Joint optimization of high-speed train timetables and speed profiles: a unified modeling approach using space-time-speed grid networks

Leishan Zhou
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, 100044, China
Email: lshzhou@bjtu.edu.cn

Lu Tong*
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, 100044, China
Email: ltong@bjtu.edu.cn

Junhua Chen
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, 100044, China
Email: cjh@bjtu.edu.cn

Jinjin Tang
School of Traffic and Transportation
Beijing Jiaotong University
Beijing, 100044, China
Email: jinjintang@hotmail.com

Xuesong Zhou
School of Sustainable Engineering and the Built Environment,
Arizona State University, Tempe, AZ, 85281, USA
Email: xzhou74@asu.edu
Tel.: +1 480 9655827

Submitted for publication in Transportation Research Part B

Abstract:
This paper considers a high-speed rail corridor that requires high fidelity scheduling of train speed for a large number of trains with both tight power supply and temporal capacity constraints. This research aims to systematically integrate problems of macroscopic train timetabling and microscopic train trajectory calculations. We develop a unified modeling framework using three-dimensional space-time-speed grid networks to characterize both second-by-second train trajectory and segment-based timetables at different space and time resolutions. The discretized time lattices can approximately track the train position, speed, and acceleration solution through properly defined spacing and modeling time intervals. Within a Lagrangian relaxation-based solution framework, we propose a dynamic programming solution algorithm to find the speed/acceleration profile solutions with dualized train headway and power supply constraints. The proposed numerically tractable approach can better handle the non-linearity in solving the differential equations of motion, and systematically describe the complex connections between two problems that have been traditionally handled in a sequential way. We further use a real-world case study in the Beijing-Shanghai high-speed rail corridor to demonstrate the effectiveness and computational efficiency of our proposed methods and algorithms.

Keywords: train trajectory planning; train timetabling; energy consumption; space-time-speed network

1 Introduction

In general, high-speed rail offers a fast and comfortable transportation mode with a high carrying capacity. Throughout the world, many new high-speed rail lines are being operated or designed to serve increasing inter-city passenger flow along economic corridors. On the other hand, due to its increased operating speed, high-speed train units have higher energy demand than conventional locomotives to overcome extra aerodynamic resistance. For high-speed rail operators, the design of time-efficient train timetables for passengers and energy-efficient train speed profiles under tight power supply and track capacity constraints becomes extremely important. Faced with service competitions from emerging fuel-efficient automobiles and state-of-the-art aircraft (Chester and Horvath, 2012), it is important for the rail industry to improve its own ridership in the multimodal inter-city transportation market, reduce the overall rail system operating cost, and contribute to societal sustainability in a long run.

In the last few decades, there has been a steady move towards the systematical use of computer simulation models and optimization tools to evaluate and further improve the energy use for single or groups of trains running on a corridor. Planners and dispatchers of high-speed trains need to target an ever-widening range of analysis goals, each with a particular set of constraints and focus areas. To name a few, commonly used optimization criteria for train operations include time-optimal, energy-optimal, or waiting-time-minimal. A typical train timetabling application focuses on a time-optimal schedule to minimize passenger in-vehicle travel time to meet the high mobility requirement of high-speed rail riders. However, demand-oriented timetable solutions might lead to higher energy consumption compared to energy-optimal schedules that typically have sophisticated, pre-designed acceleration and breaking profiles at various grade segments and speed-limit constraints. As a result, for emerging high-speed rail scheduling applications, the commonly used sequential processing of train timetabling and speed control is difficult in its own right to achieve the full system benefit under tight resource constraints.

The focus of our literature review, presented herein, is mainly on two related research areas, train timetabling, and train speed control. Cordeau et al. (1998), Lusby et al. (2011), and Cacchiani and Toth (2012) offered comprehensive surveys on both train routing and scheduling problems, including early literature summaries with both constant and variable train speed at segment levels. Recent studies, Törnquist and Persson (2007), D'Ariano et al. (2007, 2008), Corman et al. (2010), Meng and Zhou (2014), Samà et al. (2016), Fu and Dessouky (2016) have examined various important topics of train scheduling, including timetabling under a complex network or N-track condition and real time scheduling to mitigate delay propagation under various degrees of disturbances. Corman and Meng (2013) offered a systematic review on real-time train dispatching, rescheduling, and disposition under stochastic and dynamic conditions. In order to describe many complex but practically important train operational constraints, and focus on Chinese railroad timetabling applications, early work by Zhou et al. (1998) applied a Discrete Event Dynamic System (DEDS) method for offline train scheduling applications. Medanic and Dorfman (2002) and Dorfman and Medanic (2004) discussed time-efficient and energy-efficient scheduling methods using a general DEDS modeling framework for prioritizing train movement events for a group of trains running along a rail corridor.

Early studies on energy-efficient train movement control (e.g., Milroy, 1980; Benjamin et al., 1989) typically focused on controlling single train speed under various track and geometric conditions, where a locomotive has several levels of traction control gears, each corresponding to different energy consumption rates. Cheng and Howlett (1992, 1993) studied critical speed thresholds to minimize the fuel consumption of individual trains on horizontal lines. Howlett (1996) proposed various optimal train speed control strategies for a train on segmented constant grades, followed by the study by Howlett and Cheng (1997) on optimal driving strategies on a track with continuously varying grades, and the paper by Cheng et al. (1999) on segmented constant gradient with speed limits. As one of the key studies in this area, Liu and Golovitcher (2003) proposed an analytical model for the energy-efficient train speed/acceleration control along long distance rail lines with various speed limits and grade changes. They considered continuously changing control variables to reduce the computational complexity associated with different resolutions. Wong and Ho (2003) studied coast control strategies and related switch points of subway train movement using

hierarchical genetic algorithms. Bai and Mao (2009) discussed the energy consumption performance with respect to different coasting distances under various practical considerations such as braking/lower speed restrictions and train speed uniformity. Aiming for a globally optimal strategy, the study by Howlett et al. (2009) calculates the critical switching points on a track with steep grades through a local energy minimization principle. Recent notable processes along this research line have been made by Bocharnikov et al. (2010), Domínguez et al. (2012), Rodrigo et al. (2013) to optimize energy through the fully utilized regenerative energy sources. Though a comprehensive perturbation analysis for local optimal points (satisfying necessary optimality conditions), the studies by Albrecht et al. (2011, 2013) show that the optimal switching points can be uniquely determined for each steep section of track.

There has been a number of emerging studies presenting methods to perform train timetabling and speed profile optimization simultaneously. To achieve a better energy performance for subway trains, Su et al. (2013) proposed an iterative sequential optimization model to take into account both train speed profiles and trip times. Li and Lo (2014a) proposed a nonlinear integer optimization model and a genetic algorithm-based solution method to jointly consider timetabling and speed control between switching points along a metro rail line. They also specifically considered (1) how to synchronize the acceleration and breaking events and (2) how to utilize regenerative energy to minimize energy consumption. Li and Lo (2014b) developed a simplified but innovative analytical model that assumes two switching points per segment for dynamic train control and schedule adjustment, in order to construct a linear approximation-based convex programming model. Using a discretized space-time network modeling framework proposed by Yang and Zhou (2014), Yang et al. (2015) considered to optimize energy-efficient train timetables through variable segment speed expressed as different segment travel times. Using a multi-train simulator, Zhao et al. (2015) recently developed a number of solution search algorithms, such as enhanced brute force, ant colony optimization, and genetic algorithm, to optimize multiple train trajectories, with joint goals of minimizing energy consumption and delay. Recently, Goverde et al. (2016) recognized and highlighted the importance of a consistent cross-resolution representation for macroscopic aggregated railroad network optimization, microscopic train timetabling, and fine-tuning train trajectory computation. Furthermore, Besinovic et al. (2015) proposed a hierarchical framework for robust timetable design that produced feasible timetables through a heuristic micro-macro interaction algorithm. Table 1 provides a comparison of recent studies in integrated train timetabling and speed profile optimization.

Table 1 Recent studies on integrated train timetabling and speed profile optimization.

Major decision variables for timetabling	Model	Algorithm	Publication
Trip time; speed between switching points	Nonlinear constraints and objective functions	Iterative algorithm to solve macroscopic and microscopic problems sequentially	Su et al. (2013)
Arriving/departure time; speed between switching point	Nonlinear constraints and objective functions	Genetic algorithm	Li and Lo (2014a)
Arriving/departure time; speed between switching point	Convex optimization model	Analytical method	Li and Lo (2014b)
Arriving/departure time; average segment speed	Discretized space-time network to represent variable speed	Commercial optimization software GAMS	Yang et al. (2015)
Train trajectories	Time-step based multi-train simulation with continuous speed, space, time representation	Enhanced brute force, ant colony optimization, and Genetic algorithm	Zhao et al. (2015)
Arriving/departure time; running time	Integer linear programming (for macroscopic problem)	Randomized multi-start greedy heuristic	Goverde et al. (2016)
Discretized space-time-speed grid	Multi-flow network	Dynamic programming	This paper

	network, multi-commodity variables representing both timetable and speed profile characteristics	model with linear objective function and constraints

Most of the existing studies on train speed control for energy-efficient operations focus on optimizing single train movement or a cluster of train speed profiles with simple overtaking relationships. It is more challenging to construct mathematically tractable models and computationally efficient algorithms for a rail line with a number of switching points, especially on long rail segments with non-regular gradient changes. Improving both travel time and energy measures through joint train timetabling and speed control optimization requires a broader understanding of complex constraints and operating processes at different spatial and temporal regimes. In particular, the lack of a unified theoretical model that simultaneously represents these two coupled problems makes the existing sequential optimization models difficult to work together in a seamless manner.

One important challenge in operating a large number of high-speed trains on an electrified high-speed rail corridor is that the trackside traction power supply facilities need to supply the voltage required by different trains on a second-by-second basis. The conversional train timetabling stage, typically operated at a minute-by-minute resolution, is inflexible to meet the overall energy supply constraints and consider the total power consumption optimization for trains spatially distributed at the power supply segments of a corridor.

The relatively long modeling time interval (e.g., min by min) used in timetabling is difficult to connect to the high-fidelity spatial and temporal second-by-second representation for the subsequent train movement control stage. In this research, we construct a space-time-speed (STS) network that seamlessly maps train trajectories to: (1) the space-time network for scheduling (subject to time headway constraints), and then to (2) space-speed network for detailed train control that involves train position, speed, and acceleration for accurate energy calculation. Through a customized spatial and temporal discretization scheme, our research creates a cell-based network with a flexible representation of a number of constraints for different train space-time-speed trajectories. These key constraints include: (1) the spatial and temporal safety headway for timetabling signal systems and (2) total energy consumption across trains running on the same power supply segments but possibly on different tracks. It should be remarked that, one closely related study by Miyatake et al. (2009) solve the train optimal speed control problem using Bellman's dynamic programming method and described the discretized, linearized state in a three-dimensional graph. In comparison, our research focuses more on how to establish an internally consistent three-dimensional space-time-speed network to connect train speed control and timetabling problems.

Furthermore, we will develop a dynamic programming-based algorithmic solution framework for single-train subproblems. In particular, the dynamic programming (DP) algorithm can work well for the acyclic STS network with only forward moving arcs along the time axis, and the Lagrangian relaxation (LR) method can dualize the safety headway and energy consumption constrains. In a decomposed sub-problem, each single-train subproblem searches its own STS network with generalized costs from LR multiplier, which allows us to effectively solve the joint optimization problem.

The reminder of this paper is organized as follows. The next section describes the joint train timetable and speed profile optimization problem and provides the conceptual illustration of a space-time-speed network. Section 3 develops a linear integer programming model and Lagrangian relaxation based dynamic programming algorithm. Section 4 evaluates our proposed model and algorithm with some real-world medium-scale and large-scale examples, and the final section concludes the paper.

2 Problem Statement and Conceptual Illustration

2.1 Problem statement

This paper considers a high-speed rail line as illustrated in Fig 1, with a series of bi-directional track segments within a power supply district. The power supply district has an autotransformer in the middle, a

section post on the left, and a traction/power substation on the right. Through contact wires, the power supply district covers three train stations and two track segments. In this high-speed train corridor, we consider a regular traffic signal system, where trains can follow each other on a track segment when satisfying minimum time headway requirements. For simplicity, only one type of train is included in the scheduling process and different trains can pass each other only at station track sidings. The traction substation supplies power for all trains within its coverage territory. In general, an intercity high-speed rail line consists of a number of power supply districts. A power supply district typically has a service range from 50 to 300 km.

A typical passenger train scheduling process consists of the following sequential steps: passenger demand survey, train service plan development, and train timetabling. This paper focuses on the third step, which receives input from the train service plans including desired train departure time windows and train stop plans (Yue et al., 2016). Indirectly, passenger demands are considered in this timetabling process as the number of trains to be scheduled, with preferred departure/arrival times and stop plans (intermediate stops). Given train service plans, the problem under consideration needs to find optimal train schedules that can minimize energy consumption subject to essential safety headway constraints and power supply resource constraints.

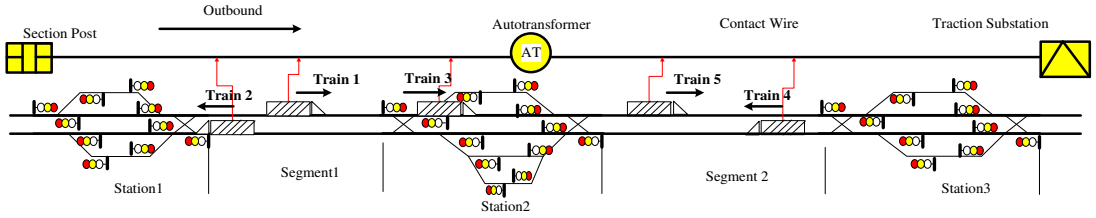


Fig 1. The network representation of a high-speed rail line within a power supply district (Tang, 2012)

The space-time diagram or the train string diagram in Fig 2 shows a schedule with 5 trains numbered from 1 to 5. If we only consider the minimum safety headway h between each pair of trains at stations, this schedule is feasible without any safety conflicts. When we need to consider the power supply resource constraint across all trains running within the same power supply district, it is extremely complex to consider tractive effort and line voltage for a set of high-speed trains, as briefly discussed in Appendix A.

