Ambulance routing for inter-hospital patient transfers in Sri Lanka

Sudheeraka Wickramarachchi, Kazuki Hasegawa, and Wei Wu

Graduate School of Integrated Science and Technology, Shizuoka University, Japan
sudheeraka.22@shizuoka.ac.jp

Keywords: Inter-hospital ambulance routing problem, Minimum risk time, Machine learning, Vehicle routing problem with time windows.

1 Introduction

The ambulance routing problem (ARP) involves finding optimal routes for ambulances to reach emergency sites, taking into account variables like traffic congestion, geographical limitations, and urgency of patient needs [2]. Essentially, the effectiveness of emergency medical transport system (EMTS) relies on the efficiency in defining and solving the ARP. Most existing studies investigated the scenario of reaching an emergency site and transporting a patient to an appropriate hospital. However, to the best of our knowledge, there are no studies that focus on designing and analyzing inter-hospital ARPs. In this study, we define an unconventional ARP in the optimization of inter-hospital transport, taking into account the particularities of the EMTS in Sri Lanka. This ARP incorporates a range of characteristics such as assignment, scheduling, and routing. To obtain good solutions for this new ARP, we propose a mathematical programming model and three two-phase heuristic approaches. Two of the heuristic approaches use machine learning techniques in the first phase.

2 Inter-hospital ambulance routing problem in Sri Lanka

The public health service of Sri Lanka is one of the major services provided free of charge to the public. Based on to the facilities that can be provided, the hospitals in Sri Lanka are categorized into 6 levels: National Hospitals (Level 1), Teaching Hospitals (Level 2), Provincial General Hospitals (Level 3), Base Hospitals (Level 4), Divisional Hospitals (Level 5), and Primary Medical Care Units (Level 6). In ascending order of levels (Level 1 to 6), space facilities, personnel with expertise availability, intensive care and surgery facilities are limited. According to the recommendations of medical experts, some patients should be transferred to other hospitals that are better facilitated, depending on the current hospital and their conditions.

In this study, we consider Colombo district, which has the highest population density of Sri Lanka. Transferring a patient may be an ‘immediately incurred transfer’ (IIT) or a ‘scheduled transfer’ (ST). An IIT can be clarified as a random occurrence such as an accident patient. In each type of transferring scenario, medical experts recommend not to exceed the minimum risk time (MRT). In this case study, we focus on the ST case, and identify the MRT as a highly significant factor, which can be expressed in terms of a due time for each patient. We follow the ST case to conduct our analysis, and try to obtain a good mathematical solution for the routing system.

3 Problem formulation

In this section, we first describe the ARP arising in Sri Lanka, and then propose a mixed integer programming (MIP) model which can well address the objective of minimizing the maximum tardiness that can occur in a patient transfer system.

In our ARP, we are given a set of patients $V_s = \{1, 2, \ldots, m\}$, a set of available beds $V_d = \{m+1, m+2, \ldots, m+n\}$, and a set of ambulances $K = \{1, 2, \ldots, k_{\max}\}$. The ambulance depot is denoted by 0. Each bed (demand) is associated with its level l_j (i.e., the bed is in a level-l_j hospital). Each patient (supply) i is required to be transferred to a bed (in a hospital) with a level of l_i or less, before the MRT d_i of patient i. A penalty weight w_i occurs, if the transfer request of patient i is ignored (not scheduled). The maximum limit of patient ignorance for the system is denoted by W. The value c_{ij} indicates the traveling time when transferring patient i to bed j. In this study, we consider all ambulances are homogeneous in traveling time. The inter-hospital ambulance routing problem (IH-ARP) aims to generate an ambulance routing solution so as to minimize the maximum tardiness among transferred patients.

From the problem input, we consider to generate a directed graph $G = (V, E)$ with $V = \{0\} \cup V_s \cup V_d$ and $E = E_{sd} \cup E_{ds} \cup \{(0, j) \mid j \in V_s\} \cup \{(i, 0) \mid i \in V_d\} \cup \{(0, 0)\}$, where $E_{sd} = \{(i, j) \mid i \in V_s, j \in V_d, l_j \le l_i\}$ and $E_{ds} = \{(i, j) \mid i \in V_d, j \in V_s\}$. We use c_{sum} to denote the value $c_{\text{sum}} = \sum_{(i,j)\in E} c_{ij}$. Next, we introduce the variables used in the proposed model. The variable x_{ijk} indicates whether the edge (i, j) is visited by the ambulance k. We use p_i to denote the departure time of patient i, and p_j to denote the arrival time at bed j, and $T_{\max}$ to denote the maximum tardiness. By using these variables, the IH-ARP can be modeled as,

$$\min \quad T_{\max} \tag{1}$$

$$\text{s.t.} \quad \sum_{i\in V} x_{jik} = \sum_{i\in V} x_{ijk} \qquad \forall j \in V, \forall k \in K \tag{2}$$

$$\sum_{k\in K} \sum_{j\in V_d} x_{ijk} \le 1 \qquad \forall i \in V_s \tag{3}$$

$$\sum_{k\in K} \sum_{i\in V_s} x_{ijk} \le 1 \qquad \forall j \in V_d \tag{4}$$

$$\sum_{i\in V_d\cup\{0\}} x_{0ik} = 1 \qquad \forall k \in K \tag{5}$$

$$p_0 = 0 \tag{6}$$

$$\sum_{i\in V_s} w_i \left(1 - \sum_{k\in K} \sum_{j\in V_d} x_{ijk}\right) \le W \tag{7}$$

