Algorithms for a Special Class of State-Dependent Shortest Path Problems with an Application to the Train Routing Problem

Lunce Fu and Maged Dessouky
Daniel J. Epstein Department of Industrial & Systems Engineering
University of Southern California
Los Angeles, California 90089
maged@usc.edu
(213) 740-4891

Abstract

We study the state-dependent shortest path problem and focus on its application to the Single Train Routing Problem consisting of a rail network with only double track segments, where the objective is to route one train through an empty network as fast as possible. We show that the Single Train Routing Problem is NP-hard. We investigate the solution properties and present sufficient conditions for optimality. Different conditions on the parameters are given to guarantee that certain local route selection is optimal. Then a dynamic programming heuristic is introduced and conditions when the proposed heuristic can obtain the optimal solution in polynomial time are also discussed. Experimental results show the efficiency of the proposed heuristics for general problem settings.

Keywords

Train routing, State-dependent shortest path problem, NP-hard, heuristics

1. Introduction

We define a state-dependent shortest path problem as the problem of finding the path with the minimal travel time in a graph where the transit time over each edge is dependent on its end nodes' states. Formally, in a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where $\mathcal{N}$ is the node set and $\mathcal{E}$ is the edge set, the transit time over edge $(i, j) \in \mathcal{E}$ is $t(\mathfrak{s}_i, \mathfrak{s}_j, \ell_{ij})$, where $\mathfrak{s}_i$ and $\mathfrak{s}_j$ are states for node i and j, and ℓ_{ij} is the length of edge (i, j). Given two nodes $s, e \in \mathcal{N}$ together with their states $\mathfrak{s}_s, \mathfrak{s}_e$, the objective is to find a path connecting s and e with the minimal travel time. If the arrival time is treated as the state, then this problem is related to the well-known time-dependent shortest path problem (see e.g., the papers by Cooke and Halsey, 1966; Dreyfus, 1969; Orda and Rom, 1990; Orda and Rom, 1991; Kaufman and Smith, 1993; Philpott, 1994; Cai et al, 1997; Chabini, 1998; Pallottino and Scutella, 1998; Ahuja et al, 2003; Fischer and Helmberg, 2014, Koch and Nasrabadi, 2014). In most models for the time-dependent shortest path problem, the transit time over an arc is positive. Therefore the "states" along the nodes in a path is strictly increasing for the time-dependent shortest path problem.

The main difference between the state-dependent shortest path problem and the classical shortest problem is that the transit time is not a constant number but a function. Therefore the classical method like Dijkstra's algorithm cannot be directly applied. However, if we can approximate the transit time $t(\mathfrak{s}_i, \mathfrak{s}_j, \ell_{ij})$ with a constant number, then the solution to the classical shortest path problem can work as a heuristic solution for the state-dependent shortest path problem. See section 5.2 for the comparison between this kind of heuristic method and our proposed method.

One possible application for the state-dependent shortest path problem is the train routing (pathing) problem where the "states" are the velocities at nodes. In contrast to the time-dependent shortest path problem, the "states" in the train routing problem, i.e. velocities may either decrease or increase along the nodes in a path.

In urban centers such as the Southern California area, the rail network is rather complex consisting of multiple track segments with various speed limits with a number of different alternative routes for a given origin-destination pair. This complexity makes it difficult to determine the fastest route for a given origin-destination pair among all the alternative routes. In this paper, we focus on the single line double track railway system, which is very common in many portions of the Southern California network. In a single line double track railway system, train stations and junctions along a single line are connected by a double track and trains can switch to the other track only at each train station or junction. We call the railway track between two consecutive junction points a segment. Even though its configuration is simple, the nature of the routing problem persists and many conclusions and results for the double track system can also be applied to other railway systems, such as single track with sidings, triple track rail systems or even a more general rail network.

The routing decision is which track of the double track segment (i.e., upper or lower) should the train travel on for each segment between adjacent junctions. This is not a trivial decision since even for this rail configuration the upper and lower tracks may differ in length and speed limits. With finite acceleration and deceleration rates, the train routing problem can be considered as a shortest path problem where the travel time over an arc is dependent on the entering velocity and the exiting velocity.

This paper extends the results of the train routing (pathing) problem studied by Nagarajan and Ranade (2008), where a new train is introduced to a network with some pre-scheduled trains. The objective is to find a route for the new train without affecting the existing schedules with the acceleration and deceleration processes taken into consideration. They showed that the routing problem is not only NP-complete, but also NP-hard to approximate this problem to any factor that is at most exponential to the input size. The definition of one node's profile is introduced as the set of all possible tuples of arrival time and velocity. An algorithm is also provided based on the expansion of the profile over an arc. We generalize the conclusions of Nagarajan and Ranade (2008) by considering the problem of routing a single train through an empty network and provide a more practical approach to solve the train routing problem. More specifically, the contributions of this paper are:

- Show the computational complexity of the problem of routing one train through an empty network when its acceleration and deceleration rates are finite constants.
- Provide different conditions on the problem parameters for the single line double track railway system to guarantee the optimality of local track selection. Experimentally show that these local optimality conditions can be used as a basis for a heuristic for real time train routing.
- Propose a dynamic programming heuristic to solve the problem of routing one train

through an empty network and also provide conditions such that the proposed heuristic can give the optimal solution in polynomial time.

The rest of the paper is organized as follows. In Section 2, we investigate the computational complexity of the single train routing problem through an empty network. Then the local optimality conditions for track selection are provided in Section 3. Section 4 presents a dynamic programming based heuristic method for the problem under general parameter settings. Section 5 experimentally compares the proposed heuristics against the optimal solution. Finally, we conclude the paper in Section 6.

2. Single Train Routing Problem

In this section, we give a formal statement of the problem of routing a single train through an empty rail network with the objective of minimizing the total travel time. Given a constant acceleration rate r_a and deceleration rate r_d, an origin node s and a destination node d in the network, and a starting velocity v_s and an ending velocity v_d, we want to find a path from s to d with the least travel time such that the train's starting velocity at node s is v_s and the ending velocity at node d is v_d. For simplicity we do not consider the train's length, (i.e. treat each train as a point). We refer to this problem as the "Single Train Routing Problem".

Since there are no other trains in the railway network, the Single Train Routing Problem is a special case of the problem considered in Nagarajan and Ranade (2008) where there exist some other trains in the network. We first show that even under this special case the problem is still NP-hard. The decision problem for the Single Train Routing Problem can be stated as follows. "Given the constant acceleration rate r_a and deceleration rate r_d, the origin node s and the destination node d in the network, and the starting velocity v_s and the ending velocity v_d, does there exist a path from s to d with the travel time less than a given value $\tilde{t}$ such that the train's starting velocity at node s is v_s and the ending velocity at node d is v_d?"

We first present a linear programming formulation to show that the decision problem above is in NP, i.e. it only takes polynomial time to check whether a given route (path) is actually a solution to the decision problem. And then show that the Single Train Routing Problem is NP-hard. Therefore the decision problem for the Single Train Routing Problem is NP-complete since it is in NP and is NP-hard.

Given a path, if we can get its travel time in polynomial time then we can tell whether this given path is a solution to the decision problem. Therefore in order to show that the decision problem is in NP, we need to show that it only takes polynomial time to compute the travel time for a given path. If the acceleration rate and deceleration rate are infinity, we can obtain the travel time in linear time by simply summing up the travel time for the different segments. With finite acceleration rate and deceleration rate, we need to compute the velocities at the beginning and the end of each segment to determine the travel time. Lu et al. (2004) provided a polynomial time algorithm to compute the travel time. Here we provide another way using a linear programming formulation. Even though for one single path the computational complexity of the linear programming formulation may be worse than the method given by Lu et al. (2004), the linear programming formulation provides a systematic perspective for the problem such that it is more robust to the parameters and provides more information when we want to solve a new problem.

Assume the given path consists of *k* segments where the length is l_i and the speed limit is $\bar{v}_i$ for the i^{th} segment. The acceleration rate and deceleration rate for the train are r_a and r_d respectively. The starting velocity is v_s and the ending velocity is v_e.

Then a direct and simple formulation for the problem is

$$\min \sum_{i=1}^{k} t(v_{i-1}, v_i, l_i, \bar{v}_i) \tag{1}$$

$$v_0 = v_s \tag{2}$$

$$v_k = v_e \tag{3}$$

$$v_i \leq \min(\bar{v}_i, \bar{v}_{i+1}) \qquad \forall i = 1, \dots, k-1 \tag{4}$$

where $t(v_{i-1}, v_i, l_i, \bar{v}_i)$ is the travel time over an arc and v_i is the velocity at the end of the i^{th} segment. We use the same definition of Lu et al (2004) of $t(v_{enter}, v_{exit}, l, \bar{v})$ in this paper.

1. If $v_{enter} \leq v_{exit} \leq \bar{v}, v_{exit}^2 - v_{enter}^2 = 2r_a l$

$$t(v_{enter}, v_{exit}, l, \bar{v}) = \frac{v_{exit} - v_{enter}}{r_a}$$

2. If $v_{exit} \leq v_{enter} \leq \bar{v}, v_{enter}^2 - v_{exit}^2 = 2r_d l$

$$t(v_{enter}, v_{exit}, l, \bar{v}) = \frac{v_{enter} - v_{exit}}{r_d}$$

3. If $v_{exit} \leq \bar{v},\ v_{enter} \leq \bar{v},\ v_{exit}^2 - v_{enter}^2 \leq 2r_a l,\ v_{enter}^2 - v_{exit}^2 \leq 2r_d l,$

$\sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}} \leq \bar{v}$

$$t(v_{enter}, v_{exit}, l, \bar{v}) = -\frac{v_{enter}}{r_a} - \frac{v_{exit}}{r_d} + (\frac{1}{r_a} + \frac{1}{r_d})\sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}}$$

4. If $v_{exit} \leq \bar{v},\ v_{enter} \leq \bar{v},\ v_{exit}^2 - v_{enter}^2 \leq 2r_a l,\ v_{enter}^2 - v_{exit}^2 \leq 2r_d l,$ $\sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}} > \bar{v}$

$$t(v_{enter}, v_{exit}, l, \bar{v}) = \frac{\bar{v} - v_{enter}}{r_a} + \frac{\bar{v} - v_{exit}}{r_d} + \frac{1}{\bar{v}}(l - \frac{\bar{v}^2 - v_{enter}^2}{2r_a} - \frac{\bar{v}^2 - v_{exit}^2}{2r_d})$$

5. Otherwise

$$t(v_{enter}, v_{exit}, l, \bar{v}) = +\infty$$

In the first scenario, the train keeps accelerating from v_{enter} to v_{exit} where condition $v_{exit}^2 - v_{enter}^2 = 2r_a l$ means that the distance is just enough for the train to accelerate from v_{enter} to v_{exit}. In the second scenario the train keeps decelerating from v_{enter} to v_{exit} and similarly condition $v_{enter}^2 - v_{exit}^2 = 2r_d l$ enforces that the distance is just enough for the train to decelerate from v_{enter} to v_{exit}. For the conditions in the third and fourth scenarios, $v_{exit} \leq \bar{v},\ v_{enter} \leq \bar{v},\ v_{exit}^2 - v_{enter}^2 \leq 2r_a l,\ v_{enter}^2 - v_{exit}^2 \leq 2r_d l$ are feasibility conditions, i.e. conditions to guarantee $t(v_{enter}, v_{exit}, l, \bar{v}) < +\infty$. Intuitively, these feasibility conditions ensure that the entering and exiting velocities are below the speed limit and that the travel distance is long enough for the entering velocity to accelerate/decelerate to the exiting velocity. $\sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}}$ is the maximum velocity without considering the speed limit. In the third scenario when the maximum velocity is less than the speed limit, the train first accelerates from v_{enter} to the maximum velocity, and then decelerates to v_{exit}. In the fourth scenario when the maximum velocity is greater than the speed limit, the train accelerates from v_{enter} to the speed limit, and then travels at a constant speed, which is the speed limit, and then decelerates from the speed limit to v_{exit}. Scenario 5 represents that the given v_{enter}, v_{exit}, l and $\bar{v}$ are infeasible, which implies the travel time should be infinity. Also we can get the following property immediately.

Property 1 $t(v_{enter}, v_{exit}, l, \bar{v}) < +\infty$ holds if and only if

$$v_{exit} \leq \bar{v},\ v_{enter} \leq \bar{v},\ v_{exit}^2 - v_{enter}^2 \leq 2r_a l,\ v_{enter}^2 - v_{exit}^2 \leq 2r_d l.$$

The drawbacks of this formulation are

- the constraints on the decision variables are implicit since $t(v_{i-1}, v_i, l_i, \bar{v}_i)$ is defined as positive infinity when the decision variables are infeasible.
- the objective function, even within the feasible region, i.e. when $t(v_{i-1}, v_i, l_i, \bar{v}_i) < +\infty$, is still nonlinear.

Our linear programming formulation is based on the intuition that the velocities v_i $(i = 1, \dots, k-1)$ are as large as possible in order to obtain the least travel time. In order to show that this intuition is correct, we first state a lemma and corollary (proofs provided in the Appendix).

Lemma 1 The function $t(v_{enter}, v_{exit}, l, \bar{v})$, when it is finite, is a strictly decreasing function of v_{enter} and v_{exit}, a strictly increasing function of l and a non-increasing function of $\bar{v}$.

Corollary 1 If $t(v_{enter}, v_{exit}, l, \bar{v}) < +\infty$, then we have

$$t(v_{enter}, v_{exit}, l, \bar{v}) \geq \max(\frac{v_{enter} - v_{exit}}{r_d}, \frac{v_{exit} - v_{enter}}{r_a})$$

We are now ready to show that our intuition is correct, i.e. the velocities v_i $(i = 1, \dots, k-1)$ are as large as possible in order to obtain the least travel time.