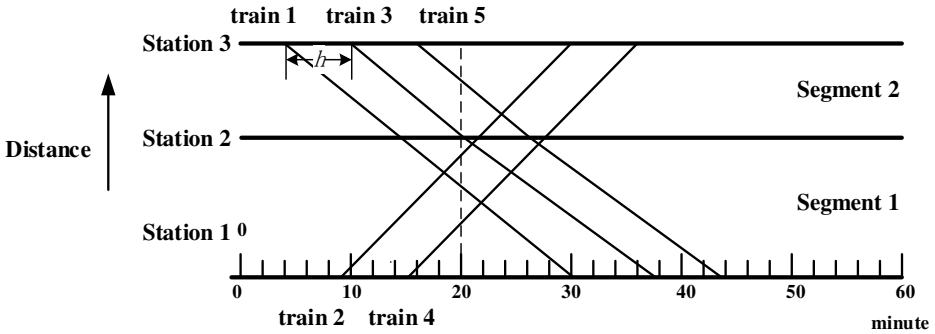


Fig 2. A sample high-speed train schedule (Tang, 2012)

Fig 3 shows a piecewise linear curve between maximum available tractive effort and line voltage supplied from a power substation for high-speed trains. The tractive effort at the wheel/rail interface is a nonlinear function of train speed, acceleration, mass, as well as aerodynamic and friction resistance under specific grade and curve conditions. If the supplied line voltage from the power supply system to a specific train is insufficient, the available power is reduced and impacts the tractive effort, resulting in a narrower

acceleration range and/or lower maximum speed.

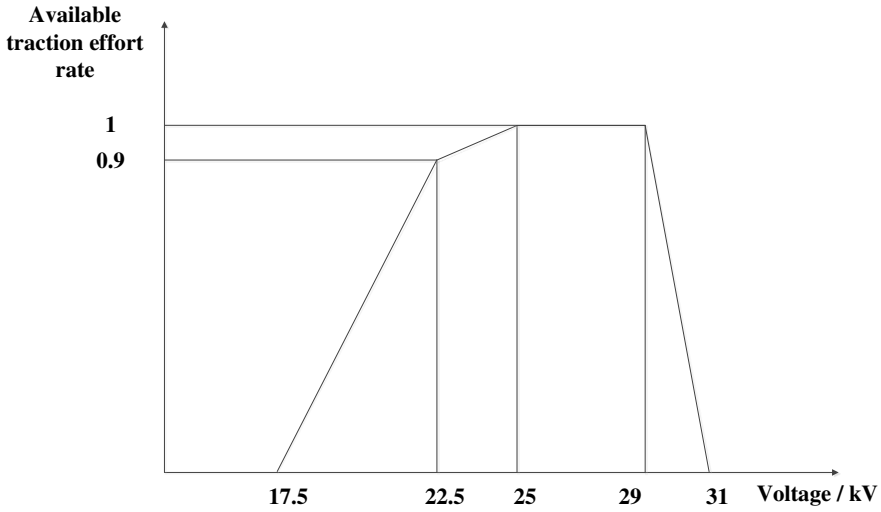


Fig 3. Relationship between pantograph-catenary voltage and traction effort rate at wheel rim (from CRRC Corporation Limited, China)

Consider an example using Fig 2, where the total power supplied to trains running in both directions is insufficient at a particular time stamp (e.g., min 20). In this scenario, high-speed trains cannot run at their target speed and this may result in arrival time delays at stations. In a real-time scheduling application, a train (e.g., train 1) with an initial delay sometimes needs to accelerate and speed up in order to return to the planned schedule as soon as possible. Its additional power demand has the potential to affect other trains operating within the same power supply district. A method to fully synchronize the (spatially distributed) train movements and schedules, more specifically, starts and stops of trains, is an extremely complex process as it has to deal with the complicated interconnections between multiple trains at very high time and space resolutions.

2.2 Continuous functional form for train motion

To derive detailed energy consumption based on the kinematics of train movements, we now list a number of related differential train motion equations and the notations in Table 2. For notational convenience, we omit the train index.

Table 2 Parameters used in the equation of motion

Symbol	Definition
$\boldsymbol{H}$	Set of time clocks for analysis, $\{1, \dots, T\}$
$\boldsymbol{t}$	Indices of time clocks, $t \in H$
$\boldsymbol{d(t)}$	Distance along the track at time t
$\boldsymbol{v(t)}$	Train speed at time t
$\boldsymbol{m}$	Train mass
$\boldsymbol{acc(t)}$	Acceleration associated with train traction force at time t, that is the ratio of tractive force F and mass m.
$\boldsymbol{dec(t)}$	Deceleration associated with train breaking force at time t
$\boldsymbol{w(v(t))}$	Resistive acceleration due to train motion at speed $v(t)$
$\boldsymbol{h(d(t))}$	Resistive acceleration due to the track gradient and curve at location $d(t)$

Eq. (1) defines the vehicle speed $v(t)$ from distance change. Eq. (2) indicates that the resultant acceleration/deceleration is affected by train traction $acc(t)$ or braking force $dec(t)$, speed-related train motion resistance $w(v(t))$ and location-related track grade and curve resistance $h(d(t))$.

$$\frac{\mathrm{d}d(t)}{\mathrm{d}t} = v(t) \tag{1}$$

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = acc(t) - dec(t) - w(v(t)) - h(d(t)) \quad (2)$$

The energy cost of the train journey is the total energy supplied to the train, which can be calculated from the work done by the traction force. The total energy cost or the total mechanical power at the wheel/rail interface J of the journey from time 1 to time T is formulated in Eq. (3).

$$J = \int_{t=1}^{T} m \cdot acc(t) \cdot v(t)\,\mathrm{d}t \quad (3)$$

2.3 Constructing a space-time-speed grid network for joint train routing, timetabling and trajectory optimization as a path finding problem

Before constructing the space-time network representation, we now start discretizing the time horizon to a sequence of time intervals with a certain length (e.g., 1 second) and the time index can be denoted as $t = 1, 2, \ldots, |H|$.

A space-time network representation has been widely used in transportation route optimization, various scheduling applications, and general dynamic network flow modeling. Interested readers are referred to general survey papers by Powell et al. (1995), and recent development for time-dependent stochastic shortest path by Yang and Zhou (2014). Typically, for a physical rail segment (i, j) connecting a pair of physical nodes i and j, we can create the corresponding time-expanded arcs (i, j, t, s) where t is entering time, s is the exit time of a traveling arc, and $s - t$ corresponds to the train running time for a certain train k on link/segment (i, j). For a waiting arc at node i, one can use $(i, i, t, s = t + 1)$ to denote the waiting step from time t to time $s = t + 1$ at the same node i.

As shown in Eq. (3) for calculating train energy consumption, one of the key elements is the acceleration $acc(t)$ as a difference of $v(t)$ at different times. In a conventional space-time network, for a pair of space-time vertices (i, t) and (j, s), there are an infinite number of trajectories connecting these two points, each with the same speed but different possible acceleration/deceleration strategies. In this case, it is difficult to directly use the attributes at two adjacent points (i, t) and (j, s) to calculate the resulting train motion characteristics and energy consumption rates that highly depend on the train speed u at each time clock.

For a set of physical nodes/stations N and a set of physical rail links L, this paper defines a directed STS network $G = (Q, A)$ with each STS vertex $(i, t, u) \in Q$ and each STS arc $(i, j, t, s, u, v) \in A$ that indicates the train movement from node i to node j at entering time t with speed v and exit time s with speed u. The notations based on STS network are listed in Table 3.

Table 3 Subscripts, parameters, and variables used in mathematical formulations.

Symbol	Definition
$\boldsymbol{K}$	Set of trains
$\boldsymbol{H}$	Set of time stamps in the planning time horizon
$\boldsymbol{N}$	Set of nodes/stations
$\boldsymbol{L}$	Set of links
$\boldsymbol{V}$	Set of speed values
$\boldsymbol{Q}$	Set of STS vertices
$\boldsymbol{A}$	Set of STS arcs
$\boldsymbol{k}$	Index of trains, $k \in K$
$\boldsymbol{t, s, \tau}$	Time indices, $t, s, \tau \in H$
$\boldsymbol{i, j}$	Indices of nodes, $i, j \in N$
$\Delta\boldsymbol{t}, \Delta\boldsymbol{d}$	Time increment and space increment
$\boldsymbol{(i, j)}$	Index of physical link, $(i, j) \in L$
$\boldsymbol{u, v}$	Values of instantaneous speed, $u, v \in V$
$\boldsymbol{(i, t, u)}, \boldsymbol{(j, s, v)}$	Indices of space-time-speed vertexes, $(i, t, u), (j, s, v) \in Q$
$\boldsymbol{(i, j, t, s, u, v)}$	Index of space-time-speed arcs indicating the train movement from node i to node j at entering time t with speed v and leaving time s with speed u, arc $(i, j, t, s, v, u) \in A$
$\overline{\boldsymbol{acc}}\boldsymbol{(i, j, t, s, u, v)}$	Average acceleration associated with train traction force on STS arc (i, j, t, s, u, v)
$\boldsymbol{e(i, j, t, s, u, v)}$	Mechanical energy consumption amount on traveling arc (i, j, t, s, u, v)

To succeed in meeting a number of study objectives for joint scheduling of train schedule and movement, the construction of the STS network will be guided by the following set of network construction principles.

1. Similar to the finite difference method used in cellular automata models for traffic flow simulation (Daganzo, 2006), the space along the rail corridor is discretized in increment Δd, so the location of any node i can be denoted as $n^i \cdot \Delta d$, where n^i is an integer number. As a result, the distance between any node pair (i, j) should be an integer multiple of Δd.
2. The time along the planning horizon is discretized in increment Δt, the mapped time stamp for index t is $t \cdot \Delta t$.
3. If we consider the high-speed train constantly accelerates/decelerates during time interval or time step Δt, the speed along the speed dimension is discretized in increment $\Delta d/\Delta t$. This property can meet the following proposition: for any arc (i, j, t, s, u, v) between time index t to time index $s = t + \Delta t$, when v is a feasible value (as an integral multiple of $\Delta d/\Delta t$), then the ending speed u is uniquely determined as a feasible value as well. During speed holding process with the acceleration of zero, the speed u equals to v.
4. The upper bound speed at node i for train k is the minimum value of the maximum speed of train k and speed limit at node i.
5. For node i which is not an intermediate stop for train k, the corresponding STS vertex $(i, t, 0)$ for any time t and its connected arcs should be set to invalid for train k. That is the vertex and arc are eliminated from train k's STS network and thus not usable by train k.
6. To model the must-stop train activity, for node i' which is an intermediate stop for train k, we should eliminate any non-stop passing arc (i, j, t, s, u, v), where the intermediate stop node i' is between node i and j (i.e., $i < i' < j$). Additionally, when a desired stop time window $[t', s']$ is given, the arcs with arrival time earlier than t' or departure time later than s'at node i' should be eliminated.
7. There are three types of arcs: (i) traveling arcs (i, j, t, s, u, v) moving from node i to node j with a unit time interval from t to time s; (ii) waiting arc $(i, i, t, s, 0, 0)$ for node i that is the origin or destination node of train k; (iii) dwell arc $(i, i, t, s, 0, 0)$ for node i where i is an intermediate stop for train k and time duration $(s - t)$ is the required stopping time.
8. A traveling arc with a speed change from u to v is only feasible when the corresponding acceleration/deceleration value falls in the range between maximum acceleration determined by traction power, and maximum deceleration determined by braking force. Traveling arcs consist of four subclasses of arcs, namely full power traction, cruising, coasting and full braking arcs.

The proof for guiding principle 3 can be briefly described as follows. The travel distance between (i, j) is $(n^j - n^i) \times \Delta d$, the average speed is $(u + v)/2$, so that $(u + v)/2 \cdot \Delta t = (n^j - n^i) \cdot \Delta d$, that is $u = 2(n^j - n^i) \cdot (\Delta d/\Delta t) + v$. Since we have setup the space-time lattice with the space increment Δd and the time increment Δt, we can easily show that if v is an integer multiple of $(\Delta d/\Delta t)$ then u must definitely be an integer multiple of $(\Delta d/\Delta t)$.

By recognizing the above network construction principles, we now present a few remarks on the STS grid network. By adding the speed dimension into the space-time network, where each vertex is now associated with location, time, and speed, the acceleration/deceleration rate between two vertexes can be uniquely determined. More importantly, from a state-space representation perspective (typically for dynamic programming), the STS network can clearly define the speed state and the associated stop status of the train movement. This enables modelers to be flexible in assigning the feasible arcs or arc transitions for different practical rules such as speed limits, maximum acceleration/deceleration rates, as well as the required stopping stages at the origin, destination and intermediate stations.

Within the proposed STS framework, train overtaking is allowed at intermediate stations as part of the network building process, namely principle 6. First, at the scheduled intermediate stops, we build dwell STS arcs and eliminate the non-stopping STS arcs. When a train stays at its scheduled intermediate stops

(by selecting the corresponding dwell arcs), other trains can overtake it (by using a non-stopping STS arc through the station). When deciding the pattern of overtaking, the optimization process selects a solution that could lead to a better system performance measure (e.g., energy consumption).

Specifically, an STS network is able to systematically incorporate many modeling methods based on (i) the space-time network used in the general area of train scheduling and (ii) the train speed profiles used in the area of an optimal train control problem. For example, the space-time network based train routing and scheduling algorithm (e.g., Meng and Zhou, 2014) can be further applied to the STS network representation, while the specific constraints on train motion and energy cost functional form (e.g., Yang et al., 2012) can be also adopted.

In the current space-time discretization scheme, which linearly interpolates inside the grid, the cost functions associated with the trajectories might be discontinuous at the boundary between two rectangles. More importantly, there could be different degrees of truncation errors in the space and time discretization, especially where there are non-uniform acceleration and higher order derivatives along the trajectories. We refer to Kushner and Dupuis (1992) on various techniques for discretizing a continuous time and space using finite-element (FE) or finite difference (FD) methods, and future research should address the consistency, stability, and convergence issues associated with different numerical approximation schemes. One promising direction is to use a variable resolution policy and value function representations proposed by Munos and Moore (2002) to dynamically implement the state-space triangulation using a tree structure.

