Assessing Transit Unrealized Accessibility with Space-time Prisms

Luyu Liu, Adam Porr, Harvey Miller

Uncertainties in public transit systems’ travel time have been a major obstacle to make transit more accessible and reliable for commuters. Due to delays caused by traffic and road conditions, the actual accessibility derived from real-time data can be very different from the one promised by the schedule. However, very few prior accessibility studies addressed this discrepancy explicitly with time geography methods. In this paper, we use a well-established time geography method, space-time prism (STP), to measure the *accessibility reliability*. Accessibility reliability is defined as the difference between STP derived from retrospective real-time data and STP derived from schedule data. The methods will use two mobility datasets of large volumes: General Transit Feed Specification (GTFS) real-time data, which produce retrospective real-time STP, and GTFS schedule data, which produce scheduled STP. We will also investigate the reliability measure’s connections to social equity factors, such as different social, demographic, and economic factors. The paper will conduct a case study in the Central Ohio Transit Authority (COTA) bus system, a public transit agency in Columbus Ohio. The analysis will focus on the spatial and temporal patterns of the reliability measure from 2018 – 2021 across the city of Columbus, especially the changes before and during the COVID-19 pandemic. This can provide insights about possible impacts of the pandemic on the reliability of the transit accessibility in different communities. All the analyses and results will be visualized in a public web-based platform. This paper provides a scalable time-geography approach to gauge the reliability of transit accessibility with very large datasets; some results can also reveal new empirical patterns of transit accessibility’s impact on social equity.

1. Introduction

As transit planning’s focus shifting from mobility to accessibility (Banister, 2008), accessibility becomes a crucial determinant to promote transit policies and foster transit use. Accessibility measures the ease of reaching destinations: a transit system with higher accessibility can have better usability and user experience for transit users; therefore, authorities aim to increase accessibility as one of their priorities. As public schedule dataset like General Transit Feed Specification (GTFS) becomes more available, researchers and transit planners start to calculate accessibility with static scheduled data. Standard data formats introduce a convenient way to access the transit system’s network information, including stops as nodes, trips as links, and scheduled travel time as travel cost. This has greatly helped us understand the accessibility of transit system.

However, accessibility of transit systems is highly dynamic and time-dependent due to traffic condition and delay that the actual performance can be very different from the schedule (Park, Mount, Liu, Xiao, & Miller, 2020). Many prior studies have examined the possible problems of using scheduled data for modelling and benchmarking (Bills & Carrel, 2021; Wessel, Allen, & Farber, 2017; Wessel & Farber, 2019): actual accessibility can have significant deviations from the accessibility calculated from schedule data. There are several factors that contribute to this deviation: first, road conditions and congestions can slow the vehicles down and human operators can leave the stops earlier than the scheduled time. Second, only travel time at timepoint stops is explicitly defined in the official timetables of many transit systems; travel time at non-timepoint stops is derived from interpolation, which is not strictly followed in practice.

This paper introduces a time geography approach based on space-time prism (STP) to understand *accessibility reliability*, retrospective accessibility’s deviation from scheduled accessibility. This measure represents the difference between the actual physical accessible space and the expected accessible space calculated from schedule data during the same time budget; the aggregated version of this measure can also show consistency and quality of the transit service, which is vital for administrative and planning purposes. We use GTFS schedule and GTFS real-time data, then produce retrospective real-time STP and scheduled STP, and finally derive the difference between the two STPs. The origin-level deviation measure has been reported to be non-uniform (Wessel & Farber, 2019) and non-random (Wessel et al., 2017); however, few studies examined its connections to social equity factors, such as different social, demographic, and economic factors. This paper fills in this blank by using geostatistics model to investigate the potential correlations.

We conduct a case study in the Central Ohio Transit Authority (COTA) bus system, a public transit agency in Columbus Ohio. The analysis focus on the spatial and temporal patterns in different levels from 2018 – 2021 across Columbus, especially during the COVID-19 pandemic. This paper provides a scalable time geography approach to measure the reliability of transit accessibility with very large datasets and investigate its implications on social equity.

1. Background

We discuss the background of this paper in this section from three perspectives: 1) the evolution of space-time prism, 2) the development of transit accessibility measurements, and 3) unreliability of schedule-based accessibility measures.

