, and I. Nouaouri, "Local Search Analysis for a Vehicle Routing Problem with Synchronization and Time Windows Constraints in Home Health Care Services," IFAC-PapersOnLine, vol. 49, no. 12, pp. 1210–1215, jan 2016.

[33] D. Bredström and M. Rönnqvist, "Combined vehicle routing and scheduling with temporal precedence and synchronization constraints," European Journal of Operational Research, vol. 191, no. 1, pp. 19–31, nov 2008.

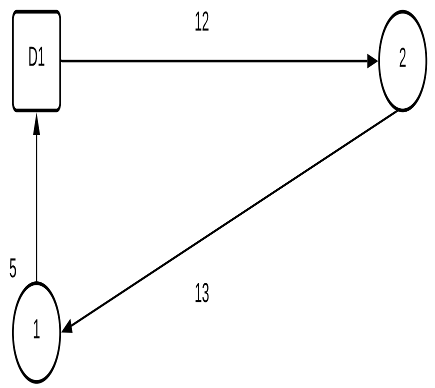
[34] S. R. Haddadene, N. Labadie, and C. Prodhon, "NSGAII enhanced with a local search for the vehicle routing problem with time windows and synchronization constraints," IFAC-PapersOnLine, vol. 49, no. 12, pp. 1198–1203, 2016.