Proposition 1 If $u_i (i = 0, \dots, k)$ is the optimal solution for problem (1)-(4), then for any feasible solution for (1)-(4) $v_i (i = 0, \dots, k)$, we have $u_i \geq v_i \ \forall i = 0, \dots, k$.

Proof: We show this proposition by contradiction. Suppose $\exists\, j \in \{1, \dots, k-1\}$ such that $u_j < v_j$. We consider the solution formed by replacing u_j with v_j, i.e. $(u_0, \dots, u_{j-1}, v_j, u_{j+1} \dots, u_k)$.

If the constructed solution is feasible, i.e. $t(u_{j-1}, v_j, l_j, \bar{v}_j) < +\infty$ and $t(v_j, u_{j+1}, l_{j+1}, \bar{v}_{j+1}) < +\infty$. From Lemma 1 and $u_j < v_j$, we have

$$t(v_j, u_{j+1}, l_{j+1}, \bar{v}_{j+1}) < t(u_j, u_{j+1}, l_{j+1}, \bar{v}_{j+1})$$
$$t(u_{j-1}, v_j, l_{j+1}, \bar{v}_{j+1}) < t(u_{j-1}, u_j, l_{j+1}, \bar{v}_{j+1})$$

which implies that $(u_0, \dots, u_{j-1}, v_j, u_{j+1} \dots, u_k)$ is a strictly better solution than $(u_0, \dots, u_k)$. It is a contradiction to the fact that $u_i (i = 1, \dots, k)$ is the optimal solution. Therefore, $(u_0, \dots, u_{j-1}, v_j, u_{j+1} \dots, u_k)$ is an infeasible solution.

If $t(u_{j-1}, v_j, l_j, \bar{v}_j) = +\infty$, from Property 1 either $v_j^2 - u_{j-1}^2 > 2r_a l_j$ or $u_{j-1}^2 - v_j^2 > 2r_d l_j$ holds. Since $u_{j-1}^2 - v_j^2 \le u_{j-1}^2 - u_j^2 \le 2r_d l_j$ is valid, we have

$$v_j^2 - u_{j-1}^2 > 2r_a l_j \tag{5}$$

Since $(v_0, \dots, v_k)$ is a feasible solution, we have

$$v_j^2 - v_{j-1}^2 \le 2r_a l_j < v_j^2 - u_{j-1}^2$$

which implies that $u_{j-1} < v_{j-1}$. We can continue this argument from $u_{j-1} < v_{j-1}$ to $u_{j-2} < v_{j-2}$ until (5) is not satisfied i.e. $\exists j_0 \in \{1,2,\dots,j\}$ $s.t.$ $v_{j_0}^2 - u_{j_0-1}^2 \le 2r_a l_{j_0}$. (The existence of j_0 is guaranteed due to the fact that $v_1^2 - u_0^2 = v_1^2 - v_0^2 \le 2r_a l_1$).

Similarly if $t(v_j, u_{j+1}, l_{j+1}, \bar{v}_{j+1}) = +\infty$, from Property 1 either $u_{j+1}^2 - v_j^2 > 2r_a l_{j+1}$ or $v_j^2 - u_{j+1}^2 > 2r_d l_j$ holds. Since $u_{j+1}^2 - v_j^2 \le u_{j+1}^2 - u_j^2 \le 2r_a l_{j+1}$ is valid, we have

$$v_j^2 - u_{j+1}^2 > 2r_d l_{j+1} \tag{6}$$

Since $(v_0, \dots, v_k)$ is a feasible solution, we have

$$v_j^2 - v_{j+1}^2 \le 2r_d l_{j+1} < v_j^2 - u_{j+1}^2$$

which implies that $u_{j+1} < v_{j+1}$. We can continue this argument from $u_{j+1} < v_{j+1}$ to $u_{j+2} < v_{j+2}$ until (6) is not satisfied i.e. $\exists j_1 \in \{j, j+1, \dots, k-1\} \text{s.t.}\, v_{j_1}^2 - u_{j_1+1}^2 \le 2r_d l_{j_1+1}$. (The existence of j_1 is guaranteed due to the fact that $v_{k-1}^2 - u_k^2 = v_{k-1}^2 - v_k^2 \le 2r_d l_k$).

Thus the solution $(u_0, \dots, u_{j_0-1}, v_{j_0}, \dots, v_{j_1}, u_{j_1+1}, \dots, u_k)$ is feasible. Also we have $u_i < v_i \quad \forall i \in \{j_0, \dots, j_1\}$. From Lemma 1 we know $t(v_{enter}, v_{exit}, l, \bar{v})$ is a strictly decreasing function of v_{enter} and v_{exit}. Therefore the total travel time for the solution $(u_0, \dots, u_{j_0-1}, v_{j_0}, \dots, v_{j_1}, u_{j_1+1}, \dots, u_k)$ is strictly less than the travel time for $(u_0, \dots, u_k)$, which contradicts the fact that $(u_0, \dots, u_k)$ is the optimal solution for problem (1)-(4).□

Consider the formulation as follows:

$$\min \sum_{i=0}^{k} v_i^2 \tag{7}$$

$$v_0^2 = v_s^2 \tag{8}$$

$$v_k^2 = v_e^2 \tag{9}$$

$$v_i^2 \le \min(\bar{v}_i^2, \bar{v}_{i+1}^2) \qquad \forall\, i = 1, \dots, k-1 \tag{10}$$

$$v_i^2 - v_{i-1}^2 \le 2r_a l_i \qquad \forall i = 1, \dots, k \tag{11}$$

$$v_{i-1}^2 - v_i^2 \le 2r_d l_i \qquad \forall i = 1, \dots, k \tag{12}$$

Now we are ready to show (7)-(12) can give the same solution as (1)-(4).

Proposition 2: If the optimal solution for problem (7)-(12) is $v_i^*(i = 0, \dots, k)$, then $|v_i^*| = \sqrt{v_i^{*2}}$ $(i = 0, \dots, k)$ is also the optimal solution for problem (1)-(4).

Proof: Without loss of generality, we assume $v_i^* \geq 0$. First we show the feasibility of $v_i^*(i = 0, \dots, k)$ for problem (1)-(4). Constraints (2)-(4) are the same as constraints (8)-(10). From Property 1 and constraints (10)-(12) we have

$$t(v_{i-1}^*, v_i^*, l_i, \bar{v}_i) < +\infty \qquad \forall i \in \{1, \dots, k\}$$

Therefore $v_i^*(i = 0, \dots, k)$ is a feasible solution for problem (1)-(4).

We now show the optimality of $v_i^*(i = 0, \dots, k)$ for problem (1)-(4). Suppose $u_i(i = 0, \dots, k)$ is the optimal solution for problem (1)-(4). From Proposition 1 we know that

$$u_i \geq v_i^* \quad \forall i = 0, \dots, k$$

On the other hand, $u_i(i = 1, \dots, k)$ is also a feasible solution for problem (7)-(12). So

$$\sum_{i=0}^{k} v_i^{*2} \leq \sum_{i=0}^{k} u_i^2 \leq \sum_{i=0}^{k} v_i^{*2}$$

where the first inequality is due to the optimality of of $v_i^*(i = 0, \dots, k)$ for problem (7)-(12) and the second inequality is due to $u_i \geq v_i^* \quad \forall i = 0, \dots, k$. Therefore we have

$$u_i = v_i^* \quad \forall i = 0, \dots, k \square$$

If we take $x_i = v_i^2 \;\; (i = 0, \dots, k)$, then problem (7)-(12) becomes the following linear programming model.

$$\min \sum_{i=0}^{k} x_i \tag{13}$$

$$x_0 = v_s^2 \tag{14}$$

$$x_k = v_e^2 \tag{15}$$

$$x_i \leq \min(\bar{v}_i^2, \bar{v}_{i+1}^2) \qquad \forall\, i = 1, \dots, k-1 \tag{16}$$

$$x_i - x_{i-1} \leq 2 r_a l_i \qquad \forall i = 1, \dots, k \tag{17}$$

$$x_{i-1} - x_i \leq 2 r_d l_i \qquad \forall i = 1, \dots, k \tag{18}$$

$$x_i \geq 0 \qquad \forall i = 0, \dots, k \tag{19}$$

The above linear programming problem computes the optimal speed at each segment given a route (path). We next show that when the path is not given and there are no other trains in the network, the problem is NP-hard. In contrast, Nagarajan and Ranade (2008) showed a more constrained version of the Single Train Routing Problem (i.e., with some pre-scheduled trains in the network) is NP-hard.

Proposition 3 The Single Train Routing Problem is NP-hard.

Proof: We show that the Subset Sum problem, which is known to be NP-complete, is polynomial reducible to the Single Train Routing Problem. The Subset Sum Problem is defined as follows:

Given a set of positive integers $C = \{a_0,\ a_1, \dots, a_{m-1}\}$ and a positive target integer A, does there exists a subset $S \subset \{0, \dots, m-1\}$ such that $\sum_{i \in S} a_i = A$?

Given an instance of a Subset Sum Problem, we construct a corresponding Single Train Routing Problem instance as follows: consider the following network with m+1 nodes and 2m arcs. We index the nodes by $n_0, n_1, \dots, n_m$. $\forall\ i = 1, \dots, m$, there are two arcs between n_{i-1} and n_i. The upper arc has length of one and the lower arc has length of $1 + a_i$.

Figure 1 Railway Configuration

Also we set

$$r_a = r_d = 0.5$$

$$v_s = \sqrt{m + A}\,, v_e = 0$$

The speed limit for each arc is big enough so that the train can never reach its speed limit. For example we can set the speed limit for each arc to be $\sqrt{3m + A + \sum_{i=0}^{m-1} a_i}$, which the train can only reach if it travels through all the arcs in the network and keep accelerating starting from v_s.

First we show that if the Subset Sum Problem has a solution, i.e. $\exists\, S \subset \{0, \dots, m-1\}$ such that $\sum_{i \in S} a_i = A$, the optimal travel time for the corresponding Single Train Routing Problem is $2\sqrt{m+A}$. From node n_i to node $n_{i+1}(i = 0, \dots, m-1)$ we pick the upper arc, i.e. the one with length of one, if $i \notin S$ and pick the lower arc, i.e. the one with length of $1 + a_i$, if $i \in S$. Then the path consisting of all the picked arcs has a length of

$$m + \sum_{i \in S} a_i = m + A$$

which is just enough for the train to decelerate from v_s to v_e since

$$v_s^2 - v_e^2 = 2r_d(m + A)$$

Now we show this picked path is the optimal solution for the Single Train Routing Problem. For any other path whose length is less than $m + A$, it is not enough for the train to decelerate from v_s to v_e, which implies its infeasibility. Then we consider any other path whose length is greater than $m + A$. Since the speed limit for each arc is the same, we can view a path as an arc and use $t(v_{enter}, v_{exit}, l, \bar{v})$ to compute the travel time. From Lemma 1 we know $t(v_{enter}, v_{exit}, l, \bar{v})$ is a strictly increasing function of the travel length. So compared to the picked path, it always takes more time to travel through a path with length greater than $m + A$. Therefore the picked path is the optimal solution for the constructed instance of the Single Train Routing Problem.

Then we consider the scenario when the given instance of the Subset Sum Problem has no solution, i.e. $\forall\, S \subset \{0, \dots, m-1\}, \sum_{i \in S} a_i \neq A$. From Corollary 1, the optimal time for the corresponding Single Train Routing Problem will be no less than $\frac{v_s - v_e}{r_d} = 2\sqrt{m+A}$. By the same logic we showed above, the optimal path for the corresponding Single Train Routing Problem has length greater than $m + A$ (feasibility) and its length is closest to $m + A$ (optimality) among all feasible paths.

In summary, if we can solve the Single Train Routing Problem in polynomial time, then the Subset Sum Problem can be solved in polynomial time. Therefore the Single Train Routing Problem is NP-hard. □

3. Solution Properties

In Section 2 we showed that the Single Train Routing Problem is NP-hard, in this section we investigate the properties of the optimal solution and provide conditions on which the problem can be solved in polynomial time.

Suppose there are m+1 junction points and they can be train stations or simply railway junctions where trains can switch to another track. We denote the junction points by $n_0, n_1, \dots n_m$. Let the length and speed limit for the upper track between node n_{i-1} and node $n_i (i = 1, \dots, m)$ be l_{i1} and $\bar{v}_{i1}$ respectively. And the length and speed limit for the corresponding lower track are l_{i2} and $\bar{v}_{i2}$ respectively.

If $l_{i1} = l_{i2}$ for some i, the optimal path will always pick the track with the higher speed limit for the i^{th} segment no matter which track it picks for the other segments since $t(v_{enter}, v_{exit}, l, \bar{v})$ is a non-increasing function of the speed limit. So we only consider the scenario when $l_{i1} \neq l_{i2}$. Without loss of generality, we assume $l_{i1} > l_{i2} \quad (\forall i = 1, \dots m)$.

Intuitively, the optimal path has the incentive to select the track with shorter distance and higher speed limit to reduce the travel time. It takes less time to travel through a shorter distance and a higher speed limit enables the train to accelerate to a higher velocity. However in order to accelerate to a higher velocity to take advantage of the higher speed limit, a longer distance is needed for the acceleration. Therefore the two conditions, shorter distance and higher speed limit, are opposing each other in some way. Therefore our first claim is on the condition $\bar{v}_{i1} \leq \bar{v}_{i2}$, but we need a lemma before we give this statement.

