Production, Manufacturing and Logistics

A practical vehicle routing problem with desynchronized arrivals to depot

Zahra Naji-Azimi¹, Majid Salari¹, Jacques Renaud⁴, Angel Ruiz⁴

¹Department of Management, Ferdowsi University of Mashhad, P.O. Box 91779-48551, Mashhad, Iran
²Department of Industrial Engineering, Ferdowsi University of Mashhad, P.O. Box 91779-48551, Mashhad, Iran
³Département opérations et systèmes de décision, Faculté des sciences de l’administration, Pavillon Palais-Prince, Université Laval, 2325, rue de la Terrasse, Québec, G1V 0A6, Canada
⁴Centre interuniversitaire de recherche sur les réseaux d’entreprise, la logistique et le transport (CIRRELT), Québec, Canada

A R T I C L E   I N F O

Article history:
Received 12 June 2015
Accepted 7 April 2016
Available online 13 April 2016

Keywords:
Routing
Multiple routes
Time windows
Desynchronized arrivals

A B S T R A C T

The transportation of biomedical samples is a key component of healthcare supply chains. The samples are collected, consolidated into cooler boxes, and then transported to be analyzed in a specialized laboratory. Since many hospitals and samples’ collection points are assigned to the same laboratory, it is important to manage the flow of samples arriving to the laboratory to avoid congestion. In other words, it is preferable to try to desynchronize the samples’ arrivals by managing the vehicles’ departure times and the routes ordering. We propose a mathematical model and a multi-start heuristic to minimize the route duration times and the maximum number of samples’ boxes arriving at the laboratory within a given time period. Based on real data, we demonstrated that both the model and the heuristic are very efficient in solving real size instances.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Vehicle routing and scheduling are key components of the efficiency of modern supply chains, where quantities of goods and raw materials should be continuously exchanged in the seamless way possible. In its most common version, the vehicle routing problem (VRP) is used to plan the distribution of goods from a depot to a set of customers, (Köc, Bekta, Jabali, & Laporte, 2016, Laporte, 2009, Semet, Toth, & Vigo, 2014) and subject to some constraints, like time windows (Bráysy & Gendreau, 2005a, 2005b), vehicle restrictions (Semet, 1995) and many other practical and industrial considerations (Coelho, Renaud, & Laporte, 2016). In other situations, products are exchanged between the customers, which leads to the pickup and delivery VRP (Berbeglia, Cordeau, & Laporte, 2010, Gschwind, 2015; Berbeglia, Cordeau, Gribkovskaia, & Laporte, 2007). Vehicles can also be used to bring back the customers’ goods to a central depot, or consolidation point, like in the waste collection problem (Ghiani, Lagana, Manni, Musmanno, & Vigo, 2014). These practical routing problems deal with many real-world constraints and are often referred as rich VRP (Lahyani, Khemakhem, & Semet, 2015).

In this article, we focus on a situation arising in the healthcare supply chain context which corresponds to the daily transportation of biomedical samples from hospitals or clinics, which will be referred to as collection points (CP), to a single laboratory (Lab) where they will be analyzed. Transportation is done by a fleet of vehicles performing multiple routes during the day. As the lifespan of these samples is limited, CPs often require multiple collection requests on a given day, and each collection request is generally bounded by a time window. Samples are consolidated in cold boxes that, once collected, must arrive to the Lab within a limited specified time to preserve samples’ integrity. The boxes are opened at the Lab, and each sample is manually registered in the system and bar-coded for tracking purposes during the analysis process. These manual tasks are time consuming and according to our partner’s experts, the Quebec’s Ministère de la santé et des services sociaux (Ministry of Health and Social Services – MSSS), they constitute a bottleneck of the samples’ supply chain. In fact, if too many boxes arrive in a short period of time, samples are queued and may suffer long wait, which might exceed their remaining lifetime making them unsuitable to be analyzed. Our observations on the ground confirmed that some days or periods are more popular for sample collection. Thereby, reducing the maximum number of samples boxes arriving within a time period is desirable in order to normalize the Lab workload and minimize sample losses. Thus, the objective of

http://dx.doi.org/10.1016/j.ejor.2016.04.007
0377-2217/© 2016 Elsevier B.V. All rights reserved.
this vehicle routing problem with desynchronized arrivals at the depot (VRP-DA) can be defined as minimizing the sum of the routes’ traveling time as well as the maximum number of boxes arriving at the Lab within any time period of the considered planning day. This work is motivated by the reengineering of the health supply chain performed by the Quebec’s government under the name of the Optilab project (MSSS, 2012). One of the first objective was to centralize the transportation needs of the collection points to obtain better tariffs from the carriers (Anaya Arenas, Chabot, Renaud, & Ruiz, 2016). Now the MSSS seeks at reducing the number of laboratories in order to better use the strategic ones. In order to do so, it clearly appears that a good management of samples arriving to the Lab would be an important issue. To the best of our knowledge, our contribution is the first one considering desynchronizing the vehicles’ arrivals to the depot as this is closely related to samples’ limited lifetime, which is not the case for most of classical or industrial goods.

Instead of regularizing the arrivals, many works have tried to balance the vehicles’ workload. Jozefowiez, Semet, and Talbi (2009) considered a bi-objective VRP which minimized the total length of routes and the balance of routes, which is defined as the difference between the maximal route length and the minimal route length. Kritikos and Ioannou (2010) studied a VRP with time windows (VRPTW) where they balanced the load carried by each active vehicle. Baños, Ortega, Gil, Márquez, and de Toro (2013) also considered a multi-objective VRPTW with both distance and load imbalances. López-Sánchez, Hernández-Díaz, Vigo, Caballero, and Molina (2014) solved a balanced open VRP where the maximum time spent on the vehicle must be minimized. We note that, if all vehicles depart at the same time, balancing the routes will indeed concentrate the vehicle arrival times at the depot.

