Exploring the Personalized Journey Planning Problem in the Timetable-scheduled Railway Network

Jiamin Zhang*

College of Transportation, Shandong University of Science and Technology, Qingdao 266590, China

*zjmsunrise@126.com

Abstract

It is recognized that finding the optimal itinerary or journey plan for a traveler/passenger using public transport (including the railway transportation) is an algorithmic challenge. As a key component of transport service information, the journey planning problem has recently gained interest for the large-scale high-speed railway network in China and the development of MaaS (Mobility as a Service) system. The itinerary planning is considered as the one-origin-node to one-destination-node journey planning problem in our study. At first, the personalized door-to-door travel costs for each O-D pair query on the basis of the detailed requirements analysis are formulated. Secondly, the digraph of the train service network based on the timetable are descripted. Thirdly, two-phase solutions for the journey planning problem in the railway network are proposed. Fourthly, a discussion on the proposed methods are represented. Future research should focus on the more flexible journey planning problem for non-transfer scenarios under the condition of virtual coupling system and dynamic coupling system. And the large language model (LLM) in the artificial intelligence field can be employed to solve the complex joureny planning problem in the large-scale railway network in the futrue.

Keywords

Journey Planning; Travel Cost; Digraph; Shortest Path Problem; A* Algorithm; Railway.

1. Introduction

As a key component of transport service information, the journey planning problem has recently gained interest for the large-scale high-speed railway network in China and the development of MaaS (Mobility as a Service) system. Nowadays, passengers are used to plan their railway trips with electronic timetable information systems, e.g., Chinese passenger railway service system 12306. Railway journey planning has to be investigated on the basis of the established timetable. One of the challenges in journey planning is to narrow the gap between practical relevance and academic solutions via theoretical understanding. The itineraries can be modeled as paths in a train-route graph and the optimal one can be found by searching for a shortest path. A fundamental algorithm understanding of the timetable information system, e.g., the representation of timetables for querying, is indispensable for the ideal journey planning towards both passengers and railway understakings [1].

Most of the existing literatures have focused on the itinerary planning for urban public transport or multimodal transportation networks as path-finding problem [2-3]. Meanwhile, the travel time on any link of the urban transit network or multimodal transport network is uncertain (stochastic and dynamic), and the distribution function of which generally depends on the departure time from the upstream node [4]. In most cases, the characteristic of the stochastic travel time is simplified by assuming that it is time dependent, i.e., depending on the departure time from the upstream node. There are two categories of Flex-Route Transit (FRT) problem in public tranport, i.e., the static FRT, and the dynamic FRT. For the static FRT, all customers are known in advance (before the beginning of the service). For the dynamic FRT, customers can request a ride after the shuttle started its service. The requests can be issued at any time, and they are dynamically affecting the shuttle schedule. Unlike the flex-route transit problem in urban transit [5], the backtracking is not allowed to be of use in the real-time operation environment of high-speed railway system.

In contrast to the headway-based model for the itinerary planning problem [2], the timetable-based journey planning is equivalent to schedule-based approach to a certain degree, which is suitable to employ the labeling algorithm [6], i.e., label correcting and label setting algorithms, for solving this kind of problems. The itinerary on the timetabled railway network is a category of scheduled transit. It is recognized that finding the optimal itinerary or journey plan for a traveler/passenger using public transport (including the railway transportation) is an algorithmic challenge [7], which is also a problem of routing passengers. For exact timetable queries, i.e., optimal itinerary query, [7] proposed the efficient algorithms for the earliest arrival question by using time-dependent networks as a model to seek optimal itineraries/journey plans for railway travelers.

The most simplified or fundamental basic version of journey planning problem can be taken as solving the timetable information issue, i.e., given an origin/departure station, a destination/arrival station, and a departure time, the task is to seek the connection/itinerary that arrives as desirable (e.g., early) as possible at the arrival station. There are two branches for solving the timetable information queries [8-9], i.e., the heuristic two-phase approaches, and the direct shortest-path approaches. For the heuristic two-phase approaches, heuristics are used to keep the search spaces small enough, and the first phase heuristically restricts the search space. However, sometimes it can't guarantee the resulting connections are optimal. The typical electronic timetable information systems introducing the heuristic two-phase approaches include HAFAS, EFA, TRAINS, and ARIADNE. One of the most prominent heuristics using the geometric embedding of the graph is the leverage of goal directed search in the artificial intelligence literature, i.e., A^*-algorithm, so as to speed up the computation procedure towards the optimal shortest path.

