Measuring Real-time Transit Apps’ Impact on Public Transit Waiting Time Using High-resolution Real-time data

Luyu Liu

The Ohio State University

**Abstract:**

Real-time information (RTI) and Real-time transit apps (RTA) have been proven to have significant impact on passengers’ waiting time and user experience. Although precious research thoroughly surveyed and simulated the overall impact, few studies investigate the impact’s geographic and temporal pattern. In this paper, we utilize General Transit Feed Specification (GTFS) real-time data for the buses’ high-resolution real-time status. We first introduce the concept of *volunteered optimization* and several trip planning relaxation strategies to simulate RTA and non-RTA users. Then, we optimize the best *prudent relaxation* strategy based on RTI and we calculate the difference between different relaxation strategies to measure RTA users’ waiting time reduction. Moreover, we visualize and analyze the waiting time difference and find the universal presence of marginalized stops near the originating stops. The results prove that RTI and RTA can significantly decrease waiting time for users and it demonstrates great variation geographically and temporally. The methodology also shows great potential in future research and industry application.

**Keywords:** Transit real-time data; GTFS; waiting time; volunteered optimization.

1. Introduction

Mobile technologies are changing people’s life in different ways, also for transportation especially public transit. Real-time information (RTI) and real-time transit apps (RTA) are reshaping our way to take public transit (PT) system. With all the benefits of PT system, many people still felt reluctant to take public transit. Topmost reasons why people do not take PT are: long travel time especially long waiting time, lack of comfort, lack of control/certainty, and unreliability (Beirão & Cabral, 2007). Correspondingly, RTA can improve PT user experience in three factors: strengthening control over the timetable, optimizing waiting time, and adapting to unreliability.

RTA provides users ability to comprehend the sophisticated timetable in a PT system, for both scheduled and real-time timetable. Paper timetable are inflexible and limited to schedule, and traditional phone call and text service are also limited to scheduled static timetable. RTA provide both scheduled and real-time support for PT users with smart phone through user-friendly interface.

RTA decreases the waiting time. Waiting time is a critical factor affecting people’s preference of transportation (Beirão & Cabral, 2007). RTA can plan best time for users to leave for the public transit based on the walking time and PT timetable. Ideally, RTA can diminish the waiting time to 0, which means as soon as users arrive at the stop, the bus arrives. Many studies investigated the waiting time reduction by real-time information: 91% percent of RTI users self-reported spending less waiting time in Seattle, 2010 (Ferris, Watkins, & Borning, 2010), and RTI users can save 2.4 minutes in Seattle, 2011 (Watkins, Ferris, Borning, Rutherford, & Layton, 2011) and 1.79 minutes in Tampa, 2014 (Brakewood, Barbeau, & Watkins, 2014) according to self-reported survey.

RTA adapts to PT system’s unreliability when optimizing waiting time (Brakewood et al., 2014). For all PT system, delay is inevitable. When delay happens, real-time transit apps (RTA) users can use their smart phone to get access to the real-time status of buses and plan their trips accordingly. With the help of RTA, users can have an anticipation about the on-time transit’s delay, thus postponing the departure time to reduce the waiting time at the stop. Many RTAs provide real-time trip planner to help users to schedule a best time to take public transit, such as Google Map and Transit.

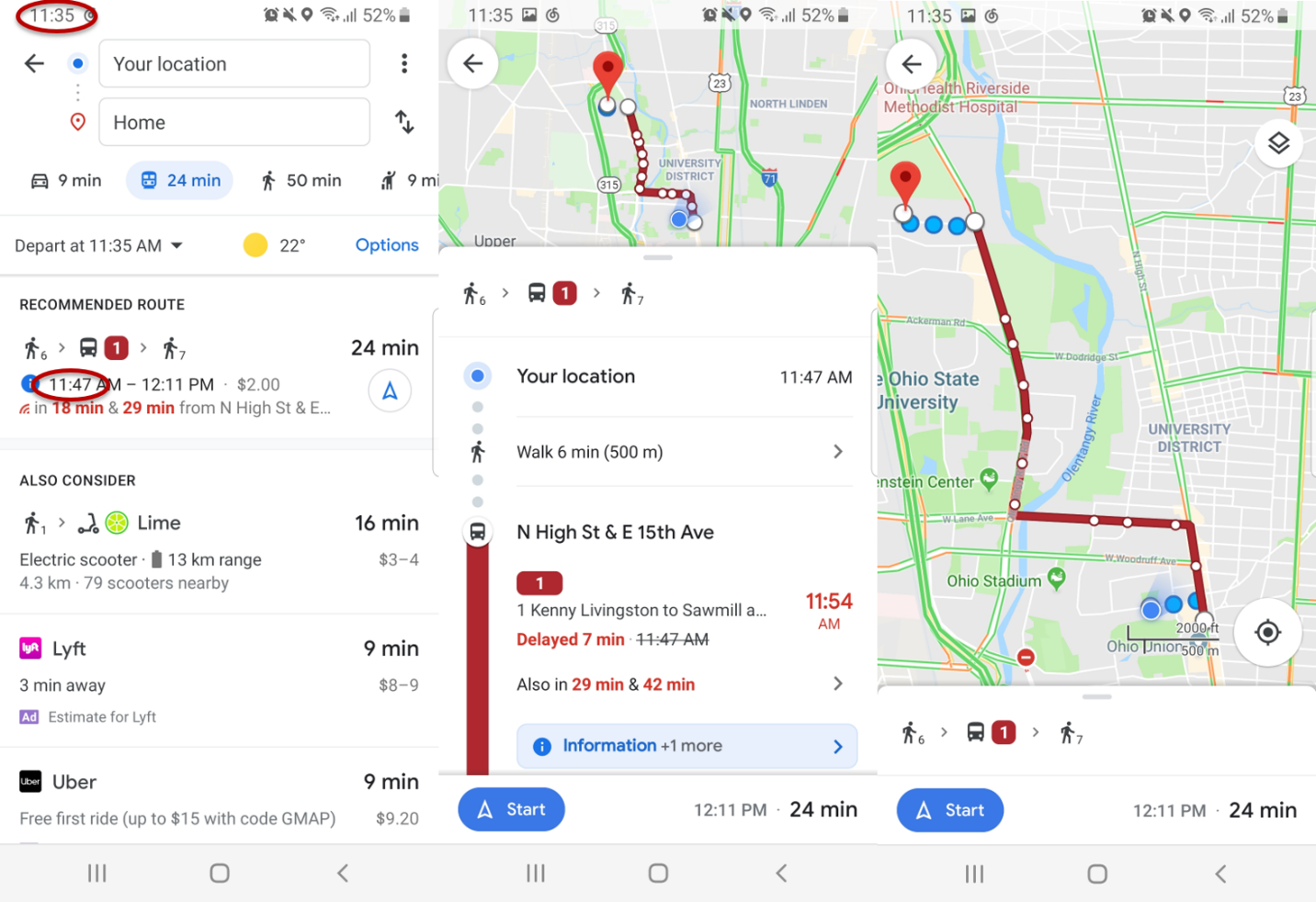


Figure 1 A typical interface of a real-time transit app.

However, besides the benefit RTA produces, the smart trip strategies can lead to several undesirable results. The optimization of waiting time also comes with high risk of missing a bus. Figure 1 shows the changing trend of waiting time along with user’s arrival time at the stop, supposing the buses are all running absolutely punctually. The optimal point, where waiting time is 0, is exactly the point with highest risk of missing the scheduled bus, where the waiting time is the headway. Meanwhile, with the walking time increases, the instability of the PT trip’s real-time performance also increases. During the time when users are walking from home to the stop, a delayed bus can speed up to catch up the former delay, which can result in missing the bus for RTA users. Therefore, we would like to calculate and compare the missing rate for both non-RTA and RTA users.

Figure 2 Hypothetical waiting time's changing pattern along with user’s arrival time.

Nevertheless, even with more and more technology breakthroughs in this area, there are still economic and intellectual barriers for certain population to get access to the real-time public transit information. People who do not use smart phone application or do not own a smart phone cannot know the real-time status. These ordinary users have to plan their trips according to the schedule, their daily experience, or even random. Based on this fact, we would like to assess the average waiting time difference between non-RTA and RTA users and social justice issue behind the difference.

In conclusion, the project is to assess and optimize the waiting time reduction and consequent missing risk between non-RTA and RTA users by consulting real-time bus feed. Based on the results, we can prove that whether it is statistically useful for users to follow different possible trip planning strategies derived by RTA’s trip planner, and how reliable different RTA strategies are. We will adopt several reasonable assumptions to build mathematical models and implement the models with General Transit Feed Specification (GTFS, real-time transit data standard) data provided by Central Ohio Transit Authority (COTA) in Columbus, Ohio.

