Measuring the Impacts of Disruptions on Public Transit Accessibility and Reliability

Luyu Liu [1, 2, *], Adam Porr [3], Harvey J. Miller [1, 21]

[1] Department of Geography, The Ohio State University, Columbus, OH, USA

[2] Center for Urban and Regional Analysis, The Ohio State University, Columbus, OH, USA

[3] Mid-Ohio Regional Planning Commission, Columbus, OH, USA

Public transit systems are facing higher risk of system degradation from external disruptions, affecting their ability to deliver reliable accessibility to transit users. Therefore, resilience, the ability to maintain functions during a disruption, becomes a crucial assessment of public transit systems. In this paper, we calculate two space-time prism-based measures with General Transit Feed Specification real-time (GTFS-RT) data: realizable real-time accessibility, a conservative real-time accessibility measure that can be achieved by users subject to delays, and scheduled accessibility, accessibility based on schedule. We also define accessibility unreliability, the deviation between realizable accessibility and scheduled accessibility, to measure the reliability of delivered accessibility. We use the two measures to conduct two case studies of short- and long-term disruptions, namely Ohio State football games and the COVID-19 pandemic, on the Central Ohio Transit Authority (COTA) bus system in Columbus, Ohio. We find there are two peaks of high unreliability before and after each football games, with the stadium as the geographic center of the disruption. The after-game peaks are shorter and more intense than the before-game. We also find COVID-19 had persistent negative impacts on accessibility and reliability: Realizable accessibility universally declined during the pandemic, but only part of cities experienced unreliability increase, primarily in urban perimeters and suburbs. Improved traffic conditions during the pandemic may help to reduce unreliability, but the later service cuts increased unreliability. The two case studies prove the effectiveness of the method to detect system disturbances and provide important guidance for public transit system operation and planning.

Keywords: Realizable accessibility; Accessibility reliability; Public transit system; Resilience; Football game; the COVID-19 pandemic.

[*] Corresponding author, email: liu.6544@osu.edu
ORCID: Luyu Liu (0000-0002-6684-5570), Adam Porr (0000-0002-4776-5575), Harvey J. Miller (0000-0001-5480-3421)

1 Introduction

As extreme weather, pandemics, and social unrest and other disruptions shake our world, public transit systems are operating in more unstable environments, challenging their ability to deliver reliable accessibility to their clientele. Therefore, the resilience of a public transit system – the ability to maintain its functions during after external shocks and disruptions – should be a major focus for public transit research and planning. Accessibility is a primary indicator of a public transit system's utility, as it determines people's ability to reach opportunities and resources given the limited time available to conduct essential and discretionary activities (Tong et al., 2015). Deviation between scheduled and delivered service propagate through routes and can spread through the system due to interconnections among equipment and operators (Park et al., 2020). Delays degrade user experience and the usefulness of the transit system and have negative consequences for transit-dependent riders who may miss work, medical appointments and other time-critical events. This makes reliability one of the most important factors that affect people's preference and use of public transit (Chakrabarti & Giuliano, 2015; Erhardt et al., 2022).

Public transit faces both short-term and long-term disruptions. Short-term disruptions are temporary events that do not fundamentally alter the service and infrastructure. Prominent examples are traffic jams, weather events, and major entertainment events such as concerts, sports events, games, and street festivals. Short-term disruptions affect accessibility primarily by influencing the on-time performance, in the form of delayed or sometimes early arrivals. Long-term disruptions have persistent impacts on the system; examples include pandemics, damaging weather events, and infrastructure failures. Long-term disruptions can create more nuanced patterns of unreliability beyond direct and temporary impacts on on-time performance, including budget constraints and subsequent route and schedule changes.

There are still large gaps in research on the reliability and resilience of public transit-based accessibility. First, prior studies focus on system resilience based on travel time, ridership, and capacity (Mudigonda et al., 2019), rather than accessibility and accessibility reliability. Second, studies of short-term disruptions are lacking. The historical lack of reliable high-resolution data sources made it hard to conduct empirical analysis on short-term disruptions. Finally, few papers discussed the recoverability of transit accessibility after a disruption; this is a major aspect of system resilience. Due to the recent availability of high-resolution real-time data, we now can address these gaps. In this paper, we use *realizable real-time accessibility* – a space-time prism-based measure (H. J. Miller, 1999, 2017) that conservatively measures the accessibility that can be achieved by users in a public transit system subject to delays (Liu et al., 2022). Whereas accessibility measures based exclusively on schedule information fail to account for delays and measures based on real-time vehicle locations make unrealistic assumptions about the information available to users, realizable accessibility avoids both limitations, thereby simulating the trip-planning process that can be realized by users in real-world scenarios. The gap between scheduled accessibility and reliable accessibility is an indicator of the reliability of transit accessibility (Liu et al., 2022). We use the realizable accessibility and reliability measures to conduct two case studies of the impacts of short-term and long terms disruptions on public transit: college football games whose traffic impacts result in bus delays; and the COVID-19 pandemic that persistently altered the system's schedules and routes. We conduct these case studies using data from Columbus, Ohio, USA and its bus-based public transit system, the Central Ohio Transit Authority (COTA).

2 Background

We review relevant literature in this section. We first introduce the transportation resilience and its two core features. We then review the development of public transit reliability from travel time-based reliability to accessibility reliability. We finally assess disruptions and their impacts on public transit accessibility and reliability.

2.1 Resilience

Resilience is the capacity of a system to maintain its functions during a disruption (Azolin et al., 2020; Holling, 1973). As climate change, pandemics, and energy crises increase the risk and frequency of disruptions, transportation resilience becomes a new focus of transportation focus. However, the definition of transportation resilience can be heterogenous and nuanced. Most prior research agrees that resilience includes two core features: robustness and recoverability (Azolin et al., 2020; Gu et al., 2020; Wan et al., 2018).

Robustness – some researchers also use the terms adaptability and reliability, or vulnerability and unreliability as antonyms – is the ability to maintain service during a disruptive event. An ideally robust transport system should still maintain a minimum required performance in the face of a disruptive event. Robustness is measured by the decline of a system performance (Gu et al., 2020; Wan et al., 2018). Recoverability – some researchers also use resilience or resiliency – is the ability for the system to return to its previous state in a timely manner (Wan et al., 2018). It is usually measured by the time from the disruptive event happens to the time when the performance recovers to pre-disruption level (Gu et al., 2020; Wan et al., 2018). The two aspects determine the transport system's ability to resist, adapt to, and recover from the disruption.

2.2 Accessibility Reliability of Public Transit systems

Reliability can be defined as the variation of a public transit system's performance (Gu et al., 2020); however, its specific definition can be nuanced, depending on the performance measures. Most of the prior research investigated travel time reliability (Gu et al., 2020; Kathuria et al., 2020). Carrion & Levinson (2012) categorized this concept into three categories: 1) centrality-dispersion, which measures the variation of travel time around the mean value; 2) scheduling delays, which measures the difference between preferred travel time and actual travel time; 3) average delays, which measures the difference between scheduled time and actual time, i.e., on-time performance of a public transit system. Travel time reliability represents the fidelity of the transit service; higher reliability means that a user can expect their incoming trips to abide by the scheduled or average performance.

Due to the direct link between travel time and accessibility, the reliability of accessibility can also be defined as its variation over time. However, depending on the standard of comparison, i.e., average accessibility or scheduled/expected accessibility, the data and methods used by different studies can still vary. D'este & Taylor (2003) and

Taylor & D'Este (2007) first introduce reliability and vulnerability related to accessibility. These traditional studies based on vulnerability analysis utilized road network data, including road capacity and geometries, and empirical traffic flow data to simulate the reliability of the transportation system. However, due to the unique time-dependent and schedule-based nature of public transit systems and their accessibility (Gendreau et al., 2015; Pereira, 2019), the simulation methods based on road network can be very inaccurate. Meanwhile, due to the lack of empirical high-fidelity public transit data, these traditional data are not grounded in real-world scenarios, which could lead to even more significant misestimation of the systems' performance.

