Visualizing Mobility of Public Transportation System

Wei Zeng, Student Member, IEEE, Chi-Wing Fu, Member, IEEE, Stefan Müller Arisona, Alexander Erath, and Huamin Qu, Member, IEEE

Abstract—Public transportation systems (PTSs) play an important role in modern cities, providing shared/massive transportation services that are essential for the general public. However, due to their increasing complexity, designing effective methods to visualize and explore PTS is highly challenging. Most existing techniques employ network visualization methods and focus on showing the network topology across stops while ignoring various mobility-related factors such as riding time, transfer time, waiting time, and round-the-clock patterns. This work aims to visualize and explore passenger mobility in a PTS with a family of analytical tasks based on inputs from transportation researchers. After exploring different design alternatives, we come up with an integrated solution with three visualization modules: isochrone map view for geographical information, isotime flow map view for effective temporal information comparison and manipulation, and OD-pair journey view for detailed visual analysis of mobility factors along routes between specific origin-destination pairs. The isotime flow map linearizes a flow map into a parallel isoline representation, maximizing the visualization of mobility information along the horizontal time axis while presenting clear and smooth pathways from origin to destinations. Moreover, we devise several interactive visual query methods for users to easily explore the dynamics of PTS mobility over space and time. Lastly, we also construct a PTS mobility model from millions of real passenger trajectories, and evaluate our visualization techniques with assorted case studies with the transportation researchers.

Index Terms—Mobility, public transportation, visual analytics

1 INTRODUCTION

Public transportation system (PTS) is an important infrastructure in most modern cities. From the perspective of city management and urban planning, PTS is more than a service provider. It has several significant impacts to a city: economically, since PTS could reduce the overall transport cost of the city [20, 11]; socially, since PTS ensures all members of the city are able to travel [35]; and environmentally, since PTS generally saves more energy than private transport [9].

From the perspective of passengers, public transportation systems provide not only shared passenger transport services available to everyone in the general public, but also rapid transit services via trains and subways, thus capable of moving large volume of people efficiently across a city. This is particularly important for big cities in Asia, where private cars and taxis are not the major modes of transport, and most people rely on PTS to travel.

Hence, studying the efficiency of a PTS is highly beneficial to both individuals as well as to the entire city as a whole. Thanks to recent availability of various forms of public transportation data, including the passenger journey data collected via RFID cards, transit schedule data, and transportation network data, we now can study and explore the efficiency of a PTS by modeling and integrating these real-world data rather than relying on simulations. By then, we can further design and develop visual analytics methods to explore these data and serve the transportation researchers and urban planners. In particular, this work focuses on exploring and visualizing passenger mobility in a PTS, e.g., how fast passengers can travel by PTS, which is a highly crucial factor that impacts the overall PTS efficiency.

However, developing visual analytics methods to meet this goal is a highly challenging task due to the following issues:

- First, public transportation systems are increasingly complex to meet the population growth, e.g., metropolises like London and New York have 270 underground stations and 460 subway stations, respectively, offering more than a billion passenger trips per year. If we also consider buses and other transport modes, the PTS network would be overly complex for exploration and analysis. This motivates us to study PTS mobility models [29, 19] from the transportation research community to analyze routes started from a common origin in a complex network.

- Second, existing works in visualizing public transportation networks mostly employ network visualization methods and focus on presenting the network topology across stops. They ignore
various mobility-related factors, e.g., riding time, transfer time, waiting time, etc., that affect the PTS efficiency. Hence, novel methods have to be developed to meet the needs of exploring and analyzing these factors.

- Lastly, the mobility-related factors to be explored are not static. They vary dynamically with both time and space, and could also exhibit round-the-clock patterns. Hence, spatio-temporal visualization strategies have to be considered to maximize the visual analytics capability of a method.

To address the above issues, we present in this work a visual analytics framework to visualize and explore mobility-related factors in a public transportation system with three visualization modules:

- isochrone map view, which presents geographical regions accessible from a given starting location within certain duration;
- isotime flow map view, a novel strategy that linearizes a flow map in a parallel isotime fashion, enabling visualization and exploration of various mobility-related factors; and
- OD-pair journey view, which enables interactive exploration of various mobility-related factors, and their temporal variations, along the origin-destination journeys.

To develop the above visualization modules, we first analyzed the problem with the help of two transportation researchers, and identified the related analytical tasks (Section 3). Then, we constructed a PTS mobility model from different pieces of real data including transport network data and passenger RFID card data with several million trips (Section 4). After that, we developed the three visualization modules mentioned above, and refined their visual designs with the transportation researchers (Sections 5.1 to 5.3). During the design phase, we also implemented and explored different design alternatives for presenting mobility-related factors (Section 5.4). Lastly, to evaluate our visual analytics framework, we explored two case studies with transportation researchers who are currently working on public transportation planning and management, and also presented expert reviews received from two transportation researchers (Section 6).

2 RELATED WORK

This section discusses related works in the following five categories:

2.1 Isochrone Maps

Isochrone maps [39], or isochrone diagrams, are traditional visual representations used in transportation and urban planning for showing areas of equal travel time from a given starting location. Isochrone maps usually employ contour lines/colors in its representation, and can be easily overlaid on geographical maps for depicting information such as accessibility [26, 12]. However, isochrone maps alone are insufficient for revealing the exact travel time and routes from the starting location, as well as other mobility-related factors.