1. Literature review

The idea of measuring real-time information’s impact on public transit system and waiting time was first introduced even before the 21st century. In 1995, Reed (1995) investigated signage and telephone’s real-time information’s impact on passengers’ waiting time (Reed, 1995). After the widespread application of smart personal devices, numerous studies examined new technologies’ influence on users’ behaviors. Recently, real-time information is becoming more accessible due to less expensive automated vehicle location system and the open data policy. Correspondingly, the body of literature is steadily growing and there are numerous studies investigating real-time information’s impact on public transit users (Brakewood et al., 2014; Brakewood, Macfarlane, & Watkins, 2015; Cats & Gkioulou, 2017; Ferris et al., 2010; Fries, Dunning, & Chowdhury, 2011; Liu, Shi, & Jian, 2017; Papangelis, Nelson, Sripada, & Beecroft, 2016; Watkins et al., 2011).

Brakewood and Watkins (2018) systematically reviewed and categorized most research studying real-time transit information impact on the public transit system and users (Brakewood & Watkins, 2018), including waiting time, path choice, and ridership. The paper classifies all studies according to five dimensions: information media, mode of transit, location, methodology, and findings. Similarly, this literature review will inspect studies that investigated RTA’s impact on the waiting time, based on their real-time information media and method.

* 1. Real-time information media

We can categorize the research according to their information media, including static signage, telephone and text services, and smart phone application. Signage and at-stop displays can provide PT users useful information and reduce actual and perceived wait time (Dziekan & Vermeulen, 2006; Reed, 1995). Moreover, at-stop displays’ psychological effect is even more important: systems showing the next train or bus’s departure time can greatly release anxiety (Dziekan & Vermeulen, 2006). However, due to its static nature, user can only get the information at the stop, which limits its actual effectiveness. Telephone service and text services are more useful than signage (Reed, 1995), since for the first time, people can get access to the real-time data out of the station.

However, it is after the introduction and wide application of the real-time transit information that people’s transit experience is actually changed by smart phone and real-time transit apps.

* 1. Method

Brakewood and Watkins (2018) categorized all studies based on the methods used: Survey-based methods, simulation models, and aggregate-level econometric analysis (Brakewood & Watkins, 2018). Survey-based methods include on-board surveys (Fan, Guthrie, & Levinson, 2016), before-after surveys (Chow, Block-Schachter, & Hickey, 2014), web-based surveys (Ferris et al., 2010), in-person surveys (Watkins et al., 2011), interviews and observations (Papangelis et al., 2016), and stated preference surveys (Liu et al., 2017).

[Survey-based method]

Survey-based method is definitely the majority among all RTI impact studies. Surveys sample respondents with the same questions; they measure many variables, test hypotheses, and conclude temporal sequence from questions about past behavior, experiences, or characteristics (Neuman & Robson, 2014). Without the support of automatic real-time data, passenger surveys are the most direct methods to assess PT system use, especially for user experience and perceptions. Meanwhile, survey data can quantitatively assess different attributes using self-reported data (Neuman & Robson, 2014), which partially guarantees generalizability and authenticity. Besides, in contrast to the automatic generated data, the surveys’ data also point to users, instead of vehicles. The human-centered nature of survey data also guarantees its direct and close connection with human per se.

However, survey methods have their problems: Some survey-based methods rely on user’s self-reported information, which lacks precision and authenticity especially for non-cognitive value. Surveys measure the user’s perceptual estimation of the assessed value. Compared with value obtained by physical sensors, the self-reported information may be obfuscated and biased by threatening questions (Bradburn et al., 1979) and its self-evident nature (Goyder, 1986).

Survey’s sampling strategy and measurement is often questionable. First, the survey’s size, especially for some traditional data collection methods, is critically small due to high cost of data collection (Goyder, 1986). Second, some methods, especially IT (Information technology)-based methods, are often biased and dubious. Mail survey (Rossi, Wright, & Anderson, 2013), text/phone call survey, and internet-based survey (Wright, 2005) can significantly reduce the time and economic cost of the survey. However, these methods face a same problem: it is hard to access a representative sample (Rossi et al., 2013; Wright, 2005). For public transport system, not all users can get access to these services while the survey will only sample these specific people.

[Simulation]

Mathematical simulation is often used to investigate and solve problems that are too difficult or costly to measure directly. For example, Cats and Gkioulou (2014) adopted an agent-based model to simulate the influence of PT reliability and real-time information on waiting time uncertainty (Cats & Gkioulou, 2017). Agent-based model simulation usually adopts several assumptions and represents the simultaneous actions and interactions of various agents. The simulation tries to imitate and predict the performance of a complex system such as PT system (Gkioulou, 2013). However, the effectiveness of the model is debatable, and the adopted assumptions could also be inconsistent with the reality.

Although the overall impact of RTI on waiting time is well explored, few studies investigated the variance (Brakewood & Watkins, 2018), including spatial and temporal patterns. Meanwhile, ironically, there are no studies using actual real-time transit data source due to the lack of these data and corresponding theory supports. We would like to address these gaps of waiting time problem in the real-time transit data context and implement it using actual real-time transit data.

1. Methodology

In this section, we first introduce the data source and corresponding manipulation processes. Moreover, we theorize the synchronization process during the procedure of transit and the concept of *delay reclamation*. Based on the synchronization theory, we propose and model several trip planning strategies during the process of decision making and optimize the RTA user’s strategy based on deterministic real-time data. We also calculate the waiting time difference between RTA users’ deterministic process and non-RTA users’ probabilistic process.

* 1. Data source

Introduced by Google first in 2006, GTFS is a collection of two data types: GTFS static and GTFS real-time expansion. GTFS static indicates the schedule data of a PT system in several separate tables (Google Developers, 2016). GTFS static is the current *de facto* standard for PT system schedules and PT geographic information (Google Developers, 2016). As a standard for open data, it is easy to share and access for the public, open-source programmer, and researchers. For this paper, we collected and organized all history schedule data in MongoDB and Python environment from Application Programming Interface (API) provided by Central Ohio Transit Authority (COTA) bus system in Columbus, Ohio.

Besides schedule data, GTFS real-time data provides vehicles’ geographic data with high temporal resolution. GTFS real-time includes two parts: buses’ location real-time data and the trip updates, which specifies the fluctuations in the real-time timetable (Google Developers, 2018). Unlike many traditional transit data, GTFS real-time data have large volume, high resolution, normalized standard, and full system coverage. For this study, we collected GTFS real-time data from API from COTA bus system with the frequency of 1 minute for more than 1 year.

* 1. Synchronization

Taking a bus could be conceptualized as a synchronization process between the walking trip to the target stop and the target bus’s *trip sequence array*. Trip sequence array is defined as the collection of trips running on the same route as the target bus in the direction of target bus.

Depending on user’s arrival time at the stop *t*, the actual bus that user will take can be different from the scheduled one. We use the same concept in the transfer synchronization process: *desynchronization degree* (DD), to measure the desynchronization between the bus and user at the stop. DD is an integer indicator that represent how many buses the user loses in the trip sequence array.

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

If the user arrives at the stop between bus DD = n – 1 and bus DD = n, then the user will take bus DD = n.

The process of walking is linear: the users can strictly control the walking time by selecting their home departure time (HDT). For a user, the relationship between HDT and arrival time is linear.

Nevertheless, the actual real-time performance of the bus is non-linear: first, the users cannot directly control the boarding time by selecting their HDT, and the relationship between HDT and user’s boarding time is non-linear; second, the bus will not run at a fixed velocity and the expected time of arrival of bus at the stop is constantly changing. If the bus is delayed when the user departs home, during the walking time, the bus may catch up a part of delay by accelerating.

We therefore define the concept of *reclaimed delay (RD).* It is the time difference between the actual time of arrival (ATA) and the expected time of arrival (ETA) at the stop.

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

Where: S is the collection of stops on the route, i is the target stop, t is the current time.

The reclaimed delay could be tiny but critical: consequently, the user will lose the bus and suffer waiting time penalty for a relatively long time. Thus, the synchronization of these two processes is highly unstable. Figure 3 shows a space-time diagram of the synchronization process.

