TRB Annual Meeting

Impact of Unplanned Rail Disruption on Urban Transit Systems

--Manuscript Draft--

Full Title:	Impact of Unplanned Rail Disruption on Urban Transit Systems
Abstract:	This study proposes a general incident analysis framework from supply and demand sides using automated fare collection (AFC) and automated vehicle location (AVL) data in public transit systems. Specifically, on the supply side, we propose an incident-based network redundancy index to analyze the network's ability to provide alternative services under a specific rail disruption. The impacts on the operation are analyzed through the headway changes in multiple rail lines. On the demand side, we calculate the demand changes of different rail lines, rail stations, bus routes, and bus stops to show the passenger flow redistribution under incidents. Individual behavior is analyzed using a binary logit model based on inferred passengers' mode choices and socio-demographics using AFC data. The public transit system of Chicago Transit Authority is used as a case study. Two rail disruption cases are analyzed, one of which has high network redundancy and low for another. Results show that the service frequency of the incident line was largely reduced during the incident time. Depending on the incident location, the network's redundancies are different, as well as the passengers' behavior. In the low redundancy scenario, most passengers choose to use nearby buses to move, either to their destinations or to the nearby rail lines. In the high redundancy scenario, most of the passengers transferred to nearby lines.
Manuscript Classifications:	Public Transportation; Rail Transit Systems AP065; Equipment (Cars, Fare Collections, Signage); Management; Subway; Sustainability and Resilience; Transportation Systems Resilience AMR00; Disaster Response, Recovery, and Business Continuity AMR20; Emergency Response; Incident Management; Traveler Behavior
Manuscript Number:	TRBAM-21-03587
Article Type:	Presentation
Order of Authors:	Baichuan Mo
	Max Y von Franque
	Haris N. Koutsopoulos
	John P. Attanucci
	Jinhua Zhao

Impact of Unplanned Rail Disruption on Urban Transit Systems

Baichuan Mo (Corresponding Author)
Research Assistant
Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 20139
Tel: +1 857-999-5906; Email: baichuan@mit.edu

Max Y von Franque
Department of Urban Studies and Planning
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 20139
Tel: +1 415-847-4877; Email: maxvf@mit.edu

Haris N. Koutsopoulos, Ph.D.
Professor
Department of Civil and Environmental Engineering
Northeastern University
Boston, MA 02115, United States
Tel: +1 617-373-6263; Email: h.koutsopoulos@northeastern.edu

John P. Attanucci
Research Associate
Center for Transportation & Logistics
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139
Email: jattan@mit.edu

Jinhua Zhao, Ph.D.
Associate Professor
Department of Urban Studies and Planning
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 20139
Tel: +1 617-253-7594; Email: jinhua@mit.edu

Word count:6,723 words text + 3 table $\times$ 250 words (each) = 7,473 words

Submission Date: August 2, 2020

ABSTRACT
This study proposes a general incident analysis framework from supply and demand sides using automated fare collection (AFC) and automated vehicle location (AVL) data in public transit systems. Specifically, on the supply side, we propose an incident-based network redundancy index to analyze the network's ability to provide alternative services under a specific rail disruption. The impacts on the operation are analyzed through the headway changes in multiple rail lines. On the demand side, we calculate the demand changes of different rail lines, rail stations, bus routes, and bus stops to show the passenger flow redistribution under incidents. Individual behavior is analyzed using a binary logit model based on inferred passengers' mode choices and socio-demographics using AFC data. The public transit system of Chicago Transit Authority is used as a case study. Two rail disruption cases are analyzed, one of which has high network redundancy and low for another. Results show that the service frequency of the incident line was largely reduced during the incident time. Depending on the incident location, the network's redundancies are different, as well as the passengers' behavior. In the low redundancy scenario, most passengers choose to use nearby buses to move, either to their destinations or to the nearby rail lines. In the high redundancy scenario, most of the passengers transferred to nearby lines.

Keywords: Incident analysis; Rail disruptions; Smart card data;

INTRODUCTION

Urban transit systems play a crucial role in cities worldwide, transporting people to jobs, homes, outings, and a vast variety of other activities. Millions rely on urban transit systems to provide them with transportation. However, no system is perfect, and all are susceptible to unplanned delays and service disruptions, whether that comes from equipment, weather, passengers, or other internal and external factors.

Mitigating the impact of unplanned service disruptions is an important task for urban transit agencies. For this reason, it is important to recognize how a transit system is affected by service disruptions. The main analysis framework for incident impacts can be summarized in Table 1. The two main dimensions of analysis, supply and demand, can further be broken down into network property and operations for supply analysis and passenger flow and individual behavior for demand analysis. The network property analysis applies graph theory-based techniques to calculate indicators related to incidents (*1–4*), such as network resilience, vulnerability, redundancy, and a variety of other factors. The operations analysis focuses on changes in an agency's operation during the incident period (*5–7*), including headway, routing, staffing, and other operator-controlled factors. From the demand aspect, passenger flow analysis investigates the demand changes of different stations, lines, or regions of a network, presenting passenger's choices and flow re-distribution after service disruptions. The individual behavior analysis focuses on studying the individual's response behavior (such as mode choices, waiting time tolerance) to the incident (*8, 9*) and its relationship to individual's characteristics (e.g., travel histories, demographics). Surveys are usually used for such studies.

TABLE 1: Analysis framework for incident impacts

	Analysis tasks	Description
Supply	Network properties	Calculate indicators such as resilience, vulnerability, redundancy
	Operations	Analyze changes of agency's operations (e.g., headway, routing)
Demand	Passenger flows	Analyze demand changes of different stations, lines
	Individual behavior	Analyze people's response behavior under incidents

Previous research has used a variety of methods to analyze the impact of service disruptions. Of these methods, the three most notable are graph theory-based, survey-based, and simulation-based. Graph theory-based methods usually derive resilience or vulnerability indicators based on the network topology (*1–4*). These methods are effective for understanding high-level network properties related to incidents. Survey-based methods investigate passengers about their behavior and opinions during the incident period (*10–14*). They can analyze and understand passengers' individual-level behavior in-depth using econometric models. Simulation-based methods simulate passenger flows on the transit network under certain hypothetical incident scenarios (*15–17*). These studies can output many metrics of interest such as the vehicle loads changes and additional travel delays caused by incidents.

Recently, automated data collection systems in transit networks enable data-driven analysis on service disruptions. The two major sources are automatic fare collection (AFC) and automatic vehicle location (AVL) data. AFC data are collected when passengers tap their transit cards on smart card readers (in buses or rail station gates). The records include time, card ID, and location of the tap in. Depending on whether the fare system requires passengers to tap out, AFC data may only include tap-in records or both tap-in and tap-out records. AVL data record vehicle's (bus and train) time-dependent locations based on GPS and train tracking systems, from which the headway can be extracted. Multiple studies have been conducted using AFC and AVL data to look at transit disruptions (*5–7*). For example, Mojica (*5*) used AFC data to analyze changes in the travel behavior of a sample of rail commuters during a large scale maintenance project. Sun et al. (*7*) analyzed three types of abnormal passenger flows during rail disruptions using AFC data with both tap-in and tap-out records. Tian and Zheng (*6*) proposed a classification model to predict whether commuters switch from rail to other transportation modes because of unexpected travel delays using six months of AFC data.

However, despite rich studies on incident analysis, there are still research gaps. First, for the graph theory-based approaches, the network indicators such as redundancy are usually defined for the whole network, an OD pair, or a link. Incidents usually cause service interruptions for multiple links depending on the power system and rail track structure. An incident-based indicator should be proposed to reflect the network's redundancy under an actual incident. Second, for the individual-behavior analysis methods, rarely has previous studies leveraged the AFC data to analyze passengers' mode choices. This problem needs to extract both individual choices and socio-demographic information from the AFC data. Third, most of the previous studies on incident analysis only addressed one or two tasks in Table 1 with a single incident case study. A comprehensive study that analyzes all four tasks with multiple comparable case studies using AFC and AVL data is needed.