Let p_k denote the subpath starting from the beginning of the first segment to the end of the k^{th} segment, i.e. $p_k = (l_1, \dots, l_k, \bar{v}_1, \dots, \bar{v}_k)(k = 1, \dots, m)$, where the length of the i^{th} segment in the subpath is l_i and the speed limit for the i^{th} segment in the subpath is $\bar{v}_i$ for $i = 1, \dots, k$.Note that when $k = m$, p_m gives the complete path. Let q_k denote the subpath starting from the beginning of the k^{th} segment to the end of the m^{th} segment, i.e. $q_k = (l_k, \dots, l_m, \bar{v}_k, \dots, \bar{v}_m)$,$(k = 1, \dots, m)$. Note that when $k = 1$, q_1 gives the complete path and is equivalent to p_m. Let $T(v_s, v_e, p_k)$ be the minimum travel time through path p_k starting with velocity of v_s and ending up with a velocity of v_e. Also take $p'_k = (l_1, \dots, l'_k, \bar{v}_1, \dots, \bar{v}_k)$ and $q'_k = (l'_k, \dots, l_m, \bar{v}_k, \dots, \bar{v}_m)$.

The following lemma (proof provided in the appendix) will provide the lower bound and the upper bound for the travel time if the starting velocity or the exiting velocity changes.(The starting and exiting velocities are given as inputs, but in order to investigate the solution properties and apply them to a subpath p_k we suppose their values can vary.) In order to achieve the lower bound and the upper bound, this lemma also proves the intuition that the travel time will be raised if the length of the last piece of track (or the first piece of track) is extended.

Lemma 2 Given $l'_m > l_m$, then we have:

If $T(0, v_e, p_m) < +\infty$, then $\forall v \in [0, v_e)$ we have

$$T(0, v_e, p_m) < T(0, v, p_m) < T(0, v_e, p_m) + \frac{v_e - v}{r_d} \tag{20}$$

$$T(0, v_e, p_m) < T(0, v_e, p'_m) < +\infty \tag{21}$$

If $T(v_s, 0, q_1) < +\infty$, then $\forall v \in [0, v_s)$ we have

$$T(v_s, 0, q_1) < T(v, 0, q_1) < T(v_s, 0, q_1) + \frac{v_s - v}{r_a} \tag{22}$$

$$T(v_s, 0, q_1) < T(v_s, 0, q'_1) < +\infty \tag{23}$$

Now we are ready to show our claim about the conditions $l_{i1} > l_{i2}$ and $\bar{v}_{i1} \leq \bar{v}_{i2}$.

Condition 1

$$v_s = v_e = 0$$

Proposition 4 Under Condition 1 and the assumption $l_{i1} > l_{i2}$, if $\bar{v}_{i1} \leq \bar{v}_{i2}$ for some i, the optimal path will always select the lower track for the i^{th} segment.

Proof: Suppose there is an optimal path picking the upper track for the i^{th} segment. Let the optimal entering velocity into the i^{th} segment be v_{i-1} and the optimal exiting velocity from the i^{th} segment be v_i.

Suppose $v_i^2 - v_{i-1}^2 \leq 2r_a l_{i2}$ and $v_{i-1}^2 - v_i^2 \leq 2r_d l_{i2}$, from Property 1 we know $t(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2}) < +\infty$. From Lemma 1 we know

$$t(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2}) < t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i2}) \leq t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}).$$

So

$$\begin{aligned} &T(0, v_{i-1}, p_{i-1}) + t(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2}) + T(v_i, 0, q_{i+1}) \\ &\qquad < T(0, v_{i-1}, p_{i-1}) + t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) + T(v_i, 0, q_{i+1}) \end{aligned}$$

where the RHS denotes the travel time for the optimal path and the LHS denotes the travel time if we only switch to the lower track for the i^{th} segment and keep the rest of the optimal path the same. This is a contradiction to the fact that the optimal path has least travel time. Therefore $v_i^2 - v_{i-1}^2 > 2r_a l_{i2}$ or $v_{i-1}^2 - v_i^2 > 2r_d l_{i2}$.

If $v_i^2 - v_{i-1}^2 > 2r_a l_{i2}$, take $v_i' = \sqrt{v_{i-1}^2 + 2r_a l_{i2}} < v_i$, then $t(v_{i-1}, v_i', l_{i2}, \bar{v}_{i2}) = \frac{v_i' - v_{i-1}}{r_a}$ and from Corollary 1 we have

$$t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) \geq \frac{v_i - v_{i-1}}{r_a} = t(v_{i-1}, v_i', l_{i2}, \bar{v}_{i2}) + \frac{v_i - v_i'}{r_a}$$

From Lemma 2 we have

$$T(v_i', 0, q_{i+1}) < T(v_i, 0, q_{i+1}) + \frac{v_i - v_i'}{r_a}$$

So

$$\begin{aligned} & T(0, v_{i-1}, p_{i-1}) + t(v_{i-1}, v_i', l_{i2}, \bar{v}_{i2}) + T(v_i', 0, q_{i+1}) \\ & < T(0, v_{i-1}, p_{i-1}) + t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) - \frac{v_i - v_i'}{r_a} + T(v_i, 0, q_{i+1}) + \frac{v_i - v_i'}{r_a} \\ & = T(0, v_{i-1}, p_{i-1}) + t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) + T(v_i, 0, q_{i+1}) \end{aligned}$$

where the RHS denotes the travel time for the optimal path and the LHS denotes the travel time if we only change the velocity at the end of the i^{th} segment from v_i to v_i' and keep the rest of the velocities the same. It contradicts the optimality of v_i.

Similarly, if $v_{i-1}^2 - v_i^2 > 2r_d l_{i2}$, take $v_{i-1}' = \sqrt{v_i^2 + 2r_d l_{i2}} < v_{i-1}$, then $t(v_{i-1}', v_i, l_{i2}, \bar{v}_{i2}) = \frac{v_{i-1}' - v_i}{r_d}$ and from Corollary 1 we have

$$t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) \geq \frac{v_{i-1} - v_i}{r_d} = t(v_{i-1}', v_i, l_{i2}, \bar{v}_{i2}) + \frac{v_{i-1} - v_{i-1}'}{r_d}$$

From Lemma 2 we have

$$T(0, v_{i-1}', p_{i-1}) < T(0, v_{i-1}, p_{i-1}) + \frac{v_{i-1} - v_{i-1}'}{r_d}$$

So

$$\begin{aligned} & T(0, v_{i-1}', p_{i-1}) + t(v_{i-1}', v_i, l_{i2}, \bar{v}_{i2}) + T(v_i, 0, q_{i+1}) \\ & < T(0, v_{i-1}, p_{i-1}) + \frac{v_{i-1} - v_{i-1}'}{r_a} + t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) - \frac{v_{i-1} - v_{i-1}'}{r_a} + T(v_i, 0, q_{i+1}) \\ & = T(0, v_{i-1}, p_{i-1}) + t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) + T(v_i, 0, q_{i+1}) \end{aligned}$$

This conclusion contradicts the fact that the optimal path picks the upper track. Therefore our assumption is wrong, and any optimal path will pick the lower track for the i^{th} segment. □

Notice that Condition 1 is necessary for Proposition 4. That is, without Condition 1, the optimal path may pick the upper track for the i^{th} segment even though we have conditions $l_{i1} > l_{i2}$ and $\bar{v}_{i1} \leq \bar{v}_{i2}$. A simple example is from the proof of Proposition 3, where the speed limits for the upper track and lower track are the same and a nonzero entering velocity is given. In this case, the optimal path may need to select the upper track for some segments in order to decelerate to a full stop (velocity of zero).

Due to Proposition 4, the Single Train Routing Problem when the starting and ending velocities of the paths are zero becomes simple given

Condition 2'

$$l_{i1} > l_{i2} \text{ and } \bar{v}_{i1} \leq \bar{v}_{i2}$$

for each i. Next we focus on the scenarios when the following condition is true:

Condition 2

$$l_{i1} > l_{i2} \text{ and } \bar{v}_{i1} \geq \bar{v}_{i2}$$

When Condition 2 is true, the train can either select the upper track and take advantage of a longer distance to accelerate to a higher velocity, or select the lower track and take advantage of a shorter distance to travel through. The optimal track selection depends on the problem parameters.

Next we give conditions on the parameters such that we can determine which track to choose for a certain segment. We first start with the simple idea that if the travel time for one track is strictly less than the travel time for the other track regardless of the entering and exiting velocities, then the track with less travel time will be selected. In other words, we need conditions such that the upper bound of the travel time for one track is strictly less than the lower bound of the travel time for the other track. So we have the following condition and proposition:

Condition 3

$$t(0, 0, l_{i1}, \bar{v}_{i1}) < \frac{l_{i2}}{\bar{v}_{i2}}$$

Proposition 5 Under conditions 2 and 3, the optimal path will always select the upper track for the i^{th} segment.

Proof: Suppose there exists an optimal path such that it picks the lower track for the i^{th} segment. Let v_{i-1}^* and v_i^* be the optimal velocities at the beginning and the end of the i^{th} segment for this optimal path. Since $t(v_{i-1}^*, v_i^*, l_{i2}, \bar{v}_{i2}) < +\infty$, from Property 1 and Condition 2 we have $v_{i-1}^* \leq \bar{v}_{i2} \leq \bar{v}_{i1}, v_i^* \leq \bar{v}_{i2} \leq \bar{v}_{i1}$ and $v_i^{*2} - v_{i-1}^{*2} \leq 2r_a l_{i2} < 2r_a l_{i1}$, $v_{i-1}^{*2} - v_i^{*2} \leq 2r_d l_{i2} < 2r_d l_{i1}$, which implies $t(v_{i-1}^*, v_i^*, l_{i1}, \bar{v}_{i1}) < +\infty$. From Lemma 1 we know

$$t(v_{i-1}^*, v_i^*, l_{i2}, \bar{v}_{i2}) \geq t(\bar{v}_{i2}, \bar{v}_{i2}, l_{i2}, \bar{v}_{i2}) = \frac{l_{i2}}{\bar{v}_{i2}}$$

$$t(v_{i-1}^*, v_i^*, l_{i1}, \bar{v}_{i1}) \leq t(0, 0, l_{i1}, \bar{v}_{i1})$$

From Condition 3 we have $t(v_{i-1}^*, v_i^*, l_{i1}, \bar{v}_{i1}) < t(v_{i-1}^*, v_i^*, l_{i2}, \bar{v}_{i2})$, which implies that if we change the i^{th} segment of the optimal path to the upper track, the travel time will be reduced. This is a contradiction. Therefore any optimal path will pick the upper track for the i^{th} segment. □

Similarly we propose conditions on which the upper track will be preferred. However, since the upper track has a higher speed limit, a feasible velocity for the upper track may not be feasible for the lower track. Therefore compared to Condition 3, we need a stricter condition to compensate for the infeasibility for the lower track.

Condition 4

$$t(0, 0, l_{i2}, \bar{v}_{i2}) + \frac{\bar{v}_{i1}}{r_a} + \frac{\bar{v}_{i1}}{r_d} < \frac{l_{i1}}{\bar{v}_{i1}}$$

Proposition 6 Under conditions 1, 2 and 4, the optimal path will always select the lower track for the i^{th} segment.

Proof: Suppose there exits an optimal path such that it picks the upper track for the i^{th} segment. We name this path as the base path p_m. Let v_{i-1}^* and v_i^* be the optimal velocities at the beginning and the end of the i^{th} segment for this optimal path. Now change the i^{th} segment for path p_m to the lower track. Let $v_{i-1}'^*$ and $v_i'^*$ be the optimal velocities at the beginning and the end of the i^{th} segment for the new path. So the optimal travel time for the base path is $T(0, v_{i-1}^*, p_{i-1}) + t(v_{i-1}^*, v_i^*, l_{i1}, \bar{v}_{i1}) + T(v_i^*, 0, q_{i+1})$ and the optimal travel time for the new path is $T(0, v_{i-1}'^*, p_{i-1}) + t(v_{i-1}'^*, v_i'^*, l_{i2}, \bar{v}_{i2}) + T(v_i'^*, 0, q_{i+1})$.

If $v_{i-1}'^* < v_{i-1}^*$, from Lemma 2 we have

$$T(0, v_{i-1}'^*, p_{i-1}) < T(0, v_{i-1}^*, p_{i-1}) + \frac{v_{i-1}^* - v_{i-1}'^*}{r_d} \leq T(0, v_{i-1}^*, p_{i-1}) + \frac{\bar{v}_{i1}}{r_d}$$

If $v_{i-1}'^{*} \geq v_{i-1}^{*}$, we also have

$$T(0, v_{i-1}'^{*}, p_{i-1}) \leq T(0, v_{i-1}^{*}, p_{i-1}) < T(0, v_{i-1}^{*}, p_{i-1}) + \frac{\bar{v}_{i1}}{r_d}$$

So $T(0, v_{i-1}'^{*}, p_{i-1}) < T(0, v_{i-1}^{*}, p_{i-1}) + \frac{\bar{v}_{i1}}{r_d}$ holds. Similarly, we have

$$T(v_i'^{*}, 0, q_{i+1}) < T(v_i^{*}, 0, q_{i+1}) + \frac{\bar{v}_{i1}}{r_a}$$

From Lemma 1 we have

$$t(v_{i-1}'^{*}, v_i'^{*}, l_{i2}, \bar{v}_{i2}) \leq t(0, 0, l_{i2}, \bar{v}_{i2})$$

$$\frac{l_{i1}}{\bar{v}_{i1}} = t(\bar{v}_{i1}, \bar{v}_{i1}, l_{i1}, \bar{v}_{i1}) \leq t(v_{i-1}^{*}, v_i^{*}, l_{i1}, \bar{v}_{i1})$$

Therefore,

$$\begin{aligned} &T(0, v_{i-1}'^{*}, p_{i-1}) + t(v_{i-1}'^{*}, v_i'^{*}, l_{i2}, \bar{v}_{i2}) + T(v_i'^{*}, 0, q_{i+1}) \\ &\quad < T(0, v_{i-1}^{*}, p_{i-1}) + \frac{\bar{v}_{i1}}{r_d} + t(0, 0, l_{i2}, \bar{v}_{i2}) + T(v_i^{*}, 0, q_{i+1}) + \frac{\bar{v}_{i1}}{r_a} \\ &\quad < T(0, v_{i-1}^{*}, p_{i-1}) + \frac{l_{i1}}{\bar{v}_{i1}} + T(v_i^{*}, 0, q_{i+1}) \\ &\quad \leq T(0, v_{i-1}^{*}, p_{i-1}) + t(v_{i-1}^{*}, v_i^{*}, l_{i1}, \bar{v}_{i1}) + T(v_i^{*}, 0, q_{i+1}) \end{aligned}$$

where the second inequality is from Condition 4. It contradicts that path p_m is the optimal path. Therefore any optimal path will pick the lower track for the i^{th} segment.□

Even though Condition 3 and Condition 4 can guarantee the track choice for some certain segment, they are very strict since the range of the travel time for picking one track does not overlap with the range of the travel time for picking the other track.