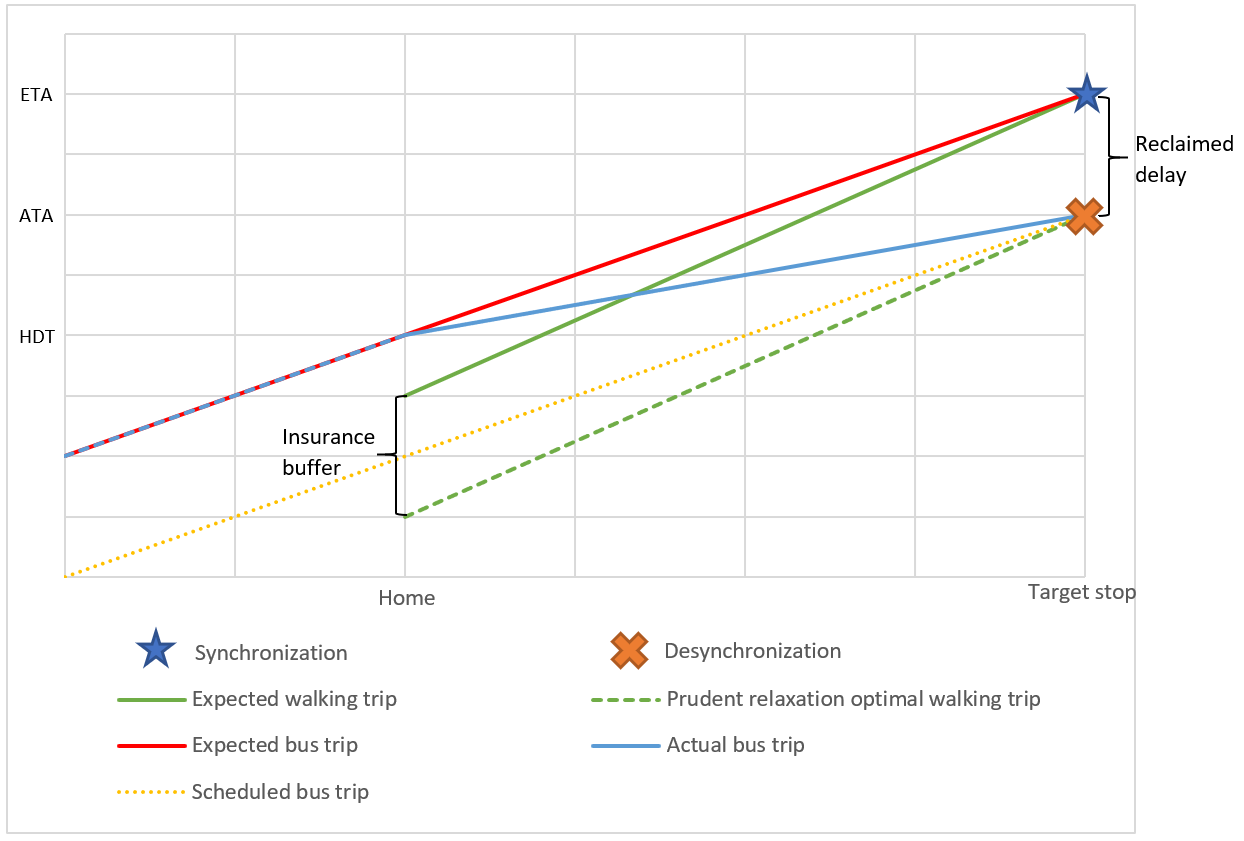


Figure 3 Space-time diagram of the synchronization and desynchronization

* 1. Trip planning strategies

A trip planning strategy (TPS) can be interpreted as a tactic for a user to plan and execute the transit trip. Assuming no disturbance during user’s walking and boarding, different TPS has only one controllable factor to determine the actual waiting time: the time to leave home for the transit (home departure time, HDT). RTA relaxes the fixed timetable in a frequently delayed PT system, thus saving waiting time for RTA users. Depending on how to determine the home departure time, there are different trip planning strategies for both RTA and non-RTA users and their different purposes. Since we have access to the buses’ real-time information, if the home departure time is given explicitly, the calculation of TPSs’ real-time performance is deterministic.

* + 1. Non-RTA users’ trip planning strategies

In traditional incidence behavior studies, users are divided into two groups: timetable-dependent passengers, who are aware of the system schedule or empirical performance, and timetable-independent passengers, who are not aware of the system schedule and performance (Frumin & Zhao, 2012). These two user groups were introduced in the context of non-RTA experience.

[**Arbitrary relaxation (AR)**]

Before the time of smart phone, text, and public real-time information, under many circumstances, PT users were not particularly planning their trips. Normally, they would walk to the stop and catch the subsequent bus arbitrarily. The major assumption is that the user’s arrival time is independent from the vehicle system (Frumin & Zhao, 2012).

Because the user’s decision-making process is random, it is reasonable to assume user’s HDT or user’s arrival time is evenly distributed among the headway between two buses. Traditionally, the average waiting time is the expectation of the random variable headway (Frumin & Zhao, 2012).

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

Where: is the average waiting time, is the average headway, is the standard variance of headway.

However, since we have access to the deterministic real-time vehicle departure time, we can calculate the average waiting time as the mean of the departure time of target bus and its subsequent bus. Thus, the users’ home departure time is:

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

Where: is the median of the target bus and its subsequent bus. is the bus’s actual real-time departure, is a small constant time, is the very first prior bus’s actual real-time departure.

Theoretically, this strategy is not very efficient. We will also calculate the waiting time difference between AR and NR. This is another good benchmark for non-RTA users.

[**Null relaxation (NR)**]

Without knowing any information about the running status, users can still follow the schedule published to the public in advance. Or if a user has urgent affairs, such as work with strict timetable and medical emergency, she/he prefers earlier final arrival time than convenience. Under this circumstance, the user will follow the scheduled timetable of the PT system regardless of waiting time.

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

Where: is the user’s home departure time (HDT), is the walking time from user’s home to the stop, is the scheduled bus departure time.

Consequently, the RTA user will not benefit from waiting time reduction. However, since no bus/train will leave earlier than the scheduled time, NR minimizes the missing risk. NR is another benchmark for waiting time reduction strategies, which has the lowest missing risk.

[**Empirical relaxation (ER)**]

If a user can get access to the historical information, either from a database or daily experience, the user can relax the HDT based on the empirical average time/maximum time without any real-time assistant.

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

Where: is the bus trip’s the average actual real-time departure time, is the bus trip’s actual departure time on day *i*, and n is the total number of days.

This is a common non-RTA strategy. We will also calculate the waiting time difference between ER and NR to confirm its efficiency.

* + 1. RTA users’ trip planning strategies

For the simulation of RTA trip planning process, most RTAs will directly use the ETAs provided by GTFS trip update for the buses’ real-time information (Google Developers, 2018; Transit app, 2019). The process of RTAs’ decision making consists of several steps: First, RTAs will provide different path choices for users and their HDT or users will find the desired bus trips/routes from the list in the RTA. Then, the trip update data will provide ETAs at the target stops for RTAs or users. Finally, RTAs or users will subtract estimate walking time and obtain estimated HDT. The estimated HDT is not constant, instead, RTAs will update it according to the real-time trip update data. Depending on the relationship between given estimated HDTs and the current time, the user will decide the actual HDT. Just like non-RTA TPSs, the essential part is the criteria to derive the actual HDT.

[**Omnipotent relaxation (OR)**]

Supposing a hypothetical omniscient public transit user who are always aware of the real-time and future status, the user can adjust their home departure time (HDT) accordingly. The user will always catch the desired bus in time without any waiting, regardless of the real-time performance of the PT system and his/her home’s distance from the stop. Practically, nobody can achieve real global optima deterministically. OR is a theoretical benchmark for all strategies, which represents the best strategy for all possible waiting time reduction strategies.

[**Greedy relaxation (GR)**]

Most transit users do not want to wait. Moreover, if a user encounters extreme weather events, especially cold weather and heavy precipitation, she/he may want to reduce the waiting time as much as possible. In practice, a greedy relaxation (GR) user will only leave if RTA tells him so according to the real-time data, when the bus’s ETA at the stop is equal to walking time plus current time:

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

Where: is the scheduled bus’s ETA at the stop given by RTA and real-time data, and is the current time when .

This strategy can achieve temporary optima. However, due to the instability of PT system, the missing risk of GR is also the highest. Due to the possible reclaimed delay, the bus will likely arrive earlier than ETA. Even a smallest disturbance during the walking process or the bus running status during user’s walking can result in a missing bus using this strategy. Consequently, the user will suffer from a long waiting time penalty, which is almost equal to a headway, the largest possible waiting time.

[**Prudent relaxation (PR)**]

If a user would like to save waiting time and keep some degree of missing risk, she/he may leave several minutes earlier than GR. Similar strategies were adopted in the context of scheduled time: some passengers will leave 2-3 minutes before scheduled service arrival in case of some unexpected “risky” events, such as delay by the elevators (Fonzone, Schmöcker, & Liu, 2015). These risk aversion reactions are usually defined as costs of unreliability (Frumin & Zhao, 2012).

Similarly, a prudent RTA passenger will leave a short buffer for the risk of missing bus and unexpected delay reclamation. This short time buffer, which is defined as insurance buffer (IB), trades some time to reduce missing risk, thus decreasing instability of the buses’ real-time performance.

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

Where: IB is the short time buffer.

Figure 3 also suggests the relationship between IB and reclaimed delay. RTA TPS is risky due to the existence of reclaimed delay, thus, the ultimate purpose of IB is to compensate for the reclaimed delay. Therefore, in the optimal PR scenario, an optimal insurance buffer should be exactly equal to the reclaimed delay to achieve 0 waiting time.