To fill the research gaps, this paper conducts a comprehensive analysis of the impact of unplanned rail disruptions on urban rail systems. Specifically, on the supply side, we propose an incident-based network redundancy index to analyze the ability of bus and rail networks to provide alternative services under a specific rail disruption. The impacts on operations are analyzed through the headway changes in multiple rail lines. On the demand side, we calculate the demand changes of different rail lines, rail stations, bus routes, and bus stops to show the passenger flow redistribution under incidents. Individual behavior is analyzed using a binary logit model based on inferred passengers' mode choices and socio-demographics using AFC data. The public transit system of the Chicago Transit Authority (CTA) is used as the case study. Specifically, two rail disruption cases are analyzed, one of which has high network redundancy and low for another.

The contribution of this paper is threefold:

- Propose an incident-based network redundancy index to reflect the ability to provide alternative services for the bus and rail system as a whole.
- Conduct a comprehensive incident analysis with two different types of incidents following a proposed incident analysis framework using AFC and AVL data.
- Propose an individual mode choice analysis method using AFC data.

The remainder of this paper is organized as followings. Section 2 reviews the literature. Section 3 elaborates the methodology used in this study. Case studies and data are described in Section 4 and results are shown in Section 5. Section 6 concludes the paper and discusses the policy implications.

LITERATURE REVIEW

There are generally four methods researchers use to analyze the impact of disrupted operations, including graph theory-based, survey-based, simulation-based, and AFC data-based methods. Each method has strengths and weaknesses depending on the context.

Graph theory-based

Graph theory-based analysis identifies key aspects of the network's properties and functions related to incidents (such as resilience, redundancy, vulnerability) based on graph theory (or complex network theory). These properties and functions show high-level topological characteristics of the network, providing intuitive conclusions for incident management. For example, Yin et al. (*1*), using the graph theory-based method, studied subway networks with respect to static disruptions, finding the weakness or key positions of the network through "network betweenness" and "global efficiency" calculation. Similarly, (*2*) built a general framework to assess the resilience of large and complex metro networks by quantitatively analyzing its vulnerability and recovery rapidity with proposed graph theory-based definitions.

Redundancy is one of the important indicators for analyzing the network's properties related to incidents. It was defined as "the extent to which elements, systems, or other units of analysis exist that are substitutable, i.e., capable of satisfying functional requirements in the event of a disruption, degradation, or loss of function". Redundancy has been widely studied, not just in transportation analysis. These areas are not limited to but include reliability engineering (*18*), communications (*19*), water distribution systems (*20*), and supply chain and logistics (*21*). Transportation specific resiliency and redundancy studies include (*4*), who developed a qualitative framework and basic concepts for vulnerability, resilience, and redundancy. Other studies, like Wilson-Goure et al. (*22*), Murray-Tuite (*23*), and Goodchild et al. (*24*), each defined redundancy in the context of a specific transportation area of study. However, nearly all previous studies defined redundancy for networks, links, and OD pairs. This study proposed an incident-based redundancy index to evaluate the network's ability to satisfy functional requirements under a specific incident. Moreover, the bus system, which is an important alternative for rail but rarely considered in previous studies, is included for analyzing rail disruptions.

Survey-based

Many studies utilized survey-based methods to examine incident impacts. Surveys, including revealed preference (RP) and stated preference (SP), are a good medium to understand individual choices and demand. Surveys can be divided into two categories.

Examples of transit-oriented RP studies include Currie and Muir (*10*), who conducted an RP survey to understand rail passengers' behavior, perceptions, and priorities in response to unplanned urban rail disruption in Melbourne, Australia. Murray-Tuite et al. (*11*) used a web-based RP survey to understand the long term impacts of a deadly metro rail collision in Washington DC. Tsuchiya et al. (*25*) conducted an RP survey in Japan that looked at passenger choices of four alternative routes. Pnevmatikou and Karlaftis (*26*) used RP survey data to analyze the effect of a pre-announced closure of an Athens Metro Line. SP survey studies include Kamaruddin et al. (*27*), who studied the modal shift behavior of rail users after incidents. Fukasawa et al. (*12*) investigated the effect of providing information such as estimated arrival time, arrival order, and congestion level on passengers' modal shift behavior in response to an unplanned transit disruption. This was contrasted with another SP survey conducted by Bai and Kattan (*28*), which found that light rail transit passengers in Calgary, Canada had a higher willingness to switch travel modes given less information. The major drawback of survey-based methods is time-consuming and labor-inefficient. Hence, it is important to develop individual behavior analysis methods using AFC data as alternatives.

Simulation-based

Simulation-based studies evaluate the impact of the incident by running simulation models. Usually, different hypothetical incident scenarios are tested. The system performance metrics such as travel delays and vehicle loads are output to analyze the incident effect. For example, Balakrishna et al. (*15*) used a simulation-based framework to model transportation network performance under emergency conditions. Traffic densities and travel times are compared for different emergency scenarios. Suarez et al. (*16*) looked at the effects of climate change on Boston's urban transportation using a simulation model, suggesting almost a doubling in delays and lost trips. Another example is Hong et al. (*17*), who simulated passenger flows in a metro station during an emergency.

Automatic fare collection data-based

AFC data-based methods are powerful in examining the impact of actual incidents. While relatively new, these methods have been gaining traction in the previous decade. Mojica (*5*) used AFC data to analyze the impact of station closures and service reductions. Using a case in Chicago, they concluded that 79% of all riders continued to use rails, 8.4% switched to buses, and 4% percent switched to non-transit methods. Tian and Zheng (*6*) looked at unexpected train delay effects on Singapore's MTR customers. Using AFC data, researchers built a classification model to predict whether commuters switch from MRT to other transportation modes because of unexpected train delays. Sun et al. (*7*) used AFC data to quantify three choices passengers could make in the face of delay: leave the system, take a detour, or continue the journey but be delayed. This model was applied to the Beijing metro network.

Though there are existing studies on AFC data, they usually only addressed one or two tasks in Table 1, and majorly in the demand side. This study aims to conduct a comprehensive analysis of both the demand and supply sides using AFC data.

METHODOLOGY

In this section, we illustrate the analysis framework for an unplanned incident. On the supply-side, a method to calculate the network redundancy index under a certain incident is proposed, which reflects the network's ability to provide alternative routes when incidents occur. To analyze the agency's operations, we specify how we calculate the headway distribution using AVL data. On the demand-side, we describe how to analyze passenger flows under incidents using AFC data, and how to use AFC data to analyze passengers' mode choice using a binary logit model.

To extract the effect of the incident, we need to compare the information on the incident day and normal days. A "normal day" is defined as the recent same weekdays without incidents happening in the incident period. The headway and passenger flows in the incident day are compared to those of normal days to reveal the counterfactual.

Supply analysis

Network redundancy under incidents

As mentioned in Section 2, since redundancy is used to evaluate network's functional responses in the event of disruptions, it is important to develop an incident-specific redundancy index (opposite to network, link or OD pair-specific in the literature), that is, for a given incident, evaluating the network's ability to provide alternative services under this incident. Besides, given the substitutional relationship between bus and urban rail systems, the redundancy index in this study also explicitly considers the complementary role of bus and rail systems during the incident.

Redundancy is usually evaluated by the number of available paths for each OD pair because more available paths correspond to more opportunities of realizing the impacted trips when encountering service disruptions (*3*). Hence, the network redundancy under incidents (NRUI) should capture the transport capacity of alternative paths after the incident. The typical path capacity is defined as the max number of passengers transported per hour (i.e. service frequency times the vehicle capacity). It is a time-insensitive value, which means the travel times of paths are not considered. However, for the redundancy calculation, the path travel times matter because passengers may not be successfully transported to the destination during the incident period if they choose paths with long travel time. This means the time-insensitive path capacity cannot reflect the actual ability of paths to move passengers. Hence, a time-sensitive path capacity should be used for the calculation. Moreover, since the NRUI is incident-specific, the impact of the incident on the network capacity should be captured (i.e. for different incidents, the NRUI should be different depending on the conditions of the incident). In this study, we propose a normalized NRUI to meet the above requirements.

