Inferring passenger responses to urban rail disruptions using smart card data: A probabilistic framework

Baichuan Mo [a,*], Haris N. Koutsopoulos [b], Jinhua Zhao [c]

[a] *Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States of America*
[b] *Department of Civil and Environmental Engineering, Northeastern University, Boston, MA 02115, United States of America*
[c] *Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA 20139, United States of America*

ARTICLE INFO

ABSTRACT

This study proposes a probabilistic framework to infer passengers' responses to unplanned urban rail service disruptions using smart card data in tap-in-only public transit systems. We first identify 19 possible response behaviors that passengers may have based on their decision-making times and locations (i.e, the stage of their trips when an incident happened), including transferring to a bus line, canceling trips, waiting, delaying departure time, etc. A probabilistic model is proposed to estimate the mean and variance of the number of passengers in each of the 19 behavior groups using passengers' smart card transactions. The 19 behavioral responses can be categorized from two aspects. From the behavioral aspect, they can be grouped into 5 aggregated response behaviors including using bus, using rail (changing or not changing route), not using public transit, and not being affected. The inference of the 19 behaviors can be classified into four cases based on the information used (historical trips vs. subsequent trips) and the context of the observed transactions (direct incident-related vs. indirect incident-related). The public transit system (bus and urban rail) of the Chicago Transit Authority (CTA) is used as a case study based on a real-world rail disruption incident. The model is applied with both synthetic data and real-world data. Results with synthetic data show that the proposed approach can estimate passengers' behavior well. The mean absolute percentage error (MAPE) for the estimated expected number of passengers in each behavior group is 20.5%, which outperforms the rule-based benchmark method (60.3%). The estimation results with real-world data are consistent with the incident's context. An indirect model validation method using demand change information and incident log data demonstrates the reasonableness of the results.

1. Introduction

Urban rail transit plays an important role in urban mobility. However, with aging systems, continuous expansion, and near-capacity operations, service disruptions often occur. Disruptions can range from short-term delays at some stations to shutdowns of entire subway lines over an extended period. These incidents may result in delays and cancellation of thousands of trips as well as economic and opportunity losses (Cox et al., 2011).

Consequently, there is growing research interest and literature in the area of rail disruption analysis and management. These efforts can be classified into two types: supply-oriented and demand-oriented (Leng et al., 2018). The supply-oriented research focuses on analyzing the network vulnerability and improving network resilience from the supply and operation perspectives. The

---

* Corresponding author.
E-mail address: baichuan@mit.edu (B. Mo).

literature in rail transit network vulnerability based on complex network theory is very intensive. It explores the vulnerability of network topology when some nodes or links of the network are failed. Degree, betweenness, centrality measures, and connectivity methods are usually used (Derrible and Kennedy, 2010; Zhang et al., 2015; Dimitrov and Ceder, 2016; López et al., 2017). From the operations point of view, many studies look at adjusting the timetable (Veelenturf et al., 2016), managing rolling stock (Nielsen et al., 2012), and designing shuttle buses (Jin et al., 2016) during urban rail disruptions to ensure operational feasibility and improve system efficiency.

Demand-oriented research focuses on understanding and modeling passengers' behavior under rail disruptions. Transit users' behavior can be significantly different in the event of service disruptions and vary depending on the stage of the trip at the time of the disruption (Lin et al., 2018). A better understanding of passengers' behavior in the event of disruption is important for operators to recommend alternative routes, adjust the capacity of rail lines, and provide shuttle services (Adelé et al., 2019). However, nearly all of the previous research investigated passenger behavior using survey-based methods (Bai and Kattan, 2014; Teng and Liu, 2015; Murray-Tuite et al., 2014). For example, Lin et al. (2018) used a joint revealed and stated preference (SP) survey to estimate transit user mode choice in response to a transit service disruption in the City of Toronto. Rahimi et al. (2020) utilized survey data collected in the Chicago Metropolitan Area to analyze how transit users respond to unplanned service disruptions and the factors that affect their behavior. Survey-based methods are usually time-consuming and labor-intensive. Besides, SP surveys require passengers to respond to hypothetical situations, which may not reflect the actual travel choices of passengers (Sun et al., 2016).

Recently, thanks to the widely adopted automated fare collection (AFC) system, passengers' travel information is recorded in the AFC data, providing opportunities to capture individual choices under rail disruptions using data-driven approaches. However, studies using AFC data to explore the impact of unplanned disruptions on individual responses are limited. Silva et al. (2015) proposed a method to analyze large-scale mass transportation systems during unplanned disruptions. They estimated the disruption effects on passenger volumes during incidents using smart card data. van der Hurk (2015) developed a model based on smart cards to forecast the route choices of passengers impacted by disruptions under different scenarios. The study shows that operators can help passengers minimize their overall inconvenience by providing individual advice. Sun et al. (2016), using AFC data, estimated three groups of passengers (leaving the system, detouring, and continuing to travel) during the rail disruption. Recently, Liu et al. (2021) also proposed a data-driven approach to evaluate disruption impacts on system performance and individual responses in urban railway systems using AFC data. They considered four groups of passengers: performing trips, changing travel time, changing stations, and changing modes.

However, there are some limitations in the previous studies. First, the approaches of identifying passenger responses in previous studies are rule-based and deterministic, meaning that they directly map the observed AFC records to a specific response behavior. The rule-based method ignores uncertainty and randomness in passengers' behavior (i.e., the observed AFC records may be due to behavior randomness, rather than the impact of incidents), which may introduce estimation bias. Also, deterministic methods cannot quantify the uncertainty (i.e., variance) in the estimated results. Second, most of the previous studies are based on data from closed AFC systems with both tap-in and tap-out information, which does not apply to many open transit systems where only tap-in information is available (such as the transit systems in Chicago, Boston, and New York). Third, most of the previous studies only considered three or four possible response behaviors. In this paper, we show that passenger's responses are diverse depending on where they are when the incident happens. There are 19 possible responses identified in this study.

To fill the research gap, this paper proposes a probabilistic passenger behavior estimation framework under rail disruptions using tap-in-only AFC data. The historical travel trajectories before the incident and the subsequent travel records after the incident are both used for inference and capturing the uncertainty in passengers' behavior. We first identify 19 possible response behaviors that passengers may have based on their decision-making times and locations[1] (i.e, the stage of their trips when an incident happened), including transferring to a bus line, canceling trips, waiting, delaying departure time, etc. A statistical inference model is proposed to estimate the mean and variance of the number of passengers in each of the 19 behavior groups using passengers' AFC data. The urban bus and rail system operated by the Chicago Transit Authority (CTA) is used as a case study. The proposed model is validated with a synthetic data set and applied using an actual data set from CTA. Results show that the proposed model can estimate passengers' travel behavior after the rail disruption accurately and outperform the rule-based benchmark model.

The identified 19 behavioral responses can be classified from two aspects. From the behavioral aspect, they can be grouped into 5 aggregated response behaviors including using bus, using rail (changing or not changing route), not using public transit, and not being affected. These five aggregated response behaviors are general and applicable for the incident analysis for any other public transit system. From the methodological aspect, the inference of the 19 behaviors can be classified into four cases based on the information used (historical trips vs. subsequent trips) and the context of the observed transactions (direct incident-related vs. indirect incident-related).

The main contributions of the paper are as follows:

- Provide a comprehensive framework of passengers' behavior under service disruptions. A total of 19 possible behavior groups for passengers at different stages of their trips are considered, which enables a more detailed modeling framework. The behavior identification is based on when and where passengers are making their decisions during a disruption. The method is general and can be used for other transit systems (the resulting possible behaviors may vary according to the context of the system, i.e., not necessarily 19)

---

[1] The proposed model is not restricted to the 19 behaviors. The way of recognizing possible responses is general and can be extended to different case studies. See Section 2.1 for details.
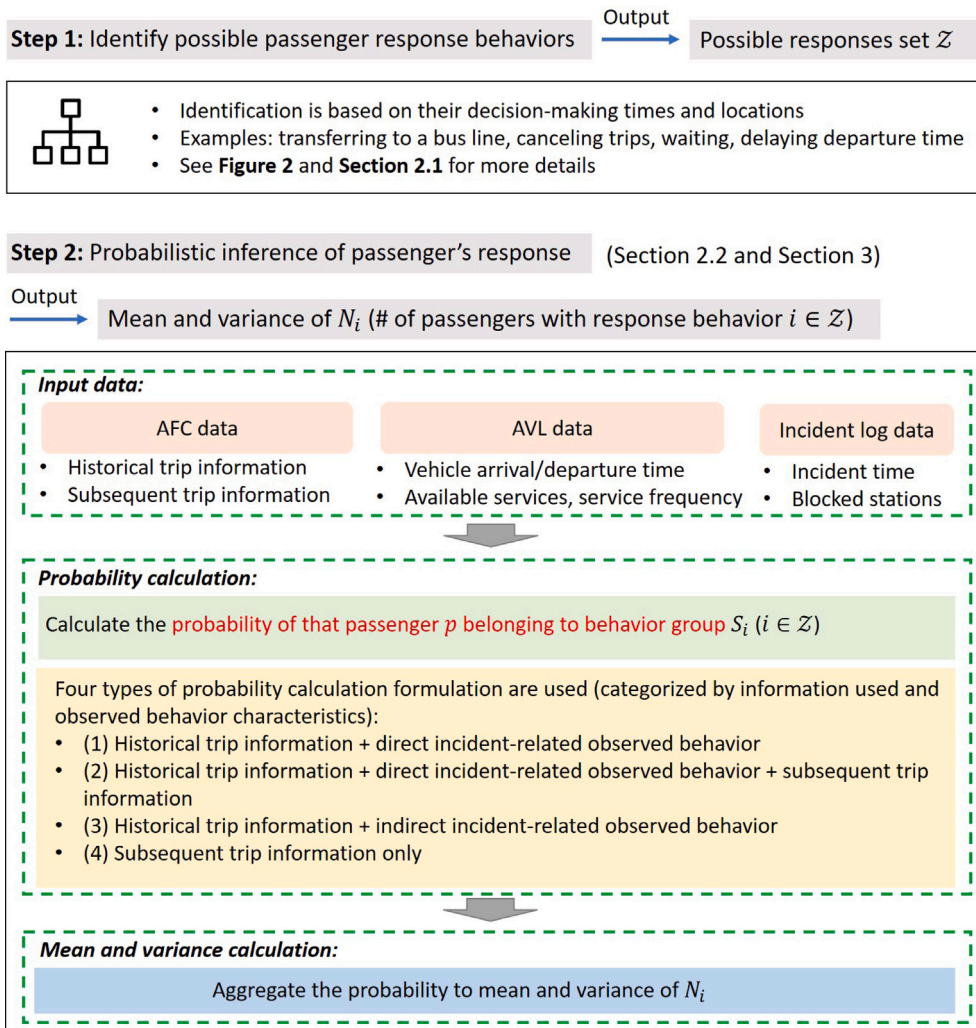
Step 1: Identify possible passenger response behaviors → **Output** → Possible responses set $\mathcal{Z}$

- Identification is based on their decision-making times and locations
- Examples: transferring to a bus line, canceling trips, waiting, delaying departure time
- See **Figure 2** and **Section 2.1** for more details

Step 2: Probabilistic inference of passenger's response   (Section 2.2 and Section 3)

Output → Mean and variance of N_i (# of passengers with response behavior $i \in \mathcal{Z}$)

Input data:

AFC data	AVL data	Incident log data

- Historical trip information
- Subsequent trip information

- Vehicle arrival/departure time
- Available services, service frequency

- Incident time
- Blocked stations

Probability calculation:

Calculate the probability of that passenger p belonging to behavior group S_i ($i \in \mathcal{Z}$)

Four types of probability calculation formulation are used (categorized by information used and observed behavior characteristics):
- (1) Historical trip information + direct incident-related observed behavior
- (2) Historical trip information + direct incident-related observed behavior + subsequent trip information
- (3) Historical trip information + indirect incident-related observed behavior
- (4) Subsequent trip information only

Mean and variance calculation:

Aggregate the probability to mean and variance of N_i

Fig. 1. Framework of the methodology.

- Propose a probabilistic behavior inference model with a specific formulation for each of the 19 behavior groups. The model enables the estimation of the mean and variance of the number of passengers in each group to capture passenger's behavior uncertainty. To the best of the authors' knowledge, this is the first article providing the estimation for both mean and variance of post-incident behaviors using AFC data.
- Leverage both passengers' historical travel trajectories and their subsequent tap-in records after the incident to facilitate behavior inference. This is contrary to previous studies where only the AFC data on the incident day is used.

The rest of the paper is organized as follows. Sections 2 and 3 present the methodology of this study. Section 4 discusses the case study for model application and the corresponding results. Section 5 concludes the paper and discusses future research directions.