PR and GR can be categorized as a PR family, for GR is a special case of PR with IB = 0. With different IBs, each PR strategy can vary in actual waiting time. We would like to optimize IB and find the best PR strategy with max average waiting time reduction.

* + 1. Proportions of user groups

Since we assume the walking time is constant, HDT can also be linearly transformed into users’ arrival time. Numerous studies investigated the pattern of users’ actual arrival time or incidence behavior. A common observation obtained by these studies is that users’ actual waiting time is less than the pure random average waiting time (AWT) (Frumin & Zhao, 2012; Luethi, Weidmann, & Nash, 2007). Jolliffe and Hutchinson (1975) and Bowman and Turnquist (1981) contributed to the classification of three classes of passengers, which can be translated into respectively relaxation strategies: omnipotent relaxation with proportion of , null relaxation with proportion of , arbitrary relaxation with proportion of (Bowman & Turnquist, 1981; Jolliffe & Hutchinson, 1975). The value of determines the overall pattern of average waiting time.

Based on this classification, we also theorize each user group’s proportion of different TSP. Unlike traditional classfication, we know that omnipotent relaxation cannot be achieved by any user. Specifically, we will concentrate on RTA users: all RTA users’ proportion in all the transit users is , while each TPS of PR family has one unique best empirical insurance buffer (IB) with proportion of in all RTA users. For non-RTA users, empirical relaxation users’ proportion in non-RTA users is and arbitrary relaxation users’ proportion in non-RTA users is . Table 1 shows different TPSs’ availability for each users and theoretical ratio.

|  |  |  |  |
| --- | --- | --- | --- |
| Strategy | Non-RTA users | RTA users | Theoretical ratio |
| Omnipotent relaxation | X | X | 0 |
| Null relaxation | ✓ | ✓ |  |
| Greedy relaxation | X | ✓ |  |
| Prudent relaxation | X | ✓ |  |
| Empirical relaxation | ✓ | ✓ |  |
| Arbitrary relaxation | ✓ | ✓ |  |

Table 1 Different strategies' availability for non-RTA and RTA users and theoretical ratio

However, we do not have access to the actual proportions of each user group. Consequently, we will only calculate the PR optimal TPS and the waiting time difference between various RTA and non-RTA TPSs.

* 1. Measures and optimization

We would like to measure the difference of waiting time and risk of missing bus/train between the RTA users and non-RTA users. Therefore, we present the definitions of two indexes: missing risk (MR) and average waiting time (AWT).

[**Missing risk**]

|  |  |  |
| --- | --- | --- |
|  |  | (9) |

Where: n is total number of trips;

[**Waiting time**]

|  |  |  |
| --- | --- | --- |
|  |  | (10) |

Where: is the actual waiting time, is the bus’s actual real-time departure time, and is user’s arrival time at the stop. is user’s HDT, and is user’s walking time from home to the stop.

So, the bus’s actual real-time departure time is:

|  |  |  |
| --- | --- | --- |
|  |  | (11) |

Thus:

|  |  |  |
| --- | --- | --- |
|  |  | (12) |

Equation 12 proves that the only factor that user can control and can affect waiting time is .

Besides single trip’s waiting time, we can also calculate the mean of waiting time’s distribution of all trips:

|  |  |  |
| --- | --- | --- |
|  |  | (13) |

Where: n is total number of trips.

[**Volunteered optimization and PR optimal**]

Previous research concentrated on PT system’s optimization in the stage of planning and operation. Due to the lack of authority and information access, ordinary users were rarely considered as a part of optimization process. However, with RTI, although users still cannot directly improve the real-time systematic performance (delay, ridership) of the system per se, optimization can be conducted in the individual level to reduce waiting time.

Correspondingly, we introduce the concept of volunteered optimization: despite PT systems’ instability and uncertainty, users with RTA can adapt and optimize each trip according to the delay and real-time information. Volunteered optimization is independent from the scheduled timetable; instead, it is based on the real-time status and decentralized. With the maximization of waiting time reduction in the individual level, the overall waiting time penalty will also be diminished. To achieve volunteered optimization, it is necessary to optimize each individual trip, which is from PR strategy family. Correspondingly, we need to find an optimal PR strategy: *PR optimal*.

For PR or GR strategy, the users will plan their HDT according to the bus’s ETA instead of the real arrival time. Here we define expected waiting time:

|  |  |  |
| --- | --- | --- |
|  |  | (14) |

Where: is the ETA, t is the user’s arrival time at the stop, is PR strategy’s HDT, and is the walking time.

Combining (5) and (13), we have:

|  |  |  |
| --- | --- | --- |
|  |  | (15) |

For PR family, the expected waiting time is equal to the insurance buffer.

Figure 3 shows the theoretical relationship between users’ HDT and expected waiting time, missing risk, and waiting time.

Figure 4 Theoretical relationship between user HDT and average expected waiting time, average actual waiting time, and average risk.

To find the , empirically, we simulated the users’ real-time waiting time using different IB. We constitute the optimization problem in the following formula.

|  |  |  |
| --- | --- | --- |
|  |  | (16) |

Subject to:

|  |  |  |
| --- | --- | --- |
|  |  | (17) |
|  |  | (18) |
|  |  | (19) |
|  |  |  |

Where: is the actual waiting time for the user live walking time from the stop who intends to catch trip . is the collection of all trips, and is the stops on trip , is the designated walking time range. is the actual departure time of target bus, and is the actual arrival time at the stop for the user. is the expected waiting time as well as the insurance buffer. is the expected departure time of the scheduled bus.

Figure 5 is the flow chart of PR optimization: We will calculate the performance for all buffers for optimization, then we will find the smallest waiting time and the corresponding buffer. If there are multiple options, designate the one with smaller buffer. After getting optimal buffers for every day and every trip and stop, we reduce all buffers into one day’s buffers by finding the maximum of the optimal buffers. In this way, we guarantee the optimality of obtained buffers: first, obtained buffers in each day have the least waiting time; second, obtained buffers are the smallest one among the buffers with the least waiting time; third, reduced buffers are the maximum buffer, which guarantees the synchronization for each day when recalculating the performance.

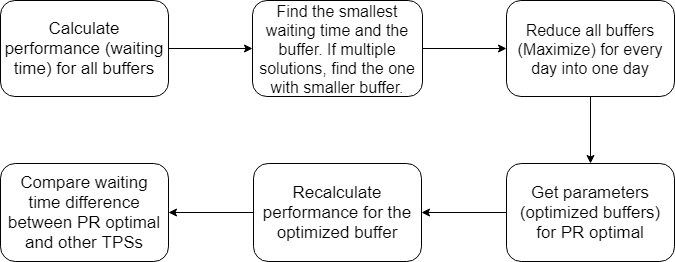


Figure 5 Flow chart of PR optimization algorithm

For the first step of the optimization process, the complexity is , where b is the number of possible buffers, d is the number of dates, t is the number of GTFS trips, s is the number of stops in a trip, w is the number of walking time, n is the number of the GTFS feeds for the stop. Thus, the total complexity is polynomial and of high power. The optimization process of produces massive number of parameters . We minimize waiting time over , which will have a different IB for each day, each trip, each stop, and each walking time. In practical, to reduce computation load, we calculated the AWT for every stop on COTA bus route No. 2 for different IB, from 0 to 300 seconds with interval of 10 seconds. We also parallelized the outmost loop to improve computation performance.

[**Waiting time difference**]

For non-RTA users who has no access to the real-time data, the most rational and practical strategy is NR or AR. And for RTA users, the most beneficial strategy is PR with an optimal insurance buffer. We introduce the difference between NR's waiting time and ’s waiting time:

|  |  |  |
| --- | --- | --- |
|  |  | (20) |

In practice, we will calculate for GR,, ER, AR, and compare these strategies’ efficiency.

* 1. Implementation and analysis

We collected GTFS real-time trip update data and corresponding GTFS static schedule data of the COTA (Central Ohio Transit Authority) bus system in Columbus, Ohio from January 2018. The data is stored in a MongoDB database. The GTFS trip update contains the ETA and ATA for every stop and every minute. We will develop the algorithm in Python environment and maintain our smart transit database.

The apps will calculate the walking distance between the user and the target stop and get expected HDT by subtracting ETA of bus and the calculated walking time. In all our maps, we assume that people’s walking is based on Euclidean distance and their walking speed is 1.4 meters per seconds.

Our analysis includes:

1. Find the PR optimal in the PR family;
2. Calculate waiting time difference for GR, PR optimal, ER, AR;
3. Visualize the waiting time difference for different strategy, based on different stops and different walking time;
4. Visualize the temporal pattern.
5. Analysis

In this section, we focus on the geographic and temporal analysis of buffer, waiting time, and waiting time difference between different TPSs. For the study area and target, we choose COTA bus No. 2 from Southeast to Northwest from Feb 2018 to Feb 2019. The bus route has two schedules: the frequent one originates from blue circled stop (frequent originating stop) in Figure 8 with headway of 15 minutes or better, while the standard on originate from red circled stop (standard originating stop) in Figure 8 with headway of 15 – 30 minutes (COTA, 2019).

