Title: Integrating Management of Truck and Rail Systems in Los Angeles

Project Number: 14-3.1a
Year: 2013-2017

FINAL REPORT
September 2018

Principal Investigator
Maged Dessouky

PhD Students
Lunce Fu
Shichun Hu

MetroFreight Center of Excellence
Viterbi School of Engineering
University of Southern California
Los Angeles, California, USA

Table of Contents

Title: Integrating Management of Truck and Rail Systems in LA

Abstract

This project establishes models to optimize the balance of freight demand across rail and truck modes. In real life situations, trains often travel at different speeds (i.e. passenger trains and freight trains share the same rail network). This incurs train delay whereby reducing the efficiency of the rail network. To provide a solution for this problem, we develop heuristic algorithms to improve conventional dispatching rules to reduce the average train delay. Then we build a control model and provide the solution procedure to adapt a dynamic headway concept inspired by new signaling technology like Positive Train Control (PTC). Rail network data of the Southern California region is collected to perform a detailed simulation analysis. The simulation results show significant improvement of network efficiency brought by our model and algorithms: as high as 21% reduction in average train delay with our best dispatching policy while with the dynamic headway control model, the average train delay is reduced by 40%. The railway network is therefore shown to have the potential to increase throughput capacity by 20%.

Introduction

Railway has always been an effective mode to transport both people and goods. Freight trains are about four times more fuel efficient than trucks and passenger trains are popular because they can comfortably transport people to their destinations on time at a lower cost while reducing greenhouse gas emissions. Serving as a major part of the transportation service in the United States, railway moves about 40% of freight measured in ton-miles in 2012 [1], generating approximately $70 billion of revenue in 2016. 41.6% (by units) of America's international freight is transported through railway. According to the statistics given by the Federal Railroad Administration, total U.S. freight shipments will see a 41% rise from an estimated 18.0 billion tons in 2015 to 25.3 billion tons in 2045 [2]. However, considering the difficulties and high cost of extending the current railway infrastructure, this increase in rail demand will no doubt bring challenges to the current railway network, therefore requiring better rail traffic operation to mitigate the growing freight demand.

In addition, from the passengers' perspective, they desire quick and punctual transportation services. With the advancement of technology, passenger trains are able to travel at a much faster speed compared to freight trains, but the limitation of the rail infrastructure makes it more practical and cost effective to allow passenger trains to share some portions of the railway tracks with slower freight trains. If a faster passenger train catches up with a freight train on the railway track and there are no crossover junctions, the nature of the railway transportation determines that the passenger train has to follow the freight train at the speed of the freight train while keeping a safety headway from the freight train.

To tackle the issues mentioned above where it is often the case that fast passenger trains share the same rail network with slow freight trains, we need dispatching rules to efficiently guide both freight and passenger trains through the network so that train delays are minimized and the railway network capacity is improved. New communication technologies have the potential to improve railway operations, especially through more efficient train scheduling and dispatching. Positive Train Control (PTC) is introduced as a system of monitoring and controlling the movement of trains to increase security by reducing human operation. With PTC, trains can communicate with other trains to share information.

Previously trains are 'blind' and controlled by the signals which are operated by experienced human dispatchers. With PTC, each train can have information of trains near it ('locally') and even trains far away from it ('globally'). Utilizing this new technology, a train's velocity can be monitored and controlled in real time, the concept of dynamic headway (the track segment between two consecutive signals) is introduced to increase railway network efficiency. Moreover, effective communication between the trains will facilitate better dispatching controls for passenger trains as well as freight trains to minimize train delays.

Project Objective

The purpose of this research is to further the state-of-the-art of the train scheduling and routing problem taking into consideration the new capabilities that the newly introduced technologies such as PTC provide. Specifically, the contribution of this research is (1) we develop three heuristic dispatching rules considering multiple train speeds and blocking time that effectively reduce average train delay, (2) we establish a simulation framework to represent dynamic headway, and develop an algorithm to determine the optimal velocities given the headway distance and the speed, and (3) we use these models to estimate the additional amount of freight that the rail system can handle if the developed control rules are used to control rail movement in Southern California.

Project Description

Mu and Dessouky [6] study the dispatching policies for a double-track segment when there are only two train speeds. They developed an analytical model to predict the expected delay with a switchable policy. Train length and safety headways were both set to zero to keep the analytical model tractable. For better accuracy, we extend the study for the double-track to the case where trains are travelling at multiple speeds while considering train length and safety headways in our analysis.

In this project, we first develop three heuristic dispatching rules for the control of trains travelling on double-track railway segments with heterogeneous traffic and experimentally compare the performance of the heuristic rules using simulation. The first switchable policy is to switch the fast train, if it has potential delay on its designated track, to the opposite direction track if that track is empty. The second policy has a smarter condition to dispatch the trains by considering the speed of the attempting switching train. The last policy adds the consideration of blocking time in the dispatching rule. That is, fast trains are switched to the opposite direction track if they do not extend the current busy period on the opposite direction track for a tolerable length of time. We also investigate how these dispatching rules will improve if crossovers are placed in the segment. Since all three heuristic dispatching rules are based on fixed headway distance, we then advance the current train scheduling and routing problem through dynamic headway control facilitated by new technologies such as PTC.

Research Approach

In this section, we first introduce the three heuristic dispatching rules we developed and then describe our dynamic headway model and heuristics for solving the problem.