The example in Fig 4 depicts a simple train motion trajectory in an STS grid network, where a train departing at time t from node i with an initial speed 0 stops at node $i+4$ at time $t+3$. By using this STS network representation, we can also account for the location-specific attributes related to train running states, such as gradient resistances, curve resistances, speed limits, and power supply constraints. For instance, for varying speed limits at different rail segments, one can simply eliminate the STS vertexes beyond the speed limits from the feasible solution domain. These unreachable vertexes cannot be chosen during the train movement optimization process. Fig 4, where green cycles represent the candidate STS vertexes and gray circles represent the infeasible vertexes, illustrates the spatial speed limit requirement of $v = 2$ (e.g., due to construction zone) at location $i+2$ and $i+3$, compared to the other segments with a normal speed limit of $v = 5$.

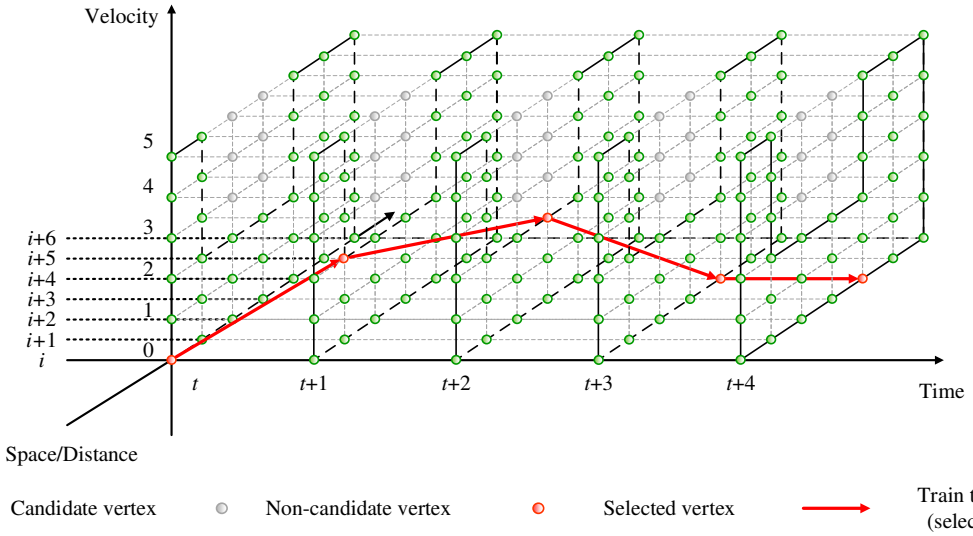


Fig 4. A train trajectory in an STS grid network (with a speed limit between nodes $i+2$ and $i+3$)

Using a post-processing step, we can further map the (optimized or selected) three-dimensional train trajectory from the STS network to the two-dimensional driver-oriented train speed profiles (one speed-space plane) and scheduling-relevant train timetables (on space-time plane). As shown in Fig 5, the projection of the train trajectory onto the 2-dimensional space-speed network clearly shows the acceleration,

cruising, and deceleration stages in the train speed profile for various sections of rail track. Similarly, the projection of the train trajectory onto the 2-dimensional space-time plane indicates train timetables.

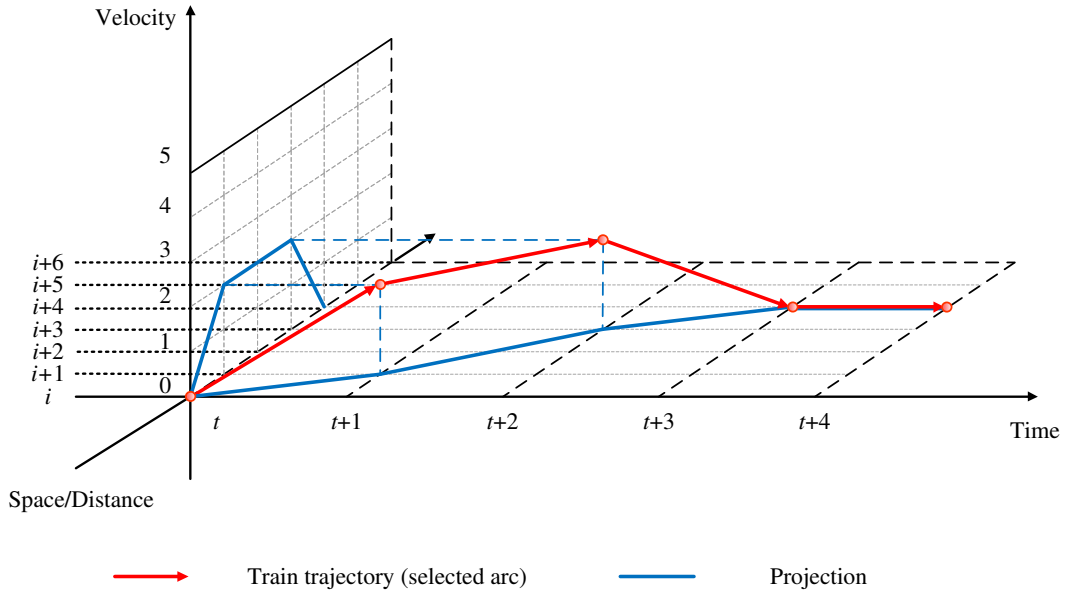


Fig 5. Projections of a train STS trajectory to the space-time and speed-space planes

For a better illustration, Table 4 lists sample values associated with STS traveling arcs. Given a constant speed limit of 7 units, different types of arcs are built by following the STS network construction principles and listed in Table 4. 1) During the acceleration process, a train travels 230 distance units and takes 42 time units for acceleration from standstill to the maximum speed. The full power traction arcs $(0{,}230, t, t+42, 0{,}7)$ are built at node 0 for each time t, whose cost equals the energy consumption for acceleration. 2) Cruising arcs $(i, i+8, t, t+1, 7{,}7)$ describe the speed-holding stage where a train moves 8 distance units within 1 time unit at the maximum speed. The arc cost equals the energy consumed against resistance. 3) Train coasting occurs before the braking. Its corresponding arcs, e.g., $(i, i+106, t, t+12, 7{,}6)$, are constructed at each node. 4) Full braking arcs allow trains to stop at their destinations. The time and distance taken vary according to when and where the train begins to brake. Moreover, as the mechanical power is not required in the last two types of arcs, their arc costs are set to zero in our proposed model.

Table 4 Sample values associated with STS arcs

Arc type	Arc						Sample Arc cost
	i	j	t	s	u	v	
Full power traction	0	230	t	$t+42$	0	7	1772
Cruising	i	$i+8$	t	$t+1$	7	7	24
Coasting	i	$i+106$	t	$t+12$	7	6	0
Full braking	i	$i+41$	t	$t+11$	6	0	0

2.4 Energy consumption function for train traveling arcs

By entering the corresponding values for STS traveling arc (i, j, t, s, u, v) for the continuous presentation of speed change in Eq. (2), that is, $\frac{\mathrm{d}v(t)}{\mathrm{d}t} \cong \frac{v-u}{s-t}$; $w(v(t)) \cong w\left(\frac{u+v}{2}\right)$; $h(d(t)) \cong \frac{h(i)+h(j)}{2}$, we can derive the average acceleration $\overline{acc}(i, j, t, s, u, v)$ by Eq. (4).

$$\overline{acc}(i, j, t, s, u, v) = max\left\{\left(\frac{v-u}{s-t} + w\left(\frac{u+v}{2}\right) + \frac{h(i)+h(j)}{2}\right), 0\right\} \quad (4)$$

According to Eq. (4), we can compute the mechanical energy consumption $e(i,j,t,s,u,v)$ on STS arc (i,j,t,s,u,v) of each lattice in Eq. (5) (as an integral from time t to time s).

$$e(i,j,t,s,u,v) = m \times \overline{acc}(i,j,t,s,u,v) \times \frac{u+v}{2} \times (s-t) \quad (5)$$

For train waiting and stopping arcs, the corresponding energy consumption is assumed to be zero. For travel arcs with deceleration, we assume zero consumed energy. For a more sophisticated train operating mode with regenerative energy sources, one can further determine a negative value for $e(i,j,t,s,u,v)$ to feed the energy from braking back to the power supply system.

3 A Model and an algorithm for the joint train trajectory and scheduling optimization

3.1 Model

The optimization model aims to incorporate train speed control into the timetable design process considering realistic operating constraints. Table 5 lists notations used in the optimization model.

Table 5 Notations used in the joint train control and scheduling optimization model

Symbols	Definition
$\boldsymbol{o_k, d_k}$	Origin and destination of train k
$\boldsymbol{EDT_k, LAT_k}$	Earliest departure time and latest arrival time of train k
$\boldsymbol{P}$	Set of power supply districts
$\boldsymbol{p}$	Index of power supply district, $p \in P$
$\boldsymbol{R_p}$	Power supply capacity per time unit (e.g., KW/h) of power supply district p
$\boldsymbol{G_{p,t}}$	Set of vertices in power supply district p at time t
$\boldsymbol{h}$	Safety time headway
$\boldsymbol{\varphi(i,j,t)}$	Set of incompatible STS arcs of link (i,j) at time t
$\boldsymbol{c_{i,j,t,s,u,v}}$	Cost on arc (i,j,t,s,u,v)
$\boldsymbol{x_{i,j,t,s,u,v}(k)}$	=1 if the STS arc (i,j,t,s,u,v) is used in the trajectory of train k =0 otherwise

Objective function

The joint train control and scheduling problem aims to find the speed profiles and timetables under specific optimization goals (e.g., minimizing energy consumption). Within the STS network framework, the objective function to minimize the total cost of all high-speed trains is stated in Eq. (6), where the seven-dimensional binary variable $x_{i,j,t,s,u,v}(k)$ is 1 if the STS arc (i,j,t,s,u,v) is used by train k and 0, otherwise. The parameter $c_{i,j,t,s,u,v}$ represents the cost when the corresponding STS arc (i,j,t,s,u,v) is selected.

$$\min\ Z = \sum_{k \in K} \sum_{(i,j,t,s,u,v) \in A} \left(c_{i,j,t,s,u,v} \cdot x_{i,j,t,s,u,v}(k) \right) \quad (6)$$

The generic cost structure in our model provides a flexible way to apply the STS network modeling framework to various objective functions, such as energy consumption and time consumption. For the energy-optimal objective under consideration in this paper, $c_{i,j,t,s,u,v}$ is identical to the mechanical power consumption $e(i,j,t,s,u,v)$ defined in Section 2.4. In fact, the high-dimensional label cost matrix from the final DP solution provides a vector of energy use cost across time index $t = EAT_k$ to LAT_k at the final destination station.

STS flow balance constraint

To depict a feasible train trajectory in the STS network, a set of flow balance constraints is constructed as follows. For each individual train k, its feasible flow needs to be satisfied at all vertices, starting from the source vertex $(o_k, EDT_k, 0)$ and ending at the sink vertex $(d_k, LAT_k, 0)$. We can state the flow balance constraints formally as follows.

$$\sum_{(j,s,v)\in Q} x_{i,j,t,s,u,v}(k) - \sum_{(j,s,v)\in Q} x_{j,i,s,t,v,u}(k) = \begin{cases} 1, & i = o_k, t = EDT_k, v = 0 \\ -1, & i = d_k, t = LAT_k, v = 0 \\ 0, & otherwise \end{cases} \quad for\ all\ k \in K \qquad (7)$$

Power supply constraint

Due to the capacity limitation of transformers illustrated in Section 2.2, the total mechanical power required by trains cannot exceed the power supply capacity R_p at any time in power supply district p. Let vertex set $G_{p,t}$ represent all STS vertices in power supply district p at time t. Then we present the power supply constraints by Eq. (8). The unit on two sides of Eq. (8) could be KW per hour or per second. On the other hand, Appendix A intends to convey the fact that the practical computation of railway power supply is quite complex as it involves current, voltage and resistance states of different vertical wires in a sophisticated multi-conductor traction network. This complexity calls for a simulation-based solution algorithm in which we are able to model a large number of nonlinear transformations and constraints realistically and practically.

$$\sum_{k\in K} \sum_{(i,j,t,s,u,v)\in G_{p,t}} \frac{e(i,j,t,s,u,v)}{s-t} \cdot x_{i,j,t,s,u,v}(k) \le R_p \quad for\ all\ p \in P, t \in H \qquad (8)$$

Train safety constraint (time headway)

As the time headway h must be enforced between any pair of vehicle trajectories according to railroad safety operational rules, we define the time headway-related train safety constraints using the incompatible arc set $\varphi(i,j,t)$. That is, if a train uses an STS arc (i,j,t,s,u,v), denoted as arc a_0, other trains are not permitted to use any STS arcs from the set of arcs on physical link (i,j), say arcs $a_1, a_2, \cdots, a_M$, within the time range $[t-h, t+h]$ at node i and time range $[s-h, s+h]$ at node j, respectively. Thus, the incompatible arc set $\varphi(i,j,t)$ includes all the potentially conflicting STS arcs $a_0, a_1, a_2, \cdots, a_M$, and only one of the arcs can be selected. The definition of $\varphi(i,j,t)$ is similar to the concept of cliques with incompatible arcs due to safety headway constrains introduced by the seminar paper by Caprara et al. (2002). That is, in our paper, a certain pair of different STS arcs associated with different trains cannot be selected together as shown in constraint (10), while Caprara et al. (2002) previously defined the similar set to denote incompatible space-time arcs, but without speed dimension.

$$\sum_{(i,j,t,s,u,v)\in\varphi(i',j',\tau)} \sum_{k\in K} x_{i,j,t,s,u,v}(k) \le 1 \ for\ all\ (i',j') \in L, \tau \in H \qquad (9)$$

Binary variable constraint

There is also binary definitional constraints for variables $x_{i,j,t,s,u,v}(k)$.

$$x_{i,j,t,s,u,v}(k) \in \{0,1\} \ for\ all\ k \in K, (i,j,t,s,u,v) \in A \qquad (10)$$

3.2 Problem decomposition through a set of single-train optimization problems based on Lagrangian relaxation

Compared with classic mathematic models for the shortest path problem, our joint train control and scheduling optimization problem has two additional side constraints (8) and (9). Instead of solving this problem directly, we relax these difficult constraints by incorporating them into the objective function with power supply multipliers $\pi_{p,t}$ and safety multipliers $\rho_{i',j',\tau}$ following the Lagrangian relaxation procedure. The Lagrangian relaxation problem derived in Eq. (11) can be treated as a generalized least cost/shortest path problem in a time-expended network (Caprara et al., 2002; Cacchiani et al., 2012; Fischer and Helmberg, 2014). The dualized cost $c'_{i,j,t,s,u,v}(k)$ is defined in Eq. (12).

