Measuring risk of missing transfers in public transit systems using high-resolution schedule and real-time bus location data

Abstract

The emergence of urban Big Data creates new opportunities for deeper understanding of transportation within cities, revealing patterns and dynamics that were previously hidden. Public transit agencies are collecting and publishing high-resolution schedule and real-time vehicle location data to help users schedule trips and navigate the system. We can use these data to generate new insights into public transit delays, a major source of user dissatisfaction. Leveraging open General Transit Feed Specification (GTFS) and administrative Automatic Passenger Counter (APC) data, we develop two measures to assess the risk of missing bus route transfers and the consequent time penalties due to delays. *Risk of Missing Transfers* (RoMT) measures the empirical probability of missed transfers and *Average Total Time Penalty* (ATTP) shows overall time loss compared to the schedule. We apply these measures to data from the Central Ohio Transit Authority (COTA), a public transit agency serving the Columbus, Ohio, USA metropolitan area. We aggregate, visualize, and analyze these measures at different spatial and temporal resolutions, revealing patterns that demonstrate the heterogeneous impacts of bus delays. We also simulate the impacts of dedicated bus lanes reducing missing risk and time penalties. Results demonstrate the effectiveness of measures based on high-resolution schedule and real-time vehicle location data to assess the impacts of delays and guide planning and decision making that can improve on-time performance.

Keywords: Public transit; risk of missing transfer; General Transit Feed Specification (GTFS) data; Automatic Passenger Counter (APC) data

1. Introduction

Transfers between routes are often necessary in public transit systems. The expansion of city footprints can make direct routes difficult to provide (Knoppers and Muller, 1995). Configuring systems that include route transfers allows public transit providers to cover more space and time with fewer vehicles (Walker, 2012). However, transit delays, defined as a positive deviation of a transit vehicle's actual arrival time from the scheduled time, are inevitable due to traffic, equipment malfunctions, external events and other circumstances. Transit delays causing users to miss intended transfers between routes imposes potentially significant time penalties on users, making the system less functional and desirable to use.

Until recently, scientific analysis of transit transfers has been limited, focusing on the users' experience and transfer node design (Guo and Wilson, 2004, 2011; Sun et al., 2007, 2010), with limited investigation into real-time behavior and system performance (Jang, 2010; Nesheli and Ceder, 2015; Nishiuchi et al., 2015). However, the context for scientific understanding of public transit is changing due to the emergence of location-aware and wireless communication technologies. These technologies allow public transit

agencies to capture real-time data on vehicle locations across the entire transit system. Many agencies are publishing these data along with schedule data to enable public transit webpages and apps that make navigating the system easier for users. These data can also be leveraged to enable new scientific insights into public transit dynamics, including public transit transfers and the impacts caused by delays.

In this paper, we develop measures for the evaluation of transfer performance in public transit systems using high-resolution schedule and real-time vehicle location datasets. *Risk of Missing Transfers* (RoMT) measures the empirical probability of missing transfers based on historical data, while the *Average Total Time Penalty* (ATTP) shows overall time loss compared to the scheduled trip. Unlike many former composite scores, RoMT and ATTP are intuitive to compute and understand; the measures are also easier to aggregate into different levels of spatial and temporal resolution and expand to other systems with real-time support. We implement these measures using high-resolution schedule and real-time vehicle location data from the Central Ohio Transit Authority (COTA) bus system in Columbus, Ohio, including both open data published in the General Transit Feed Specification (GTFS) format and administrative data derived from Automated Passenger Counters (APCs). We explore the patterns of RoMT and ATTP for different spatial and temporal resolution, and simulate the impact of dedicated bus lanes on RoMT and ATTP. The results demonstrate the ability of RoMT and ATTP measures to leverage large-volume big data and assess the impacts of delays on transfers; the measures can not only guide planning and decision making to improve on-time performance, but also provide important information for ordinary users about transfers' empirical performance.

In the next section of this paper, we review previous research on transfers from the perspectives of the data utilized and the analysis conducted. In the subsequent section we introduce our data sources and methods. Following this, we show the results of the spatial and temporal analyses. We conclude this paper with a discussion of the strengths and limitations of this study, and steps for future research.

2. Background

This literature review covers two dimensions of the development of measuring and analyzing public transit transfers: data and measures. We first discuss two types of data sources: *small data* and *big data*. Following this, we discuss existing measures that use these data for transfer measurement and their different tasks and benefits.

2.1. From small data to big data

Small data. Traditionally, studies of public transit transfer properties and behaviors use data collected for specific research questions, often using dedicated and survey instruments and location tracking via Global Positioning System (GPS) receivers. We call these datasets *small data* not only due to smaller volume but also since it is often collected for a specific purpose. While small data are valuable and easy to interpret, there are several issues that limit their usefulness. First and most importantly, although most small data are carefully sampled, they cannot provide blanket coverage of an entire system, making it

difficult to see broader temporal and spatial patterns. Since most small data are purposely created, they are also expensive and time-consuming to collect. For example, Guo and Wilson (2011) created and maintained special purpose station inventory and field survey databases. This requires substantial time and resources, often for a relatively small volume of data.

Another issue is the lack of universal standards and definitions, making comparisons difficult. Different transfer studies have varying definitions of transfers and their data (Guo and Wilson, 2004), limiting comparability. In addition, sampling frames are fragile, meaning that data collected for one set of questions cannot easily be repurposed for other questions (Miller and Goodchild, 2015).

An example of small data is stated preference (SP) data, used widely to support mode choice models (Guo and Wilson, 2011). Although many transfer assessment studies use SP data, the choice dimension is typically small, meaning that SP data may not be able to capture the full diversity of transfer situations (Bovy and Stern, 2012). Other semi-quantitative data collecting methods, such as on-board questionnaires, can also lack precision and reliability. The result of these imprecise data sources is that most studies provide a partial assessment of the system since it is difficult to have a detailed assessment at high spatial or temporal resolutions across the entire system (Guo and Wilson, 2011).

Big data. In the past, detailed, real-time performance data about public transit was difficult to acquire (Dessouky et al., 1999). However, this has changed due to the development of new data collection and sharing technologies. The widespread application of new information and communication technologies (ICTs) provide the technical support for what is often labeled *big data* (Hilbert, 2016). The definition of big data is diverse; a commonly accepted definition encompasses the "three Vs": large volume, high variety, fast velocity (Chen et al., 2014). However, as Miller and Goodchild (2015) argue, in many applications, especially in urban science, the unique and valuable characteristic of Big Data is ubiquity: its widespread coverage and availability, often as a byproduct of digitally-enabled operations and activities.