Traffic information systems are among the most prominent real-world applications of Dijkstra's algorithm for shortest paths. In practice, this problem is usually solved by heuristical variations of Dijkstra's algorithm, which do not guarantee optimality. Heuristic variations of Dijkstra's algorithm and their speed-up techniques are usually used to solve the traffic information systems for shortest path and process the numerous on-line queries for desired trip connections in practice of wide-area railroad networks [10]. In order to find all potentially attractive connections/alternatives for passengers in railway systems, i.e., meeting the query, [11] introduced the concept of relaxed Pareto dominance (Pareto version of Dijkstra's algorithm) with fare estimations for solving the algorithmically multi-criteria (e.g., travel time, ticket costs, and number of interchanges) shortest path problems, and implemented a prototypal information server PARETO on the basis of time-expanded model in a cooperation with Deutsche Bahn Systems.

With regards to the shortest-path problem, the point-to-point journey planning problem in the transportation network belongs to the class of constrained shortest path problem in most cases [12], which is regarded as the category of "hard" problem with non- polynomial computation time in nature. In order to handle the computationally intractable problem in realistic large-scale railway network, [13] proposed the Adaptive Large Neighborhood Search (ALNS) metaheuristic algorithm for the intertwined network design, line planning and capacity study in realistic size scenarios. By extending the time-dependent approach to support a vast number of on-line queries, [14] proposed models for three realistic versions of optimal itinerary planning problems, i.e., the Earliest Arrival problem with Non-Zero Transfer Time (by using the modified Dijkstra algorithm), Minimum Number of Transfer problem, and a combination of them.

For the direct shortest-path approaches, a query can be answered by applying some shortest path algorithm to a suitably defined graph, i.e., mapping a timetable information query to a single shortest-path problem in an appropriately constructed graph, and using the variants of Dijkstra's classical shortest-path algorithm with the speedup-technique called the multi-level graph approach. For modeling the timetable information in railway systems (and other public transportation systems) in weighted graphs, the main shortest-path approaches include the time-expanded approach (every event at a station, e.g., the train departure, is modeled as a node in the graph) and the time-dependent approach (the edge weights of the graph depend on time or are functions of time) [15]. When constructing the transformed graph models incorporating time, many details concerning feasible journeys/itineraries have to be considered, e.g., rules for train transfers. Meanwhile, criteria of optimality has to be clearly defined as edge weights, e.g., earliest arrival at the destination, and minimisation of train transfers, for achieving the Pareto-optimal shortest paths.

For the first time, [16] studied the single objective dynamic programming approaches for a time-dependent shortest-path problem, which was extended by [17] towards multiple criteria later. A detailed theoretical analysis of operation counts proposed by [7] proved that the time-expanded approach is less efficient than the time-dependent approach, i.e., the time-dependent approach would perform less CPU work. While on the other hand, it turns out that the time-expanded approach is more robust for modeling more complicated scenarios. Time-dependent shortest path algorithms [18] are required to handle the variability of travel times caused by traffic congestion/delay when producing optimal journeys involving both private vehicle legs and public transport on the public transport network. [18] compared various time-dependent shortest path algorithms, among of which the Dynamic A^* algorithm and Dynamic Adaptation of Dijkstra's algorithm are included.

Some of the existing literatures have focused on the earliest arrival time for the departure time with label correcting algorithms [19]. The quality of a passenger itinerary can be judged by more than one criterion, e.g., the number of transfers, the duration, the ticket/monetary cost, and even the convenience. Additionally, passengers may have various preferences with different individual weighting of the criterion towards a favorable trip planning. With regards to these, it would be desirable to present as many as required the journey planning options with Pareto-optima, i.e., the different itineraries that trade-off the travel objectives and constraints, e.g., minimize the duration with some constraints as the number of transfers and ticket cost. Two types of dynamic shortest path problems in the network can be distinguished [20], i.e., the fastest paths problem, and the minimum-cost paths problem. The cost of the former one depends on the travel time, while the cost of the latter one can be of general form. With regards to this, the optimization of the multicriteria journey planning problem can be ascribed to the latter one.

For the first time, [21] proposed the A^* algorithm, which is an algorithm with high efficiency to explore the one-to-one shortest path problem in both static and dynamic networks [22]. Later, the A^* algorithm was further extended and studied. [22] provided a brief introduction to the A^* algorithm for the readers from the transportation area. Up till now, A^* algorithm has been widely used in graph traversal and pathfinding [23], which can be adopted as the base search algorithm. For computing fastest paths in a class of dynamic networks, i.e., time-dependent shortest paths problem, [22] presented the adaptations of the A^* algorithm with high efficiency in the intelligent transportation systems, which examined the one-origin-node to one destination-node problem variant for a given departure time, regardless of one origin-node to all destination-nodes shortest path problem and the all-nodes to one-destination-node problem variant. A^* algorithm is a classic goal-directed shortest path search. The goal of A^* is to seek the optimal path from a starting point to the terminal goal within a graph, by considering a comprehensive combination of time and money savings perceived by the passengers. In order to support a distributed search in an integrated multimodal traffic information server system, [24] presented two extensions of A^* algorithm, i.e., the cooperative A^* algorithm, and the hierarchical A^* algorithm. And they suggested that a combination of hierarchical A^* and cooperative A^* algorithm could be more convenient for journeys spanning longer distances.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202406_10(6).0019

The main contributions of this study are summarized as follows.