Closer to our context, Doerner, Gronalt, Hartl, Kiechle, and Reimann (2008) and Doerner and Hartl (2008) studied a blood collection problem which shares some characteristics with our problem. However, in their study they considered that the blood deterioration process begins right after the donation, and they used interdependent time windows. They also did not consider arrival times at the depot. Liu, Xie, Augusto, and Rodríguez (2013) considered the pickup and delivery of four demand types in a home health care system. They dealt with many specific constraints, but they did not consider a time limit for the samples’ delivery to the depot, nor did they manage vehicles’ arrival times. Sahinyyazan, Kara, and Taner (2015) studied a system composed of bloodmobiles and shuttles which brought the collected blood to the depot to prevent spoilage. In order to maximize the collected blood and minimize the transportation cost, they managed the activities of the bloodmobiles and shuttles over a time horizon. However, they did not consider a time limit for the blood’s return to the depot. Finally, and as stated before, this particular context was introduced by Anaya Arenas et al. (2016) which minimized traveling distance, but did not consider desynchronizing the arrivals at the Lab.

The remainder of this article is as follows. In Section 2 we propose the problem formulation and some valid inequalities designed to improve its solvability. A multi-start heuristic is developed in Section 3. The efficiency of this formulation and of the heuristic and the impact of desynchronized arrivals are evaluated in Section 4, based on a set of real instances obtained from the MSSS of Quebec, Canada. We also demonstrate how this new formulation and heuristic improve upon those of Anaya Arenas et al. (2016), where desynchronized arrivals were not considered. Our conclusions are presented in Section 5.

2. Problem definition and formulation

In order to formulate the vehicle routing problem with desynchronized arrivals, we need to identify both the locations of the CPs and the collections requests. The n CPs are represented by \( V = \{ v_1, \ldots, v_n \} \) and each CP \( i \) requires \( Q_i \) collection requests, leading to a total of \( p = \sum_{i=1}^{n} Q_i \) requests. Each request is composed of one box that contains several samples. Then we define a complete graph \( G = (V, E) \), where \( V = \{ v_0, v_1, \ldots, v_p \} \) is the set of nodes in the network, which includes the laboratory as nodes \( \{ v_0, v_{p+1} \} \) where every route must start and end, and the set \( P = \{ v_1, v_2, \ldots, v_p \} \), being the \( p \) transportation requests. Also, we note \( R_i \) as the set of request nodes in \( V \) which corresponds to the same CP location. We consider the arc set \( A = \{(v_i, v_j) : v_i, v_j \in V, i \neq j, i = 0, \ldots, p, j = 1, \ldots, p+1 \} \) and a travel time \( (t_{ij}) \) and a travel distance \( (d_{ij}) \) are assigned to each arc \( (v_i, v_j) \). Clearly, \( t_{ij} \) and \( d_{ij} \) are equal to zero for every \( (v_i, v_j) \) if \( i = 0 \) and \( j \neq 0 \) (i.e., \( i = 0 \) and \( j \) correspond to two requests from the same CP). In addition, each request needs to be served within a time window \([a_j, b_j]\). Finally, any two requests related to the same collection point cannot be on the same route.

\( K \) uncapacitated vehicles are available for satisfying the transportation requests, and each vehicle can perform multiple routes \((r = 1, \ldots, R)\) within a work shift, but a limit on the length of the working day \( T_k \) must be respected. In addition, we need to consider a loading time \( (\tau_r) \) for each transportation request, as well as the vehicle’s unloading time \( (\tau_l) \) at the Lab before a new route can be started. Furthermore, let \( T_{max} \) be the maximal transportation time for the samples associated to request \( i \). The objective is to minimize the total traveling time of the vehicle, including the waiting times, plus a weighted penalty \( \theta \) associated to the maximum number of boxes arriving at the depot during the most visited time period. In the following sections we present the VRP-DA formulation followed by some valid inequalities to strengthen it.

2.1. VRP-DA formulation

The following decision variables are needed to define the VRP-DA:

- \( x_{ijk}^r \) Binary variable equal to 1 if vehicle \( k \) travels from request \( i \) to request \( j \) in its route \( r \); 0 otherwise.
- \( u_{ikr} \) Continuous variable that indicates the visit time (start of loading) of transportation request \( i \) by vehicle \( k \) in route \( r \).
- \( y_{ikr} \) Binary variable equal to 1 if the request \( i \) performed by the \( r \)th route of the vehicle \( k \), arrives at the Lab within the \( r \)th time period; 0 otherwise.
- \( w \) Highest number of boxes arriving to the Lab within any given time period (the busiest one).

The model VRP-DA reads as follows.

\[
\begin{align*}
\text{Min} \quad & \sum_{k=1}^{K} \sum_{r=1}^{R} \left( u_{p+1kr} - u_{0kr} \right) + \theta w \\
\text{Subject to:} \\
\sum_{k=1}^{K} \sum_{r=1}^{R} x_{ijk}^r &= 1 \quad j = 1, \ldots, p \\
\sum_{j=1}^{n} \sum_{k=1}^{K} x_{ijk}^r &\leq 1 \quad l = 1, \ldots, n; \ k = 1, \ldots, K; \ r = 1, \ldots, R \\
\sum_{i=0}^{p+1} \sum_{l=1}^{p} x_{ijk}^r - \sum_{i=0}^{p+1} \sum_{l=1}^{p} x_{ijl}^r &= 0 \quad j = 1, \ldots, p; \ k = 1, \ldots, K; \ r = 1, \ldots, R \\
\sum_{j=1}^{p} x_{0jk}^r &\leq 1 \quad k = 1, \ldots, K; \ r = 1, \ldots, R
\end{align*}
\]
\[
\sum_{j=1}^{p} x_{0jkr} - \sum_{j=1}^{p} x_{j,p+1kr} = 0 \quad k = 1, \ldots, K; \quad r = 1, \ldots, R
\]  
(6)

\[
a_{j} - T_{k} \left(1 - \sum_{i=0}^{p} x_{jkr}\right) \leq u_{jkr} \leq b_{j} + T_{k} \left(1 - \sum_{i=0}^{p} x_{jkr}\right) \quad j = 1, \ldots, p; \quad k = 1, \ldots, K; \quad r = 1, \ldots, R
\]  
(7)

\[
u_{jkr} - u_{jkr} + (b_{j} + \tau_{r} + t_{ij} - a_{j}) x_{jkr} \leq b_{j} - a_{j} \quad i = 0, \ldots, p; \quad j = 1, \ldots, p + 1;
\]  
(8)