* 1. The evolution of space-time prism

Space-time prism (STP) is a well-established time geography method to measure physical accessibility in public transit systems (Miller, 1991; Wu & Miller, 2001). We can witness an obvious trend of STP framework progressing with increasing data volume and data standardization, with which STP can gauge accessibility with less error and more details. Hägerstrand (1970) first proposes the conceptual framework of STP and then Lenntorp (1976) provides the very first operational definition of space-time physical accessibility measures. With the rising popularity of geographic information system (GIS), H. J. Miller (1991) first introduces a generic GIS-based procedure to derive space-time prism concepts. H. J. Miller (1999) later introduces computational procedures to apply STP to transportation network. However, these classic time geography works can be conceptually abundant but may not provide enough analytical insights (Miller, 2017). Despite later measures become more people-oriented and detailed, the data collection process is still inconvenient and inconsistent. For example, Kwan (1999) explores individual accessibility and applies the method in Franklin County, Ohio with a travel diary dataset manually collected by the author.

New data technologies motivate more analytical and empirical studies, such as Global Positioning System (GPS) (M.-P. Kwan, 2000) and Automatic Vehicle Location (AVL) (Tang, Song, Miller, & Zhou, 2016). This is a further step towards more comprehensive measurement of accessibility with STP. More abundant data also motivate more conceptual innovations and more diverse and detailed interpretations of human mobility pattern. For example, Lee & Miller (2019) introduce an analytical time geographic method to calculate a collective-level representative STP. Lee & Miller (2020) also introduce a robust STP to incorporate travel time uncertainty in accessibility measurement.

* 1. The evolution of Transit accessibility measurements

Better datasets also motivate the innovations of transit accessibility measurements. Malekzadeh & Chung (2020) conclude there are two major trends for transit accessibility studies: 1) towards capturing travelers’ behavior and their stochasticity; 2) towards developing more disaggregated transit accessibility measurements. Both trends exemplify how larger, more detailed, and more accessible datasets impact the formulation of transit accessibility models.

Due to its multimodal and nonlinear nature, early transit accessibility models usually adopt simple assumption based on travel time estimations, which significantly reduces their computation load (Malekzadeh & Chung, 2020). For example, some early system accessibility models only consider the proximity to transit stops (Hsiao, Lu, Sterling, & Weatherford, 1997; Zhao, Chow, Li, Ubaka, & Gan, 2003), which is a major simplification. As transit-related datasets become more detailed and accessible, models can better capture the travelers’ behavior and their stochasticity, such as system-facilitated models (Tribby & Zandbergen, 2012) and access-to-destination models (Farber, Bartholomew, Li, Páez, & Habib, 2014; Owen & Levinson, 2015). More detailed models can provide a better understanding of users’ actual travel experience from users’ perspective.

Another trend of transit accessibility is more disaggregated transit accessibility measurements. Larger and detailed datasets, higher computational ability, and better visualization methods help to improve the fidelity of transit accessibility. The rise of standard data format General Transit Feed Specification (GTFS) also marks another boom of accessibility studies (Wessel & Farber, 2019). GTFS data have a well-defined structure for scheduled data and are often released publicly by transit authorities (Barbeau & Antrim, 2013). Therefore, many recent studies use GTFS to derive STP in a larger scale without compromising the fine details of transit systems (Lee & Miller, 2018; Tasic, Zhou, & Zlatkovic, 2014).

* 1. Unreliability of schedule-based accessibility measures

As recent studies focus more on capturing users’ stochasticity, unreliability becomes the center of the discussion: how well can the accessibility measurement capture the actual experience of a user in the system? Unreliability can be defined as an accessibility measurement’s deviation from a standard benchmark, which usually represents the actual or experienced accessibility. Due to the lack of accessible real-time data source, most traditional accessibility measures are calculated based on transit schedule (Wessel & Farber, 2019); therefore, many schedule-based accessibility measures may have very large unreliability.