Recent progress in data curation and collection techniques enables the measurement of accessibility reliability patterns with higher frequency and fidelity. This transformative development enhances our ability to accurately assess and understand the performance dynamics of public transit systems, especially accessibility and reliability. Due to the availability of these new datasets, such as General Transit Feed Specification real-time (GTFS-RT) data and Automatic Vehicle Location (AVL) data, many studies explored the possibilities of defining the concept of real-time accessibility and accessibility reliability, which is similar to actual travel time and travel time reliability, respectively. Following the same logic, accessibility reliability can be defined as the variation between expected/scheduled accessibility and actual accessibility delivered by the system based on its performance. For example, Wessel et al. (2017) and Wessel & Farber (2019) investigated the accuracy of schedule-based accessibility by calculating the difference between delivered accessibility and scheduled accessibility. They use retrospective GTFS-RT data to estimate the delivered accessibility in a public transit system based on vehicle location data. They show that schedule-based accessibility overestimates delivered accessibility. However, their retrospective accessibility measure assumes transit users have complete *a priori* knowledge of actual arrival time of vehicles; this requires clairvoyance or a perfect, idealistic real-time bus information system. In practice, real-time bus information has more complex impacts on waiting and travel times: it can increase waiting and travel times due to the temporal granularity of updates combined with bus operators attempting to make up for delays (Liu & Miller, 2020a). This implies that the retrospective measure is also an overestimate of delivered accessibility.

To resolve this issue, Liu et al. (2022) introduce realizable real-time accessibility as a more conservative measure of transit-based accessibility. They compare this measure to scheduled accessibility as a measure of accessibility reliability – the difference between scheduled and realizable accessibility. This represents the degree to which expected measure overestimate actual accessibility, as well as the fidelity of public transit systems to deliver an accurate and reliable service.

Reliability can also be used to measure resilience, namely robustness and recoverability of a transit system. Robustness as the increase of accessibility unreliability during a disruption, while recoverability can be measured by the recovery period of accessibility reliability after the disruption to a previous baseline. We will use this theoretical framework in our analysis.

2.3 Disruptions and transit reliability

A major factor in the reliability of accessibility delivered by a public transit system is both chronic and occasional disruptive events. Depending on the effects, persistency, and frequency of the event, we can categorize disruptions by: 1) Short-term and long-term (Lin et al., 2016), 2) planned and unplanned (Zhu & Levinson, 2012), and 3) Recurring and non-recurring (Lin et al., 2016; Park et al., 2020). These three categorizations are highly correlated with each other but not the same. In this paper, we adopt the short/long-term categorization based on the dimension of recoverability as we discussed above; we review the factors affecting public transit reliability in following paragraphs.

Short-term disruption. We define short-term disruption as the event that: 1) are short in time span: typically, not exceeding a single day, which is the time unit of the operation of most transit systems; 2) do not fundamentally change the schedule of the transit system. In that sense, short-term disruptions usually influence the unreliability by only on-time performance, i.e., delays and early arrival.

A primary example is traffic. As many public transit systems use buses and trams that share roads with other vehicles, traffic on roads can significantly impact the on-time performance (Carrion & Levinson, 2012; Park et al., 2020). Other examples include weather (e.g., heavy rain) (Mesbah et al., 2014; Pender et al., 2014) and major social events (e.g., concerts, sporting events, festivals, protests) (Berche et al., 2009). However, due to the ephemeral nature of these events and a previous lack of reliable high-resolution data, the research on this topic is still lagging.

Long-term disruptions. We define long-term disruptions as events that: 1) are longer in time span, which can last from several days to months and years; 2) affect both the on-time performance and the schedule; 3) may result in a new normal, rather than returning to the pre-disruption state. The studies and data on long-term disruptions are more abundant due to their more profound and persistent effects compared to short-term disruptions.

The COVID-19 pandemic is a major long-term disruption, if not the most important one in this century, that has huge impacts on human mobility in the entire world; almost all countries have witnessed major changes in traffic (Lee et al., 2020), working from home rate (Beck et al., 2020), and public transit ridership (Liu et al., 2020). Despite having a V-shaped recovering trend in people’s mobility during the early stage of the pandemic (Kim & Kwan, 2021), COVID-19 has persistent and nuanced negative implications on public transit accessibility. For public transit, Kar et al. (2021) studied the public transit accessibility to essential services in 22 US cities in 2021 and found significant declines; the paper also pointed out that the pandemic-related decline primarily impacts marginalized communities. In response to the disruption, transit authorities and government also enacted policies and system adjustments to resist the negative impacts. For example, Singh et al. (Singh et al., 2022) found COVID-19 pandemic has negative impact on the transit accessibility in Winnipeg, Canada but a new BRT system helps to increase the accessibility for underprivileged populations.

Extreme weather events can also incur persistent disruption to public transit and transit accessibility. A prime example is flood and sea level rising caused by climate change. Li et al. (Li et al., 2018) simulated the potential effect of a 100-year pluvial flood on Shanghai Metro, China and found universal decrease in accessibility. He et al. (He et al.,

2021) found flood disruptions lead to increase in headways and loss of job accessibility in Kinshasa, Democratic Republic of the Congo. Despite many existing discussions on disruptions' impacts on public transit and its accessibility, very few papers offered a holistic and high-fidelity analysis on public transit accessibility and reliability.

3 Method

3.1 Case Study Site and Data

We choose the city of Columbus, Ohio, as our case study site. Columbus is the state's capital, the largest city, and a major metropolis in the US Midwest. It is also the home to The Ohio State University and its college football team. The Central Ohio Transit Authority (COTA) is the city's public transit system; it serves the area with more than 1.2 million residents and generated 19 million trips in 2019. Figure 1 shows the population density and COTA bus routes with their corresponding frequency.