2.2 Schematic Maps

Public transportation networks are increasingly complex in cities. They can comprise hundreds of subway stations and thousands of bus stops. Exploring such a complex structure is a challenging task. Hence, schematic maps [23] were proposed to overview the network’s essential topological relations across stops, aiming to help users navigate the network and identify paths to bring them from their origins to destinations. One early work is the London underground map designed by Harry Beck in 1933 [14]. This pioneering idea has further inspired the design of many metro maps around the world [27], as well as many automated methods [24, 36, 40].

Through a user study, Bartram [10] confirmed that users can navigate the transport network efficiently and accurately with the help of schematic maps. Meilinger et al. [22] further explored and confirmed the value of schematic maps for better wayfinding and self localization. In addition, these studies also show that users can still acquire valuable information about the PTS even though exact geographical information is not available in the visual presentation.

2.3 Visual Analytics of Movement


Recently, some works focused on visual analytics of movement data related to urban traffic. Archambault et al. [7] explored local-connection patterns between airports in a worldwide flight dataset by browsing the proximity and paths around particular airports. Wang et al. [37] designed a multiple-view visual analysis system to study traffic congestion in taxi data. Ferreira et al. [13] proposed a novel visual query model, allowing users to interactively explore and compare results obtained from millions of taxi trips. Zeng et al. [41] proposed a novel visualization technique called the interchange circos diagram to reveal interchange patterns in massive public transport trips.

2.4 Time-Oriented Data Visualization

Time-oriented data, or data with a temporal dimension, is ubiquitous in many fields. Extensive amount of effort has been devoted to develop visualization methods to meet the needs of analyzing and understanding such kind of data, see [2, 3] for comprehensive surveys. In particular, Aigner et al. [2] highlighted that the time dimension could be linear or cyclical, so different visualization strategies could be adopted for different situations accordingly. For example, to show sunshine intensity variation over days, which is a kind of cyclical time data, we can employ spiral graphs [38] to reveal the cyclical patterns.

The data set we used contains both linear and cyclical temporal patterns, e.g., travel time from a point to all locations on a map is linear, whereas travel time variation between an origin-destination pair over a day is cyclical. To account for different types of time attributes, this work adopts both lineal line and circular graph visualization strategies.

2.5 Flow Map Visualization

Flow map [17] is an effective visualization technique, commonly employed to show the movement of objects from one origin location to many destinations, e.g., people migration across a country [28, 16]. A pioneering work was done by Tobler [32], who joined origins and destinations with lines and arrows, and revealed flow volumes with line width or color. However, it has certain issues: first, visual clutter would easily occur when applied to large data sets; second, longer flow lines can easily overlap and occlude shorter flow lines.

To address these issues, spatial generalization and aggregation methods [16, 5] have been proposed to reduce the number of locations pairs before presenting the flow map. Another effective method to resolve these issues is edge bundling. Phan et al. [28] proposed the flow map layout method to automatically cluster nodes into a tree-like hierarchical structure, and then bundle neighboring flow lines to present the general flow trend, while Verbeek et al. [34] later proposed a spiral-tree-based method to compute crossing-free flows.

In this work, our goal is to develop a visual analytics interface to present and explore passenger mobility in a public transportation system. In particular, we design the isotime flow map that linearizes a flow map into a parallel isoline representation, maximizing the visualization of mobility information while presenting clear and smooth pathways from origin to destinations. We further develop the OD trip journey view overlaid on this parallel isoline representation, enabling users to visualize and analyze mobility information, as well as their temporal variation, along specific routes in the flow map.

3 OVERVIEW

In this section, we first introduce relevant terminologies from the transportation research, and present the problem definition. After that, we describe the related analytical tasks, the mobility-related factors, and the input data set, and then give an overview of the system workflow.
3.1 Terminologies and Problem Definition

In the following, we list down common terminologies about public transportation system (PTS) [25] to facilitate the discussion:

- A transportation network consists of roads and subways, and is usually modeled as a directed graph data structure, where nodes are stops (metro station platforms and bus stops) with geographical locations, and directed edges connect neighboring nodes;
- A transit route is a sequence of nodes and edges, starting and ending at bus/subway terminals;
- A transit line is a public transportation service offered by a certain transport mode, e.g., a bus line and a subway line. There are two kinds of transit lines: bidirectional with two transit routes between two distinct terminals, and cyclical with one single transit route starting and ending at the same terminal;
- A trip refers to an individual transit route service taken between two stops/terminals;
- A transfer refers to a change of transit route services; it could happen at the same location (e.g., a bus stop), or between two different but neighboring locations (e.g., between different subway platforms or from subway to bus); and
- A journey is a passenger travel from an origin to a destination in the PTS; it could comprise multiple trips and transfers.

In land-use and public transport planning, transportation researchers would like to explore the level of connections, or the travel efficiency, from a particular location to other parts of the city, given the existing land use and transport network. By this, they can quickly identify which part(s) of the city is/are less developed, find out what facilities are lacking, and explore inefficient usage of public transportation resources. This problem is also related to optimal routes algorithms [29, 19] in transportation research, where transportation researchers study routes starting from a common origin to different points of interests in the city.