$$p_j \ge p_i + c_{ij} - (1 - x_{ijk}) c_{\text{sum}} \qquad \forall (i, j) \in E_{ds} \cup E_{sd}, \forall k \in K \tag{8}$$

$$T_{\max} \geq \sum_{j \in D} \sum_{k \in K} c_{ij} x_{ijk} - d_i - c_{\text{sum}} \left(1 - \sum_{k \in K} \sum_{j \in V_{\text{d}}} x_{ijk}\right) \qquad \forall i \in V_{\text{s}} \tag{9}$$

$$p_i \geq 0 \qquad \forall i \in V \tag{10}$$

$$T_{\max} \geq 0 \tag{11}$$

$$x_{ijk} \in \{0, 1\} \qquad \forall (i, j) \in E, \forall k \in K. \tag{12}$$

Constraints (2) confirm the flow conservation. Constraints (3) ensure that each patient can be transferred to at most one bed. Constraints (4) show similar relations for each bed. Constraint (5) states that each ambulance departs from the depot. The starting time is defined by constraints (6). The maximum limit of patient ignorance is denoted by the constraints (7). Constraints (8) to (9) define linkage between arrival time of each visited node and status of the MRT value.

4 Heuristic approaches

We design three heuristic approaches, all of which consist of two phases.

Approach based on local search (αLS): In the first phase, we assign patients to convenient beds depending on an assignment cost $c'_{ij} = c_{ij} + \alpha d_i$ that combines the traveling time c_{ij} and MRT d_i of each patient i, where α is an algorithm parameter determined by trial and error. In the resulting assignment problem, we may ignore some patients because of the limited number of beds. This supply-demand unbalance can be resolved by introducing dummy nodes.

We consider the output (patient-bed assignment) of the first phase as task nodes to construct a graph for the second phase. Each task node v_{ij} represents a transfer of patient i to bed j. We consider a directed complete graph consisting of all task nodes and a special node v_{00} representing the depot. The problem in the second phase can be identified as a vehicle routing problem with time windows (VRP-TW) depending on the MRTs of patients.

For solving this VRP-TW, we first construct an initial solution by utilizing a nearest neighbor method with time window restrictions. Thereafter, to obtain an improved solution, we propose a local search (LS) algorithm using generalized OR-opt operations extended from the classical OR-opt operations [1].

Approach based on ML incorporated with MIP (ML-MIP): In the second approach, we use ML methods for patient selection. After the first phase, we can determine which patients must to be transferred (without being ignored) in the system. Then, the route construction is performed using a simplified version of the MIP with a reduced feasible region of the original model (1)–(11).

Approach combining machine learning and local search (ML-LS): This approach improve the first phase of the αLS approach. In the first phase, we utilize ML to obtain multi-class (OneVsRest) predictions, that is, probability β_{ih} of assigning patient

i to hospital h (including a dummy hospital to denote the ignorance). Then, we set $c''_{ij} = 1 - \beta_{ih}$ if bed j is in hospital h, and use this modified cost c''_{ij} for the assignment problem in the first phase.

5 Computational results

We performed computational experiments on 5 instances from real-world data for the proposed model (MIP) in Section 3, and three approaches (αLS, ML-MIP, ML-LS) in Section 4. When machine learning is involved, we employed logistic regression (LR), gradient boosting (GB), random forest (RF), artificial neural network (ANN). Preliminary experiments showed that GB and LR outperformed the others in ML-MIP and ML-LS approaches, respectively. Thus, we report the results with their best machine learning models. The time limit for each instance was set to 300 seconds. The parameter α used in αLS is set to $\alpha = 5$.

Table 1 shows the running times in seconds (time) and objective function values (obj), where notation '—' indicates that no feasible solution was obtained in the time limit. From Table 1, we observe that MIP and ML-MIP showed similar performance in terms of both running time and objective function value. The αLS and ML-LS approaches obtained good solutions for large-scale instances within a small time compared to MIP and ML-MIP. Compared to ML-LS, for some small-scale instances, αLS failed to obtain an optimal solution within the time limit due to the poor assignment obtained in the first phase.

Table 1: Comparison of four approaches.

instance			MIP		ML-MIP		αLS		ML-LS	
m	n	$k_{\max}$	time	obj	time	obj	time	obj	time	obj
25	17	4	26.1	0	25.4	0	0.3	2	0.7	0
25	22	5	300.0	33	300.0	43	0.4	0	0.5	0
45	42	9	37.6	0	13.1	0	0.9	9	0.8	0
80	77	16	—	—	—	—	2.0	0	1.9	5
96	94	22	—	—	—	—	2.7	0	2.8	0

References

1. G Babin, S.D., Laporte, G.: Improvements to the or-opt heuristic for the symmetric travelling salesman problem. Journal of the Operational Research Society **58**(3), 402–407 (2007)
2. Tlili, T., Harzi, M., Krichen, S.: Swarm-based approach for solving the ambulance routing problem. Procedia Computer Science **112**, 350–357 (2017)