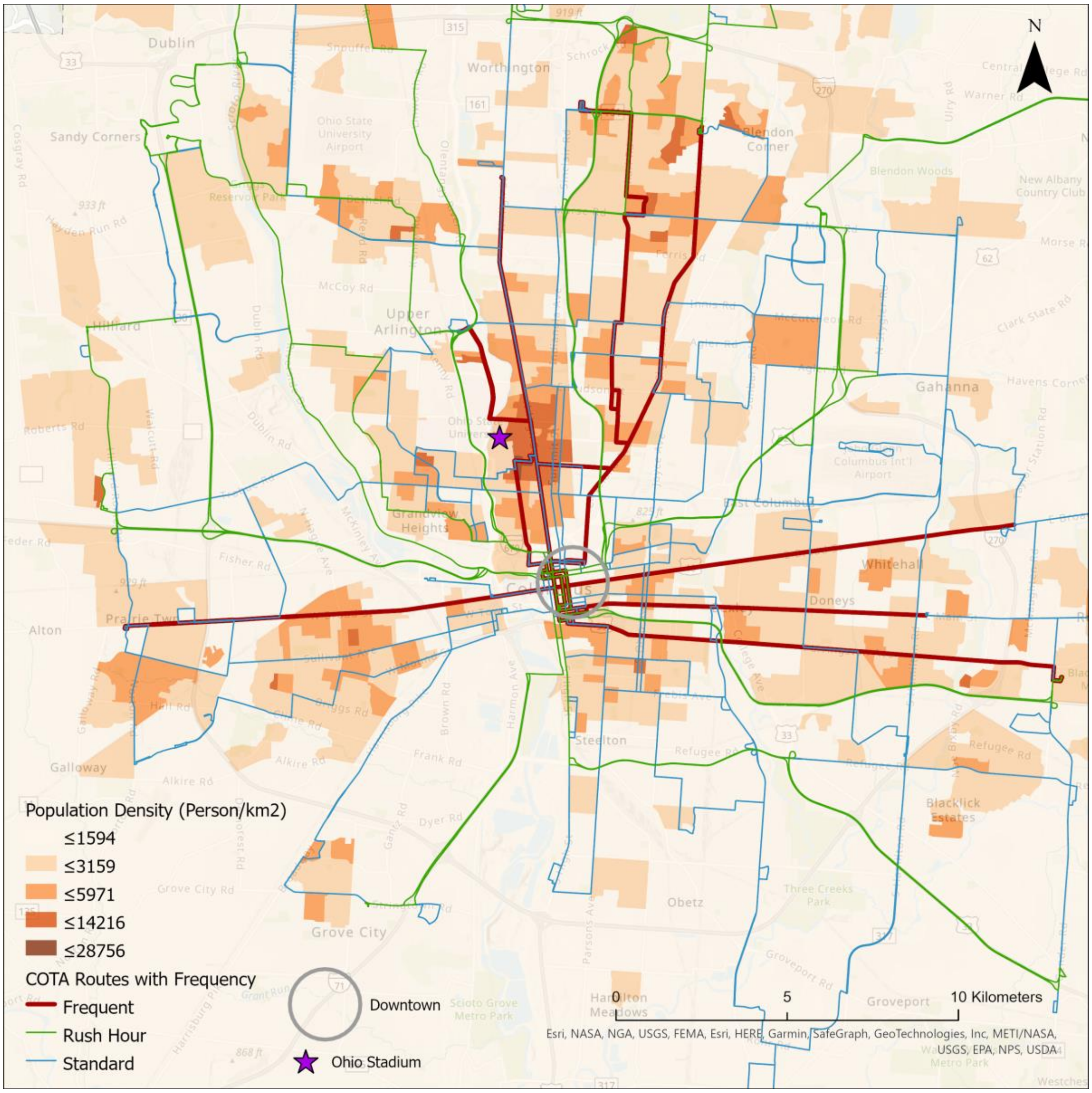

Figure 1: The population density of Columbus, OH and COTA bus routes with corresponding frequencies. Gray circle is downtown, and the purple star is Ohio Stadium.

The primary data source in this paper is General Transit Feed Specification (GTFS) data. It is the de facto standard to transmit real-time information (Antrim & Barbeau, 2017; Liu & Miller, 2020b). The data conforms to two standards, GTFS static and GTFS real-time data, which contain the schedule timetable and real-time timetable, respectively (Google, 2021; Google Developers, 2020). Based on the two datasets, we can calculate the past scheduled and actual arrival time for any bus at any stop. We collected GTFS static and GTFS real-time data from COTA's application programming interface (API) from May 2018 until January 2022.

3.2 Accessibility Measure

Accessibility is a diverse concept that can measure different aspects of mobility (E. J. Miller, 2018). In this paper, we focus on the measure of physical accessibility in a transit system. Physical accessibility measures the upper limits on the reachability of locations by a transit user given a time budget; in other words, the distance a user can travel using a transit service given a travel time budget such as 30 minutes.

We use a well-established time geography concept – the space-time prism (STP) – to quantify the physical accessibility (Hägerstrand, 1970; H. J. Miller, 2017). It represents the envelope of all possible space-time paths in three possible scenarios: i) travel from an origin to all possible destinations, ii) travel from all possible origins to a destination, or, iii) travel between an origin destination pair. In each scenario, the space-time prism is a function of departure and/or arrival times, a time budget, and the speed afforded by the mobility modes, including multimodal trips such as walking and public transit (H. J. Miller, 2017). In our analysis, we treat each bus stop as a single origin and calculate the prisms from each single origin to all possible destinations at a particular departure time. We calculate the implicit STP based on public transit stops – this is the number of accessible stops from an origin stop given a time budget (Liu et al., 2022). First, we introduce a decision variable to determine if a user can arrive at a stop within the time budget.

$$\delta_{ij\tau\phi} = \begin{cases} 1, if\ t_{ij\phi} \leq \tau \\ 0, if\ t_{ij\phi} > \tau \end{cases} \tag{1}$$

where $\delta_{ij\tau\phi}$ represents whether a user can arrive at stop j from stop i starting from time point ϕ within the time budget τ, and $t_{ij\phi}$ is the shortest travel time between stops i and j starting from a time point ϕ. We thus define the implicit STP as:

$$S_{i\phi} = \left\{ \sum_{j \in S} \delta_{ij\tau\phi} \,|\forall \tau \in \mathrm{T} \right\} \tag{2}$$

where $S_{i\phi}$ represents the implicit STP from stop i at time point ϕ, while T is the set of all time budgets and S is the set of stops. The implicit STP measures the accessibility to network nodes.

Note that transit networks are time-dependent: the travel times for transit passengers in each transit link are determined by their arrival time at the origin stops (Gendreau et al., 2015; Wang et al., 2019), because passengers must wait for a bus and cannot move without one. For example, a person who arrives early will not leave earlier than a person who arrives later if they take the same bus; meanwhile, a person who misses a bus will take significantly longer time in a same transit link. The dynamic weights of public transit network add on the difficulties and computational costs of the problem. To calculate the travel time, we developed a time-dependent Dijkstra algorithm to solve this special routing problem. We use a first-in-first-out (FIFO) rule to make the static Dijkstra algorithm compatible to a transit network with dynamic costs (Ahn & Shin, 1991; Ichoua et al., 2003). The rule assumes a public transit vehicle leaving an origin stop will never arrive later at the destination stop than a public transit vehicle on the same route that is scheduled later. One vehicle overtaking another in violation of the FIFO restriction is a rare event: we estimate from COTA data that 95% of the buses meet this restriction.

3.3 Unreliability measures

We define unreliability of transit accessibility as the deviation between the schedule-based accessibility and the delivered accessibility. Schedule-based accessibility represents the promise that the transit authorities make to users, which cannot be perfectly kept under most circumstances due to on-time performance loss.

However, the definition of actual experienced physical accessibility can be nuanced. As we already discussed in the previous sections, retrospective real-time STPs are not feasible for ordinary users to achieve in practice without clairvoyance or a perfect real-time bus information system. Liu et al. (2022) introduce a realizable real-time accessibility measure that assumes no real-time information as the system operates and deviations from schedules occur. It is calculated in two steps – planning and implementation. During the calculation process, the algorithm will first plan the trip according to buses' scheduled arrival time and then implement the plan using the actual arrival time instead (Liu et al., 2022).

We measure accessibility unreliability as the normalized difference between scheduled and realizable accessibility:

$$U_i^{\phi} = \frac{S_i^{\phi} - R_i^{\phi}}{R_i^{\phi}} = \left\{ \frac{s_{i\tau}^{\phi} - r_{i\tau}^{\phi}}{r_{i\tau}^{\phi}} \,|\forall\tau \in \mathrm{T} \right\} \tag{3}$$

where S_i^{ϕ} is the scheduled STP starting from a time point ϕ, R_i^{ϕ} is the realizable STP, $s_{i\tau}^{\phi}$ is the schedule-based number of accessible stops, and $r_{i\tau}^{\phi}$ is the realizable-based number of accessible stops. Both scheduled and realizable STPs are calculated with equation (2) but with different travel times calculated from the Dijkstra algorithm.

As discussed in the background section, there are two major aspects of resilience, namely robustness and recoverability. We implement the two concepts with accessibility unreliability. We can define robustness as the change of accessibility unreliability before and during the disruption and recoverability as the duration of the disruption's impact. Again, it is noteworthy that long-term disruptions will be very likely to land on a new normal, and COVID-19 is still an ongoing event; both factors make it impossible to define the recoverability of systems against COVID-19. We introduce specific definitions and analyses in the following sections with each case study.