This is a collaborative work with two transportation researchers specialized in public transport systems. Based on their inputs, the following visualization problem is defined:

- Input: an origin A in a given public transportation network, time $t_0$, and a certain duration $T$;
- Output: a set of destinations B (and related routes) that are reachable from A at $t_0$ within $T$; and
- Goal: we aim to present and explore mobility-related factors (see Section 3.2) associated with the routes from A to B.

3.2 Analytical Tasks and Mobility-Related Factors

To address the problem, here are the basic analytical tasks that our visual analytics interface should support:

- T1: Given the input information $A$, $t_0$ and $T$, extract and present reachable destinations B;
- T2: Present clear pathways/routes from A to B; and
- T3: Examine and compare the travel time and travel efficiency of the routes from A to B.

The above basic tasks focus on presenting and exploring routes starting from A at a given time $t_0$. Additionally, we would need to allow the users to select specific destination nodes, say $B_i \in B$, and then:

- T4: Present detailed path information from A to $B_i$, i.e., various mobility-related factors, see below for details; and
- T5: Examine the mobility-related factors and their round-the-clock pattern, i.e., their temporal variations over a day.

In this work, the following mobility-related factors are considered:

- Waiting time at a bus stop or subway platform for a route service;
- Riding time on a vehicle for traveling between two stops;
- Transfer time for walking between neighboring stops; and
- Travel efficiency measures the efficiency of traveling between a specific pair of origin and destination relative to the average efficiency (speed) of travel in the PTS.

The above analytical tasks and mobility-related factors are the baseline requirements for our visual design to be presented in Section 5.

3.3 Input Data Sets

Passenger RFID card data records passenger journeys in the Singapore public transportation system over a typical working day. This includes the metro system called the mass rapid transit (MRT) and the public bus network, where passengers use their personalized RFID cards to tap on card readers on buses and MRT station entries to go in/out of the PTS. The card reader system records every tap in/out action, and also considers transfers between bus and MRT service: if the transfer time is $\leq 30$ minutes, the two trips will be sequenced together. Each trip record includes an anonymous card ID, tap-in/out stops, tap-in time, riding time, transportation type (bus or MRT), and transfer number. The tap-in time works slightly different for buses and MRT: for buses, passengers tap-in when they hop on buses, while for MRT, they tap-in when entering the station gates; after that, they still need to proceed to the platform and wait on the platform before getting into a train. We account for this difference and devise different strategies in estimating mobility-related factors for bus and MRT, see Section 4. The whole data set has 5.31M trip records in total.

Transportation network data includes the Singapore public road and the MRT network, which are modeled as a directed graph.

Transit line schedule data includes transit routes, stop facilities, and schedule information. For stop facilities, there are 4.8k bus and MRT stops, each with geographical position, name, reference ID, and related edge connection in the PTS network. For schedule information, it is a timetable showing when each bus/train leaves its starting terminal, and reaches each stop along its transit route.

3.4 System Workflow

The workflow of our visual analytics framework is illustrated in Fig. 2. It has two major phases: The data preprocessing phase loads and integrates various input data, estimates mobility-related factors over space and time, and constructs the PTS mobility model, see Section 4. To support near real-time determination of routes from a given origin A to destinations B, we index the mobility-related factors on stops and
transit routes. Note that this preprocessing phase is a one-time offline process, after which we store the precomputed information on hard disk, and load them in the next phase.

The visual exploration and analysis phase starts with our main interface with three modules: isochrone map view (Section 5.1), iso-time flow map view (Section 5.2), and OD-pair journey view (Section 5.3), which complement one another and work together to present the mobility-related factors and support the various analytical tasks.

4 MODELING PTS MOBILITY

4.1 PTS Mobility Model

In reality, public transport stops are usually not the origin or destination of a passenger journey; one often needs an initial walk, say from home/office/shop to a public transport stop, before the PTS trips and transfers, as well as a final walk to reach the destination. Since we have no data about the initial and final walks, we consider passenger journeys to start and end at stops in the public transport network.

One key factor that affects how passengers plan their journeys is travel time, which is also a crucial factor that affects the overall efficiency of a public transportation system. Hence, we choose to construct a PTS model [29, 19] that focuses on time-efficient journeys. In detail, we model a passenger journey with $n$ trips and $n-1$ transfers, so the overall travel time of the journey is modeled as

$$T_{\text{journey}} = \sum_{i=1}^{n} T_{\text{trip}}^i + \sum_{i=1}^{n-1} T_{\text{trans}}^i, \quad n \geq 1,$$

where $T_{\text{trip}}^i$ is the travel time for the $i$th trip, and $T_{\text{trans}}^i$ is the transfer time between the $i$th and $(i+1)$th trips. Since waiting time is often needed before boarding a vehicle, e.g., train and bus, we further divide $T_{\text{trip}}^i$ into waiting time $T_{\text{wait}}^i$ and riding time $T_{\text{ride}}^i$:

$$T_{\text{journey}} = \sum_{i=1}^{n} [T_{\text{wait}}^i + T_{\text{ride}}^i] + \sum_{i=1}^{n-1} T_{\text{trans}}, \quad n \geq 1.$$  

4.2 Estimating Mobility-Related Factors

In the data preprocessing stage, we first clean the raw passenger trip data by removing incomplete and erroneous data records, e.g., some passengers went out of buses without tap-out, some passengers stayed exceptional long inside the metro system compared to normal travel time needed to go between their tap-in and tap-out stations, etc.