Descriptions of dispatching policies

Figure 1 shows a typical double-track railway segment between two major intersections. The length of the track segment is denoted by D. There are multiple types of trains travelling on the track segments. Each type of train is identified by its speed and we assume that the arrival of each train type at each end of the double-track segment is an independent Poisson Process. The upper and lower tracks of the segment can be travelled in both directions. The free running time of the train is defined to be the minimum traveling time of the train assuming there is no other traffic in the network. The delay time of the train can be calculated as:

Delay = Completion time – Arrival time at the segment – Free running time
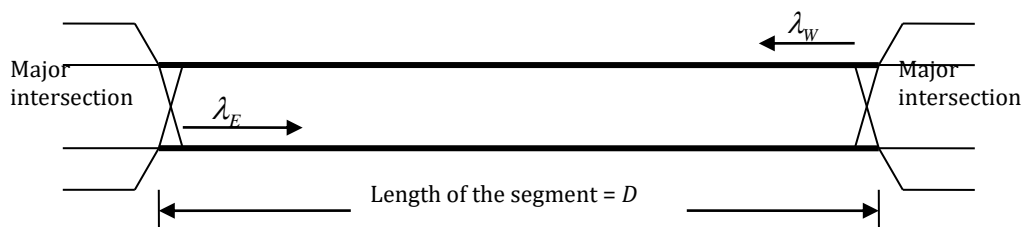

Figure 1. Double-track railroad segment

A delay can occur (1) when a faster train catches up with a slower train traveling in the same direction so that the faster train has to travel at the speed of the slower train, while keeping the required headway, or (2) when a train arrives at the track segment and it has to wait for the track to be cleared from being occupied by trains traveling in the reverse direction.

Probably the easiest dispatching policy for a double-track segment is to dedicate each track to traffic in one direction (i.e., all eastbound trains travel on the lower track while all westbound trains travel on the upper track). We refer to this policy as a dedicated policy. The drawback of the dedicated policy is that it can be likely for a fast train to catch up with a slower train on its dedicated track. If the fast train catches up with a slower train, it has to keep a safety distance between the slower train and travel at the speed of the slower train. Thus, the fast train can experience a significant delay if there is a large difference in train speeds.

Next we introduce three dispatching rules that allow trains to travel in either track segment. Without loss of generality, for the following dispatching rules, let the lower track be the designated track for trains traveling eastbound and let the upper track be the designated track for trains travelling westbound. We refer to the three dispatching rules as Switchable2-I, Switchable2-II and Switchable2-III policy. The Switchable2-I policy considers the potential delay of the arriving train when deciding to

3

switch the train. Besides the potential delay, the Switchable2-II policy also considers the speed of the arriving train when deciding to switch the train. The faster the speed is, the more likely the train will switch. To increase the switching frequency, the Switchable2-III policy switches the train to the other track in cases where the other track is not completely empty of trains. A more detailed description of these heuristic rules is in [7].

<u>Switchable2-I policy</u>

1. Upon arrival, if the designated track is occupied by trains travelling in the opposite direction, the arriving train waits for the opposing moving trains to finish traveling on the segment before proceeding on its designated track.

2. Upon arrival, if the designated track is not occupied by trains travelling in the opposite direction, and if the potential length of delay of the arriving train traveling on its designated track is less than ω, the arriving train will start traveling on its designated track.

3. Upon arrival, if the designated track is not occupied by trains travelling in the opposite direction, and if the potential length of delay of the arriving train traveling on its designated track is greater than ω, the arriving train will attempt to switch to the other track. The arriving train will use the reverse direction track only if it is empty. If the reverse direction track is occupied by trains travelling in either direction, the arriving train will use its designated track.

The Swtichable2-I policy is similar to the switchable policy described in Mu and Dessouky [6], where they derive analytical equations to measure the delay assuming a no headway requirement since they assume the trains are infinitesimally small. The Switchable2-I policy will attempt to switch an arriving train if it will catch up with a slower train on its designated track and its potential delay is greater than ω. The optimal value of ω ranges from 0 to the time difference between the free running times of the slowest and fastest train. If a fast train catches up with a slower train near the end of its designated track, the potential delay of the fast train might not be significant enough for the policy to switch the fast train to the other track, since usage of the other track might block the traffic in the other direction. When a train attempts to switch, the policy only allows it to switch when the reverse direction track is empty. If the reverse direction track is occupied by another switched train which is traveling in the same direction as the train attempting to switch, the policy prohibits the train from switching so not to extend the reverse direction busy period for the other track.

With many different train speeds, even a slow train can catch up with another slower train and experience delay on its designated track. If we only consider the potential delay on the designated track as the criterion to switch the train, we might tend to switch some slow trains. The switched slow trains will block the other track for a long time, which is not desired for the traffic in the other direction. Intuitively, if both a relatively fast and a relatively slow train have the same potential delay on the designated track, the relatively fast train should be switched because the fast train can finish traveling on the reverse direction track in a shorter amount of time. Thus a higher potential delay and

a higher train speed should lead to a higher chance to switch. Let D_p denote the potential length of delay an arriving train will experience on its designated track. Let S_{ar} denote the speed of the arriving train. In the Switchable2-II policy, instead of having $D_p \geq \omega$ as the criterion to switch the arriving train, a more complicated criterion $\alpha D_p + \beta S_{ar} \geq \delta$ is used, where α, β and δ are parameters. A good assignment of the values of α, β and δ can be obtained by discretizing them and enumerating the parameters in multiple simulation runs.

Switchable2-II policy

1. Upon arrival, if the designated track is occupied by trains travelling in the opposite direction, the arriving train waits for the opposing moving trains to finish traveling on the segment before proceeding on its designated track.

2. Upon arrival, if the designated track is not occupied by trains travelling in the opposite direction, and if $\alpha D_p + \beta S_{ar} < \delta$, the arriving train will start traveling on its designated track.