$$\min\ Z' = \sum_{k\in K} \sum_{(i,j,t,s,u,v)\in A} \left(c_{i,j,t,s,u,v}(k) \cdot x_{i,j,t,s,u,v}(k)\right) + \sum_{p\in P} \sum_{t\in H} \left[\pi_{p,t} \left(\sum_{k\in K} \sum_{(i,j,t,s,u,v)\in G_{p,t}} e(i,j,t,s,u,v) \cdot x_{i,j,t,s,u,v}(k) - R_p\right)\right] + \sum_{(i',j')\in L} \sum_{\tau\in H} \rho_{i',j',\tau} \left(\sum_{k\in K} \sum_{(i,j,t,s,u,v)\in\varphi(i',j',\tau)} x_{i,j,t,s,u,v}(k) - 1\right)$$

$$= \sum_{k\in K} \sum_{(i,j,t,s,u,v)\in A} c'_{i,j,t,s,u,v}(k) \cdot x_{i,j,t,s,u,v}(k) - \sum_{p\in P} \sum_{t\in H} \pi_{p,t} \cdot R_p - \sum_{(i',j')\in L} \sum_{\tau\in H} \rho_{i',j',\tau} \quad (11)$$

$$c'_{i,j,t,s,u,v}(k) = c_{i,j,t,s,u,v}(k) + \sum_{p\in P} \sum_{t\in H:(i,j,t,s,u,v)\in G_{p,t}} \pi_{p,t} \cdot e(i,j,t,s,u,v) + \sum_{(i',j')\in L} \sum_{\tau\in H:(i,j,t,s,u,v)\in\varphi(i',j',\tau)} \rho_{i',j',\tau} \quad (12)$$

3.3 Dynamic programming algorithm

As the proposed STS network has three dimensions, the search space of the joint timetabling and train control problem becomes extremely large. In order to find the optimal solution in reasonable time, it is important to design an efficient algorithm. We refer interested readers to a number of important studies on how to solve the time-dependent shortest path problem on a network with time-dependent arc costs (Ziliaskopoulos and Mahmassani, 1993; Chabini, 1998; Fischer and Helmberg, 2014). Our proposed DP method is essentially a state-dependent and time-dependent least cost algorithm where the speed is treated as an additional state dimension embedded in the path search process in time-expanded networks. In our study, a forward dynamic programming algorithm is designed to solve the large scale joint train timetabling and speed control problem. Let $\lambda_k(i,t,u)$ denote the label cost of train k at STS vertex (i,t,u). We also define LB^*, and UB^*, respectively, as the objective functions obtained from the best available lower bound/upper bound solutions. The proposed solution algorithm includes the following five steps.

Algorithm 1: DP algorithm based on Lagrangian relaxation

Input: STS network $G = (Q, A)$ built specifically for train k

Desired earliest departure vertex $(o_k, EDT_k, 0)$ and latest arrival vertex $(d_k, LAT_k, 0)$ of each train $k \in K$, where EDT_k and LAT_k denote the earliest departure time and latest arrival time of train k, repectively

Output: The least cost train trajectory in an STS network for each train $k \in K$ (i.e., train running speed profile and timetable)

Step 1. Initialization

Set label cost on each vertex $\lambda_k(i,t,u)$ to $+\infty$ and set its predecessor $pre_k(i,t,u)$ to $(-1,-1,-1)$ for each train $k \in K$

Set label cost on vertex $\lambda_k(o_k, EDT_k, 0) = 0$ for each train $k \in K$

Set the values of Lagrangian multipliers: $\pi_{p,t} = 0$ for all $p \in P, t \in T$, $\rho_{i',j',\tau} = 0$ for all $(i',j') \in L', \tau \in H$

Set Lagrangian iteration number $q = 0$, best lower bound $LB^* = -\infty$, and the best upper bound UB^* can be set to $+\infty$ or obtained from a published train timetable

Step 2. Label updating in dynamic programming

For train $k \in K$ **do**
For $t = EDT_k$ to LAT_k **do**
For each vertex $(i,t,u) \in V$ **do**
For each arc $(i,j,t,s,u,v) \in A$ **do**
If $\lambda_k(i,t,u) + c'_{i,j,t,s,u,v}(k) < \lambda_k(j,s,v)$
Then $\lambda_k(j,s,v) = \lambda_k(i,t,u) + c'_{i,j,t,s,u,v}(k)$
$pre_k(j,s,v) = (i,t,u)$
End // for each arc
End // for each vertex
End // for each time
End // for each train

Step 3. Obtaining the least cost STS train trajectory

Step 3.1: Obtain optimal DP solutions

For train $k \in K$ **do**

Find the sink vertex $(d_k, LAT_k, 0)$ and back trace to generate space-time trajectory and the solution results of variable $x_{i,j,t,s,u,v}(k)$

End // for each train

Step 3.2: Update the lower bound

By substituting the solution results that include $x_{i,j,t,s,u,v}(k)$ for all trains into the dualized problem, the lower bound LB^q is generated; if $LB^q > LB^*$, set $LB^* = LB^q$

Step 3.3: Update the upper bound

- Sort all trains in train set K based on a priority rule (e.g., a train having earlier earliest departure time EDT_k or latest arrival time LAT_k ranked as a higher priority)
- Set previously scheduled train set $PST = \emptyset$
- **For** each train $k \in K$, check the DP solution **do**

(1) Generate feasible solution space for train k

(1.1) Check headway conflicts with each train $k' \in PST$

If trains k and k' have conflicts with each other, for train k' with higher priority keep the same solution, mark the infeasible time vertex and infeasible arcs for the other trains in the restricted STS space, denoted as $RS(k)$, associated with the safety constraint. That is, in this $RS(k)$, the STS vertices/arcs violating time headway constraints are removed or disabled

(1.2) Check power supply conflicts for each power supply district $p \in P$, time $t \in H$

Based on the scheduled train set ST, calculate the cumulative power usage $R_t(p)$ at power supply district p time t, and obtain the available power that can be used by train k, denoted by $RS_t(p,k) = R_p - R_t(p)$

(2) Update train k's schedule

Then compute time-dependent STS least cost path and the cost by calling DP algorithm within the feasible solution space, which could lead to a shift in train k's departure time at the previous intermediate station or adjustment of incoming speed profiles in the restricted STS space $RS(k)$ considering available power $RS_t(p,k)$

(3) Add train k to the previously scheduled train set PST

End // for each train

- Update upper bound objective function UB^q

Compute the actual transportation costs along the path solution for vehicle k, update upper bound objective function UB^q

If $UB^q < UB^*$, set $UB^* = UB^q$

Step 3.4: Calculate solution gap

Find the relative solution gap between LB^* and UB^* by $\frac{UB^* - LB^*}{UB^*}$

If the gap = 0, an optimal solution is found and exit

Step 4. Updating the Lagrangian multipliers based on sub-gradient calculation

For each power supply district $p \in P$ at time $t \in H$ **do**

$$\pi_{p,t} = \pi_{p,t} + \theta^q\left(\left[\sum_{k \in K} \sum_{(i,j,t,s,u,v) \in G_{p,t}} e(i,j,t,s,u,v) \cdot x_{i,j,t,s,u,v}(k)\right] - R_p\right)$$

End // for each power supply district at each time

For each basic physical track segment $(i', j') \in L'$ at time $\tau \in H$

$$\rho_{i',j',\tau} = \rho_{i',j',\tau} + \theta^q\left(\left[\sum_{k \in K} \sum_{(i,j,t,s,u,v) \in \varphi(i',j',\tau)} x_{i,j,t,s,u,v}(k)\right] - 1\right)$$ **do**

End // for each basic physical track segment at each time

(θ^q is the Lagrangian sub-gradient step size at iteration q (Fisher, 1985; Ahuja et al., 1993)

Step 5: Termination condition test

If q is greater than a predetermined maximum value, terminate the algorithm; otherwise $q = q + 1$ and go back to **Step 2**.

In the lower bound solution, the complex constraints are only recognized through LR multipliers, so we do not expect the DP algorithm could resolve all conflicts, and in many cases, we still obtain possibly infeasible lower bound solutions. Note that, in the upper bound updating process in Step 3.3, by working on a partial complete schedule with previously scheduled trains, the DP algorithm is able to find the headway-feasible and energy-feasible train STS paths for a single train to be scheduled. Due to the nature of Lagrangian relaxation, we also do not expect the final lower bound is exactly equal to the best upper bound available, but this framework offers a systematic way to quantify the optimality gap, and to support other path search-based heuristic method development in the future.

3.4 Search space reduction

We now analyze the complexity of the dynamic programming algorithm of the label updating procedure. In this multi-loop framework, the algorithm first performs train selection operations for $|K|$ times and for each train it searches along the time, space and speed axes. Therefore, an apparent estimation for the worst computational time could be $O([|H| \times |N| \times |V|]^2)$ for each train. We need to recognize the following fact to further precisely estimate the complexity. The structure of STS transportation network leads to a very limited outgoing degree for each STS node. For instance, the change of time index from t to s along its outgoing arcs is simply defined as a constant number of options, namely, second-by-second travel time and dwell time. The change of space index and speed index is defined in the same way. As a result, instead of having a complicated quadratic form that considers all three dimension changes, the complexity in a real-life example should be $O(|H| \times |E|)$, where $|E|$ is the maximum outgoing degree of the STS nodes.

High-dimensional time-expanded networks are typically complex to build, and storing such a network externally as an algorithm input for a general-purpose optimization package such as CPLEX, could lead to prohibitive memory space requirements. The computational challenges for large-scale experiments can be tackled through a search region reduction process by adopting the space-time prism concept (Tang et al., 2016). In this space-time-speed prism based reduction framework, a potential STS trajectory area for a train consists of all accessible STS vertices and arcs that the train can reach when traveling from its desired earliest departure vertex to its latest arrival vertex. That is, an STS vertex (i, t, u) is within train k's accessible STS trajectory area if 1) train k can reach this vertex from its original vertex $(o_k, EDT_k, 0)$ and 2) train k can arrive at the destination vertex $(d_k, LAT_k, 0)$ from this vertex. By finding the feasible directed out-tree from the original vertex $(o_k, EDT_k, 0)$ and in-tree towards the destination vertex $(d_k, LAT_k, 0)$, we determine the forward energy label cost $\pi^F_{i,t,u}(k)$ and backward energy label cost $\pi^B_{i,t,u}(k)$. Accordingly, the set of possible vertices $V(k)$ and $A(k)$ for train k can be defined by the following inequalities and sets, where Θ denotes a large number of energy cost.

$$V(k) = \{(i,t,u) \in V \mid \pi^F_{i,t,u}(k) + \pi^B_{i,t,u}(k) < \Theta\} \quad (13)$$

$$A(k) = \{(i,j,t,s,u,v) \in A \mid (i,t,u) \in V(k) \; and \; (j,s,v) \in V(k)\} \quad (14)$$

The process of determining potential STS trajectory areas is illustrated in Algorithm 2.

Algorithm 2: Search region reduction algorithm based on STS feasible prism

Step 1. Initialization

Set $\pi^F_{i,t,u}(k) = +\infty, \; \pi^B_{i,t,u}(k) = +\infty$ for all $i \in N, t \in H, u \in V$

Set $V(k) = \emptyset, A(k) = \emptyset$

Set $\pi^F_{o_k,EDT_k,0}(k) = 0, \pi^B_{d_k,LAT_k,0}(k) = 0$

Set $V(k) = \{(o_k, EDT_k, 0), (d_k, LAT_k, 0)\}$

Step 2. Forward label updating

For each time $t \in [EDT(k), LAT(k)]$ **do**

For each STS arc $(i,j,t,s,u,v) \in A$ **do**

If $\pi^F_{j,s,v} > \pi^F_{i,t,u} + e(i,j,t,s,u,v)$

Then $\pi^F_{j,s,v} = \pi^F_{i,t,u} + e(i,j,t,s,u,v)$

End // for each STS arc

End // for each time

Step 3. Backward label updating

For each time $t \in [LAT(k), EDT(k)]$ **do**

For each STS arc $(i,j,t,s,u,v) \in A$ **do**

If $\pi^F_{i,t,u} > \pi^F_{j,s,v} + e(i,j,t,s,u,v)$

Then $\pi^F_{i,t,u} = \pi^F_{j,s,v} + e(i,j,t,s,u,v)$

End //for each STS arc

End // for each time

Step 4. Construct set of accessible STS vertices

For each STS vertex $(i,t) \in V$ **do**

If $\pi^F_{i,t,u} + \pi^B_{i,t,u} < \Theta$

Then $V(k) = V(k) \cup \{(i,t,u)\}$

End // for each vertex

Step 5. Construct set of accessible STS arcs

For each STS arc $(i,j,t,s,u,v) \in A$ **do**

If $(i,t,u) \in V(k)$ and $(j,s,v) \in V(k)$

Then $A(k) = A(k) \cup \{(i,j,t,s,u,v)\}$

End // for each STS arc

It is also important to understand that the proposed optimization framework can find an optimal solution only if the feasible solution region is non-empty. In contrast, if a feasible solution does not exist, with respect to the all defined constraints, the proposed dynamic programming based search algorithm can report empty solution as the ending boundary condition cannot be accessible for any space-time-speed feasible trajectories. To further reduce the search space, one can also adopt shooting heuristics, which have been widely used in the vehicle trajectory numerical optimization field. Interested readers in shooting heuristics are referred to the survey by Von Stryk and Bulirsch (1992), a dissertation by Tang (2012) on a full-scale simulation-based platform, and a recent study by Zhou et al. (2016) for emerging automated vehicle trajectory optimization applications.

4 Numerical experiments

4.1 Small-scale illustrative examples

We first illustrate the proposed space-time-speed grid network concept using a simple example on a small rail network (3 stations, 2 sections, and 1 power supply district). Consider a case where the safety headway between trains is set to 180 s (i.e., 3 minutes) and one power supply district can support a maximum of 10 energy units (i.e., 10 kwh). The discretized spacing along the space axis is 100 m, 10 sec along time axis, which further leads to a speed of 36 km/h. The algorithms described in this paper were implemented in C++ and the experiments were performed on a computer workstation running two Xeon E5-2680 processors clocked at 2.80 GHz with 20 cores and 192GB RAM running Windows Server 2008 x64 Edition.