The vehicle must return to the depot after a certain number of time periods. Constraints (10) ensure that, if request \( i \) has been visited by the \( r \)th vehicle, the variable \( y_{jkr} \) can take the value one. Constraints (14) and (15) force the relation between the flow and \( y \) variables and discretize time into periods of \( \omega \) units of time. When using these constraints in the case where \( y_{jkr} = 1 \), we have \( \omega t \leq u_{jkr} < \omega t + 1 \), meaning that the vehicle must be visited within the \( t \)th time period, thus within time \( \omega t \) and \( \omega t + 1 \). Constraints (16) calculate the Lab’s maximum workload during the available time periods. As expressed by (16), the number of boxes arriving during a given time period corresponds to the number of locations visited by the vehicle during the period. Assuming that a sample box is collected at each request location, the workload then corresponds to the number of boxes arriving at the depot within the period. Domains of the variables are given by (17)–(20).

### 2.2. Valid inequalities

The solvability of the model (1)–(20) can be improved by the addition of the following groups of inequalities.

\[
\sum_{j=1}^{p} x_{0jkr} - \sum_{j=1}^{p} x_{0jkr-1} \leq 0 \quad k = 1, \ldots, K; \quad r = 2, \ldots, R
\]  
(21)

\[
u_{jkr} - \sum_{j=0}^{K} \sum_{r=1}^{R} (a_{j} + \tau_{r} - a_{i} + t_{ij}) x_{jkr} \geq a_{i} \quad i = 1, \ldots, p + 1; \quad k = 1, \ldots, K; \quad r = 1, \ldots, R
\]  
(22)

\[
\sum_{j=1}^{p} x_{0jkr} - \sum_{j=1}^{p} x_{jkr} - 1 \leq 0 \quad k = 2, \ldots, K
\]  
(23)

\[
\sum_{j=1}^{R} \sum_{r=1}^{R} x_{0jkr} - \sum_{j=1}^{R} \sum_{r=1}^{R} x_{jkr} - 1 \leq 0 \quad j = 1, \ldots, p; \quad k = 2, \ldots, K
\]  
(24)

\[
\sum_{i=1}^{R} \sum_{r=0}^{T} y_{itkr} \leq M \sum_{i=1}^{R} \sum_{r=1}^{T} y_{itkr-1} \quad k = 2, \ldots, p
\]  
(25)

\[
\sum_{i=0}^{R} \sum_{r=0}^{T} y_{itkr} \leq M \sum_{i=1}^{R} \sum_{r=1}^{T} y_{itkr-1} \quad k = 1, \ldots, K; \quad r = 2, \ldots, R
\]  
(26)

\[
\sum_{j=1}^{K} \sum_{r=1}^{R} x_{jkr} \leq |S| - 1 \quad S \subseteq \{1, 2, \ldots, p\}, |S| = 2 \text{ or } 3
\]  
(27)

\[
\sum_{k=1}^{K} \sum_{r=1}^{R} x_{jkr} = 0 \quad \forall i, j \in V \setminus \{0, p + 1\} \left(\tau_{i} + \tau_{j} + t_{ij} > b_{j}\right)
\]  
(28)

\[
\sum_{k=1}^{K} \sum_{r=1}^{R} x_{jkr} = 0 \quad \forall i, j \in V \setminus \{0, p + 1\} \left(\tau_{i} + \tau_{j} + t_{ij} + p + 1 > T_{\max}\right)
\]  
(29)
\[
\sum_{k=1}^{K} \sum_{r=1}^{R} \sum_{l=1}^{t-1} v_{ijk} = 0 \quad \forall i \in V \setminus \{0, p + 1\} \left( a_{i} + r_{i} + t_{i, p+1} \geq \omega t \right) ; \\
t = 1, \ldots, T - 1 \tag{30}
\]

To avoid symmetry, active routes are ordered by constraint \((21)\). Constraint \((22)\) puts a lower bound on the minimum value of each variable \(u_i\). Essentially, when visiting the arc \((i, j)\) by the \(r\)th route of vehicle \(k\) we have \(x_{ijk} = 1\) and consequently, \(u_{ijk} - t_i + t_j \leq u_{jk} \). In case that \(x_{ijk} = 0\) we have \(u_{ijk} - t_i + t_j \leq b_i - a_j\) which is always valid. Constraints \((23)\) and \((24)\) are adaptations of the symmetry breaking constraints proposed by Coelho and Laporte (2014). Essentially, they eliminate many symmetric solutions of same values by ordering the use of the vehicles and the assignment of requests to vehicles. Constraints \((25)\) and \((26)\) are symmetry defeating constraints and are only introduced to enhance the model by breaking the symmetry caused by variables \(y\). In particular, none of the requests can be performed by routes of the vehicle \(k\), whenever the first route of the vehicle \(k\) – 1 has not already been used. In addition, \((26)\) states that route \(r\) of the vehicle \(k\) can be used to deliver a request to the Lab only if route \(r - 1\) of the same vehicle has already been used. Constraint \((27)\) is a classical sub-tours elimination constraint generated for subsets of two and three requests. Constraints \((28)\) to \((30)\) are generated only if specific conditions are respected. Constraint sets \((28)\) and \((29)\) state that we cannot visit arc \((i, j)\) by different routes of available vehicles when \(a_i + r_i + t_{ij} > b_j\) and \(a_j + r_j + t_{ji, p+1} - b_i \geq T_{max}\), respectively. In fact, \((28)\) and \((29)\) respectively remove redundant arcs violating the time window and maximum sample travel time. Constraint \((30)\) shows the relation between the earliest arrival time of a request and the possible time periods where its corresponding sample can be delivered at the Lab. Essentially, when \(a_i + r_i + t_{ip+1} \geq \omega t\), request \(i\) cannot be delivered to the Lab sooner than the \(r\)th\hbox{th} time period.