2. Model framework

Fig. 1 shows an overview of the model framework. There are two steps for inferring passenger's responses. At the first step, we aim to identify all possible passenger response behaviors to the incident based on their decision-making times and locations. Details of step 1 are shown in Section 2.1. At step 2, based on the results of step 1, we aim to associate each passenger to a specific response behavior by calculating the corresponding probabilities based on the observed passenger AFC records and his/her travel histories. The input data for the inference include AFC, AVL (automated vehicle location), and incident log. There are four different formulations for the probability calculation, which are categorized by the used information and properties of observed AFC records. Then, we aggregate the probabilities to the mean and variance of the number of passengers in the different response behavior groups. Details of the step 2 are shown in Sections 2.2 and 3.
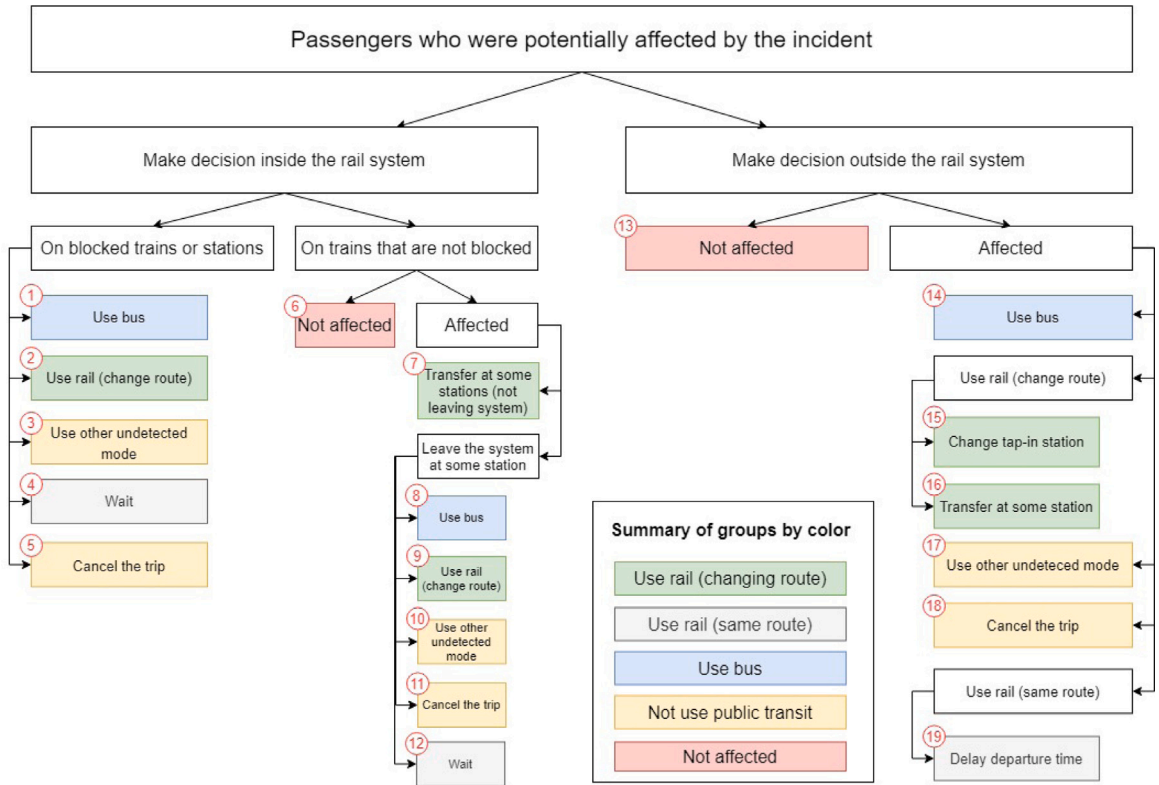
Fig. 2. Passenger responses to an unplanned rail disruption.

2.1. Passenger behavior under disruptions

A prerequisite for behavior inference is to identify possible options passengers may have during the disruption. According to Sun et al. (2016), passenger responses to a service disruption are generally triggered when the delay time is long enough (e.g., greater than 30 min). Hence, for a meaningful analysis, this study focuses on substantial unplanned service disruptions (i.e., blockage or shutdown of service as opposed to reduced capacity or frequency) so that there are observable behavior changes.

For an incident beginning at T_1 and ending at T_2, we consider the analysis time period as $[T_s, T_e] = [T_1 - \delta_1, T_2 + \delta_2]$, where δ_1 is set as the maximum travel time in the system because all passengers tapping in before $T_1 - \delta_1$ are not affected. δ_2 is the recovery time for the system after the incident ends, which can be pre-calculated based on the smart card data (Mo et al., 2022) (i.e., we assume that after $T_2 + \delta_2$ the system is fully recovered). We only consider passengers who were potentially affected by the incident, defined as passengers who had (or were supposed to have) tap-in records during the analysis period ($[T_s, T_e]$) on the incident day. Passengers who are supposed to tap in are those with historical trips indicating that they may have a rail trip during this period, though we do not observe them on the incident day AFC data. These passengers are considered because they may cancel their trips or use other undetected modes (details can be found in the following sections).

Fig. 2 summarizes possible passenger behaviors under different cases. A total of 19 possible response behaviors are considered. The general approach to characterize these behaviors is elaborated on below. The approach can be applied to other public transit systems to identify a similar set of possible response behaviors.

Passengers' behavior may vary a lot depending on the stage of their trips at the time of service disruption (Lin et al., 2016). Therefore, all potentially affected passengers are first divided into two groups: (a) passengers in and (b) out of the rail system. The first group of passengers was on a train or inside a station platform when the incident happened, while passengers in the second group have not entered the system yet (e.g., at home).

When the disruption happens, some of the stations in the rail system are blocked (i.e. trains are not allowed to move in these stations) due to the incident. Passengers who are in the blocked stations/trains are forced to leave the system. These passengers have five options: changing to a bus line, changing to another rail line or station, waiting until the system is restored, canceling the trip, or changing to other undetected modes. It is worth noting that if they choose transit services (rail or bus) again, they need to re-tap to use the alternative services. The undetected modes include Transit Network Companies (TNC), walking, bicycling, etc. It is worth noting that using a shuttle bus that was deployed to mitigate the incident impacts can be categorized into "bus" or "undetected mode" depending on whether passengers are required to tap their fare card or not. If passengers are not in the blocked stations,
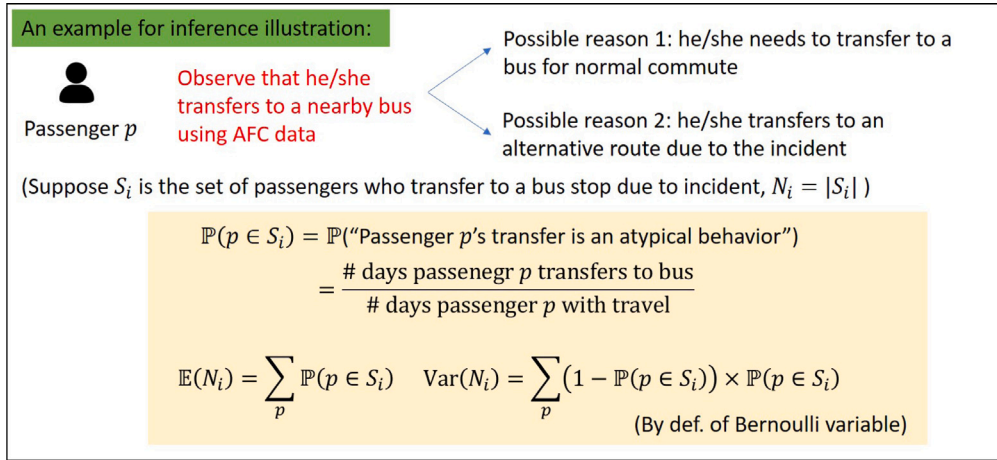
Fig. 3. Illustration example of the probabilistic behavior inference.

their trains could still move. Hence, they may not be affected by the incident. Or if they were affected, compared with passengers in the blocked stations, they would have one more option: transferring halfway to another line without leaving the system.

For passengers out of the system when the incident happens, if their travel routes on the rail system are not blocked, they are not affected. Otherwise, instead of following the original route, they may choose to use buses, use rails but change the tap-in station, use rail by transferring at a halfway station, use other undetected modes, cancel the trip, or delay their departure time until the system recovers.

All these behaviors can be summarized into five groups: use rail (changing route), use rail (same route), use buses, not use public transit, and not being affected. Note that these five alternatives are general for different transit systems and can be used to guide the potential behavior identification. To better describe these behaviors, we assign a specific ID to each (i.e., numbers in the red circle in Fig. 2). These behaviors are inferred separately based on their characteristics in the AFC data.

2.2. Probabilistic behavior inference

We propose a probabilistic framework to infer passengers in each behavior group using AFC data. The probabilistic framework facilitates the inference of whether a specific observed behavior for a passenger is due to the incident, or is typical. In this study, we focus on open public transit systems where only tap-in information is available. The AFC data include both bus and rail boarding records.

The key idea of the inference framework is to identify (1) whether an observed AFC data record (e.g., transfer to bus) is atypical or not and (2) whether the atypical behavior is owing to the incident or behavioral randomness. These two questions are answered probabilistically (i.e., obtaining the corresponding probabilities). And the corresponding probabilities are used to calculate the mean and variance of the number of passengers in each behavior group.

Fig. 3 presents an explanatory example for the probabilistic behavior inference method. Consider a passenger p in the system. We observe that he/she has a transfer record to a nearby bus stop from the incident line. In typical rule-based method (Sun et al., 2016; Liu et al., 2021), this passenger will be directly identified as "transferring to bus due to incident". However, in the probabilistic framework, we consider two possible reasons for this observed record: (1) he/she transfers to a bus for a normal commute. (2) he/she transfers to an alternative route due to the incident. We should only account for the second reason as the impact of incidents. Therefore, we use historical data to calculate the probability that "this transfer is an atypical behavior" (i.e., due to the incident). Then, the mean and variance of the number of passengers with a specific response behavior can be obtained from this probability (by the definition of the Bernoulli random variable).

2.2.1. Notation

Denote S_i as the set of passengers who have behavior i **in response to the incident**, $i \in \mathcal{Z} = \{1, 2, \ldots, Z\}$ (behavior IDs are shown in Fig. 2, for example, "Behavior 1" means offloading from the train when an incident happens and using bus to respond to the incident). Let N_{S_i} be the number of passengers in set S_i. The day when the incident happens is referred to as the **incident day**. A **normal day** is defined as a day without (substantial) incidents in the analysis period and area and with the same day of the week as the incident day. For example, if an incident happens on Friday [8:00~9:00] at Line X, then a normal day can be all Fridays in the last 2 months where there are no substantial incidents occurring during [8:00$-\delta_1$ ~ 9:00$+\delta_2$] at Line X.

Note that we use the term "no substantial incidents" due to the high frequency of various types of incidents in a public transit system and it may be hard to find an "absolute normal day" without any incidents. The selection of normal days is a trade-off between sample sizes and accuracy. A larger number of normal days can provide more observations to estimate the habitual behaviors of

Table 1
Notation.

Variable	Type	Description
Z	Constant	Total number of behaviors considered.
M	Constant	Total number of normal days considered.
N_{S_i}	Random variable	Number passengers in set S_i.
$\mathcal{P}$	Set	The set of all potentially affected passengers.
$\mathcal{P}^H$	Set	The set of passengers with reliable history trips.
$\mathcal{P}^F$	Set	The set of passengers with future trips after the incident on that day.
S_i	Set	The set of passengers with behavior i (see Fig. 2).
B_{S_i}	Set	A set of passengers that is defined to infer S_i.
o_{p_k}	Constant	Origin of the kth trip for passenger p within the analysis period.
t_{p_k}	Constant	Tap-in time of the kth trip for passenger p within the analysis period.
m_{p_k}	Constant	Travel mode of the kth trip for passenger p within the analysis period.
K_p	Constant	Total number of public transit trips for passenger p within the analysis period.
M_p	Constant	Total number of normal days passenger p has public transit trips.
T_s	Constant	Start time of the analysis period.
T_e	Constant	End time of the analysis period.
T_1	Constant	Incident start time.
T_2	Constant	Incident end time.
TT_d	Constant	Threshold to identify transfer trips for two consecutive tap-ins.
d_r	Constant	Maximum walking distance for transferring to a rail station.
d_b	Constant	Maximum walking distance for transferring to a bus.
$D(s, s')$	Constant	A function which returns the walking distance between station s and s'.
$\mathbb{1}_{\{\cdot\}}$	Constant or Random variable	Indicator function which returns 1 if the event is true and 0 otherwise.
$\mathcal{W}$	Set	The set of all blocked rail stations during the incident.
$\mathcal{W}_b$	Set	The set of bus stops within walking distance from any of the blocked stations.
$\mathcal{W}_r$	Set	The set of all unblocked rail stations within walking distance from any of the blocked stations.
$\tilde{d}_{p_k}$	Random variable	Inferred original destination for trip k for passenger p.
D_{p_k}	Set	The set of all possible original destinations for trip k for passenger p.
$s_p(T, d)$	Constant	Location of passenger p at time T if his/her destination is d.

passengers. However, it may also include days with incidents that can introduce bias. Usually, we aim to have normal days with consistent demand and supply characteristics, and are significantly different from the incident day (as shown in the case study, Section 4.4).

Suppose that we have collected the AFC data of the incident day and a total of M normal days. Let $\mathcal{P}$ be the set of all potentially affected passengers, which is defined as the set of all passengers with at least one AFC data record in $[T_s, T_e]$ on the incident day or any of the M normal days. Let $\mathcal{P}^H \subseteq \mathcal{P}$ be a subset of passengers with reliable history trips and M_p be the number of normal days that passenger p has trips on ($M_p \leq M$). Then $\mathcal{P}^H = \{p \in \mathcal{P} : M_p \geq M^R\}$, which means passengers with more than M^R normal days with travel, where M^R is a predetermined threshold to recognize passengers with reliable history trips. In future studies, a more complicated method to define $\mathcal{P}^H$ can be explored considering the travel regularity (Goulet-Langlois et al., 2017).