[35] M. Cissé, S. Yalçındağ, Y. Kergosien, E. Şahin, C. Lenté, and A. Matta, "Or problems related to home health care: A review of relevant routing and scheduling problems," Operations Research for Health Care, vol. 13, pp. 1–22, 2017.
[36] H. Allaoua, S. Borne, L. Létocart, and R. Wolfler Calvo, "A matheuristic approach for solving a home health care problem," Electronic Notes in Discrete Mathematics, vol. 41, pp. 471–478, 2013.
[37] L. En-nahli, H. Allaoui, and I. Nouaouri, "A multi-objective modelling to human resource assignment and routing problem for home health care services," IFAC-PapersOnLine, vol. 48, no. 3, pp. 698–703, 2015.
[38] K. Braekers, R. F. Hartl, S. N. Parragh, and F. Tricoire, "A bi-objective home care scheduling problem: Analyzing the trade-off between costs and client inconvenience," European Journal of Operational Research, vol. 248, no. 2, pp. 428–443, jan 2016.
[39] R. Linfati, J. W. Escobar, and J. Escalona, "A Two-Phase Heuristic Algorithm for the Problem of Scheduling and Vehicle Routing for Delivery of Medication to Patients," Mathematical Problems in Engineering, vol. 2018, 2018. [Online]. Available: https://www.hindawi.com/journals/mpe/2018/8901873/
[40] J. Decerle, O. Grunder, A. H. El Hassani, and O. Barakat, "A two-phases matheuristic for the home care routing and scheduling problem," IFAC-PapersOnLine, vol. 49, no. 12, pp. 1484–1489, jan 2016.
[41] C. Fikar and P. Hirsch, "Home health care routing and scheduling: A review," Computers & Operations Research, vol. 77, pp. 86–95, 2017.
[42] A. Blum, P. Chalasani, D. Coppersmith, B. Pulleyblank, P. Raghavan, and M. Sudan, "The minimum latency problem," in Proceedings of the Annual ACM Symposium on Theory of Computing, vol. Part F129502. Association for Computing Machinery, may 1994, pp. 163–171.
[43] M. M. Silva, A. Subramanian, T. Vidal, and L. S. Ochi, "A simple and effective metaheuristic for the Minimum Latency Problem," European Journal of Operational Research, vol. 221, no. 3, pp. 513–520, sep 2012.
[44] J. N. Tsitsiklis, "Special cases of traveling salesman and repairman problems with time windows," Networks, vol. 22, no. 3, pp. 263–282, may 1992. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/net.3230220305
[45] M. Fischetti, G. Laporte, and S. Martello, "Delivery man problem and cumulative matroids," Operations Research, vol. 41, no. 6, pp. 1055–1064, dec 1993. [Online]. Available: https://pubsonline.informs.org/doi/abs/10.1287/opre.41.6.1055
[46] G. Mattos Ribeiro and G. Laporte, "An adaptive large neighborhood search heuristic for the cumulative capacitated vehicle routing problem," Computers and Operations Research, vol. 39, no. 3, pp. 728–735, mar 2012.
[47] P. Chen, X. Dong, and Y. Niu, "An iterated local search algorithm for the cumulative capacitated vehicle routing problem," Advances in Intelligent and Soft Computing, vol. 136, pp. 575–581, 2012. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-27711-5_76
[48] L. Ke and Z. Feng, "A two-phase metaheuristic for the cumulative capacitated vehicle routing problem," Computers and Operations Research, vol. 40, no. 2, pp. 633–638, feb 2013.
[49] F. B. Ozsoydan and A. Sipahioglu, "Heuristic solution approaches for the cumulative capacitated vehicle routing problem," Optimization, vol. 62, no. 10, pp. 1321–1340, oct 2013. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/02331934.2013.841158
[50] L. Ke, "A brain storm optimization approach for the cumulative capacitated vehicle routing problem," Memetic Computing, vol. 10, no. 4, pp. 411–421, dec 2018. [Online]. Available: https://doi.org/10.1007/s12293-018-0250-0
[51] S. Nucamendi-Guillén, F. Angel-Bello, I. Martínez-Salazar, and A. E. Cordero-Franco, "The cumulative capacitated vehicle routing problem: New formulations and iterated greedy algorithms," Expert Systems with Applications, vol. 113, pp. 315–327, dec 2018.