Model \((1)-(30)\) extends and strengthens the one proposed by Anaya Arenas et al. (2016) who used constraints \((2)-(7), (9)-(11), (21), (23), (24)\) and the following connectivity constraint:

\[
u_{ikr} + r_i + t_{ij} - u_{jk} \leq T_k \left( 1 - x_{ijk} \right) \quad i = 0, \ldots, p; j = 1, \ldots, p + 1; \quad k = 1, \ldots, K; r = 1, \ldots, R \tag{31}
\]

We improved this constraint by using \((8)\) and \((22)\). In their model, they also minimized the total traveled distance which is a different objective function than \((1)\). Nonetheless, the formulation is only able to solve efficiently small to medium sized instances, as it will be shown in the section devoted to numerical experiments. The development of a fast and efficient solving method was therefore necessary to deal with the larger real-life instances provided by our partner.

### 3. Heuristic algorithm

In this section, we develop a multi-start algorithm for the introduced problem. In the next sections we first describe the algorithm, followed by a detailed description of how visiting times are updated efficiently to handle the time windows and the maximal sample transportation time constraints.

#### 3.1. Multi-start algorithm

The multi-start algorithm is based on three procedures, namely Construction, Extraction-Reinsertion and Swap. To ensure that different executions lead to different solutions, the algorithm sets a level of randomization which modifies two important parameters of the problem: the maximal sample transportation time \((T_{max})\) and the maximum length of a vehicle working day \((T_k)\). These parameters impact the Construction algorithm, which may provide different initial feasible or unfeasible solutions. Parameters are adjusted according to the feasibility of the solution produced at the previous execution. Only feasible solutions are passed to the improvement steps.

**Construction procedure**

Initial solutions are built by a constructive method in which nodes are sequentially added to the routes. We use the following rules to select \(n_i\), the first node to be visited by a route.

- \(N_1 = \text{argmax}_{i} \{ t_0 \} \) i.e., the set of nodes \(i \in P = \{ v_1, v_2, \ldots, v_p \}\) whose travel time from the Lab is the greatest.
- \(N_2 = \text{argmin}_{i \in N_1} \{ b_i \} \) i.e., the set of nodes \(i \in N_1\) whose time window upper bound is the lowest.
- \(n_1 = \text{Choose a node from } N_2\) randomly.

At each step, to add a new node to the set of visited nodes of the working route, we verify the possibility of adding each unvisited node to every insertion place of the current route and select the insertion leading to the smallest increase in the route’s total travel time. For instance, if node \(i\) is inserted between nodes \(j\) and \(k\), the detour is computed as \(t_j + t_k - t_{ij}\). The general framework of the construction phase is sketched in Algorithm 1, where \(K\) is the number of vehicles and \(R\) the maximum number of routes per vehicle.

**Extraction-Reinsertion procedure**

In this procedure, the goal is to reduce the value of the objective function \((1)\) by repositioning some of the nodes. Thus, the length of the time period \(\omega\) must be considered in the calculation of this cost. To try improving the solution, the procedure extracts a node from its location and reinserted into its best feasible position. Starting from the first route of the first vehicle, all the nodes are repositioned in all possible locations. A move is accepted as soon as it leads to an improvement, and the whole procedure is repeated until no improvement can be reached. During this procedure, we are allowed to create a new route or close an already existing one.

**Swap**

Following their order in the current solution, we consider each pair of nodes and their corresponding positions are swapped. This swap is applied to all possible combinations of two nodes over all the vehicles’ routes. As soon as a move improves the solution’s cost, it is accepted and the procedure stops whenever the swapping of all available nodes offers no more improvement.

**Multi-start heuristic algorithm**

The general framework of the multi-start algorithm is provided in Algorithm 2 and consists of two loops. During the inner loop’s execution, the goal is to construct an initial feasible solution. Within this loop, we run the Construction procedure by setting different temporary values for the maximum vehicle travel time \((T_k)\) and the sample travel time \((T_{max})\) parameters until a feasible solution is obtained (the Repeat - Until loop). Essentially, the algorithm is initialized at the first iteration by setting \(T_{surv} = T_{max}\) and \(T_{det} = T_k\), which are the initial feasible parameters of the instance. For the other iterations, if the generated solution produced by applying the Construction procedure is feasible, the values of \(T_{surv}\) and \(T_{det}\) are decreased by a factor \(\alpha\) which is an input parameter (set to 0.1 in our computational experiments):

\[
T_{surv} = T_{surv} - \alpha T_k \\
T_{det} = T_{det} - \alpha T_{max}
\]

Otherwise, in order to increase the chance of obtaining a feasible solution, the corresponding values of the parameters are
increased by applying the updates:

\[ T_{\text{max}} = \min \{ T_{\text{max}} + \alpha T \text{max} : T \text{max} \} \]

Finally, the values to be used as the maximum travel and sample times when running the Construction procedure are set by applying the following relations in which \( \text{rand} (0, x) \) is a random integer number between 0 and x.

\[ \overline{T}_k = \text{rand} (0, T_k - \text{Temp}_k) + \text{Temp}_k \]

\[ T_{\text{max}} = \text{rand} (0, T_{\text{max}} - \text{Temp}_{\text{max}}) + \text{Temp}_{\text{max}} \]

Upon obtaining an initial feasible solution, we apply the Extraction-Reinsertion and Swap procedures to try to improve the quality of the initial solution. To do so, we use the original values for the sample travel time (i.e. \( T_{\text{max}} \)) and vehicle travel time (i.e. \( T_k \)). The algorithm stops after a given number of iterations (i.e. \( \text{Max}_{\text{iter}} \)).

3.2. Information update

In order to efficiently manage the time windows and the maximum sample transportation time constraints, we need rules to update the earliest and the latest start time of nodes in the solution. In the following, we offer an adaptation of the method proposed by Campbell and Savelsbergh (2004) for the VRP with time windows.