In public transit, inexpensive GPS receivers and wireless communication allow widespread tracking of vehicle locations in real-time. These data are collected automatically on an ongoing basis by public transit authorities, meaning they are readily available without additional and prohibitive cost or effort. Meanwhile, the World Wide Web combined with data services allow sharing schedule and real-time vehicle location data. This technology revolution allows the possibility of more detailed investigation of transfer performance across an entire transit system.

However, big data also has limitations. As suggested by the characteristic "high variety," these data are often heterogeneous in terms of structure, quality and support: diverse data sources, lack of metadata, and lack of quality control all make big data challenging from a scientific perspective (Miller and Goodchild, 2015). Accordingly,

standard protocols for sharing transit schedule and real-time data, such as General Transit Feed Specification (GTFS) (Google Developers, 2016, 2018), help to solve these problems.

2.2. From non-real-time measures to real-time measures

Non-real-time measures. Along with the data source's development from small data to big data, we also witness the progress of measures from non-real-time measures to real-time measures. Non-real-time measures have a relatively low *temporal accuracy*; similar to Firmani et al. (2016)'s definition, we define this as: how accurate is the measure's recorded time compared to the actual time. It represents the systematic error caused by the temporal delay of measurement.

For example, many non-real-time measures do not measure actual performance. Instead, they try to gauge static features, like *transfer nodes' design* and *transfer connectivity*. For example, Guo & Wilson (2011) assess transfer cost based on both users' and operators' perspective; they develop an index that measures each transfer node's effectiveness based on additional transfer time and fare and apply it to the London Underground system. Hadas & Ranjitkar (2012) combine transfer connectivity and travel time to represent the effectiveness of transfers. Although the non-real-time measures have been proven to be extremely useful to assess the static qualities of system design, the results, based on schedules alone cannot represent the actual performance.

Some non-real-time measures also use second-hand data sources like stated preference surveys; their temporal accuracy is also low since the surveys are usually conducted long after the actual trips have been undertaken. For example, many researchers analyze users' perceptions and attitudes about transfers (Guo and Wilson, 2004; Liu et al., 1997). These studies focus on measuring *transfer penalties*, namely how much and why people prefer not to take transfer trips; these penalties encompass a broad range of factors such as walking time, number of transfers in a single trip, waiting time, ticket fare, and other environmental factors.

Real-time measures. On-board questionnaires can be considered as the first real-time measure, however, they are not widely used to study transfers due to the difficulties of acquiring abundant data. After the emergence of big data, automated big data real-time measures based on first-hand information with *high velocity* and *large volume* create new opportunities for transfer studies. For example, Nishiuchi et al. (2015) used data envelop analysis to derive measures to evaluate the efficiency of user transfers across transportation systems. The measure concentrates on the transfer stations' commuting efficiency using users' smart card real-time data.

Real-time measures have two major advantages. First, the recorded time has higher trueness. The recorded time is closer to the time of actual events: this is especially important for temporal analyses. Second, as we discussed in the last section, the measured value has higher trueness. Given the same measuring precision, since higher temporal accuracy can reduce temporal systematic error, it also suggests higher value accuracy. However, few research studies prior to the advent of Big Data have sought to assess

transfers based on real-time performance and missing risk due to delays. Progress in data availability, real-time monitoring and other smart city technologies are making this topic an area of active investigation: researchers can now conduct more detailed analysis and develop more precise measures and models of public transit transfer performance (Hadas and Ranjitkar, 2012; Kujala et al., 2018). Real-time measures also requires more responsive in-situ censors, additional data supports such as standard format and data streaming pipeline, and administrative support. All of these requirements take extra technological and economic costs.

From this review, we conclude that big data and real-time measures are the future direction of transfer studies. This paper contributes to this literature by developing measures of missing risk and transfer time penalties using high-resolution real-time big data sources. The new measures we propose are amongst the first to focus on transfers' real-time performance due to delay and the first to use bus systems' actual real-time big data to calculate it.

Our new measures can demonstrate detailed patterns for any geographic and temporal resolution. While the small data sources can only present a homogeneous average pattern (Guo and Wilson, 2011), smart card data can analyze patterns during different hours (Nishiuchi et al., 2015) for both single specific and aggregated trips, but also on a daily, weekly, or monthly basis.

3. Methodology

This section discusses the methodology in our study. We first describe our data sources; then we define public transit transfers from a space-time perspective and conceptualize the impact of vehicle delays as a problem in transfer synchronization. Then, we discuss the methods involved in missing risk measurement and analysis.

3.1. Data sources

In this paper, we leverage two datasets for measures and analytics.

General Transit Feed Specification (GTFS) data. General Transit Feed Specification (GTFS) is a combination of two data standards: *GTFS static* and *GTFS real-time expansion*. GTFS static reports the schedule data for a public transportation system. GTFS static is now the *de facto* standard for public transportation schedules and associated geographic data (Google Developers, 2016). Public transit system administrations are encouraged to share their GTFS static publicly: by 2010, almost 85% of transit miles traveled in the U.S were covered by open data published by transit authorities (Antrim and Barbeau, 2013).

GTFS real-time expansion provides frequently updated vehicle location data. GTFS real-time includes two components: vehicles' location and the trip updates, which contains vehicles' arrival and departure time at every sequential stop. Moreover, the temporal resolution can be as high as 1 min (Kujala et al., 2018). GTFS overcomes the disadvantages of both traditional data and unclean big data: it is high-volume, frequently updated, publicly

accessible, standardized format, and covers the entire public transit system. These features make GTFS a good big data source to conduct real-time measures.

However, despite all the valuable features, GTFS data has limitations. In particular, its temporal accuracy is lower than other sources such as automated passenger count data (discussed below). This is because GTFS real-time data feeds are updated based on a fixed temporal interval, not based on the actual events of a public transit vehicle entering and leaving a designated stop. Consequently, the actual arrival/departure time may be different from the times recorded in the GTFS data.

Automated Passenger Count (APC) data. Automated Passenger Counting (APC) data are generated by devices that are installed on vehicles to track and report transit ridership (Chu, 2010; Transit Wiki, 2019). These data often contain arrival time and departure time at each stop. A major advantage of APC data compared to GTFS data is the arrival and departure time is recorded at the stop as the events occur instead of being updated according to a specified temporal interval. However, APC data are not open. As administrative data, APC data are not available for the public and transit mobile applications. Moreover, APC data does not typically offer widespread coverage of the whole system. Instead, a subset of public transit vehicles are installed with APC devices rather than blanket coverage as with GTFS data.