[52] N. Smiti, M. M. Dhiaf, B. Jarboui, and S. Hanafi, "Skewed general variable neighborhood search for the cumulative capacitated vehicle routing problem," International Transactions in Operational Research, vol. 27, no. 1, pp. 651–664, jan 2020. [Online]. Available: http://doi.wiley.com/10.1111/itor.12513
[53] J. Lysgaard and S. Wøhlk, "A branch-and-cut-and-price algorithm for the cumulative capacitated vehicle routing problem," European Journal of Operational Research, vol. 236, no. 3, pp. 800–810, aug 2014.
[54] X. Wang, T. M. Choi, H. Liu, and X. Yue, "A novel hybrid ant colony optimization algorithm for emergency transportation problems during post-disaster scenarios," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 48, no. 4, pp. 545–556, apr 2018.
[55] X. Wang, T.-M. Choi, Z. Li, and S. Shao, "An Effective Local Search Algorithm for the Multidepot Cumulative Capacitated Vehicle Routing Problem," IEEE Transactions on Systems, Man, and Cybernetics: Systems, pp. 1–11, sep 2019.
[56] I. Martínez-Salazar, F. Angel-Bello, and A. Alvarez, "A customer-centric routing problem with multiple trips of a single vehicle," Journal of the Operational Research Society, vol. 66, no. 8, pp. 1312–1323, aug 2015. [Online]. Available: www.palgrave-journals.com/jors/
[57] J. C. Rivera, H. Murat Afsar, and C. Prins, "Mathematical formulations and exact algorithm for the multitrip cumulative capacitated single-vehicle routing problem," European Journal of Operational Research, vol. 249, no. 1, pp. 93–104, feb 2016.
[58] J. C. Rivera, H. M. Afsar, and C. Prins, "A multistart iterated local search for the multitrip cumulative capacitated vehicle routing problem," Computational Optimization and Applications, vol. 61, no. 1, pp. 159–187, may 2015. [Online]. Available: https://link.springer.com/article/10.1007/s10589-014-9713-5
[59] S. Nucamendi-Guillén, D. Flores-Díaz, E. Olivares-Benitez, and A. Mendoza, "A Memetic Algorithm for the Cumulative Capacitated Vehicle Routing Problem including Priority Indexes," Applied Sciences, vol. 10, no. 11, p. 3943, jun 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/11/3943
[60] R. Liu and Z. Jiang, "A hybrid large-neighborhood search algorithm for the cumulative capacitated vehicle routing problem with time-window constraints," Applied Soft Computing Journal, vol. 80, pp. 18–30, jul 2019.
[61] M. Huang, K. R. Smilowitz, and B. Balcik, "A continuous approximation approach for assessment routing in disaster relief," Transportation Research Part B: Methodological, vol. 50, pp. 20–41, apr 2013.
[62] M. E. Bruni, S. Khodaparasti, and P. Beraldi, "The selective minimum latency problem under travel time variability: An application to post-disaster assessment operations," Omega (United Kingdom), vol. 92, apr 2020.
[63] C. Sabo, M. Kumar, K. Cohen, and D. Kingston, "VRP with minimum delivery latency using Linear Programming," in AIAA Infotech at Aerospace Conference and Exhibit 2012, 2012.
[64] D. R. Gaur, A. Mudgal, and R. R. Singh, "Routing vehicles to minimize fuel consumption," Operations Research Letters, vol. 41, no. 6, pp. 576–580, 2013.
[65] D. R. Gaur and R. R. Singh, "Cumulative vehicle routing problem: A column generation approach," in Algorithms and Discrete Applied Mathematics, S. Ganguly and R. Krishnamurti, Eds. Cham: Springer International Publishing, 2015, pp. 262–274.
[66] ——, "A heuristic for cumulative vehicle routing using column generation," Discrete Applied Mathematics, vol. 228, pp. 140–157, 2017.
[67] T. Liu, X. Tang, H. Wang, H. Yu, and X. Hu, "Adaptive hierarchical energy management design for a plug-in hybrid electric vehicle," IEEE Transactions on Vehicular Technology, vol. 68, no. 12, pp. 11 513–11 522, 2019.
[68] Y. Qin, X. Tang, T. Jia, Z. Duan, J. Zhang, Y. Li, and L. Zheng, "Noise and vibration suppression in hybrid electric vehicles: State of the art and challenges," Renewable and Sustainable Energy Reviews, vol. 124, p. 109782, 2020.
[69] E. Uchoa, D. Pecin, A. Pessoa, M. Poggi, T. Vidal, and A. Subramanian, "New benchmark instances for the Capacitated Vehicle Routing Problem," European Journal of Operational Research, vol. 257, no. 3, pp. 845–858, 2017.