(1) The personalized door-to-door travel costs for each O-D pair query on the basis of the detailed requirements analysis are formulated.

(2) The digraph of the train service network based on the timetable are descripted.

(3) Two-phase solutions for the journey planning problem in the railway network are proposed.

The remainder of the paper is organized as follows. Section 2 descripts the journey planning problem in more details, with regards to the personalized door-to-door travel cost for each O-D pair query on the basis of the detailed requirements analysis. Section 3 proposed the digraph representation of the train service network based on the railway timetable. The two-phase solutions to the railway journey planning are proposed in Section 4. Section 5 presents the discussions on this study. In the end, Section 6 concludes this research.

2. Problem Description

The itinerary planning is considered as the one-origin-node to one-destination-node journey planning problem in our study. If the intermediate node along the itinerary is not incorporated as a mandatory visit, a query [23] for a passenger-oriented complete trip chain in the timetable-scheduled railway system can be defined as a tuple $\{o(city,\ \mathrm{position}),\ \mathrm{d(city, position), dep(date, time), arr(date, time), seat}\}$ consisting of city and position of origin $o(city,\ \mathrm{position})$, city and position of destination $\mathrm{d(city, position)}$, date and time of departure $\mathrm{dep(date, time)}$, date and time of arrival $\mathrm{arr(date, time)}$, and class of seat. On the other hand, the specified intermediate node can be added to the query tuple. In this study, the default for the query is the first case.

The general purpose of the journey planning is to offer some trip plans and guide the passengers to travel from an origin to a destination in certain "optimal" way, e.g., the minimal travel time [10]. [25] investigated two single-criterion (i.e., the earliest arrival (EA), and the minimum number of transfers (MNT)) and a few bicriteria optimization problems with EA and the MNT as the two criteria. The traditional algorithms such as Dijkstra 's (the cornerstone for shortest path problems in graph with non-negative arc/edge weights) and A^* have laid the groundwork for journey planning under less complex railway network conditions. By introducing heuristics to guide the search towards the goal node more efficiently, A^* is regarded as an extension of the traditional Dijkstra's algorithm.

Mobility patterns depend on the passenger utility. Passengers always follow the best suitable path to reach their destination in terms of utility. The passenger-centric solution considering end-user experience and preferences, marking a significant advancement in modern transportation paradigm driven by the age background of rapid technological growth, e.g., big data analytics, AI (such as reinforcement learning algorithms, and even the Large Language Models), and IoT [26]. The door-to-door travel time has a strong influence on the mode selection[27], without the exception of the journey route choice. In pre-trip planning, both the access time and the egress time have to be taken into account, which relates to the choices of the departure station and arrival station in the transportation hub of the origin and destination.

The target for the personalized journey planning in the timetable-scheduled railway network is to minimize the travel cost under real-time train operation scenarios for each O-D pair query, and to obtain the fast query times with the associated planning algorithms. so as to seek the optimal trip plans for the passengers with regards to the choices of the travel route, train, and the seat. In comparison with the transit itinerary planning on the road [28], railway journey planning integrates the path calculation and schedules simultaneously under the timetabled network framework, given the origin, destination, and other circumstance specifications like access and egress calculations for the entire trip.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202406_10(6).0019

The personalized journey planning consists of three parts, i.e., the route-path choice (A route connection between each pair of station stops in the timetable can be determined as a route-path), the selection of the train (including the departure/arrival time), and the preferred seat selection. By considering the time-money-tradeoff and the real-time train operation scenarios, the personalized door-to-door travel cost for each O-D pair query on the basis of the detailed requirements analysis can be computed as formula (1).

$$ptc = \alpha_1 * tto * td + \alpha_2 * mc + \alpha_3 * tp * nt + \alpha_4 * \beta_1 * ddp + \alpha_5 * \beta_2 * tdp + \alpha_6 * \beta_3 * sdp \tag{1}$$

Time duration can be estimated as formula (2).

$$td = eat + rtt + eet + stt \tag{2}$$

Transfer time can be estimated as formula (3).

$$t_{ij}(l_u, l_v) = \begin{cases} t_{l_v}^{d}(j) - t_{l_u}^{a}(i) + ctt & \text{if } t_{l_v}^{d}(j) - t_{l_u}^{a}(i) \le T_0 \\ t_{l_v}^{d}(j) - t_{l_u}^{a}(i) & \text{if } t_{l_v}^{d}(j) - t_{l_u}^{a}(i) > T_0 \end{cases} \tag{3}$$

Monetary cost can be estimated as formula (4).

$$mc = am + sttc + em \tag{4}$$

Where:

ptc denotes the personalized travel cost.

tto denotes tradeoff between time and money, which depends on the value time of the passenger.

td denotes the time duration of the entire journey.

mc denotes the monetary cost of the entire journey, e.g, associated ticket price.