To leverage the positive features of both data sources, we merge raw APC data and GTFS data into a new combined dataset. We enumerated all GTFS trips while querying the APC database: if the corresponding trip is in the APC database, we override the record in the GTFS database to take advantage of the higher temporal accuracy of the APC data. Below, we will provide results based on the merged APC-GTFS dataset.

3.2. Transfer synchronization, desynchronization and time penalties

Public transit transfers link *generating trip* and *receiving trip*. A user first boards a bus to start the generating trip, then alights to catch the next bus to start the receiving trip. The transfer itself can be characterized as: i) a street-crossing transfer; ii) a sidewalk-based transfer, and; iii) a non-walk transfer at the same stop (Hadas and Ranjitkar, 2012). Based on this categorization, we can generalize the transfers as: i) *non-walking transfer*, which does not require a walking for the transfer, and; ii) *walking transfer*, which requires walking from the generating trip's stop (which we label the *generating stop*) to the receiving trip's stop (the *receiving stop*).

Synchronization and desynchronization. We further conceptualize transfer as a process of synchronization: i) the generating trip brings passengers to the generating stop; ii) users then transition to receiving stop; iii) the receiving trip picks up passengers at receiving stop.

Transit delays can result in inconsistent arrival and departure times hence the desynchronization of scheduled generating trip and receiving trip. For each transfer, we can measure the time penalty when the receiving bus is leaving; this is the time point when

the desynchronization happens. Due to desynchronization between the generating and receiving trip at the receiving stop, the actual transfer can differ from the schedule according to the relative temporal order of the two trips arrival/departure time. Figure 1 illustrates this process using a time-space diagram.
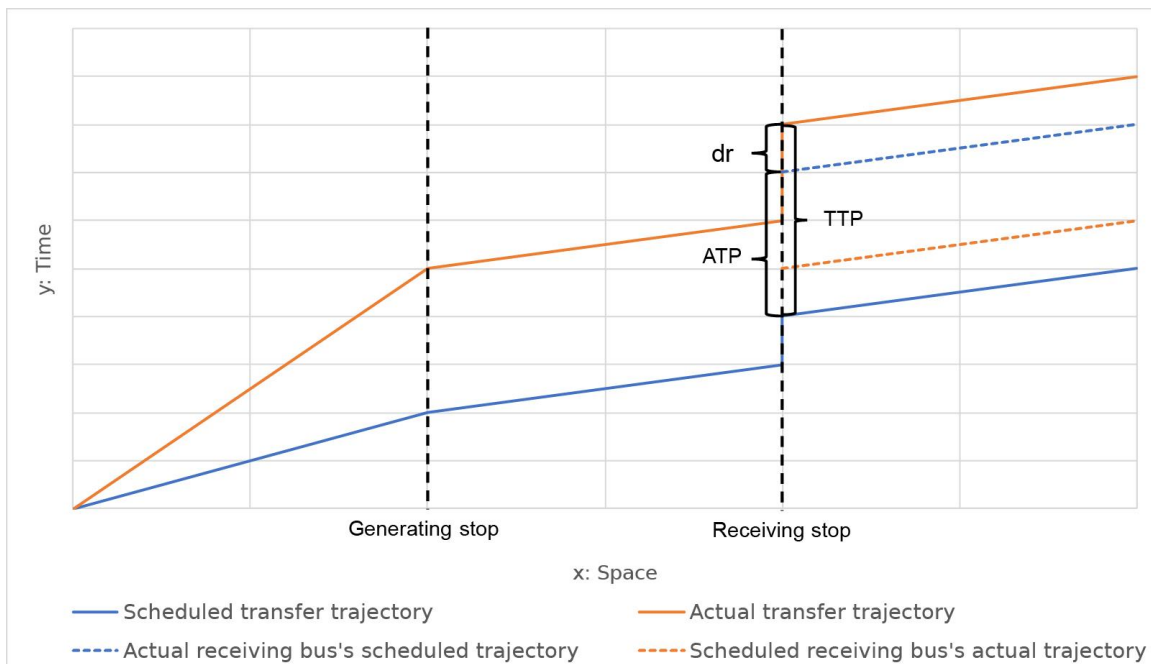

Figure 1: Time-space diagram of a delayed transfer and the corresponding scheduled transfer

Due to the desynchronization, the actual receiving bus can be different from the scheduled receiving bus. We can conceptualize the schedule of all buses running on the same route as an array of trips (a *trip sequence array*). We assume the passenger will always take the first available bus. If the generating bus is sufficiently late, the passenger will miss the scheduled bus and need to take a later scheduled bus. Likewise, if the receiving buses are sufficiently late, the passenger can catch an earlier receiving bus in the trip sequence array. The *desynchronization degree* (DD), an integer variable, measures a transfer's desynchronization in the trip sequence array: it represents the order number of the actual bus before/after the scheduled bus. For example, if the actual bus is the *n-th* bus after the scheduled bus, the DD is *n*; if the actual bus is the *n-th* bus before the scheduled bus, the DD is *-n*; if the actual bus is the scheduled bus, then the DD is 0.

Transfer time penalties. We calculate two types of potential time penalties for each transfer. The first is *total time penalty* (TTP):

$$TTP = T_r^n - t_r^0 \tag{1}$$

where: T_r^n is the actual departure time of actual receiving bus (DD = n) and t_r^0 is the scheduled departure time of scheduled receiving bus (DD = 0). TTP represents the total

time loss compared to the schedule at the receiving stop. The value shows the total delay when the receiving trip starts, which encompasses both the generating bus and receiving bus time loss. However, since the synchronization process involves two vehicles, it is also important to determine the corresponding time loss caused by each bus. For example, a large TTP could be because of the receiving bus's large delay but the synchronization is not disturbed; on the other hand, a large TTP could be also because of the first bus's delay, which results in desynchronization and thus long waiting time.

To quantify these two different types of delay, we decompose TTP as follows:

$$TTP = T_r^n - t_r^0 = (T_r^n - t_r^n) + (t_r^n - t_r^0) = d_r + (t_r^n - t_r^0)$$
$$= d_r + ATP$$

(2)

where: t_r^n is the scheduled departure time of the actual receiving bus (DD = n), t_r^0 is the scheduled departure time of the scheduled receiving bus (DD = 0), d_r is the delay of the actual receiving bus at the receiving stop. The second part of the decomposition $(t_r^n - t_r^0)$ is defined as *additional time penalty* (ATP), which represents the time cost caused by the transfer desynchronization.