3. Upon arrival, if the designated track is not occupied by trains travelling in the opposite direction, and if $\alpha D_p + \beta S_{ar} \geq \delta$, the arriving train will attempt to switch to the other track. The arriving train will use the reverse direction track only if it is empty. If the reverse direction track is occupied by trains travelling in either direction, the arriving train will use its designated track.

Switchable2-III is based on the Switchable2-II policy. For the first two policies, the attempted switching trains will switch if the reverse direction track is empty. The idea being that if another train is allowed to travel on the reverse direction track when one is already traveling on it may cause a significant amount of time that segment is blocked for a train traveling on its designated direction. However, in the case of multiple speeds, if the reverse direction track is occupied by a switched train, it will do no significant harm if we switch another faster train to the reverse direction track, given the latter switched train can catch up with the former switched train. To extend this idea, in the Switchable2-III policy, if the reverse direction track is occupied by a switched train, a newly arriving train can switch to the reverse direction track only if the arriving train will extend the busy period on the reverse direction track by no longer than μ time units. A good value of μ can be found by discretizing it and enumerating in multiple simulation experiments.

Switchable2-III policy

1. Upon arrival, if the designated track is occupied by trains travelling in the opposite direction, the arriving train waits for the opposing moving trains to finish traveling on the segment before proceeding on its designated track.

2. Upon arrival, if the designated track is not occupied by trains travelling in the opposite direction, and if $\alpha D_p + \beta S_{ar} < \delta$, the arriving train will start traveling on its designated track.

3. Upon arrival, if the designated track is not occupied by trains travelling in the opposite direction, and if $\alpha D_p + \beta S_{ar} \geq \delta$, the arriving train will attempt to switch to the other track. The arriving train will use the reverse direction track if it is empty. If the reverse direction track is occupied by trains travelling in the same direction as the arriving train and if the arriving train extends the current busy period on the reverse direction by no longer than μ time units, the arriving train will switch to the reverse direction track. Otherwise, the arriving train travels on its designated track.

Now suppose the double-track segment has a crossover in the middle of the segment. The introduction of the crossover at the middle can significantly increase the effectiveness of the switchable policy. With the help of the crossover, trains can switch to the other track at the beginning of the track and then switch back in the middle. Also, trains can switch to the other track in the middle of the track. In both ways, the switched trains do not have to travel through the entire segment of the other track. Thus the double track segment could be better utilized. Next, we are going to describe a switchable policy (namely, Switchable2-w/cross) which is based on the Switchable2-III policy. Treating the crossover in the middle of the segment as a station connecting two halves of the segment, the Switchable2-w/cross policy dispatches trains almost the same as what the Switchable2-III policy will do for those two double-track segments connected together. The Switchable2-w/cross policy is designed to dispatch trains with multiple speeds on a double-track segment with crossovers in the middle. The description of the Switchable2-w/cross policy below focuses on the eastbound trains and uses the notations in Figure 4. Let D_p^{EB1} denote the potential delay of the arriving train on track segment EB1. Let D_p^{EB2} denote the potential delay of the train on track segment EB2 as it arrives at EB2.
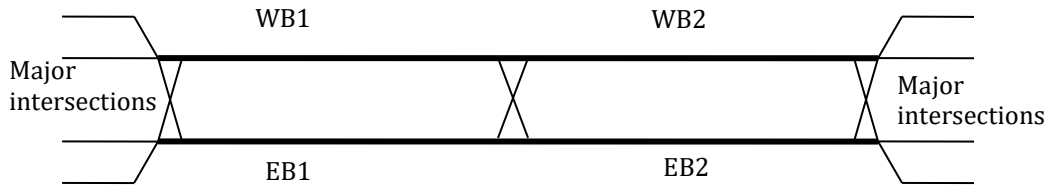

Figure 2. Double-track railroad segment with crossovers

Switchable2-w/cross policy

1. Upon arrival, if EB1 is occupied by trains travelling in the opposite direction, the arriving train waits for the opposing moving trains to finish traveling on the segment before proceeding on EB1.

2. Upon arrival, if EB1 is not occupied by trains travelling in the opposite direction, and if $\alpha_1 D_p^{EB1} + \beta_1 S_{ar} < \delta_1$, the arriving train will start traveling on EB1.

3. Upon arrival, if EB1 is not occupied by trains travelling in the opposite direction, and if $\alpha_1 D_p^{EB1} + \beta_1 S_{ar} \geq \delta_1$, the arriving train will attempt to switch to WB1. The arriving train will use WB1 if it is empty and if EB2 is not occupied by westbound trains. If WB1 is occupied by

6

eastbound trains and if the arriving train extends the current busy period on WB1 by no longer than μ_1 time units, the arriving train will switch to WB1. Otherwise, the arriving train travels on EB1.

4. When an eastbound train reaches the end of track EB1, if EB2 is occupied by trains travelling in the opposite direction, the train at the end of EB1 waits for the opposing moving trains to finish traveling on EB2 before proceeding on EB2. When an eastbound train reaches the end of track EB1, if EB2 is not occupied by trains travelling in the opposite direction, and if $\alpha_2 D_p^{EB2} + \beta_2 S_{ar} < \delta_2$, the eastbound train will start traveling on EB2. But if $\alpha_2 D_p^{EB2} + \beta_2 S_{ar} \geq \delta_2$, the eastbound train will attempt to switch to WB2. The eastbound train will use WB2 if it is empty. If WB2 is occupied by eastbound trains and if the arriving train extends the current busy period on WB2 by no longer than μ_2 time units, the train at the end of EB1 will switch to WB2. Otherwise, the train travels on EB2.