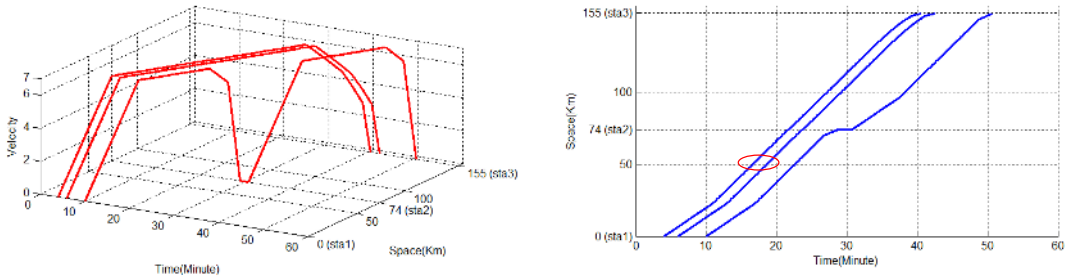
Table 6 presents four scenarios examined with various number of trains and different planned stops and departure time patterns. Specifically, in scenario I, three trains depart from station 1 with very tight time interval and trains 1 and 2 have the same origin-destination pairs with potentially overlapping time

windows. In this case, we can obtain energy-optimal timetables for individual trains, but need to resolve potential safety headway conflicts within their preferred time windows. In scenario II, train 2 may overtake train 1 at station 2 in order to get a better train trajectory. In scenario III, the power supply district constraint needs to be met by adjusting individual train trajectories/departure times, as serving four operational trains in one district may be impossible. Finally, in scenario IV with five trains, conflicts due to both the headway and power supply constraints should be considered.

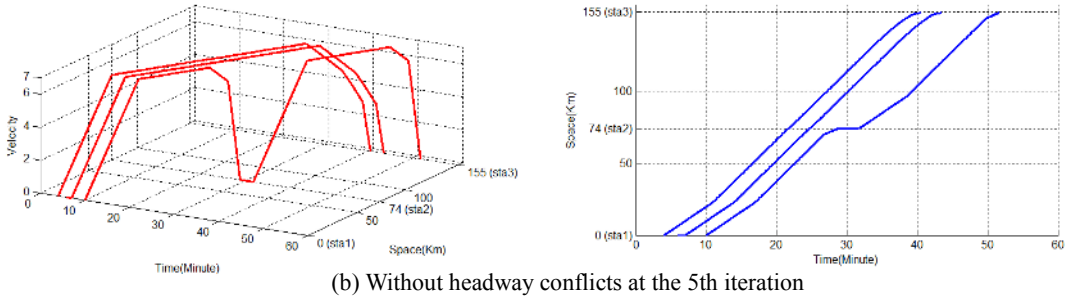
Table 6 Input of the three-station illustrative example including number of trains, pre-planned train stop pattern (Y denoting yes, N denoting no), and desired time windows (by minute)

Scenario	**I**	**II**	**III**	**IV**
Number of trains	3	3	4	5
Stop pattern of train 1	Y-N-Y	Y-Y-Y	Y-N-Y	Y-Y-Y
Stop pattern of train 2	Y-N-Y	Y-N-Y	Y-N-Y	Y-N-Y
Stop pattern of train 3	Y-Y-Y	Y-N-Y	Y-Y-Y	Y-N-Y
Stop pattern of train 4	-	-	Y-N-Y	Y-Y-Y
Stop pattern of train 5	-	-	-	Y-N-Y
EDT_k and $LAT_k$of train 1	4, 51	4, 58	4, 51	4, 58
EDT_k and LAT_k of train 2	6, 53	6, 52	8, 54	6, 52
EDT_k and LAT_k of train 3	10, 61	12, 57	12, 62	10, 56
EDT_k and LAT_k of train 4	-	-	20, 66	18, 69
EDT_k and LAT_k of train 5	-	-	-	25, 72
Departure time window of train 1	[4, 9]	[6, 9]	[6, 9]	[6, 9]
Departure time window of train 2	[6, 11]	[6, 11]	[8, 13]	[6, 11]
Departure time window of train 3	[10, 15]	[11, 16]	[12, 22]	[10, 15]
Departure time window of train 4	-	-	[20, 30]	[18, 28]
Departure time window of train 5	-	-	-	[24, 34]
Potential safety headway conflict	Y	Y	N	Y
Potential power supply conflict	N	N	Y	Y
Potential overtaking at station	N	Y	Y	Y

Fig 6-9 present train STS trajectories and timetables obtained through the proposed DP algorithm. In scenario I, the first iteration yields a headway conflict between trains 1 and 2 at a distance of 50 km along the track. This conflict is resolved through five iterations by delaying the departure time of train 2 so that headway constraints are satisfied. For scenario II, travel time in iteration 5 (compared to iteration 4) increases from 7120 to 7420 but the energy cost remains at 15991. The reason is that the cost here is only associated with mechanical power, as a result when a train stops/dwells at stations, there is no extra energy cost. In the last iteration, the departure time of the third train is postponed in order to avoid safety conflicts. As the train speed profiles do not change within the sections, the energy cost keeps the same as the value in iteration 4. For scenarios III and IV with power supply conflicts, the energy consumption value is also provided in figures 8 and 9, respectively. In the first iteration, the power consumption of both scenarios exceeds the capacity of the power supply district. By optimizing the train STS trajectories, the power consumption in the final solution falls below the available power threshold of 10 energy units throughout the time horizon. As a specific example, we obtain a feasible timetable for scenario IV, after 5 iterations, by trading off train 1's travel cost for the overall system cost, that is, letting it wait at station 2 for an extended time period. Table 7 and Fig 10 also demonstrate the computational efficiency and solution quality of our developed algorithm.

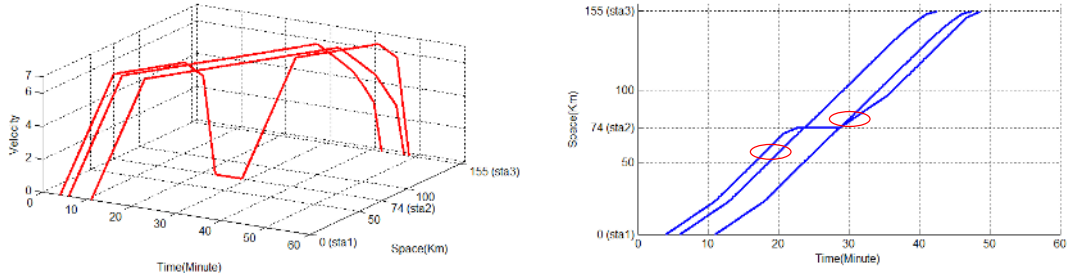


(a) With headway conflicts at the 1st iteration

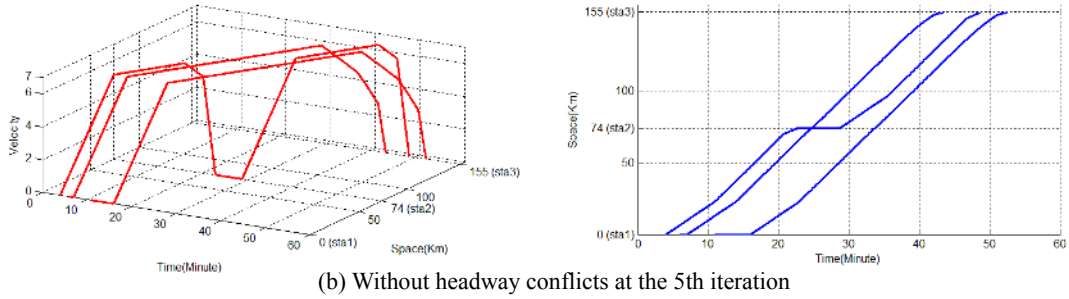


(b) Without headway conflicts at the 5th iteration

Fig 6. Train STS trajectories for scenario I

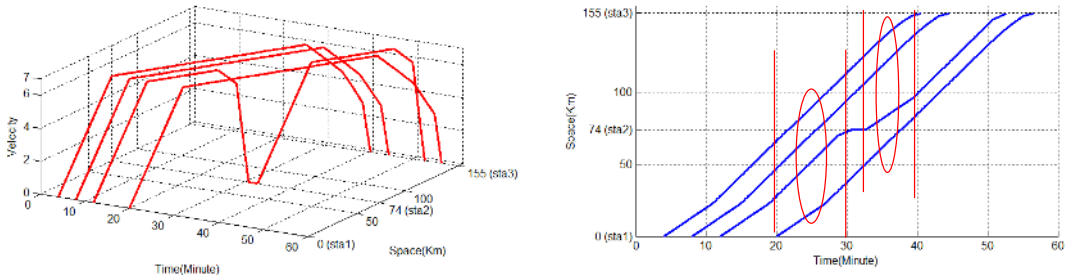


(a) With headway conflicts at the 1st iteration

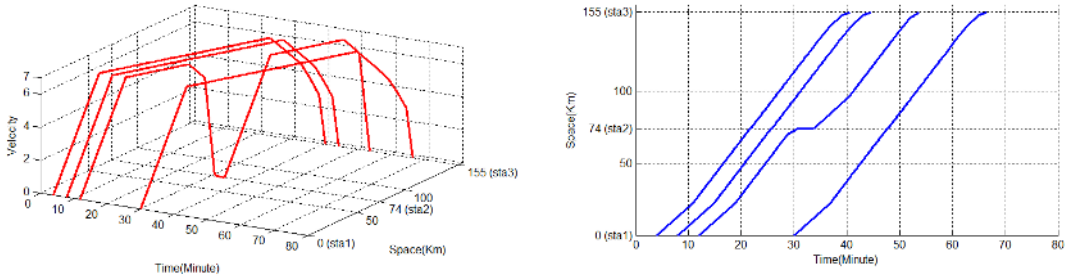


(b) Without headway conflicts at the 5th iteration

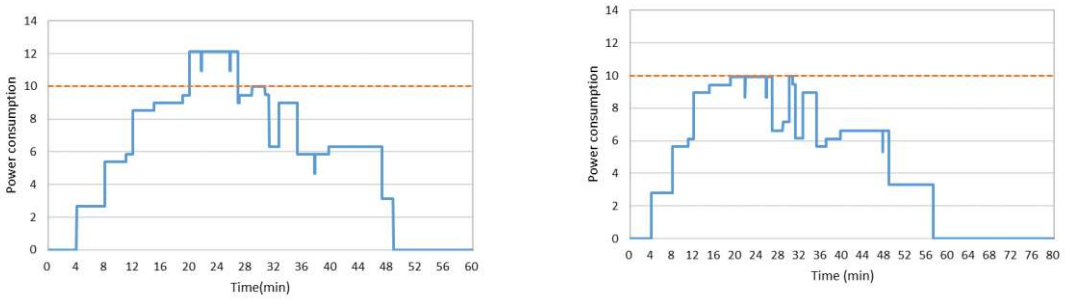
Fig 7. Train STS trajectories with overtaking for scenario II



(a) Train trajectories with power supply conflicts at the 1st iteration

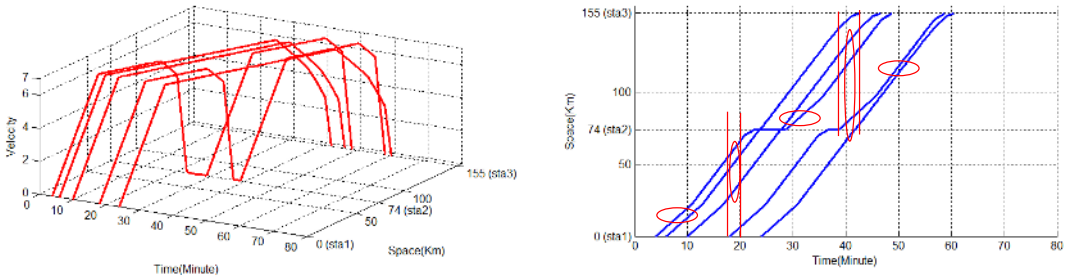


(b) Train trajectories without power supply conflicts at the 5th iteration (by shifting train departure time)

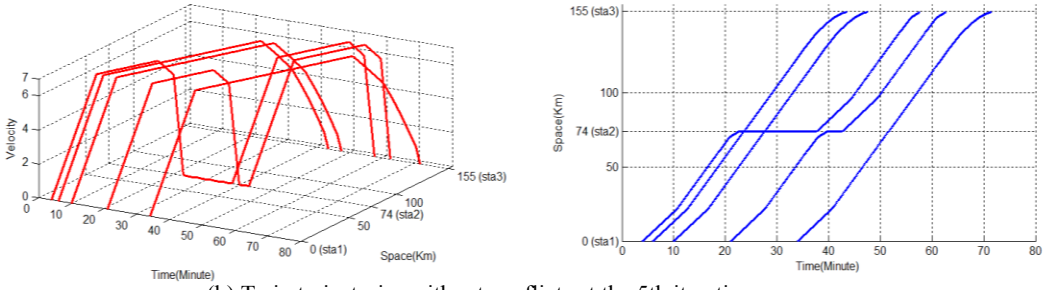


(c) Train power consumption (left: with conflicts at the 1st iteration; right: without conflicts at the 5th iteration)

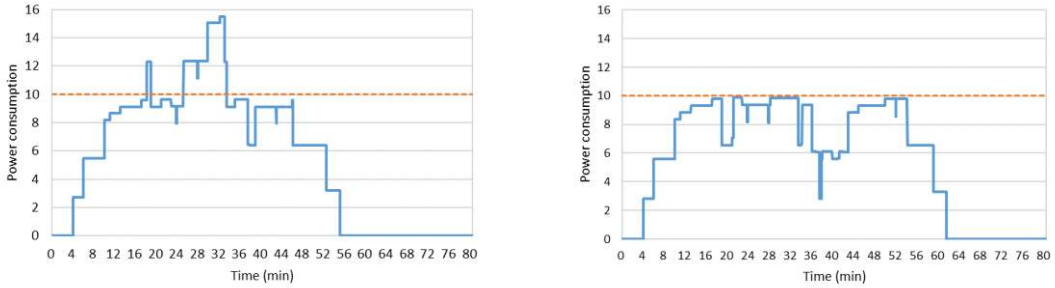
Fig 8. Train STS trajectories and energy consumption for scenario III