Consider a passenger $p \in \mathcal{P}$ with a public transit trip chain $\{ (o_{p_1}, t_{p_1}, m_{p_1}), (o_{p_2}, t_{p_2}, m_{p_2}), \ldots, (o_{p_{K_p}}, t_{p_{K_p}}, m_{p_{K_p}}) \}$ within the analysis time period. o_{p_k} is the origin of the kth trip. t_{p_k} is the start time (transaction time) of the kth trip. And m_{p_k} is the mode of kth trip ($m_k \in \{\text{rail, bus}\}$). It holds that $T_s \leq t_{p_1} < t_{p_2} < \cdots < t_{p_{K_p}} \leq T_e$. We define $\mathcal{P}^F \subseteq \mathcal{P}$ as the subset of passengers with subsequent trips after the incident on the incident day, that is, $\mathcal{P}^F = \{p \in \mathcal{P} : p \text{ has trips after } T_e \text{ on the incident day}\}$. According to previous destination estimation studies for tap-in only systems (Barry et al., 2002; Zhao et al., 2007; Gordon et al., 2013), the destination of the trip $(o_{p_k}, t_{p_k}, m_{p_k})$ can be inferred using information of the next trip $(o_{p_{k+1}}, t_{p_{k+1}}, m_{p_{k+1}})$ (i.e., the trip chain method). The basic idea is to use the next tap-in location to estimate the destination of the current trip. Hence, for $p \in \mathcal{P}^F$, we can obtain the destination of the trip $(o_{p_{K_p}}, t_{p_{K_p}}, m_{p_{K_p}})$. It is worth noting that if the incident happened in the evening, we would extend $\mathcal{P}^F$ to include passengers with trips in the next morning.

As mentioned above, when a disruption happens, some of the stations in the rail system are blocked. The set of all blocked rail stations due to the disruption is denoted as $\mathcal{W}$.

The notation used in this study is summarized in Table 1.

2.2.2. Conceptual framework

We first outline the framework of the general inference model. For a specific behavior S_i, we define B_{S_i} as the set of passengers with related **observable** behavior that can be identified from the AFC data. The word "observable" indicates that $\mathbb{1}_{\{p \in B_{S_i}\}}$ is a known constant, where $\mathbb{1}_{\{\cdot\}}$ is an indicator function which returns 1 if the event is true and 0 otherwise. For example, B_{S_i} can be a set of passengers with a bus transfer trip during the incident period, or a set of passengers with a rail tap-in trip during the incident period, etc. The definition of B_{S_i} should satisfy that $S_i \subseteq B_{S_i}$. The goal is to identify S_i from B_{S_i}.

The specification of B_{S_i} depends on to what extent passengers in S_i can be observed in the AFC data. If the behavior of S_i generates many special AFC records, B_{S_i} can be defined in more detail. In this case, $|B_{S_i}|$ is relatively small, which reduces the scope for inferring S_i. On the other hand, if the behavior of S_i does not generate special AFC records, B_{S_i} can only be defined in a general way (e.g., passengers with a rail trip during the incident), bringing challenges in extracting S_i.

According to the context of S_i, there are two types of B_{S_i} regarding their relationship to the incident. For a passenger $p \in B_{S_i}$, historical information can be used to infer whether the behavior that passenger p is showing in B_{S_i} is atypical or not. However, "atypical" may not be enough to conclude whether p is affected by the incident or not. For example, B_{S_i} may be defined as passengers with a bus trip during the incident period. "atypical" only indicates the bus trip is a change of the passenger's habitual behavior. However, the behavioral change on that particular day may be caused by many reasons, not necessarily the incident. To conclude $p \in S_i$, p's behavior needs to satisfy both "atypical" and "change is due to the incident". This type of B_{S_i} is referred to as "indirect incident-related". However, sometimes, if B_{S_i} is specified based on a lot of information related to the incident, we can infer that "atypical" is equivalent to "affected by the incident". For example, if B_{S_i} are passengers with a transfer to bus stops close to the blocked rail stations after the incident, and this behavior is atypical, we can assume this change is due to the incident because B_{S_i} is based on direct incident-related information (i.e., the transfer bus stops are close to the blocked rail stations). This type of B_{S_i} is referred to as "direct incident-related".

Besides historical information, the subsequent trips information after the incident can also be used. As mentioned before, the subsequent tap-in information can be used to infer trip destinations using the trip chain method (Barry et al., 2002; Zhao et al., 2007; Gordon et al., 2013). Though recent studies also use historical information to infer passenger's destination (Cheng et al., 2020), for the purpose of this study, only subsequent tap-in information is used as the destination estimation part is not the focus of this study. Note that the proposed probabilistic framework is quite general and any destination estimation model can be used as long as the probability of each candidate destination can be obtained (see Section 3.2 for details). Although passengers may have multiple path choices in rail systems (Mo et al., 2021), we assume that all passengers follow the schedule-based shortest path to simplify the formulation (Barry et al., 2009). This assumption can be relaxed by summing over all paths with corresponding path choice probabilities in the formulation, instead of only considering a single path. For a passenger p, we obtain his/her original path in the rail system as the shortest path to the inferred destination d. Based on the characteristics of the path (explained below), we define a related event $Y_p(d)$. Since the path is known given d, $\mathbb{1}_{\{Y_p(d)\}}$ is a known constant. S_i can thus be inferred based on $\mathbb{1}_{\{Y_p(d)\}}$ (i.e. the property of the original path). For example, B_{S_i} can be a set of passengers without transfer trips during the incident period. $Y_p(d)$ can be the event that the original path for p is blocked and a transfer is not available. Then if $Y_p(d)$ is true, a passenger $p \in B_{S_i}$ can only use other undetected modes or cancel trips.

In summary, historical trips and subsequent trips after the incident are two types of available information to infer S_i. Therefore, from model formulations perspective, we can characterize the inference model in two dimensions: (1) historical trip information vs. subsequent trip information, (2) indirect incident-related B_{S_i} vs. direct incident-related B_{S_i}. We summarize the formulation in each case as follows:

- **(1) "Historical trip information + direct incident-related B_{S_i}"**: In this case, we have "atypical" = "affected by the incident". Therefore,

$$\mathbb{E}[\mathbb{1}_{\{p \in S_i\}}] = \mathbb{1}_{\{p \in B_{S_i}\}} \cdot \mathbb{P}(\text{"Behavior atypical"} \mid p \in B_{S_i}) \tag{1}$$

  S_1, S_2, S_4, and S_{12} belong to this case. $\mathbb{P}(\text{"Behavior atypical"} \mid p \in B_{S_i})$ is estimated based on the context of S_i using historical trip information. Details of the formulation can be found in Section 3.1.
- **(2) "Historical trip information + indirect incident-related B_{S_i}"**: In this case, we need to satisfy both "atypical" and "the change is due to the incident" in order to identify S_i. Therefore,

$$\mathbb{E}[\mathbb{1}_{\{p \in S_i\}}] = \mathbb{1}_{\{p \in B_{S_i}\}} \cdot \mathbb{P}(\text{"Behavior atypical"}, \text{"Change is due to the incident"} \mid p \in B_{S_i}) \tag{2}$$

  $S_5, S_{11}, S_{14}, S_{15}, S_{17}, S_{18}$, and S_{19} belong to this case. The joint probability $\mathbb{P}(\cdot, \cdot \mid p \in B_{S_i})$ is not estimated directly. Instead, we show that it can be estimated using the difference of the marginal probabilities between normal and incident days. Details of the formulation can be found in Section 3.3.
- **(3) "Subsequent trip information only"**: In this case, the event of path properties (as a function of the inferred destination d) can help to identify S_i. Hence,

$$\mathbb{E}[\mathbb{1}_{\{p \in S_i\}}] = \sum_d \mathbb{1}_{\{p \in B_{S_i}\}} \cdot \mathbb{1}_{\{Y_p(d)\}} \cdot \mathbb{P}(\text{"Original destination is } d\text{"} \mid p \in B_{S_i}) \tag{3}$$

  S_3, S_7, S_{10}, and S_{16} belong to this case. Some behavior assumptions are made when two groups are indistinguishable by the above formulation. $\mathbb{P}(\text{"Original destination is } d\text{"} \mid p \in B_{S_i})$ is estimated based on a destination inference model (Gordon et al., 2013) with subsequent trip information. Details of the formulation can be found in Section 3.4.
- **(4) "Historical trip information + direct incident-related B_{S_i} + Subsequent trip information"**: This scenario is a combination of historical and future information. Hence, we combine Eqs. (1) and (3):

$$\mathbb{E}[\mathbb{1}_{\{p \in S_i\}}] = \sum_d \mathbb{1}_{\{p \in B_{S_i}\}} \cdot \mathbb{P}(\text{"Behavior atypical"} \mid p \in B_{S_i})$$
$$\cdot \mathbb{1}_{\{Y_p(d)\}} \cdot \mathbb{P}(\text{"Original destination is } d\text{"} \mid p \in B_{S_i}) \tag{4}$$

  S_8 and S_9 belong to this case. Details of the formulation can be found in Section 3.2.

The above cases and the corresponding formulations are used to infer whether a specific passenger belongs to a certain group. The expected number of passengers in the group is calculated as

$$\mathbb{E}[N_{S_i}] = \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in S_i\}}] \tag{5}$$

It is worth noting that there are no explicit criteria to assign the inference of S_i to one of the four cases. There is a trade-off between including more information and dealing with sample sparsity. For example, one may argue that both historical and subsequent trip information should be included for all inferences. However, many passengers do not have reliable history trips or future trips (i.e. $p \notin \mathcal{P}^H \cap \mathcal{P}^F$). The inference for those passengers can only be approximated by the results of $p \in \mathcal{P}^H \cap \mathcal{P}^F$ (details in Section 3). Hence, simply including more information will lead to higher approximation errors due to sample sparsity, which is the reason that we have four types of formulations and some of them only include either future or history information, but not both. Determining the formulation for an S_i needs empirical knowledge and numeral tests to judge which kinds of information are more critical for the inference.

2.2.3. Uncertainty

In this study, we estimate the variance of the N_{S_i} ($\text{Var}[N_{S_i}]$) to quantify the uncertainty. $\text{Var}[N_{S_i}]$ captures the behavioral randomness of passengers in B_{S_i}. The behavior of a passenger in B_{S_i} is atypical or not (i.e., $\mathbb{1}_{\{\text{"Behavior atypical"}|p \in B_{S_i}\}}$) is an indicator random variable. High behavioral randomness indicates high variance of N_{S_i} because we cannot easily conclude whether a passenger's observed behavior in the incident day is typical or not. In this case, $\mathbb{P}(\text{"Behavior atypical"} \mid p \in B_{S_i})$ is close to 0.5 (where $\text{Var}[N_{S_i}]$ reaches the maximum), which implies that the passenger's behavior pattern is hard to estimate from the historical trips.

Besides passengers' inherent travel irregularity, $\text{Var}[N_{S_i}]$ is also determined by the definition of B_{S_i}. If B_{S_i} is specified narrowly, such as a set of passengers with a transfer trip to bus stops near the blocked rail stations after the incident, passengers may seldom have this "complicated" behavior on normal days. If a passenger has this behavior in the incident day, it is highly likely to be atypical (i.e., $\mathbb{P}(\text{"Behavior atypical"} \mid p \in B_{S_i})$ is close to 1). In this case, the $\text{Var}[N_{S_i}]$ is relatively low. However, if B_{S_i} has a very broad definition, such as a set of passengers with a bus trip in the incident period, $\mathbb{P}(\text{"Behavior atypical"} \mid p \in B_{S_i})$ may be close to 0.5 because passengers may use different modes on different normal days and it is difficult to infer having a bus trip is atypical or not on the incident day. In this case, the $\text{Var}[N_{S_i}]$ is relatively high. Since the definition of B_{S_i} is according to S_i, $\text{Var}[N_{S_i}]$ provides the information about whether S_i is easy to be inferred by the AFC data or not (low variance means S_i can be inferred more precisely).

3. Model formulation

In this section, we elaborate on the inference formulation for every behavior group. The section is organized by the formulation cases mentioned in Section 2.2.2. However, due to the tedious derivations and some formulation duplication, we only present the formulations for a part of behavior groups. The complete formulations can be found in Appendix.

3.1. Historical trip information + direct incident-related B_{S_i}: Inferring S_1 and S_2

By definition, passengers in S_1 and S_2 have at least one rail tap-in record before T_1 because they were in the blocked stations/trains when the incident happened. Since passengers who decide to use the public transit system again after alighting need to re-tap in, passengers in S_1 have another bus tap-in record after T_1, and passengers in S_2 have another rail tap-in record after T_1.

As passengers in S_1 and S_2 left the rail system from the blocked stations, the re-tap-in bus/rail stations should be close to the blocked stations and the time difference between two consecutive tap-ins should not be too large. Otherwise, they may be two separate trips instead of a transfer. Let TT_d be the tap-in time difference threshold for transferring. We assume that if $t_{p_k} - t_{p_{k-1}} < TT_d$, trip k is a transfer trip following trip $k-1$.[2] Denote the walking distance threshold for passengers transferring to a bus (resp. rail) as d_b (resp. d_r). Then the set of bus (resp. rail) stops close to the blocked stations is defined as $\mathcal{W}_b = \{s : s \text{ is a bus station and } \exists s' \in \mathcal{W} \text{ s.t. } D(s, s') \leq d_b\}$ (resp. $\mathcal{W}_r = \{s : s \notin \mathcal{W} \text{ is a rail station and } \exists s' \in \mathcal{W} \text{ s.t. } D(s, s') \leq d_r\}$), where $D(s, s')$ returns the walking distance between stations s and s'.