Let $\mathcal{W}$ be the set of all OD pairs of the rail network. For an OD pair $w \in \mathcal{W}$, let $\mathcal{P}_w$ be the set of available paths for *w before the incident*. As we consider both bus and urban rail systems, a path $p \in \mathcal{P}_w$ may include segments of bus trips. $\mathcal{P}_w$ can be obtained in several ways, such as route choice surveys, Google Map API, and k-shortest paths. Let A_p be the time-sensitive path capacity of p under incident I, which is defined as

$$A_p = \sum_{k=1}^{\lfloor D_I/H_p \rfloor} \frac{\min\{D_I - (k-1)H_p, L_p\}}{L_p} \cdot C_p \tag{1}$$

where D_I is the duration of incident I. H_p is the headway of path p (defined as the maximum headway of each segment of path p). C_p is the vehicle (i.e. train or bus) capacity of path p (defined as the minimum vehicle capacity of each segment of path p). L_p is the travel time of path p. $\lfloor D_I/H_p \rfloor$ is the number of vehicles dispatched for path p during the incident period. The first term in Eq. 1 is the proportion of k-th trip which finished during the incident period. Hence, Eq. 1 captures the number of people successfully transported to the destination along path p (passengers who did no finished the trip are counted proportionally) during the incident.

A_p is an indicator reflecting p's ability to serve impacted trips during the incident period and assigns different weights to different paths. Therefore, $\sum_{p\in\mathcal{P}_w} A_p$ reflects the ability of the network to provide services for OD pair w. Actually, one of the network-level redundancy in the literature is defined as $\sum_{w\in\mathcal{W}} \sum_{p\in\mathcal{P}_w} A_p$ (where A_p is defined differently), which means the total capacity of paths in the network (*29*).

In this study, we need to capture the incident-specific property of the redundancy metric. Let $\mathcal{W}_I$ be the set of all OD pairs with at least one path blocked due to incident I. Mathematically, $\mathcal{W}_I = \{\ w \in \mathcal{W} : \exists p \in \mathcal{P}_w \text{ s.t. } p \text{ is blocked due to incident } I\ \}$. Then, only passengers with OD in $\mathcal{W}_I$ are affected by the incident. As the incident may augment passengers' path choice alternatives, some paths that are not chosen before the incident can be available during the incident. Hence, we define $\tilde{\mathcal{P}}_w$ as the set of available paths for $w \in \mathcal{W}_I$ *during the incident*. $\tilde{\mathcal{P}}_w$ can be seen as $\mathcal{P}_w$ without the blocked paths and adding the augmented paths. To make these two sets comparable, we assume $|\mathcal{P}_w| \geq |\tilde{\mathcal{P}}_w|$, which means the number available paths after incident is less than or equal to those before. The NRUI for incident I is formulated as:

$$R_I = \frac{\sum_{w \in \mathcal{W}_I} \sum_{p \in \tilde{\mathcal{P}}_w} A_p}{\sum_{w \in \mathcal{W}_I} \sum_{p \in \mathcal{P}_w} A_p} \tag{2}$$

where the numerator (denominator) is the total capacity of available paths after (before) the incident. Since the path sets after the incident ($\tilde{\mathcal{P}}_w$) are longer and less preferred than the those before the incident ($\mathcal{P}_w$). We usually have $\sum_{p \in \mathcal{P}_w} A_p \geq \sum_{p \in \tilde{\mathcal{P}}_w} A_p$ for $w \in \mathcal{W}_I$[1]. Therefore, $0 \leq R_I \leq 1$ by definition. $R_I = 1$ means the path capacity before and after the incident are exactly the same, suggesting that the incident does not harm the function of the network (i.e. the network is fully redundant under incident I). $R_I = 0$ means no alternative are paths available after the incident (i.e. the network is has no redundant under incident I)

Headway analysis

Headway is an important indicator of the level of service for transit systems. Analyzing headway patterns during an incident can provide straightforward information about how services are reduced by the incident. As mentioned in Section 1, AVL data provide the headway of each platform in the urban rail system. In this study, we calculate the headway temporal distribution for lines of interest to show the impact of incidents.

Let us divide the analysis time into several intervals with equal length. Denote the total number of time intervals as T, the headway on platform i of trip j as $H_{i,j}$. Suppose line l has two directions, inbound and outbound. The headway of line l *outbound* at time interval τ is calculated as

$$H^{\text{out}}_{l,\tau} = \frac{\sum_{i \in \mathcal{S}_l^{\text{out}}} \sum_{j \in \mathcal{R}_{i,\tau}} H_{i,j}}{\sum_{i \in \mathcal{S}_l^{\text{out}}} |\mathcal{R}_{i,\tau}|} \tag{3}$$

where $\mathcal{S}_l^{\text{out}}$ is set of platforms in the outbound direction of line l. $\mathcal{R}_{i,\tau}$ is the set of trips passing through platform i during time interval τ. Eq. 3 means the headway of a line is calculated as the mean of all platforms. The inbound headway, $H^{\text{in}}_{l,\tau}$, is calculated in the similar way by replacing $\mathcal{S}_l^{\text{out}}$ with $\mathcal{S}_l^{\text{in}}$. The headway distributions of both normal days and the incident day are calculated for comparison.

Demand analysis

Passenger flow analysis

AFC data record passengers' tap-in information in bus and rail systems (tap-out is not available in this study), which reflect passenger's route choices after the incident if they use the transit system again. Therefore, analyzing AFC data can help to understand the passenger flow re-distribution after the incident.

At the station level, we can count the number of tap-in passengers for the stations near the incident area, and compare the demand between normal days and incident days. Stations with high demand increase reflect passengers' choices after the incident. At the line-level, we can calculate the number of tap-in passengers for lines near the incident area. Line-level demands are calculated

[1] Sometimes there are some exceptions where $\sum_{p \in \mathcal{P}_w} A_p < \sum_{p \in \tilde{\mathcal{P}}_w} A_p$ due to the augmented paths have high vehicle capacity and low headway. We manually let $\sum_{p \in \mathcal{P}_w} A_p = \sum_{p \in \tilde{\mathcal{P}}_w} A_p$ in this case

as the sum of all station-level demands. Demand at stations belonging to multiple lines will be divided uniformly.

Individual behavior analysis

Passengers may have different mode choices after the incident. One important question is how the characteristics of the passengers related to their mode choices. This is typically studied by the survey. In this study, we propose a method of using AFC data for individual behavior analysis.

The first step is to infer individual choices using AFC data. In this study, only two choices are considered: 1) using transit and 2) not using transit (including canceling trips and using other travel modes). This is because these two options can be confidently identified using AFC data and they are important for transit operators. Due to the high travel irregularity of passengers in transit systems (*30*), it is more convenient to identify the behavioral changes of *regular passengers* (*5*). In this study, we define regular passengers as those who appear in the public transit system on every pre-determined normal day and show the same travel trajectories (i.e. same number of trips and tap-in stations). Hence, if these passengers showed different travel trajectories during the incident period, most likely they would be affected by the incident and choose a new travel mode. The mode choice after the incident for a regular passenger i is denoted as Y_i. We infer Y_i as the following:

- $Y_i =$ "Using transit" if 1) there are additional transit trips or 2) changes of tap-in stations in the incident period. The first condition implies the regular passenger may transfer to nearby rail or bus, with more transit trips than usual. The second condition implies the regular passenger may change to different rail lines or bus routes to respond to the incident.
- $Y_i =$ "Not using transit" if the transit trips that are supposed to happen disappear in the incident period. This means the regular passengers may change to other modes or cancel their trips.