5. When an eastbound train reaches the end of track WB1, if $\alpha_2 D_p^{EB2} + \beta_2 S_{ar} < \delta_2$, the eastbound train will start traveling on EB2. But if $\alpha_2 D_p^{EB2} + \beta_2 S_{ar} \geq \delta_2$, the eastbound train will attempt to continue to WB2. The eastbound train will use WB2 if it is empty. If WB2 is occupied by eastbound trains and if the train at the end of WB1 extends the current busy period on WB2 by no longer than μ_2 time units, the train at the end of WB1 will continue to WB2. Otherwise, the train travels on EB2.

Dynamic headway model and solution

After developing dispatching policies dealing with different train speeds, we now extend fixed headway to dynamic headway and look for its potential to increase rail network capacity. We next describe the steps needed to form our dynamic headway control model and provide a solution procedure for the model.

The railway track is discretized into different segments. Segments are the smallest, indivisible units in this model. All points within one segment share one speed limit. Then several segments and/or junctions are grouped into one node to formulate a network G = (V, E), where V is the set of nodes or vertices and E is the set of arcs. Notice that each arc works as linkage between two nodes, and that it may or may not correspond to a junction. An example network construction for a portion of the railway network is given as follows:
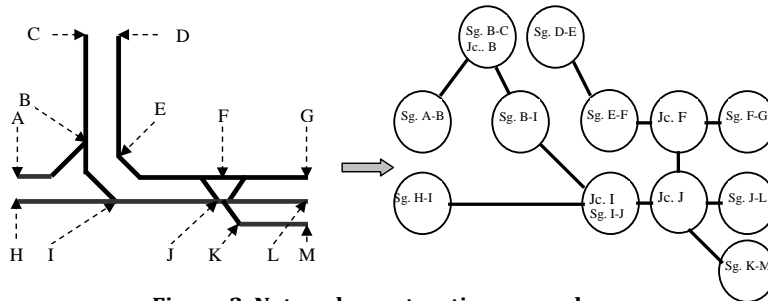

Figure 3. Network construction example

Then a train's movement through the railway network can be modelled as movement through the constructed network G. We set the capacity of each node to be one, i.e. there can be at most one train within each node at any time. Then the headway is modeled as all the available nodes between two consecutive trains. However the constraint of the previous models that each node should be long enough for trains to stop within each node makes the network representation not very efficient since the capacity for each node is set to one. In other words, the succeeding train cannot enter the next node until the preceding train leaves. Therefore the headway between two consecutive trains is at least a node's distance. Also notice that the headway is controlled by a node's distance for all trains, which implies that the standard models can only provide "fixed" headway. We develop a new reformulation for "dynamic" headway, i.e. the headway between two consecutive trains is dependent on the two trains' types and velocities.

Different from the previous model in [5], we do not constrain each node's length. The smaller each node's length is, the finer our discretization and approximation for the headway is. On the other hand, the number of nodes in the constructed network will increase as the nodes' length decrease. Therefore in practice we keep the nodes' length moderately small, for example a quarter of the nodes' length of the previous model. Also we do not enforce a train to stop within its current node. Otherwise the entering velocities must be small enough and therefore trains may never reach the speed limit. Therefore a certain number of nodes ahead of each train must be assigned to it before the train can enter the current node. We call these assigned nodes to a train as its dynamic headway. Dynamic headway should be long enough in order for the corresponding train to come to a full stop without collision.

The dynamic headway is categorized into two types. In the first scenario, there is no preceding train within the focal train's braking distance. Then the dynamic headway is no shorter than the braking distance. Let the focal train's velocity be v and its maximal deceleration rate be r_{d1}. Then the dynamic headway distance H_D is:

$$H_D \geq \frac{v^2}{2r_{d1}}$$

In the second scenario, the succeeding train's dynamic headway works as buffer between the two consecutive trains. Let the preceding train's velocity be μ, its maximal deceleration rate be r_{d2} and the response time for the preceding train be Δt_1. Then the dynamic headway distance H_D should be long enough to avoid collision, i.e.

$$H_D + \frac{\mu^2}{2r_{d2}} \geq \frac{v^2}{2r_{d1}} + \Delta t_1 * v$$

Thus,

$$H_D \geq \frac{v^2}{2r_{d1}} + \Delta t_1 * v - \frac{\mu^2}{2r_{d2}}$$

To take advantage of the PTC technology, each node length is small so that one train can occupy several nodes at the same time. Also each train can occupy some nodes ahead of it as a buffer between it and its preceding train. These nodes are defined to be "headway nodes". Since the headway nodes need to be determined by the trains' velocities and deceleration/acceleration rates, the number and the length of the headway nodes vary as the trains are traveling through the network. We name it as dynamic

headway control. The train scheduling problem for dynamic headway control seeks a path for each train, controls each train's velocity along the path and assigns headway nodes in an efficient way. We developed a mathematical formulation for this problem. However, due to the computational difficulty of solving for the optimal solution, we have developed efficient heuristics for solving the problem. A detailed model description can be found in [3].

In general, the procedure provides a dynamic control scheme for headway control. Decisions are made at certain points (at the beginning of each segment). We call them decision points. When a train reaches a decision point, the solution procedure is applied to obtain the velocity at the next decision point as well as the corresponding travel time. In summary, three decisions need to be made to guide a train through the rail network in our dynamic headway model:
1. Routing decision, i.e. selects the next headway node when facing multiple candidates
2. Headway decision, i.e. determines the number of new headway nodes a train needs to occupy
3. Velocity decision, i.e. calculates the velocity at the next decision point

For the routing decision, we assume a greedy routing algorithm which is the same as in [6]. That is, we route to the next headway node which has the highest speed limit. We next describe the headway decision, based on which the velocity decision is made.