Next we give less restrictive conditions on which the ranges of the travel time for picking the track may overlap and still can guarantee the track choice. We first present a lemma (the proof is given in the appendix) before giving the conditions. This lemma compares the travel time for picking the upper track and the travel time for picking the lower track given the same entering velocity and the same exiting velocity. Define Δ as $\Delta(v_{i-1}, v_i) = t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) - t(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2})$, on the domain $\{(v_{i-1}, v_i) | \; t(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2}) < +\infty\}$.

Lemma 3 If Condition 2 is satisfied, $\Delta(v_{i-1}, v_i)$ is a strictly decreasing function of v_{i-1} and v_i.

Based on the fact that $\Delta(v_{i-1}, v_i)$ is strictly decreasing with respect to the two velocities v_{i-1} and v_i, we have the following condition and proposition:

Condition 5

$$t(0,0,l_{i1},\bar{v}_{i1}) < t(0,0,l_{i2},\bar{v}_{i2})$$

Proposition 7 Under conditions 2 and 5, the optimal path will always select the upper track for the i^{th} segment.

Proof: Suppose there exists an optimal path picking the lower track for the i^{th} segment. Let v_{i-1}^* and v_i^* be the optimal velocities at the beginning and the end of the i^{th} segment for this optimal path. Similar to the argument in Proposition 5, we have $t(v_{i-1}^*, v_i^*, l_{i1}, \bar{v}_{i1}) < +\infty$. Define $\Delta(v_{i-1}, v_i) = t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) - t(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2})$. From Lemma 3 we know $\Delta(v_{i-1}^*, v_i^*) \leq \Delta(0,0) = t(0,0,l_{i1},\bar{v}_{i1}) - t(0,0,l_{i2},\bar{v}_{i2}) < 0$, where the last inequality is due to Condition 5. So we have $t(v_{i-1}^*, v_i^*, l_{i1}, \bar{v}_{i1}) < t(v_{i-1}^*, v_i^*, l_{i2}, \bar{v}_{i2})$, which implies the travel time will be reduced if the train switches to the upper track for the i^{th} segment in the optimal path. It is a contradiction. Therefore any optimal path will pick the upper track for the i^{th} segment.□

Condition 5 is more general than Condition 3. In fact, we have the following:

$$t(0,0,l_{i1},\bar{v}_{i1}) < \frac{l_{i2}}{\bar{v}_{i2}} = t(\bar{v}_{i2},\bar{v}_{i2},l_{i2},\bar{v}_{i2}) < t(0,0,l_{i2},\bar{v}_{i2})$$

where the second inequality is due to Lemma 1. In order to give a measure of the restrictiveness of Condition 5, we present the following proposition. (proof provided in the appendix)

Proposition 8 If Conditions 2 and 5 hold, we have $l_{i2} > \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\bar{v}_{i2}^2$ i.e. $\sqrt{\frac{2r_a r_d l_{i2}}{r_a + r_d}} > \bar{v}_{i2}$.

Notice that the condition $l_{i2} > \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\bar{v}_{i2}^2$ guarantees that the train can reach the speed limit if the train picks the lower track for the i^{th} segment. Therefore Condition 5 is more general than Condition 3, but more restrictive than the condition $l_{i2} > \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\bar{v}_{i2}^2$.

Next we give a less restrictive condition to make the lower track preferred for the i^{th} segment. First we present a lemma (proof provided in the appendix) which provides a bound for the travel time over a node.

Lemma 4 If $v'_{enter} \geq v_{enter}$, $v'_{exit} \geq v_{exit}$ and $t(v'_{enter}, v'_{exit}, l, \bar{v}) \leq t(v_{enter}, v_{exit}, l, \bar{v}) < +\infty$, then we have

$$t(v_{enter}, v_{exit}, l, \bar{v}) < t(v'_{enter}, v'_{exit}, l, \bar{v}) + \frac{v'_{enter} - v_{enter}}{r_a} + \frac{v'_{exit} - v_{exit}}{r_d}$$

And similar to Condition 4 and Proposition 6, in order to make the lower track preferred, we need conditions to compensate for the infeasibility of the entering velocity and the exiting velocity when switching to the lower track since it has a lower speed limit. So we have:

Condition 6

$$\frac{l_{i2}}{\bar{v}_{i2}} + \frac{2\bar{v}_{i1}}{r_a} + \frac{2\bar{v}_{i1}}{r_d} < t(\bar{v}_{i2}, \bar{v}_{i2}, l_{i1}, \bar{v}_{i1})$$

Proposition 9 Under conditions 1, 2 and 6, the optimal path will always select the lower track for the i^{th} segment.

Proof: Suppose there exits an optimal path such that it picks the upper track for the i^{th} segment. We name this path as path p_m. Let v_{i-1}^* and v_i^* be the optimal velocities at the beginning and the end of the i^{th} segment for this optimal path. Now change the i^{th} segment for path p_m to the lower track and keep the rest of path p_m the same. Let $v_{i-1}'^*$ and $v_i'^*$ be the optimal velocities at the beginning and the end of the i^{th} segment for the new path. So the optimal travel time for path p_m is $T(0, v_{i-1}^*, p_{i-1}) + t(v_{i-1}^*, v_i^*, l_{i1}, \bar{v}_{i1}) + T(v_i^*, 0, q_{i+1})$ and the optimal travel time for the new path is $T(0, v_{i-1}'^*, p_{i-1}) + t(v_{i-1}'^*, v_i'^*, l_{i2}, \bar{v}_{i2}) + T(v_i'^*, 0, q_{i+1})$.

If $v_{i-1}'^* < v_{i-1}^*$, from Lemma 2 we have

$$T(0, v_{i-1}'^*, p_{i-1}) < T(0, v_{i-1}^*, p_{i-1}) + \frac{v_{i-1}^* - v_{i-1}'^*}{r_d} \leq T(0, v_{i-1}^*, p_{i-1}) + \frac{\bar{v}_{i1}}{r_d}$$

If $v_{i-1}'^* \geq v_{i-1}^*$, we also have

$$T(0, v_{i-1}'^*, p_{i-1}) \leq T(0, v_{i-1}^*, p_{i-1}) < T(0, v_{i-1}^*, p_{i-1}) + \frac{\bar{v}_{i1}}{r_d}$$

So $T(0, v_{i-1}'^*, p_{i-1}) < T(0, v_{i-1}^*, p_{i-1}) + \frac{\bar{v}_{i1}}{r_d}$ holds. Similarly, we have

$$T(v_i'^*, 0, q_{i+1}) < T(v_i^*, 0, q_{i+1}) + \frac{\bar{v}_{i1}}{r_a}$$

From Lemma 3 we have $\Delta(v_{i-1}'^*, v_i'^*) > \Delta(\bar{v}_{i2}, \bar{v}_{i2})$, i.e.

$$t(v_{i-1}'^*, v_i'^*, l_{i1}, \bar{v}_{i1}) - t(v_{i-1}'^*, v_i'^*, l_{i2}, \bar{v}_{i2}) > t(\bar{v}_{i2}, \bar{v}_{i2}, l_{i1}, \bar{v}_{i1}) - t(\bar{v}_{i2}, \bar{v}_{i2}, l_{i2}, \bar{v}_{i2})$$

$$= t(\bar{v}_{i2}, \bar{v}_{i2}, l_{i1}, \bar{v}_{i1}) - \frac{l_{i2}}{\bar{v}_{i2}}$$

Also from Lemma 4 we have

$$t(v_{i-1}'^*, v_i'^*, l_{i1}, \bar{v}_{i1}) < t(v_{i-1}^*, v_i^*, l_{i1}, \bar{v}_{i1}) + \frac{v_{i-1}^* - v_{i-1}'^*}{r_a} + \frac{v_{i-1}^* - v_{i-1}'^*}{r_d}$$

$$< t(v_{i-1}^{*}, v_{i}^{*}, l_{i1}, \bar{v}_{i1}) + \frac{\bar{v}_{i1}}{r_a} + \frac{\bar{v}_{i1}}{r_d}$$

So

$$t(v_{i-1}'^{*}, v_{i}'^{*}, l_{i2}, \bar{v}_{i2}) < t(v_{i-1}'^{*}, v_{i}'^{*}, l_{i1}, \bar{v}_{i1}) - t(\bar{v}_{i2}, \bar{v}_{i2}, l_{i1}, \bar{v}_{i1}) + \frac{l_{i2}}{\bar{v}_{i2}}$$
$$< t(v_{i-1}^{*}, v_{i}^{*}, l_{i1}, \bar{v}_{i1}) + \frac{\bar{v}_{i1}}{r_a} + \frac{\bar{v}_{i1}}{r_d} - t(\bar{v}_{i2}, \bar{v}_{i2}, l_{i1}, \bar{v}_{i1}) + \frac{l_{i2}}{\bar{v}_{i2}}$$
$$< t(v_{i-1}^{*}, v_{i}^{*}, l_{i1}, \bar{v}_{i1}) - \left(\frac{\bar{v}_{i1}}{r_a} + \frac{\bar{v}_{i1}}{r_d}\right) \qquad (24)$$

where the last inequality is due to Condition 6.

Therefore

$$T(0, v_{i-1}'^{*}, p_{i-1}) + t(v_{i-1}'^{*}, v_{i}'^{*}, l_{i2}, \bar{v}_{i2}) + T(v_{i}'^{*}, 0, q_{i+1})$$
$$< T(0, v_{i-1}^{*}, p_{i-1}) + \frac{\bar{v}_{i1}}{r_d} + t(v_{i-1}'^{*}, v_{i}'^{*}, l_{i2}, \bar{v}_{i2}) + T(v_{i}^{*}, 0, q_{i+1}) + \frac{\bar{v}_{i1}}{r_a}$$
$$< T(0, v_{i-1}^{*}, p_{i-1}) + t(v_{i-1}^{*}, v_{i}^{*}, l_{i1}, \bar{v}_{i1}) + T(v_{i}^{*}, 0, q_{i+1})$$

where the last inequality is due to (24), i.e. $t(v_{i-1}'^{*}, v_{i}'^{*}, l_{i2}, \bar{v}_{i2}) + \frac{\bar{v}_{i1}}{r_a} + \frac{\bar{v}_{i1}}{r_d} < t(v_{i-1}^{*}, v_{i}^{*}, l_{i1}, \bar{v}_{i1})$.

It contradicts that path p_m is the optimal path. Therefore any optimal path will pick the lower track for the i^{th} segment. □

Unfortunately the relationship between Condition 4 and Condition 6 is not as clear as the relationship between Condition 3 and Condition 5. Even though intuitively Condition 4 is stricter than Condition 6, the inequality from Lemma 4 determines the restrictiveness of Condition 6. If this inequality is not tight enough, Condition 4 may be more restrictive than Condition 6.

All the conditions are summarized in Table 1 and all the main propositions are summarized in Table 2. Notice that even though all the conclusions are valid for the double track railway system, Propositions 4-9 can be extended to other railway systems or even a general railway network. For the single track with sidings railway system, all the conclusions still apply since all the conclusions are based on local optimality conditions, i.e. we only switch to another track for a certain segment and then compare the new path with the original path. For the same reason, we can even extend all the conclusions to a triple track railway system by performing pair-wise comparisons between all tracks and identifying if there exists any track that dominates the other two in terms of the problem parameters. In fact, this approach can be extended to any number of tracks that are connected by two junctions. Even though all the conditions and propositions discussed in this section only give local optimality conditions, it is easy to extend to the entire optimal solution if every segment can be applied by one of the propositions, which is summarized by the following corollary.

Corollary 2 Under Condition 1, the Single Train Routing problem for a single line double track railway system can be solved in polynomial time if every segment can be applied by one of the propositions 4, 5, 6, 7 and 9.