tp denotes the transfer penalty.

nt denotes the number of transfers of the entire journey.

ddp denotes the date deviation penalty for the journey, which depends on the gap between the passenger’s specified date and the recommended date of the train service.

tdp denotes the time deviation penalty for the journey, which depends on the gap between the passenger’s specified time and the recommended time of the train service.

sdp denotes the seat deviation penalty for the journey, which depends on the gap between the passenger’s specified seat and the recommended seat of the train service.

eat denotes the estimated access time for the journey.

rtt denotes the riding time on the train for the journey.

eet denotes the estimated egress time for the journey.

stt denotes the sum of transfer time for the entire journey.

am denotes access monetary cost, e.g., ticket price.

sttc denotes the sum of train ticket cost for the journey.

em denotes the egress monetary cost for the journey, e.g., ticket price.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202406_10(6).0019

$t^{a}_{l_u}(i)$ denotes the arrival time of train l_u at transfer station i.

$t^{d}_{l_v}(j)$ denotes the departure time of train l_v at transfer station j.

ctt denotes the cycle time of timetable. If the trains are running daily, $ctt = 1440$.

$t_{ij}(l_u, l_v)$ denotes the transfer time between train l_u at transfer station i and train l_v at station j.

T_0 denotes the a lower bound of time duration for a pair of two consecutive connections.

α_i denotes the associated weight coefficient, which is specified by the passengers, $i = 1,\ldots,6$.

β_i denotes the 0-1 binary parameter, if the associated deviation happens, $\beta_i = 1$, else $\beta_i = 0$, $i = 1,2,3$.

As we mainly investigated the multi-criteria optimization problem with the personalized travel cost in this study, given a dataset D composing of the candidate connections of trip plans, for a query q from passenger p, we would like to seek the K ($K \le 5$) most possible trip plan, formally defined as solving the K user-specific optimal path with the associated minimum travel cost. Two-phase solution approaches are proposed in this study. In the first phase, the candidate route connections of trip plan are generated to ensure the search efficiency. In the second phased, the purpose is to answer the query by applying certain shortest-path algorithm to a trip-plan dataset suitably constructed in the first phase. It has been recognized that no "best" shortest path algorithm exists for every situation of transportation problem [29] (e.g., the multicriteria shortest paths), especially, when the factor "time" is taken into account. In order to capture the peculiarities of the journey planning problems in railway system under various scenarios, the focus in the second phase of this study is to design and implement the ad hoc shortest path procedures [29] associated with the structure of the digraph, the problem size, order of node selection, and the cost measure leveraged for assessing the paths.

3. Digraph Representation of the Train Service Network based on the Timetable

From the perspective of theoretical computer science, the nature of journey planning lies in the data-structure question[30], whose objective is to construct a data-sturcture,e.g., the train-route graph, from the timetable, and then response the queries quickly, with two basic models (i.e., time-expanded models,and time-dependent models[31] and their variations and extensions [32]. Besides,more engineering aspects have been explored, tested and investigated [33]. At first, the train service network on the basis of the timetable is constructed to facilitate the candidate passenger trip-plan generation.

The information about the train route, the stops, and departure/arrival times can be obtained from the railway timetable. The train service network is the kind of constrained passenger-oriented network on the basis of the timetable, whose main objectives are the affordable train service and the transportation resource. The train service network based on the timetable can be represented as G=(V,A,W), and the symbol V represents the stop of the train (except the starting and terminal station, each stop node is split into two nodes: one arrival node and one departure node; and each arrival node can appear only as the source of an arc, while each departure node can appear only as the destination node of an arc,); $A \subseteq V \times V$, the symbol A represetns the boarding arc (the connection between the physical station and its associated departure node), the riding arc (the connection between two adjacent departure node and arrival node on the same train line in the running direction, which is the elementary connection), the alighting arc (the connection between the arrival node and its associated physical station), and the transfer arc (the connection between the arrival node and the departure node on the different train). The symbol represents the weight of the associated arc, involving the time, and ticket fare, etc. The train service netwrok G can be refered to as the train-route digraph, which has the characteristics of multiple nodes, multiple arcs, and complex network structure.

With the constructed travel service network graph based on timetable, we can enumerate the candidate trip plans from the access node associated with each start stop to the egress node associated with each end stop. Passengers can choose to travel between any departure station and arrival station if both are connected by a directed path or via a feasible scheduled transfer in between. The length of a trip path is defined as the sum of the arc weights in the path of the graph.