Other regular passengers without the above behaviors may not be affected by the incident or have other choices that hard to be identified (e.g., transfer to another line without leaving the system), which are not considered in the behavior analysis.

The second step is to extract samples' characteristics (i.e. demographics). As regular passengers have consistent tap-in stations at the morning peak, we can infer their home locations as the morning tap-in stations, thus obtain the median household income in the corresponding neighborhoods or census tracks. The historical add-value transactions of passengers also provide demographic information. For example, low-income people may not be able to deposit a large amount of money in the smart card at once. Hence, the added values reflect passenger's income information to some extent. AFC data also provide information on passengers' fare status (e.g., whether using a pass). Denote all this "proxy" demographic information for passenger i as X_i.

In this study, we applied a binary logit model (*31*) to analyze the relationship between X_i and Y_i. Let the utility of mode j for passenger i be U_{ij}. We specified U_{ij} as

$$U_{ij} = \beta_j X_i + \varepsilon_j \tag{4}$$

where ε_j is the error term that is assumed to be Gumbel distributed. β_j is the vector of parameters to be estimated. It is worth noting that since we only collect the demographic information, no alternative specific variables (such as travel times of a specific mode) are used. The probability of passenger i choosing mode j is

$$P_{ij} = \frac{\exp(\beta_j X_i)}{\sum_{j' \in \mathcal{C}} \exp(\beta_j X_i)} \tag{5}$$

where $\mathcal{C} = \{$“Using transit”, “Not using transit”$\}$ is the choice set.

CASE STUDIES

Chicago Transit System

We use the CTA transit system as the case study. CTA is the second-largest transit system in the United States, providing services in Chicago, Illinois, and some of its surrounding suburbs, including the trains of the Chicago "L" (rail) and CTA bus service. It operates 24 hours each day and on an average weekday provides 0.84 and 0.81 million rides on buses and trains, respectively (*32*). The map of the CTA rail system is shown in Figure 1. The rail system consists of eight lines (named by color) and the "Loop". The Loop, located in the Chicago downtown, is the 2.88 km long circuit of elevated rail that forms the hub of the Chicago rail system. Its eight stations account for around 10% weekday boardings of the CTA trains.

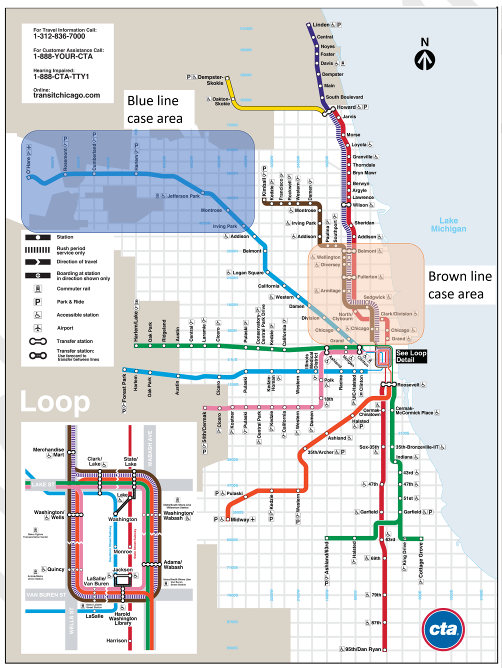


FIGURE 1: CTA rail system map

CTA’s AFC system is entry-only, meaning passengers only use their farecards when entering a rail station or boarding a bus, and so no information about a trip’s destination is directly provided. The train tracker system can provide trains arrival and departure time for each station.

According to the control center data, the CTA reported a total of 27,198 incidents in 2019. However, around 80% percent of the incidents have a duration of fewer than 10 minutes. Since small incidents may not affect the system significantly, this study focuses on substantial incidents that lasted longer than 1 hour.

Passengers who leave the rail system because of service disruptions need to re-tap-in if they decide to use other CTA normal buses or rails. And they are only charged with a small value of transfer fee. However, no tap-in is needed for shuttle bus users. So, we have no information for passengers using shuttle buses.

Rail disruption cases

Since the location of the incidents may influence the impact on the entire system, we selected two case studies of low and high redundancy, respectively, for comparison.

For this study, we also chose to look at substantial incidents. The CTA reported a total of 27,198 incidents in 2019, with an average of 75 incidents a day (*33*). However, the average incident duration was only 4.97 minutes. To widen the amount of available passenger data, we narrowed the scope to incidents that lasted longer than 1 hour.

Blue line Jefferson Park case

The incident on the blue line is a low redundancy case because the blue line is far away from other rail lines, adding difficulties in finding alternative services (see Figure 2). On February 1 (Friday), 2019, at 8:14 AM, the inbound track in the link between Harlem and Jefferson park was closed down due to physical problems. All trains in the blue line are suspended. CTA decided to use the remained single direction track to serve trains from both directions in the incident link. At 9:03 AM, 49 minutes after the incident, single track operations commence between Harlem and Jefferson Park, with shuttle service starting 7 minutes later. At 9:40 AM, all inbound trains succeeded to move under the single-track operation. At 12:09 PM, the full line is reopened. The entire incident lasted 4 hours and 9 minutes. And we set a major incident period as 8:14-9:40 AM.

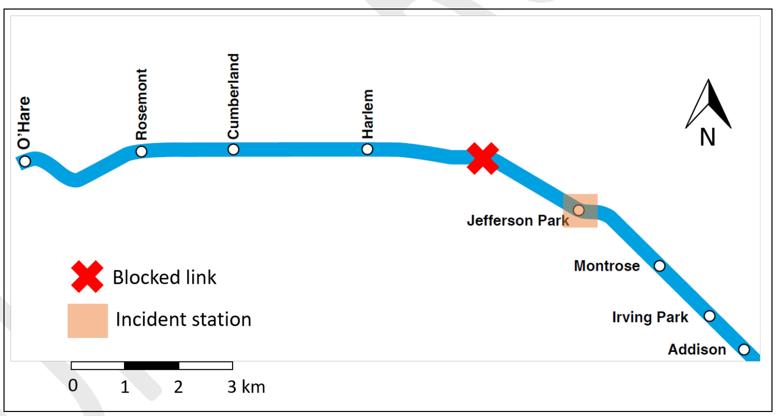


FIGURE 2: Incident diagram of blue line Jefferson Park case

Brown Line Sedgwick case

The incident on the brown line is a high redundancy case because the red line is a good substitution for the incident line (See Figure 3). On September 24 (Tuesday), 2019, at 9:09 AM, a Purple Line train collided with a Brown line train at Sedgwick station. The incident caused a sequence of stations to be blocked and closed in both brown and purple lines since these two lines share the same track in this area. This includes Fullerton and Armitage stations to the North and Chicago and Merchandise Mart (MM) Stations to the south. Southbound trains short turned at Fullerton station, while northbound trains short turned at MM station. At 9:28 AM, 19 minutes after the incident started, bus substitution service began between Fullerton to MM. Service resumes at all blocked stations at 10:19 AM, 70 minutes after the start of the incident.