Table 1 Conditions Summary

Condition 1	$v_s = v_e = 0$
Condition 2	$l_{i1} > l_{i2}$ and $\bar{v}_{i1} \geq \bar{v}_{i2}$
Condition 2'	$l_{i1} > l_{i2}$ and $\bar{v}_{i1} \leq \bar{v}_{i2}$
Condition 3	$t(0,0,l_{i1},\bar{v}_{i1}) < \frac{l_{i2}}{\bar{v}_{i2}}$
Condition 4	$t(0,0,l_{i2},\bar{v}_{i2}) + \frac{\bar{v}_{i1}}{r_a} + \frac{\bar{v}_{i1}}{r_d} < \frac{l_{i1}}{\bar{v}_{i1}}$
Condition 5	$t(0,0,l_{i1},\bar{v}_{i1}) < t(0,0,l_{i2},\bar{v}_{i2})$
Condition 6	$\frac{l_{i2}}{\bar{v}_{i2}} + \frac{2\bar{v}_{i1}}{r_a} + \frac{2\bar{v}_{i1}}{r_d} < t(\bar{v}_{i2},\bar{v}_{i2},l_{i1},\bar{v}_{i1})$

Table 2 Summary of the Propositions

Necessary Condition	Proposition 4	Proposition 5	Proposition 6	Proposition 7	Proposition 9
Condition 1	✓		✓		✓
Condition 2		✓	✓	✓	✓
Condition 2'	✓				
Condition 3		✓			
Condition 4			✓		
Condition 5				✓	
Condition 6					✓
Selected Track	Lower track	Upper track	Lower track	Upper track	Lower track

4. A Discrete Dynamic Programming based Heuristic

Even though many conditions are proposed to enforce that the Single Train Routing Problem can be solved in polynomial time, there are still many scenarios when the parameters do not meet the given conditions. Therefore in this section we propose a dynamic programming approximation algorithm to solve the Single Train Routing Problem regardless of the parameter values. In this section we first describe the algorithm and then we investigate the scenarios when the proposed algorithm only takes polynomial time to solve the Single Train Routing Problem.

The algorithm is based on the idea of recording the optimal travel time for different exiting velocities at the end of each segment sequentially. Since we do not know which exiting velocity will be used in the following segments, a simple idea is to record all the travel times for all possible exiting velocities. However it is impossible to keep a record of an infinite number of travel times. Therefore we can keep track of travel times only for a finite number of exiting velocities to obtain an approximate solution. Let

$$St_i = \{(v_{i1}, tt_{i1}), (v_{i2}, tt_{i2}), \dots, (v_{ir_i}, tt_{ir_i})\}$$

be the record we keep at the end of the i^{th} segment, where r_i is the total number of exiting velocities we record at the end of the i^{th} segment.

One way to compute St_i is to choose a discretization size d, and then discretize the velocity $\max(\bar{v}_{i1}, \bar{v}_{i2})$. In other words, take $r_i = \left\lfloor \frac{\max(\bar{v}_{i1}, \bar{v}_{i2})}{d} \right\rfloor + 1$, where $\lfloor x \rfloor$ is the largest integer no more than x, and $v_{ij} = (j-1) * d (\forall j \in \{1, \dots, r_i\})$. Then the algorithm is described as follows:

1. Set $St_0 = \{(v_s, 0)\},\ r_0 = 1$ and $i = 1$.
2. Let $r_i = \left\lfloor \frac{\max(\bar{v}_{i1}, \bar{v}_{i2})}{d} \right\rfloor + 1$. $\forall j \in \{1, \dots, r_i\}$ set $v_{ij} = (j-1) * d$ and

$$tt_{ij} = \min_{j_0 = 1, \dots r_{i-1}} \{tt_{(i-1)j_0} + \min[t(v_{(i-1)j_0}, v_{ij}, l_{i1}, \bar{v}_{i1}), t(v_{(i-1)j_0}, v_{ij}, l_{i2}, \bar{v}_{i2})]\}$$

   Then set $St_i = \{(v_{i1}, tt_{i1}), \dots, (v_{ir_i}, tt_{ir_i})\}$

3. Let $i = i + 1$. If $i = m$ go to Step 4, otherwise go to Step 2.
4. Set

$$tt^* = \min_{j_0=1,\ldots r_{m-1}} \{tt_{(m-1)j_0} + \min[t(v_{(m-1)j_0}, v_e, l_{m1}, \bar{v}_{m1}), t(v_{(m-1)j_0}, v_e, l_{m2}, \bar{v}_{m2})]\}$$

and return tt^*.

The equation in Step 2

$$tt_{ij} = \min_{j_0=1,\ldots r_{i-1}} \{tt_{(i-1)j_0} + \min[t(v_{(i-1)j_0}, v_{ij}, l_{i1}, \bar{v}_{i1}), t(v_{(i-1)j_0}, v_{ij}, l_{i2}, \bar{v}_{i2})]\}$$

represents the relationship between St_{i-1} and St_i, i.e. how to generate St_i given St_{i-1} using the idea of forward recursion. And when the velocities for the upper track are not feasible, i.e. $t(v_{(i-1)j_0}, v_{ij}, l_{i1}, \bar{v}_{i1}) = +\infty$, based on the minimum rule the lower track will be selected. If the velocities are infeasible for both tracks, i.e. $t(v_{(i-1)j_0}, v_{ij}, l_{i1}, \bar{v}_{i1}) = +\infty$ and $t(v_{(i-1)j_0}, v_{ij}, l_{i2}, \bar{v}_{i2}) = +\infty$, we have $tt_i = +\infty$.

Both the solution quality and computation time depends on the discretization size d. When d is large, the size of St_i is small and the solution computation time is short. However, the obtained solution may not be close to the optimal solution due to the coarse discretization. In contrast, with a small d, the solution can be improved and also the solution computation time becomes long.

In some certain scenarios, the Single Train Routing Problem can be solved in polynomial time if St_i is selected carefully. We consider the scenarios when each track segment is long enough to make any entering and exiting velocities feasible as long as they are less than the speed limit, i.e. $\forall v_{enter} \in [0, \bar{v}], v_{exit} \in [0, \bar{v}]$, we have $t(v_{enter}, v_{exit}, l, \bar{v}) < +\infty$. If we have $l \geq \max\left(\frac{1}{2r_a}, \frac{1}{2r_d}\right) \bar{v}^2$, then

$$v_{enter}^2 - v_{exit}^2 \leq \bar{v}^2 \leq 2r_d l$$
$$v_{exit}^2 - v_{enter}^2 \leq \bar{v}^2 \leq 2r_a l$$

From Property 1 we know $t(v_{enter}, v_{exit}, l, \bar{v}) < +\infty$. Therefore we are interested in the following conditions:

Condition 7

$$l_{ij} \geq \max\left(\frac{1}{2r_a}, \frac{1}{2r_d}\right)\bar{v}_{ij}^2 \quad \forall i = 1, \ldots, m \; \forall j = 1, 2$$

Proposition 10 Under Condition 7, the Single Train Routing problem for a single line double track railway system can be solved in polynomial time.

Proof: Suppose the optimal path picks β_i^{th} ($\beta_i \in \{1,2\}$) track for the i^{th} segment ($i = 1, \dots, m$), where the first track denotes the upper track and the second track denotes the lower track for each segment. Let the corresponding optimal velocities be v_i^* ($i = 0, \dots, m$). Then consider the new solution v_i' such that $v_0' = v_s, v_m' = v_e$ and $v_i' = \min(\bar{v}_{i\beta_i}, \bar{v}_{(i+1)\beta_{i+1}})$ $\forall i = 1, \dots, m-1$. Due to the feasibility of the optimal solution, we have $v_i^* \le v_i{}'$ ($i = 0, \dots, m$)

On the other hand, we show v_i' ($i = 0, \dots, m$) is also a feasible solution, i.e. $t(v_{i-1}', v_i', l_{i\beta_i}, \bar{v}_{i\beta_i}) < +\infty \;\; \forall i = 1, \dots, m$. From Condition 7 we have

$$v_i'^2 - v_{i-1}'^2 \le \bar{v}_{i\beta_i}^2 \le 2r_a l_{i\beta_i}$$

$$v_{i-1}'^2 - v_i'^2 \le \bar{v}_{i\beta_i}^2 \le 2r_d l_{i\beta_i}$$

From Property 1 we know $t(v_{i-1}', v_i', l_{i\beta_i}, \bar{v}_{i\beta_i}) < +\infty \;\; \forall i = 1, \dots, m$. So from Proposition 1 we know $v_i' \le v_i^*$ ($i = 0, \dots, m$).

Therefore $v_i^* = v_i' = \min(\bar{v}_{i\beta_i}, \bar{v}_{(i+1)\beta_{i+1}})$ $\forall i = 1, \dots, m-1$.

Consider the following state sets:

$$St_i = \{(v_{i1}, tt_{i1}), (v_{i2}, tt_{i2}), (v_{i3}, tt_{i3}), (v_{i4}, tt_{i4})\}$$

where $v_{i1} = \min(\bar{v}_{i1}, \bar{v}_{(i+1)1})$, $v_{i2} = \min(\bar{v}_{i1}, \bar{v}_{(i+1)2})$, $v_{i3} = \min(\bar{v}_{i2}, \bar{v}_{(i+1)1})$ and $v_{i4} = \min(\bar{v}_{i2}, \bar{v}_{(i+1)2})$. Then we have $v_i^* \in St_i$. Therefore if we perform the dynamic programming algorithm on the constructed St_i sequentially, the optimal solution can be obtained.

Since the size of St_i is constant at each step, it only takes polynomial time for the dynamic programming algorithm to terminate. □

Next we consider the following conditions:

Condition 8 $\forall i = 1, \dots, m-1$

$$l_{ij} \ge \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\bar{v}_{ij}^2 \quad \forall i = 1, \dots, m \; \forall j = 1, 2$$

Since Condition 7 is a necessary condition for Condition 8, we can have the following corollary immediately.

Corollary 3 Under Condition 8, the Single Train Routing problem for a single line double track railway system can be solved in polynomial time.

Notice that Nagarajan and Ranade (2008) also gave a proof for the polynomial time to obtain the optimal solution when Condition 8 is true. However their algorithm is based on the expansion of profiles (all possible tuples of travel time and arrival velocities) over arcs and it is difficult to implement. In contrast, the dynamic programming algorithm we proposed is easier to implement and allows a less restrictive condition (Condition 7) to guarantee the polynomial time to obtain the optimal solution.

5. Experimental Results

Previously, we gave conditions on the problem parameters when the problem is polynomial. For general parameter settings, the problem was shown to be NP-hard and we proposed a dynamic programming heuristic to solve the problem. In section 3 we also show several local optimality conditions, which can also lead to another heuristic method. In this section, we experimentally test these derived heuristics for general parameter settings.

5.1 Dynamic Programming Based Heuristic

We first experimentally test the performance of the dynamic programming heuristic presented in Section 4 on general parameter settings. The parameter settings are given as follows. The double track railway system contains 10 segments. The length for each track is generated from uniform distribution between 0.5 miles and 1.5 miles and the speed limit for each track is generated from uniform distribution between 10 miles/hour to 80 miles/hour. Both the acceleration rate and the deceleration rate are set to $2112\ \text{feet/minute}^2$. And both the entering velocity and the exiting velocity are set to 0.

We first randomly generate 10 instances and vary the step size d from 45 miles/hour to 1 mile/hour. The comparison between our proposed dynamic programming and the optimal solution is measured in the fraction of the difference with respective to the optimal travel time, i.e. $\frac{travel\ time - optimal\ travel\ time}{optimal\ travel\ time}$ and is presented in Table 3, where each row represents one instance and each column denotes a different step size in the dynamic program. The optimal solution is found through enumeration.

Table 3 The fraction difference between the dynamic programming heuristic and the optimal solution

		Step size									
		45	40	35	30	25	20	15	10	5	1
Case	1	0.6489	0.4177	0.1390	0.1384	0.1424	0.1414	0.0698	0.0435	0.0121	0.0016
	2	0.3832	0.2545	0.3010	0.3622	0.3509	0.0447	0.0762	0.0343	0.0246	0.0036

	3	0.3077	0.3310	0.3612	0.2469	0.0564	0.1350	0.0363	0.0458	0.0130	0.0034
	4	0.3620	0.3533	0.2512	0.2965	0.1914	0.2310	0.0625	0.0746	0.0256	0.0013
	5	0.2436	0.1348	0.1887	0.2246	0.1977	0.0593	0.0464	0.0312	0.0120	0.0020
	6	0.5768	0.4720	0.4127	0.3502	0.2909	0.1013	0.1153	0.0532	0.0305	0.0032
	7	0.4568	0.2822	0.3061	0.3391	0.2891	0.0341	0.0728	0.0104	0.0104	0.0025
	8	0.4031	0.1607	0.2032	0.1877	0.2432	0.0804	0.0643	0.0380	0.0131	0.0026
	9	0.1893	0.2143	0.2555	0.2502	0.2287	0.1273	0.0448	0.0410	0.0136	0.0033
	10	0.4557	0.1515	0.1923	0.2107	0.1800	0.1332	0.1471	0.0202	0.0160	0.0053

As expected, from Table 3, the fraction difference between our proposed heuristic and the optimal solution decreases as the step size decreases. When the step size is 1 mile/hour, the percentage difference becomes less than 1% for all instances.

We now vary the acceleration and deceleration rates to see how they will affect the results at different levels of the step size. The results are presented in Table 4. Each row denotes some certain acceleration/deceleration rate. For example for the first row we set the acceleration and deceleration to be $528\ \text{feet/miniute}^2$. Each column represents a different step size in the dynamic programming algorithm. We generate 50 instances and the number in each cell is the average fraction difference between the proposed dynamic programming algorithm and the optimal solution.