4. Two-phase Solution Approach

4.1 Phase I: Offline Generate Candidate Route Connections of Trip Plans based on Train Service Network

We remark that a train l can be characterized by a tuple as following: $l=(l_{numb},\{stop_k|k=1,2,\ldots,s\},\{sno_k,sn_k,city_k,sat_k,sdt_k,ds_k,cs_k|k=1,2,\ldots,s\},\{pri_h|h=1,2,3\})$

Where:

l_{numb} denotes the number of train l.

stop denotes the set of the train's stop stations.

s denotes the number of the stops.

sno_k denotes the number of the train in the running direction.

snk denotes the name of the train stop.

cityk denotes the city of the train stop.

satk denotes the arrival time of train at the stops, sat1=none.

sdtk denotes the departure time of train at the stops, sdtm=none.

dsk denotes the accumulating distance at station k on the running direction.

csk denotes the city code of station k.

prih denotes the basic fare rate associate with the class of the seat, h=1,2,3.

In most big cities, it has more than one passenger stations, i.e., there are passenger railway hubs with more than one stations within the city. According to the node match method based on the running direction, the pseudocode for pre-computing (preprocessing) the candidate trip-plan dataset in railway central passenger server is showed as follows.

Algorithm 1: pre-computing the candidate trip-plan dataset

set dirO the direction of the origin

set dirD the direction of the destination

Foreach $o \in V$ or $city(o)=city(V)$

Foreach $d \in V$ or $city(d)=city(V)$

Foreach pair(o, d), $o \in V$ or $city(o)=city(V)$, $d \in V$ or $city(d)=city(V)$

set LD the set of train number heading to dirD from the city of the origin

set LA the set of train number coming from dirO from the city of the destination.

If $LD \cap LA \neq \Phi$

set $L_{non-trans} = LD \cap LA$.

Set $Plans_{o \to d}^{non-trans} = \{l_{non-trans} \mid l_{non-trans} \in LD,\ l_{non-trans} \in LA,\ l_{non-trans} \in L_{non-trans}\}$

If $LD \cap LA = \Phi$, there is no direct trains connecting the od pair, calculate 1-transfer set as follows:

List stops of the train in the set LD $stop^{ld}$ (ld $\in$ LD)

List stops of the train in the set LA $stop^{la}$ (la $\in$ LA)

;;If $stop^{ld} \cap stop^{la} = \Phi$, there is no 1-transfer trip within the same station.

;;If $stop^{ld} \cap stop^{la} \neq \Phi$, there is 1-transfer trip within the same station.

If $city(stop^{ld}) \cap city(stop^{la}) \neq \Phi$;;there are 1-transfer trips, even the 1-transfer trip between different stations within the same city.

Set $transfer1 = stop^{ld} \cap stop^{la}$

Set $city(transfer1) = city(stop^{ld}) \cap city(stop^{la})$

$transfer1 = \{stop_i \mid stop_i \in stop^{ld}, stop_i \in stop^{la}, city(stop_i) = city(stop^{ld}) = city(stop^{la}), i = 1, 2, \cdots, |transfer1|\}$

;; rank the transfer stations in TRANSFER1 descendingly
set i=1
While i<=|transfer1|-1

If $rank(stop_i) < rank(stop_{i+1})$

Set $stp = stop_{i+1}$

Set $stop_{i+1} = stop_i$

Set $stop_i = stp$;;to set $rank(stop_i) \geq rank(stop_{i+1})$

Set i=i+1
End while
Set i=1
While i<=|transfer1|-1

Set $sdt^{l'}(stop_i)$ departure time of train l' at transfer station $stop_i$.

Set $sat^{l}(stop_i)$ arrival time of train l at transfer station $stop_i$.

If $sdt^{l'}(stop_i) - sat^{l}(stop_i) \geq T_0$, $l \in LD$, $l' \in LA$

Set $plan^{i}_{o \to d} = \left\{ <l,\ stop_i,\ l'>_j \mid j = 1, 2, \cdots \right\}$

Set i=i+1

Set $Plans^{transfer1}_{o \to d} = \left\{ plan^{i}_{o \to d} \mid i = 1, 2, \cdots, |transfer1| \right\}$;; the set of 1-transfer

End while
;;up till now, the set of non-transfer and 1-transfer have been built.
;;considering the scenario of 2-transfer

List stops of the train in the set LD $stop^{ld}$ (ld $\in$ LD)

ISSN: 2414-1895 DOI: 10.6919/ICJE.202406_10(6).0019

List stops of the train in the set LA $stop^{la}$ (la $\in$ LA)