Headway Decision

We use a simulation approach for modelling and solving the headway decision. Based on the headway, each train will be assigned several nodes ahead of it in the simulation model. When the train moves towards the next headway node, the simulation model determines whether new headway nodes are needed and assigns them to the train if necessary. Moreover, the exiting velocity of the next headway node needs to be calculated. In summary the simulation works as follows:
1. When a train enters into the first node of its schedule, it is assigned several nodes to serve as the headway between it and its preceding train.
2. When a train enters into the first headway node (nearest node to the train in the headway nodes), the routing algorithm will be called to determine the headway nodes if needed. Also the exiting velocity of the first headway node will be calculated. Then the train will be routed to the end of the first headway node according to the exiting velocity.
3. When the last headway node (farthest node to the train in the headway nodes) reaches the end node of the schedule, the train is routed to a full stop at the end node according to its minimal travel time.

A simple example in Figure 4 shows how the dynamic headway schema works for the same track configuration. Before entering Node 1, Train 1 has already pre-occupied Node 1 and Node 2. When it finally enters Node 1, our algorithm is called to decide whether Node 3 shall be added to Train 1's pre-occupied nodes. If Node 3 is added to Train 1's pre-occupied nodes, Train 1 will not need to decelerate when traveling within Node 1, thus reducing travel time compared to the fixed headway schema.
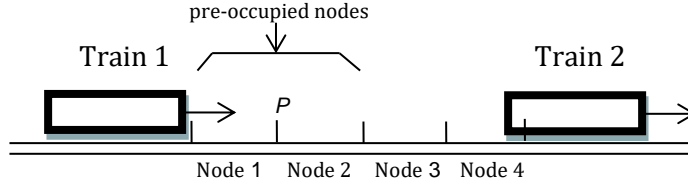
Figure 4. Dynamic headway model

Velocity Decision

The velocity decision depends on the headway decision. Since the headway distances can be categorized into two different types, we now show how the velocity is obtained.

In the scenario where the headway works as the braking distance which means no preceding trains exist, only the velocity at the end of the current node needs to be determined, and then the train travels through the current node. Given the current node n_0 and the current velocity v_0, to minimize the travel time within the current node, we want to maximize the velocity at the end of the current node. Let v_1^* be the optimal exiting speed of the train's head from node n_0, i.e. the optimal entering speed to node n_1. Suppose after making the headway decision, the headway nodes are $n_1, n_2, \dots, n_Q$, so the number of seized headway nodes is Q. Let l_i be the length and let $\bar{v}_i$ be the speed limit of the ith node of the seized nodes. Let the velocity at the beginning of node i be v_i ($i = 1, \dots, Q$) and let the velocity at the end of node n_Q be v_{Q+1}. Consider a sequence of problems indexed by β ($\beta = 1, \dots, Q$):

$$\max \quad v_\beta$$
$$\text{s.t.} \quad t(v_i, v_{i+1}, l_i, \bar{v}_i) < +\infty \quad \forall i = \beta, \dots, Q$$
$$v_i \leq \bar{v}_i \quad \quad \quad \quad \forall i = \beta, \dots, Q$$
$$v_{Q+1} = 0$$

The βth problem solves the maximal velocity at the beginning of node n_β so that the train can stop at the end of node n_Q. The first constraint simply means that the minimal travel time should be feasible for a train to travel through node i at a velocity of v_i at the beginning of the node to a velocity of v_{i+1} at the end of the node ($i = \beta, \dots, Q$).

Let $\tilde{v}_\beta$ be the optimal value for the βth problem ($\beta = 1, \dots, Q$). Intuitively, the above sequence of problems recursively solve for $\tilde{v}_1$ when the train is at the beginning of node n_1. The βth problem can take advantage of the optimal value obtained by the $(\beta + 1)^{th}$ problem instead of solving the βth problem explicitly.

In the other scenario the headway works as buffer distance between two successive trains. Again, let the current headway nodes be $n_1, n_2, \dots, n_Q$ and node n_{Q+1} is occupied by the preceding train which travels in the same direction. Let the current velocity of the preceding train be μ. Then we can obtain the potential maximal velocity $\tilde{v}_1$ at the beginning of node n_1 as follows:

$$\sum_{i=1}^{Q} l_i = \tilde{v}_1^2/(2r_{d1}) + \Delta t_1 * \tilde{v}_1 - \mu^2/(2r_{d2})$$

where Δt_1 is the response time for the succeeding train and, r_{d1} and r_{d2} are the deceleration rates for the succeeding and preceding trains respectively.

However, $\tilde{v}_1$ obtained by the equation above may not be accurate. When the succeeding train travels $\sum_{i=1}^{Q} l_i + \mu^2/(2r_{d2})$, it may stop within some node which is occupied by the preceding train. To obtain the maximal velocity v_1^*, let

$$\sum_{i=Q+1}^{R} l_i < \frac{\mu^2}{2r_{d2}} \le \sum_{i=Q+1}^{R+1} l_i$$

where nodes $n_{Q+1}, \ldots, n_R$ are occupied by the preceding train (some of them are physically occupied and others are headway nodes). Notice that R can be uniquely determined by the above inequalities, and therefore is dependent on μ. So the maximal velocity v_1^* at the beginning of node n_1 satisfies

$$\sum_{i=1}^{R} l_i = v_1^{*2}/(2r_{d1}) + \Delta t_1 * v_1^*$$

The details on how the velocity at each node is calculated at each node to satisfy the above constraints is presented in [3].