We represent the \( r \)th route of vehicle \( k \) by \( k_r = \{ c_{0k_r}^r, c_{1k_r}^r, \ldots, c_{j_kr}^r, c_{j_kr+1}^r \} \) in which \( n_{k_r} \) is the number of nodes visited by the \( r \)th route of vehicle \( k \). We denote by \( E_{c_{j_kr}}^r \) and \( L_{c_{j_kr}}^r \), the earliest and the latest time at which node \( c_{j_kr}^r \) can be visited by the \( r \)th route of vehicle \( k \), respectively.

At the beginning, for each \( k \) and \( r \), we set \( E_{c_{0k_r}}^r = E_{c_{j_kr+1}^r}^r = a_0 \) and \( L_{c_{0k_r}}^r = L_{c_{j_kr+1}^r}^r = b_0 \). In case of visiting node \( c_{j_kr}^r \) between nodes \( c_{i_kr}^r \) and \( c_{i+1k_r}^r \) we apply the relations (R1) and (R2) to update the earliest and latest visit time of \( c_{j_kr}^r \) as follows:

\[ E_{c_{j_kr}}^r = \max \{ a_{c_{j_kr}}, E_{c_{j_kr-1}^r} + \tau_{i_{j_kr}} + t_{i_{j_kr+1}^r} \} \quad (R1) \]

\[ L_{c_{j_kr}}^r = \min \{ b_{c_{j_kr}}, L_{c_{j_kr-1}^r} - \tau_{i_{j_kr}} - t_{i_{j_kr+1}^r} \} \quad (R2) \]

If \( E_{c_{j_kr}}^r \leq L_{c_{j_kr}}^r \), inserting \( c_{j_kr}^r \) between \( c_{i_kr}^r \) and \( c_{i+1k_r}^r \) will not violate the time windows, but it is still necessary to verify that both the maximum duration of the vehicle length and the maximum sample travel time are satisfied. To do so, we use the relations (R3) and (R4) to update the earliest and latest start time of nodes visited after and before \( c_{j_kr}^r \) as follows:

\[ E_{c_{s_i}}^r = \max \{ a_{c_{s_i}}, E_{c_{s_{i-1}^r}} + \tau_{i_{s_i}} + t_{i_{s_{i+1}^r}} \} \quad (R3) \]

\[ L_{c_{s_i}}^r = \min \{ b_{c_{s_i}}, L_{c_{s_{i-1}^r}} - \tau_{i_{s_i}} - t_{i_{s_{i+1}^r}} \} \quad (R4) \]

Finally, the minimum length of the \( r \)th route of vehicle \( k \) requires serving the route at the latest feasible time, by setting it to \( v_{c_{s_{n_{k_r}^r}+1}^r} - v_{c_{s_i}^r} \). In this expression, \( v_{c_{s_i}^r} \) is the real visit time of
First, we compare our formulation and heuristic to those in Anaya Arenas et al. (2016). Then, we evaluate the model’s behavior and the heuristic in the context of desynchronized arrivals.

### 4.1 Results on biomedical sample transportation instances

The proposed model VRP-DA of Section 2 can be used to solve the BSTP if we use constraints (2)–(15) and replace the objective function (1) by the distance minimization objective

\[
\sum_{i=0}^{n-1} \sum_{j=1}^{K} d_{ij} x_{ij}.
\]

These results are reported under columns VRP-DA in the following tables. Columns BSTP-EG correspond to the Anaya Arenas et al. (2016) model. Table 1 reports data on the small and medium instances. These instances have up to 10 collection points and 20 transportation requests.

As can be observed, both formulations solved all the instances to optimality. For medium instances, the BSTP-EG proposed by Anaya Arenas et al. (2016) required 98 seconds in averages, while our formulation required only 0.96.

Table 2 reports the results for the larger instances having up to 20 collection points and 50 requests. We ran the BSTP-EG model, but it was never able to produce any feasible solution after 10 800 seconds of computing. This behavior was also observed by Anaya Arenas et al. (2016) and this is why they initialized their model with the best solution given by their heuristics. Thus, column BSTP-EG in Table 2 reports their original results (distance and final Cplex gap in percentage after 10 800 seconds). Column VRP-DA reports our results after 3 600 and 10 800 seconds of computing time, respectively. Values marked by an asterisk indicate instances for which Cplex ran out of memory and the best feasible solutions for them are reported.

The new model clearly offers a better performance, obtaining better solutions for all the cases except instances 33 and 38. The average solution cost was reduced from 1388 for the previous model to respectively 1322 and 1290 after 3600 and 10 800 seconds of computing time with the new VRP-DA model. Average optimality gap was also reduced from 21.51 percent for BSTP-EG to 8.77 percent for VRP-DA. Thus, the new formulation is shown to be faster and more effective.

Table 3 reports heuristics’ results for all the 38 instances. Column “Arenas-Arenas” reports their best found results over different heuristic combinations. Columns “Heuristic” report the results of the algorithm developed in Section 4 after 1 and 100 iterations, respectively. Computing times are not reported, as they are negligible. Our multi-start algorithm can be applied to the BSTP by using the travel time as a cost function for the Extraction-Reinsertion and Swap procedures and by setting the value of \( \theta = 0 \). For each heuristic, we report the distance (Dist) and the gap (Gap), with

### Table 1

Performance of the mathematical models on small and medium instances.