The two parts of TTP's decomposition, ATP and d_r, correspond to the time penalty caused by desynchronization and normal delay of the actual receiving bus. The value of ATP depends on the passenger's actual arrival time at receiving stop and the receiving buses' schedules. If the passenger's actual arrival time is before the scheduled bus's departure time (DD = 0), there will be no additional time penalty; if the actual arrival time is after the *nth* bus's departure time (DD = n), which can be noted as *n-th* receiving bus, then there is an additional time penalty which is worth the sum of *n* receiving buses' headways.

Beyond a single transfer's time penalty, we can expand the measure to a collection of transfers. The collection can have different spatiotemporal definitions depending on different purposes, such as transfers between two routes during an hour every day, or transfers at a stop during a year. We can measure the *average total transfer time penalty* (ATTP) for a collection of transfers C:

$$ATTP_C = E_C(TTP_i) = \frac{1}{|C|} \sum_{i=1}^{|C|} TTP_i$$

(3)

where: TTP_i is the measurement of transfer i's total time penalty.

Transfers: The good, the bad, and the ugly. We classify all transfers into three types according to their real-time synchronization performance. We can distinguish them by the receiving bus's desynchronization degree.

- *The good: normal transfers* (DD = 0). A passenger getting on a normal transfer will catch the same bus as the scheduled transfer. Under this circumstance, ATP = 0, which means there is no additional time penalty, while the performance can be still different from the schedule due to the normal delay of the receiving trip.

8

- *The bad: missed transfers* (DD > 0). The passenger will take a bus after the scheduled bus, hence will suffer from additional time penalty other than normal delay. Under this circumstance, ATP > 0. The missed transfers have several scenarios: 1) generating trip is delayed such that the user cannot catch the scheduled receiving bus; 2) the scheduled receiving bus is out of service; 3) the scheduled receiving bus is severely delayed after another receiving bus. Scenario 1 is the most common circumstance. For scenario 2, if the scheduled receiving trip is no longer running, the passenger must take the next bus. Likewise, for scenario 3, a severely delayed bus can be caught up by another bus on the same route scheduled after it. It is natural for users to take the closest bus despite the buses being out of sequence.
- *The ugly: preemptive transfers* (DD < 0). During a preemptive transfer, instead of the scheduled bus, the user will get on a bus which should have arrived earlier than the passenger at the receiving stop. This is due to delays in the receiving buses. The passenger will naturally take the nearest bus regardless of the schedule. The ATP's value can be negative, zero or positive, however, a negative ATP will not necessarily suggest a better performance since the TTP can be positive meanwhile. In fact, a preemptive transfer's TTP does not guarantee to be worse than a normal transfer; it may achieve better, same, or worse performance depending on the synchronization process. With some license, we refer to these as "ugly" since they are they are unpredictable with respect to impact.

Accordingly, we measure each transfer with a binary value t that represents whether it is a missed transfer. Based on the assessment of single transfer, we define *risk of missing transfer*: it is the proportion of missed transfers in a collection of transfers, based on the empirical schedule and real-time vehicle location data; we can interpret this as an empirical probability of a missed transfer in this collection. Mathematically, in a collection C, missing risk is:

$$RoMT_C = E_C(t) = \frac{1}{|C|} \sum_{i=1}^{|C|} t_i \tag{4}$$

$$t_i = \begin{cases} 1, if\ DD_i > 0 \\ 0, \text{otherwise} \end{cases}$$

where: C is the collection of transfers and t_i is the binary measurement variable of each transfer.

3.3. Determining valid transfers

There are four policies for transfer scheduling in public transit systems: 1) Unscheduled transfers; 2) Scheduled transfers without vehicles waiting; 3) Single holding strategy that lower frequency vehicles wait for higher frequency vehicles; 4) Double holding transfer that both vehicles hold for transfers (Ceder, 2016; Knoppers and Muller, 1995). Many transit authorities, especially those that rely on buses, use an unscheduled transfer policy, meaning there are few explicitly scheduled transfers in the GTFS static data. Moreover, in

reality, transit users' transfer behavior and transit real-time apps will not strictly follow the scheduled transfers.

Consequently, we have to search empirically for possible transfers from the GTFS static data. Theoretically, any two trips at two stops which are proximal enough for users to access can be regarded as a valid transfer. This can be refined with passenger data that shows actual transfers; this is likely to be a subset of the valid transfers. However, the danger with this approach is we may miss a potential transfer if it did not occur in the data.

Based on the data structure in the GTFS data, we define three levels of aggregation: *stop*, *route*, and *trip*. Every trip is run according to a fixed schedule by a bus at a specific time. Trips with the same schedule can be aggregated into a route, and some routes can be bound to a stop. To find transfer schedule from GTFS schedule, we developed a hierarchical searching algorithm in the Python and MongoDB environment. Using the algorithm, we derived all possible stops combinations, route combinations, and GTFS trip combinations. Only those combinations with near distance (Euclidean distance < 100 meters) and unique routes are selected for the transfer schedule: if there are multiple transfers with the same route combination and same generating stop, we only keep the one with least walking distance to remove some redundancy, since passengers are most likely to walk to the closest stop for a transfer.

4. Analysis

We conducted a case study using data from Central Ohio Transit Authority (COTA) bus system in Columbus, Ohio from February 2018 to January 2019. We acquired the General Transit Feed Specification (GTFS) schedule and real-time data via the COTA public application programming interface. COTA shared the Automated Passenger Count (APC) dataset from May 2018 to January 2019. When merging the APC and GTFS datasets, 45.06% of the total records were matched on average, meaning that roughly half of the GTFS data was updated with the more accurate APC data. Based on the GTFS alone, the average risk of missing (RoMT) over the study period is 7.14% (σ =25.75%) and the average total time penalty (ATTP) is 3.74 ($\sigma = 12.97$) minutes; based on the merged APC-GTFS, the average RoMT is 8.55% (σ =27.96%) and the ATTP is 4.57 (σ =15.44) minutes. Although the mean value is relatively small, the standard deviation is substantially large, which suggests the temporal and spatial variation is large. This suggests that GTFS data alone underestimates missing risk and time penalties, although it provides reasonable estimates.

We archived the data using a MongoDB database. The GTFS real-time data, APC data, and their auxiliary databases total nearly one terabyte. Due to large database size, we optimized and parallelized our code to deal with the computational burden. We also developed different summary measures based on varying spatial or temporal aggregations.