Table 4 The average fraction difference between the dynamic programming heuristic and the optimal solution

		Step size									
		45	40	35	30	25	20	15	10	5	1
Acceleration/deceleration	528	1.9332	1.9332	1.9332	1.9332	1.9332	0.5758	0.3312	0.2445	0.0964	0.0140
	1056	1.7131	1.7131	1.7131	1.1109	0.4100	0.3191	0.2448	0.1317	0.0607	0.0076
	1584	1.5356	1.5356	0.8989	0.5210	0.3002	0.2761	0.1801	0.0971	0.0350	0.0059
	2112	1.3005	0.8944	0.5847	0.3965	0.2827	0.2151	0.1359	0.0772	0.0297	0.0048
	2640	0.8807	0.6492	0.4660	0.3306	0.2500	0.1751	0.1127	0.0642	0.0228	0.0042
	3168	0.7106	0.5098	0.3951	0.3110	0.2163	0.1486	0.1008	0.0571	0.0210	0.0038
	3696	0.5779	0.4330	0.3440	0.2870	0.1909	0.1281	0.0905	0.0520	0.0180	0.0034
	4224	0.4842	0.3801	0.3213	0.2636	0.1709	0.1176	0.0822	0.0481	0.0167	0.0031

Table 4 shows that as the acceleration/deceleration rates increase, the average fraction difference decreases, which implies the dynamic programming algorithm performs better at higher rates since as the acceleration and deceleration rates approach infinity the problem becomes polynomial solvable and the dynamic program becomes trivial to solve. The results in Table 3 and Table 4 show that with a step size of 1 mile/hour the proposed algorithm is at least within 1.5% of the optimal even when the acceleration and deceleration rates are very small.

Next we vary the distribution of the length and the speed limit by changing their upper bounds in order to test the robustness when we set the step size to 1 mile/hour. The results are shown in Tables 5 and 6. Table 5 shows how the average difference between the dynamic programming algorithm and the optimal solution changes as the acceleration and deceleration rates and the upper bound for each track length change. Each row represents one acceleration and deceleration rate. And each column corresponds to one upper bound for each track length. For example, for the first column, we generate 50 instances where the track length is from a uniform distribution between 0.5 miles and 1 mile and then the average fraction difference of the 50 instances is calculated for each cell.

Table 5 The average fraction difference between the dynamic programming heuristic and the optimal solution as acceleration/deceleration rate and upper bound for the length change

	Upper bound for the track length										
Acceleration/deceleration rate	1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
528	0.0159	0.0171	0.0166	0.0143	0.0156	0.0141	0.0142	0.0132	0.0135	0.0132	0.0120
1056	0.0096	0.0098	0.0088	0.0093	0.0081	0.0081	0.0080	0.0078	0.0074	0.0075	0.0074
1584	0.0070	0.0070	0.0059	0.0062	0.0071	0.0064	0.0055	0.0068	0.0056	0.0053	0.0058
2112	0.0057	0.0061	0.0049	0.0054	0.0054	0.0049	0.0044	0.0053	0.0044	0.0043	0.0050
2640	0.0050	0.0049	0.0038	0.0047	0.0048	0.0040	0.0036	0.0043	0.0036	0.0036	0.0042
3168	0.0043	0.0042	0.0034	0.0041	0.0040	0.0039	0.0032	0.0039	0.0033	0.0032	0.0036
3696	0.0039	0.0037	0.0032	0.0038	0.0037	0.0035	0.0028	0.0034	0.0030	0.0029	0.0032
4224	0.0036	0.0034	0.0029	0.0035	0.0033	0.0032	0.0025	0.0031	0.0027	0.0028	0.0029
4752	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

From Table 5, as the upper bound for each track length increases the average fraction difference decreases. Intuitively, as the upper bound for each track length increases and the acceleration and deceleration rates increase, the probability of Condition 7 being satisfied becomes larger and the proposed dynamic programming algorithm can provide a better approximation. Also the percentage difference is below 2% for all scenarios, which implies that the step size of 1 mile/hour is enough to obtain a relatively good solution.

Table 6 shows how the average fraction difference between the dynamic programming algorithm and the optimal solution as the acceleration and deceleration rates and the upper bound for the speed limit for each track change. Each row represents one acceleration and deceleration rate. And each column corresponds to one upper bound for each speed limit. For example, for the first column, we generate 50 instances where the speed limit is from a uniform distribution between 10 miles/hour and 55 miles/hour and then the average fraction difference of the 50 instances is calculated for each cell.

Table 6 The average fraction difference between the dynamic programming heuristic and the optimal solution as acceleration/deceleration rate and upper bound for the speed limit change

		Upper bound for the speed limit										
		55	60	56	70	75	80	85	90	95	100	105
Acceleration/deceleration rate	528	0.0100	0.0124	0.0139	0.0133	0.0143	0.0145	0.0159	0.0138	0.0141	0.0137	0.0150
	1056	0.0069	0.0066	0.0079	0.0078	0.0080	0.0081	0.0088	0.0088	0.0092	0.0083	0.0090
	1584	0.0052	0.0051	0.0060	0.0058	0.0058	0.0063	0.0068	0.0067	0.0065	0.0065	0.0068
	2112	0.0044	0.0044	0.0049	0.0046	0.0044	0.0051	0.0050	0.0053	0.0051	0.0051	0.0048
	2640	0.0038	0.0037	0.0042	0.0039	0.0038	0.0046	0.0042	0.0042	0.0044	0.0046	0.0042
	3168	0.0034	0.0032	0.0037	0.0035	0.0034	0.0039	0.0040	0.0038	0.0039	0.0038	0.0038
	3696	0.0030	0.0030	0.0034	0.0032	0.0030	0.0036	0.0035	0.0035	0.0035	0.0035	0.0036
	4224	0.0027	0.0027	0.0031	0.0030	0.0028	0.0032	0.0032	0.0031	0.0033	0.0032	0.0032
	4752	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

From Table 6, as the upper bound for the speed limit of each track increases the average fraction difference increases. Intuitively, as the upper bound for the speed limit increases and the acceleration and deceleration rates decrease, the probability of Condition 7 being satisfied becomes smaller and the proposed dynamic programming algorithm performs worse. Still the percentage difference is still below 2% for all scenarios, which implies that a step size of 1 mile/hour is robust.

5.2 Local Based Heuristics

The drawback with a dynamic programming based approach is that they are not amenable for real time control. Therefore, if a real time based approach for train routing is preferred the local optimality conditions presented in Section 3 (Conditions 1-6) can provide a framework for an efficient heuristic. A simple intuitive heuristic for real time control is to ignore the acceleration and deceleration process, i.e. treat the travel time as the ratio of the length and the speed limit and select the segment with the smaller ratio. We call this method as the Simple Greedy heuristic. To make a comparison, we can combine this Simple Greedy heuristic with the proposed conditions 1-6. We name the new method as the Extended Greedy heuristic. The detailed procedure of the Extended Greedy heuristic is given as follows:

For segment *i* with upper track $(l_{i1}, \bar{v}_{i1})$ and lower track $(l_{i2}, \bar{v}_{i2})$:

1. If $l_{i1} = l_{i2}$, select the track with higher speed limit.
2. If $l_{i1} > l_{i2}$,
   2.1 If $\bar{v}_{i1} \leq \bar{v}_{i2}$, select the lower track.
   2.2 Else if condition 2 and 5 are satisfied, select the upper track.
   2.3 Else if condition 2 and 4 are satisfied, select the lower track.
   2.4 Else if condition 2 and 6 are satisfied, select the lower track.
   2.5 Else select the track with lower ratio of the length and the speed limit.
3. If $l_{i1} < l_{i2}$, the same as scenario 2 , but swap the lower track with the upper track, and vice versa.

We use the same experimental settings as in Section 5.1 except the sampling is done to ensure the average difference between the two track lengths within the segment is 0.5 mile and the average difference between the two speed limits is 30 miles/hour. Each track length is uniformly generated between 0 to its upper bound while each speed limit is generated between 10 mile/hour to its upper bound. The average differences mentioned above are set by using a different upper bound for the two tracks within the same segment.

We first vary the upper bound for the track length and the acceleration/deceleration rate. For each specified upper bound for track length and specified acceleration/deceleration rate, 50 random networks are generated. Comparison between the Simple Greedy heuristic, the Extended Greedy heuristic and the optimal solution (obtained through enumeration) is made and the results are summarized in Tables 7 and 8. The unit for acceleration/deceleration rate is feet/min*min and the unit for track length is miles.

Table 7 Average fraction difference between the Simple Greedy heuristic and the optimal solution as the acceleration/deceleration rate and upper bound for track length change

		Upper Bound for track length										
		1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
Acceleration/deceleration rate	264	0.3425	0.2687	0.2299	0.2104	0.2198	0.1799	0.1973	0.1777	0.1757	0.1744	0.1706
	528	0.3063	0.2289	0.1948	0.1898	0.1806	0.1561	0.1621	0.1377	0.1442	0.1432	0.1368
	792	0.2789	0.2064	0.1659	0.1720	0.1613	0.1436	0.1384	0.1227	0.1269	0.1186	0.1221
	1056	0.2518	0.1899	0.1467	0.1573	0.1457	0.1325	0.1249	0.1122	0.1134	0.1012	0.1171
	1320	0.2296	0.1743	0.1305	0.1447	0.1348	0.1250	0.1167	0.1048	0.1034	0.0908	0.1132
	1584	0.2111	0.1608	0.1198	0.1346	0.1272	0.1197	0.1092	0.1016	0.0961	0.0850	0.1103
	1848	0.1959	0.1488	0.1126	0.1267	0.1202	0.1153	0.1037	0.0992	0.0912	0.0840	0.1079
	2112	0.1838	0.1393	0.1077	0.1209	0.1146	0.1102	0.0990	0.0974	0.0880	0.0830	0.1055

Table 8 The average fraction difference between the Extended Greedy heuristic and the optimal solution as the acceleration/deceleration rate and upper bound for track length change

		Upper Bound for track length										
		1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
Acceleration/deceleration rate	264	0.0031	0.0054	0.0222	0.0286	0.0150	0.0230	0.0369	0.0398	0.0448	0.0440	0.0493
	528	0.0157	0.0292	0.0405	0.0561	0.0407	0.0549	0.0588	0.0502	0.0651	0.0630	0.0604
	792	0.0312	0.0537	0.0401	0.0640	0.0579	0.0701	0.0635	0.0595	0.0663	0.0721	0.0681
	1056	0.0428	0.0682	0.0488	0.0707	0.0664	0.0781	0.0681	0.0530	0.0625	0.0734	0.0743
	1320	0.0481	0.0742	0.0508	0.0654	0.0706	0.0807	0.0724	0.0541	0.0645	0.0686	0.0793
	1584	0.0532	0.0772	0.0561	0.0657	0.0729	0.0831	0.0748	0.0595	0.0654	0.0667	0.0829
	1848	0.0559	0.0740	0.0610	0.0636	0.0705	0.0847	0.0766	0.0635	0.0647	0.0698	0.0846
	2112	0.0553	0.0734	0.0647	0.0642	0.0716	0.0838	0.0744	0.0668	0.0603	0.0710	0.0848

From Table 7, the Simple Greedy heuristic method improves as the acceleration/deceleration rate increases and the upper bound for the track length increases. This makes sense considering that the Simple Greedy heuristic method can give the optimal solution as the acceleration/deceleration rate goes to infinity or the acceleration/deceleration time is negligible compared to the total travel time. Table 8 shows that the Extended Greedy heuristic method is stable across different acceleration/deceleration rates and upper bounds for track length. It performs much better than the Simple Greedy Heuristic method, especially when the acceleration/deceleration rate is low and the upper bound for the track length is small.

Tables 9 and 10 show the same results when we vary the upper bound for the speed limit instead of the upper bound of the track length. The same conclusions as before can also be drawn from these results. That is, the Simple Greedy heuristic performance improves as the acceleration rates and speed limits increase and the Extended Greedy heuristic consistently outperforms the Simple Greedy heuristic.

Table 9 The average fraction difference between the Simple Greedy heuristic and the optimal solution as the acceleration/deceleration rate and upper bound for speed limit change

		Upper Bound for speed limit										
		55	60	65	70	75	80	85	90	95	100	105
Acceleration/deceleration rate	264	0.2314	0.2294	0.2131	0.1948	0.1929	0.1956	0.2025	0.1824	0.1593	0.1843	0.1653
	528	0.1789	0.1891	0.1800	0.1648	0.1657	0.1631	0.1677	0.1587	0.1326	0.1487	0.1431
	792	0.1573	0.1646	0.1530	0.1417	0.1476	0.1480	0.1487	0.1373	0.1139	0.1304	0.1290
	1056	0.1447	0.1458	0.1346	0.1242	0.1320	0.1379	0.1367	0.1244	0.1029	0.1192	0.1174
	1320	0.1371	0.1347	0.1217	0.1128	0.1192	0.1308	0.1312	0.1174	0.0995	0.1115	0.1076
	1584	0.1331	0.1294	0.1130	0.1063	0.1101	0.1251	0.1283	0.1125	0.0961	0.1051	0.1004
	1848	0.1327	0.1268	0.1062	0.1017	0.1037	0.1211	0.1242	0.1080	0.0918	0.0994	0.0955
	2112	0.1336	0.1250	0.1017	0.0974	0.0996	0.1181	0.1201	0.1045	0.0882	0.0950	0.0916

Table 10 The average fraction difference between the Extended Greedy heuristic and the optimal solution as the acceleration/deceleration rate and upper bound for speed limit change

		Upper Bound for speed limit										
		1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
Acceleration/deceleration rate	264	0.0330	0.0307	0.0411	0.0452	0.0363	0.0228	0.0369	0.0336	0.0354	0.0264	0.0268
	528	0.0638	0.0606	0.0763	0.0598	0.0669	0.0494	0.0499	0.0609	0.0544	0.0460	0.0478
	792	0.0806	0.0686	0.0806	0.0697	0.0757	0.0704	0.0627	0.0610	0.0572	0.0576	0.0591
	1056	0.0812	0.0682	0.0740	0.0705	0.0780	0.0809	0.0702	0.0591	0.0599	0.0645	0.0618
	1320	0.0832	0.0735	0.0723	0.0664	0.0759	0.0838	0.0798	0.0628	0.0651	0.0651	0.0581
	1584	0.0870	0.0800	0.0709	0.0685	0.0732	0.0867	0.0863	0.0702	0.0636	0.0660	0.0556

	1848	0.0886	0.0816	0.0688	0.0624	0.0716	0.0858	0.0887	0.0761	0.0644	0.0673	0.0572
	2112	0.0909	0.0860	0.0696	0.0604	0.0715	0.0882	0.0881	0.0804	0.0636	0.0672	0.0554

6. Conclusions

In this paper we considered the Single Train Routing Problem, i.e. routing one train through an empty railway network. We proved that this problem is NP-hard and the decision problem is NP-complete. Also we investigated different scenarios when the parameters satisfy certain conditions such that we know the local choice for the optimal solution. Also we proposed a dynamic programming heuristic framework to solve the problem. Several scenarios when the proposed algorithm can obtain the optimal solution and only take polynomial time are investigated. Experimental analysis showed the efficiency of the proposed heuristics for general parameter settings.