To identify passengers in S_1, we define a passenger set $B_{S_1} = \{p : \exists k \in \{1, \ldots, K_p - 1\} \text{ s.t. } t_{p_k} \leq T_1 < t_{p_{k+1}}, t_{p_{k+1}} - t_{p_k} < TT_d, m_{p_k} = \text{rail}, m_{p_{k+1}} = \text{bus}, o_{p_{k+1}} \in \mathcal{W}_b\}$. B_{S_1} represents passengers with a rail tap-in record before the incident and a bus transferring tap-in record after the incident. And the second tap-in station is within the walking distance of the blocked stations. As we described above, passengers in S_1 should also in B_{S_1} ($S_1 \subseteq B_{S_1}$). However, S_1 and B_{S_1} are not necessarily equivalent because passengers in B_{S_1} may transfer to a bus stop as a normal routine, that is, they did not transfer to a bus line in response to the rail disruption. Denote the event that p was affected by the incident as A_p. Then we have

$$\mathbb{E}[N_{S_1}] = \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in S_1\}}] = \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in B_{S_1}\}} \cdot \mathbb{1}_{\{A_p | p \in B_{S_1}\}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_1}\}} \cdot \mathbb{P}(A_p \mid p \in B_{S_1}) \tag{6}$$

---

[2] This is a typical way for tap-in only public transit systems to determine transfer trips for fare calculation. Future study may include tap-out time estimation model to better define a transfer trip.

Note that $\mathbb{1}_{\{p\in B_{S_1}\}}$ is a constant because for every p, we observe whether it belongs to B_{S_1} or not using the AFC data from the incident day. $\mathbb{P}(A_p \mid p \in B_{S_1})$ is calculated as

$$\mathbb{P}(A_p \mid p \in B_{S_1}) = 1 - \underbrace{\frac{\#\text{ normal days } p \text{ showing trip records described in } B_{S_1}}{M_p}}_{\text{Prob. that transferring to a bus stop near blocked station is a typical behavior}} \qquad \forall p \in \mathcal{P}^H. \tag{7}$$

Eq. (7) means that given a passenger with the observed behavior described in B_{S_1} on the incident day, the probability that this behavior is atypical[3] equals to 1 minus the relative frequency that the passenger has the same behavior on normal days. For example, if p transferred to a bus stop in $\mathcal{W}_b$ on every normal day, then transferring to the bus stop in $\mathcal{W}_b$ is highly likely to be a routine, rather than a change in behavior due to the incident (i.e., $\mathbb{P}(A_p \mid p \in B_{S_1}) = \frac{0}{M_p} = 0$). Then, p will not be counted into S_1.

If history information of p is unavailable or very limited (i.e., $p \notin \mathcal{P}^H$), Eq. (7) may fail to work. In this scenario, we assume

$$\mathbb{P}(A_p \mid p \in B_{S_1}) = \frac{\sum_{p' \in \mathcal{P}^H} \mathbb{P}(A_{p'} \mid p' \in B_{S_1})}{|\mathcal{P}^H \cap B_{S_1}|} \qquad \forall p \notin \mathcal{P}^H \tag{8}$$

which estimates the corresponding probability of passengers with little historical information using that of passengers with enough historical information. This is a typical way to estimate behavior of passengers without enough information in the AFC data (Gordon et al., 2013), though it may be biased considering different behavior patterns for $p \in \mathcal{P}^H$ and $p \notin \mathcal{P}^H$. There is no better way to address this issue given data limitations.

As $\mathbb{1}_{\{A_p|p\in B_{S_1}\}}$ is a Bernoulli random variable with probability $\mathbb{P}(A_p \mid p \in B_{S_1})$, the corresponding variance of N_{S_1} can be calculated as

$$\begin{aligned}
\mathrm{Var}[N_{S_1}] &= \sum_{p \in \mathcal{P}} (\mathbb{1}_{\{p\in B_{S_1}\}})^2 \cdot \mathrm{Var}[\mathbb{1}_{\{A_p|p\in B_{S_1}\}}] \\
&= \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p\in B_{S_1}\}} \cdot [\mathbb{P}(A_p \mid p \in B_{S_1}) - \mathbb{P}(A_p \mid p \in B_{S_1})^2]
\end{aligned} \tag{9}$$

Similarly, for passengers in S_2, similarly, we can define $B_{S_2} = \{p : \exists k \in \{1, \ldots, K_p - 1\} \text{ s.t. } t_{p_k} \le T_1 < t_{p_{k+1}}, t_{p_{k+1}} - t_{p_k} < TT_d, m_{p_k} = \text{rail}, m_{p_{k+1}} = \text{rail}, o_{p_{k+1}} \in \mathcal{W}_r\}$. Then we have

$$\mathbb{E}[N_{S_2}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p\in B_{S_2}\}} \cdot \mathbb{P}(A_p \mid p \in B_{S_2}) \tag{10}$$

where $\mathbb{P}(A_p \mid p \in B_{S_2})$ can be calculated in the same way as Eqs. (7) and (8) by replacing B_{S_1} with B_{S_2}. And the variance of N_{S_2} can be calculated as

$$\mathrm{Var}[N_{S_2}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p\in B_{S_2}\}} \cdot [\mathbb{P}(A_p \mid p \in B_{S_2}) - \mathbb{P}(A_p \mid p \in B_{S_2})^2] \tag{11}$$

3.2. Historical trip information + direct incident-related B_{S_i} + subsequent trip information: Inferring S_8 and S_9

Passengers in groups S_8 and S_9 continued to use the public transit system after the incident. Hence, they have at least one tap-in record before T_1 and at least one tap-in record after T_1. The difference between S_8, S_9 and S_1, S_2 is that passengers in S_8 and S_9 leave the rail system at some upstream station before the blocked stations. To differentiate S_8 and S_9 with other normal transfer passengers, we need to infer their original route and consider whether the route is blocked. If their original routes are not blocked, the transfers are not due to the incident.

We first identify S_8. Consider a passenger $p \in \mathcal{P}^F$. Suppose $\exists k \in \{1, \ldots, K_p - 1\}$ s.t. $t_{p_k} \le T_1 < t_{p_{k+1}}, t_{p_{k+1}} - t_{p_k} < TT_d, m_{p_k} = \text{rail}, m_{p_{k+1}} = \text{bus}$, and $o_{p_{k+1}} \notin \mathcal{W}_b$, which means p has a rail trip before the incident and a bus transfer trip after the incident, and the boarding stop of bus trip is not close to the blocked stations (otherwise he/she is already considered in the inference of S_1). If p was affected by the incident, the transferring trip $k + 1$ would be an atypical behavior for p.

Denote the event "transferring is atypical for p" as TA_p. Let $(o_{p_{k*}}, t_{p_{k*}}, m_{p_{k*}})$ be the next non-transfer trip of trip $k + 1$. Mathematically, $k* = \min\{k' > k + 1 : t_{p_{k'}} - t_{p_{k+1}} > TT_d\}$. Given TA_p, if without any incident, the original trip chain for passenger p is $\{\ldots, (o_{p_k}, t_{p_k}, m_{p_k}), (o_{p_{k*}}, t_{p_{k*}}, m_{p_{k*}}), \ldots\}$, the observed transfer bus trip $(o_{p_{k+1}}, t_{p_{k+1}}, m_{p_{k+1}})$ is caused by the disruption. Our goal is to use trip $k*$ to infer the original destination of trip k (i.e. the destination under normal condition). This can be done from the destination estimation model using the trip chain method (Barry et al., 2002; Zhao et al., 2007; Gordon et al., 2013). Let the set of all possible *original* destinations for trip k be $\mathcal{D}_{p_k}$ and $\tilde{d}_{p_k}$ the random variable representing the original destination of trip k. The destination estimation model provides $\mathbb{P}(\tilde{d}_{p_k} = d)$ for any $d \in \mathcal{D}_{p_k}$.

However, trip $k*$ may not exist for some p because the subsequent trip information may not be available (e.g., $p \notin \mathcal{P}^F$). For $p \notin \mathcal{P}^F$, the destination distribution can be approximated by $p \in \mathcal{P}^F$ (Gordon et al., 2013):

$$\mathbb{P}(\tilde{d}_{p_k} = d) = \frac{\sum_{p' \in \mathcal{P}^F : o_{p_k} = o_{p'_k}} \mathbb{P}(\tilde{d}_{p'_k} = d)}{|\{p' \in \mathcal{P}^F : o_{p_k} = o_{p'_k}\}|} \qquad \forall p \notin \mathcal{P}^F, d \in \mathcal{D}_{p_k}. \tag{12}$$

---

[3] Formulation type 1, "atypical" = "affected by the incident" in this case.

Eq. (12) means that the probability of $\tilde{d}_{p_k} = d$ for $p \notin \mathcal{P}^F$ is estimated as the average value of $p \in \mathcal{P}^F$ with the same origin.

As we assume that, for a given $\tilde{d}_{p_k}$, passengers follow the shortest path (Barry et al., 2009), the original route for p from o_{p_k} to $\tilde{d}_{p_k}$ can be obtained. Using automated vehicle location (AVL) data and a transit loading model (Zhu et al., 2017; Mo et al., 2020), we can further infer the location of passenger p in the rail system at time T_1 for a given $\tilde{d}_{p_k}$. Suppose that at time T_1, p was in location $s_p(T_1, \tilde{d}_{p_k})$ (which corresponds to a station or some middle point between two stations). Then, if the remaining route segment from $s_p(T_1, \tilde{d}_{p_k})$ to $\tilde{d}_{p_k}$ was blocked, p would be affected by the incident. Let the event that the original route of p is blocked given the original destination is d be $RB_p(d)$.

We define $B_{S_8} = \{p : \exists k \in \{1, \ldots, K_p - 1\} \text{ s.t. } t_{p_k} \leq T_1 < t_{p_{k+1}}, \; t_{p_{k+1}} - t_{p_k} < TT_d, \; m_{p_k} = \text{rail}, \; m_{p_{k+1}} = \text{bus}, o_{p_{k+1}} \notin \mathcal{W}_b\}$, which represents passengers with a rail tap-in record before the incident and a bus transferring tap-in record after the incident. Then we have $\mathbb{1}_{\{p \in S_8\}} = \mathbb{1}_{\{p \in B_{S_8}\}} \cdot \mathbb{1}_{\{TA_p\}} \cdot \sum_{d \in D_{p_k}} \mathbb{1}_{\{RB_p(d)\}} \cdot \mathbb{1}_{\{\tilde{d}_{p_k} = d\}}$. Note that $\mathbb{1}_{\{TA_p\}}$ and $\mathbb{1}_{\{RB_p(d)\}}$ are independent because the former is determined by the historical trips while the later is determined by the subsequent trips after the incident. Therefore, the number of passengers in S_8 can be calculated as:

$$\mathbb{E}[N_{S_8}] = \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in S_8\}}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_k}} \mathbb{1}_{\{p \in B_{S_8}\}} \cdot \mathbb{P}(TA_p \mid p \in B_{S_8}) \cdot \mathbb{1}_{\{RB_p(d)\}} \cdot \mathbb{P}(\tilde{d}_{p_k} = d) \tag{13}$$

$\mathbb{1}_{\{RB_p(d)\}}$ is a constant because given the original destination and path, we can conclude whether the path is blocked or not. $\mathbb{P}(TA_p \mid p \in B_{S_8})$ can be calculated in the same way as Eqs. (7) and (8) by replacing B_{S_1} and A_p with B_{S_8} and TA_p, respectively.

The variance of N_{S_8} can be calculated as

$$\text{Var}[N_{S_8}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_k}} \mathbb{1}_{\{p \in B_{S_8}\}} \cdot \mathbb{1}_{\{RB_p(d)\}} \cdot \left[\mathbb{P}(TA_p \mid p \in B_{S_8}) \cdot \mathbb{P}(\tilde{d}_{p_k} = d) - \mathbb{P}(TA_p \mid p \in B_{S_8})^2 \cdot \mathbb{P}(\tilde{d}_{p_k} = d)^2 \right] \tag{14}$$

Similarly, for passengers in S_9, we have $B_{S_9} = \{p : \exists k \in \{1, \ldots, K_p - 1\} \text{ s.t. } t_{p_k} \leq T_1 < t_{p_{k+1}}, \; t_{p_{k+1}} - t_{p_k} < TT_d, \; m_{p_k} = \text{rail}, \; m_{p_{k+1}} = \text{rail}, o_{p_{k+1}} \notin \mathcal{W}_r\}$. Then:

$$\mathbb{E}[N_{S_9}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_k}} \mathbb{1}_{\{p \in B_{S_9}\}} \cdot \mathbb{P}(TA_p \mid p \in B_{S_9}) \cdot \mathbb{1}_{\{RB_p(d)\}} \cdot \mathbb{P}(\tilde{d}_{p_k} = d) \tag{15}$$

$$\text{Var}[N_{S_9}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_k}} \mathbb{1}_{\{p \in B_{S_8}\}} \cdot \mathbb{1}_{\{RB_p(d)\}} \cdot \left[\mathbb{P}(TA_p \mid p \in B_{S_9}) \cdot \mathbb{P}(\tilde{d}_{p_k} = d) - \mathbb{P}(TA_p \mid p \in B_{S_9})^2 \cdot \mathbb{P}(\tilde{d}_{p_k} = d)^2 \right] \tag{16}$$

3.3. Historical trip information + indirect incident-related B_{S_i}: Inferring S_5 and S_{11}

S_5 and S_{11} are people who were already in the rail system and decided to cancel their trips because of the rail disruption. The AFC records of passengers in S_5 and S_{11} can be described as $B_{S_{5,11}} = \{p : t_{p_{K_p}} \leq T_1, \; m_{p_{K_p}} = \text{rail}\}$, which means passengers having at least one rail tap-in record before T_1 and no tap-in record between T_1 and T_e.