Since mobility-related factors are time-varying, it is not feasible to estimate their continuous changes over time even with millions of passenger trip records. Hence, we divide the temporal dimension into seventy-two 15-minute time bins from 6am to midnight, which is the normal PTS operating period of a day. Then, we integrate various input data, and estimate the average value of each mobility-related factor (per stop or stop connection) per time bin. Note also that since passenger tap-in and tap-out mechanisms are slightly different for buses and metro services, we may need to consider bus and metro independently when estimating the mobility-related factors. Moreover, we assume that metro services always follow the transit line schedule while bus services may not (due to road sharing and local traffic).

Waiting Time is a per transit route, per stop and per time bin quantity. To estimate it for bus services, we extract all bus trips from the RFID card data. For each stop of each bus transit route, we first compute the average time over all tap-in tap-out times at the same stop of the same bus to estimate when the bus stays at each stop, say $t_{\text{bus}}^{\text{stop}}$. Hence, we can obtain all $t_{\text{bus}}^{\text{stop}}$ for all bus services (same transit route) at a given stop over the day, and then compute the time interval between successive $t_{\text{bus}}$ to estimate the bus frequency (interval) at each stop per bus line per time bin; half of such a value is the expected waiting time.

For MRT services, though transit line schedule data is accurate, actual waiting time may sometimes be longer than the time interval between successive trains since during the peak hours, passengers may not be able to board a train immediately after reaching the MRT platform. Hence, we estimate MRT waiting time as follows. First, we extract all MRT trips without MRT-to-MRT transfer since having a transfer could bias the computation below. Then, for each trip, we extract the tap-in time, and search for the next train that the passenger should have boarded at the tap-in station. By this, we can look up the transit line schedule to obtain the ride time required for him/her to reach the destination station, and estimate the related waiting time as: (tap-out time - tap-in time) - scheduled riding time. Since we can obtain multiple waiting time from different passenger records, we further compute their average as the expected waiting time.

Riding Time is a per successive stops (along the service route) and per time bin quantity. For bus services, after we estimate $t_{\text{bus}}$, at the stops of each bus line (see above), we can estimate the riding time of each bus between successive stops of the same bus. Again, we average multiple instances of such a value to obtain the expected riding time per successive stops and per time bin. For MRT services, we obtain riding time simply by looking up the transit line schedule data.

Transfer Time. There are three cases of transfer: First, it is from MRT to MRT. If the transfer happens between nearby platforms, we assume zero transfer time. However, in some cases, one may have to walk a fairly long distance from one platform to another. Since there are no card tapping activities during the transfer, we estimate transfer time by taking advantage of the data massive-ness: 1) extract all MRT journeys with only one transfer; 2) estimate the transfer time of the journey as: (tap-out time - tap-in time) - $T_{\text{trans}}^i$; and 3) again, average the results from different journeys per time bin.

The second case is from bus/MRT to bus. If the two bus stops are the same (same reference ID), we assume zero transfer time. Otherwise, we need a walk to the next bus stop, so we estimate the transfer time as tap-in time (next bus) - tap-out time (prev. bus) - wait time (next bus). Note that MRT to bus is slightly different from bus to bus since it requires a walk from MRT platform to the tap-out gate. However, since we have no information about such a walk, we ignore it and estimate transfer time in the same manner as bus to bus.

The last case is from bus to MRT, where we estimate the transfer time as tap-in time (MRT) - tap-out time (bus).

Travel Efficiency differs from the general concept of speed since it considers also waiting and transfer time in addition to riding time. Moreover, it describes the relative efficiency of travel along a specific route as compared to the mean mobility of the entire PTS network.

Before computing the travel efficiency of a specific route, we first determine the mean mobility of the entire PTS by: 1) compute the mobility of each passenger journey in the RFID data as: total journey distance divided by $T_{\text{journey}}$; and 2) compute the mean mobility $\mu$ and also its standard deviation $\sigma$ over all the journeys. By this, the travel efficiency of a given route (started at a given time) is obtained by normalizing its mobility value against $\mu$ and $\sigma$. Fig. 3 plots the travel efficiency of all bus services in the 8:00-8:15 time bin. Here we quan-
In reality, public transport stops are usually not the origin or destination of a passenger journey; one often needs an initial walk, say from home to a bus stop. Hence, we choose to consider passenger travel time, which is also a crucial factor that affects the overall efficiency of a public transportation system. Therefore, we estimate MRT waiting time as follows. First, we assume zero transfer time. However, in some cases, one may have to wait for the next bus after reaching the MRT platform. We can estimate transfer time in the same manner as bus to bus.