In summary, whenever a train enters into a node, two decisions will be made: how many new headway nodes are needed (headway decision) and the exiting speed for the current node (velocity decision). As a result, the train will seize these new headway nodes. And given the speed limit for the current node and the train's deceleration and acceleration rates, we can calculate the minimal traveling time within the current node. The detailed steps are:
1. A train enters into a node, which is a pre-occupied headway node.
2. Is a new headway node available? If yes, go to 3. If no, go to 5.
3. Can the train reach the speed limit outside the current node given the current headway nodes? If yes, go to 6. If no, go to 4.
4. Seize one new headway node. Go to 2.
5. Is the new headway node seized by a train traveling in the same direction? If yes, go to 7. If no, go to 6.
6. Compute the exiting speed which can make the train stop within the current headway, go to 8.
7. Compute the exiting speed which can avoid collision with the preceding train.
8. Route the train to the end of its current node, where the train reaches the exiting speed.

Analysis and Results

In this section, we first introduce the data sets and the parameters we use in our experiment and then present the experimental results of the heuristic dispatching rules we developed. Lastly, we present the impact of our dynamic headway control model.

Data

For the simulation of our heuristic dispatching rules, we build a model using Arena [4]. For the dispatching policy, a double-track segment that is eight miles long is used and the safety headway between two consecutive trains is set to be one mile. The maximum speeds and lengths for the different types of trains are listed in Table 1. The reason the faster trains are shorter is that they are more likely to be passenger trains instead of freight trains.

Table 1. Maximum train speeds and train lengths

Train Speed (mile/hour)	Train Length (feet)
50	5,000
70	6,000
90	1,000
120	1,000
140	1,000

For the dynamic headway control analysis, we used the actual railway network in the Los Angeles area that was collected in year 1 of the project. The chosen part is from Downtown Los Angeles to Pomona. Figure 5 illustrates the mileage between the stations. At the intermediate station (El Monte), there is a crossing for trains traveling in the north/south direction. Notice that this figure only provides mileage information and does not show the actual railway trackage configuration. The railway trackage configuration in this area consists of single-, double- and triple- tracks.
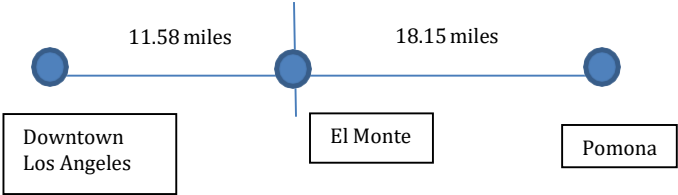
Figure 5. Mileage information

Also, two types of trains (freight train and passenger train) are tested on this area. The detailed information about these two types of trains can be found in Table 2.

Table 2. Characteristics of the passenger and freight trains

	Length (feet)	Max speed (feet/min)	Acceleration rate (feet/min**2)	Deceleration rate (feet/min**2)
Freight Train	6000	6160	1584	1584
Passenger Train	1000	6952	2112	2112

Analysis of the dispatching policies

For the Switchable2-I policy, the optimal value of ω needs to be determined for each scenario. The possible values for ω range from 0 to $D / S_{sl} - D / S_{fa}$, where S_{sl} and S_{fa} denote the slowest and fastest train speeds possible on the track, respectively. In the numerical experiment, we discretize the value of ω into steps of 0.1 and enumerate all the possible values. The value of ω which gives the smallest average delay for all the trains is used in the Switchable2-I policy.

For the Switchable2-II policy, the values of α, β and δ need to be determined. Without loss of generality, the value of α can be fixed at 1 and a good assignment of the values of β and δ can be obtained by discretization and enumeration. There are no obvious upper bounds for β and δ. In the experiments, the upper bound of β is set to 2 and the upper bound of δ is set to $1\left(D/S_{sl} - D/S_{fa}\right) + 2S_{fa}$. The best values of β and δ are always found to be far below their upper bounds. The best values of β and δ which produce the lowest average train delay are used in the Switchable2-II policy.

The values of α, β and δ in the Switchable2-III policy are determined the same way as in the Switchable2-II policy. The extra parameter μ has an upper bound of $(D + 1.136) / S_{sl}$ (where 1.136 accounts for the longest train length in unit of miles) and a lower bound of 0. The parameter μ is also discretized and enumerated together with the other two parameters.

In the Switchable2-w/cross policy, the speed and length of each arriving train have the same characteristics as in the previous dispatching policies. The crossover is located in the middle of the eight-mile long track segment. Suitable values of the parameters α_1, β_1, δ_1, μ_1, α_2, β_2, δ_2 and μ_2 in the Switchable2-w/cross policy can be obtained as in the Switchable2-III policy.

Table 3 shows the average delays of the four policies when the arrival rate is 0.16 trains per minute. The average is based on 10 simulation runs. By choosing a good switching threshold value, the Switchable2-I policy is able to significantly reduce the average delay from the dedicated policy. However, the more complex switching condition function of Switchable2-II policy reduces the delay some more and the Switchable2-III policy is able to further reduce the average train delay. Compared to the dedicated policy, the best switchable policy, the Switchable2-III policy reduces the average train delay by 21%.

Table 3. Comparisons of the four policies

Arrival rate = 0.16	Average train delay (min)	Standard deviation (min)
Dedicated policy	0.9666	0.0013
Switchable2-I	0.8202	0.0011
Switchable2-II	0.7923	0.0012
Switchable2-III	0.7623	0.0012

Table 4 shows the delay of the fastest trains (e.g. high priority passenger trains) under different dispatching policies on the 8-mile long track segment. The Switchable2-III policy is able to reduce the delay under the dedicated policy by 0.484 minutes. Considering that the normal route length of the passenger train is much longer than 8 miles (e.g., in the downtown Los Angeles area, the route length of passenger trains can be as high as 40 miles), the potential reduction of delay for passenger trains over their entire routes could be significant.