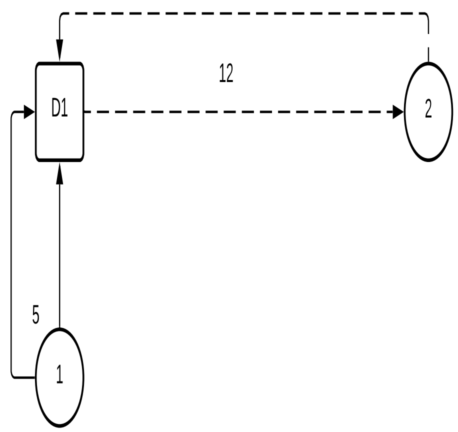
ACKNOWLEDGMENT

To Universidad del Bío-Bío (UBB), Universidad Andrés Bello (UNAB), and Universidad de Santiago de Chile (USACH), which supported this research.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access



(a)



(b)

FIGURE 1. Optimal solution by objective function. (**a**) Minimum total cost/time. (**b**) Minimum latency.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

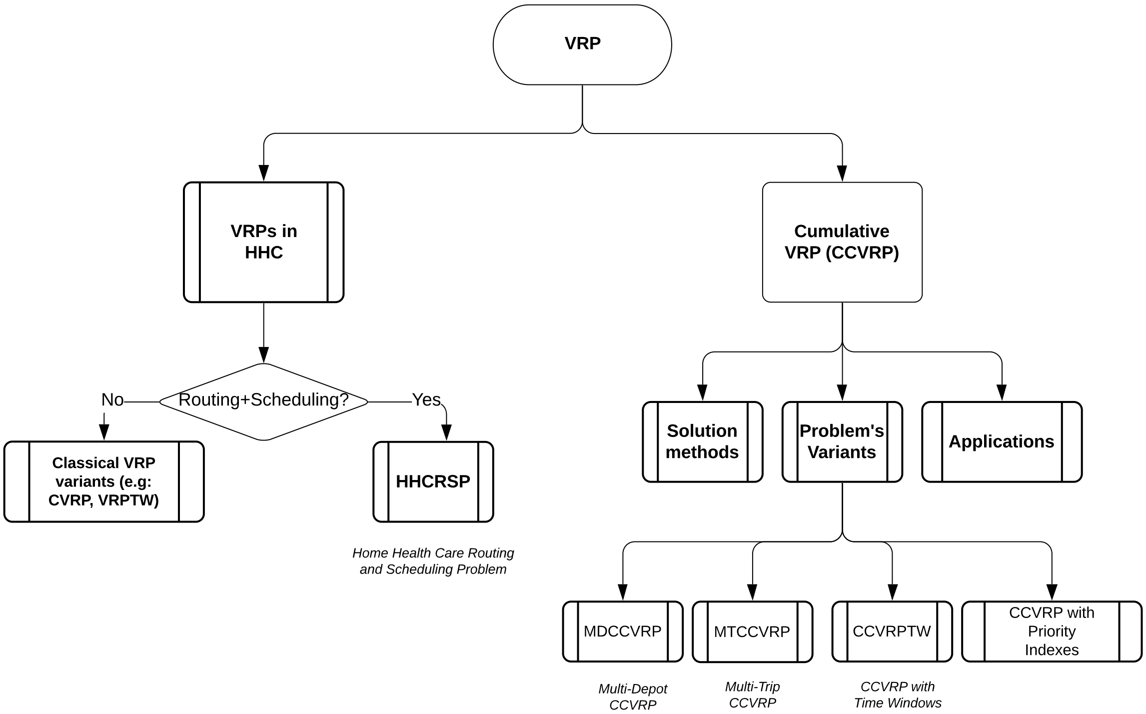


FIGURE 2. Structure of the literature review.

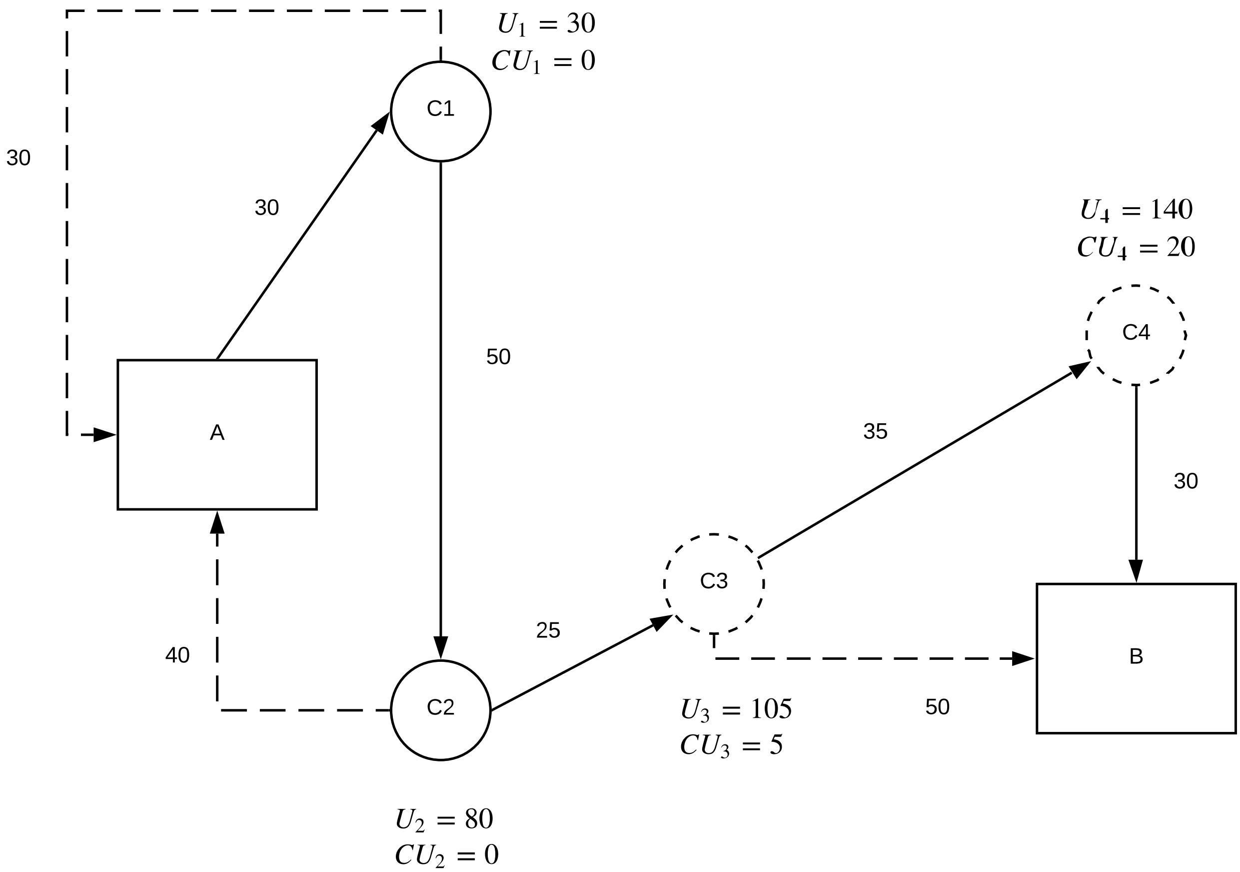


FIGURE 3. Explanatory example of the problem.