Consider a passenger $p \in \mathcal{P}^F \cap \mathcal{P}^H$. Let $(o_{p_{k*}}, t_{p_{k*}}, m_{p_{k*}})$ be the next non-transfer trip following trip K (i.e., $k^* = \min\{k' > K : t_{p_{k'}} - t_{p_{K_p}} > TT_d\}$). As k^* is the next non-transfer trip right after K, p had no non-transfer trips within $[t_{p_{K_p}}, t_{p_{k*}}]$ on the incident day. We use an example to illustrate the AFC records that may help to identify S_5 and S_{11}. Consider a passenger who plans to go to the supermarket on the incident day. He/she was in the system when the incident happened. Suppose that he/she decided to cancel his/her trip and return home. Then he/she would not have the typical returning trip from the supermarket. In this situation, k^* may be some other trips late in the evening or the first trip in the next day. However, in the historical AFC records. the typical trip right after K_p should be the returning trip from the supermarket. Therefore, we can assume that if passenger p has high probability of having trips within $[t_{p_{K_p}}, t_{p_{k*}}]$ on normal days, he/she is very likely to cancel the trip K_p because the typical following trip for K_p that is supposed to occur in $[t_{p_{K_p}}, t_{p_{k*}}]$ does not exist on the incident day.

However, it is worth noting that since we only have public transit trip records, passengers who do not cancel trips but use other travel modes to replace both trip K_p and the returning trip may also be identified as "cancel trips". Consider the example above, if a passenger takes Uber to the supermarket and then takes Uber back. He/she would be identified as "cancel trips". However, the information in AFC data is not enough to differentiate these two groups of passengers. Hence, in this study, we assume that the incident only changes passengers' mode choices of trips in the analysis period, which implies that the returning trip travel mode for the passenger will be public transit if he/she usually uses public transit. Note that this assumption can be relaxed if we focus on estimating the number of passengers "not using public transit" in an aggregated framework (see Fig. 2).

Denote the event that passenger $p \in B_{S_{5,11}}$ canceled trip K_p after the incident as CT_p. Based on the assumption above, we can derive the probability as

$$\mathbb{P}(CT_p \mid p \in B_{S_{5,11}}) = 1 - \frac{\# \text{ normal days } p \text{ has rail trips in } [T_s, T_1] \text{ with origin } o_{p_{K_p}} \text{ but no trip in } [t_{p_{K_p}}, t_{p_{k*}}]}{\# \text{ normal days } p \text{ has rail trips in } [T_s, T_1] \text{ with origin } o_{p_{K_p}}} \quad \forall p \in \mathcal{P}^H \cap \mathcal{P}^F \tag{17}$$

The second term in Eq. (17) represents the conditional probability that there is no trip in $[t_{p_{K_p}}, t_{p_{k*}}]$ on normal days given that the passenger already has a rail trip in $[T_s, T_1]$ with origin $o_{p_{K_p}}$. The lower is this probability, the higher is the probability that this behavior is atypical (i.e. the passenger actually cancels his/her trip) on the incident day.

For $p \notin \mathcal{P}^H \cap \mathcal{P}^F$, similar to Eq. (8), we can approximate the probability by that of passengers in $\mathcal{P}^H \cap \mathcal{P}^F$:

$$\mathbb{P}(CT_p \mid p \in B_{S_{5,11}}) = \frac{\sum_{p' \in \mathcal{P}^H \cap \mathcal{P}^F} \mathbb{P}(CT_{p'} \mid p \in B_{S_{5,11}})}{|\mathcal{P}^H \cap \mathcal{P}^F \cap B_{S_{5,11}}|} \quad \forall p \notin \mathcal{P}^H \cap \mathcal{P}^F \tag{18}$$

As mentioned before, passengers may cancel trips due to many reasons, not necessarily because of the incidents. Therefore, we need to consider the event $CT_p \cap A_p$, which represents passengers canceling trips because of the incident. However, directly calculating $\mathbb{P}(CT_p, A_p \mid p \in B_{S_{5,11}})$ is difficult. The following equations show an aggregate calculation approach:

$$\begin{aligned}
\mathbb{E}[N_{S_5} + N_{S_{11}}] &= \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in S_5 \cup S_{11}\}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{5,11}}\}} \cdot \mathbb{P}(CT_p, A_p \mid p \in B_{S_{5,11}}) \\
&= \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{5,11}}\}} \cdot \mathbb{P}(CT_p \mid p \in B_{S_{5,11}}) \cdot \mathbb{P}(A_p \mid CT_p, p \in B_{S_{5,11}}) \\
&= \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{5,11}}\}} \cdot \mathbb{P}(CT_p \mid p \in B_{S_{5,11}})(1 - \mathbb{P}((A_p)^c \mid CT_p, p \in B_{S_{5,11}})) \\
&= N_{CT} - \tilde{N}_{CT}
\end{aligned} \tag{19}$$

where

$$N_{CT} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{5,11}}\}} \cdot \mathbb{P}(CT_p \mid p \in B_{S_{5,11}}) \tag{20}$$

$$\tilde{N}_{CT} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{5,11}}\}} \cdot \mathbb{P}(CT_p \mid p \in B_{S_{5,11}}) \cdot \mathbb{P}((A_p)^c \mid CT_p, p \in B_{S_{5,11}}) \tag{21}$$

N_{CT} is the expected number of passengers who canceled trips on the incident day (not necessarily due to the incident) and $\tilde{N}_{CT}$ is the expected number of passengers who canceled trips on the incident day and the reason is not the incident. We can approximate $\tilde{N}_{CT}$ as the number of passengers canceling trips on normal days. Specifically, denote $\tilde{N}_{CT}^{(j)}$ as the number of canceling-trip passengers calculated by applying Eq. (20) to the AFC data of jth normal day. Then we have

$$\mathbb{E}[N_{S_5} + N_{S_{11}}] = N_{CT} - \tilde{N}_{CT} = N_{CT} - \frac{\sum_{j=1}^M \tilde{N}_{CT}^{(j)}}{M} \tag{22}$$

To calculate the variance of $N_{S_5} + N_{S_{11}}$, we assume

$$\mathbb{P}(A_p \mid CT_p, p \in B_{S_{5,11}}) = \frac{N_{CT} - \tilde{N}_{CT}}{N_{CT}}. \quad \forall p \in B_{S_{5,11}} \tag{23}$$

Eq. (23) means the probability that p's behavior is atypical given that he/she canceled trips equals the expected number of passengers canceling trips due to the incident divided by the total expected number of passengers canceling trips (not necessary due to the incident). It implies that we are using population statistics to approximate the individual probability. Then, we can calculate the variance as:

$$\text{Var}[N_{S_5} + N_{S_{11}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{5,11}}\}} \cdot \left[\mathbb{P}(CT_p \mid p \in B_{S_{5,11}})\mathbb{P}(A_p \mid CT_p, p \in B_{S_{5,11}}) - \mathbb{P}(CT_p \mid p \in B_{S_{5,11}})^2 \mathbb{P}(A_p \mid CT_p, p \in B_{S_{5,11}})^2 \right] \tag{24}$$

It is worth noting that the number of passengers canceling trips due to the incident is expected to be small. Therefore, Eq. (22) may be smaller than zero due to variations in the AFC data. This means that there is no big difference between the number of canceling trips passengers (Eq. (20)) on the incident day and on normal days, implying the number of passengers who canceled trips due to the incident is negligible. In this situation, we simply let $\mathbb{E}[N_{S_5} + N_{S_{11}}] = 0$ and $\text{Var}[N_{S_5} + N_{S_{11}}] = 0$.

3.4. Subsequent trip information only: Inferring S_3, S_7 and S_{10}

Identifying S_3, S_7 and S_{10} is similar to identifying S_8 and S_9. The inference leverages the subsequent trip information to infer the original routes. We consider these three groups together because they have the same AFC records in the incident day (i.e., at least one rail tap-in record before T_1 and no tap-in record between T_1 and T_e.). We define the corresponding set as $B_{S_{3,7,10}} = \{p : t_{p_{K_p}} \leq T_1, m_{p_{K_p}} = \text{rail}\}$.

Passengers in S_7 are those who transfer at some upstream stations (not go out) if their original rail route is blocked. For p in $B_{S_{3,7,10}} \cap \mathcal{P}^F$, let k^* be his/her next non-transfer trip after trip K_p. Then, using the same way as in Section 3.2, we can infer the destination distribution for trip K_p (i.e. obtain $\mathbb{P}(\tilde{d}_{p_{K_p}} = d)$ for any $d \in D_{p_{K_p}}$), as well as their locations when the incident happened (i.e. $s_p(T_1, \tilde{d}_{p_{K_p}})$). For a given $\tilde{d}_{p_{K_p}}$, if the original route from $s_p(T_1, \tilde{d}_{p_{K_p}})$ to $\tilde{d}_{p_{K_p}}$ is blocked, as we do not observe another tap-in record in $[T_1, T_2]$, p would only have three options: (1) transferring to alternative routes from $s_p(T_1, \tilde{d}_{p_{K_p}})$ to $\tilde{d}_{p_{K_p}}$ without going out of the rail system (S_7), (2) using other undetected modes ($S_3 + S_{10}$), and (3) canceling the trip ($S_5 + S_{11}$). This section focuses on the first two behaviors. It is worth noting that passengers can transfer only if there exist alternative routes from $s_p(T_1, \tilde{d}_{p_{K_p}})$ to $\tilde{d}_{p_{K_p}}$

within the rail system. Given an inferred original destination $d \in D_{p_{K_p}}$, we denote the event that p's original route is blocked but transfer is available as $RBTA_p(d)$.

We assume that passengers would not cancel trips when alternative routes were available. Then, if $RBTA_p(d)$ was true, p could either use intra-system transferring or use other undetected modes. However, given the data limitations, there is no available information to differentiate these two behaviors. We, thus, assume that the probability of using rail if a transfer is available, given the destination d of passenger p, is $\alpha_{p,d}$, that is, $\mathbb{P}(\mathbb{1}_{\{URTA_p|d\}} = 1) = \alpha_{p,d}$, where $URTA_p|d$ is the event that passenger p will use rail if a transfer is available given the destination is d. $\alpha_{p,d}$ can be estimated using a discrete choice model (DCM) with the utility expressed as a function of the travel cost, travel time of different travel modes (including transfer by rail, TNC, etc.) (Ben-Akiva et al., 1985). Given d, travel cost and travel time of different travel modes can be obtained from the Google Map API, and the parameters in the DCM can be estimated from survey data (Rahimi et al., 2020).

Notice that $\mathbb{1}_{\{URTA_p|d\}}$ is independent of $\mathbb{1}_{\{\tilde{d}_{p_K} = d\}}$ (i.e., the conditional independence). And based on the above assumptions, we have

$$\mathbb{E}[N_{S_7}] = \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in S_7\}}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_K}} \mathbb{1}_{\{p \in B_{S_{3,7,10}}\}} \cdot \mathbb{1}_{\{RBTA_p(d)\}} \cdot \alpha_{p,d} \cdot \mathbb{P}(\tilde{d}_{p_K} = d) \tag{25}$$

And the corresponding variance is

$$\text{Var}[N_{S_7}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_K}} \mathbb{1}_{\{p \in B_{S_{3,7,10}}\}} \cdot \mathbb{1}_{\{RBTA_p(d)\}} \cdot \left[\alpha_{p,d} \cdot \mathbb{P}(\tilde{d}_{p_K} = d) - \alpha_{p,d}^2 \cdot \mathbb{P}(\tilde{d}_{p_K} = d)^2\right] \tag{26}$$

Passengers in $B_{S_{3,7,10}}$ whose original routes were blocked and a transfer is *not* available have two options: (1) using other undetected modes or (2) canceling trips. Hence, we can use the total number of transfer-unavailable passengers minus the number of canceling-trip passengers to represent passengers using other undetected modes ($S_3 + S_{10}$). Note that when a transfer is available, passengers with probability $1 - \alpha_{p,d}$ may choose other undetected modes, and should also be counted into $S_3 + S_{10}$. Given an inferred original destination $d \in D_{p_{K_p}}$, denote the event that p's original route is blocked and transfer is not available as $RBTN_p(d)$. Then,

$$\mathbb{E}[N_{S_3} + N_{S_{10}}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_{K_p}}} \mathbb{1}_{\{p \in B_{S_{3,7,10}}\}} \cdot \mathbb{1}_{\{RBTA_p(d)\}} \cdot (1 - \alpha_{p,d}) \cdot \mathbb{P}(\tilde{d}_{p_{K_p}} = d) +$$

$$\sum_{p \in \mathcal{P}} \sum_{d \in D_{p_{K_p}}} \mathbb{1}_{\{p \in B_{S_{3,7,10}}\}} \cdot \mathbb{1}_{\{RBTN_p(d)\}} \cdot \mathbb{P}(\tilde{d}_{p_{K_p}} = d) - \mathbb{E}[N_{S_5} + N_{S_{11}}] \tag{27}$$

The first term in Eq. (27) indicates passengers with available intra-system transfer routes but still choosing other undetected modes. The second term represents the total number of passengers without intra-system transfer routes. And the third term ($\mathbb{E}[N_{S_5} + N_{S_{11}}]$) is the number of passengers canceling trips, which is calculated in Section 3.3.