Set i=1

While i<=| $stop^{ld}$ |-1

Foreach $stop_i \in stop^{ld}$

If $rank(stop_i) < rank(stop_{i+1})$

Set $stp = stop_{i+1}$

Set $stop_{i+1} = stop_i$

Set $stop_i = stp$;;to set $rank(stop_i) \geq rank(stop_{i+1})$

i=i+1

End for

End while

Set i=1

While i<=| $stop^{la}$ |-1

Foreach $stop_i \in stop^{la}$

If $rank(stop_i) < rank(stop_{i+1})$

Set $stp = stop_{i+1}$

Set $stop_{i+1} = stop_i$

Set $stop_i = stp$;;to set $rank(stop_i) \geq rank(stop_{i+1})$

i=i+1

End for

End while

Set i=1

While i<=| $stop^{ld}$ |-1

Foreach $stop_i \in stop^{ld}$

j=1

While j<=| $stop^{la}$ |-1

Foreach $stop_j \in stop^{la}$

if $LDI \cap LAJ \neq \Phi$

Set $Ltransfer2 = LDI \cap LAJ$;;a line connecting two transfer stations

Set $transfer2 = \left\{ < stop_i, stop_j > \right\}$;;two transfer station pair

Volume 10 Issue 6, 2024
DOI: 10.6919/ICJE.202406_10(6).0019

If $sdt^{l_2}(stop_i) - sat^{l_1}(stop_i) \geq T_0 \quad l_1 \in LD,\ l_2 \in Ltransfer2,\ stop_i \in LD$

And

$sdt^{l_3}(stop_j) - sat^{l_2}(stop_j) \geq T_0 \quad l_3 \in LA,\ l_2 \in Ltransfer2,\ stop_j \in LA$

Set $plan_{o \to d}^{i-j} = \left\{ <l_1,\ stop_i,\ l_2,\ stop_j,\ l_3>_m \mid m = 1, 2, \cdots \right\}$

i=i+1

j=j+1

set $Plans_{o \to d}^{transfer2} = \left\{ plan_{o \to d}^{i-j} \mid i = 1, 2, \cdots;\ \text{j=1,2,}\cdots \right\}$

End foreach

End foreach

End while

End while

Create dataset Buffer0 in central server

Create dataset Buffer1 in central server

Create dataset Buffer2 in central server

Save $\underset{\substack{o \in V \text{ or } city(o)=city(V) \\ d \in V \text{ or } city(d)=city(V)}}{Plans_{o \to d}^{non-trans}}$ in Buffer0

Save $\underset{\substack{o \in V \text{ or } city(o)=city(V) \\ d \in V \text{ or } city(d)=city(V)}}{Plans_{o \to d}^{transfer1}}$ in Buffer1

Save $\underset{\substack{o \in V \text{ or } city(o)=city(V) \\ d \in V \text{ or } city(d)=city(V)}}{Plans_{o \to d}^{transfer2}}$ in Buffer2

End foreach pair(o,d).

4.2 Phase II: Real-time Journey Planning under Practical Train Operation Scenarios

The task of this second phase is to solve the real-time transit itinerary planning problem for a given planning request on the basis of candidate trip plans generated in the first phase, by focusing on the efficiency of the search algorithm. The ultimate goal of this second phase is to provide a small set (a set with a number of K associated elements) of highly attractive and worthy connections as a reply to a passenger query. For the trip plan with transfers, it is a Cartesian product of two more connections in the train service network. The real-time itineray planning is a kind of scenario-based issue, including the scenario associated with the non-transfer trip plan to query the dataset Buffer0, the scenario associated with the 1-transfer trip plan to query the dataset Buffer1, and the scenario associated with the 2-transfer trip plan to query the dataset Buffer2. As for the associated query algorithm, the adaptive A^* metaheuristic [13], and the problem-specific A^* search methods (employ a problem-specific heuristic based on A^* search) can be referred to guide the search procedure heuristically to the dataset of candidate route connections of trip plans. And the personalized door-to-door travel cost for each O-D pair query (formula (1)) can be used as the evaluation function of the variants of the A^* algorithm, which can be estimated dynamically.

Providing the trip demand at its various stations within the city of the origin and destination, the corresponding passengers will prefer the journey route plan when the departure station and the arrival station are located in their vicinity. Correspondingly, the concept of vicinity depends on the

ISSN: 2414-1895 DOI: 10.6919/ICJE.202406_10(6).0019

estimation of the access time and egress time resulting from the feeder transport of railways, e.g., metro, bus, car, bicycle, or foot. Meanwhile, all passenger stations within the city of the origin and the destination should be identified, with regards to the optimization of the complete personalized railway journey planning. For the access and egress estimations, the information system of the urban transit, e.g., Mobility as a Service (MaaS) platform [34], is suggested to be incorporated into the railway journey planning system, or to integrate with the train timetable information system (e.g., 12306 passenger service system in China).

5. Discussion

There has been a number of published papers on timetable queries, but most of them considered only very simplified scenarios [9]. In this study, we have incorporated as many as possible scenarios, which have been embodied in the formula of the personalized travel cost estimation, i.e., formula (1). Though the side-constraints are not considered, the feasibility of the solution can be ensured because the candidate trip plans in the first phase are pre-computed based on the train service network and filtered out by the personalized travel cost. By the end of the year 2023, the national railway running mileage has achieved 159000 km in China, and the length of high-speed railways in service has reached 45000 km in China. This length of railway operating kilometrage can guarantee that most of the itineraries can be fulfilled within twice transfers in the railway system.