Unreliability can be summarized into two main factors: *uncertainty* and *accuracy*. *Uncertainty* refers to the stochastic variation of the accessibility measure, due to on-time performance and measuring error. Public transit systems are constantly changing – i.e., early or late time because of unexpected external or internal factors, such as traffic, weather, vehicle conditions, or operator conditions. Hall (1983) was among the first to consider uncertainty when formulating and calculating accessibility. Similar to the development of STP, more studies are dedicated to discussing the unreliability of accessibility measures with better datasets. For example, Kim & Song (2018) discuss an integrated measure of accessibility and reliability for transit systems; Zhang, Dong, Zeng, & Li (2018) introduce a time-dependent reliability modelling approach based on GPS trajectories to address traditional measures’ overestimation problem.

Another factor that can contribute to a schedule-based accessibility measure’s unreliability is *accuracy*. It can be defined as the systematic deviation of an accessibility measure from the standard benchmark. Some papers discuss the topic with empirical evidence: Wessel et al. (2017) constructed a retrospective transit timetable from real-time automatic vehicle location data to better capture the dynamic nature of transit system. The paper also provided a case study in Toronto Transit system and pointed out that real-time based accessibility does have significant deviation from the scheduled, and the pattern of the deviation does not seem random. Wessel & Farber (2019) moreover explore the accuracy of schedule-based accessibility in Jacksonville, Massachusetts Bay, and San Francisco. The paper concluded that scheduled-based accessibility measures overestimate on average by 5 to 15 percent or more, and it may not be sufficient to use schedule data alone to access transit accessibility for most transit systems.

Both factors can contribute to

1. Method

We introduce the method to calculate the accessibility unreliability in this section. We first introduce the two main transit datasets we use in this paper. Then, we demonstrate a time-dependent Dijkstra algorithm to calculate the two versions of space-time prisms. Last, we propose the data source and geostatistic model to investigate the connection between accessibility reliability and other social and demographic factors.

* 1. Data source

We use General Transit Feed Specification (GTFS) data as the main data source for time geography analyses in this paper. GTFS is a data standard that let transit authorities publish transit data and developers/researchers consume the data (Google Developers, 2020). GTFS includes two parts: GTFS static and GTFS real-time data.

We first use GTFS static data to acquire the transit schedule data. GTFS static data has been the *de facto* standard for transit schedule and associate geographic data (Google Developers, 2020). The data are organized into several separate relational tables, which define a transit system’s stops, trips, routes, arrival and departure time, and other schedule information (Google Developers, 2020). GTFS static data are very convenient for researchers and developers to obtain complete, systematic, and official schedule data directly from the transit authorities.

We collect both GTFS static and real-time trip update data from the official application programming interface (API) of the Central Ohio Transit Authority (COTA) from February 2018 to July 2020 (Central Ohio Transit Authority, 2021). We save the GTFS static data whenever there are any changes in the dataset. Besides the major schedule adjustments such as the route redesign in 2018 (Lee & Miller, 2018; Schmitt, 2018) and COVID-19-related adjustments in 2020 (Liu, Miller, & Scheff, 2020) as well as the three seasonal schedule adjustments in every January, May, and September, we also record every minor change that happen day by day.

We use GTFS real-time data to obtain the real-time or retrospective time in the system. GTFS real-time extension is a feed specification for transit authorities to provide real-time updates about their fleet to the public (Google, 2021). GTFS real-time data include two main datasets: *trip update*, which contains the expected arrival/departure time of each trip at each stop in the transit system, and *vehicle position*, which is similar to automatic vehicle location (AVL) data and shows the location of active vehicle in the system. GTFS real-time data are broadcasted by transit authorities at a regular interval from 10 second to 90 seconds (Liu & Miller, 2020a). We derived the actual arrival time of each trip at each stop from trip update data for: 1) it is very convenient and has relatively high accuracy; 2) most trip planning apps, which are the source of transit information for many passengers, are using trip update as the source for arrival time.

We also collect GTFS real-time trip update data from COTA’s official API. We collect trip update feeds at the interval of 60 seconds, which is a common frequency for US transit systems (Liu & Miller, 2020a). The timespan covers February 2018 to July 2021 and the total data volume exceeds 1 terabyte. It is challenging to maintain such a huge dataset and conduct scientific calculations in traditional relational database environment; therefore, we use noSQL and MongoDB database to adapt to these difficulties.