3.4 Short-term Disruption: College Football Games

Ohio Stadium hosts college football games from September to December every one or two weeks, which attracted more than a hundred thousand viewers to the stadium before the pandemic (Kaufman, 2021). Home games attract large amounts of traffic to and from the Ohio Stadium before and after the game, creating short-term disruptions to both private and public transportation. Away games (i.e., those that take place at another university) also attract traffic to the vicinity of Ohio Stadium due to the desire of fans to associate with one another while watching the game on television but not to the same degree as a home game.

Ohio State football games are a good example of a short-term disruption because: 1) most football games last around 3 hours, which is short in time span compared to other disruptive events; 2) transit systems recover from the disruption relatively quickly as crowds and traffic disperse; 3) football games do not change the schedule of transit system in a fundamental way, despite some short-term rerouting in some areas. We select all home and away game days in 2018 and 2019 from September to December and calculate the accessibility unreliability; we also collected the weather data, including temperature and precipitation, and found no significant anomalies in the record as shown in Table 2 in the appendix. We also calculate this measure for other Saturdays without a home or away game in the same time period for comparison.

First, we investigate the temporal trend of accessibility reliability before and after the event time. Each game can have a different start time and impacts can thus occur at different hours; therefore, we categorize games based on their start time. There are three start time slots: 12:00, 15:30/16:00, and 19:30. There are 9 home games at 12:00, 4 home games at 15:30/16:00, and 1 home game at 19:30. Also, since the impacts of football games are spatially heterogenous, we map the accessibility unreliability at each stop across the whole city of Columbus.

3.5 Long-term Disruption: the COVID-19 Pandemic

Since Jan 2020, the COVID-19 pandemic has had persistent and significant impacts on transit systems across the whole United States. For this case study, we choose the COVID-19 pandemic as an example of a long-term disruptive event, the city of Columbus as our study area, and COTA as our transit system of interest.

The city of Columbus reported its first three cases on March 9, 2020; local authorities declared the state of emergency on March 11, 2020, and enacted lockdown and curfew shortly after the date (Chow, 2020), which resulted in immediate decline of the ridership (Liu et al., 2020). The plunge in ridership also led to service cuts and schedule changes to adapt to staff shortage and economic difficulties (Van Niel, 2021). To investigate the distinctive impacts of different stages of the pandemic, we select all the Wednesdays during the period of March 2019 to January 2022. We first calculate the average realizable accessibility, i.e., the average number of accessible stops, and the accessibility reliability for each date.

To measure the robustness of the system with respect to COVID, we calculate the changing rate of realizable accessibility and the difference of accessibility unreliability between the first year of COVID (March 1, 2020 – March 1, 2021) and the year before COVID (March 1, 2019 – March 1, 2020). The two measures gauge the disruption's impacts on accessibility and unreliability, which represent the extent and quality of the public transit service respectively. We map the two measures for every stop in the city of Columbus and explore their spatial pattern.

4 Results

4.1 Short-term Disruption: Football games

We calculate accessibility and unreliability of every stop and every hour from 8:00 to 22:00 for every game day from 2018 to 2019.

Table 1: Average scheduled accessibility, realizable accessibility, and unreliability for each game day.

	Scheduled Accessibility (# accessible stops)	Realizable Accessibility (# accessible stops)	Unreliability
9/1/2018	225	153	47.79%
9/8/2018	225	158	42.23%
9/22/2018	224	157	42.28%
10/6/2018	224	151	47.93%
10/13/2018	224	162	37.99%
11/3/2018	225	159	42.10%
11/24/2018	225	167	34.83%
9/7/2019	204	139	47.22%
9/21/2019	194	133	45.36%
10/5/2019	214	145	47.41%
10/26/2019	187	131	42.91%
11/9/2019	188	136	37.73%
11/23/2019	203	144	41.00%

We aggregate all game days based on their start time; Figure 2 shows the hourly profile of the average accessibility unreliability. Higher unreliability means that the deviation of realizable accessibility from the scheduled accessibility is larger; for example, unreliability of 100% means that system users can only access half of the stops in the scheduled scenario. All game days, except the 19:30 game (discussed later), have two unreliability peaks before and after the game, which represent the traffic to and from the stadium respectively.

Therefore, we name the unreliability peak before game before-game peak and the unreliability peak after game after-game peak.

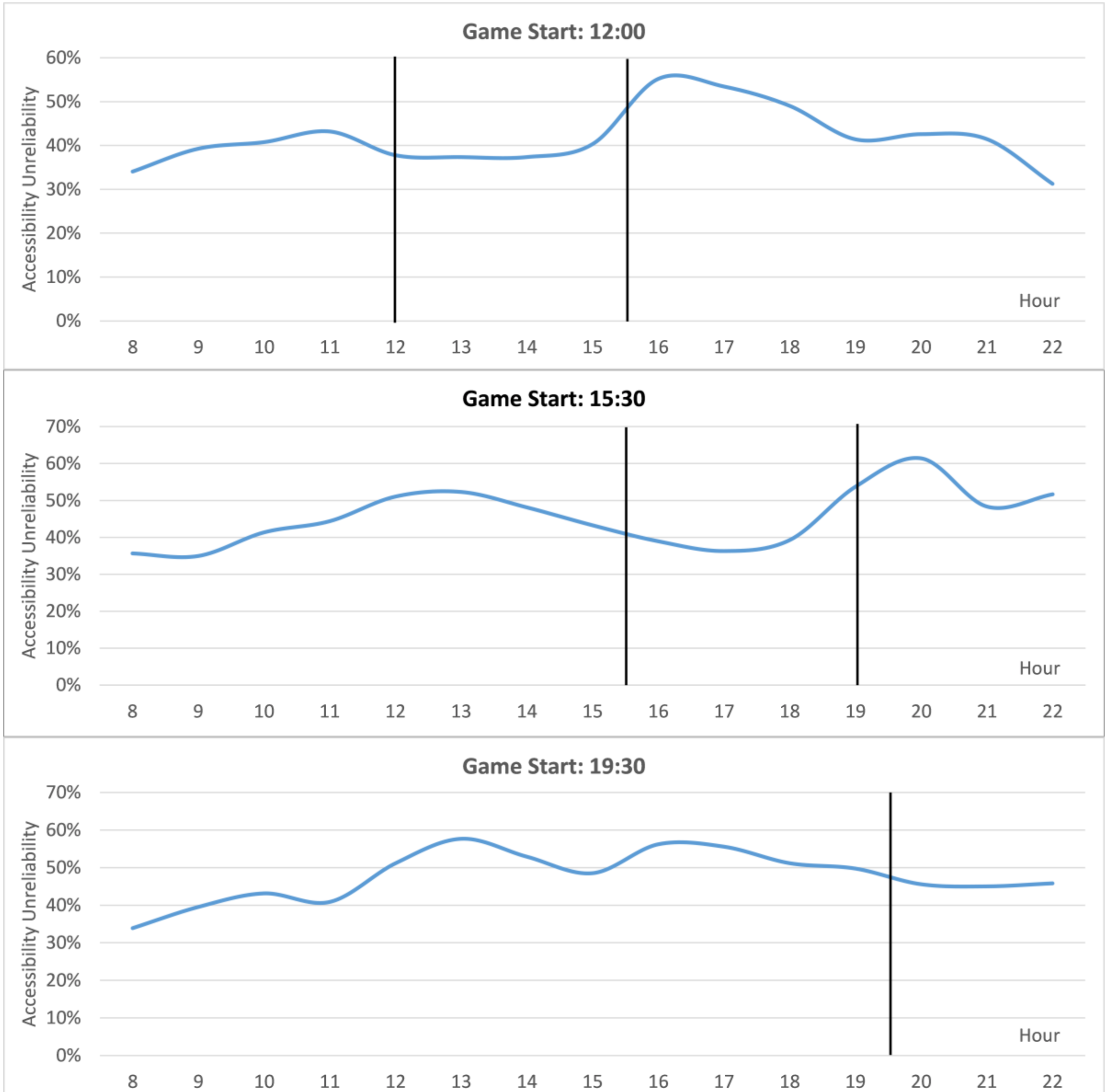

Figure 2: Average hourly profile of accessibility unreliability in three different game start time slots. Black line indicates game start and end time. Note that the 19:30 game ended at 23:03, which is outside the scale of the graph and normal operation hours.