(a) Train trajectories with both headway and power supply conflicts at the 1st iteration



(b) Train trajectories without conflicts at the 5th iteration



(c) Train power consumption (left: with conflicts at the 1st iteration; right: without conflicts at the 5th iteration)

Fig 9. Train STS trajectories and energy consumption with overtaking for scenario IV

Table 7 Results of the 3-station railroad corridor on 4 different scenarios

Scenario	Iteration	LB^*	UB^*	Gap (%)	Total coasting time (s)	Power consumption (kwh)	Travel time (s)	CPU time (s)
I	1	22388	40000	44.0%	960	15568	6820	4.369
	2	22513	22513	0.0%	960	15573	6940	
	3	22513	22513	0.0%	960	15573	6940	
	4	22513	22513	0.0%	960	15573	6940	
	5	22513	22513	0.0%	960	15573	6940	
II	1	22633	40000	43.4%	980	15573	7060	4.398
	2	22885	23011	0.5%	980	15585	7300	
	3	22706	23011	1.3%	980	15586	7120	
	4	22711	23011	1.3%	980	15591	7120	
	5	23011	23011	0.0%	980	15591	7420	
III	1	29393	40000	26.5%	1440	20383	9010	6.039
	2	30405	30993	1.9%	1320	20795	9610	
	3	30593	30993	1.3%	1440	20983	9610	
	4	30993	30993	0.0%	1440	21383	9610	

	5	30993	30993	0.0%	1440	21383	9610	
IV	1	38016	40000	5.0%	1440	26326	11690	6.971
	2	38941	39619	1.7%	1320	26651	12290	
	3	39287	39619	0.8%	1120	26757	12530	
	4	39219	39619	1.0%	1320	26869	12350	
	5	39619	39619	0.0%	1440	26969	12650	

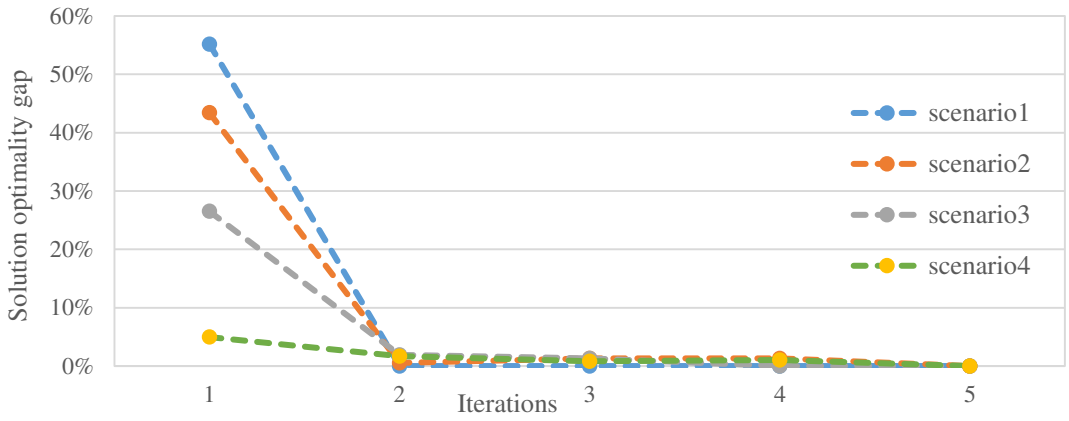


Fig 10. 5-iteration series of solution optimality gaps for the illustrative example

As our proposed model aims to represent train movements in a time-space-speed network on discretized states, it is critically important to understand the effect of finer discretization on output and potential impacts on different coarseness levels. For scenario IV, we further evaluate several finer resolutions of distance and time and obtain the computational efficiency and solution optimality, which have been presented in Table 8. Obviously, a finer space and/or time discretization leads to increased memory space and CPU time. For instance, a setting of 10 meters and 1 second uses about 188G RAM, and takes about 12 minutes to find optimal solutions (after 5 iterations). Compared to the base setting of 100 meters and 10 seconds, the discretization of the former scenario is about 10 times finer, while it requires approximately 100 times more memory and computational time. An extreme discretization case (1 meter and 1 second) results in insufficient memory with the existing computer configuration of 192 GB RAM. As it is relatively difficult to construct unified cost functions across different space-time-speed resolutions, comparison of solution quality improvement associated with a finer discretization is beyond the scope of this paper.

Table 8 RAM usage and CPU time consumption test for finding exact solution, for scenario IV with various resolutions of distance and time

Unit of space (meter)	Unit of time (second)	Unit of speed (km/h)	RAM usage (GB)	CPU time(s)
100	10	36	2.38	6.971
50	5	36	8.19	18.921
20	2	36	43.01	124.338
10	1	36	188.45	722.87
50	10	18	5.02	13.492
20	10	7.2	12.633	47.981
10	10	3.6	29.04	64.31
1	1	3.6	out of memory	--

By applying the search region reduction algorithm, the number of possible STS arcs for each train decreases by approximately 85% on average. Table 9 provides the results of search space reduction for four scenarios. The joint train trajectory and scheduling optimization model is also implemented in the general-purpose optimization package GAMS (Rosenthal, 2015). The sample GAMS codes for four scenarios can be downloaded at https://www.researchgate.net/publication/310328190_Sample_GAMS_codes_for_joint_optimization_of_high-speed_train_time-tables_and_speed_profiles.

Table 9 Search space reduction results based on space-time-speed prism

Scenario	Train	Time range	Original number of arcs	Number of arcs after reduction	Percentage of reduction
I	Train 1	[0,65]	59689	8650	85.51%
	Train 2			8763	85.32%
	Train 3			8542	85.69%
II	Train 1	[0,60]	59401	8748	85.27%
	Train 2			8890	85.03%
	Train 3			8632	85.47%
III	Train 1	[0,70]	79683	11718	85.29%
	Train 2			12093	84.82%
	Train 3			11689	85.33%
	Train 4			12585	84.21%
IV	Train 1	[0,75]	98521	15347	84.42%
	Train 2			14520	85.26%
	Train 3			14421	85.36%
	Train 4			13914	85.88%
	Train 5			14463	85.32%

4.2 Large-scale example using the Beijing-Shanghai high-speed rail corridor

The following large-scale example focuses on the Beijing-Shanghai (Jinghu) high-speed rail corridor, shown in Fig 11. This corridor is well established starting commercial train services on June 30, 2011 and has served more than 200 million passengers in April, 2014 (He et al., 2015). Fig 12 shows a simplified traction characteristic and the running resistance of trains used in our example, with the speed limit of 300 km/h and the power supply capacity of 50 mega volt amps (MVA) for each power supply district.

Fig 11. Spatial coverage of Beijing-Shanghai high-speed railroad corridor

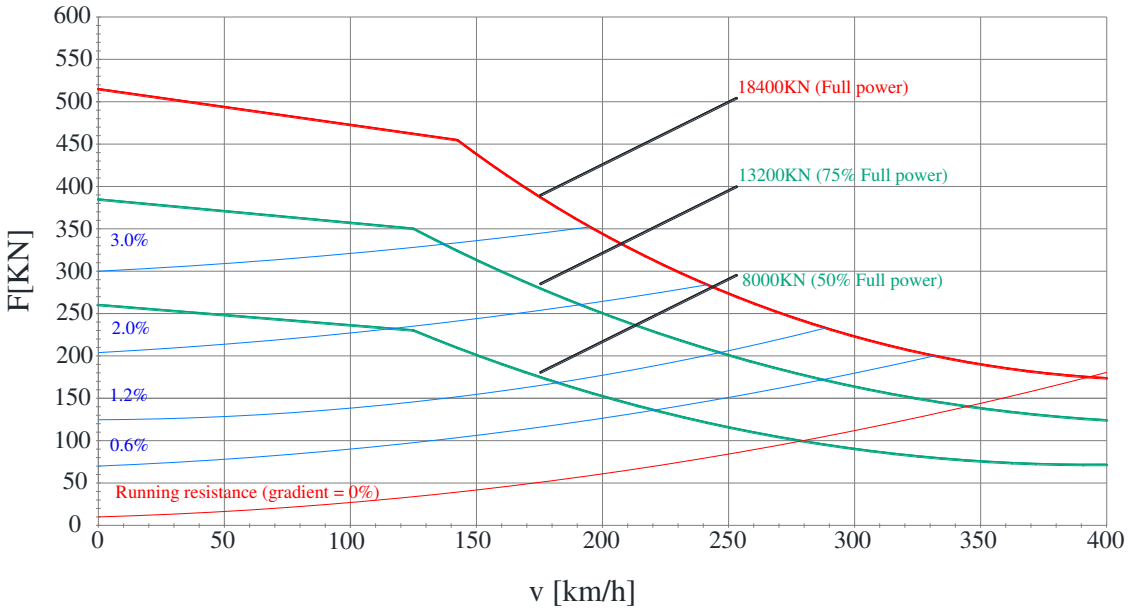


Fig 12. Traction characteristics of high-speed trains (from CRRC Corporation Limited, China)

We now consider the entire Beijing-Shanghai high-speed rail line that is approximately 1,300 km, has 23 passenger stations, powered by 26 substations. The corresponding real-world timetable shown in Fig. 12, specifies the train normal running time per segment and required dwell times for our optimization algorithm. A fully discretized space-time-speed grid network (i.e., 1 meter and 1 second) with the exact dynamic programming algorithm requires extremely demanding computational space and time. For this case, we need to reduce the search space and improve the search speed dramatically to achieve a solution

in a feasible timeframe and computational resources. An approximate dynamic programming method is adopted within a rolling horizon-based decomposition to optimize the STS trajectories by synchronizing train departure and dwell time. The overall CPU time for this heuristic-oriented algorithm (Tang, 2012) is about 2 hours for a total of 188 trains. An outline of this search algorithm is presented in Appendix B.

Given a real-world schedule of Beijing-Shanghai high-speed rail line shown in Fig 13, we generate an optimized real-world train timetable in Fig 14. The overall total energy consumption is reduced from 6,169.85 MWh to 5,429.47 MWh (with an equivalent energy-savings of about 12%).

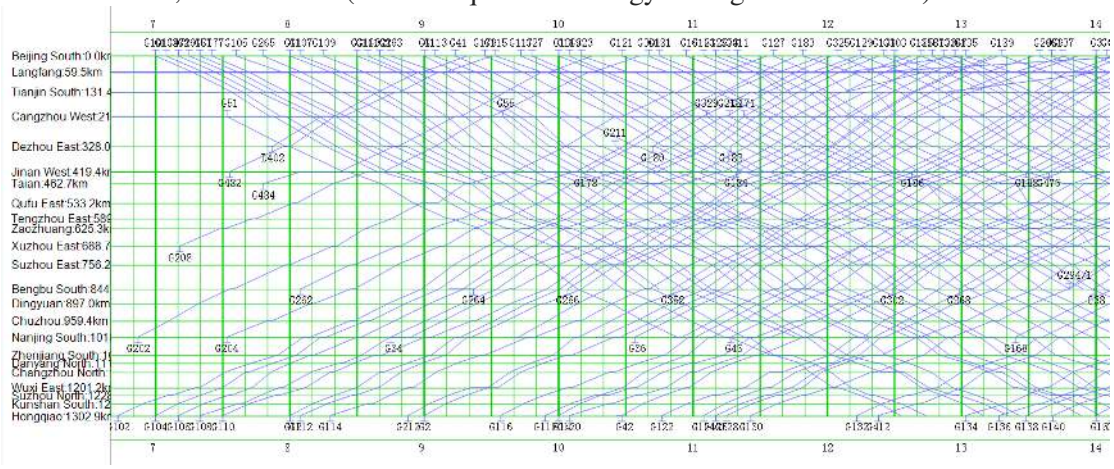


Fig 13. A real-world timetable of Beijing-Shanghai high-speed rail line

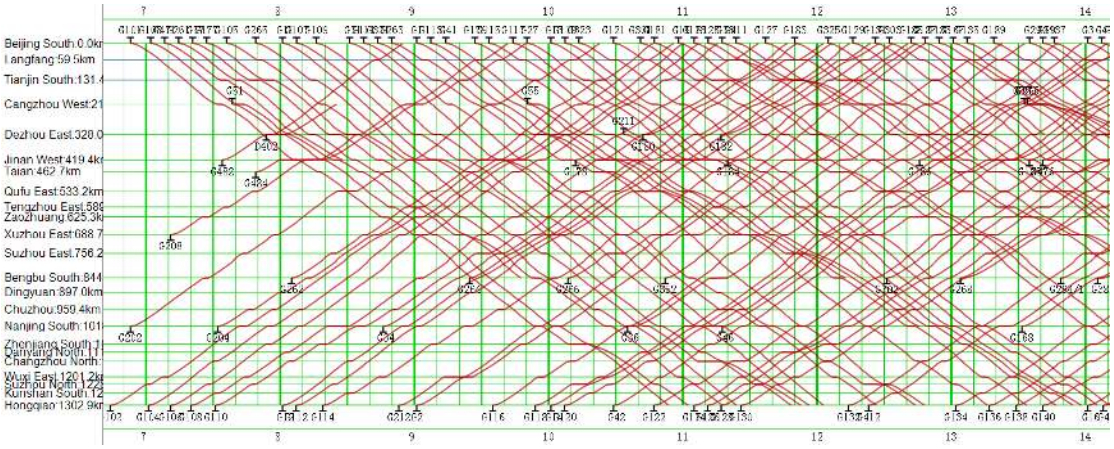


Fig 14. One optimized train schedule solution for Beijing-Shanghai high-speed rail line

Under different speed limit conditions, Table 10 details the per-train optimization changes in terms of end-to-end travel time (i.e., traveling speed) and used energy, with an energy saving range of 7.2% to 15.5%. The case "before optimization" is the real-world timetable used in year 2013. Using this real-world timetable as a starting point, we allow about 10% increases in trip travel times for each train, and then use the approximate DP algorithm to adjust 82 trains' departure times and smooth their driving trajectories to improve energy efficiency. It is relatively complicated to report average travel times as there are different OD pairs along the corridor with different travel distance, so we highlight the energy saving percentages and average speed change percentages in the last two columns of Table 10. We can observe an interesting tradeoff from this set of experiments: by allowing a slight reduction in average speed by 5%, we are able to reach a significant energy saving with an average of 12.5%. Future research needs to exploit the tradeoff curves between traveler time values and overall energy savings to obtain better Pareto optimal solutions for both system operators and users, as shown in a multi-criterial scheduling approach by Zhou and Zhong (2005).