* 1. Time-dependent routing problem

As we introduced in the last section, we use space-time prism, a well-established time geography method, to measure accessibility in public transit systems (Miller, 1991; Wu & Miller, 2001). In practice, we first calculate the shortest travel time between the origin stop to other stops to produce the contour lines of different travel time at an equal interval, and then derive the STP.

However, it is hard to obtain accurate travel times in a transit network, even with a complete archive of retrospective arrival times. A major reason is because transit networks are *discontinuous* and *time-dependent* (Gendreau, Ghiani, & Guerriero, 2015; Wang, Yuan, Ma, & Wang, 2019). Unlike private vehicle or pedestrian network, a user cannot move in the network unless there is an available vehicle, which is scheduled to arrive at only some specific time points. Therefore, the network costs of a transit work can vary depending on the passenger’s arrival time at the origin node of a transit link. This variation also applies to components of the travel time, including wait time and in-vehicle time.

The time-dependent routing problem can be generally categorized into two models: deterministic and stochastic model (Gendreau et al., 2015). Stochastic models use a stochastic factor to measure or predict the time-varying travel time. They are more useful to capture the randomness caused by congestion, weather, accidents, and road maintenance (Gendreau et al., 2015); however, due to the random nature of these models, the results include many uncertainties, even with deterministic and retrospective travel time records. Because we collected all the arrival times at all the stops and aim for more accurate travel time, we use a deterministic approach to address the time-dependent routing problem.

We use a modified Dijkstra algorithm with dynamic costs to solve the problem. Dijkstra algorithm is a classic and efficient algorithm to solve the shortest path routing problem (Golden, 1976). It uses a greedy heuristic search strategy to find the shortest path from the origin node to every other nodes (Xie, Zhu, Yan, Yuan, & Zhang, 2012), which is very useful and efficient to calculate the STP.

However, traditional Dijkstra algorithm’s correctness is based on non-negative static costs, which time-dependent transit networks do not satisfy. Due to *passing*, when a later start time may result in an earlier arrival time (Gendreau et al., 2015), the results generated by Dijkstra algorithm with dynamic costs may not be the global optimal answer. Therefore, many prior studies introduced *no-passing* or *first-in-first-out (FIFO)* rule to make Dijkstra algorithm compatible with the time-dependent requirement (Ahn & Shin, 1991; Ichoua, Gendreau, & Potvin, 2003). FIFO rule assumes a vehicle leaving the origin stop earlier will never arrive later at the destination stop than another vehicle. FIFO rule is a prerequisite to use Dijkstra to calculate routing problem in a transit system.

Therefore, we test if COTA system satisfies the FIFO rule. We calculate whether each bus in the transit system can indeed pass the subsequent bus in the same route. The average ratio of no-passing bus is [ADD NUMBER]; therefore, we can conclude that there are very few passing occurrence in the COTA system, and the FIFO rule generally applies to the system.

* 1. Three space-time prisms

We calculate the shortest travel time between any stops in the system with the introduced algorithm based on the scheduled and retrospective GTFS data, and then produce space-time prisms (STPs). We derive STP by calculating the number of accessible bus stops. We first introduce a decision variable to represent whether a user starting from stop at time point can arrive at another stop within the time budget :

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

Where: is the shortest travel time between stop and . Therefore, the number of accessible stops with the time budget can be written as:

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

Where: represents the number of accessible bus stops from stop at the time point with the time budget , and is the set of all stops. We can then introduce the definition of STP:

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

Where: represents the STP from stop at a time point , while is the set of all time budgets.

However, despite we have a time-dependent deterministic algorithm to calculate accurate shortest travel time, the actual travel time can still vary depending on the data source and the trip planning strategy (Liu & Miller, 2020a). Therefore, we produce three versions of travel time and corresponding space-time prism: scheduled, posteriori real-time, and priori real-time STP.

**Scheduled STP.** Scheduled STPs are calculated based on the scheduled time and the GTFS static dataset. It represents the expected physical accessibility that a passenger can achieve in theory. However, like we discussed above, the actual travel time and accessibility may vary due to on-time performance deviations; therefore, scheduled STP may be very different from the accessibility experienced by an actual passenger.