Several phenomena prove that the peaks in game days are not random or caused by ordinary commuting traffic. Figure 3 visualizes the hour the relationship between the positions of the two peaks and the game start and end time. We can clearly witness that the positions of the peaks are shifting along with the changing game start time. The consistency strongly suggests that the peaks are caused by the football games. Figure 4, moreover, reaffirms this correlation between football games and unreliability. The graph shows the hourly profile of accessibility unreliability for home game, away game, and non-game days in the same time period. The unreliability in home game days is higher than away game days and non-game days, while away game days are higher than non-game days.

We also measure the two factors of resilience as we introduce in the background and method section, namely robustness and recoverability. In term of robustness, unreliability at the before-game peak is 8.7% higher than the average, while unreliability at the after-game peak is 25.4% higher than the average, showing the impact on transit service's reliability. In term of recoverability, the duration of football games' impact, i.e., the gap between before- and after-game peaks, is 6.8 hours. Note that this measure does not completely encompass all the affected period, as the traffic forms before the peak; instead, it reflects the most affected period and the core of the disruption.

We can also divide the whole period into three sections: 1) before-game gap, i.e., the gap between the before-game peak and the game start, 2) game duration, and 3) after-game gap, i.e., the gap between the game end and the after-game peak. The average before-game gap is 2.2 hours, and average after-game gap is 1.1 hours. Note that there is no after-game peak in Oct 5, 2019 from the third graph in Figure 2 and Figure 3, which started at 19:30. The game ended at 23:03, and given the average after-game gap, the after-game peak would have been at the midnight, which is outside the normal operating hours of COTA buses.

We can see the before-game impacts have longer duration but less disruptive effects, while after-game impacts have shorter duration but larger disruptive effects. This is likely due to people arriving at the event at different times, but leaving the event at the same time, creating a more intense but less extensive disruption.

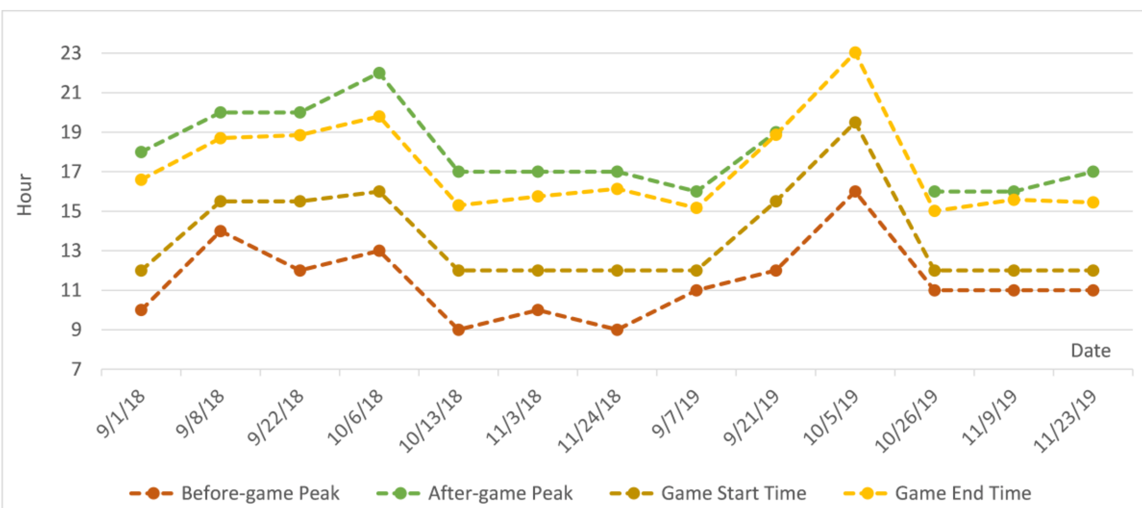

Figure 3: The relationship between positions of before-game peak, game start time, game end time, and after-game peak. Before-game peak's value represents the hour when the unreliability reaches a climax for the first time, and after-game peak's value represents the time when the unreliability reaches a climax for the second time.

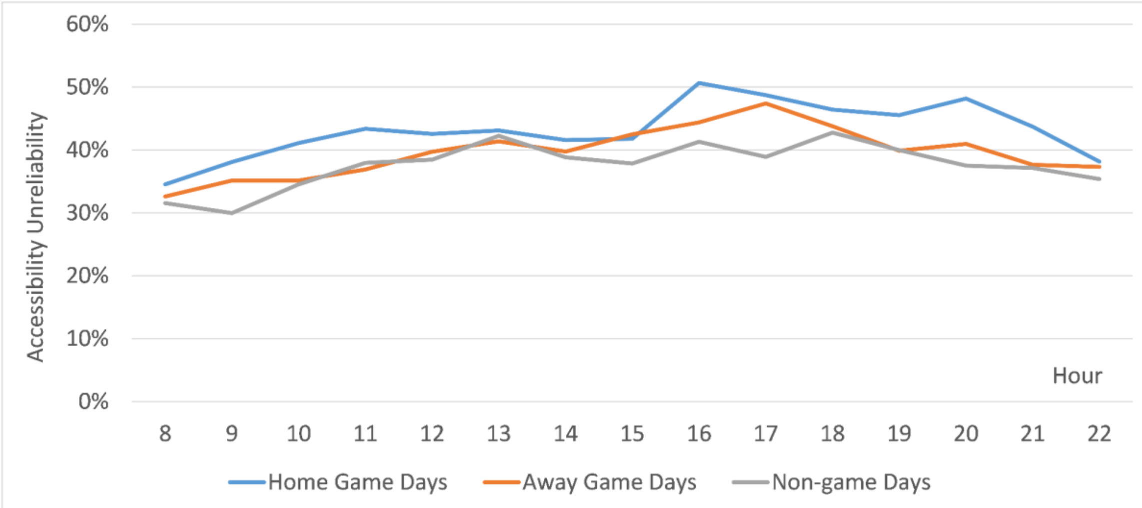

Figure 4: The average hourly profile of accessibility unreliability for home game, away game and non-game days in the same time period.

Figure 5 visualizes the spatial pattern of unreliable accessibility from the disruptive events. These maps show the unreliability value at each stop – i.e., the highest accessibility unreliability value during the game day – of the before-game and after-game peaks for all 9 games that started from 12:00. Public transit unreliability shows a strong clustering pattern. Both before-game and after-game peaks values are clustered around the Ohio Stadium, which is the main site of the football games. We also conduct same analysis for away game days and non-game days, and we find no high clusters around the stadium. This, together with the evidence we present above, strongly suggests the causality between football home games and high public transit unreliability.

Figure 5 also presents the before- and after-game gap at each stop. Stops near the stadium immediately reached the peak as soon as the game ends, while they reach the before-game peak later. This, again, reflects a shockwave-like pattern of football games. Before the game, as viewers and most traffic are coming to the site, the event's impacts would spread from the perimeter to the center; as soon as the football game ends, the impacts would spread from the center and reach neighboring stops first and spread to the perimeter. Note that the before-game gaps' pattern is much more heterogeneous than the after-game gaps. This is likely due to incoming traffic before the game being more diverse and dispersed, while outgoing traffic after the game being more concentrated and intense. This is also consistent with our findings on Figure 3 above.