It is important to note that different speed limit settings could be associated with different minimum or preferred cruising speed requirements. That is, the search space for a higher speed limit does not always contain the search space for a lower speed limit. It is possible that the solution with a higher speed limit could be worse than that with a lower speed limit, as they operate at overlapping but different feasible regions. What we could observe from Table 10 is that, with the increased average speed, energy consumption per train per kilometer also significantly increases, due to the fact that a higher speed limit could lead to significantly shorter end-to-end travel time and more energy consumption demands.

Table 10 Per-train measure of effectiveness for entire corridor test case under different speed limit conditions

Speed limit (km/h)	**Before Optimization**		**After Optimization**		**Percentage decrease of average speed**	**Percentage reduction of energy consumption**
	Average speed (km/h)	Energy consumption per train per kilometer (kwh)	Average speed(km/h)	Energy consumption per train per kilometer (kwh)		
200	178.5	35.8	163.6	30.8	8.3%	13.9%
220	192.7	40.1	182.3	33.9	5.4%	15.5%
240	223.1	43.1	212.8	37.4	4.6%	13.2%
260	244.3	46.5	242.5	41.7	0.7%	10.2%
280	263.7	50.2	251.6	46.6	4.6%	7.2%
300	286.6	62.8	270.8	55.1	5.5%	12.3%

5 Conclusions and future research

Many recent research aims to address the important problem of joint train trajectory and timetable optimization method to reduce system-wide energy consumption with sufficient flexibility within the schedule constraints. This paper presents a carefully discretized space-time-speed network framework to characterize both train timetables and speed profiles simultaneously at different space and time resolutions. To the best of our knowledge, there are a few prior studies using space-time-speed constructs to illustrate the importance of tracking vehicle location, movements, and speed, jointly. For example, Kuijpers et al. (2011) proposed kinetic space-time prisms to consider acceleration limits on the top of classical space-time prisms. Zhao et al. (2015) used an interesting 3-dimensional continuous space-time-speed representation to graphically illustrate the detailed time-step based vehicle movement simulation. In our research, by fully discretizing the speed levels and synchronizing the entering and leaving speed levels u and v according to the space-time grid network, we are able to enable a feasible high-dimensional search path algorithm, through Lagrangian relaxation, for timetabling and trajectory optimization applications. This network construction method could offer a good unified representation scheme for simultaneous optimization of train timetables and energy consumption in a railway network.

Through using a dynamic programming based algorithm and priority rule-based upper generation methods, we tested the proposed approach on a number of small-scale hypothetic examples. This study also adapts the proposed STS network search framework to construct and evaluate a heuristic approximation DP algorithm in a case study for the Beijing-Shanghai high-speed rail line. The simulation results demonstrate that the obtainable energy consumption saving reaches a range of 7% to 15%.

There are a few important remarks about possible extensions of the proposed research. First, the trajectories to be optimized in our proposed space-time-speed grid network can be mapped to both speed profile and space-time dimensions, and the space-time network can be systematically extended from a single corridor to n-track or multi-route cases. Thus, a future research extension could be how to construct a joint optimization model with train routes, timetable and speed-based trajectories, so that one can improve solution optimality with additional feasibility. To consider moving block signaling, it is also possible to further define a dynamic headway of moving blocks based on the state variables (distance and speed) defined in the STS network, but handling of the dynamic headway is a very challenging modeling issue by its own. Interested readers can find a recent paper by Fu and Dessouky (2016) on optimization models and algorithms for dynamic headway control.

The proposed space-time-speed network modeling framework can be also extended to consider as an enhancement for many existing studies on the green vehicle routing problem, e.g., Bektaş and Laporte (2011), Dekker et al. (2012) and Fukasawa et al. (2015). Specifically, our proposed framework offers a flexible way to explicitly model both speed and acceleration as major energy consumption factors, while the acceleration rate is represented as a derivative of the speed changes in the STS network. One can also consider heterogeneous traffic with different types of trains with distinctive driving characteristics, and each train type corresponds to its own STS graph with different configurations of traveling arcs, speed limits and acceleration constraints. In this case with different types of trains, we might need to carefully design the space-time-speed discretization scheme so that the headway and priority constraints between trains can be matched across networks with different resolutions. Our constructed STS grid network can be also viewed as a sophisticated adaption of the broader state-space-time network-based modeling framework proposed by Mahmoudi and Zhou (2016) for the vehicle routing problem with vehicle carrying capacity with time window constraints.

In our future research, we are interested in developing different effective heuristics to further reduce the search space and memory requirement for maintaining high-dimensional networks in the memory, for example, using a hash table-based data structure or using a continuous or semi-discretized time-expanded network representation (Boland et al, 2015). A more practically useful integrated optimization process should consider the complicated energy consumption calculations and constraints, different stochastic disturbances, and other realistic train operating scenarios. In this paper, by taking a simplified set of assumptions about energy consumption, the developed STS based approach aims to provide important theoretical insights on the joint optimization of high-speed train timetables and speed profiles. Our future research needs to refine the value function to better resolve the complicated energy consumption calculation and other practical constraints. The measure of timetable robustness has not been integrated into the proposed model in this study. We can improve the reliability of high-speed train timetables and speed profiles by considering buffer time and probabilistic delay propagation together with the high-resolution STS network. Currently, our research teams are also working on extending the STS modeling framework to better control and smooth speed changes in emerging applications of autonomous vehicle routing, platooning in highway and urban arterial networks (Li et al. 2015; Lu et al. 2016; Ruan et al. 2016).

Acknowledgements

This research is supported by China Ministry of Transport Construction Project in Science and Technology (No. 2015318223010), Fundamental Research Funds for the Central Universities (No. 2015LBZ011) and China Railway Corporation (No. 2016X005-D). The third author is partially supported by National High Technology Research and Development Program of China (No.2015AA016404). The last author is partially funded by National Science Foundation–United States under NSF Grant No. CMMI 1538105 "Collaborative Research: Improving Spatial Observability of Dynamic Traffic Systems through Active Mobile Sensor Networks and Crowdsourced Data". The authors would like to thank Dr. Jeffrey Stempihar at Arizona State University for his valuable comments.

Appendix A: Calculation of traction power supply in real-world applications

The practically useful integrated train scheduling and traction power supply optimization program involves a large number of parameters across different knowledge domains. Based on terminologies from power engineering, Fig A1 and Table A1 below list a set of important parameters and briefly describe matrix-oriented equations in calculating traction power supply in real-world applications.

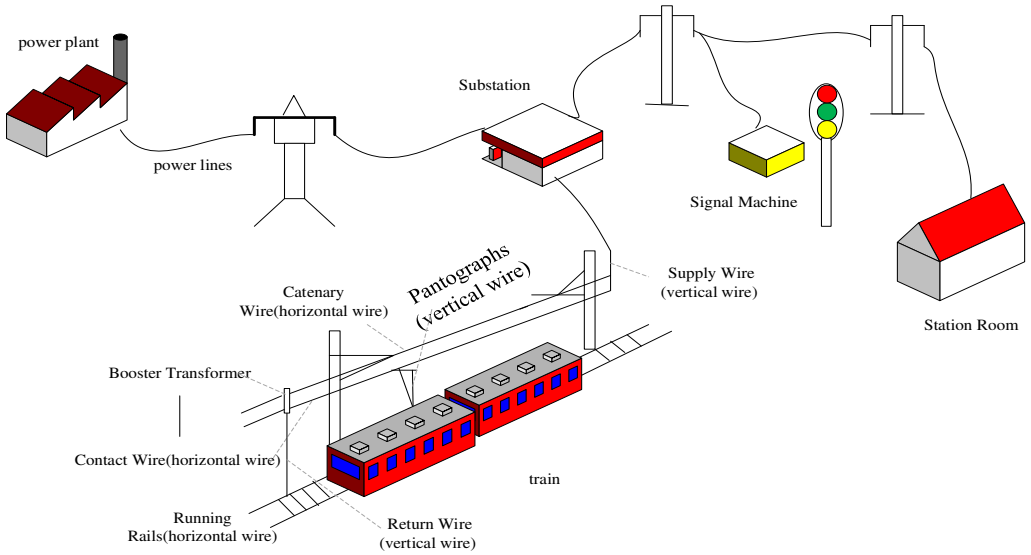


Fig A1. Electric power system to electrified railway power supply

Table A1 Definition of symbols in high-speed rail power supply system

Symbol	Definition
$\boldsymbol{i, j}$	Indices of horizontal wires (include catenary wire, contact wire, running rails)
$\boldsymbol{k}$	Index of vertical wires (include pantographs, substation supply wire, return wire)
$\boldsymbol{m}$	Number of horizontal wires
$\boldsymbol{n}$	Number of vertical wires
$\boldsymbol{z_k^{ij}}$	Resistance per unit length, if $(i = j)$, wire $i$is self-resistance, if $(i \neq j)$, there is a mutual resistance between wires $i$and j
$[\boldsymbol{Y}]_{\boldsymbol{k}}$	Admittance matrix of vertical wire k, $= \{Y_k^1, Y_k^2, \cdots, Y_k^m\}$
$[\boldsymbol{Z}]_{\boldsymbol{k}}$	Resistance matrix of vertical wire k, $= \{Z_k^1, Z_k^2, \cdots, Z_k^m\}$
$[\boldsymbol{\psi}]_{\boldsymbol{k}}$	Current vectors of vertical wire k, $= \{\psi_k^1, \psi_k^2, \cdots, \psi_k^m\}$
$[\boldsymbol{U}]_{\boldsymbol{k}}$	Voltage vectors of vertical wire k, $= \{U_k^1, U_k^2, \cdots, U_k^m\}$

First, a multi-conductor model is adapted to consider different traction net distribution forms, and a parallel transmission line theory can be used to calculate variable length and precision of the traction net impedance. Specifically, the unit distance traction net impedance needs to be represented as a $m \times m$ complex matrix. Fig A2 shows a multi-conductor circuit net which is a transformation form of the power supply system in Fig A1.

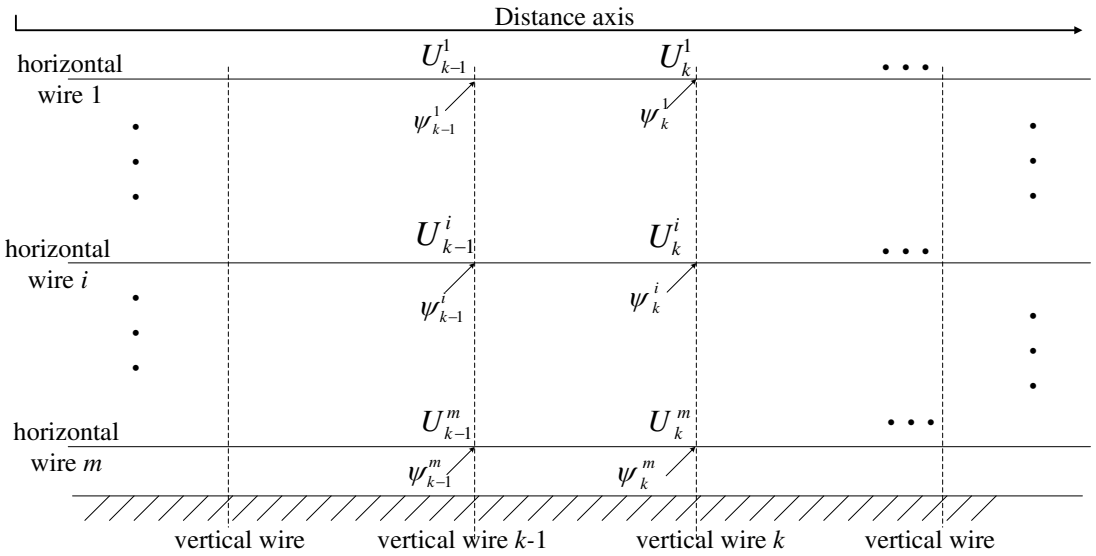


Fig A2. Multi-conductor traction net

Based on a multi-conductor traction net, a current analytical model is used to calculate train catenary voltage. We further transform a multi-conductor traction net into a dynamic abstract network. Based on a node voltage method, this method can quickly find solutions for the traction power supply system. Interested readers can refer to the classical paper by Carson (1926) and more recent textbooks such as Grainger and Stevenson (1994) to calculate and update matrices $[Z]_k$ and $[Y]_k$ using Eqs. (A1) and (A2).

$$[Z]_k = \begin{bmatrix} z_k^{11} & \cdots & z_k^{1m} \\ \vdots & \ddots & \vdots \\ z_k^{m1} & \cdots & z_k^{mm} \end{bmatrix}_{m\times m}, \quad [Y]_k = \begin{bmatrix} y_k^{11} & \cdots & y_k^{1m} \\ \vdots & \ddots & \vdots \\ y_k^{m1} & \cdots & y_k^{mm} \end{bmatrix}_{m\times m} \tag{A1}$$

$$\begin{bmatrix} [Y]_1+[Z]_1^{-1} & [Z]_1^{-1} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ [Z]_1^{-1} & [Z]_1^{-1}+[Y]_2+[Z]_2^{-1} & -Z_2^{-1} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & -Z_2^{-1} & [Z]_2^{-1}+[Y]_3+[Z]_3^{-1} & -[Z]_3^{-1} & 0 & \cdots & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot & \cdot & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & -[Z]_{N-2}^{-1} & [Z]_{N-2}^{-1}+[Y]_{N-1}+[Z]_{N-1}^{-1} & -[Z]_{N-1}^{-1} \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & -[Z]_{N-1}^{-1} & [Z]_{N-1}^{-1}+[Y]_N \end{bmatrix} \begin{bmatrix} [U]_1 \\ [U]_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ [U]_{N-1} \\ [U]_N \end{bmatrix} = \begin{bmatrix} [\psi]_1 \\ [\psi]_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ [\psi]_{N-1} \\ [\psi]_N \end{bmatrix} \tag{A2}$$

Appendix B: Approximate dynamic programming algorithm

For $t = 0$ to T **do**

For train $k \in K$ **do**

Step 1: Select possible state vertex from promising set at time t

Step 2: Cost label updating based on forward dynamic programming

$C_{k,s,v,g'} = \min\{c_{i,j,t,s,u,v,g,g'}(k) + C_{k,t,u,g}\}$ for promising or non-dominated vertex set $\Omega_t = \{k, t, u, g\}$ and possible gear level actions

Step 3: Applying priority rules for ensuring safety headway and power supply limit across multiple trains

If safety conflict or energy supply conflict exists, apply priority rule-based heuristics to adjustment promising vertex state set in Ω_t

Step 4: Check next available state for each train to ensure a safe stop at next station along fuel-efficient trajectories

Apply single-shooting rules to predict or simulate the train trajectories from state

vertex k, t', u, g to the end of power supply district, if the next available state $\{k, t', u, g\}$ is infeasible, then adjust this state by either changing the driving speed or adjusting to a different possible gear level.