Table 4. Comparisons of the four policies (fastest train delay)

Arrival rate = 0.16	Average train delay (min)	Standard deviation (min)
Dedicated policy	1.7178	0.0013
Switchable2-I	1.3855	0.0011
Switchable2-II	1.3313	0.0014
Switchable2-III	1.2335	0.0013

Figure 6 shows the performance of the four policies as the arrival rate varies. The relative performances between the four policies remain the same as seen in Table 3. As expected, the delay increases as the arrival rates increase but the gap between the dedicated policy and the switchable policy reduces at the increased rates since there is less opportunity for switching when there are more trains in the network. Figure 7 shows the performance of the four policies as the track length varies. As the figure shows, the average delay relationship with the track segment length is similar to the arrival rates since there is less opportunity for switching with longer segments assuming no crossovers within the segment.
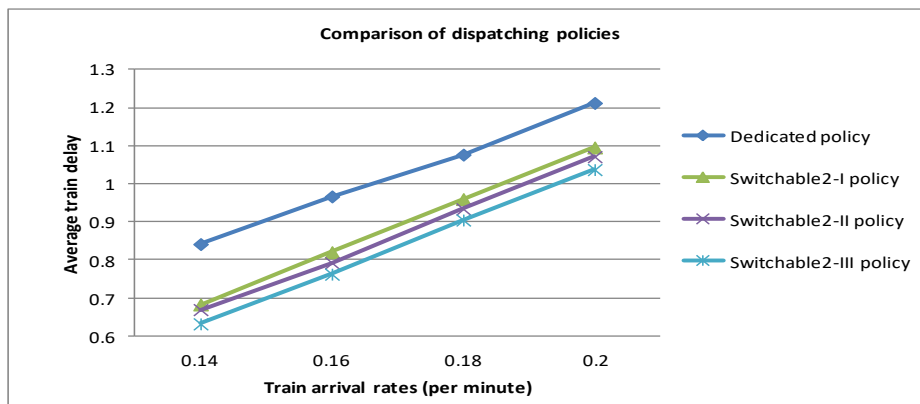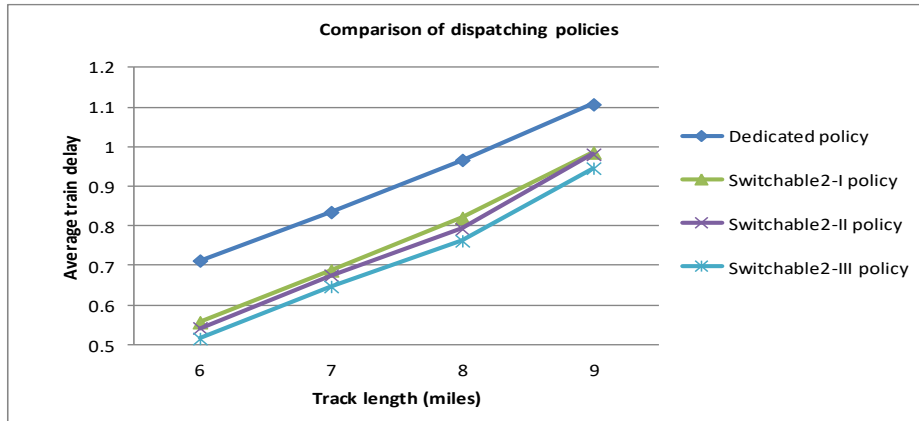

Figure 6. Varying arrival rates

14

Figure 7. Varying track lengths

Figure 8 shows the average delay under both the dedicated policy and the Switchable2-w/cross policy. The results clearly show that the Switchable2-w/cross policy dominates the dedicated policy as the arrival rates vary. In this numerical experiment, the Switchable2-w/cross policy can reduce the average train delay by as high as 41.9%.
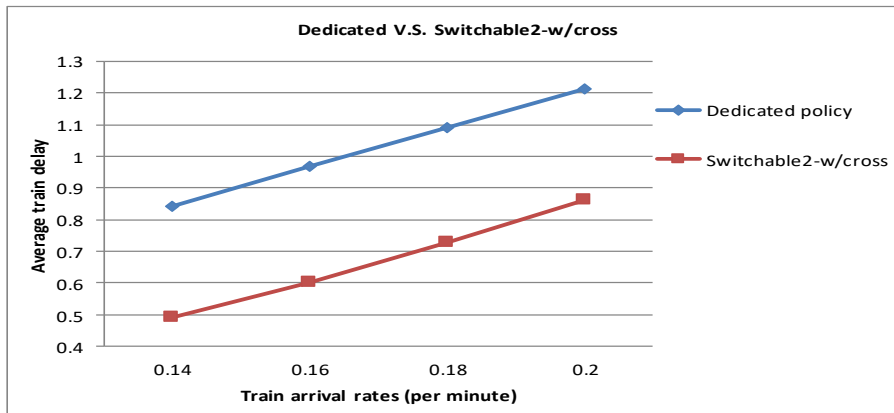

Figure 8. Dedicated vs. Switchable2-w/cross

Analysis of dynamic headway control

We compared our dynamic headway modeling approach with the constant headway approach. The average node length in the constant headway model is 1.63 miles while it is 0.83 and 0.55 miles in the dynamic headway model respectively.

We first tested the scenario with all trains travelling in the same direction (westbound from Pomona to Downtown Los Angeles). Simulation results are shown in Table 5. In these experiments, the number of trains is evenly divided between passenger and freight trains. For example, in the row where the number of trains per day is 10 in Table 5, the arrival rate in the simulation is 5 per day for freight trains and 5 per day for passenger trains. The arrival process for both freight trains and passenger trains was assumed to follow a Poisson Process.