<table>
<thead>
<tr>
<th>#</th>
<th>Dist</th>
<th>BSTP-EG Second</th>
<th>VRP-DA Second</th>
<th>#</th>
<th>Dist</th>
<th>BSTP-EG Second</th>
<th>VRP-DA Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>193.9</td>
<td>0.05</td>
<td>0.04</td>
<td>13</td>
<td>754.4</td>
<td>2.76</td>
<td>0.26</td>
</tr>
<tr>
<td>2</td>
<td>125.3</td>
<td>0.05</td>
<td>0.04</td>
<td>14</td>
<td>2303.9</td>
<td>0.49</td>
<td>0.21</td>
</tr>
<tr>
<td>3</td>
<td>311.8</td>
<td>0.06</td>
<td>0.08</td>
<td>15</td>
<td>2340.0</td>
<td>1.32</td>
<td>0.28</td>
</tr>
<tr>
<td>4</td>
<td>235.8</td>
<td>0.07</td>
<td>0.07</td>
<td>16</td>
<td>1260.0</td>
<td>1.30</td>
<td>0.59</td>
</tr>
<tr>
<td>5</td>
<td>324.2</td>
<td>0.13</td>
<td>0.08</td>
<td>17</td>
<td>193.0</td>
<td>0.78</td>
<td>0.52</td>
</tr>
<tr>
<td>6</td>
<td>270.7</td>
<td>0.19</td>
<td>0.18</td>
<td>18</td>
<td>193.0</td>
<td>0.87</td>
<td>0.36</td>
</tr>
<tr>
<td>7</td>
<td>279.9</td>
<td>0.51</td>
<td>0.30</td>
<td>19</td>
<td>284.7</td>
<td>4.44</td>
<td>0.39</td>
</tr>
<tr>
<td>8</td>
<td>267.9</td>
<td>0.27</td>
<td>0.17</td>
<td>20</td>
<td>301.3</td>
<td>3.71</td>
<td>0.78</td>
</tr>
<tr>
<td>9</td>
<td>184.0</td>
<td>0.23</td>
<td>0.16</td>
<td>21</td>
<td>154.9</td>
<td>15.87</td>
<td>1.70</td>
</tr>
<tr>
<td>10</td>
<td>556.8</td>
<td>0.14</td>
<td>0.12</td>
<td>22</td>
<td>2301.0</td>
<td>10.08</td>
<td>0.69</td>
</tr>
<tr>
<td>11</td>
<td>618.9</td>
<td>0.07</td>
<td>0.14</td>
<td>23</td>
<td>931.3</td>
<td>883.61</td>
<td>0.72</td>
</tr>
<tr>
<td>12</td>
<td>199.4</td>
<td>0.29</td>
<td>0.17</td>
<td>24</td>
<td>995.3</td>
<td>286.93</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>991.3</td>
<td>66.68</td>
<td>4.51</td>
</tr>
<tr>
<td>Avg.</td>
<td>0.17</td>
<td>0.13</td>
<td></td>
<td></td>
<td>98.37</td>
<td>98.37</td>
<td>0.96</td>
</tr>
</tbody>
</table>

### Table 2

Performance of the models on large instances.

<table>
<thead>
<tr>
<th>#</th>
<th>BSTP-EG (10 800 seconds)</th>
<th>VRP-DA (max 3 600 seconds)</th>
<th>VRP-DA (max 10 800 seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cost</td>
<td>Gap</td>
<td>Cost</td>
</tr>
<tr>
<td>26</td>
<td>1193.2</td>
<td>26.8</td>
<td>1193.0</td>
</tr>
<tr>
<td>27</td>
<td>1832.8</td>
<td>1.2</td>
<td>1832.8</td>
</tr>
<tr>
<td>28</td>
<td>2108.9</td>
<td>56.2</td>
<td>1932.9</td>
</tr>
<tr>
<td>29</td>
<td>497.0</td>
<td>8.3</td>
<td>495.5</td>
</tr>
<tr>
<td>30</td>
<td>523.7</td>
<td>10.1</td>
<td>484.8</td>
</tr>
<tr>
<td>31</td>
<td>636.8</td>
<td>13.7</td>
<td>578.8</td>
</tr>
<tr>
<td>32</td>
<td>1700.7</td>
<td>23.3</td>
<td>1602.5</td>
</tr>
<tr>
<td>33</td>
<td>586.7</td>
<td>6.9</td>
<td>586.9+</td>
</tr>
<tr>
<td>34</td>
<td>1787.0</td>
<td>29.7</td>
<td>1660.7</td>
</tr>
<tr>
<td>35</td>
<td>1883.1</td>
<td>28.5</td>
<td>1725.8</td>
</tr>
<tr>
<td>36</td>
<td>360.8</td>
<td>25.6</td>
<td>1867.9</td>
</tr>
<tr>
<td>37</td>
<td>2022.3</td>
<td>25.0</td>
<td>1929.0</td>
</tr>
<tr>
<td>38</td>
<td>445.1</td>
<td>16.3</td>
<td>*</td>
</tr>
<tr>
<td>Avg.</td>
<td>1388.4</td>
<td>21.5 percent</td>
<td>1322.0</td>
</tr>
</tbody>
</table>
respect to the best solution obtained by Cplex (solution to instances 1 to 27 are proven optimal). For small and medium size instances 1 to 25, the proposed heuristic with 100 iterations produced all of the optimal solutions. For the larger instances, we improved the average gap produced by the Anaya Arenas et al. (2016) heuristic, 8.25 percent, to 4.67 percent when only 1 iteration of the heuristic is run, and to 0.91 percent after 100 iterations. Even more, the multi-start heuristic produced solutions that improved the best known ones (produced by Cplex) in 5 times.

### 4.2. Results for the VRP with desynchronized arrivals

In this section we analyze the ability of the VRP-DA formulation and the multi-start heuristic to minimize the total number of boxes arriving to the Lab during any time period. Our computational experiments were still based on the Anaya Arenas et al. (2016) set of instances, as they are the practical foundation of this work. However, for each instance, we set the opening and closing times of the laboratory and then divided the working hours into time periods of $\omega$ minutes. Clearly the length of these periods not only influences the size of the model, but also the maximal number of boxes arriving within a period. Managing the time periods adds a new complexity level to the problem and impacts the size of the instances that can be solved to optimality. This is why the instances in this section are divided into two sets, with instances 1 to 19 in the first one and instances 20 to 38 in the second one. The VRP-DA model can only be solved to optimality by Cplex for the latter set.

Let us first analyze the behavior of model VRP-DA when the synchronization factor is not considered. To this end, we set the penalty factor $\theta = 0$ in the objective function (1). The model optimizes only the sum of the route’s durations and we observe for each instance the number of boxes arriving to the Lab at the busiest period.