Figure 5: Before-game and after-game peaks' unreliability value and gap

4.2 Long-term Disruption: COVID-19

COVID-19 has persistent negative impacts on public transit accessibility and accessibility reliability. Figure 6 (top) visualizes the temporal pattern of schedule-based accessibility and realizable accessibility; both significantly declined during the lockdown (March – June 2020) and remained lower than the pre-COVID level during the post-lockdown era. The decline of accessibility measures reflects the deterioration of transit service due to travel restriction and schedule change. Note that the rapid decline of accessibility is not perfectly synchronous with the start of the pandemic. The major schedule change made by transit authority, which aimed to adapt to the plunging ridership and financial difficulties, were enacted in May 2020, rather than immediately after the outbreak.

Meanwhile, unreliability is also impacted by the pandemic. As we introduced in the background section, long-term disruption can impact unreliability by affecting both the on-time performance and the schedule change. This means that these two factors can conflict with each other and produce nuanced patterns. Accessibility unreliability during the lockdown first declined and then increased as Figure 6 (bottom) shows. The decline could be because the lockdown eliminated most commuting travel and reduced roadway congestion (Lee et al., 2020), perhaps resulting in better on-time performance. Meanwhile, the schedule for the first few weeks remained unchanged, resulting in less unreliability. Following the service cut in May 2020, both accessibility measures rapidly declined but scheduled accessibility declined faster, resulting higher unreliability than usual. However, we do not observe major change in the global average of unreliability after the lockdown compared to pre-COVID conditions.

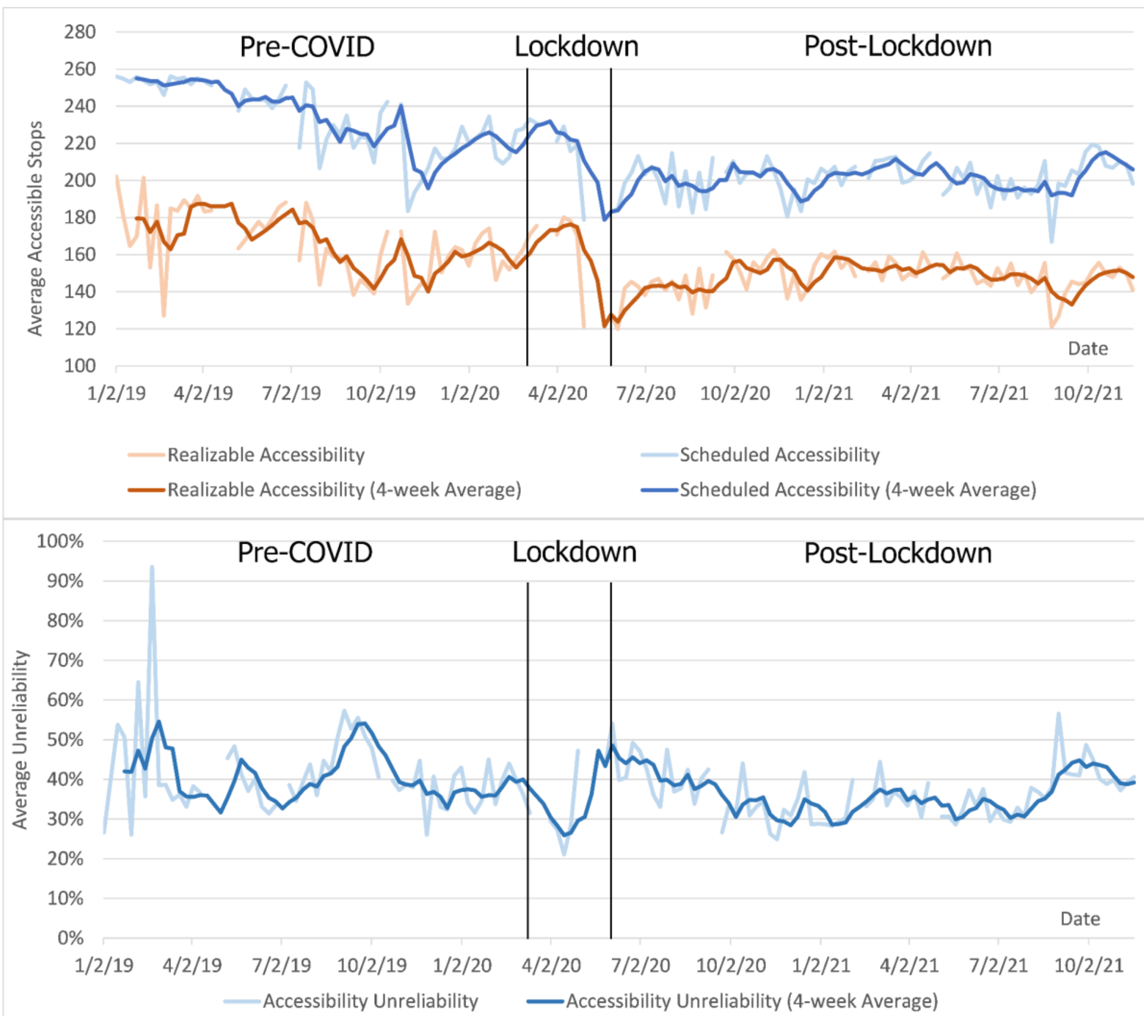

Figure 6: Temporal pattern of accessibility and unreliability

COVID-19's impacts on realizable accessibility are also spatially heterogeneous. Figure 7 shows the changing rate of realizable accessibility (left) and the difference of accessibility unreliability (right) between the year before and after the COVID-19 outbreak. The red color means more system performance loss, and the blue color means less performance loss. The downtown area, which accounts for most ridership in the system and experienced the fewest service cuts, has less accessibility and reliability loss. The decline of unreliability can also be explained by the reduction of general traffic. However, urban perimeters and suburban areas experienced more unreliability and more accessibility loss due to service cuts.

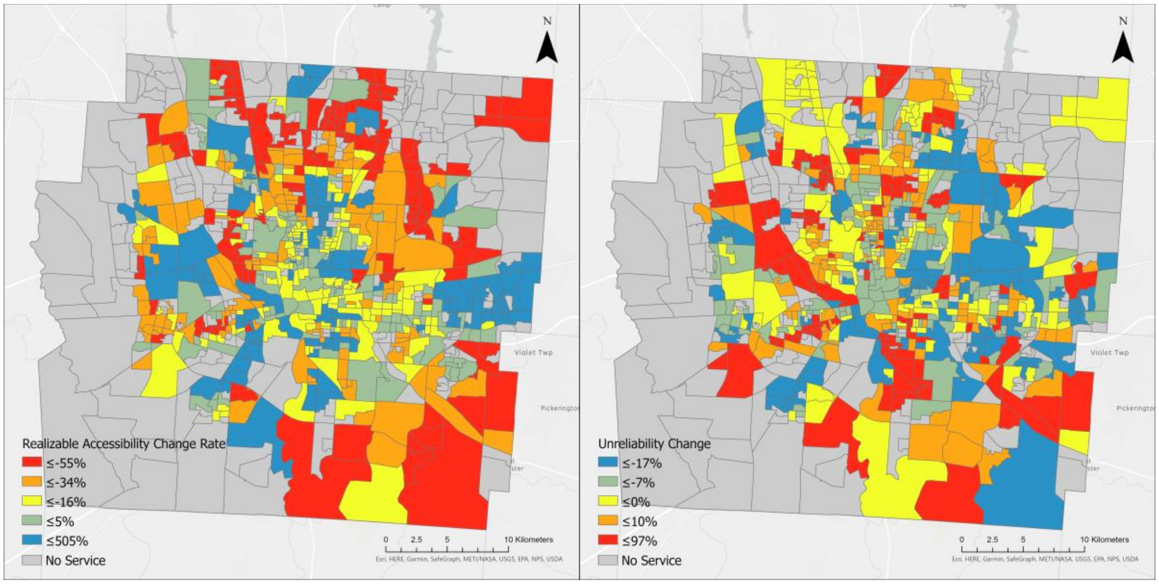

Figure 7: The change rate of realizable accessibility and unreliability after COVID-19. Red color indicates worse performance in both maps.