According to Section 3.3, $N_{S_5} + N_{S_{11}} = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{5,11}}\}} \cdot \mathbb{1}_{\{CT_p, A_p | p \in B_{S_{5,11}}\}}$. And $\mathbb{1}_{\{URTA_p|d\}} \cdot \mathbb{1}_{\{\tilde{d}_{p_{K_p}} = d\}}$ is independent of $\mathbb{1}_{\{CT_p, A_p | p \in B_{S_{5,11}}\}}$ because the choice behavior ($\mathbb{1}_{\{URTA_p|d\}}$) is estimated from survey data, the destination inference ($\mathbb{1}_{\{\tilde{d}_{p_{K_p}} = d\}}$) is based on subsequent trip information, while the estimation of canceling trips ($\mathbb{1}_{\{CT_p, A_p | p \in B_{S_{5,11}}\}}$) is based on historical trip information. So, the variance can be calculated as

$$\text{Var}[N_{S_3} + N_{S_{10}}]$$
$$= \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_K}} \mathbb{1}_{\{p \in B_{S_{3,7,10}}\}} \cdot \mathbb{1}_{\{RBTA_p(d)\}} \cdot [(1 - \alpha_{p,d}) \cdot \mathbb{P}(\tilde{d}_{p_K} = d) - (1 - \alpha_{p,d})^2 \cdot \mathbb{P}(\tilde{d}_{p_K} = d)^2] +$$
$$\sum_{p \in \mathcal{P}} \sum_{d \in D_{p_K}} \mathbb{1}_{\{p \in B_{S_{3,7,10}}\}} \cdot \mathbb{1}_{\{RBTN_p(d)\}} \cdot [\mathbb{P}(\tilde{d}_{p_K} = d) - \mathbb{P}(\tilde{d}_{p_K} = d)^2] + \text{Var}[N_{S_5} + N_{S_{11}}]$$

where $\text{Var}[N_{S_5} + N_{S_{11}}]$ is obtained in Section 3.3.

4. Case study

4.1. Chicago Transit system

We use data from the CTA transit system as the case study because this paper focuses on open public transit systems with only tap-in information and CTA is an open system.

CTA is the second-largest transit system in the United States, providing services in Chicago, Illinois, and some of its surrounding suburbs. The transit network consists of the Chicago "L" (rail) and CTA bus services. It operates 24 h each day and on an average weekday provides 0.84 and 0.81 million rides on buses and trains, respectively (CTA, 2019). The map of the CTA rail system is shown in Fig. 4. The rail system consists of eight lines (named by color) and the "Loop". The Loop, located in the Chicago downtown area, 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% of all weekday boardings on the CTA trains.

Two data sources are used in this study: the AFC transaction data and train tracker (or AVL) data. CTA's AFC system is entry-only as passengers only use their fare cards when entering a rail station or boarding a bus. No information about a trip's destination is directly provided. The AVL system provides trains' arrival/departure times at each station.
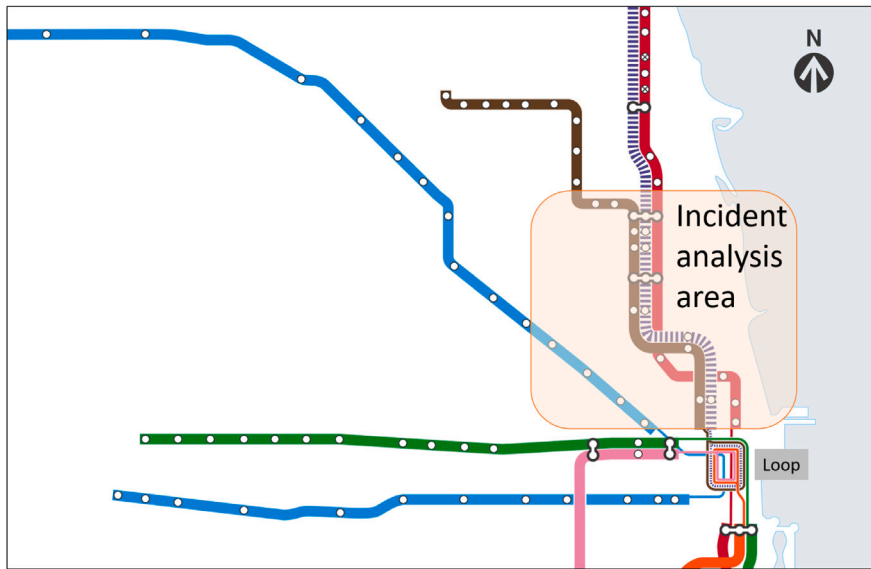
Fig. 4. CTA rail system map. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.2. Disruption background

The rail disruption used in this study happened on September 24, 2019. At 9:09AM, two trains collided at the Sedgwick station on the Brown line (see Figs. 4 and 5). This collision caused an interruption in service with five stations near Sedgwick on both Purple and Brown lines which are paralleled in this area being blocked (Fig. 5). The disruption lasted for 70 min and ended at 10:19 AM when trains returned to normal operations.

The reasons for choosing this incident are as follows: (1) It is a substantial unplanned service disruption that can trigger observable behavior changes. (2) The incident area has enough alternative services (such as nearby rail lines, bus routes, etc.) to cover 19 possible behaviors so that we can illustrate the proposed model's performance.

When the incident happened, passengers who were in blocked stations and trains were cleared out of the system. The station closure sign was placed outside the fare collection gate in blocked stations, reminding passengers about the service suspension. CTA informed passengers about the incident from both the Ventra app (CTA user app to manage and pay fares on CTA) and CTA Tweets right after the disruption. All passengers in the system were informed of train and platform announcements.

During the service interruption, CTA provided bus shuttle services between Fullerton and Merchandise Mart. People who were forced to leave their trains from the blocked stations would re-tap-in if they decided to use CTA normal bus or rail services and were only charged a small transfer fee.[4] However, no tap-in is needed for shuttle bus users. Hence, the shuttle bus is defined as an undetected mode in this study.

4.3. Parameter settings

Based on the incident information, the incident start time is $T_1 = 9$:09AM and end time $T_2 = 10$:19AM. $\delta_1 = \delta_2 = 60$ min is used according to the network scale and the analysis of system recovery time (Mo et al., 2022). Therefore, the analysis period is from $T_s = 8$:09AM to $T_e = 11$:19AM. The normal days are selected as all Fridays (except for the incident day) in September and October, 2019.

The time threshold for transferring $TT_d = 2$ h is used based on the CTA fare system. The walking distance threshold for bus and rail systems are set as $d_b = 0.7$ km and $d_r = 1.2$ km, respectively. These two numbers are slightly higher than the typical public transit transfer distance (Peng et al., 2009) so as to capture the increase in wiliness-to-walk during service disruptions.

As discussed before, $\alpha_{p,d}$ and β_p can be calculated based on the passenger's travel time and travel cost for different choices (including canceling trips) using DCM. The parameters in the DCM can be estimated from survey data or extracted from previous survey-based studies (Lin et al., 2018; Rahimi et al., 2020). The reason for using $\alpha_{p,d}$ and β_p is that, from AFC data alone, some groups of passengers cannot be identified as they have the same AFC transactions. AFC data only allows estimating $N_{S_3} + N_{S_7} + N_{S_{10}}$ (the number of passengers using intra-system transfers or not using public transit) and $N_{S_{17}} + N_{S_{18}}$ (the number of passengers out of the

---

[4] Sometimes there is no need to re-tap-in, depending on whether the control center has informed the CTA staff working in rail stations and bus drivers to allow free rides, and whether passengers asked for free rides due to the incident. In this study, we assume all passengers would re-tap-in according to the observation in the AFC data.
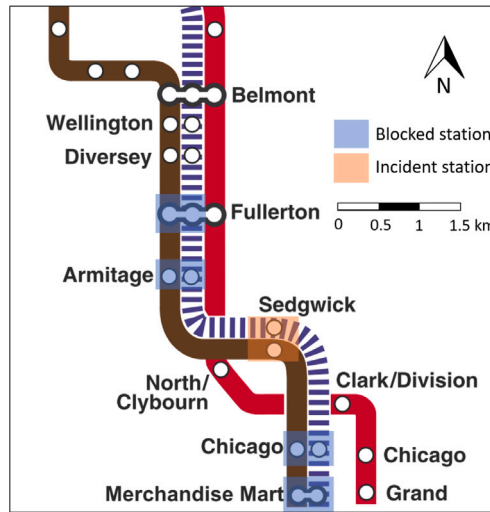
Fig. 5. Rail disruption case. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

system when the incident happens and not using public transit) as a whole. Model-based inferences are necessary for differentiating these groups. In this study, as we focus on a data-driven approach, the model-based parameters are set as $\alpha_{p,d} = 0.95$, $\beta_p = 0.9$ for all p and d for simplicity. These values are based on the sample statistics of CTA riders who participated in the survey about travel mode choices during incidents (Rahimi et al., 2020).

4.4. Descriptive analysis

For a better understanding of the incident, we show the demand patterns of three rail lines (Brown, Purple, and Red) and bus stations around the incident area (i.e., $\mathcal{W}_b$). The line-level demand is calculated as the sum of all station demands in the line.

Fig. 6 shows the comparison of the number of tap-in passengers on the incident day and normal days (aggregated by 15-min interval). We observe that the normal day demand patterns are relatively consistent compared to the incident day, which enables us to differentiate behavioral discrepancy on the incident day. As expected, the demand on the Brown and Purple Lines (interrupted by the incident) both decreased during the incident (Figs. 6a and 6b). And it gradually returned to normal with the end of the incident. As the Red Line runs adjacent to the Brown and Purple Lines for a large portion (see Fig. 5) in the incident area and it is not suspended, we see a significant increase in demand during the incident period with a return to normal after the incident is over (Fig. 6c). In terms of the nearby bus stops, the demand pattern is similar to that of the Red Line.

In terms of the demand change numbers, we see that the demand increase on the Red Line (1413) is much higher than that in the nearby bus stations, implying that most of the passengers choose the Red Line as the alternative. Note that the total demand decrease in Brown and Purple lines is slightly smaller than the total demand increase due to the fact that some passengers may first tap into the incident lines and then leave. This means that the actual demand decrease is higher than $680 + 506 = 1186$.

4.5. Rule-based benchmark models

We choose the rule-based deterministic method that has been used in previous studies (Sun et al., 2016; Liu et al., 2021) as the benchmark model. The rule-based method directly maps passengers with observed behavior (B_{S_i}) to those who are influenced by the incident (S_i). Note that as this paper considers different behavior sets from those of previous studies, we cannot use their rules to classify passengers. For a fair comparison, we use the rule defined in our paper (B_{S_i}) as the criterion. Recall that there are four formulations in Section 2.2.2 to infer passengers' responses. For the rule-based model, the number of formulations reduces to two because some cases share the same formulation:

- **"Historical trip information + direct incident-related B_{S_i}"** and **"Historical trip information + indirect incident-related B_{S_i}"**: In the rule-based method, we assume $p \in B_{S_i}$ is equivalent to $p \in S_i$. Therefore, eliminating the probability component in Eqs. (1) and (2), we have

$$\mathbb{1}_{\{p \in S_i\}} = \mathbb{1}_{\{p \in B_{S_i}\}} \tag{28}$$

$S_1, S_2, S_4, S_{12}, S_5, S_{11}, S_{14}, S_{15}, S_{17}, S_{18}$, and S_{19} belong to this case.