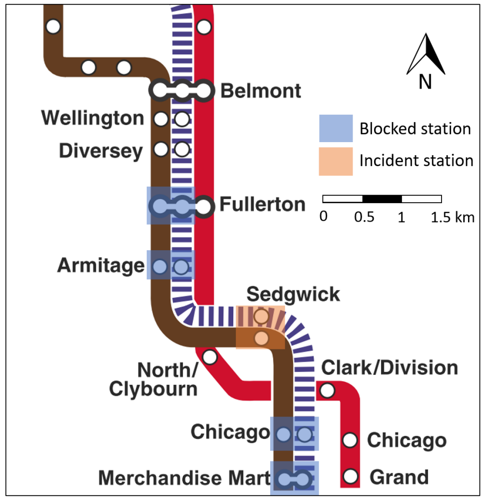


FIGURE 3: Incident diagram of brown line Sedgwick case

RESULTS

Supply analysis

Redundancy index

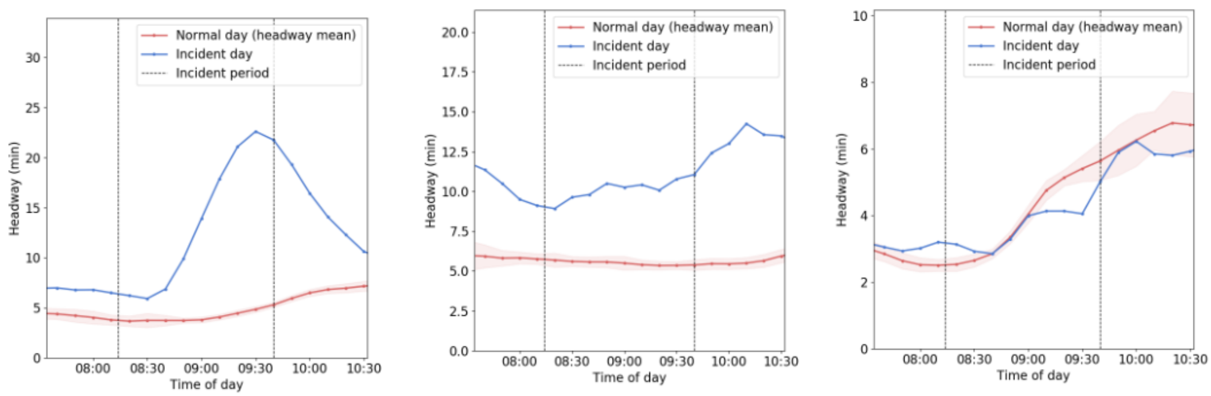
The NURI (Eq. 2) for the blue line case is 0.093, for the brown line case is 0.747. The values mean that the transit system maintains 9.3% (74.7%) transporting capacity for the blue (brown) line incident during the incident period.

The low redundancy in the blue line incident is due to the lack of alternative rail lines. Though there are some nearby bus services (see Figure 7), the capacity of buses is much lower than that of rails. Also, most of the bus routes are not directly connected to downtown, which increases the travel time for passengers using buses. The high redundancy of the brown line incident is as expected. In the incident area, the red line is almost parallel with the brown and purple lines. Also, there exist many south-bound bus routes going to the downtown (see Figure 9).

Headway analysis (Blue Line Jefferson Park incident)

The headway analysis results of the two incidents are shown in Figure 4 and 5. The shade around normal day lines is standard deviations (same for all following figures). For each case, we select three line+direction of interests to analyze.

Looking at the blue Line inbound (Figure 4a), the headway was a little bit longer than usual at the start of the day due to unknown reasons. As the incident starts, the headway increases immediately for inbound trips. The rise in headway is steeper before 9:30 AM, which is understandable since before that time the CTA is working on changing the system to single-track operation. Once the single-track operation successfully worked for all inbound trains, the headway plateaued, and then gradually decreased after 9:30 AM. Figure 4b shows the headway change of blue line outbound. Similarly, the headway was a little bit longer than usual at the start of the day. As the incident starts, headways gradually increased. However, though the single-track operation starts at 9:03 AM, the outbound headway still remains higher than the normal condition. This may be because CTA allowed more inbound trains to cross the single track area as they serve the major demands in the morning peak, which caused delays of outbound trains. The headway of the brown line is also shown in Figure 4c because the brown line can be a possible alternative for passengers in the south part of the blue line. The headway remains relatively unchanged throughout the Blue Line incident, which means the incident of the blue line did not affect the operation in the brown line.



(a) Blue Line (Inbound) (b) Blue Line (Outbound) (c) Brown Line (Inbound)

FIGURE 4: Headway temporal distribution (Blue Line Jefferson Park incident)

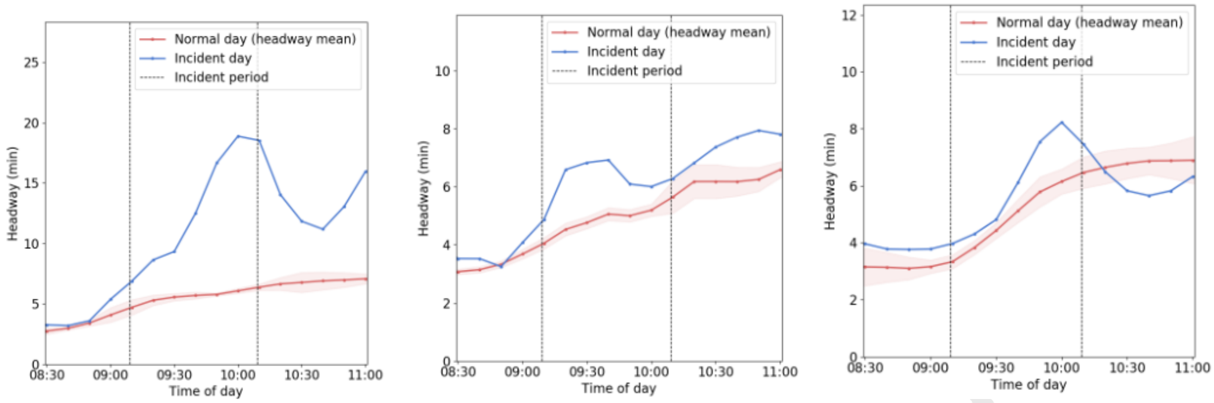
Headway analysis (Brown Line Sedgwick incident)

For the brown line Sedgwick case (see Figure 5), we can see a variety of different trends. Recall that the Brown and Purple Line share tracks in the incident area, while the Red Line runs on separate tracks in the incident area but shares tracks further up the line.

Looking at the brown line inbound (Figure 5a), we see a rise in inbound headways of both the Brown and Purple line trains. This is as expected because these two lines are blocked due to the incident and it took some time before trains began terminating early and turning around. We see a slight decrease in headways towards the end of the incident, as the system normalizes to the premature terminations of both parts of the line. Once the incident clears and the line is fully reopened, we see a further decrease in headways, with slight rising for both the Brown and Purple, most likely due to unevenly spaced trains throughout the line.

For the purple line (Figure 5b), As most of the local shops are much farther from the incident area, the service is less disrupted on a line level, despite sharing tracks with the brown line. This explains why headways deviated less from the average.

The red line experiences little deviation from its normal deal service for the first half of the incident (Figure 5c), largely because it does not share tracks at the incident location and could run largely uninterrupted. However, halfway through the incident, there is a spike in headway time. This could be due to two possible factors. Because of the bad service on brown and red lines, people chose to take the red line northbound instead, leading to more passengers and thus the longer time at the stations loading and unloading passengers. A second explanation could be that because trains on brown and purple lines were terminating and reversing farther up in areas that the red line shares tracks and stations, these actions could be causing increased congestion, increasing headway between trains.