4.1.Spatial patterns

To investigate the spatial pattern of missing risk, the first thing is spatial aggregation, since *trip patterns* (each vehicle trip; the finest level of resolution) are too specific and not representative of broader patterns. We can aggregate in different ways. Naturally, *route patterns* are useful, which aggregate the trip combinations based on their route schedules. *Stop patterns* are also useful since the quality of transfers between stops is assessed and stop combinations are geographically distinguishable, making it especially crucial for visualization. We concentrate on stop patterns in our analyses.

Figure 2 show the spatial pattern of the RoMT and ATTP. It shows some differences between RoMT and ATTP's spatial distribution, especially on High Street (the major north-south thoroughfare in Columbus, indicated by a red ellipse in Figure 2), and the northern downtown area (indicated by a green rectangle in Figure 2). Stops along High Street have relatively higher missing risk but also have relatively lower average total time penalty. This is likely due to traffic and other disturbances on this route elevating the risk, although headway between buses is short meaning the time penalty is small. Similarly, the high ATTP clusters on some roads in downtown area and some peripheral roads that do not have higher RoMT. Although the RoMT is low, the time delay can be high, especially for downtown, due to longer headways.
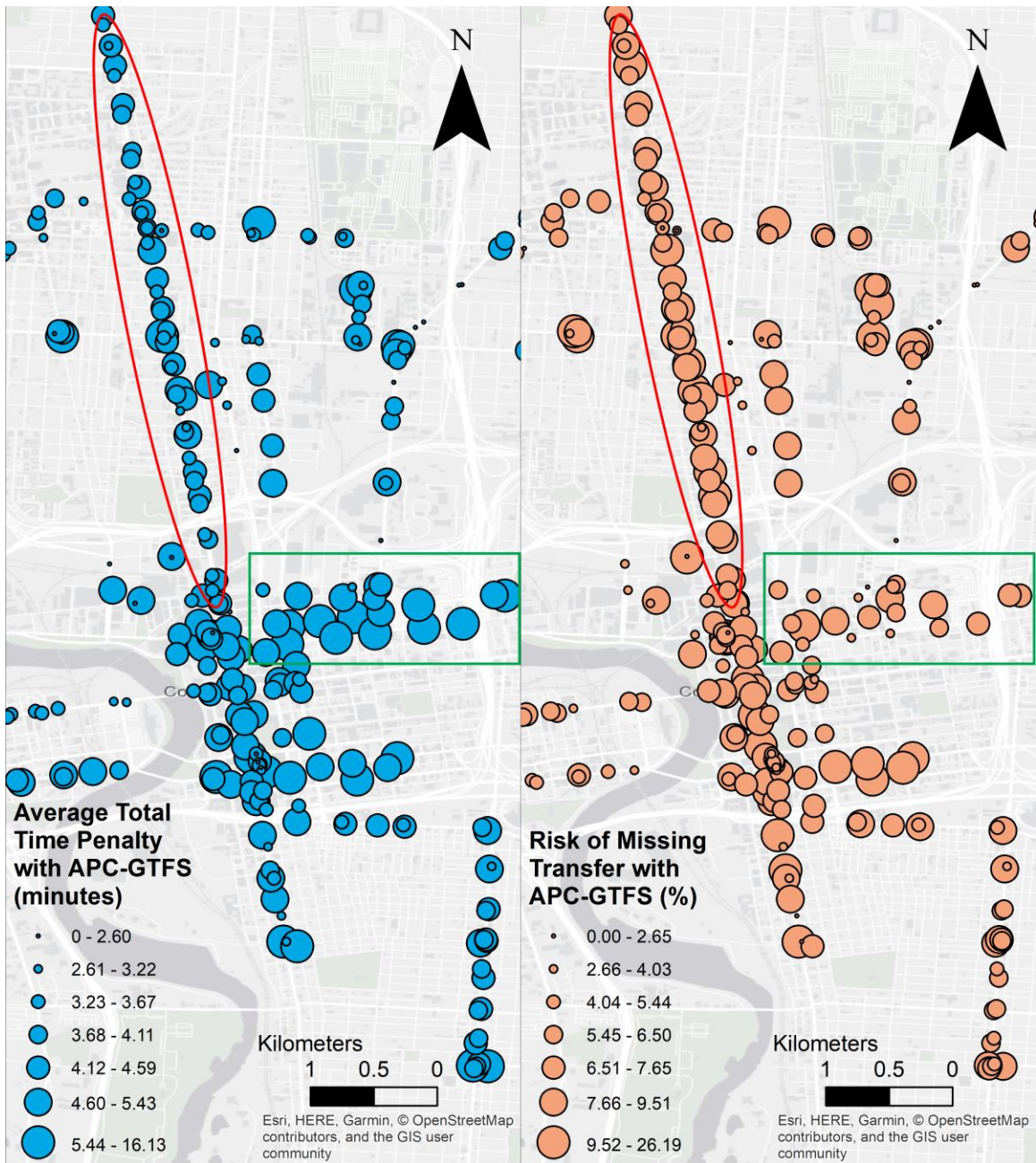
Figure 2: Spatial pattern of ATTP and RoMT (in quantile classification)