For our lack of real-world data, the analysis of our approach is limited to a theoretical method without computational performance. Besides the number of nodes, it is proved in [2] that the computational performance of the labelling algorithms they proposed is a function of the time window width, i.e., the width of the time window specified by the earliest departure time from the origin and the latest arrival time at the destination in the multimodal transportation networks. On a single processor with conservative estimate, the overall CPU time taken to complete all the computation in phase I (i.e., generating promising candidate route-paths) for a case of large-scale transit itinerary planning under uncertainty is about several months [3]. In phase I, the candidate route connections of trip plans can be offline generated by repeatedly solving the pre-compute candidate trip-plans and stored into a database. In nature, this is the kind of enumeration of paths, which may be more computationally burdensome with larger networks. For large-scale railway networks, phase I generally takes significant amount of CPU time, but it typically needs to be performed only once for the purpose of query in the phase II.

6. Conclusion

In this paper, we proposed two-phase solution approach for journey planning in the railway system, considering the operation scenarios of non-transfer itinerary, 1-tarnsfer itinerary, and 2-transfer itinerary. The first phase is a robust feasible path approach that is suitable for most of the cases. The second phase is a problem-specific heuristic search which concerns the efficiency of the real-time search algorithm. In the phase I, we offline generate the candidate trip plans and store them into the associated dataset, i.e., dataset Buffer0 for non-transfer, dataset Buffer1 for 1-transfer, and dataset Buffer2 for 2-transfer. Whenever there is no direct train between two stations, a scheduled transfer ensures that a train moving for the destination station departs from a transfer station shortly after a train originating from the origin station has arrived. On a long-term horizon, transfers can be avoided with the technique of virtual coupling, and even the dynamic coupling (trains that run in relative braking distance are allowed mechanical coupling and decoupling while driving). Future research should focus on the more flexible journey planning problem for non-transfer scenarios under the condition of virtual coupling system and dynamic coupling system. Meanwhile, the large language model (LLM) in the artificial intelligence field can be employed to solve the complex joureny planning problem in the large-scale railway network in the futrue.

Acknowledgments

No acknowledgement was addressed in this study.

References

[1] Gatto, M., Jacob, R., Peeters, L., Weber, B., & Widmayer, P. (2004). Theory on the tracks: A selection of railway optimization problems. Bulletin of EATCS, 84, 41-70.

[2] Zografos, K. G., & Androutsopoulos, K. N. (2008). Algorithms for itinerary planning in multimodal transportation networks. IEEE Transactions on Intelligent Transportation Systems, 9(1), 175-184.

[3] Li, J. Q., Kong, N., Hu, X., & Liu, L. (2015). Large-scale transit itinerary planning under uncertainty. Transportation Research Part C: Emerging Technologies, 60, 397-415.

[4] Pattanamekar, P., Park, D., Rilett, L. R., Lee, J., & Lee, C. (2003). Dynamic and stochastic shortest path in transportation networks with two components of travel time uncertainty. Transportation Research Part C: Emerging Technologies, 11(5), 331-354.

[5] Shahin, R., Hosteins, P., Pellegrini, P., Vandanjon, P. O., & Quadrifoglio, L. (2024). A survey of Flex-Route Transit problem and its link with Vehicle Routing Problem. Transportation Research Part C: Emerging Technologies, 158, 104437.

[6] Hamacher, H. W., Ruzika, S., & Tjandra, S. A. (2006). Algorithms for time-dependent bicriteria shortest path problems. Discrete optimization, 3(3), 238-254.

[7] Brodal, G. S., & Jacob, R. (2004). Time-dependent networks as models to achieve fast exact time-table queries. Electronic Notes in Theoretical Computer Science, 92, 3-15.

[8] Müller-Hannemann, M., Schulz, F., Wagner, D., & Zaroliagis, C. (2007). Timetable information: Models and algorithms. In Algorithmic Methods for Railway Optimization: International Dagstuhl Workshop, Dagstuhl Castle, Germany, June 20-25, 2004, 4th International Workshop, ATMOS 2004, Bergen, Norway, September 16-17, 2004, Revised Selected Papers (pp. 67-90). Springer Berlin Heidelberg.

[9] Schulz, F. (2005). Timetable information and shortest paths (Doctoral dissertation, Karlsruhe Institute of Technology.

[10] Empirical, A. (2003, May). Dijkstra's Algorithm On-Line. In Algorithm Engineering: 3rd International Workshop, WAE'99 London, UK, July 19-21, 1999 Proceedings (p. 110). Springer.

[11] Müller-Hannemann, M., & Schnee, M. (2007). Finding all attractive train connections by multi-criteria pareto search. In Algorithmic Methods for Railway Optimization: International Dagstuhl Workshop, Dagstuhl Castle, Germany, June 20-25, 2004, 4th International Workshop, ATMOS 2004, Bergen, Norway, September 16-17, 2004, Revised Selected Papers (pp. 246-263). Springer Berlin Heidelberg.

[12] Ziegelmann, M. (2001). Constrained shortest paths and related problems, Dissertation, Saarland University.