(a) Brown Line

(b) Purple Line
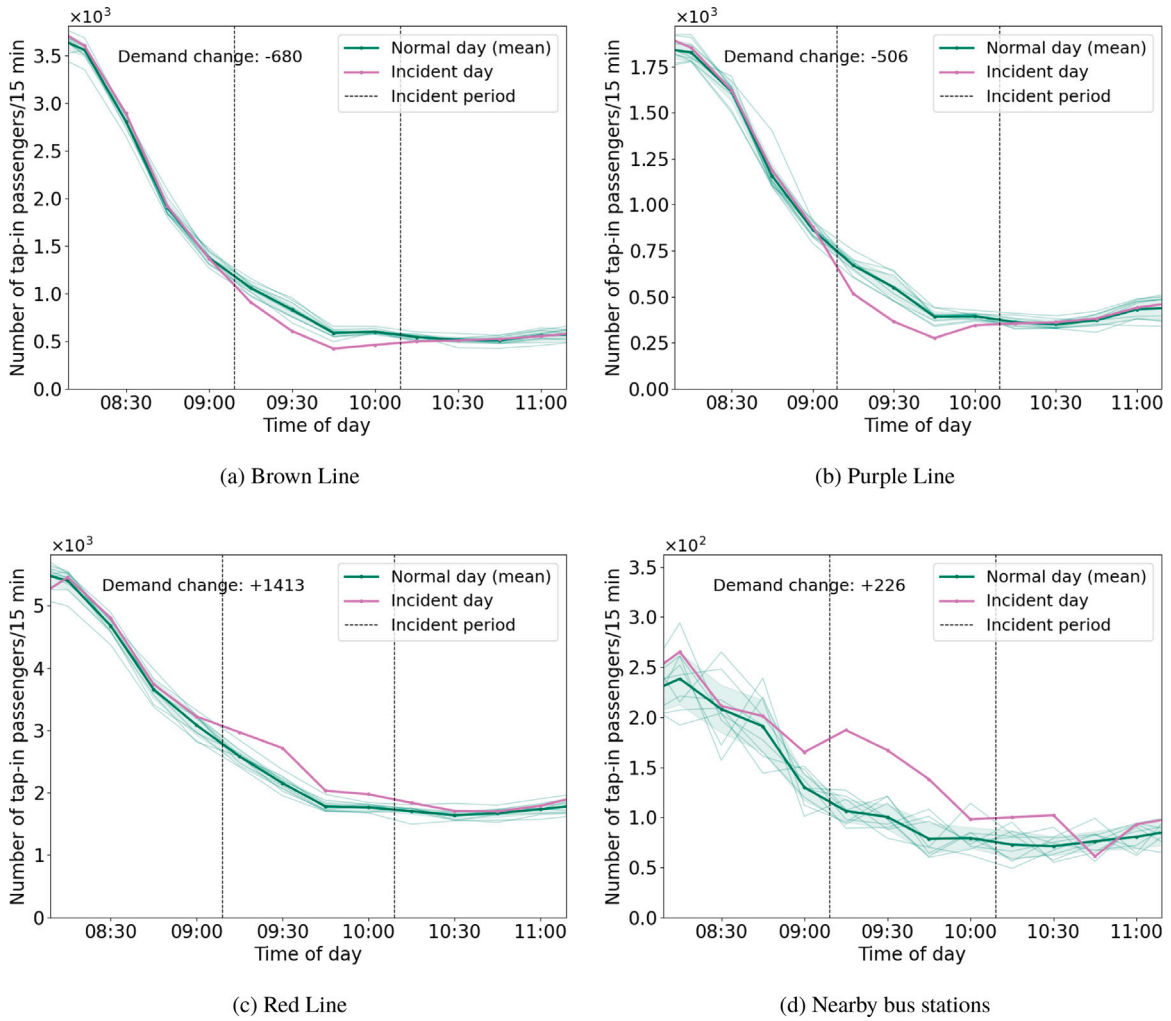

(c) Red Line

(d) Nearby bus stations

Fig. 6. Demand comparison for normal days and the incident day. A green thin line represents the demand curve for a single normal day. The green shade areas represent the ±standard deviation. The demand change is calculated as the total number of tap-ins during the incident period (9:09–10:19 AM) on the incident day minus that of the normal day average. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- "**Subsequent trip information only**" and "**Historical trip information + direct incident-related** B_{S_i} **+ Subsequent trip information**": In this case, we first infer a destination d for the passenger. Then, eliminating the probability component in Eqs. (3) and (4), we have

$$\mathbb{1}_{\{p \in S_i\}} = \mathbb{1}_{\{p \in B_{S_i}\}} \cdot \mathbb{1}_{\{Y_p(d)\}} \tag{29}$$

$S_3, S_7, S_{10}, S_{16}, S_8,$ and S_9 belong to this case.

The estimated number of passengers in group S_i is calculated as.

$$\hat{N}_{S_i} = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in S_i\}} \tag{30}$$

Since this is a deterministic method, variance information is not available.

4.6. Model validation with synthetic data

4.6.1. Synthetic data generation

Since there are no available observations for passengers' actual choices, the model validation is conducted with a simulation-based synthetic data set generated from the actual AFC data. The generation process is as follows. And the illustration diagram is shown in Fig. 7.
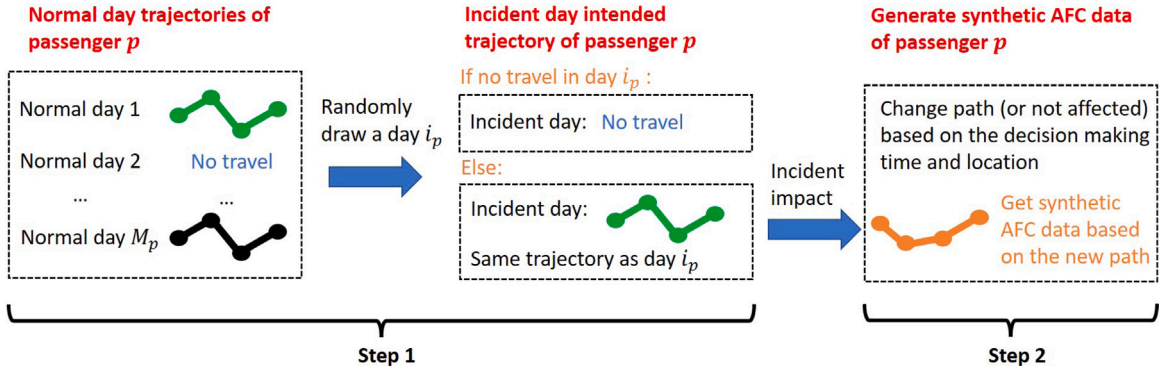
Fig. 7. Diagram for synthetic data generation.

Step 1: Sample intended trajectories. For each passenger p who has used the CTA system in any of M normal days (i.e., $M_p \geq 1$), we randomly sample one normal day ID (from 1 to M), denoted i_p. If the passenger does not have an AFC record on the i_pth normal day, we assume that he/she did not use public transit on the incident day. Otherwise, the AFC records on the i_pth normal day are treated as his/her intended trajectory. We assume that, on the incident day, passenger p would follow the same travel trajectory as the i_pth normal day (i.e., tap-in and tap-out records) if there was no incident. For all intended trajectories, the public transit trip destinations are inferred from the destination estimation model (Gordon et al., 2013).

Step 2: Generate synthetic AFC data for the incident. The data from step 1 are the passenger "intended" trajectories under normal conditions. We also need to generate the "actual" AFC records subject to the incident at Sedgwick station (see Section 4.2). Specifically, with the intended trajectories of all passengers, we can infer their locations when the incident occurs based on a transit assignment model (see Section 3.2). For the purpose of model validation, we assume that passengers' behavior follows the diagram in Fig. 2. From passengers' locations and intended routes, we can identify all affected and unaffected passengers based on whether their original routes are blocked or not. For all affected passengers, we first enumerate their possible choices based on the stage of the trip they are at when the incident occurs (e.g., at the blocked stations, in the system but not at blocked stations, outside the system, etc.) and availability of different travel modes. Then, each passenger is assigned an available mode based on the choice probabilities. For this application, the choice probabilities are calculated using the behavior model in Rahimi et al. (2020) and Lin et al. (2018). If the passenger is assigned with public transit, we find the available nearby bus or alternative rail lines for him/her and calculate his/her tap in time based on the walking distance. The new tap-in record is added to the synthetic data on the incident day. For passengers who decide to wait until the system recovers, we assume they all wait outside the blocked stations and tap in right after T_2. Then new AFC records are added to the synthetic data. If the passenger is out of the system when the incident happens and is assigned with an undetected travel mode or canceling trips, we remove his/her AFC transaction in $[T_s, T_e]$. For passengers in the system deciding to cancel their trips, we remove their subsequent AFC transactions (i.e., returning trips) as assumed in Section 3.3. The new AFC records are treated as synthetic data on the incident day (where the incident does happen).

The synthetic AFC data on the incident day and passengers' "true" choices are then used as the ground truth for model validation. Data generation and model estimation processes are replicated 15 times.

4.6.2. Validation criteria

Since the proposed model can output the expected number of passengers in each behavior group (i.e., $\mathbb{E}[N_{S_i}]$) and corresponding variance (i.e., $\mathrm{Var}[N_{S_i}]$), it is worth validating both estimates. The validation of $\mathbb{E}[N_{S_i}]$ is straightforward. As in the synthetic data we have the "true" value of N_{S_i}, a comparison between the "true" N_{S_i} and the estimated $\mathbb{E}(N_{S_i})$ can be conducted (For the benchmark model, the comparison is against $\hat{N}_{S_i}$). Since the data generation and model estimation processes are replicated 15 times, the "true" average of N_{S_i} and estimated $\mathbb{E}[N_{S_i}]$ are compared (Fig. 8).

To validate $\mathrm{Var}[N_{S_i}]$, we notice that the "true" N_{S_i} in each replication of the synthetic data can be seen as a sample drawn from the underlying behavioral distribution. This distribution is a reflection of passenger's choice probabilities and inferred destination distribution. Therefore, the sample variance of N_{S_i} over the 15 replications can be seen as the "true" $\mathrm{Var}[N_{S_i}]$, which is compared with the estimated $\mathrm{Var}[N_{S_i}]$ (Fig. 9). Note that since we have 15 estimated $\mathrm{Var}[N_{S_i}]$ from different replications, the average value is used for comparison.

To quantify the estimation errors over all behavior groups, we calculate the root mean square error (RMSE) and mean absolute percentage error (MAPE) as follows:

$$\mathrm{RMSE}(\mathbb{E}[\cdot]) = \sqrt{\frac{\sum_{i=1}^{Z}(\bar{N}_{S_i} - \bar{\mathbb{E}}[N_{S_i}])^2}{Z}}, \tag{31}$$

$$\mathrm{MAPE}(\mathbb{E}[\cdot]) = \frac{1}{Z}\sum_{i=1}^{Z}\frac{|\bar{N}_{S_i} - \bar{\mathbb{E}}[N_{S_i}]|}{\bar{N}_{S_i}}, \tag{32}$$
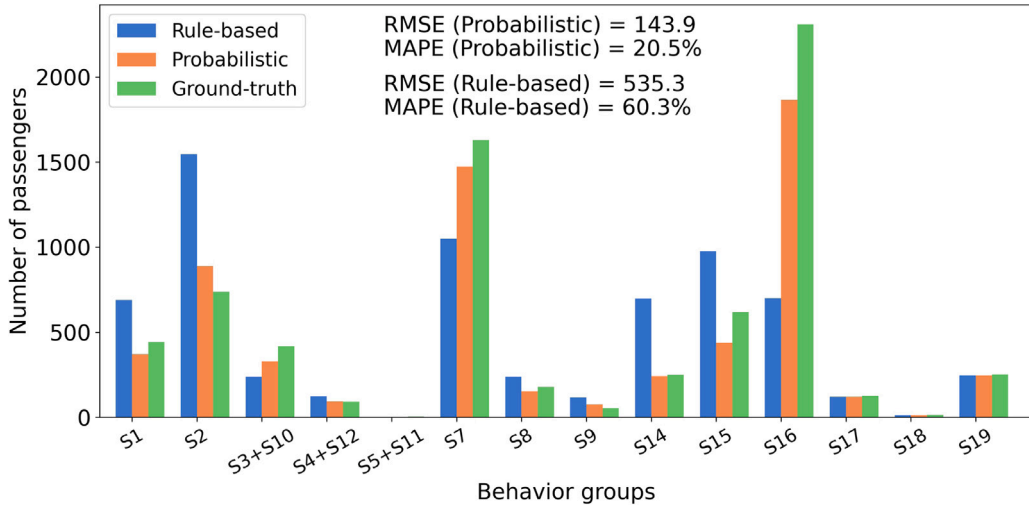
Fig. 8. Estimation results of expectations with synthetic data.

where $\bar{N}_{S_i}$ (resp. $\hat{\mathbb{E}}[N_{S_i}]$) is the average value of the "true" N_{S_i} (resp. estimated $\mathbb{E}[N_{S_i}]$) over the 15 replications. The RMSE and MAPE of $\text{Var}[N_{S_i}]$ are calculated in a similar way.

4.6.3. Results

Model estimation results with synthetic data are shown in Figs. 8 and 9. Note that we exclude the results of N_{S_6} and $N_{S_{13}}$ (number of not affected passengers) in the graph as their values are too large and may distort the comparison. In Fig. 9, the standard deviations (i.e., $\sqrt{\text{Var}[N_{S_i}]}$) are shown instead of variance for unit consistency.

Fig. 8 presents the estimated results of $\mathbb{E}[N_{S_i}]$ (probabilistic model) and $\hat{N}_{S_i}$ (rule-based). Results show that the probabilistic model can estimate passenger's response behaviors with an RMSE = 144 and MAPE = 20.5%. It significantly outperforms the rule-based benchmark model (RMSE = 536 and MAPE = 60.3%). The absolute errors of the probabilistic model are relatively large for $\mathbb{E}[N_{S_{16}}]$ and $\mathbb{E}[N_{S_7}]$. This may be due to the fact that there are around 30% passengers without future information for destination inference. Their destination distribution is approximated by the inferred population (Eq. (12)), leading to estimation errors. In terms of the rule-based model, it has a system error (overestimation) because it does not account for the fact that some observed behaviors are due to behavior randomness, rather than the impact of incidents.

Fig. 9 presents the estimation results for $\sqrt{\text{Var}[N_{S_i}]}$. Note that the rule-based model cannot output estimated variance, thus is not plotted in the figure. Results show that the probabilistic model can capture the patterns of standard deviation for different behavior groups well. The RMSE is 4.4 and MAPE is 69.8%, which is higher compared to the error of the expected values. This is reasonable because variance is the second moment which in general is harder to estimate than the first moment (i.e., expectation).

4.7. Model application with real-world data

4.7.1. Results

In the real-world data, we only implement the probabilistic method. Table 2 summarizes the estimation results for the real-world data from the CTA system. Overall, most of the passengers (97.43%) are not affected by the incident. This is reasonable because the incident only affected a small area. 69.51% of all affected passengers choose to use rail by changing routes. This is expected because the Red Line is a good substitution for the blocked Brown and Purple lines. Most of the OD pairs can be connected by the Red line when the Brown and Purple lines do not work. 6.57% of passengers choose to wait or delay their departure times (i.e., using rail without changing routes). 15.72% choose to use buses while 8.09% choose to not use public transit.