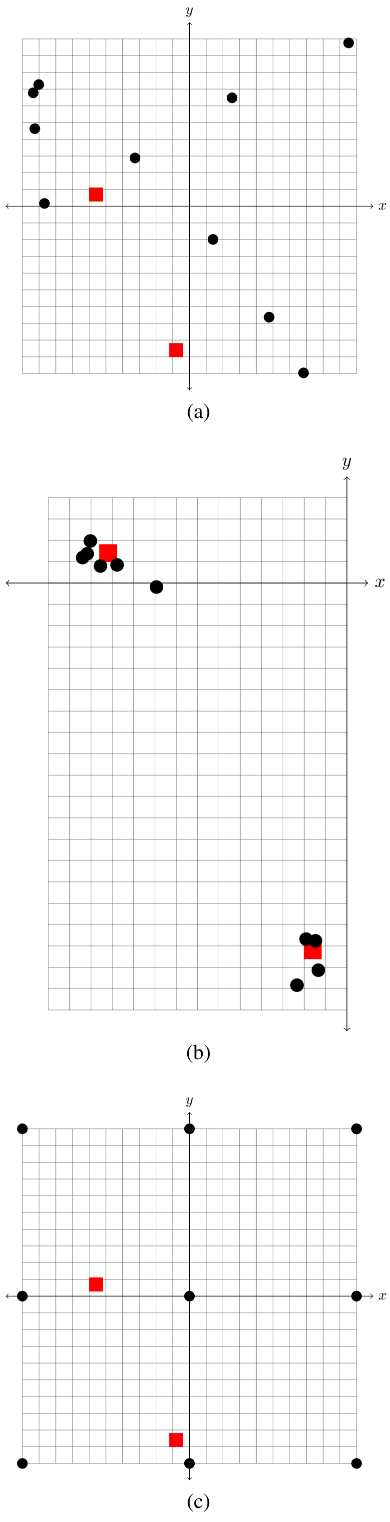


FIGURE 4. Geographical distribution types. (**a**) Random. (**b**) Cluster. (**c**) Geometric.