The visual exploration and analysis phase starts with our main interface with three modules: isochrone map view (Section 5.1), OD-pair journey view (Section 5.2), and temporal path comparison (Section 5.3). Note that MRT to bus is slightly different from bus to bus since it differs from the general concept of reachability. Hence, every stop will be surrounded by a circular region; we further union all these regions to determine time-reachable regions on the map (note: such union is done by rendering without tedious geometric computation). After that, we can plot the related contour lines and areas, see Fig. 5 for an example with a city center location as A. Here we use a red icon on the map to show the location of A, and highlight the contour regions in blue: dark blue for [0, 30) min., light blue for (30, 60] min., and white for >60 min. Moreover, we present in gray the set of all reachable edges, which is a subset of the entire PTS network, and adjust their line width to reveal the amount of time-efficient journeys that pass through each edge. By this, main branches can be emphasized.

5 Viszualization Design

In this section, we describe how we support the analytical tasks defined in Section 3.2 through the following three visualization modules:

5.1 Isochrone Map View

For Task 1, our goal is to extract and present all reachable locations B given A, t₀, and T. To handle it, we first compute time-efficient journeys from A to every single stop in the PTS using the estimated mobility-related factors. This is done by a real-time breadth-first-like mechanism (single-source shortest time-efficient paths) that iteratively identifies and expands time-reachable stops (nodes in the network graph) over the geographical map before T is reached.

In addition, at every reachable stop (including A), we consider “passenger walk” from the stop by using the remaining journey time at the stop within T, and assume a constant walking speed of 5km/h without encountering obstacles like buildings and roads. Hence, every stop will be surrounded by a circular region; we further union all these regions to determine time-reachable regions on the map (note: such union is done by rendering without tedious geometric computation). After that, we can plot the related contour lines and areas, see Fig. 5 for an example with a city center location as A. Here we use a red icon on the map to show the location of A, and highlight the contour regions in blue: dark blue for [0, 30) min., light blue for (30, 60] min., and white for >60 min. Moreover, we present in gray the set of all reachable edges, which is a subset of the entire PTS network, and adjust their line width to reveal the amount of time-efficient journeys that pass through each edge. By this, main branches can be emphasized.

5.2 Isochrone Flow Map View

To handle Tasks 2 & 3, i.e., to present clear routes from A to B and to examine and compare their travel efficiencies, the isochrone map view alone is insufficient. If we apply colors to B in this view to show the travel efficiency, the colors we employed would easily mess up with the isochrone colors. Moreover, it is hard to present clear pathways for examining and comparing time-efficient journeys in the isochrone map view, particularly with numerous pathways from A to B.

Hence, handling Tasks 2 & 3 is nontrivial, so we first explore different design alternatives in a pilot study, see Section 5.4 for detail. After comparing these alternatives, we propose the isochrone flow map view, which is a novel visualization strategy, that presents a flow map visualization in a parallel isotime fashion, see Fig. 4 for an example. The followings detail its construction procedure:

1. Parallel isotime model. First, we arrange a horizontal timeline on the bottom of the view to show the journey time from t₀ and t₀ + T (left to right). In this view, A is the red dot on the left while all destinations and nodes to B (in fact, all visual elements) are tagged with time. Thus, we can quickly look at the horizontal coordinate of any visual element w.r.t. the time axis to find out the related journey time from A. To further facilitate such visual examination and comparison, we draw an array of vertical gray lines in the background of the view to show different amount of time intervals.

2. Tree structure. To present time-efficient journeys from A to B,
Fig. 6. The OD-pair journey view focuses on Tasks 4 & 5, presenting detailed mobility-related information along routes from \( A \) to user-specified nodes (left), as well as their round-the-clock variations using our proposed visual representation: mobility wheel (right). Note also that we employ the standard colors of Singapore MRT lines to show the riding time on MRT, and encode bus lines by yellow.

Fig. 7. Left: A tree structure that includes all time-efficient journeys from starting node \( A \). Right: the result of applying our spatial layout algorithm to arrange the nodes in the visualization view.

rather than showing them one by one, we present them as a tree structure, which is a subgraph of the entire PTS rooted at \( A \). Such tree is constructed by examining the journeys and identifying branch nodes (transfer stops) and leaf nodes among the journeys. Here we also count the number of time-efficient journeys (as a weight factor) that go through each branch node.

3. Spatial layout. To present the time-efficiency journeys as a tree structure, we take the flow map visualization approach and layout the tree according to the parallel isotime model. Hence, the horizontal coordinates of all the nodes in the tree are fixed according to the related journey time from \( A \), see Fig. 4, and so, our main task in this step is to determine the vertical coordinates of all the nodes in the tree. As for this, we devise the following recursive method, which helps to avoid visual clutter and promote tree balancing:

- Given \( A \), we first extract all child branch nodes of \( A \).
- To improve the tree balancing, we sort these branch nodes as follows: Given \( k \) nodes, we first find out the node with the highest weight factor and assign it as \( n_1 \); the node with the 2nd highest weight as \( n_2 \); the node with the 3rd highest weight as \( n_3 \); then as \( n_{k-1}, \) etc. Let \( w_j \) be the corresponding weight of node \( n_j \).
- Then, we divide the vertical range from \( A \) into sub-ranges according to \( w_j \). See Fig. 7(right): node \( B_3 \) with the highest weight on the top, \( B_2 \) with the 2nd highest weight on the bottom, etc.
- Lastly, we repeat the above procedure for each child branch nodes of \( A \) till reaching the leaf nodes.