* 1. The parameters of Prudent Relaxation optimal TPS

We first calculated and analyzed the optimal insurance buffer (IB) for Prudent relaxation optimal strategy. Figure 5 shows the geographic distribution of PR optimal’s insurance buffer. It shows an extreme geographically diverse pattern for IBs. So, we cannot assign a single IB for all stops and different users with different walking distance.

Figure 6 shows the miss risk for the PR optimal strategy, and it shows an unnatural high risk for PR optimal strategy. This indicates that even if users adopt the optimal IB from 0 to 300 seconds with the shortest waiting time, they still have a high risk of missing a bus. The actual optimal IB exceeds the limit of 300 seconds in the optimization process. Thus, for IB between 0 and 300 seconds, it will not help users to save time and they will need even larger IB.

For destination stop (blue circled) in the suburban area, traffic’s impact on the performance is less, thus buses are usually more punctual and stable and they have less delay to reclaim. Therefore, the IB is less and people will be less vulnerable to miss risk. Figure 7 also shows the miss risks of the stops near destination for PR optimal strategy is the lowest among all stops.

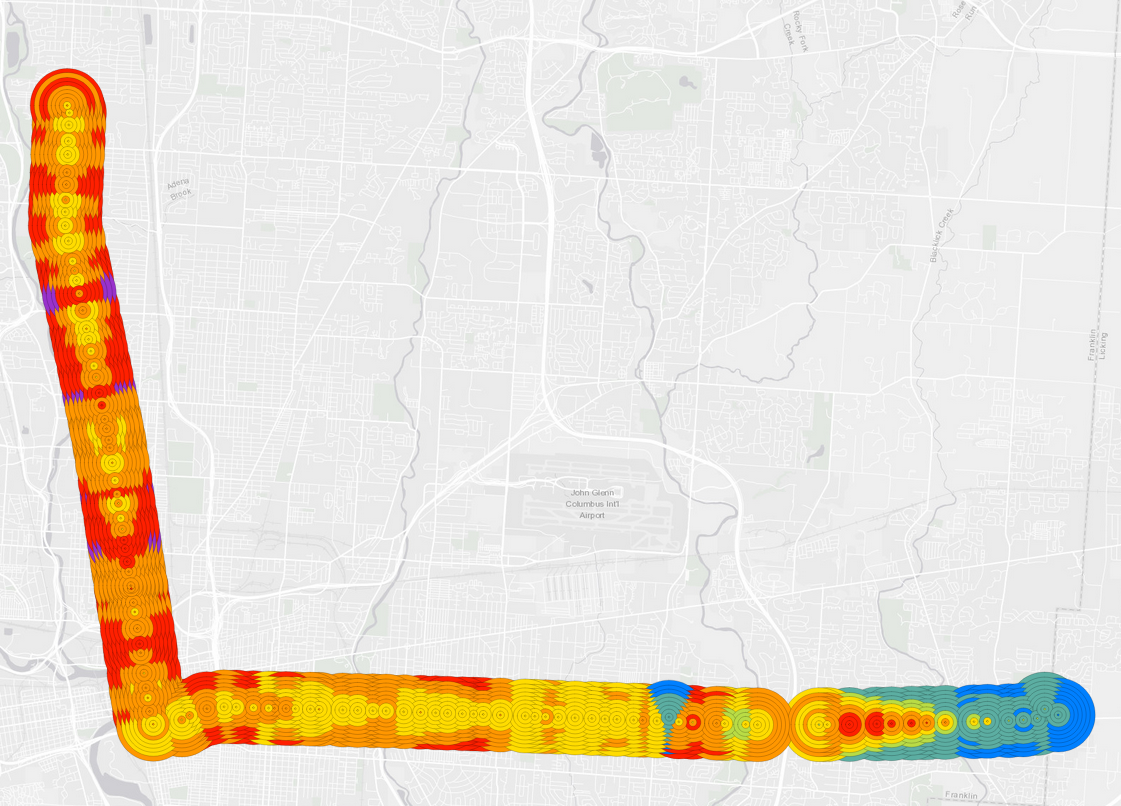


Figure 6 PR optimal’s insurance buffer for each stop and walking time in COTA bus route No. 2 from Southeast to Northwest in 2018.

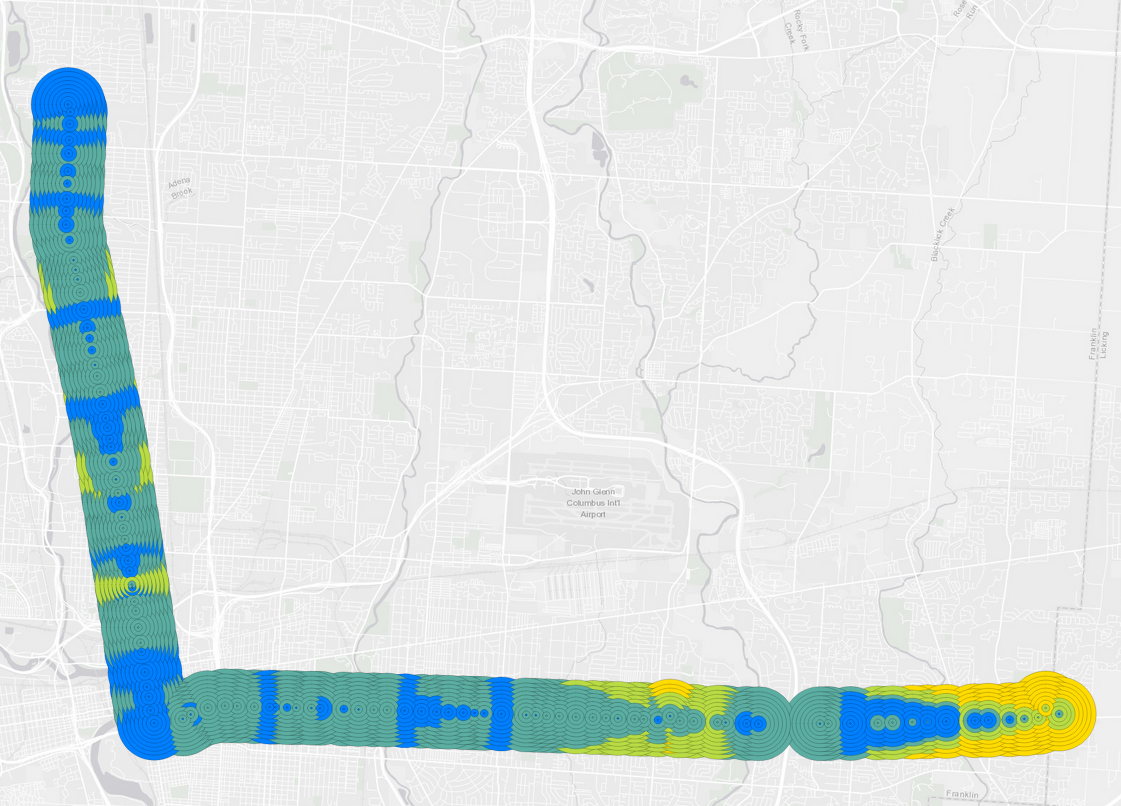


Figure 7 PR optimal’s average miss risk for each stop and walking time in COTA bus route No. 2 from Southeast to Northwest in 2018.

* 1. PR optimal’s optimality

We moreover calculated the waiting time of PR optimal and waiting time difference between PR optimal and other TPSs for walking time from 0 to 10 minutes.

[PR optimal’s waiting time]

Noticeably, there are two significant high clusters near the two originating stops (red circled and blue circled) in the standard and frequent No.2 bus schedule. Figure 8 shows the average waiting time on COTA bus route No.2 from Southeast to Northwest. Because the actual optimal buffer exceeds the limits of 300 seconds in the optimization process, there are a significant high waiting time cluster near the standard originating stop (red circled). Besides the high miss risk observed in Figure 7, the headway in these stops are larger in the bus schedule. Thus, the standard originating stop have almost twice larger waiting time than the frequent originating stop (blue circled). On the contrary, the stops near the destination stop have significantly low waiting time. Figure 8 is also the waiting time difference between PR optimal and omnipotent relaxation (OR) TPS, since OR will always achieve 0 waiting time.