(a) Brown Line (Inbound) (b) Purple Line (Inbound) (c) Red Line (Inbound)

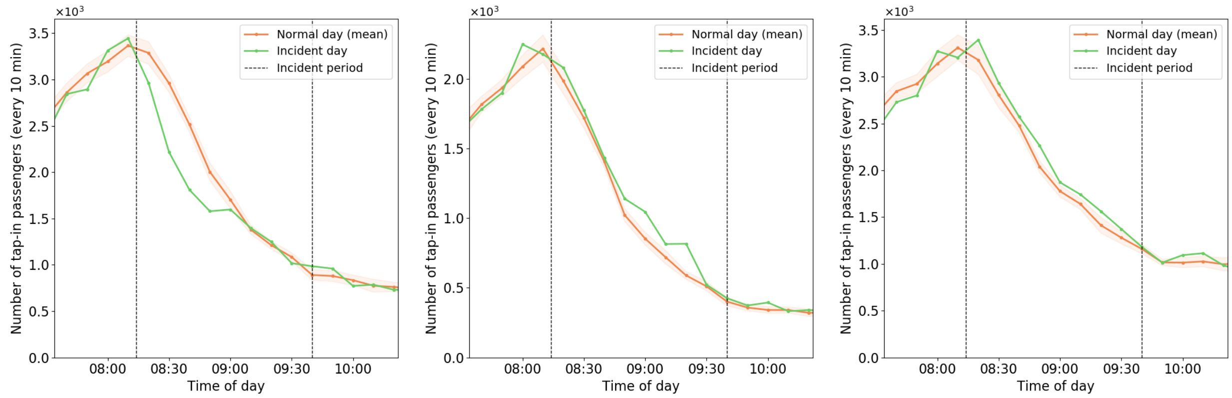
FIGURE 5: Headway temporal distribution (Brown Line Sedgwick incident)

Demand analysis

Passenger flow analysis allows us to look at how passengers reacted to unplanned disruptions. Specifically, we look at the line-level and station-level demand changes during an incident. This allows us to understand how individual lines and stations react demand-wise to both incidents.

Passenger flow analysis (Blue line Jefferson park incident)

We first analyze the demand patterns of the blue, brown, and red lines for the blue Line Jefferson Park incident (see Figure 6. The blue line (Figure 6a) initially experiences a drop in passengers entering the blue line immediately after the incident, which is as expected because passengers were informed of the incident and chose to not tap in. As single tracking opens up, tap ins started to return to regular levels as the system's backlog slowly begins to clear. By 9:40 AM, single tracking is in full operation. Hence, the number of tap in passengers is closer to the average. For the brown line (Figure 6b), we see a slight spike of demand about 30 minutes after the blue Line incident. This is due to the brown line being relatively close to the incident area. Passengers most likely either walked or took alternative bus routes and then transferred onto the brown Line to continue their commute into downtown Chicago. We also observe a consistent demand increase in the red line for the entire major incident period (Figure 6c). This reason is the same as that of the brown line because red and brown lines are largely overlapped near the incident area.



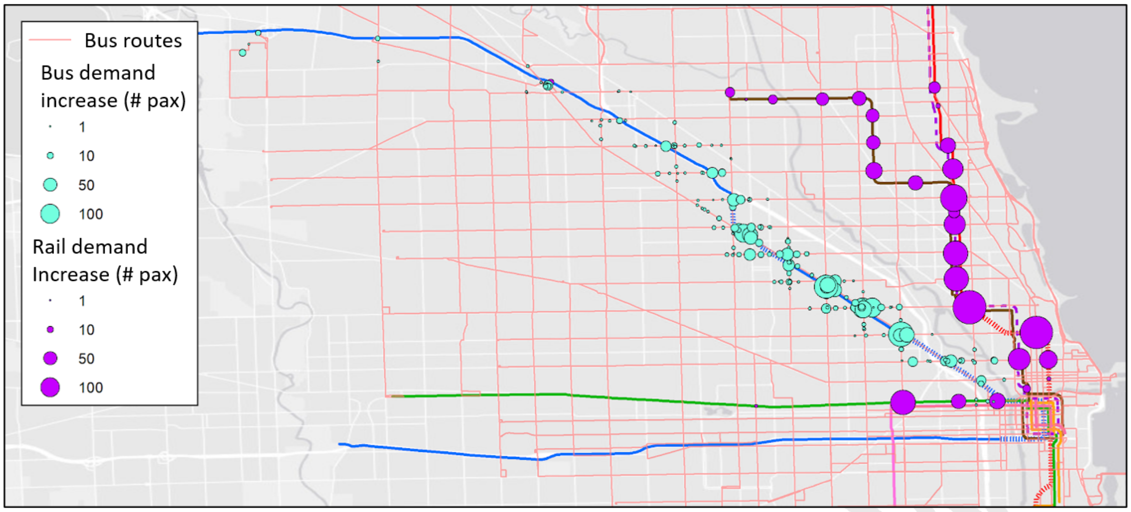
(a) Blue Line (blocked) (b) Brown Line (not blocked) (c) Red Line (not blocked)

FIGURE 6: Line level passenger flow analysis (Blue Line Jefferson Park Incident)

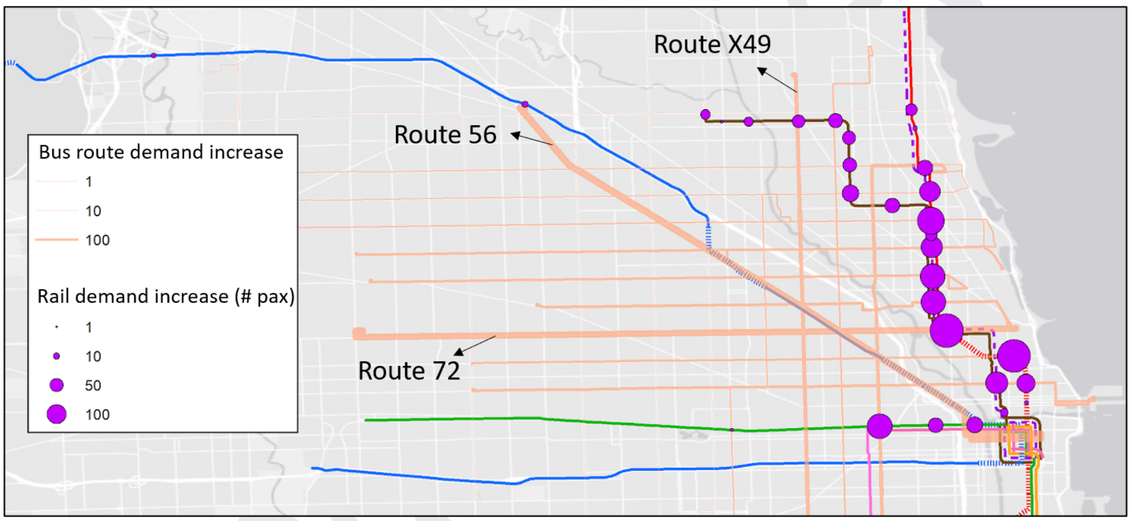
Demand changes in rail stations and bus system are shown in Figure 7. In Figure 7a, we see a rise in bus stations that are close to the blue line, which means many passengers chose to use bus services instead after the incident. We also observe a substantial increase in ridership on the nearby brown and red lines. This is presumably from passengers taking buses near the blue line and transferring to the brown and red lines. However, we see little increase in ridership on the green line in comparison. This shows that some passengers did not make good choices, as the green line is close to the blue line are provides service to downtown as well.

In Figure 7b, we calculate demand changes at bus route level. The demand for multiple bus lines increased, top 3 of which are routes 56, 72, and X49. Route 56 increased most likely because it runs parallel to much of the blue Line and connects directly to downtown. The increase in route 72 was most likely due to passengers transferring to that route and using it to connect to the brown and red lines further east in order to make their journey to downtown. Since there is little increase in the green line where the X49 connects, most of the increased ridership in route X49 was probably passengers with destinations in the south that route X49 directly heads to.

The total decrease of the number of tap-in passengers in the blue line is 2,219, while the increases in nearby bus stations, brown line, and red line are 2,426, 845, and 1,125, respectively. It is worth noting that passengers may tap-in the blue line then get out to use buses due to long waiting times. This causes the demand decrease in the blue line is actually larger than 2,219. By the number comparison, we observe that most of the decreased demand in the blue line change to use buses instead. For these passengers using buses, around 80% of them transferred to nearby red and brown lines.