4.2. Temporal patterns

We now examine aggregate temporal patterns of risk of missing transfer (RoMT) and average total time penalties (ATTP). Figure 3 provides the monthly trends of RoMT and ATTP for both datasets. July, December, and January show an overall low time penalty pattern. This can be due to better overall traffic conditions during summer and holiday season vacation. August is the worst month to take a transfer; this may be due to the start of an academic year in a city with a massive university campus near the city center.
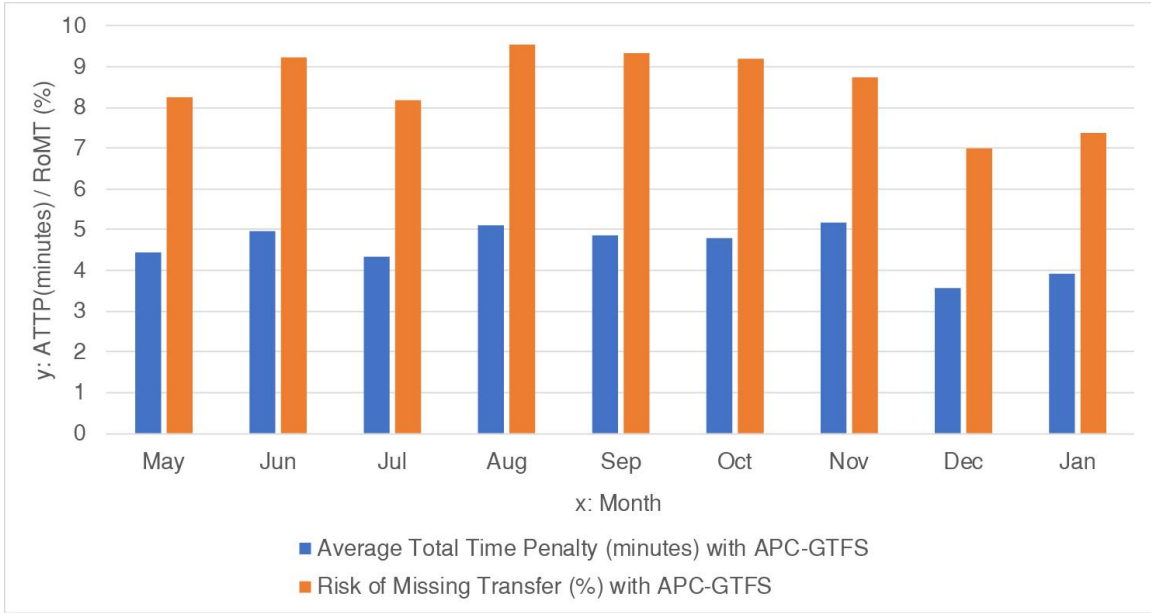
Figure 3: Overall monthly RoMT and ATTP trends.

Figure 4 provides the aggregate trends by day of the week and frequency. We can see the overall RoMT and ATTP peak on Friday; Wednesday, Thursday, and Friday exhibit higher levels of risk and time penalties, likely due to the overall traffic pattern in this city. Both measures are relatively low on weekends, as would be expected due to lower traffic congestion. RoMT and ATTP are relatively low on Mondays, possibly due to flexible working schedule and long weekends for some residents, leading to less commuting. However, we observe Sundays have the lowest ATTP. Intuitively, frequency can be a significant factor accounting for the measures. We conducted a Pearson correlation analyses, ATTP and RoMT and found no significant correlation with daily frequency: p-values are 0.38 for ATTP and 0.118 for RoMT.
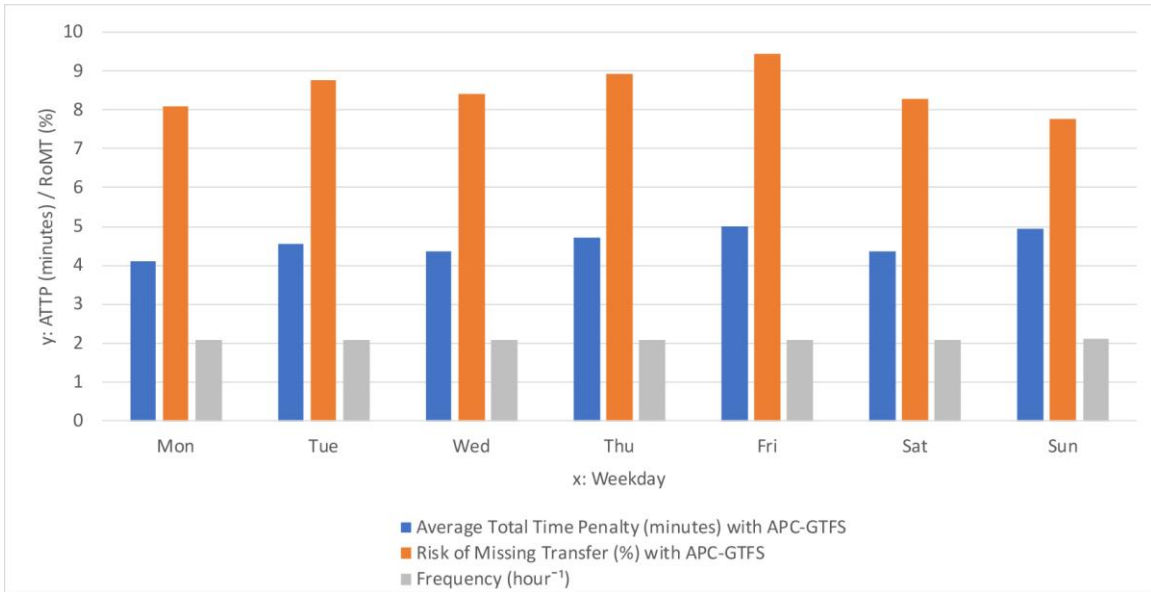
Figure 4: Overall weekday RoMT and ATTP trends and daily frequency.

Figure 5 illustrates the hourly trend and there are three major time periods when missing risk and penalties are high: mornings (8:00–10:00), afternoon (17:00–19:00), and night hours (22:00–24:00). High risk and penalties during the morning and afternoon periods can be explained by overall traffic pattern during these busy hours. However, nighttime with lower traffic also displays high risk and high total time penalty. At night, as the risk increases and service frequency decreases, the time penalties are higher due to sparser scheduled service. In terms of frequency impact, according to the Pearson correlation analyses between each measure and hourly frequency shown in Figure 6, ATTP has significant negative correlation with the frequency, while RoMT has no significant correlation with the frequency.
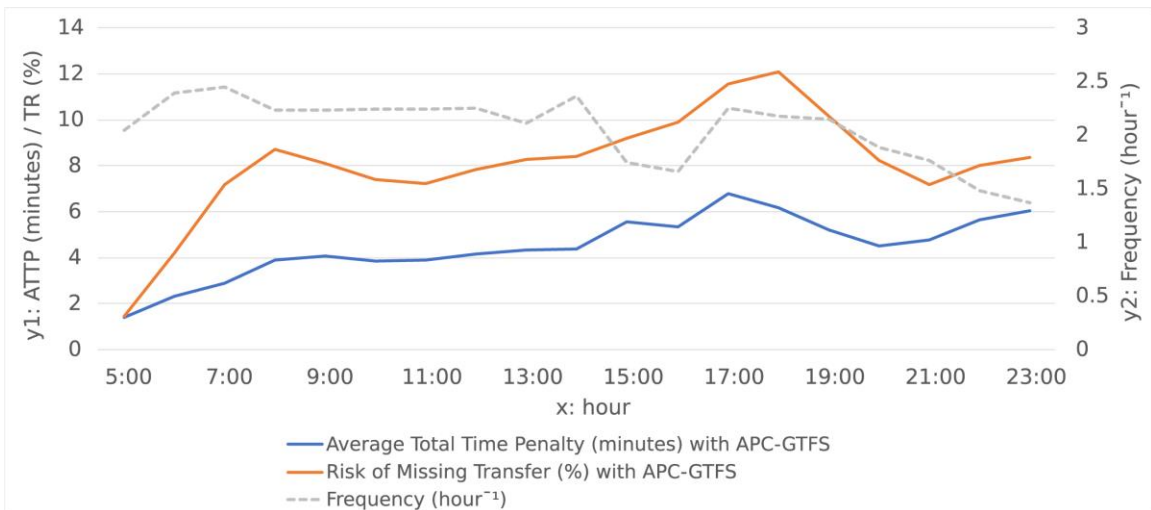