5 Conclusion

Public transit systems are facing higher risks of system degradation and failures caused by disruptions, such as climate change and pandemics. Despite myriad discussions on transit accessibility, reliability, and resilience, few papers integrate accessibility and its reliability into the study of resilience of public transit systems against disruptions. To fill in the gaps, we use two measures in this paper: realizable accessibility, which represents the accessibility that can be actually achieved by a transit user (Liu et al., 2022), and scheduled accessibility, which represents the expected useability of the transit system. Based on the two measures, we define accessibility reliability as the difference between delivered accessibility and scheduled accessibility to measure the variation of transit system's performance. Our paper provides a new way for future research and planning to understand a public transit system's resilience against different types of disruptions. The method uses the change in realizable accessibility and accessibility reliability before, during and after disruptive events as two measures of system resilience. We choose two examples, namely the Ohio State football games and the COVID-19 pandemic, to exemplify short-term and long-term disruption, respectively.

We find that the presence of football games is correlated with exceptional high unreliability in local public transit system. Days with Ohio State home game days have significantly higher unreliability than away game and non-game days, while days with away games have higher unreliability than non-game days. There were two peaks of unreliability before and after each game, the relationship between game times and peak times was consistent in all cases. Spatial analysis also shows that Ohio Stadium was the center of the high unreliability cluster, while other days did not show similar patterns. All evidence strongly suggests that the high unreliability was caused by the football games, rather than random fluctuations or daily commuting.

Analyses of COVID-19 and subsequent service cuts show that transit users experienced universal decline in realizable accessibility. However, improved traffic conditions caused by the lockdown during the early stages of the pandemic may have helped to temporarily reduce unreliability, but unreliability later increased again after the service cuts. The shrinking service schedule and improved traffic conditions, as two contradicting forces, created an intricate pattern of unreliability. Our spatial analysis also reveals that the city center, which has the most ridership and accessibility, experienced the least accessibility and reliability loss, while most of the urban perimeters and suburbs witnessed substantial decline in system performance and service quality.

The contribution of the proposed methods and results is threefold. First, both case studies show the effectiveness of realizable accessibility and accessibility unreliability to detect system disturbances, and their effectiveness for both measuring short- and long-term disruptions with high spatial and temporal resolution. We suggest that more public transit systems should use real-time accessibility and unreliability measure to monitor system performance and guide future system operation and planning. Second, the results regarding football games reveal patterns of large social events' impacts on public transit accessibility and reliability. First, our results show that the impacts of football games extend beyond the local vicinity; instead, more areas outside the OSU campus were affected, and there is a strong tendency of spatial correlation between unreliability and distance of football game site. Meanwhile, we also witness a strong formality of its temporal patterns, with the start and the end of each game accompanied by spikes of unreliability. In response to this, transit authorities can plan and broadcast rerouting in advance in acknowledgement of average before- and after-game peak hours, while preserving the normal schedule for other hours. Furthermore, it is noteworthy that the paper may provide useful insights for many other college towns and cities that hosts football games; considering Columbus is a more populous city, the effect of football games in other college towns or smaller cities may be lower. Meanwhile, for other cities without football games, the findings of this study can still provide important implications related to the unreliability of transit service affected by other similar major social events. Finally, major long-term disruptions, such as COVID-19, have effects on both on-time performance and schedule. Although the global unreliability of the system was not negatively impacted in the long run, many parts of the transit network suffered a permanent decrease in realizable accessibility. This shows unreliability alone cannot capture the whole picture of the transit experience. Instead, unreliability must be considered in tandem with other measures, such as accessibility, to understand system performance holistically.

Our paper also has limitations. First, despite strong indications of causality, we cannot make definite conclusions on the causality between high unreliability and the disruptive events; there can be other confounding factors. Second, we do not address the implications of the disruptions on social equity and individual experience. Third, our study does not address the heterogeneity of accessibility to different opportunities such as jobs, childcare, or parks; we only measure accessibility with the number of accessible stops. Our study is also limited to one community. Our methods should be applied to other communities with available real-time and schedule GTFS data. We would like to see more case studies on accessibility reliability in different settings in the future. Fourth, without behavioral studies, we do not know if the calculated travel time would faithfully reflect the accessibility experience of transit users. Finally, our stop-based analyses can be subject to modifiable area unit problem (MAUP) and produce unreliable results (Javanmard et al., 2023).

6 Declarations

6.1 Ethical Approval

Not applicable.

6.2 Authors' contributions

Luyu Liu, Adam Porr, and Harvey J. Miller designed the study and conception. Luyu Liu and Adam Porr collected the data. Luyu Liu conducted the analysis. Luyu Liu wrote the manuscript, and Adam Porr and Harvey J. Miller reviewed and revised the manuscript. Luyu Liu prepared figures 1 - 7. All authors reviewed the results and approved the final version of the manuscript.

6.3 Funding

The authors received no specific funding for this work. The authors declare no financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

6.4 Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon request.

7 Appendix

Table 2: Date, time, temperature, and precipitation of home game days in 2018 and 2019 (Weather Underground, 2023) at John Glenn Columbus International Airport Weather Station, Columbus, Ohio.

Date	Game Start Time	Average Temperature (°C)	Total Precipitation (cm)	Game Start Temperature (°C)	Game Start Precipitation (cm)
2018-09-01	12:00 PM	24.2	0.0	29.4	0.0
2018-09-08	3:30 PM	18.3	1.2	18.3	0.1
2018-09-22	3:30 PM	16.6	0.3	17.2	0.0
2018-10-06	4:00 PM	24.4	0.0	30.6	0.0
2018-10-13	12:00 PM	7.2	0.2	7.8	0.0
2018-11-03	12:00 PM	6.4	0.2	10.0	0.0
2018-11-24	12:00 PM	8.1	1.6	8.9	0.0
2019-08-31	12:00 PM	21.0	0.0	22.8	0.0
2019-09-07	12:00 PM	20.9	0.0	22.8	0.0
2019-09-21	3:30 PM	23.6	0.1	28.3	0.0
2019-10-05	7:30 PM	15.4	0.0	19.4	0.0
2019-10-26	12:00 PM	12.2	0.0	11.7	0.2
2019-11-09	12:00 PM	0.2	0.0	1.7	0.0
2019-11-23	12:00 PM	1.9	0.0	5.0	0.0

8 Reference

Ahn, B.-H., & Shin, J.-Y. (1991). Vehicle-routeing with time windows and time-varying congestion. *Journal of the Operational Research Society*, *42*(5), 393–400.

Antrim, A., & Barbeau, S. J. (2017). *Opening the door to multimodal applications: Creation, maintenance and application of GTFS data*.

Azolin, L. G., da Silva, A. N. R., & Pinto, N. (2020). Incorporating public transport in a methodology for assessing resilience in urban mobility. *Transportation Research Part d: Transport and Environment*, *85*, 102386.

Beck, M. J., Hensher, D. A., & Wei, E. (2020). Slowly coming out of COVID-19 restrictions in Australia: Implications for working from home and commuting trips by car and public transport. *Journal of Transport Geography*, *88*, 102846.

Berche, B., Von Ferber, C., Holovatch, T., & Holovatch, Y. (2009). Resilience of public transport networks against attacks. *The European Physical Journal B*, *71*(1), 125–137.

Carrion, C., & Levinson, D. (2012). Value of travel time reliability: A review of current evidence. *Transportation Research Part A: Policy and Practice*, *46*(4), 720–741.

Chakrabarti, S., & Giuliano, G. (2015). Does service reliability determine transit patronage? Insights from the Los Angeles Metro bus system. *Transport Policy*, *42*, 12–20.