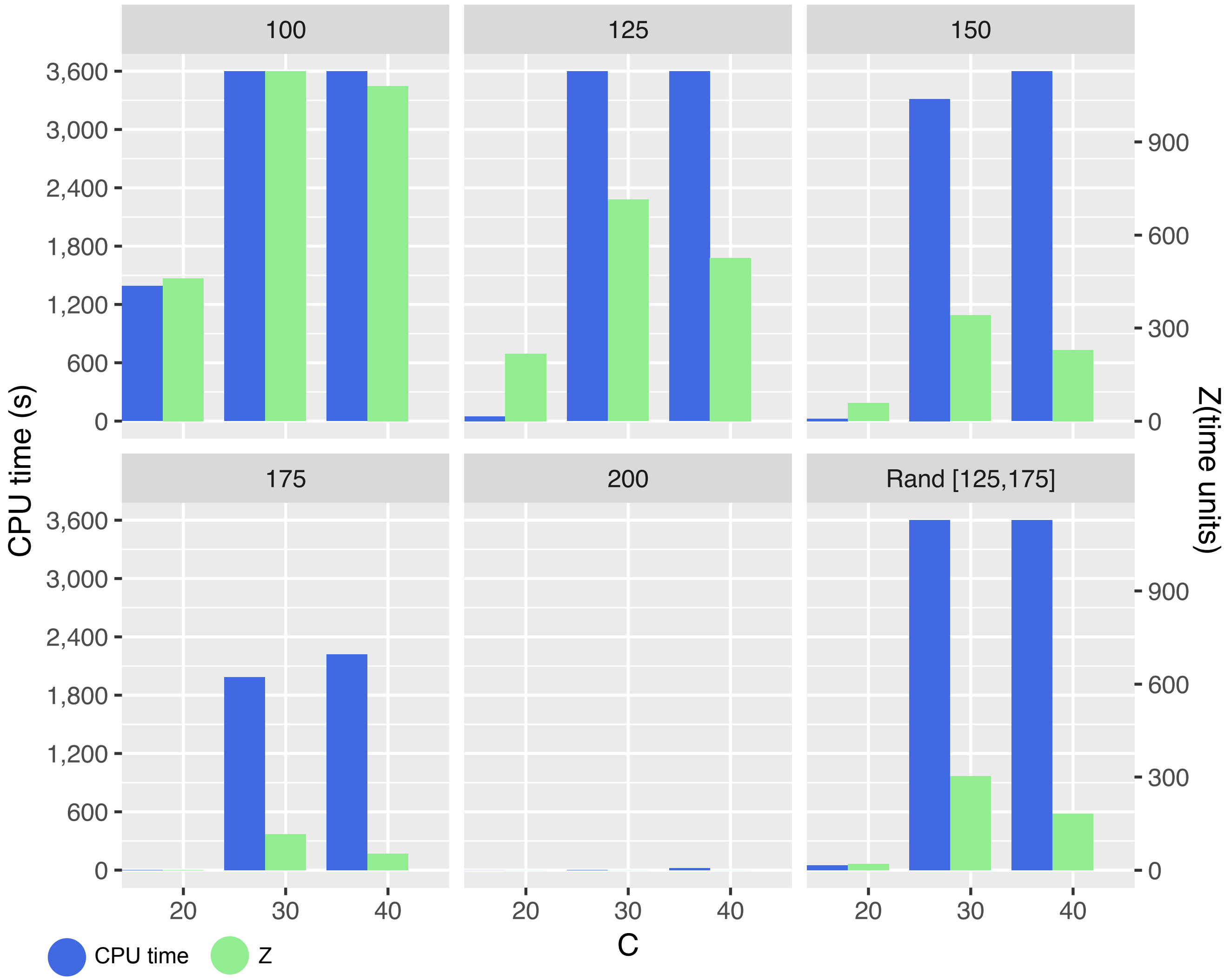


FIGURE 5. Average CPU time and objective function value by allowed quota and problem size.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