Step 5: Size reduction for promising state vertex

Reduce the size of $\Omega_{t'=t+1}$ for the next time step t'

End // for each train

End // for each time

References

Ahuja, R.K., Magnanti, T.L. and Orlin, J.B., 1993. Network flows: theory, algorithms, and applications.

Albrecht, A., Howlett, P., Pudney, P. and Vu, X., 2011, June. Optimal train control: analysis of a new local optimization principle. In *Proceedings of the 2011 American Control Conference*, pp. 1928-1933. IEEE.

Albrecht, A.R., Howlett, P.G., Pudney, P.J. and Vu, X., 2013. Energy-efficient train control: from local convexity to global optimization and uniqueness. *Automatica*, *49*(10), pp.3072-3078.

Bai, Y., Mao, B., Zhou, F., Ding, Y. and Dong, C., 2009. Energy-efficient driving strategy for freight trains based on power consumption analysis. *Journal of Transportation Systems Engineering and Information Technology*, *9*(3), pp.43-50.

Bektaş, T. and Laporte, G., 2011. The pollution-routing problem.*Transportation Research Part B: Methodological*, *45*(8), pp.1232-1250.

Benjamin, B.R., Milroy, I.P. and Pudney, P.J., 1989. Energy-efficient operation of long-haul trains. In *Fourth International Heavy Haul Railway Conference 1989: Railways in Action; Preprints of Papers*Institution of Engineers, Australia.

Besinovic, N., Roberti, R., Quaglietta, E., Cacchiani, V., Toth, P. and Goverde, R.M., 2015. Micro-macro approach to robust timetabling. In *6th International Conference on Railway Operations Modelling and Analysis–RailTokyo2015, Narashino, Japan*.

Bocharnikov, Y.V., Tobias, A.M. and Roberts, C., 2010, April. Reduction of train and net energy consumption using genetic algorithms for trajectory optimisation. In *Railway Traction Systems (RTS 2010),* pp. 1-5. IET.

Boland, N., Hewitt, M., Marshall, L. and Savelsbergh, M., 2015. *The continuous time service network design problem*. Technical report, Technical report, Optimization on-line.

Cacchiani, Valentina, Alberto Caprara, and Matteo Fischetti. "A Lagrangian heuristic for robustness, with an application to train timetabling."*Transportation Science* 46.1 (2012): 124-133.

Cacchiani, V. and Toth, P., 2012. Nominal and robust train timetabling problems. *European Journal of Operational Research*, *219*(3), pp.727-737.

Caprara, A., Fischetti, M. and Toth, P., 2002. Modeling and solving the train timetabling problem. *Operations research*, *50*(5), pp.851-861.

Carson, J.R., 1926. Wave propagation in overhead wires with ground return.*Bell system technical journal*, *5*(4), pp.539-554.

Chabini, I., 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), pp.170-175.

Cheng, J., Davydova, Y., Howlett, P. and Pudney, P., 1999. Optimal driving strategies for a train journey with non-zero track gradient and speed limits.*IMA Journal of Management Mathematics*, *10*(2), pp.89-115.

Cheng, J. and Howlett, P., 1992. Application of critical velocities to the minimisation of fuel consumption in the control of trains. *Automatica*, *28*(1), pp.165-169.

Cheng, J. and Howlett, P., 1993. A note on the calculation of optimal strategies for the minimization of fuel

consumption in the control of trains. *IEEE Transactions on Automatic Control*, *38*(11), pp.1730-1734.
Chester, M. and Horvath, A., 2012. High-speed rail with emerging automobiles and aircraft can reduce environmental impacts in California's future. *Environmental Research Letters*, *7*(3), p.034012.
Cordeau, J.F., Toth, P. and Vigo, D., 1998. A survey of optimization models for train routing and scheduling. *Transportation science*, *32*(4), pp.380-404.
Corman, F. and Meng, L., 2013, August. A review of online dynamic models and algorithms for railway traffic control. In *Intelligent Rail Transportation (ICIRT), 2013 IEEE International Conference on* (pp. 128-133). IEEE.
Corman, F., D'Ariano, A., Pacciarelli, D. and Pranzo, M., 2010. A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B: Methodological*, *44*(1), pp.175-192.
D'Ariano, A., Corman, F., Pacciarelli, D. and Pranzo, M., 2008. Reordering and local rerouting strategies to manage train traffic in real time. *Transportation Science*, *42*(4), pp.405-419.
D'ariano, A., Pacciarelli, D. and Pranzo, M., 2007. A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research*, *183*(2), pp.643-657.
Daganzo, C.F., 2006. In traffic flow, cellular automata= kinematic waves. *Transportation Research Part B: Methodological*, *40*(5), pp.396-403.
Dekker, R., Bloemhof, J. and Mallidis, I., 2012. Operations Research for green logistics–An overview of aspects, issues, contributions and challenges. *European Journal of Operational Research*, *219*(3), pp.671-679.
Domínguez, M., Fernández-Cardador, A., Cucala, A.P. and Pecharromán, R.R., 2012. Energy savings in metropolitan railway substations through regenerative energy recovery and optimal design of ATO speed profiles.*IEEE transactions on automation science and engineering*, *9*(3), pp.496-504.
Dorfman, M.J. and Medanic, J., 2004. Scheduling trains on a railway network using a discrete event model of railway traffic. *Transportation Research Part B: Methodological*, *38*(1), pp.81-98.
Fisher, M.L., 1985. An applications oriented guide to Lagrangian relaxation. *Interfaces*, *15*(2), pp.10-21.
Fischer, F. and Helmberg, C., 2014. Dynamic graph generation for the shortest path problem in time expanded networks. *Mathematical Programming*, *143*(1-2), pp.257-297.
Fu, L. and Dessouky, M., 2016. Models and Algorithms for Dynamic Headway Control. Working Paper. Accessible at http://www-bcf.usc.edu/~maged/publications/its_train1.pdf.
Fukasawa, R., He, Q. and Song, Y., 2015. A branch-cut-and-price algorithm for the energy minimization vehicle routing problem. *Transportation Science*,*50*(1), pp.23-34.
Goverde, R.M., Bešinović, N., Binder, A., Cacchiani, V., Quaglietta, E., Roberti, R. and Toth, P., 2016. A three-level framework for performance-based railway timetabling. *Transportation Research Part C: Emerging Technologies*, *67*, pp.62-83.
Grainger, J.J. and Stevenson, W.D., 1994. *Power system analysis*. McGraw-Hill.
He, G., Mol, A.P., Zhang, L. and Lu, Y., 2015. Environmental risks of high-speed railway in China: public participation, perception and trust. *Environmental Development*, *14*, pp.37-52.
Howlett, P., 1996. Optimal strategies for the control of a train. *Automatica*, *32*(4), pp.519-532.
Howlett, P.G. and Cheng, J., 1997. Optimal driving strategies for a train on a track with continuously varying gradient. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, *38*(03), pp.388-410.
Howlett, P.G., Pudney, P.J. and Vu, X., 2009. Local energy minimization in optimal train control. *Automatica*, *45*(11), pp.2692-2698.
Kushner, H. and Dupuis, P.G., 2013. *Numerical methods for stochastic control problems in continuous time*, 24. Springer Science & Business Media.
Kuijpers, B., Miller, H.J. and Othman, W., 2011, November. Kinetic space-time prisms. In *Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems*, pp. 162-170. ACM.
Li, P., Mirchandani, P. and Zhou, X., 2015. Solving simultaneous route guidance and traffic signal optimization problem using space-phase-time hypernetwork. Transportation Research Part B: Methodological, 81, pp.103-130.

Li, X. and Lo, H.K., 2014a. An energy-efficient scheduling and speed control approach for metro rail operations. *Transportation Research Part B: Methodological*, *64*, pp.73-89.

Li, X. and Lo, H.K., 2014b. Energy minimization in dynamic train scheduling and control for metro rail operations. *Transportation Research Part B: Methodological*, *70*, pp.269-284.

Liu, R.R. and Golovitcher, I.M., 2003. Energy-efficient operation of rail vehicles. *Transportation Research Part A: Policy and Practice*, *37*(10), pp.917-932.

Lu, G., Zhou, X., Peng, Q., He, B., Mahmoudi, M. and Zhao, J., 2016. Solving Resource Recharging Station Location-routing Problem through a Resource-space-time Network Representation. *arXiv preprint arXiv:1602.06889*.

Lusby, R.M., Larsen, J., Ehrgott, M. and Ryan, D., 2011. Railway track allocation: models and methods. *OR spectrum*, *33*(4), pp.843-883.

Mahmoudi, M. and Zhou, X., 2016. Finding optimal solutions for vehicle routing problem with pickup and delivery services with time windows: A dynamic programming approach based on state–space–time network representations. *Transportation Research Part B: Methodological*, *89*, pp.19-42.

Medanic, J. and Dorfman, M.J., 2002. Efficient scheduling of traffic on a railway line. *Journal of optimization theory and applications*, *115*(3), pp.587-602.

Meng, L. and Zhou, X., 2014. Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables. *Transportation Research Part B: Methodological*, *67*, pp.208-234.

Milroy, I.P., 1980. *Aspects of automatic train control* (Doctoral dissertation, © Ian Peter Milroy).

Miyatake, M., Haga, H. and Suzuki, S., 2009, September. Optimal speed control of a train with on-board energy storage for minimum energy consumption in catenary free operation. In *Power Electronics and Applications, 2009. EPE'09.* pp. 1-9. IEEE.

Munos, R. and Moore, A., 2002. Variable resolution discretization in optimal control. *Machine learning*, *49*(2-3), pp.291-323.

Powell, W.B., Jaillet, P. and Odoni, A., 1995. Stochastic and dynamic networks and routing. *Handbooks in operations research and management science*, *8*, pp.141-295.

Rodrigo, E., Tapia, S., Mera, J.M. and Soler, M., 2013. Optimizing electric rail energy consumption using the lagrange multiplier technique. *Journal of Transportation Engineering*, *139*(3), pp.321-329.

Rosenthal, R., 2015. GAMS: A User's Guide. GAMS Development Corporation.

Ruan, J.M., Liu, B., Wei, H., Qu, Y., Zhu, N. and Zhou, X., 2016. How Many and Where to Locate Parking Lots? A Space–time Accessibility-Maximization Modeling Framework for Special Event Traffic Management. *Urban Rail Transit*, pp.1-12.

Samà, M., Pellegrini, P., D'Ariano, A., Rodriguez, J. and Pacciarelli, D., 2016. Ant colony optimization for the real-time train routing selection problem. *Transportation Research Part B: Methodological*, *85*, pp.89-108.

Su, S., Li, X., Tang, T. and Gao, Z., 2013. A subway train timetable optimization approach based on energy-efficient operation strategy. *IEEE Transactions on Intelligent Transportation Systems*, *14*(2), pp.883-893.

Tang, J. (2012) Study on Global Information Simulation Optimization Method and System of High-speed Train Group Operation Process Based on Train Timetable. Dissertation. Beijing Jiaotong University, China. In Chinese.

Tang, J., Song, Y., Miller, H.J. and Zhou, X., 2016. Estimating the most likely space–time paths, dwell times and path uncertainties from vehicle trajectory data: A time geographic method. *Transportation Research Part C: Emerging Technologies*, *66*, pp.176-194.

Törnquist, J. and Persson, J.A., 2007. N-tracked railway traffic re-scheduling during disturbances. *Transportation Research Part B: Methodological*, *41*(3), pp.342-362.

Von Stryk, O. and Bulirsch, R., 1992. Direct and indirect methods for trajectory optimization. *Annals of operations research*, *37*(1), pp.357-373.

Wong, K.K. and Ho, T.K., 2003, December. Coast control of train movement with genetic algorithm. In *Evolutionary Computation, 2003. CEC'03.* Vol. 2, pp. 1280-1287. IEEE.

Yang, L., Li, K., Gao, Z. and Li, X., 2012. Optimizing trains movement on a railway network. *Omega*, *40*(5), pp.619-633.

Yang, L. and Zhou, X., 2014. Constraint reformulation and a Lagrangian relaxation-based solution algorithm for a least expected time path problem. *Transportation Research Part B: Methodological*, *59*, pp.22-44.

Yang, L., Li, S., Gao, Y. and Gao, Z., 2015. A Coordinated Routing Model with Optimized Velocity for Train Scheduling on a Single-Track Railway Line.*International Journal of Intelligent Systems*, *30*(1), pp.3-22.

Yue, Y., Wang, S., Zhou, L., Tong, L. and Saat, M.R., 2016. Optimizing train stopping patterns and schedules for high-speed passenger rail corridors.*Transportation Research Part C: Emerging Technologies*, *63*, pp.126-146.

Zhao, N., Roberts, C., Hillmansen, S. and Nicholson, G., 2015. A multiple train trajectory optimization to minimize energy consumption and delay. *IEEE Transactions on Intelligent Transportation Systems*, *16*(5), pp.2363-2372.

Zhou, F., Li, X. and Ma, J., 2015. Parsimonious shooting heuristic for trajectory control of connected automated traffic part I: Theoretical analysis with generalized time geography. *arXiv preprint arXiv:1511.04810*.

Zhou, L., Hu, S., Ma, J., & Yue, Y., 1998. Network hierarchy parallel algorithm of automatic train scheduling. *Journal of the China Railway society*, *5*, pp.15-21. In Chinese.

Zhou, X. and Zhong, M., 2005. Bicriteria train scheduling for high-speed passenger railroad planning applications. *European Journal of Operational Research*, *167*(3), pp.752-771.

Ziliaskopoulos, A.K. and Mahmassani, H.S., 1993. Time-dependent, shortest-path algorithm for real-time intelligent vehicle highway system applications. *Transportation research record*, pp.94-94.