Table 4 reports the numerical results produced for these experiments when $\omega$ was set to 60 and 30 minutes. Columns $RD$ provide the optimal duration of the routes. The routes’ duration produced by the heuristic are reported in columns $RD$. Column $w$ reports the number of boxes arriving during the busiest period, giving an idea of the values that one should obtain if desynchronization is not considered in the optimization process. Finally, columns $Second$ report the Cplex computing time in seconds. We do not report the heuristic computing time as it is always below a second.

As expected, the routes’ durations are not impacted by the length of the considered period. However, the average computing time of Cplex rises from 269 seconds to 331 seconds when the time period is reduced from 60 to 30 minutes. We can observe the excellent performance of the heuristic (with 100 iterations), which was able to find the optimal route’s duration for all instances. It is also worth to mention that both methods led to the
Table 4
Results for the VRP-DA with no penalty on the number of arrivals ($\theta = 0$).

<table>
<thead>
<tr>
<th>#</th>
<th>Cplex RD</th>
<th>Cplex w</th>
<th>Heuristic RD</th>
<th>Heuristic w</th>
<th>Cplex Second</th>
<th>Heuristic Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>270.0</td>
<td>1</td>
<td>0.2</td>
<td>270.0</td>
<td>1</td>
<td>270.0</td>
</tr>
<tr>
<td>2</td>
<td>179.0</td>
<td>2</td>
<td>0.1</td>
<td>179.0</td>
<td>2</td>
<td>179.0</td>
</tr>
<tr>
<td>3</td>
<td>378.0</td>
<td>2</td>
<td>0.4</td>
<td>378.0</td>
<td>2</td>
<td>378.0</td>
</tr>
<tr>
<td>5</td>
<td>444.0</td>
<td>2</td>
<td>2.1</td>
<td>444.0</td>
<td>3</td>
<td>444.0</td>
</tr>
<tr>
<td>6</td>
<td>444.0</td>
<td>2</td>
<td>5.6</td>
<td>444.0</td>
<td>2</td>
<td>444.0</td>
</tr>
<tr>
<td>7</td>
<td>426.0</td>
<td>2</td>
<td>3.9</td>
<td>426.0</td>
<td>2</td>
<td>426.0</td>
</tr>
<tr>
<td>8</td>
<td>431.0</td>
<td>2</td>
<td>5.0</td>
<td>431.0</td>
<td>2</td>
<td>431.0</td>
</tr>
<tr>
<td>9</td>
<td>423.0</td>
<td>3</td>
<td>9.1</td>
<td>423.0</td>
<td>3</td>
<td>423.0</td>
</tr>
<tr>
<td>12</td>
<td>311.6</td>
<td>3</td>
<td>4.8</td>
<td>311.6</td>
<td>3</td>
<td>311.6</td>
</tr>
<tr>
<td>13</td>
<td>853.0</td>
<td>4</td>
<td>77.2</td>
<td>853.0</td>
<td>4</td>
<td>853.0</td>
</tr>
<tr>
<td>14</td>
<td>347.6</td>
<td>7</td>
<td>285.0</td>
<td>347.6</td>
<td>7</td>
<td>347.6</td>
</tr>
<tr>
<td>15</td>
<td>515.0</td>
<td>3</td>
<td>9.5</td>
<td>515.0</td>
<td>3</td>
<td>515.0</td>
</tr>
<tr>
<td>16</td>
<td>306.6</td>
<td>4</td>
<td>3600.0</td>
<td>306.6</td>
<td>4</td>
<td>306.6</td>
</tr>
<tr>
<td>17</td>
<td>330.6</td>
<td>4</td>
<td>83.0</td>
<td>330.6</td>
<td>4</td>
<td>330.6</td>
</tr>
<tr>
<td>18</td>
<td>337.6</td>
<td>4</td>
<td>11.2</td>
<td>337.6</td>
<td>4</td>
<td>337.6</td>
</tr>
<tr>
<td>19</td>
<td>599.0</td>
<td>4</td>
<td>207.9</td>
<td>599.0</td>
<td>4</td>
<td>599.0</td>
</tr>
<tr>
<td>Avg.</td>
<td>412.2</td>
<td>3.3</td>
<td>269.1</td>
<td>412.2</td>
<td>3.25</td>
<td>412.2</td>
</tr>
</tbody>
</table>

Table 5
Results for the VRP-DA with penalty on the number of box arrivals ($\theta = 100$).

<table>
<thead>
<tr>
<th>#</th>
<th>Cplex RD</th>
<th>Cplex w</th>
<th>Heuristic RD</th>
<th>Heuristic w</th>
<th>Cplex Second</th>
<th>Heuristic Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>270.0</td>
<td>1</td>
<td>0.1</td>
<td>270.0</td>
<td>1</td>
<td>270.0</td>
</tr>
<tr>
<td>2</td>
<td>179.0</td>
<td>2</td>
<td>0.1</td>
<td>179.0</td>
<td>2</td>
<td>179.0</td>
</tr>
<tr>
<td>3</td>
<td>378.0</td>
<td>2</td>
<td>0.3</td>
<td>378.0</td>
<td>2</td>
<td>378.0</td>
</tr>
<tr>
<td>5</td>
<td>444.0</td>
<td>1</td>
<td>2.9</td>
<td>444.0</td>
<td>3</td>
<td>444.0</td>
</tr>
<tr>
<td>6</td>
<td>444.0</td>
<td>1</td>
<td>6.7</td>
<td>444.0</td>
<td>2</td>
<td>444.0</td>
</tr>
<tr>
<td>7</td>
<td>426.0</td>
<td>1</td>
<td>5.9</td>
<td>426.0</td>
<td>2</td>
<td>426.0</td>
</tr>
<tr>
<td>8</td>
<td>431.0</td>
<td>1</td>
<td>7.2</td>
<td>431.0</td>
<td>2</td>
<td>431.0</td>
</tr>
<tr>
<td>9</td>
<td>431.0</td>
<td>1</td>
<td>12.2</td>
<td>431.0</td>
<td>2</td>
<td>431.0</td>
</tr>
<tr>
<td>12</td>
<td>311.6</td>
<td>2</td>
<td>15.3</td>
<td>311.6</td>
<td>2</td>
<td>311.6</td>
</tr>
<tr>
<td>13</td>
<td>853.0</td>
<td>2</td>
<td>202.3</td>
<td>853.0</td>
<td>3</td>
<td>853.0</td>
</tr>
<tr>
<td>14</td>
<td>347.6</td>
<td>1</td>
<td>159.8</td>
<td>347.6</td>
<td>2</td>
<td>347.6</td>
</tr>
<tr>
<td>15</td>
<td>515.0</td>
<td>2</td>
<td>19.4</td>
<td>515.0</td>
<td>2</td>
<td>515.0</td>
</tr>
<tr>
<td>16</td>
<td>306.6</td>
<td>1</td>
<td>285.0</td>
<td>306.6</td>
<td>3</td>
<td>306.6</td>
</tr>
<tr>
<td>17</td>
<td>330.6</td>
<td>1</td>
<td>306.0</td>
<td>330.6</td>
<td>2</td>
<td>330.6</td>
</tr>
<tr>
<td>18</td>
<td>337.6</td>
<td>2</td>
<td>16.6</td>
<td>337.6</td>
<td>2</td>
<td>337.6</td>
</tr>
<tr>
<td>19</td>
<td>599.0</td>
<td>2</td>
<td>285.0</td>
<td>599.0</td>
<td>3</td>
<td>599.0</td>
</tr>
<tr>
<td>Avg.</td>
<td>412.2</td>
<td>1.3</td>
<td>439.6</td>
<td>412.2</td>
<td>1.3</td>
<td>439.6</td>
</tr>
</tbody>
</table>