[13] Canca, D., De-Los-Santos, A., Laporte, G., & Mesa, J. A. (2018). The railway network design, line planning and capacity problem: An adaptive large neighborhood search metaheuristic. In Advanced Concepts, Methodologies and Technologies for Transportation and Logistics (pp. 198-219). Springer International Publishing.

[14] Pyrga, E., Schulz, F., Wagner, D., & Zaroliagis, C. (2004). Towards realistic modeling of time-table information through the time-dependent approach. Electronic Notes in Theoretical Computer Science, 92, 85-103.

[15] Pyrga, E., Schulz, F., Wagner, D., & Zaroliagis, C. (2008). Efficient models for timetable information in public transportation systems. Journal of Experimental Algorithmics (JEA), 12, 1-39.

[16] Cooke, K. L., & Halsey, E. (1966). The shortest route through a network with time-dependent internodal transit times. Journal of mathematical analysis and applications, 14(3), 493-498.

[17] Kostreva, M. M., & Wiecek, M. M. (1993). Time dependency in multiple objective dynamic programming. Journal of Mathematical Analysis and Applications, 173, 289-289.

[18] Casey, B., Bhaskar, A., Guo, H., & Chung, E. (2014). Critical review of time-dependent shortest path algorithms: A multimodal trip planner perspective. Transport Reviews, 34(4), 522-539.

[19] Nachtigall, K. (1995). Time depending shortest-path problems with applications to railway networks. European Journal of Operational Research, 83(1), 154-166.

[20] Chabini, I. (1997). A new algorithm for shortest paths in discrete dynamic networks. IFAC Proceedings Volumes, 30(8), 537-542.

[21] Hart, P. E., Nilsson, N. J., & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. IEEE transactions on Systems Science and Cybernetics, 4(2), 100-107.

[22] Chabini, I., & Lan, S. (2002). Adaptations of the A* algorithm for the computation of fastest paths in deterministic discrete-time dynamic networks. IEEE Transactions on intelligent transportation systems, 3(1), 60-74.

[23] Wang, J., Wu, N., Zhao, W. X., Peng, F., & Lin, X. (2019, July). Empowering A* search algorithms with neural networks for personalized route recommendation. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 539-547).

[24] Preuss, T., & Syrbe, J. H. (1997). An integrated traffic information system. In Proc. 6th Int. Conf. Appl. Computer Networking in Architecture, Construction, Design, Civil Eng., and Urban Planning (europIA'97).

[25] Pyrga, E., Schulz, F., Wagner, D., & Zaroliagis, C. (2008). Efficient models for timetable information in public transportation systems. Journal of Experimental Algorithmics (JEA), 12, 1-39.

[26] Pratap, S. (2023). Transportation Transformed: A Comprehensive Review of Dynamic Rerouting in Multimodal Networks. arXiv preprint arXiv:2312.14953.

[27] Nold, M., & Corman, F. (2023). How Will the Railway Look Like in 2050? A Survey of Experts on Technologies, Challenges and Opportunities for the Railway System. IEEE Open Journal of Intelligent Transportation Systems.

[28] Trépanier M., R. Chapleau, and B. Allard,2002. Transit itinerary calculation on the Web: Based on transit user information system. Journal of Public Transportation 5 (3): 3–3.

[29] Pallottino, S., & Scutella, M. G. (1998). Shortest path algorithms in transportation models: classical and innovative aspects. In Equilibrium and advanced transportation modelling (pp. 245-281). Boston, MA: Springer US.

[30] Gatto, M., Jacob, R., Peeters, L., Weber, B., & Widmayer, P. (2004). Theory on the tracks: A selection of railway optimization problems. Bulletin of EATCS, 84, 41-70.

[31] Mosquera, F., Smet, P., & Vanden Berghe, G. (2022). An algorithm for generating a diverse set of multi-modal journeys. Algorithms, 15(11), 416.

[32] Giannakopoulou, K., Paraskevopoulos, A., & Zaroliagis, C. (2019). Multimodal dynamic journey-planning. Algorithms, 12(10), 213.

[33] Wagner, D., Willhalm, T., & Zaroliagis, C. (2004). Dynamic shortest paths containers. Electronic Notes in Theoretical Computer Science, 92, 65-84.

[34] Zhang, J. (2023). MaaS-Motivated Synergetic Design of Intelligent Mobility Service Supply Chain Network towards Seamless Traveling and Integrated Multimodal Journey Planning, International Core Journal of Engineering, 10(3): 238-255.

[35] Zhang, J. (2023). Simulation-based schedule optimization for virtual coupling-enabled rail transit services with multiagent technique. Journal of Advanced Transportation, 2023.

[36] Zhang, J. (2024). Exploring the intra-platoon planning problem for heterogeneous train traffic under virtual coupling in the railway network via simulation. Journal of Rail Transport Planning & Management, 30, 100442.