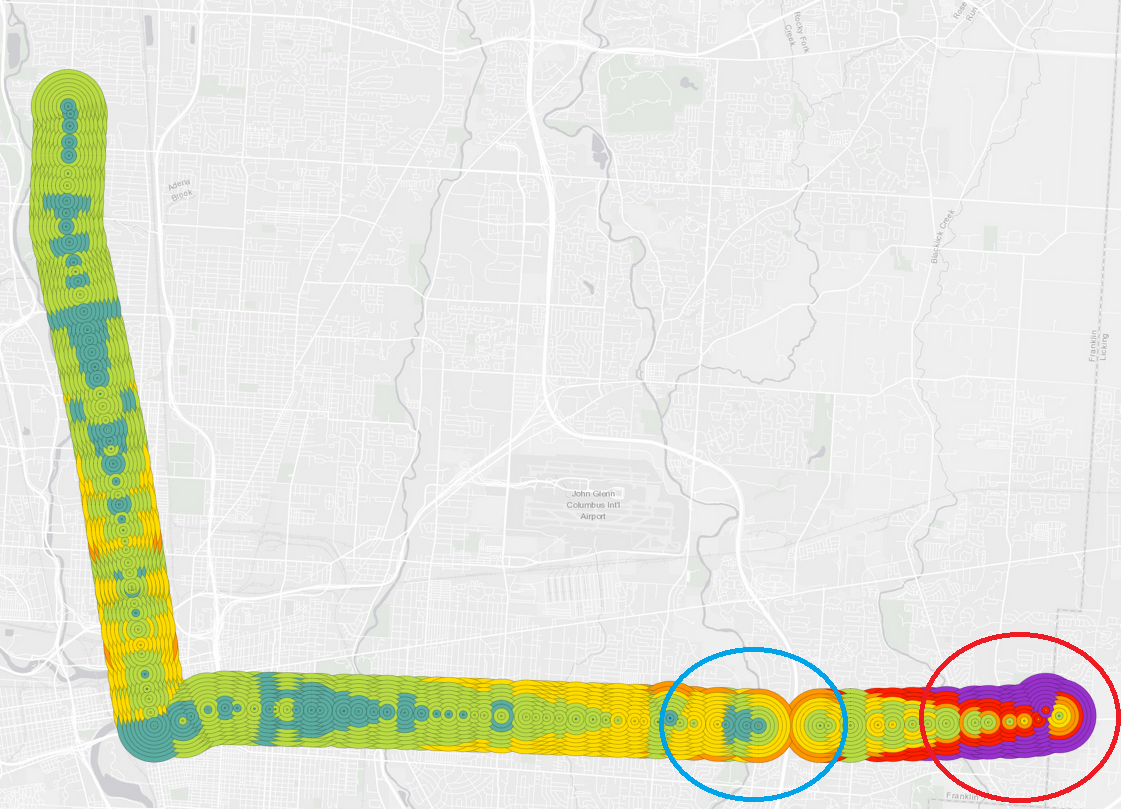


Figure PR optimal’s average waiting time for each stop and walking time in COTA bus route No. 2 from Southeast to Northwest in 2018.

[Difference from null relaxation]

Figure 9 shows the average waiting time difference on COTA bus route No. 2 from Southeast to Northwest. Red means PR optimal users’ waiting time is longer than NR users; blue means PR optimal users’ waiting time is shorter than NR users. The differences represent the distinction between performance of best RTA users (PR optimal) and best non-RTA users (NR), respectively.

First, we can observe that PR optimal does not outperform NR for all stops. In fact, for most stops, especially for those stops in the upstream near the originating stops, NR’s performance is much better than PR optimal. For stop near the standard originating stop, people who observe PR optimal TPS will have to wait [add data] minutes more than the people who follow the schedule.

There are two reasons for NR outperforming PR optimal: First, for some stops, especially stops near the originating stops, RTA users will be more likely miss the bus due to the high miss risk, thus waiting much longer time than the schedule; Second, for some stops, the obtained IB from the PR optimization process could be too large: user’s arrival time is even earlier than the scheduled time. Consequently, although IB guarantees synchronization when the user catches the bus, it makes people wait more time. In conclusion, too large or too small buffers all impair the effectiveness of PR optimal TPS, and too small buffers will especially result in desynchronization and suffer more waiting time.



Figure 9 The average waiting time difference between PR optimal and NR users on COTA bus route No. 2 from Southeast to Northwest in 2018.

[Difference from AR/ER/GR]

Figure 10 shows the waiting time difference between AR/ER/GR and PR optimal. AR and ER users will wait [insert data] more time than PR optimal.

Intuitively, arbitrary relaxation has an average waiting time of half of the headway. Therefore, we can observe a drastic change between standard schedule and frequent schedule due to the increase of headway. The performance of AR is expected to be the worst among the TPSs. However, arbitrary relaxation has the best performance among the three TPSs. In fact, greedy relaxation, as the only RTA TPS, has the worst performance and empirical relaxation using average is also worse than arbitrary relaxation. This proves that RTA users without proper advice could wait significantly longer than even arbitrary relaxation. The reason is that GR and ER (using average) will be more likely to desynchronize when catching bus thus have large miss risk. Statistically speaking, AR’s probabilistic distribution’s mean is half of the headway before the schedule; however, the mean of ER (using average)’s probabilistic distribution is all deterministic bus’s departure time’s mean. Therefore, when an ER arrives according to the average time, a considerable portion of distribution will be desynchronized with the user arrival time. Within bus departure time’s distribution, those desynchronized portion which falls before the actual departure time will incur a time penalty equal to almost a full headway. Likewise, GR’s mean is also near or even after departure time; also, due to its deterministic nature, GR could result in systematic higher chance of desynchronization, making its performance even worse than ER. This can also be observed from GR and ER’s miss rate. Figure 10 (bottom right) shows the miss risk difference between GR and PR. The average miss risk of GR is [add data] while the average miss risk of ER is [add data]. Therefore, GR and ER (using average) are not effective TPSs.

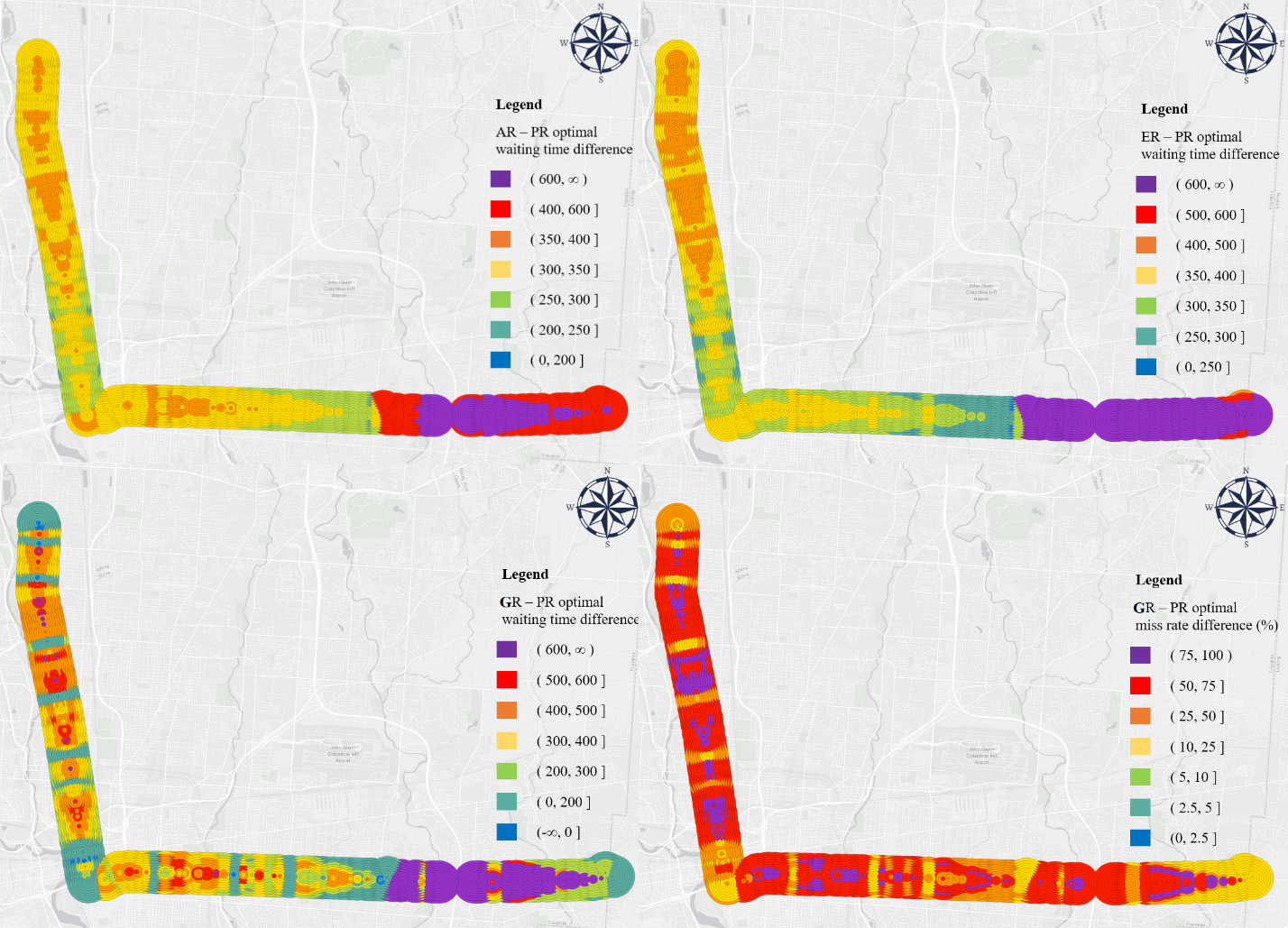


Figure 10 AR (top left), ER (top right), and GR (bottom left) - PR optimal waiting time difference and GR – PR optimal miss rate difference (bottom right)

* 1. Geographic patterns

Although many figures above visualize different variables, they share a similar geographic pattern. We investigate the geographic patterns from three directions.