Chow, A. (2020, March 9). *Ohio Confirms First Cases Of Coronavirus*. The Statehouse News Bureau. https://www.statenews.org/government-politics/2020-03-09/ohio-confirms-first-cases-of-coronavirus

D’este, G. and, & Taylor, M. A. (2003). Network vulnerability: An approach to reliability analysis at the level of national strategic transport networks. In *The network reliability of transport*. Emerald Group Publishing Limited.

Erhardt, G. D., Hoque, J. M., Goyal, V., Berrebi, S., Brakewood, C., & Watkins, K. E. (2022). Why has public transit ridership declined in the United States? *Transportation Research Part A: Policy and Practice, 161*, 68–87.

Gendreau, M., Ghiani, G., & Guerriero, E. (2015). Time-dependent routing problems: A review. *Computers & Operations Research*, *64*, 189–197.

Google. (2021). *GTFS Realtime Overview*. https://developers.google.com/transit/gtfs-realtime

Google Developers. (2020). *GTFS Static Overview | Static Transit | Google Developers*. https://developers.google.com/transit/gtfs/

Gu, Y., Fu, X., Liu, Z., Xu, X., & Chen, A. (2020). Performance of transportation network under perturbations: Reliability, vulnerability, and resilience. *Transportation Research Part E: Logistics and Transportation Review*, *133*, 101809.

Hägerstrand, T. (1970). *What about people in regional*.

He, Y., Thies, S., Avner, P., & Rentschler, J. (2021). Flood impacts on urban transit and accessibility—A case study of Kinshasa. *Transportation Research Part D: Transport and Environment*, *96*, 102889.

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 1–23.

Ichoua, S., Gendreau, M., & Potvin, J.-Y. (2003). Vehicle dispatching with time-dependent travel times. *European Journal of Operational Research*, *144*(2), 379–396.

Javanmard, R., Lee, J., Kim, J., Liu, L., & Diab, E. (2023). The impacts of the modifiable areal unit problem (MAUP) on social equity analysis of public transit reliability. *Journal of Transport Geography*, *106*, 103500.

Kar, A., Carrel, A. L., Miller, H. J., & Le, H. T. (2021). *Reducing public transit compounds social vulnerabilities during COVID-19*.

Kathuria, A., Parida, M., & Sekhar, C. (2020). A review of service reliability measures for public transportation systems. *International Journal of Intelligent Transportation Systems Research*, *18*(2), 243–255.

Kaufman, J. (2021). *Ohio State football draws crowd of only 76,540 in win over Tulsa, smallest since 1971*. The Columbus Dispatch. https://www.dispatch.com/story/sports/2021/09/18/ohio-stadium-crowd-osu-ohio-state-tulsa-football-game-76-540/8407755002/

Kim, J., & Kwan, M.-P. (2021). The impact of the COVID-19 pandemic on people's mobility: A longitudinal study of the US from March to September of 2020. *Journal of Transport Geography*, *93*, 103039.

Lee, J., Porr, A., & Miller, H. J. (2020). *Evidence of Increased Vehicle Speeding in Ohio's Major Cities during the COVID-19 Pandemic*.

Li, M., Kwan, M.-P., Yin, J., Yu, D., & Wang, J. (2018). The potential effect of a 100-year pluvial flood event on metro accessibility and ridership: A case study of central Shanghai, China. *Applied Geography*, *100*, 21–29.

Lin, T., Shalaby, A., & Miller, E. (2016). *Transit user behaviour in response to service disruption: State of knowledge*.

Liu, L., & Miller, H. J. (2020a). Does real-time transit information reduce waiting time? An empirical analysis. *Transportation Research Part A: Policy and Practice*, *141*, 167–179.

Liu, L., & Miller, H. J. (2020b). Measuring risk of missing transfers in public transit systems using high-resolution schedule and real-time bus location data. *Urban Studies*, 0042098020919323. https://doi.org/10.1177/0042098020919323

Liu, L., Miller, H. J., & Scheff, J. (2020). The impacts of COVID-19 pandemic on public transit demand in the United States. *Plos One*, *15*(11), e0242476.

Liu, L., Porr, A., & Miller, H. J. (2022). Realizable accessibility: Evaluating the reliability of public transit accessibility using high-resolution real-time data. *Journal of Geographical Systems*. https://doi.org/10.1007/s10109-022-00382-w

Mesbah, M., Luy, M., & Currie, G. (2014). Investigating the lagged effect of weather parameters on travel time reliability. *WIT Transactions on Ecology and the Environment*, *191*, 795–801.

Miller, E. J. (2018). Accessibility: Measurement and application in transportation planning. In *Transport Reviews* (Vol. 38, Issue 5, pp. 551–555). Taylor & Francis. https://doi.org/10.1080/01441647.2018.1492778

Miller, H. J. (1999). Measuring space-time accessibility benefits within transportation networks: Basic theory and computational procedures. *Geographical Analysis*, *31*(1), 187–212.

Miller, H. J. (2017). Time geography and space-time prism. *International Encyclopedia of Geography: People, the Earth, Environment and Technology*, 1–19.

Mudigonda, S., Ozbay, K., & Bartin, B. (2019). Evaluating the resilience and recovery of public transit system using big data: Case study from New Jersey. *Journal of Transportation Safety & Security*, *11*(5), 491–519.

Park, Y., Mount, J., Liu, L., Xiao, N., & Miller, H. J. (2020). Assessing public transit performance using real-time data: Spatiotemporal patterns of bus operation delays in Columbus, Ohio, USA. *International Journal of Geographical Information Science*, *34*(2), 367–392. https://doi.org/10.1080/13658816.2019.1608997

Pender, B., Currie, G., Delbosc, A., & Shiwakoti, N. (2014). Social media use during unplanned transit network disruptions: A review of literature. *Transport Reviews*, *34*(4), 501–521.

Pereira, R. H. M. (2019). Future accessibility impacts of transport policy scenarios: Equity and sensitivity to travel time thresholds for Bus Rapid Transit expansion in Rio de Janeiro. *Journal of Transport Geography*, *74*, 321–332. https://doi.org/10.1016/J.JTRANGEO.2018.12.005

Singh, S. S., Javanmard, R., Lee, J., Kim, J., & Diab, E. (2022). Evaluating the accessibility benefits of the new BRT system during the COVID-19 pandemic in Winnipeg, Canada. *Journal of Urban Mobility*, *2*, 100016.

Taylor, M. A., & D'Este, G. M. (2007). Transport network vulnerability: A method for diagnosis of critical locations in transport infrastructure systems. In *Critical infrastructure* (pp. 9–30). Springer.

Tong, L., Zhou, X., & Miller, H. J. (2015). Transportation network design for maximizing space–time accessibility. *Transportation Research Part B: Methodological*, *81*, 555–576.

Van Niel, L. (2021, September 29). *COTA faces driver shortages, adjusts services*. The Lantern. https://www.thelantern.com/2021/09/cota-faces-driver-shortages-adjusts-services/

Wan, C., Yang, Z., Zhang, D., Yan, X., & Fan, S. (2018). Resilience in transportation systems: A systematic review and future directions. *Transport Reviews*, *38*(4), 479–498.

Wang, Y., Yuan, Y., Ma, Y., & Wang, G. (2019). Time-dependent graphs: Definitions, applications, and algorithms. *Data Science and Engineering*, *4*(4), 352–366.

Weather Underground. (2023). *Columbus, OH Weather History | Weather Underground*. https://www.wunderground.com/history/daily/us/oh/columbus/KCMH

Wessel, N., Allen, J., & Farber, S. (2017). Constructing a routable retrospective transit timetable from a real-time vehicle location feed and GTFS. *Journal of Transport Geography*, *62*, 92–97.

Wessel, N., & Farber, S. (2019). On the accuracy of schedule-based GTFS for measuring accessibility. *Journal of Transport and Land Use*, *12*(1), 475–500.

Zhu, S., & Levinson, D. M. (2012). Disruptions to transportation networks: A review. *Network Reliability in Practice*, 5–20.