(a) Demand changes of nearby bus stops and rail stations levels

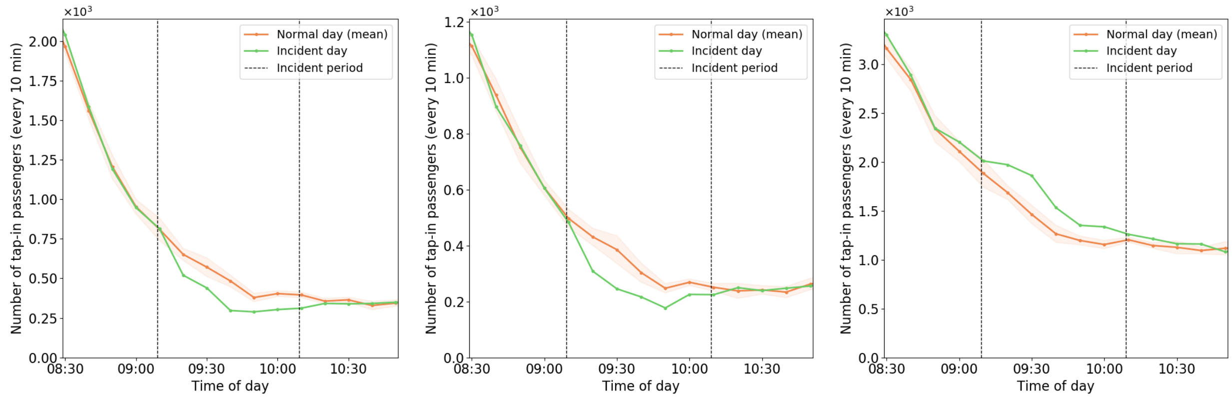


(b) Demand changes of bus routes

FIGURE 7: Station demand increase patterns (Blue Line Jefferson Park Incident)

Passenger flow analysis (Brown line Sedgwick park incident)

For the Brown Line Sedgwick case, the line-level demand changes are shown in Figure 8. As expected, demands on the brown and purple lines both decreased during the incident and normalized after the incident. As the red line runs adjacent to the brown and purple lines for a significant portion, we see an increase in demand during the incident period and normalizing after the incident back to average demand.



(a) Brown Line (blocked) (b) Purple Line (blocked) (c) Red Line (not blocked)

FIGURE 8: Demand Analysis - Line Level - Brown Line Sedgwick Incident

The demand change patterns at the station level are shown in Figure 9. During the incident, we see an increase in rail demand at Fullerton and Belmont, stations that have direct connections to the undisputed Red Line. We also see clusters of increased bus demand. Of note are the clusters outlined in red and blue. The red clusters represent increased bus demand proximal to blocked stations. These passengers may direct transfer to nearby bus stops from the blocked line. Additionally, the blue clusters represent increases in bus demand for routes that connect directly to downtown. These passengers may be those who live in the corresponding neighborhoods and change to use buses during the incident.

The total decrease in the number of tap-in passengers in the brown and purple lines are 1,141, while the increase in nearby bus stations and the red line are 696 and 1,414, respectively. This means around 800 passengers may first tap in the brown and purple lines, then transfer to the red line. For all passengers leaving the system, around one-third of them transfer to buses, and two-thirds of them transfer to the red line. However, there may be also many passengers with direct transfers in the system, which we cannot observe from AFC data.

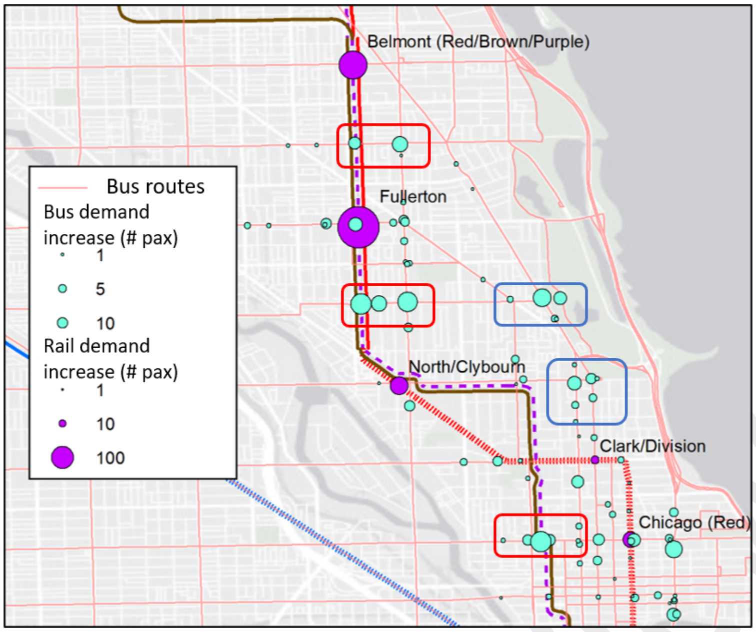


FIGURE 9: Station demand increase patterns (Brown Line Sedgwick incident)

Individual choices analysis

We sampled 539 regular passengers who are affected by the incident (see Section 3.2.2) using the AFC data during the brown line incident. The descriptive statistics of all observations are shown in Table 2.

TABLE 2: Descriptive statistics of samples

Variables	Mean	Standard deviation
House location median income ($)	79,332	29,284
Average added value in 2019 ($)	52.55	34.78
Added value times in 2019	31.83	38.54
Using pass (Yes = 1)	0.419	0.494
Reduced fare staus (Yes = 1)	0.082	0.274
Number of observations: 539		
Choices: using CTA: 144; not using CTA: 395		

The estimation results of the binary logit model are shown in Table 3. "Not using CTA" is set as the base mode. We observe that high household income has a negative effect on using CTA. This may be because high-income people can usually afford alternative modes of transportation (such as Uber/Lyft). Passengers using a pass are more likely to using CTA after the incident. This is understandable because pass users are generally commuters or those that use the system frequently. They are familiar with the system services and able to find alternative CTA services are the incident. Passengers with reduced fare status are also more likely to use CTA services. The reason may be that reduced fare status users are usually students, seniors and disabled people. They usually rely primarily on CTA to travel.

TABLE 3: Individual choice model estimation results

Parameters	Value (standard error)	
CTA: Alternative specific constant	-0.747 (0.427)	*
CTA: Average added value ($100)	-0.417 (0.441)	
CTA: Average add-value times (100)	-0.289 (0.428)	
CTA: House location median income	-0.0743 (0.0361)	**
CTA: Using pass (Yes = 1)	1.19 (0.263)	***
CTA: Reduced fare status (Yes = 1)	0.711 (0.355)	**
Not CTA: Alternative specific constant	0 (fixed)	

Number of individuals: 539. Adjusted $\rho^2 = 0.196$
*: $p < 0.1$; **: $p < 0.5$; ***: $p < 0.01$;

CONCLUSION AND DISCUSSION

This study proposes a general incident analysis framework from supply and demand sides using automatically-collected data (AFC and AVL) in public transit systems. Specifically, from the supply side, we propose an incident-based network redundancy index to analyze the network's ability to provide alternative services under a specific rail disruption. The impacts on the operation are analyzed through the headway changes in multiple rail lines. From the demand side, we calculate the demand changes of different rail lines, rail stations, bus routes, and bus stops to show the passenger flow redistribution under incidents. Individual behavior is analyzed using a binary logit model based on inferred passengers' mode choices and socio-demographics using AFC data. The CTA public transit system is used as a case study. Two rail disruption cases are analyzed, one of which has high network redundancy and low for another.

Results show that the service frequency of the incident line was largely reduced during the incident time. Depending on the incident location, the network's redundancies are different, as well as the passengers' behavior. In the low redundancy scenario, most passengers choose to use nearby buses, either directly move to their destinations or move the nearby rail lines. In the high redundancy scenario, most of the passengers directly transferred to nearby lines.