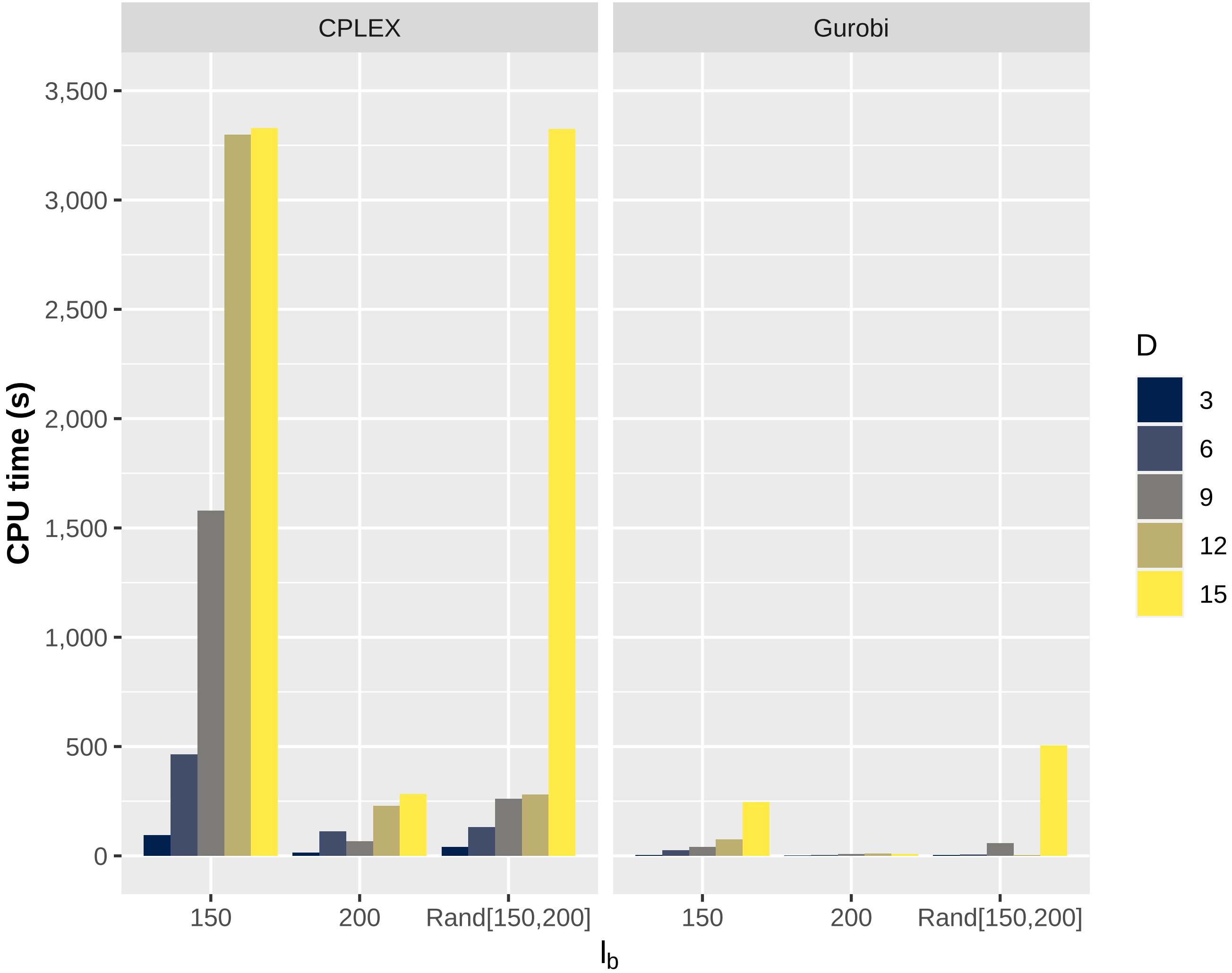


FIGURE 6. Average CPU time by cluster density, allowed quota, and solver for instances with $C = 50$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3058242, IEEE Access