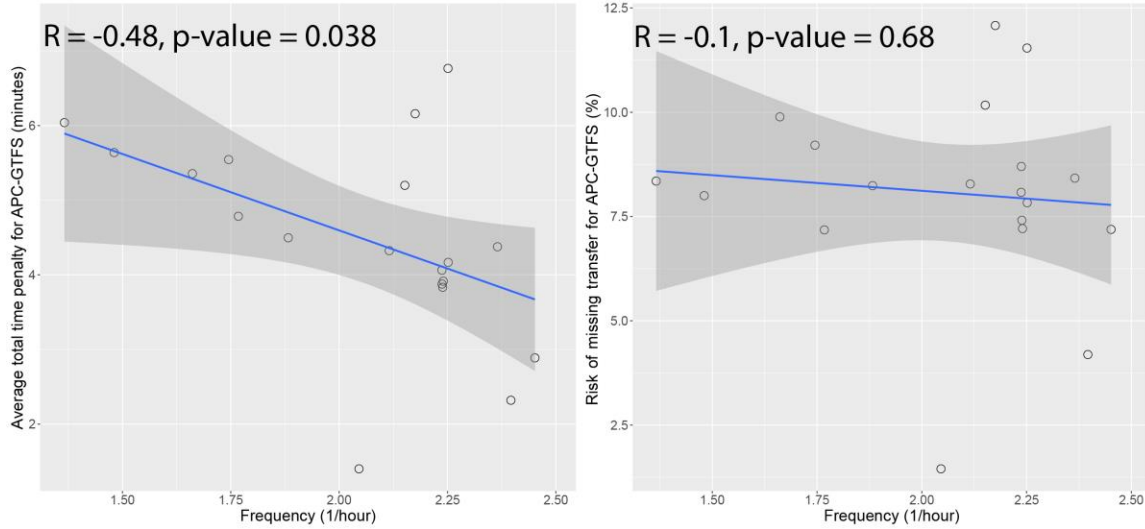Figure 5: Overall hourly RoMT and ATTP trends.

Figure 6: Scatter plots of ATTP (left side) and RoMT (right side) versus frequency.

4.3. Simulating the impacts of dedicated bus lanes

Dedicated bus lanes (DBL) can provide benefits for a bus system by reducing delays due to automobile traffic. Without the disturbance of traffic congestion, bus rapid transit systems with separated DBL can achieve rail-like performance (Li et al., 2009). We simulated the impact of DBL on delays, missing risk and time penalties using the methods in this paper.

We selected the COTA (Central Ohio Transit Authority) bus route No.2 as the target, which has the most transfers and most ridership in the system. We simulate the impact of a DBL by assume all the buses running on this route will behave according to the GTFS static schedule data after DBL is in effect (i.e., no delay). This assumption is hypothetical, and the results represent an upper bound on the actual DBL performance. We analyze RoMT and ATTP's changing trend before and after applying the assumption and the difference's spatial and temporal pattern. Across all stops on the route, the DBL will save 1.72 minutes ($\sigma = 10.09$ minutes) and Kolmogorov–Smirnov (KS) test shows the two scenarios have significantly different distributions (p-value = 0.005). Therefore, although the average time savings are modest, the impacts are statistically significant and highly differentiated across stops.

Figure 7: RoMT and ATTP difference after simulated implementation of a dedicated bus lane.

Also, we calculated different impacts on the generating trips and receiving trips. We categorized all affected transfers into two classes: transfers with generating trip on the DBL (*DBL-generating transfers*) and transfers with receiving trip on the DBL (*DBL-receiving transfers*). DBL will save DBL-generating transfers 2.25 minutes and 5.25% risk while only save DBL-receiving transfers 0.32 minutes and increase 9.03% risk. The KS tests between the two types of transfers show significant differences for both measures (p-value $< 10^{-14}$). This suggests that the DBL will eliminate delays for all transfers thus decrease all transfers' total time penalty universally; but will simultaneously decrease DBL-generating transfers' risk while increasing DBL-receiving transfers' risk, however, it will not necessarily enlarge its time penalty. Based on this simulation, we conclude that improving punctuality via a DBL can reduce ATTP, and DBL-generating transfers will benefit more than DBL-receiving transfers.

5. Conclusion

Big data creates an unprecedented opportunity for more and deeper understanding of the urban public transit systems and the study of transfers. However, due to a historical lack of attainable big data sources, few studies to date have focused on the transfers' on-time performance in a real-time context. Based on high-resolution GTFS and APC real-time and static data of huge volume, we developed risk of missing transfer (RoMT) and average total time penalty (ATTP) measures to assess transfer performance. RoMT and ATTP indicate the systemic quality of transfers and corresponding potential time cost. These measures provide important information for transit system planners and administrators concerning the transfers' feasibility, quality, and user experience. Our spatial and temporal analysis using the COTA system as a case study uncovered both general patterns like overall traffic and transit system delay, as well as some unique patterns, such as high time penalty during the nighttime due to larger headway. Additionally, we simulated dedicated bus routes' impact on the transfer performance. It suggests even a single route DBL can reduce ATTP, especially for DBL-generating transfers.