This study helps the agency understand better how passengers move and what alternatives they may pick, based on which operators can better design shuttle buses and bus routes connecting with other buses and trains. They can also increase the frequency of buses that are heavily used by passengers during the incident. Since we observe some passengers' choices may not be optimal during the incident. The operator can offer on twitter or other public announcement platforms different routes passengers can take to avoid the disruption. Integrating a better-personalized passenger profile on the route design may also help.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: study conception and design: B. Mo, H.N. Koutsopoulos, J. Attanucci, J. Zhao; data collection: B. Mo; analysis and interpretation of results: B. Mo, M. Franque; draft manuscript preparation: B. Mo, M. Franque;. All authors reviewed the results and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

The authors would like to thank the Chicago Transit Authority (CTA) for their support and data availability for this research.

REFERENCES

1. Yin, H., B. Han, D. Li, and Y. Wang, Evaluating disruption in rail transit network: a case study of Beijing subway. *Procedia Engineering*, Vol. 137, 2016, pp. 49–58.
2. Zhang, D.-m., F. Du, H. Huang, F. Zhang, B. M. Ayyub, and M. Beer, Resiliency assessment of urban rail transit networks: Shanghai metro as an example. *Safety Science*, Vol. 106, 2018, pp. 230–243.
3. Xu, X., A. Chen, S. Jansuwan, K. Heaslip, and C. Yang, Modeling transportation network redundancy. *Transportation research procedia*, Vol. 9, 2015, pp. 283–302.
4. Berdica, K., An introduction to road vulnerability: what has been done, is done and should be done. *Transport policy*, Vol. 9, No. 2, 2002, pp. 117–127.
5. Mojica, C. H., *Examining changes in transit passenger travel behavior through a smart card activity analysis*. Ph.D. thesis, Massachusetts Institute of Technology, 2008.
6. Tian, X. and B. Zheng, Using Smart Card Data to Model Commuters' Responses Upon Unexpected Train Delays. In *2018 IEEE International Conference on Big Data (Big Data)*, IEEE, 2018, pp. 831–840.
7. Sun, H., J. Wu, L. Wu, X. Yan, and Z. Gao, Estimating the influence of common disruptions on urban rail transit networks. *Transportation Research Part A: Policy and Practice*, Vol. 94, 2016, pp. 62–75.
8. Rahimi, E., A. Shamshiripour, R. Shabanpour, A. Mohammadian, and J. Auld, Analysis of Transit Users' Response Behavior in Case of Unplanned Service Disruptions. *Transportation Research Record*, Vol. 2674, No. 3, 2020, pp. 258–271.
9. Rahimi, E., A. Shamshiripour, R. Shabanpour, A. Mohammadian, and J. Auld, Analysis of transit users' waiting tolerance in response to unplanned service disruptions. *Transportation Research Part D: Transport and Environment*, Vol. 77, 2019, pp. 639–653.
10. Currie, G. and C. Muir, Understanding Passenger Perceptions and Behaviors During Unplanned Rail Disruptions. *Transportation Research Procedia*, Vol. 25, 2017, pp. 4396–4406.
11. Murray-Tuite, P., K. Wernstedt, and W. Yin, Behavioral shifts after a fatal rapid transit accident: A multinomial logit model. *Transportation research part F: traffic psychology and behaviour*, Vol. 24, 2014, pp. 218–230.
12. Fukasawa, N., K. Yamauchi, A. Murakoshi, K. Fujinami, and D. Tatsui, Provision of forecast train information and consequential impact on decision making for train-choice. *Quarterly Report of RTRI*, Vol. 53, No. 3, 2012, pp. 141–147.
13. Teng, J. and W.-R. Liu, Development of a behavior-based passenger flow assignment model for urban rail transit in section interruption circumstance. *Urban Rail Transit*, Vol. 1, No. 1, 2015, pp. 35–46.
14. Lin, T., S. Srikukenthiran, E. Miller, and A. Shalaby, Subway user behaviour when affected by incidents in Toronto (SUBWAIT) survey—A joint revealed preference and stated preference survey with a trip planner tool. *Canadian Journal of Civil Engineering*, Vol. 45, No. 8, 2018, pp. 623–633.

15. Balakrishna, R., Y. Wen, M. Ben-Akiva, and C. Antoniou, Simulation-based framework for transportation network management in emergencies. *Transportation Research Record*, Vol. 2041, No. 1, 2008, pp. 80–88.
16. Suarez, P., W. Anderson, V. Mahal, and T. Lakshmanan, Impacts of flooding and climate change on urban transportation: A systemwide performance assessment of the Boston Metro Area. *Transportation Research Part D: transport and environment*, Vol. 10, No. 3, 2005, pp. 231–244.
17. Hong, L., J. Gao, and W. Zhu, Self-evacuation modelling and simulation of passengers in metro stations. *Safety science*, Vol. 110, 2018, pp. 127–133.
18. O'Connor, P. and A. Kleyner, *Practical reliability engineering*. John Wiley & Sons, 2012.
19. Wheeler, D. C. and M. E. O'Kelly, Network topology and city accessibility of the commercial Internet. *The Professional Geographer*, Vol. 51, No. 3, 1999, pp. 327–339.
20. Kalungi, P. and T. T. Tanyimboh, Redundancy model for water distribution systems. *Reliability Engineering & System Safety*, Vol. 82, No. 3, 2003, pp. 275–286.
21. Sheffi, Y. and J. B. Rice Jr, A supply chain view of the resilient enterprise. *MIT Sloan management review*, Vol. 47, No. 1, 2005, p. 41.
22. Wilson-Goure, S., N. Houston, and A. Easton, Assessment of the State of the Practice and State of the Art in Evacuation Transportation Management. *Task Two: Literature Search for Federal Highway Administration (ITS-JPO). McLean, Virginia*, 2006.
23. Murray-Tuite, P. M., A comparison of transportation network resilience under simulated system optimum and user equilibrium conditions. In *Proceedings of the 2006 Winter Simulation Conference*, IEEE, 2006, pp. 1398–1405.
24. Goodchild, A., E. Jessup, E. McCormack, D. Andreoli, K. Pitera, S. Rose, C. Ta, et al., *Development and analysis of a GIS-based statewide freight data flow network*. Washington (State). Dept. of Transportation, 2009.
25. Tsuchiya, R., Y. Sugiyama, and R. Arisawa, A Route Choice Support System for Use During Disrupted Train Operation. In *15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual MeetingITS AmericaERTICOITS JapanTransCore*, 2008.
26. Pnevmatikou, A. and M. Karlaftis, Demand changes from metro line closures, 2011.
27. Kamaruddin, R., I. Osman, and C. A. C. Pei, Public transport services in klang valley: customer expectations and its relationship using SEM. *Procedia-Social and Behavioral Sciences*, Vol. 36, 2012, pp. 431–438.
28. Bai, Y. and L. Kattan, Modeling riders' behavioral responses to real-time information at light rail transit stations. *Transportation Research Record*, Vol. 2412, No. 1, 2014, pp. 82–92.
29. Leistritz, L., T. Weiss, K.-J. Bär, F. De VicoFallani, F. Babiloni, H. Witte, and T. Lehmann, Network redundancy analysis of effective brain networks; a comparison of healthy controls and patients with major depression. *Plos one*, Vol. 8, No. 4, 2013, p. e60956.
30. Goulet-Langlois, G., H. N. Koutsopoulos, Z. Zhao, and J. Zhao, Measuring regularity of individual travel patterns. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 19, No. 5, 2017, pp. 1583–1592.
31. Ben-Akiva, M. E., S. R. Lerman, and S. R. Lerman, *Discrete choice analysis: theory and application to travel demand*, Vol. 9. MIT press, 1985.
32. CTA, *CTA Monthly Ridership Report for September 2019*, 2019.
33. CTA, *Chicago Transit Authority - CTA Buses and Train Service*, 2019.