[Vertical – Delay propagation]

Park, Mount, Liu, Xiao, & Miller (2019) investigated COTA bus system’s on-time performance and public bus delay propagation using GTFS real-time data (Park, Mount, Liu, Xiao, & Miller, 2019). The delay propagates via bus system’s network. The paper also presents the pattern of propagating delays and models it with an exponential model:

|  |  |  |
| --- | --- | --- |
|  |  | (21) |

Where: is the difference of delays at stop *p* and stop *q*. is the delay at stop *p*. is the transit route-based distance from *p* to *q*. is the delay constant (Park et al., 2019).

In the vertical direction, the miss rate, waiting time, and waiting time difference are all correlated with the delay propagation. Larger delay difference between two subsequent stops means faster delay reclamation. Figure 3 shows that larger reclaimed delay will incur larger miss risk and consequently longer waiting time. Similarly, if a trip is frequently reclaiming delay at a certain stop, the insurance buffers/expected waiting time are higher to compensate for the expected higher risk of missing. Last, if the average reclaimed delay is extremely large at a stop, then the optimal buffer may exceed the buffer range (0 – 300 seconds) set in the optimization process. Thus, the obtained optimal buffer is not the actual optimal buffer, which can be observed in Figure 6 and Figure 7. In conclusion,

[Diagonal – Contour lines]

An obvious diagonal pattern can be observed in the map: parallel diagonal contour lines. For each point on the contour line, their waiting time difference is the same. The bottom right corner of Figure 9 shows contour lines’ consistency. This is because essentially these users are aiming for the same bus. Therefore, if no disturbance or acceleration, their trips are the same only with different walking distance. The contour lines’ formation can also be understood in a temporal sense: after the bus receives a disturbance at a stop, the disturbance will persist and spread outwards when the bus is moving to following stops.

[Horizontal – Walking time influence]

For every group of concentric circles in Figure 7, from the center to the outskirts of the stop, the waiting time difference’s changing pattern is increasing. The upper right corner of Figure 9 shows the theoretical changing pattern in the horizontal direction. The exponential model is derived from vertical and diagonal pattern.

This supports the claim that the longer distance the user lives from the stop, the more unstable their trip becomes. During the longer walking time to the stop, the bus could be more likely to accelerate to catch up the delay, making RTA users miss the target bus.

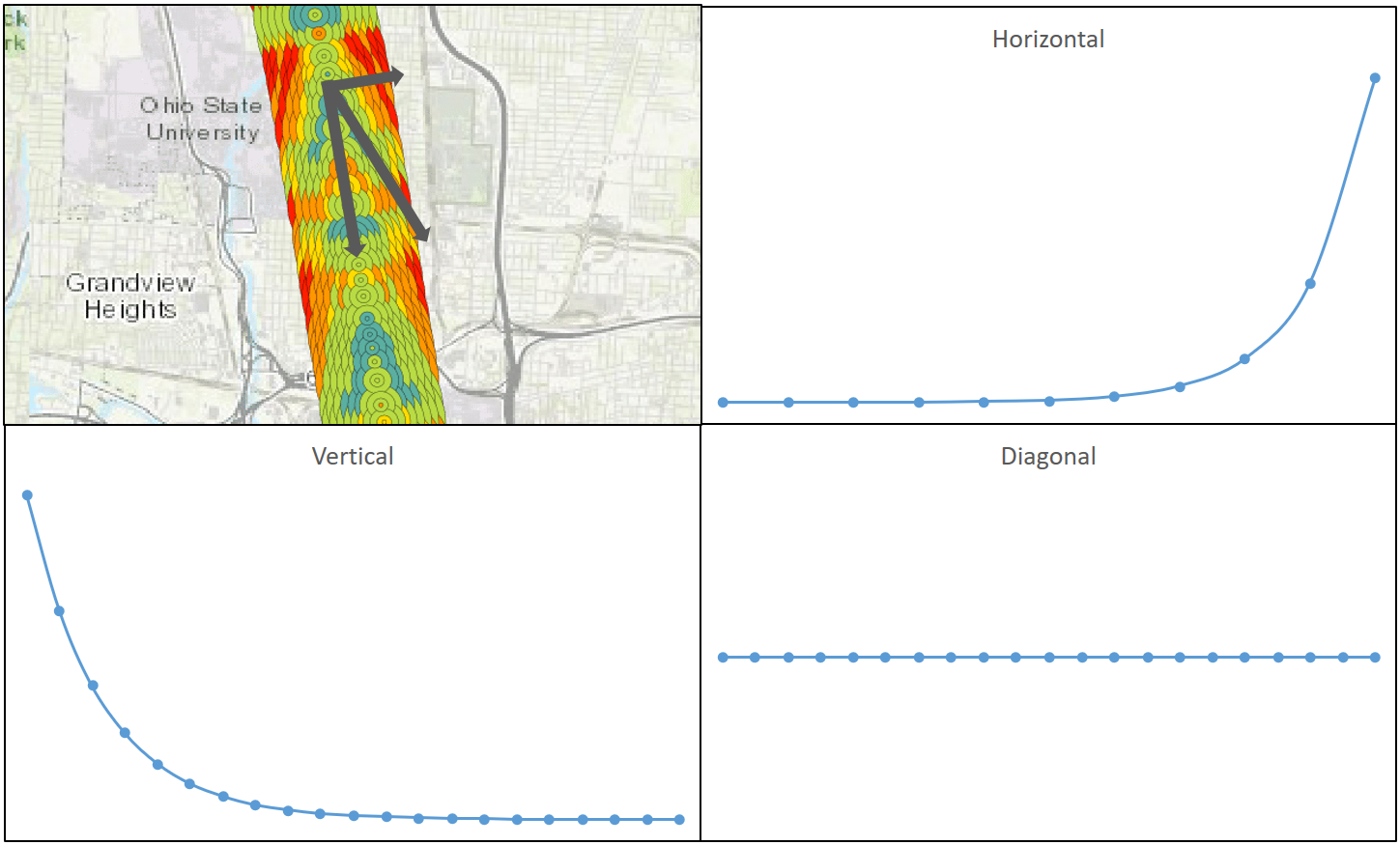


Figure 11 Waiting time difference's theoretical changing pattern on three directions

* 1. Temporal patterns
  2. Marginalized stops

We can observe that the first few start stops near the originating stop, which can be called *marginalized stops*, have a significant larger waiting time difference. It suggests that RTA strategies may be less effective at marginalized stops. In these areas, the walking time could be larger than the running time from the start stop to the target stop. These stops are marginalized, in the geographic, temporal, and social justice sense.

First, the start time at the originating station is frequently delayed, which can be called *originating delay*. Also, theoretically, these users’ HDT could be earlier than the bus’s start time at the originating station, which make the process even more risky. When the users should be heading for the stop, the bus has not left the originating stop, where users cannot have access to the real-time data in advance. Moreover, during users’ walking, the bus will likely reclaim more delay which will deprive RTA user’s advantages. After few stops when the originating delay has been fully reclaimed, the bus will regain synchronization in the following stops.

During the process of delay reclamation, the originating delay is not reclaimed gradually during all stops. Instead, to catch up the delay as soon as possible, the whole process is finished only at marginalized stops. Park et al. (2019) also present visualization proofs for the presence of the rapid originating delay reclamation (Park et al., 2019). Equation 18 and Figure 8 shows the existence of originating delays and the exponential changing trend for the delays.

Due to the universal presence of frequent originating delay, it will increase users’ missing risk in these areas. Figure 10 shows marginalized stops’ waiting time difference for different routes in COTA bus system. This phenomenon universally exists in the system.