With the support of big data, the RoMT and ATTP we have developed are a further step towards sustaining a smarter public transit systems. Compared with existing indexes and measurement systems, the spectrum of the proposed measures' audience is broad: besides academic and administrating purposes, ordinary passengers and open source developers are also potential users. Thanks to high-resolution public transit big data, we can calculate corresponding performance based on specific transfers as well as overall broad patterns:

- At the application level, urban dwellers can query each transfer's performance in their real-time transit apps and react correspondingly. Current mainstream transit apps do not show empirical risk and average time loss, especially for transfers over which users have no control. If a proposed transfer's empirical performance is shown when the apps plan the trip, urban dwellers can avoid high risk routes. This is similar to airlines apps showing the on-time performance of air routes. Unlike some composite indexes that are hard to conceptualize, RoMT and ATTP are both intuitive since they use common metrics, namely probabilities and time.
- At the management level, administrators can check the high risk and high time penalty areas and respond. With support of real-time data and the measures, transit authorities can make operational changes such as adding additional buses, enforcing bus's time table to reduce risk, and planning flexible time table adjustment accordingly. City planners can analyze spatiotemporal patterns of risk and time penalties. The patterns of proposed measures can demonstrate important information about the roads, transit system design and other transport and non-transport factors.
- At the policy-making level, policy makers should compare different public transit systems' transfer real-time performance across the US. Due to the high reusability and expandability of the indexes and the system, they can be easily implemented and applied to any transit system with published GTFS scheduled and real-time data without major modification. The common metrics also make intra-system and inter-system comparison much easier.

Future research direction can concentrate on smart and human sensors, generating abundant and high-resolution big data for analysis. Based on more precise and abundant data, there are more possibilities for more scientific planning, improvement, and knowledge derivation of transfer activities and the transit system. Moreover, there are still several limitations for this paper: though we compared datasets of different temporal accuracy, we do not have a good answer for how spatial accuracy will influence the results and how the overall impact of inaccuracy can be decomposed into the two factors. We also do not consider population and ridership factors; with transfer ridership data, we can incorporate these factors into the system.

References

Antrim A and Barbeau SJ (2013) The Many Uses of GTFS Data–Opening the Door to Transit and Multimodal Applications. *Location-Aware Information Systems Laboratory at the University of South Florida* 4.

Bovy PH and Stern E (2012) *Route Choice: Wayfinding in Transport Networks: Wayfinding in Transport Networks*. Springer Science & Business Media.

Ceder A (2016) *Public Transit Planning and Operation: Modeling, Practice and Behavior, Second Edition. Public Transit Planning and Operation: Modeling, Practice and Behavior, Second Edition*. CRC press.

Chen M, Mao S and Liu Y (2014) Big data: A survey. *Mobile networks and applications* 19(2). Springer: 171–209.

Chu X (2010) *A Guidebook for Using Automatic Passenger Counter Data for National Transit Database (NTD) Reporting*. National Center for Transit Research (US).

Dessouky M, Hall R, Nowroozi A, et al. (1999) Bus Dispatching at Timed Transfer Transit Stations Using Bus Tracking Technology. *Transportation Research Part C: Emerging Technologies* 7(4). Elsevier: 187–208.

Firmani D, Mecella M, Scannapieco M, et al. (2016) On the Meaningfulness of "Big Data Quality". *Data Science and Engineering* 1(1). Springer: 6–20.

Google Developers (2016) GTFS Static Overview | Static Transit | Google Developers. Available at: https://developers.google.com/transit/gtfs/ (accessed 8 March 2018).

Google Developers (2018) Trip Updates. Available at: https://developers.google.com/transit/gtfs-realtime/guides/trip-updates (accessed 8 April 2019).

Guo Z and Wilson NHM (2004) Assessment of the Transfer Penalty for Transit Trips: Geographic Information System-based Disaggregate Modeling Approach. *Transportation Research Record* 1872(1872). Transportation Research Board of the National Academies: 10–18. DOI: 10.3141/1872-02.

Guo Z and Wilson NHM (2011) Assessing the Cost of Transfer Inconvenience in Public Transport Systems: a Case Study of the London Underground. *Transportation*

Research Part A: Policy and Practice 45(2). Pergamon: 91–104.

Hadas Y and Ranjitkar P (2012) Modeling Public-transit Connectivity With Spatial Quality-of-transfer Measurements. *Journal of Transport Geography* 22. Elsevier: 137–147. DOI: 10.1016/j.jtrangeo.2011.12.003.

Hilbert M (2016) Big Data for Development: A Review of Promises and Challenges. *Development Policy Review* 34(1). Wiley Online Library: 135–174.

Jang W (2010) Travel Time and Transfer Analysis Using Transit Smart Card Data. *Transportation Research Record* (2144). Transportation Research Board of the National Academies: 142–149.

Knoppers P and Muller T (1995) Optimized Transfer Opportunities in Public Transport. *Transportation Science* 29(1). INFORMS: 101–105.

Kujala R, Weckström C, Mladenović MN, et al. (2018) Travel Times and Transfers in Public Transport: Comprehensive Accessibility Analysis Based on Pareto-optimal Journeys. *Computers, Environment and Urban Systems* 67. Elsevier: 41–54.

Li JQ, Song M, Li M, et al. (2009) Planning for Bus Rapid Transit in Single Dedicated Bus Lane. *Transportation Research Record* (2111). Transportation Research Board of the National Academies: 76–82.

Liu R, Pendyala RM and Polzin S (1997) Improved Reliability of Public Transportation Using Real-time Transfer Synchronization. *Transportation Research Record* (1607). Transportation Research Board of the National Academies: 74–80.

Miller HJ and Goodchild MF (2015) Data-driven Geography. *GeoJournal* 80(4). Springer: 449–461.

Nesheli MM and Ceder A (2015) Improved reliability of public transportation using real-time transfer synchronization. *Transportation Research Part C: Emerging Technologies* 60. Pergamon: 525–539.

Nishiuchi H, Todoroki T and Kishi Y (2015) A Fundamental Study on Evaluation of Public Transport Transfer Nodes by Data Envelop Analysis Approach Using Smart Card Data. *Transportation Research Procedia* 6. Elsevier: 391–401.

Sun L, Rong J, Ren F, et al. (2007) Evaluation of Passenger Transfer Efficiency of an Urban Public Transportation Terminal. In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2007, pp. 436–441. IEEE.

Sun L, Rong J and Yao L (2010) Measuring Transfer Efficiency of Urban Public Transportation Terminals by Data Envelopment Analysis. *Journal of Urban Planning and Development* 136(4). American Society of Civil Engineers: 314–319.

Transit Wiki (2019) Automated Passenger Counter. Available at: https://www.transitwiki.org/TransitWiki/index.php/Automated_passenger_counter (accessed 10 February 2019).

Walker J (2012) *Human Transit: How Clearer Thinking about Public Transit Can Enrich*

Our Communities and Our Lives. Human Transit: How Clearer Thinking About Public Transit can Enrich our Communities and our Lives. Island Press.