Acknowledgement

This research was partially supported by the Volvo Research and Education Foundation.

References

Ahuja, R. K., J. B. Orlin, S. Pallottino, and M. G. Scutella. 2003. Dynamic shortest paths minimizing travel times and costs. *Networks* 41 (4): 197-205.

Cai, X., T. Kloks, and C.K. Wong. 1997. Time-varying shortest path problems with constraints. *Networks* 29 (3): 141-50.

Chabini, Ismail. 1998. Discrete dynamic shortest path problems in transportation applications: Complexity and algorithms with optimal run time. *Transportation Research Record: Journal of the Transportation Research Board*(1645): 170-5.

Cooke, K. L., and E. Halsey. 1966. The shortest route through a network with time-dependent internodal transit times. *Journal of Mathematical Analysis and Applications* 14 (3): 493-8.

Fischer, F., and C. Helmberg. 2014. Dynamic graph generation for the shortest path problem in time expanded networks. *Mathematical Programming* 143 (1-2) (02/01): 257-97.

Kaufman, D.E., and R. L. Smith. 1993. Fastest paths in time-dependent networks for intelligent vehicle-highway systems application. *Journal of Intelligent Transportation Systems* 1 (1): 1-11.

Koch, R., and E. Nasrabadi. 2014. Continuous-time dynamic shortest path problems with negative transit times. *SIAM Journal on Control and Optimization* (01/01; 2014/08): 2449-81.

Lu, Q., M. M. Dessouky, and R. C. Leachman. 2004. Modeling train movements through complex rail networks. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 14 (1): 48-75.

Nagarajan, V., and A. G. Ranade. 2008. Exact train pathing. *Journal of Scheduling* 11 (4): 279-97.

Orda, A., and R. Rom. 1990. Shortest-path and minimum-delay algorithms in networks with time-dependent edge-length. *Journal of the ACM (JACM)* 37 (3): 607-25.

Orda, A., and R. Rom. 1991. Minimum weight paths in time-dependent networks. *Networks* 21 (3): 295-319.

Pallottino, S., and M. G. Scutella. 1998. Shortest path algorithms in transportation models: Classical and innovative aspects. In *Equilibrium and advanced transportation modelling.*, 245-281Springer.

Philpott, A. B. 1994. Continuous-time shortest path problems and linear programming. *SIAM Journal on Control and Optimization* 32 (2): 538-52.

Appendix

Proof of Lemma 1: When $t(v_{enter}, v_{exit}, l, \bar{v}) < +\infty$, $t(v_{enter}, v_{exit}, l, \bar{v})$ has two forms as follows:

1. If $v_{exit} \leq \bar{v},\ v_{enter} \leq \bar{v},\ v_{exit}^2 - v_{enter}^2 \leq 2r_a l,\ v_{enter}^2 - v_{exit}^2 \leq 2r_d l,$ $\sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}} \leq \bar{v}$

$$\mathrm{t}(v_{enter}, v_{exit}, l, \bar{v}) = -\frac{v_{enter}}{r_a} - \frac{v_{exit}}{r_d} + (\frac{1}{r_a} + \frac{1}{r_d})\sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}}$$

$$\frac{\partial \mathrm{t}}{\partial \mathrm{v_{enter}}} = \frac{1}{r_a}\left(\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}\right)^{-1/2} (v_{enter} - \sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}})$$

$$\frac{\partial \mathrm{t}}{\partial \mathrm{v_{exit}}} = \frac{1}{r_d}\left(\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}\right)^{-\frac{1}{2}} \left(v_{exit} - \sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}}\right)$$

$$\frac{\partial \mathrm{t}}{\partial \mathrm{l}} = \left(\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2 r_a r_d l}{r_a + r_d}\right)^{-1/2}$$

$$\frac{\partial \mathrm{t}}{\partial \bar{v}} = 0$$

Since $v_{enter}^2 - v_{exit}^2 \leq 2 r_d l$, we have

$$v_{enter}^2 \leq \frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2 r_a r_d l}{r_a + r_d}$$

i.e.

$$v_{enter} \leq \sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2 r_a r_d l}{r_a + r_d}}$$

which implies $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{enter}}} \leq 0$. And $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{enter}}} = 0$ only if $v_{enter}^2 - v_{exit}^2 = 2 r_d l$.

Similarly from $v_{exit}^2 - v_{enter}^2 \leq 2 r_a l$, we have

$$v_{exit} \leq \sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2 r_a r_d l}{r_a + r_d}}$$

which implies $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{exit}}} \leq 0$. And $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{exit}}} = 0$ only if $v_{exit}^2 - v_{enter}^2 = 2 r_a l$. And obviously in this case we have $\frac{\partial \mathrm{t}}{\partial \mathrm{l}} > 0$ and $\frac{\partial \mathrm{t}}{\partial \bar{v}} \leq 0$.

2. If $v_{exit} \leq \bar{v}$, $v_{enter} \leq \bar{v}$, $v_{exit}^2 - v_{enter}^2 \leq 2 r_a l$, $v_{enter}^2 - v_{exit}^2 \leq 2 r_d l$, $\sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2 r_a r_d l}{r_a + r_d}} > \bar{v}$

$$\mathrm{t}(v_{enter}, v_{exit}, l, \bar{v}) = \frac{\bar{v} - v_{enter}}{r_a} + \frac{\bar{v} - v_{exit}}{r_d} + \frac{1}{\bar{v}}\left(l - \frac{\bar{v}^2 - v_{enter}^2}{2 r_a} - \frac{\bar{v}^2 - v_{exit}^2}{2 r_d}\right)$$

$$\frac{\partial \mathrm{t}}{\partial \mathrm{v_{enter}}} = \frac{v_{enter} - \bar{v}}{r_a * \bar{v}}$$

$$\frac{\partial \mathrm{t}}{\partial \mathrm{v_{exit}}} = \frac{v_{exit} - \bar{v}}{r_d * \bar{v}}$$

$$\frac{\partial \mathrm{t}}{\partial \mathrm{l}} = \frac{1}{\bar{v}}$$

$$\frac{\partial \mathrm{t}}{\partial \bar{v}} = \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right) - \frac{1}{\bar{v}^2}\left(l + \frac{v_{enter}^2}{2r_a} + \frac{v_{exit}^2}{2r_d}\right)$$

We have $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{enter}}} \leq 0$. And $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{enter}}} = 0$ only if $v_{enter} = \bar{v}$. Similarly we have $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{exit}}} \leq 0$. And $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{exit}}} = 0$ only if $v_{exit} = \bar{v}$. And obviously we have $\frac{\partial \mathrm{t}}{\partial \mathrm{l}} > 0$.

From $\sqrt{\frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}} > \bar{v}$, we have

$$\frac{\partial \mathrm{t}}{\partial \bar{v}} = \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\frac{1}{\bar{v}^2}\left(\bar{v}^2 - \frac{r_a v_{exit}^2 + r_d v_{enter}^2 + 2r_a r_d l}{r_a + r_d}\right) < 0$$

Since $t(v_{enter}, v_{exit}, l, \bar{v})$ is a continuous function when it is finite, and from both cases we have $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{enter}}} \leq 0$ ($\frac{\partial \mathrm{t}}{\partial \mathrm{v_{enter}}} = 0$ only at finite points), $\frac{\partial \mathrm{t}}{\partial \mathrm{v_{exit}}} \leq 0$ ($\frac{\partial \mathrm{t}}{\partial \mathrm{v_{exit}}} = 0$ only at finite points), $\frac{\partial \mathrm{t}}{\partial \mathrm{l}} > 0$ and $\frac{\partial \mathrm{t}}{\partial \bar{v}} \leq 0$ Therefore $t(v_{enter}, v_{exit}, l, \bar{v})$ is a strictly decreasing function of v_{enter} and v_{exit},a strictly increasing function of l and a non increasing function of $\bar{v}$.□

Proof of Corollary 1: From Property 1, we know $v_{exit}^2 - v_{enter}^2 \leq 2r_a l$ and $v_{enter}^2 - v_{exit}^2 \leq 2r_d l$.

If $v_{enter} \geq v_{exit}$, Let $l' = \frac{v_{enter}^2 - v_{exit}^2}{2r_d} \leq l$. From Lemma 1 we have

$$t(v_{enter}, v_{exit}, l, \bar{v}) \geq t(v_{enter}, v_{exit}, l', \bar{v}) = \frac{v_{enter} - v_{exit}}{r_d}$$

We also have $t(v_{enter}, v_{exit}, l, \bar{v}) > 0 \geq \frac{v_{exit} - v_{enter}}{r_a}$. So the claim is true.

If $v_{exit} \geq v_{enter}$, Let $l' = \frac{v_{exit}^2 - v_{enter}^2}{2r_a} \leq l$. From Lemma 1 we have

$$t(v_{enter}, v_{exit}, l, \bar{v}) \geq t(v_{enter}, v_{exit}, l', \bar{v}) = \frac{v_{exit} - v_{enter}}{r_a}$$

We also have $t(v_{enter}, v_{exit}, l, \bar{v}) > 0 \geq \frac{v_{enter}-v_{exit}}{r_d}$. So the claim is true. □

Proof of Lemma 2: We first prove (20) and (21) by induction.

First we consider the scenario when $m = 1$. Since $T(0, v_e, p_1) < +\infty$, from Property 1 we have $v_e^2 \leq 2r_a l_1$. So $v^2 \leq v_e^2 \leq 2r_a l_1$, which implies $T(0, v, p_1) < +\infty$. And $v_e^2 \leq 2r_a l_1 \leq 2r_a l_1'$, which implies $T(0, v_e, p_1') < +\infty$. From Lemma 1 we know $T(0, v_e, p_1) < T(0, v, p_1)$ and $T(0, v_e, p_1) = t(0, v_e, l_1, \bar{v}_1) < t(0, v_e, l_1', \bar{v}_1) = T(0, v_e, p_1')$. Now consider a special scenario when $l_1' = l_1 + \frac{v_e^2 - v^2}{2*r_d}$. Let the train first travel through l_1 distance from velocity 0 to velocity v_e, which takes $T(0, v_e, p_1)$, and for the rest $\frac{v_e^2-v^2}{2*r_d}$ decelerate from velocity v_e to v, which takes $\frac{v_e - v}{r_d}$. $T(0, v, p_1')$ is the optimal travel time starting from velocity 0 to v with distance l_1', so we have

$$T(0, v, p_1') \leq T(0, v_e, p_1) + \frac{v_e - v}{r_d}$$

Left hand side is the optimal travel time while the right hand side is the travel time for one candidate solution. So $T(0, v_e, p_1) < T(0, v_e, p_1') \leq T(0, v_e, p_1) + \frac{v_e - v}{r_d}$. Therefore (20) and (21) are valid when m=1.

Assume the claim is true for m = k. We consider the scenario when $m = k + 1$. Since $T(0, v_e, p_{k+1}) < +\infty$, the corresponding problem (1)-(4) for this instance is feasible. Let the optimal solution for the corresponding problem (1)-(4) be v_i^* $(i = 0, \dots, k+1)$. Since $v_e^2 - v_k^{*2} \leq 2r_a l_{k+1} < 2r_a l_{k+1}'$ and $v_k^{*2} - v_e^2 \leq 2r_d l_{k+1} < 2r_d l_{k+1}'$, we have $t(v_k^*, v_e, l_{k+1}', \bar{v}_{k+1}) < +\infty$. So

$$T(0, v_e, p_{k+1}') \leq T(0, v_k^*, p_k) + t(v_k^*, v_e, l_{k+1}', \bar{v}_{k+1}) < +\infty.$$

After changing the length of the $(k+1)^{th}$ segment to l_{k+1}' ($l_{k+1}' > l_{k+1}$), let the optimal solution for the corresponding problem (1)-(4) be $v_i'^*$ $(i = 0, \dots, k+1)$. Then

$$T(0, v_e, p_{k+1}) = T(0, v_k^*, p_k) + t(v_k^*, v_e, l_{k+1}, \bar{v}_{k+1})$$

$$T(0, v_e, p_{k+1}') = T(0, v_k'^*, p_k) + t(v_k'^*, v_e, l_{k+1}', \bar{v}_{k+1})$$

If $v_k^* \geq v_k'^*$, $T(0, v_k^*, p_k) \leq T(0, v_k'^*, p_k)$ holds due to the assumption that (20) is true for m = k. Also

$$t(v_k^*, v_e, l_{k+1}, \bar{v}_{k+1}) < t(v_k^*, v_e, l'_{k+1}, \bar{v}_{k+1}) \leq t(v_k'^*, v_e, l'_{k+1}, \bar{v}_{k+1})$$

holds due to Lemma 1. So $T(0, v_e, p_{k+1}) < T(0, v_e, p_{k+1})$.