Note also that to reduce up-and-down wobbling along consecutive branches, see Fig. 8, we use simple node-based moving-window averaging to slightly shift the vertical position of the nodes.

4. Branch routing. After positioning all the nodes in the view, we next construct a Bézier curve to connect the nodes to form clear and smooth pathways. However, neighboring branches may overlap, see Fig. 9(a). To resolve this issue, we examine Bézier curves among sibling branches; if an overlap occurs, we horizontally shift the related Bézier control points, see Fig. 9(b).

5. Node color and label. Lastly, we color code each node according to its travel efficiency (see Section 4.2), and put text labels (name or stop reference ID) at nodes where on-screen radius is larger than 5 pixels in a up-down-up-down-etc. along consecutive branches.

5.1 OD-pair Journey View

Tasks 4 & 5 focus on supporting visual analysis of mobility-related factors along specific routes from \( A \). Clearly, if we color-code all flow lines in the isotime flow map view and present also their temporal variation, visual clutter would likely happen. Hence, we allow the users to click-and-select destination node(s) in the isotime flow map view, and then perform Tasks 4 & 5 through the OD-pair journey view, which is an overlay on the isotime flow map view, see Fig. 6:

- To support Task 4, we need to present detailed mobility-related information in the parallel isotime flow map: 1) we widen to highlight the branches along the user-selected route(s); 2) we color-code different portions of the flow line(s) to show the related mobility-related conditions: light blue for waiting, gray for transfer, standard colors of Singapore MRT lines for MRT riding (e.g., green for WW line), and yellow for bus riding; and 3) we highlight the starting, transfer, and ending nodes in red, gray, and dark blue, respectively, and label them with corresponding reference IDs/names.

Fig. 6(a) shows two user-selected routes: both routes have similar initial waiting time in WW16/YY3 station, but have different waiting times at their transfer nodes. For the journey to XX7 station on top, though it has no transfer time (nearby MRT platforms), it has much longer waiting time than the other route.
• To support Task 5, we design a visual cue called the mobility wheel, which is inspired by [21, 33], to show round-the-clock temporal variation of mobility-related factors, see Fig. 6(b). Our key idea here is to stack the mobility-related factors as small vertical bars, and then pack them in a time bin by time bin fashion around the mobility wheel. By this, we can visualize round-the-clock variation of all contributing mobility-related factors altogether.

In detail, we put a mobility wheel at each user-selected destination when the OD-pair journey view is brought up, see again Fig. 6(b). In addition, we use the same color coding scheme for showing the mobility-related factors as in Task 4, and highlight the current time bin (according to the main visualization) in the wheel by a thin red rectangle. The radius of the mobility wheels remain 200×200 on screen, so the user can zoom in and separate out overlapping wheels if wheel overlap occurs; moreover, the user may also click on a wheel to bring it to the top layer.

5.4 Design Alternatives
As discussed earlier in Section 5.2, when we design the isotime flow map view to handle Tasks 2 & 3, we explore different design alternatives to present the flow map and time-efficiency journeys from A to B. This section presents how we devise and implement these alternative designs, and compare them with the parallel isotime layout we have chosen. Below are the two alternative layouts we explored:

1. Time-scaled network distortion deforms the 2D map, so that distances between points on the deformed map relate to travel time [1], see Fig. 10(a). This method was popularly used in transportation to show travel time between locations, and some methods [30, 8, 31] have been proposed to perform the deformation.

In this work, we develop the time-scaled network distortion by a breadth-first visit from A; immediate child nodes of A are shifted to reflect their travel time from A, and we recursively repeat this for the branch nodes until the leaf nodes in the tree structure.

The first problem of this approach concerns with the complexity of the PTS network. When we consider many time-efficient journeys from A in the transport network, severe visual clutter would easily occur. Hence, this approach cannot present clear pathways from A to the reachable stops B (against Task 2). Second, it is difficult to accurately compare the travel efficiency of different routes in the visualization (against Task 3) even though we know that the total length of the (zigzag) routes relate to travel efficiency, see again Fig. 10(a). To the best of our knowledge, none of the existing time-scaled methods can handle these two issues.

2. Radial isotime layout. Besides the parallel isotime layout, we implemented and explored another layout alternative: a radial layout, see Fig. 10(b). It can be constructed in a way similar to the parallel isotime layout, but it positions A at the center, and B on concentric circles with increasing travel time along radial direction from A. Comparing with the parallel isotime layout, radial layout can still present clear pathways from A to the reachable stops, but after we show this early design to the transportation researchers, several negative feedbacks were received from them: First, concerning Task 3, it is not intuitive for them to examine and compare the travel time as routes and mobility-related information are arranged in a radial fashion. According to Heer et al. [18], encoding time progress from left to right along the horizontal axis can aid comparison of time-series events and their trends. Second, the radial layout cannot make full use of the screen space as the aspect ratio of common displays, e.g., 16:9. Lastly, in the parallel isotime layout, we can drag the horizontal time axis to let/right to intuitively modify the current time of the visualization, i.e., , but for radial isotime layout, such an operation is not intuitive.