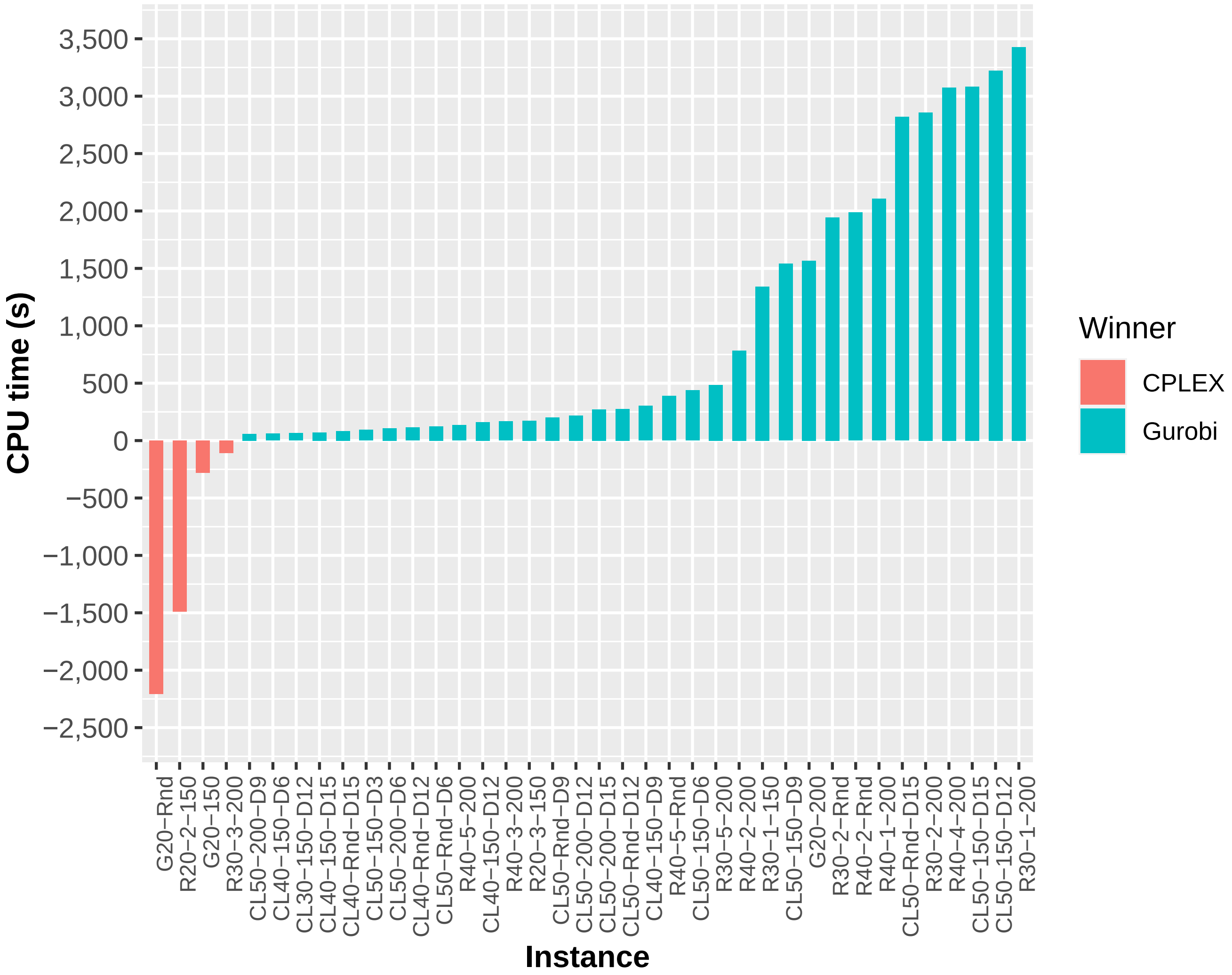


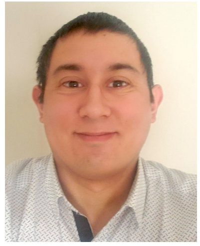
FIGURE 7. Difference in CPU time (CPLEX-Gurobi) by instance.

ALAN OSORIO-MORA was born in Concepción, Chile in 1994. He received the B.S and M.S degrees in Industrial Engineering from the University of Bío-Bío, Chile, in 2018 and 2019 respectively. From 2018 to 2020 he worked teaching and researching at the Department of Industrial Engineering, University of Bío-Bío, Chile. In 2020 he was visiting professor/researcher at the Department of Civil and Industrial Engineering, Pontifical Xavierian University of Cali, Colombia. He is currently pursuing the Ph.D. degree in Biomedical, Electrical and System Engineering with the University of Bologna, Italy. His interest research areas include logistics problems such as location and routing problems, their applications and algorithms to solve them efficiently. He is also interested on MIP and combinatorial optimization.

MURICIO SOTO-BUSTOS received the B.S degree in Industrial Engineering from the University of Bío-Bío, Chile, in 2019. From 2018 to 2019 he worked as research assistant at the Department of Industrial Engineering, University of Bío-Bío, Chile. His research interests include mathematical models for addressing real-world optimization problems in the industry, focusing on logistics.

GUSTAVO GATICA is a Computer Engineer from the University of La Serena and Ph.D in Engineering Sciences from the University of Santiago de Chile. He began his academic career in 2003, he is currently a researcher at the Andrés Bello University (Chile). His research interests include combinatorial optimization, integer programming, data science, vehicle routing, graph coloring and logistics.

PEDRO PALOMINOS holds a PhD in Industrial Engineering from Technical University of Catalonia, Spain and MS degrees in Federal University of Rio de Janeiro, Brazil. Currently he is Professor at the Department of Industrial Engineering and Director of the Smart City Program of the University of Santiago of Chile. His research interests include operations management, metaheuristics, smart cities and multi-criteria analysis.



RODRIGO LINFATI was born in Angol, Chile in 1983. He received the B.S. and M.S. degrees in industrial engineering from the University of Concepción, Chile, in 2008 and the Ph.D. degree in Operational Research from the University of Bologna, Italy, in 2012. Since 2012, he has been an Assistant Professor with the Industrial Engineering Department, University of Bío-Bío, Chile. He has published his works in prestigious international journals on the fields of operational research, transportation research, industrial engineering, manufacturing, and operations management. His research interests include mathematical models and metaheuristic algorithms for addressing real-world optimization problems in the industry, focusing on scheduling, supply chain management, cutting stock, vehicle routing, location, and logistics.