15

Table 5. Average delay (measured in minutes) for the different approaches

Number of trains per day	Constant headway method with average node length of 1.63 miles	Dynamic headway method with average node length of 0.83 miles	Dynamic headway method with average node length 0.55 of miles
10	0.40	0.44	0.42
20	0.86	0.67	0.69
40	1.94	1.76	1.59
60	3.07	3.04	2.64
80	4.55	4.73	3.83
100	6.45	7.20	5.33
150	15.28	23.10	11.86
170	28.08	88.36[1]	16.15

The delay is measured as the difference between the actual travel time and the free travel time. The free travel time is the time a train takes to travel through the network when there are no other trains. In the above example, all trains are travelling in the same direction, so only two types of delay exist. The first type is at the beginning when a new train attempts to enter the rail network. The first node (in the constant headway model) or the set of nodes representing headway (in the dynamic approach) is already occupied by a train preventing the new train from seizing the track. The second type of delay is when a fast train (passenger) catches up with a slow train (freight).

Conclusions summarized from the above results are:

1. In all cases, the average delay increases nonlinearly with increasing train arrivals. However, using a constant headway approach, the rail network is fully saturated when there are around 200 trains so we only range the number of up to 170. And if a dynamic headway approach is used the rail network capacity will be around 250 trains per day.

2. The dynamic headway approach with a smaller node size (0.55 miles) significantly outperforms the constant headway method when there is congestion in the network (the number of trains is large).

3. The dynamic headway approach provides lower average delay if we construct a network with a smaller node size. Since the railway system is represented as a node-arc discretized network, a smaller discretization of the network more closely represents the actual continuous process of the train movement.

4. Sometimes the dynamic headway approach with a larger node size performs slightly worse than the constant headway method because in the dynamic headway approach we greedily seize headway nodes until a train reaches its maximal speed. If each node is relatively large, the trains may seize more railway track resources than necessary to ensure a safe headway when seizing the last headway node.

---

[1] This large number indicates that the dynamic headway method with a node length of 0.83 miles is not stable when the number of trains per day reaches 170.

We then tested the scenarios with trains traveling in both directions (eastbound from Downtown to Pomona and westbound from Pomona to Downtown). The results are shown in Table 6. In this case, the average train count per day is equally divided by each direction and type. Thus, for the row with 10 trains per day, the average number of freight trains travelling from Downtown Los Angeles to Pomona is 2.5. With trains travelling in both directions, there is a higher chance for delay due to conflict among trains moving in opposite directions.

Table 6. Average delay (measured in minutes) for bi-direction dispatching

Number of trains per day	Constant headway method with average node length of 1.63 miles	Dynamic headway method with average node length of 0.83 miles	Dynamic headway method with average node length of 0.55 miles
10	1.40	1.15	1.21
20	2.87	2.78	2.74
40	5.49	5.29	5.27
60	9.42	8.91	8.85
80	13.29	12.80	12.69
100	20.68	17.37	16.90
150	55.44	43.72	32.02

The following conclusions are derived from the above results:
1. Compared with Table 5 where all trains are traveling in the same direction, the average delay for trains traveling in both directions generally becomes higher because of a higher chance for conflict and therefore causing congestion.
2. The dynamic headway approach, even with a larger node size, performs much better than the constant headway method. An over 40% reduction in the average delay is achieved when the number of trains reaches 150 per day and the node length is 0.55 miles. Therefore, we can clearly see that the dynamic headway method can assign the track resources more efficiently.

Implementation

The work of this project deals with problems encountered in scenarios where passenger trains and freight trains are sharing the same rail network. As a large portion of freight shipments depend on rail roads, the necessity of increasing the rail network's efficiency becomes increasingly more important. One possible solution is to better dispatch and route the trains.

As shown in this report, the dispatching policies which are easy to implement can reduce train delays and facilitate the operation of train schedules. Moreover, building crossovers are shown to further improve the network. The dynamic headway control can be realized when new technology such as PTC is in use and our solution is shown to reduce train delays and increase network capacity even more.

Conclusions

In this research project, we aim to provide solutions to problems brought by passenger trains sharing the same rail network with freight trains, namely the different train speeds that add additional delays. We first develop and compare three heuristic dispatching policies as opposed to the traditional dedicated policy. We also investigated how placing crossovers could further improve the efficiency of our proposed policies. Considering the train lengths and the fixed safety headway, these policies are based on the idea of switching faster trains to the other track to reduce the delay caused by catching up with a slow train. Simulation results validate the efficiency of these policies and the best switchable policy reduces the average train delay by 21% and this number is almost doubled with the existence of crossovers in the middle of a double-track segment, resulting in a 41.9% reduction in the average delay. We then build a dynamic headway control model utilizing the benefits of new technology such as PTC to further reduce the average delay and increase the network capacity. A solution procedure is introduced and simulation experiments based on an actual railway network are performed. The chosen railway network is from Downtown Los Angeles to Pomona consisting of single-, double- and triple- tracks with varying speed limits. The simulation results show that with dynamic headway control, the rail capacity could be increased by 20%. Also, the dynamic headway control results in 40% less average delay with 150 trains per day traveling in both directions of the rail network.

References

[1] Association of American Railroads. Overview of America's Freight Railroads. 2003

[2] Association of American Railroads. Overview of America's Freight Railroads. 2017

[3] Fu, L., & Dessouky, M. (2017). Models and algorithms for dynamic headway control. *Computers & Industrial Engineering*, *103*, 271-281.

[4] Kelton, W. D., Sadowski, R. P., & Swets, N. B. (2010). *Simulation with Arena*. New York, NY: Mc-Graw-Hill.

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

[6] Mu, S., & Dessouky, M. (2013). Efficient dispatching rules on double tracks with heterogeneous train traffic. *Transportation Research Part B: Methodological*, *51*, 45-64.

[7] Mu, S., & Dessouky, M. (2014). Control rules for dispatching trains on general networks with multiple train speeds. *Conference Proceedings Transport Research Arena*.