The $w$ value in almost all but four cases. This can be explained by the presence of an equivalent solution with respect to the route duration.

Table 5 reports the results when we set the penalty factor to $\theta = 100$, which means that the model reaches a compromise between minimizing the route's duration and the number of arrivals during the busiest period. In column $RD$, we report the lower bound on the route duration as per Table 4, which is valid for both values of $\omega$. In order to obtain a lower bound on the number of box arrivals in the busiest period, we ran the model with a very high penalty factor ($\theta = 100,000$). These results are reported in columns $\bar{w}$.

When we set $\omega = 60$ minutes and $\theta = 100$, Cplex produced optimal solutions having in average a route duration of 439.6 instead of 412.2 units (lower bound in Table 4). However, the average number of boxes arriving during the busiest period is reduced from 3.3 to 1.5 boxes. The heuristic produced average values of 435.6, and 2.3 for the same indicators. When we set $\omega = 30$ minutes, the average number of boxes arriving during the busiest period produced by Cplex decreases to 1.3, whereas the lower bound ($\bar{w}$) is equal to 1. For this case, and comparing to the results in Table 4, we can say that a better desynchronization has allowed reducing the maximal number of boxes received within the busiest period from 2.8 to 1.3 at the cost of an additional 8.2 percent in the total route duration reduction.

Table 6 reports the results produced for the larger instances. Since Cplex was unable to produce a feasible solution after 3 600 seconds of computing time, only heuristic results are reported. Experiments were run for penalty factor $\theta = 100$ and $\omega = 60$ minutes. To evaluate the heuristic's robustness and how the number of iterations influences the quality of the solutions, we report the results produced right after the construction phase of the heuristic and after 10 and 100 iterations. Computing times are only reported for 500 iterations, as they are otherwise negligible. Finally, since we do not have bounds on the route duration and the number of arrivals, we ran 10 000 iterations of the heuristic to try and get “good” bounds. To this end, first we set $\theta = 0$ to minimize the route's duration (column $RD$), and then we set $\theta = 100,000$ to minimize the number of arrivals during the busiest period (column $w$).
Table 6 shows that the lowest route duration was achieved with 10 iterations with a deviation of 4.1 percent from the minimum (obtained with $\theta = 0$). With 500 iterations, the heuristics improve the maximum number of arrivals to 4.3, while the minimum (obtained with $\theta = 100,000$) was 4.2, however the routes duration was increased slightly. Thus, running the heuristic with 500 iterations seems a good compromise while the computing time remains below five seconds.

5. Conclusions

This article deals with an important practical transportation problem encountered in the health system of the Quebec province. This problem requires the transport of biomedical samples from collection points to a laboratory where they are analyzed. As the laboratory is the bottleneck of the system, it is important to avoid congestion by balancing the arrival of samples. To this end, we discretize the working hours into periods and try to minimize the number of samples boxes arriving to the laboratory in the busiest period. We modeled the problem as a vehicle routing problem with desynchronized arrivals (VRP-DA). To the best of our knowledge, it is the first time that this problem has been addressed and modeled. We formulated it as a MIP and developed a heuristic to solve it. The formulation and the heuristic were adapted to deal with a similar routing problem in which desynchronization was not considered, and both outperformed existing methods. Our computational results demonstrated also that the problem with desynchronized arrivals is much more difficult to solve to optimality. For medium size real instances, the formulation was solved to optimality within few seconds. However, for larger instances, the Cplex was unable to obtain any feasible solution, although the heuristic proved to be efficient in minimizing both the route’s duration and the number of arrivals during the busiest period.

The algorithms developed in this research have been applied to four administrative regions in the Quebec province. On an annual basis, more than 2.1 million kilometers are involved in the new routes configurations (Renaud, Ruiz, Chabot, Anaya Arenas, & Zue Ntoutoume, 2014). We are currently working with the Ministry to reorganize transportation operations of the other 13 administrative regions as part of the global laboratories optimization project.

From a broader perspective, we believe that this practical situation, where the route planner might avoid congestion at depots or other facilities, justifies additional research.

Acknowledgments

This research was partially supported by grants OPG 0293307 and OPG 0172633 from the Canadian Natural Sciences and Engineering Research Council (NSERC), by grant 37348 from the Research Deputy of Ferdowsi University of Mashhad and by the Ministry of Health and Social Services of Quebec under contract GI04976. This support is gratefully acknowledged. We express our gratitude to Dominic Morand, conseiller à la direction de la logistique et des équipements, of the MSSS for his collaboration on this project and for providing us with real data. The authors are grateful to the three anonymous referees for their useful and constructive comments and suggestions.

References