6 Case Studies
This section presents 1) two case studies on exploring PTS mobility with our interface and 2) feedback from transportation researchers.

6.1 Case Study 1: Spatial Variation of PTS Mobility
In land-use and transportation planning, researchers and urban planners are interested in exploring the travel efficiency from a selected location to other parts of the city, as in Tasks 1 & 3. By the visual analysis, they can know what is lacking and also how to improve.

In this case study, we demonstrate how our system facilitates the exploration of PTS mobility over different locations in a city. Fig. 11 presents the isochrone map views and isotime flow map views related to two different locations on the map: an MRT station (left) and a rural-area bus stop (right). Though the starting time for both visualizations is 8am, the isochrone map views reveal very different sizes of reachable dominions from the two locations. In fact, the starting location for Fig. 11(a) is an interchange MRT station, which is the (WW24/X11) station in Singapore, with two MRT service lines and a bus terminal nearby. With rapid MRT services, passengers can reach very long distances from this location within a short period of time. In contrast, the starting location for Fig. 11(b) is a rural-area bus stop with only two bus lines available. From the isochrone map view on the right, we can clearly see the reachable regions along the two major directions from the bus stop corresponding to the two available buses. Moreover, the bus service line towards the south-east direction relates to more time-efficient journeys since its line width is wider than the other direction from that bus stop. Besides, the node colors in the isotime flow map view also confirm that the local travel efficiency here is very low, as compared to that in Fig. 11(a).

6.2 Case Study 2: Analyzing Mobility-related Factors
Our interface can also allow users to analyze and compare various mobility-related factors that affect the PTS efficiency, as in Tasks 4 & 5. Fig. 12(a) presents two user-selected routes from the source location MRT station WW24/X1 to destinations Bus stop 67852 and MRT station ZZ13/Y12: the red and blue icons on the isochrone map in Fig. 11(a) show their geographical locations, respectively.
From the isochrone map view, we can see that the physical travel distance from WW24/XX1 to ZZ13/YY12 is much longer than WW24/XX1 to Bus stop 67852. However, Fig. 12(a) clearly shows that their travel time are almost the same with a difference of just a few minutes. To understand how this happens, we can refer to the detailed mobility information shown along the two selected routes in Fig. 12(a). Here we can find that the main overhead of traveling to Bus stop 67852 is the riding time spent in the bus trip (yellow segment in the figure) since the bus is relatively slow.

We can further explore the temporal variation of the mobility-related factors using the mobility wheels shown in Fig. 12(a), see also their zoomed views in Fig. 12(b & c). In Fig. 12 (b), we can find huge variations of travel time over a day, with three peaks that correspond to the morning peak hour, evening peak hour, and late night period. The interesting observation here is that riding time in late night period does not vary too much, similar to that of the non-peak hours, but the related waiting time suddenly increases. This finding is also confirmed by the local transportation agency since both bus and train service frequencies are halved after 22:30.

On the other hand, the travel time from WW24/XX1 to ZZ13/YY12 shows much less variations compared to the other route. Fig. 12(c) also reveals that the MRT services (green and purple) are relatively more stable over time as compared to the bus services (yellow) shown in Fig. 12(b). Concerning the peak hours in the morning and evening, the local transportation agency told us that the MRT service frequency is doubled during these periods, so waiting time could be reduced. However, the visualizations here show that the waiting time during these peak hours are actually longer than that of the normal periods because passengers may not be able to board a train immediately after reaching the platform during peak periods.

6.3 Experts Review
We interviewed two transportation researchers who specialized in public transportation systems, and obtained their feedbacks of our interface. One of them is from university E (Expert A), while another from a Singapore PTS agency (Expert B). In detail, we first explained to them our system workflow and visual encoding, and then demonstrated to them the two case studies we presented above. They both thought that our interface can be a useful tool for planners and operational managers. Their feedbacks are summarized as follows.

**Interactive Visual Design.** Both of them were impressed by the visual design, especially the isotime map view. Expert B commented “It is an excellent idea to display the information of a public transport system in multiple presentation formats under an integrated and interactive manner. Differing the visualization based on the nature of the information item of major concern would greatly enhance the users’ understanding.” Expert A added “the isotime flow map view makes it very easy to identify the time-shortest routes to all destinations that can be reached within a certain travel time threshold, and parallel isotime model makes it easy for him to compare the travel time and efficiency.”

Both experts appreciated the idea of exploring the PTS mobility. Expert A specifically recognized that the choices to select destinations and visualize the detailed mobility-related factors are very useful. He is particularly interested in visualizing transfer information in our interface, as they are strongly negatively perceived travel elements. Expert B pointed out that the ability to switch between different views can greatly enhance users’ understanding of the major information.