**Posteriori real-time STP**. As we can access all the historical arrival time from GTFS real-time archive, we can calculate a retrospective version of STP with the same algorithm by changing all the scheduled arrival time to corresponding retrospective real-time arrival time (Wessel et al., 2017; Wessel & Farber, 2019). It represents the actual physical accessibility that a passenger can achieve in theory. However, when a user is planning her/his trips, the user may not know the actual arrival time of each bus as a *priori* (Wessel & Farber, 2019).

Therefore, we name this retrospective version of STP *posteriori real-time STP*. *Posteriori* means a user needs to have the experience or knowledge of the event to make the decision, and the STP is calculated from information that can only be obtained *after* the event happens, such as the actual arrival time. Despite it can be a useful reference for transit agencies and users, *posteriori STP*, or more generally *posteriori accessibility*, may not be feasible for actual users. It can overestimate users’ accessibility because it assumes users have omniscient knowledge of the transit system, even events that happen in the future. There are some very unnatural and infeasible results caused by this overestimation: in the posteriori model, a user can decide to take a very different combination of trips and routes that will not be possible without predicting the future. For example, a bus is delayed so much that the user can use it to catch other unexpected transfers to save significant amount of time. This is a possible combination for a scientist reviewing a GTFS archive; however, a user or a transit administrator usually cannot know this is a possible trip during operation.

We can deconstruct the scheduled and posteriori real-time accessibility from a perspective of decision-making: the mentioned two accessibility systems do not separate the decision-making and implementation process. For a user, the decision-making process always happens *before* the implementation process since people always plan their trips before actually taking the transit. However, the two accessibility systems naturally assume the two processes are happening simultaneously: the user can always know, board, and finish all the trips without any troubles that can incur extra waiting time or risk of missing a vehicle. Such an assumption is very unrealistic (Liu & Miller, 2020b; Park et al., 2020).

**Priori real-time STP.** We therefore define *priori STP* to better simulate transit users’ actual decision-making process. *Priori* means the user does not have to experience the event to make the decision, and the STP is calculated from only information that is obtained *before* the event happens, such as the scheduled time or expected time. From the perspective of decision making, the trip planning process happens *before* the event happens, which makes it possible for a user to implement.

Priori real-time STP can also vary depending on which trip planning strategy the user may take (Liu & Miller, 2020a). In this paper, we will assume the user will schedule all the trips based on the schedule, while implementing these trips in the transit system with actual arrival time. Therefore, the trajectories of scheduled and priori real-time STP are exactly the same but with different travel time.

In practice, the calculation of priori real-time STP involves two phases to simulate transit users’ two-step transit experience. We first calculate the scheduled STPs and their corresponding trajectories and travel time between stops. Then, we revisit each trajectory and recalculate each synchronization between buses with actual arrival time. For example, in the scheduled STP between two stops, the trajectory is {A, B, C}, where A, B, C represent subsequent stops. The user is scheduled to take bus 1 from stop A to B, then transfer at stop B to another bus 2, and finally arrive at stop C. However, because bus 1 is delayed, the user arrives late at the transfer stop B and misses the scheduled transfer bus 2. We will then find the next bus from stop B to C and record the new arrival time at stop C and travel time between stop A and C.

Priori real-time STP is a more conservative and realistic accessibility measure compared with its scheduled and posteriori real-time counterparts; it can better represent actual users’ transit experience.

* 1. Unrealized accessibility

Scheduled STP can be perceived as the promise the transit system makes, while posteriori and priori STPs are the actual experience the transit system delivers. The Difference between the expectation and the experience can be defined as the *unrealized accessibility*, which represents the part of accessibility the transit system loses during the operation compared with the schedule. Based on the STP definition we give, we can define unrealized accessibility as:

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

Where: is the scheduled STP, is the posteriori or priori real-time STP, is the scheduled travel time, is the posteriori or priori real-time travel time.

We also introduce normalized unrealized accessibility:

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

1. Analysis
   1. Spatial pattern of unrealized accessibility

We introduce two types of unrealized accessibility - posteriori and priori unrealized accessibility. We present the spatial patterns of both measures in this section.

* 1. Temporal pattern of unrealized accessibility