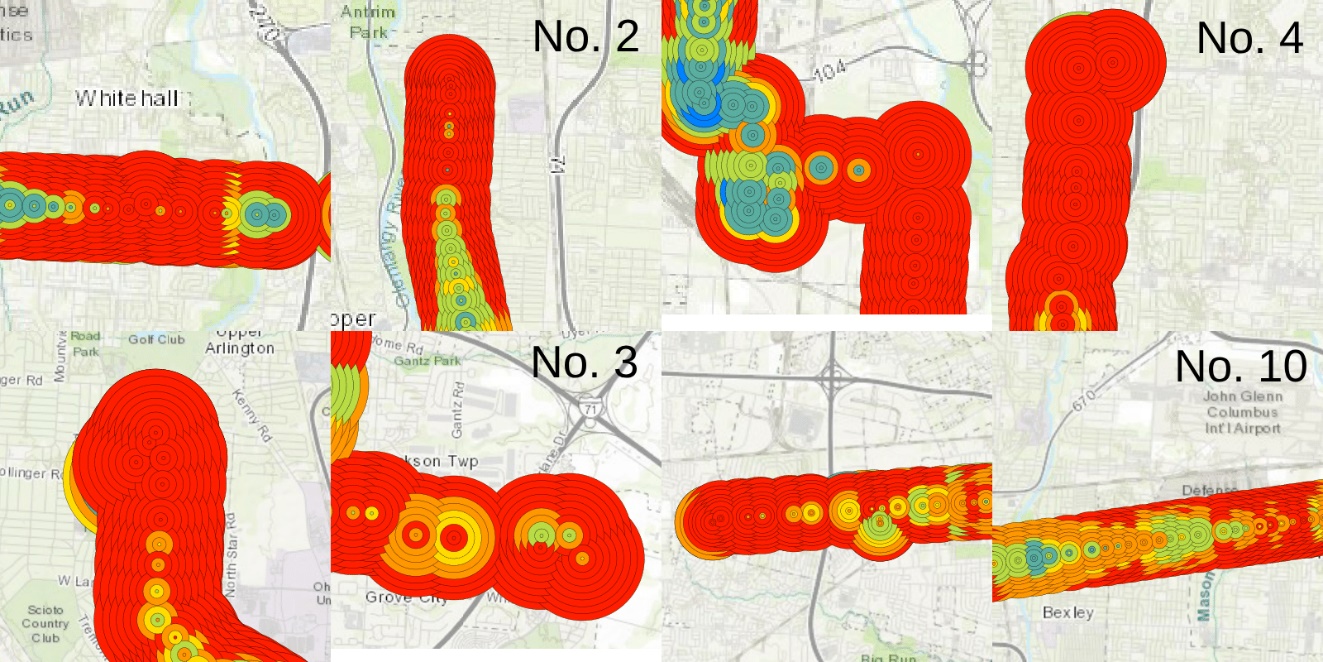


Figure 12 Marginalized stops' waiting time difference for both directions.

1. Conclusion

Real-time information (RTI) and real-time transit apps (RTA) can significantly decrease transit users’ waiting time (Brakewood & Watkins, 2018). However, few studies systematically investigated waiting time reduction’s spatiotemporal patterns and optimized the system performance from users’ perspective. In this study, using GTFS real-time data, we first developed *volunteered optimization* theory and calculated RTA and non-RTA users’ real-time performance. We theorized different waiting time relaxation strategies for trip planning under different scenarios and user groups. Then, we optimized RTA users’ *prudent relaxation* strategy with optimal insurance buffer, so that RTA users can minimize waiting time in the individual level. Based on the PR optimal strategies, we calculated the waiting time difference between RTA and non-RTA users. Moreover, we observed the presence of originating delay, *marginalized stops* and injustice distribution of delay reclamation. The empirical results and the volunteered optimization system can provide vital information for transit users, planners, and administrators.

Reference:

Beirão, G., & Cabral, J. A. S. (2007). Understanding attitudes towards public transport and private car: A qualitative study. *Transport Policy*, *14*(6), 478–489.

Bowman, L. A., & Turnquist, M. A. (1981). Service frequency, schedule reliability and passenger wait times at transit stops. *Transportation Research Part A: General*, *15*(6), 465–471.

Bradburn, N. M., Sudman, S., Blair, E., Locander, W., Miles, C., Singer, E., & Stocking, C. (1979). *Improving interview method and questionnaire design: Response effects to threatening questions in survey research*. Jossey-Bass San Francisco.

Brakewood, C., Barbeau, S., & Watkins, K. (2014). An experiment evaluating the impacts of real-time transit information on bus riders in Tampa, Florida. *Transportation Research Part A: Policy and Practice*, *69*, 409–422.

Brakewood, C., Macfarlane, G. S., & Watkins, K. (2015). The impact of real-time information on bus ridership in New York City. *Transportation Research Part C: Emerging Technologies*, *53*, 59–75.

Brakewood, C., & Watkins, K. (2018). A literature review of the passenger benefits of real-time transit information. *Transport Reviews*, 1–30.

Cats, O., & Gkioulou, Z. (2017). Modeling the impacts of public transport reliability and travel information on passengers’ waiting-time uncertainty. *EURO Journal on Transportation and Logistics*, *6*(3), 247–270.

Chow, W., Block-Schachter, D., & Hickey, S. (2014). Impacts of real-time passenger information signs in rail stations at the Massachusetts Bay Transportation Authority. *Transportation Research Record*, *2419*(1), 1–10.

COTA. (2019). 2 N HIGH / E MAIN. Retrieved June 3, 2019, from https://www.cota.com/wp-content/themes/gotravel-child/PDF/lines/2.pdf

Dziekan, K., & Vermeulen, A. (2006). Psychological effects of and design preferences for real-time information displays. *Journal of Public Transportation*, *9*(1), 1.

Fan, Y., Guthrie, A., & Levinson, D. (2016). Waiting time perceptions at transit stops and stations: Effects of basic amenities, gender, and security. *Transportation Research Part A: Policy and Practice*, *88*, 251–264.

Ferris, B., Watkins, K., & Borning, A. (2010). OneBusAway: results from providing real-time arrival information for public transit. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1807–1816). ACM.

Fonzone, A., Schmöcker, J.-D., & Liu, R. (2015). A model of bus bunching under reliability-based passenger arrival patterns. *Transportation Research Procedia*, *7*, 276–299.

Fries, R. N., Dunning, A. E., & Chowdhury, M. A. (2011). University traveler value of potential real-time transit information. *Journal of Public Transportation*, *14*(2), 2.

Frumin, M., & Zhao, J. (2012). Analyzing passenger incidence behavior in heterogeneous transit services using smartcard data and schedule-based assignment. *Transportation Research Record*, *2274*(1), 52–60.

Gkioulou, Z. (2013). Evaluating the impact of waiting time uncertainty on passengers´ decisions.

Google Developers. (2016). GTFS Static Overview | Static Transit | Google Developers. Retrieved March 8, 2018, from https://developers.google.com/transit/gtfs/

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

Goyder, J. (1986). Surveys on surveys: Limitations and potentialities. *Public Opinion Quarterly*, *50*(1), 27–41.

Jolliffe, J. K., & Hutchinson, T. P. (1975). A behavioural explanation of the association between bus and passenger arrivals at a bus stop. *Transportation Science*, *9*(3), 248–282.

Liu, Y., Shi, J., & Jian, M. (2017). Understanding visitors’ responses to intelligent transportation system in a tourist city with a mixed ranked logit model. *Journal of Advanced Transportation*, *2017*.

Luethi, M., Weidmann, U., & Nash, A. (2007). Passenger arrival rates at public transport stations. In *86th Transportation Research Board Annual Meeting (TRB 2007)*. Institute for Transport Planning and Systems, ETH Zurich.

Neuman, W. L., & Robson, K. (2014). *Basics of social research*. Pearson Canada Toronto.

Papangelis, K., Nelson, J. D., Sripada, S., & Beecroft, M. (2016). The effects of mobile real-time information on rural passengers. *Transportation Planning and Technology*, *39*(1), 97–114.

Park, Y., Mount, J., Liu, L., Xiao, N., & Miller, H. J. (2019). Assessing public transit performance using real-time data: spatiotemporal patterns of bus operation delays in Columbus, Ohio, USA. *International Journal of Geographical Information Science*, 1–26.

Reed, T. B. (1995). Reduction in the burden of waiting for public transit due to real-time schedule information: a conjoint analysis study. In *Pacific Rim TransTech Conference. 1995 Vehicle Navigation and Information Systems Conference Proceedings. 6th International VNIS. A Ride into the Future* (pp. 83–89). IEEE.

Rossi, P. H., Wright, J. D., & Anderson, A. B. (2013). *Handbook of survey research*. Academic Press.

Transit app. (2019). How we shrank our trip planner till it didn’t need data. Introducing public transit’s fastest, tiniest, offline-capable trip planner. Retrieved December 5, 2019, from https://medium.com/transit-app/how-we-shrank-our-trip-planner-till-it-didnt-need-data-84984ca56663

Watkins, K. E., Ferris, B., Borning, A., Rutherford, G. S., & Layton, D. (2011). Where Is My Bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. *Transportation Research Part A: Policy and Practice*, *45*(8), 839–848.

Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication*, *10*(3), JCMC1034.