If $v_k^* < v_k'^*$, $T(0, v_k'^*, p_k) < T(0, v_k^*, p_k) < T(0, v_k'^*, p_k) + \frac{v_k'^* - v_k^*}{r_d}$ holds due to the assumption that (20) is true for m = k.

Also since $T(0, v_k^*, p_k) < +\infty$, $\forall v \in (v_k^*, v_k'^*)$ we have $T(0, v, p_k) < T(0, v_k^*, p_k) < +\infty$. From Proposition 1 we know v_k^* is the largest velocity among all the feasible velocities at the end of the k^{th} segment, which implies v is not a feasible velocity i.e. $t(v, v_e, l_{k+1}, \bar{v}_{k+1}) = +\infty$. From $t(v_k^*, v_e, l_{k+1}, \bar{v}_{k+1}) < +\infty$ and $(v, v_e, l_{k+1}, \bar{v}_{k+1}) = +\infty \forall v \in (v_k^*, v_k'^*)$, from Property 1 we know

$$v_k^{*2} - v_e^2 = 2r_d l_{k+1}$$

$$t(v_k^*, v_e, l_{k+1}, \bar{v}_{k+1}) = \frac{v_k^* - v_e}{r_d}$$

Also from Corollary 1, we have

$$t(v_k'^*, v_e, l'_{k+1}, \bar{v}_{k+1}) \geq \frac{v_k'^* - v_e}{r_d} = t(v_k^*, v_e, l_{k+1}, \bar{v}_{k+1}) + \frac{v_k'^* - v_k^*}{r_d}$$

i.e. $t(v_k^*, v_e, l_{k+1}, \bar{v}_{k+1}) \leq t(v_k'^*, v_e, l'_{k+1}, \bar{v}_{k+1}) - \frac{v_k'^* - v_k^*}{r_d}$. So

$$\begin{aligned} T(0, v_e, p_{k+1}) &= T(0, v_k^*, p_k) + t(v_k^*, v_e, l_{k+1}, \bar{v}_{k+1}) \\ &< T(0, v_k'^*, p_k) + \frac{v_k'^* - v_k^*}{r_d} + t(v_k'^*, v_e, l'_{k+1}, \bar{v}_{k+1}) - \frac{v_k'^* - v_k^*}{r_d} \\ &= T(0, v_k'^*, p_k) + t(v_k'^*, v_e, l'_{k+1}, \bar{v}_{k+1}) \\ &= T(0, v_e, p'_{k+1}) \end{aligned}$$

$\forall v \in [0, v_e]$, take $l'_{k+1} = l_{k+1} + \frac{v_e^2 - v^2}{2r_d}$. We can break the $(k+1)^{th}$ segment into two pieces such that the length of the first piece is l_{k+1} and that the length for the second piece is $\frac{v_e^2 - v^2}{2r_d}$. So the path from the beginning of the first segment to the break point at the $(k+1)^{th}$ segment is exactly the same as the original path. Due to the optimality of $T(0, v, p'_{k+1})$, we have

$$T(0, v, p'_{k+1}) \leq T(0, v_e, p_{k+1}) + \frac{v_e - v}{r_d}$$

which implies

$$T(0, v, p_{k+1}) < T(0, v, p'_{k+1}) \leq T(0, v_e, p_{k+1}) + \frac{v_e - v}{r_d}$$

Let v_k' satisfy $T(0, v, p_{k+1}) = T(0, v_k', p_k) + t(v_k', v, l_{k+1}, \bar{v}_{k+1})$, i.e. v_k' is the optimal velocity at the end of the k^{th} segment when the train travels from the beginning of the first segment at velocity 0 to the end of the $(k+1)^{th}$ segment at velocity v.

If $v_k' > v_k^*$, then $v_e^2 - v_k'^2 \le v_e^2 - v_k^{*2} \le 2r_a l_{k+1}$ and $v_k'^2 - v_e^2 \le v_k'^2 - v^2 \le 2r_d l_{k+1}$, which implies $t(v_k', v_e, l_{k+1}, \bar{v}_{k+1}) < +\infty$. So v_k' is a feasible velocity at the end of the k^{th} segment. However, from Proposition 1 we know $v_k^* \ge v_k'$, which is a contradiction. So $v_k' \le v_k^*$, and we have

$$\begin{aligned} T(0, v_e, p_{k+1}) &= T(0, v_k^*, p_k) + t(v_k^*, v_e, l_{k+1}, \bar{v}_{k+1}) \\ &< T(0, v_k', p_k) + t(v_k', v, l_{k+1}, \bar{v}_{k+1}) \\ &= T_{k+1}(0, v) \end{aligned}$$

Therefore (20) and (21) are valid when m = k+1.

In summary, $\forall m \in N^+$

$$T(0, v_e, p_m) < T(0, v, p_m) < T(0, v_e, p_m) + \frac{v_e - v}{r_d}$$

$$T(0, v_e, p_m) < T(0, v_e, p_m') < +\infty$$

Using the same argument we can have that if $T(v_s, 0, q_1) < +\infty$, then $\forall v \in [0, v_s)$ we have

$$T(v_s, 0, q_1) < T(v, 0, q_1) < T(v_s, 0, q_1) + \frac{v_s - v}{r_a}$$

$$T(v_s, 0, q_1) < T(v_s, 0, q_1') < +\infty \qquad \square$$

Proof of Lemma 3: Given $t(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2}) < +\infty$ from Property 1 we have $t(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1}) < +\infty$. (this statement is shown to be valid in the proof of Proposition 5). So the definition of $\Delta(v_{i-1}, v_i)$ is valid.

We first consider the scenario when $\sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i2}}{r_a + r_d}} > \bar{v}_{i2}$ i.e., the following holds.

$$\left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\bar{v}_{i2}^2 < \frac{v_{i-1}^2}{2r_a} + \frac{v_i^2}{2r_d} + l_{i2} \tag{25}$$

$$t(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2}) = \frac{\bar{v} - v_{i-1}}{r_a} + \frac{\bar{v} - v_i}{r_d} + \frac{1}{\bar{v}}(l - \frac{\bar{v}^2 - v_{i-1}^2}{2r_a} - \frac{\bar{v}^2 - v_i^2}{2r_d})$$

$$\frac{dt(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2})}{dv_{i-1}} = \frac{v_{i-1}}{r_a \bar{v}_{i2}} - \frac{1}{r_a}$$

Also if $\sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i1}}{r_a + r_d}} > \bar{v}_{i1}$,

$$\frac{dt(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1})}{dv_{i-1}} = \frac{v_{i-1}}{r_a \bar{v}_{i1}} - \frac{1}{r_a}$$

$$\frac{d\Delta(v_{i-1}, v_i)}{dv_{i-1}} = \frac{v_{i-1}}{r_a \bar{v}_{i1} \bar{v}_{i2}} (\bar{v}_{i2} - \bar{v}_{i1}) < 0$$

where the last inequality is from Condition 2.

If $\sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i1}}{r_a + r_d}} \leq \bar{v}_{i1}$,

$$\frac{dt(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1})}{dv_{i-1}} = \frac{v_{i-1}}{r_a} \sqrt{\frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i1}}} - \frac{1}{r_a}$$

$$\frac{d\Delta(v_{i-1}, v_i)}{dv_{i-1}} = \frac{v_{i-1}}{r_a} \left(\sqrt{\frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i1}}} - \frac{1}{\bar{v}_{i2}} \right) < 0$$

since

$$\frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i1}} - \frac{1}{\bar{v}_{i2}^2}$$

$$= \frac{2 r_a r_d}{\bar{v}_{i2}^2 (r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i1})} \left(\left(\frac{1}{2r_a} + \frac{1}{2r_d} \right) \bar{v}_{i2}^2 - \left(\frac{v_{i-1}^2}{2r_a} + \frac{v_i^2}{2r_d} + l_{i1} \right) \right)$$

$$< \frac{2 r_a r_d}{\bar{v}_{i2}^2 (r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i1})} \left(\left(\frac{1}{2r_a} + \frac{1}{2r_d} \right) \bar{v}_{i2}^2 - \left(\frac{v_{i-1}^2}{2r_a} + \frac{v_i^2}{2r_d} + l_{i2} \right) \right) < 0$$

where the last inequality is from (25).

So we have $\frac{d\Delta(v_{i-1}, v_i)}{dv_{i-1}} < 0$ when $\sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i2}}{r_a + r_d}} > \bar{v}_{i2}$ holds.

Then we consider the scenario when $\sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i2}}{r_a + r_d}} \leq \bar{v}_{i2}$ i.e. where the following holds.

$$\left(\frac{1}{2r_a} + \frac{1}{2r_d} \right) \bar{v}_{i2}^2 \geq \frac{v_{i-1}^2}{2r_a} + \frac{v_i^2}{2r_d} + l_{i2} \tag{26}$$

$$\mathrm{t}(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2}) = -\frac{v_{i-1}}{r_a} - \frac{v_i}{r_d} + (\frac{1}{r_a} + \frac{1}{r_d}) \sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i2}}{r_a + r_d}}$$

$$\frac{dt(v_{i-1}, v_i, l_{i2}, \bar{v}_{i2})}{dv_{i-1}} = \frac{v_{i-1}}{r_a} \sqrt{\frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2 r_a r_d l_{i2}}} - \frac{1}{r_a}$$

Therefore if $\sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i1}}{r_a + r_d}} \leq \bar{v}_{i1}$,

$$\frac{dt(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1})}{dv_{i-1}} = \frac{v_{i-1}}{r_a}\sqrt{\frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i1}}} - \frac{1}{r_a}$$

$$\frac{d\Delta(v_{i-1}, v_i)}{dv_{i-1}} = \frac{v_{i-1}}{r_a}\left(\sqrt{\frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i1}}} - \sqrt{\frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i2}}}\right) < 0$$

since $l_{i1} > l_{i2}$.

If $\sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i1}}{r_a + r_d}} > \bar{v}_{i1}$,

$$\frac{dt(v_{i-1}, v_i, l_{i1}, \bar{v}_{i1})}{dv_{i-1}} = \frac{v_{i-1}}{r_a \bar{v}_{i1}} - \frac{1}{r_a}$$

$$\frac{d\Delta(v_{i-1}, v_i)}{dv_{i-1}} = \frac{v_{i-1}}{r_a}\left(\frac{1}{\bar{v}_{i1}} - \sqrt{\frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i2}}}\right) < 0$$

since

$$\frac{1}{\bar{v}_{i1}^2} - \frac{r_a + r_d}{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i2}}$$

$$= \frac{2r_a r_d}{\bar{v}_{i1}^2(r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i2})}\left(\left(\frac{v_{i-1}^2}{2r_a} + \frac{v_i^2}{2r_d} + l_{i2}\right) - \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\bar{v}_{i1}^2\right)$$

$$< \frac{2r_a r_d}{\bar{v}_{i1}^2(r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i2})}\left(\left(\frac{v_{i-1}^2}{2r_a} + \frac{v_i^2}{2r_d} + l_{i2}\right) - \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\bar{v}_{i2}^2\right) \leq 0$$

where the last inequality is due to (26).

So we have $\frac{d\Delta(v_{i-1}, v_i)}{dv_{i-1}} < 0$ when $\sqrt{\frac{r_a v_i^2 + r_d v_{i-1}^2 + 2r_a r_d l_{i2}}{r_a + r_d}} \leq \bar{v}_{i2}$ holds. Therefore $\Delta(v_{i-1}, v_i)$ is a strictly decreasing function of v_{i-1}. Similarly we can also show $\frac{d\Delta(v_{i-1}, v_i)}{dv_i} < 0$ and that $\Delta(v_{i-1}, v_i)$ is a strictly decreasing function of v_i. □

Proof of Proposition 8: Suppose it is not true, i.e. $\sqrt{\frac{2r_a r_d l_{i2}}{r_a + r_d}} \leq \bar{v}_{i2}$, then

$$t(0,0, l_{i2}, \bar{v}_{i2}) = (\frac{1}{r_a} + \frac{1}{r_d})\sqrt{\frac{2r_a r_d l_{i2}}{r_a + r_d}}$$

which is not dependent on $\bar{v}_{i2}$. So $t(0,0,l_{i2},\bar{v}_{i2}) = t(0,0,l_{i2},\bar{v}_{i1})$ since $\sqrt{\frac{2r_a r_d l_{i2}}{r_a+r_d}} \leq \bar{v}_{i2} \leq \bar{v}_{i1}$ holds. From Lemma 1 and Condition 2

$$t(0,0,l_{i2},\bar{v}_{i2}) = t(0,0,l_{i2},\bar{v}_{i1}) < t(0,0,l_{i1},\bar{v}_{i1})$$

which contradicts Condition 5. So we have $l_{i2} > \left(\frac{1}{2r_a} + \frac{1}{2r_d}\right)\bar{v}_{i2}^2$. □

Proof of Lemma 4: Take $l' = \frac{v'^2_{enter} - v^2_{enter}}{2r_a} + l + \frac{v'^2_{exit} - v^2_{exit}}{r_d}$. Then

$$t(v_{enter}, v_{exit}, l, \bar{v}) < t(v_{enter}, v_{exit}, l', \bar{v})$$

$$\leq t(v'_{enter}, v'_{exit}, l, \bar{v}) + \frac{v'_{enter} - v_{enter}}{r_a} + \frac{v'_{exit} - v_{exit}}{r_d}$$

The first inequality is due to Lemma 1 and the second inequality is based on the optimality of $t(v_{enter}, v_{exit}, l', \bar{v})$. □