**Improvements.** Both reviewers gave several fruitful comments to improve our system, including providing more visual encoding options for users to select and adding more icons for users to recognize the nodes. Expert A suggested that it might make sense to indicate the level of passenger capacity or the actual passenger numbers along the pathways, instead of showing the number of time-efficient journeys currently in our isotime flow map view. He also recommended us to “show icons or pictograms indicating nearby landmark buildings next to the stops in the isotime flow map view to help user localize the various branches and identify which facilities can be reached.” Expert
service frequencies are halved after 22:30. Stop 67852

Fig. 12. Case Study #2: Analyzing detailed mobility-related factors along two different user-specified journeys starting from MRT station “WW24/XX1” to destinations Bus stop 67852 and MRT station “ZZ13/YY12.” (a) Estimated waiting, riding, and transfer time can be clearly presented along the pathways using the color coding scheme we described in Section 5.3; we can also see together the related MRT and bus trips, as well as the transfer points. (b & c) show the zoomed views of the mobility wheels in (a) for presenting the round-the-clock mobility patterns.

B further pointed out that the multiple presentation formats could be explored to emphasize many transportation factors, like the boarding and alighting patterns at stops.

Applicability. Expert A also commented: “the tool is highly suitable to support location choice decision processes such as the choice of residence or place for setting up a business.” As shown in Case Study 1, the system allows users to evaluate different location options by comparing the ease of traveling with public transport from one to many places. If land use data is also included, it could be extended to be a very powerful tool for site selection in real estate industry. An application example might be that the user would pre-select in what type of places of interest she/he is interested in and the tool would generate an overview of how easy it is to get to various location options.

7 Discussion

The case studies demonstrate the applicability of our interface in showing mobility-related factors in a PTS. Our current model extracts mobility-related information from massive amount of passenger RFID card data, enabling transportation researchers to analyze the efficiency of a PTS based on real data rather than simulations. However, PTS efficiency is affected by many dynamic factors that pose difficulties for transportation researchers to recognize and compare. Hence, our interactive method presents various mobility-related factors in an intuitive visualization, allowing researchers to evaluate and compare travel efficiency, as well as to analyze round-the-clock variations and patterns of these factors. Moreover, transportation experts can explore the relative travel efficiency of a PTS, and apply the results to land-use and transportation planning. Furthermore, such results can also help passengers to make better travel planning through the PTS.

Future work. In our preprocessing stage, we estimate the mobility-related factors from the transportation data sets with certain assumptions. Our PTS model may fail to show MRT interruption events since we use the MRT transit line schedule data in the estimation. Regarding this, we plan to derive improved methods by incorporating more real-world data into the estimation, e.g., platform to platform distances in MRT stations, gate to platform distances, residential and land use data, etc.

Our current method presents only waiting time, riding time, transfer time, and travel efficiency along time-efficient journeys. In the future, we would like to explore other transportation attributes such as vehicle capacity and passenger composition (e.g., senior, student, and disabled). These are also important elements in a PTS. Moreover, we would also like to study the uncertainty of various factors by analyzing the massive RFID card data. Furthermore, we plan to extend our system for real-time analysis of PTS mobility, so that we may deliver adaptive journey planning for passengers. Lastly, we would also like to explore the various suggestions given in the expert feedback.

Limitations. First, our current approach assumes no train interruption, so that we can use the MRT transit line schedule data to estimate the mobility factors. Second, we ignore the initial and final walks taken by the passengers to and from the PTS stops in estimating the mobility factors. Third, our method lacks global perspectives on the transportation data, e.g., congestion patterns that affect certain areas of the PTS at certain times, which is an interesting aggregation condition to be explored. Fourth, our current method focuses on one source to many destinations (one-to-many) rather than many sources to many destinations (many-to-many), which is in fact a very challenging problem. Thanks for the reviewer comment that suggests a global overview to browse through possible origins, we will explore and study about it. Lastly, currently we only consider how fast people move, and ignore other factors that affect people’s choice, e.g., comfort-ness, cost, etc.

8 Conclusion

This paper presents a collaborative work with transportation researchers, aiming to help them visualize and explore passenger mobility in a public transportation system (PTS). In this work, after we define the problem and the analytical tasks, we then introduce and construct a PTS mobility model that characterizes the passenger mobility, and derive methods to estimate various mobility-related factors, including waiting time, riding time, transfer time, and travel efficiency, from the massive passenger RFID card data. Our visual analytic interface is an integrated solution with three visualization modules: isochrone map view, isotime flow map view, and OD-pair journey view, enabling us to efficiently perform the five analytical tasks concerning time-efficient journeys originated from a given starting location. Particularly, the isotime flow map view is a novel visualization strategy, which linearizes a flow map in a parallel isotime layout, thereby presenting clear and smooth pathways from the given origin to destinations as well as maximizing the visualization and comparison of various mobility-related factors along the routes. To come up with this design, we also explore and compare two other design alternatives. Moreover, we propose the visual strategy called the mobility wheel in the OD-pair journey view for examining round-the-clock temporal variation with all the contributing mobility information. In the end, we also explore two case studies with the transportation researchers, and present their expert feedbacks on the interface design.

Acknowledgments

We thank all reviewers for their effort and constructive comments, and also the two transportation experts who provided us with the expert review. This work was mainly established at the Singapore-ETH Centre for Global Environmental Sustainability (SEC), co-funded in part by the Singapore National Research Foundation (NRF) and ETH Zurich, the MOE Tier-2 fund (MOE2011-T2-2-041 (ARC 5/12)), the MOE Tier-1 fund (RG 29/11), and the National Basic Research Program of China (973 Program) under Grant No. 2014CB340304.
References