The variance in Table 2 captures the behavioral randomness of S_i and how much information of S_i can be captured in B_{S_i} by AFC data (see Section 2.2.3). Generally, the variances are proportional to the means. The coefficients of variation for N_{S_1} and N_{S_2} are low, meaning these two behaviors are relatively easy to be captured by AFC data. This is reasonable because multiple tap-in records are generated by these behaviors, leading to the direct incident-related B_{S_i}. Canceling trips and using other undetected modes have a relatively high coefficient of variation. This means these two behaviors are hard to be estimated using the AFC data.

Fig. 10 shows the behavior distribution for passengers **in** the rail system when the incident happened. 46% of those passengers choose the inside rail transfer (i.e. transfer without leaving the rail system). This is reasonable because passengers coming from stations north of the blocked stations (main morning peak demand) have multiple rail transfer stations (such as Belmont and Fullerton) that connect the suspended Brown and Purple lines to the Red line (see Fig. 5). This allows passengers to conveniently continue to use the rail system without leaving the system. 19% and 23% of passengers choose to leave the system and transfer to
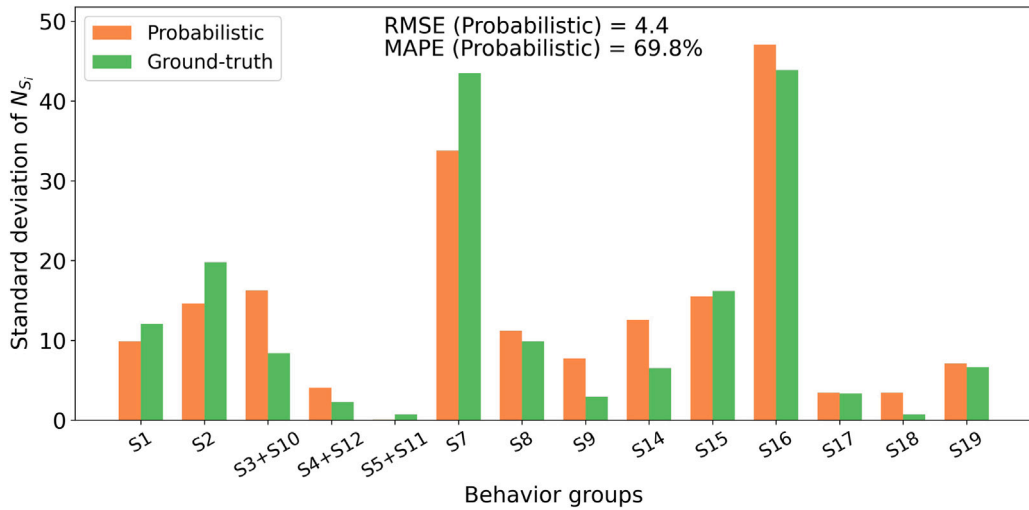
Fig. 9. Estimation results of standard deviations with synthetic data.

Table 2
Passenger behavior estimation results.

Behavior (Prop.; Impacted Prop.[b])	Group	Mean	Variance (Coeff. of variation[a], %)	Proportion (%)	Proportion (Impacted[b], %)
Use rail changing route (1.79%; 69.51%)	S2	595	157.4 (2.11)	0.25	9.61
	S7	1282	1005.7 (2.47)	0.53	20.71
	S9	56	49.0 (12.5)	0.02	0.9
	S15	831	675.5 (3.13)	0.35	13.43
	S16	1538	2639.4 (3.34)	0.64	24.85
Use rail not changing route (0.17%; 6.57%)	S4+S12	48	11.5 (7.07)	0.02	0.78
	S19	365	295.9 (4.71)	0.15	5.9
Use bus (0.40%, 15.72%)	S1	315	87.8 (2.97)	0.13	5.09
	S8	202	170.5 (6.46)	0.08	3.26
	S14	456	412.9 (4.46)	0.19	7.37
Not use public transit (0.21%; 8.09%)	S3+S10	291	255.8 (5.5)	0.12	4.7
	S17	180	180.2 (7.46)	0.07	2.91
	S5+S11	10	10.4 (32.18)	0	0.16
	S18	20	20.0 (22.39)	0.01	0.32
No impact (97.43%, N.A.)	S6	63 503	1748.1 (0.07)	26.37	N.A.
	S13	171 085	4223.9 (0.04)	71.06	N.A.

[a]Coefficient (Coeff.) of variation is calculated as the standard division divided by the mean.
[b]Impacted proportion (prop.) is the proportion within all affected passengers (excluding S_6 and S_{13}).

a bus line and other rail stations, respectively. Around 10% of passengers choose to use other undetected modes. And only a small proportion of passengers choose to wait (2%) or cancel their trips (0.3%). Overall, the estimated proportions of different behaviors are reasonable.

Fig. 11 shows the behavior distribution for all affected passengers **out of** the rail system when the incident happened. Similar to the results above, most of those passengers (45%) chose to transfer to another rail line without leaving the system. 25% of them changed tap-in stations and 13% chose to use buses. We also observe that 11% of passengers delayed their departure time and 5% used other undetected modes. Only around 1% of passengers canceled their trips. Compared to the results above, we find there is a decrease in the percentage of passengers using buses and other undetected modes and an increase in using rail. This is reasonable because when passengers are out of the system, they are more flexible in choosing rail routes, thus more likely to keep using rail services.

4.7.2. Analysis of real-world results

Though there is no direct validation for the estimation results using real-world data, we propose two indirect approaches to discuss the reasonableness of the results.

The first is to compare the ridership increase in bus stops and rail stations that are close to the blocked stations (i.e., $\mathcal{W}_b$ and $\mathcal{W}_r$). The ridership increase at these bus stops and rail stations should be similar to (slightly larger than) N_{S_1} and N_{S_2}, respectively.
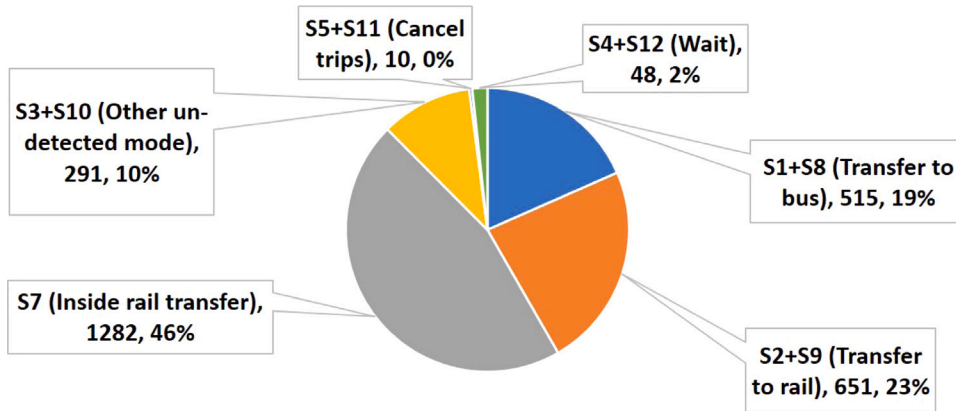
Fig. 10. Behavior distribution for passengers in the rail system when the incident happened (texts in boxes are behavior description, number of passengers, and proportion, respectively).
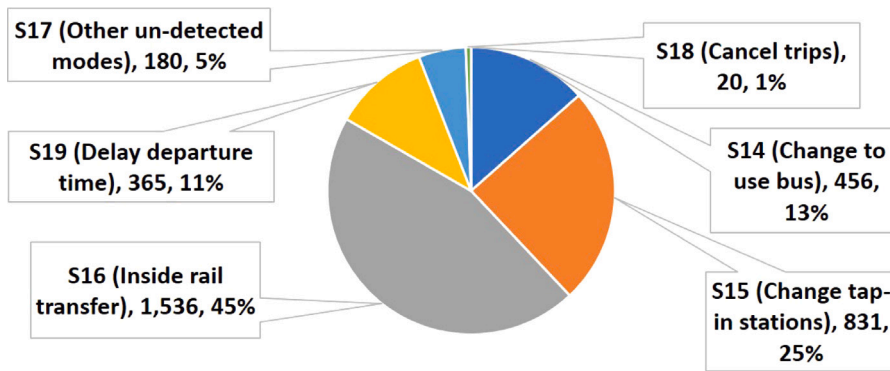

Fig. 11. Behavior distribution for passengers out of the rail system when the incident happened.

"Slightly larger" is because some ridership increase may be passengers living in the nearby neighborhoods, which do not belong to S_1 and S_2. The ridership increase is calculated as the number of tap-in passengers during the incident period minus the mean on normal days. The ridership increase for nearby bus stops is 401 passengers (slightly larger than the estimated $\mathbb{E}[N_{S_1}] = 315$), and for rail stations 720 passengers (slightly larger than the estimated $\mathbb{E}[N_{S_2}] = 595$), which is as expected.

The second approach is based on the CTA incident logs. CTA incident logs report that "run 505 (Purple line) unloads around 300 customers" and "run 416 (Brown line) unloads around 500 customers". According to the AVL data, these two trains are the only trains that unloaded passengers. Assuming that passengers who entered the blocked stations between T_1 and the time of the last train departure waited at the platforms, there were a total of 437 waiting passengers on the platforms of the blocked stations when the incident happened (based on the AFC and AVL data). According to Fig. 2, the total number of unloaded and waiting passengers should be equal to the number of passengers at the blocked stations (i.e., $\sum_{i=1}^5 N_{S_i}$). Hence, the estimated value of $\sum_{i=1}^5 \mathbb{E}[N_{S_i}]$ should be close to $300 + 500 + 437 = 1237$ passengers. However, the inference model provides estimates for $\mathbb{E}[N_{S_1}]$ and $\mathbb{E}[N_{S_2}]$, but not $\mathbb{E}[N_{S_3}]$, $\mathbb{E}[N_{S_4}]$, and $\mathbb{E}[N_{S_5}]$ (because $\mathbb{E}[N_{S_3} + N_{S_{10}}]$, $\mathbb{E}[N_{S_4} + N_{S_{12}}]$, and $\mathbb{E}[N_{S_5} + N_{S_{11}}]$ are estimated as a whole). Since $\mathbb{E}[N_{S_4} + N_{S_{12}}]$ and $\mathbb{E}[N_{S_5} + N_{S_{11}}]$ are relatively small, $\sum_{i=1}^5 \mathbb{E}[N_{S_i}] + \mathbb{E}[N_{S_{10}}] + \mathbb{E}[N_{S_{11}}] + \mathbb{E}[N_{S_{12}}]$ should be slightly greater than 1237 and $\mathbb{E}[N_{S_1}] + \mathbb{E}[N_{S_2}]$ slightly smaller than 1237 if the estimates are correct. A simple calculation leads to

$$\mathbb{E}[N_{S_1} + N_{S_2}] = 910 < 1237 < \sum_{i=1}^5 \mathbb{E}[N_{S_i}] + \mathbb{E}[N_{S_{10}}] + \mathbb{E}[N_{S_{11}}] + \mathbb{E}[N_{S_{12}}] = 1259, \qquad (33)$$

supporting the validity of the estimation results.

5. Conclusion and discussion

This study proposes a probabilistic framework to infer passengers' response behavior to an unplanned rail service disruption using smart card data in a tap-in-only public transit system. We enumerate 19 possible behaviors that passengers may have based on the stages of their trips when the incident happened. A probabilistic model is proposed to estimate the mean and variance of the number

of passengers in each of the 19 groups using passengers' historical and subsequent trip information. Based on the information used and the context of the behavior, four cases of formulations are used in the probabilistic model. Data from the CTA public transit system (bus and urban rail) is used for the case study with a rail incident. The model is implemented with both synthetic data (consistent with the CTA AFC data) and real-world data. The main conclusions of this study are as follows:

- The proposed approach can estimate passengers' behavior well and outperform the rule-based benchmark model. Results with synthetic data show that the RMSE and MAPE for the estimated expected number of passengers in each behavior group are 143.9 and 20.5%, respectively. The RMSE and MAPE for the estimated standard deviation are 4.4 and 69.8%, respectively. The estimation results with real-world data are consistent with the incident's context. An indirect model validation using ridership change information and incident log data demonstrates the reasonableness of the results.
- Results with real-world data find that most of the passengers (97.43%) are not affected by the incident. This is reasonable because the incident only affected a small area. The incident we analyzed has high service redundancy with the Red line substituting the blocked Brown and Purple lines. Our model results show that in the high redundancy case, most of the affected passengers (69.51%) choose to use rail by changing routes.
- Based on the results, CTA operators can confirm that the Red line is a good alternative and quantify the impact. To relieve the incident impact, operators can increase service frequency in the Red line. The model indicates that only 8.1% of passengers choose to leave the public transit system. This number can help CTA conduct the service loss analysis due to the incident.

The proposed model has several practical significances. First, The model is data-driven. Compared to the conventional survey-based methods, the proposed approach can effectively estimate passengers' responses without collecting data manually. Second, the output results can help transit operators better understand passengers' choices during a disruption, based on which they can design better operating strategies on the supply side to mitigate the impact of incidents. For example, for heavily used alternative services during the disruption, operators can increase the service frequency or provide shuttle buses with similar routes. Third, based on the results, operators can identify congestion in the network. They can disseminate information (e.g., route recommendation) to passengers, or conduct flow control at the gate level, to avoid overloaded routes.

Future studies can focus on the following directions. (1) Estimate the estimation error (i.e., $\text{Var}[\mathbb{E}[N_{S_i}]]$). The estimation error is another type of uncertainty. It comes from the fact that we are using sample data to estimate a specific probability. For example, as $\mathbb{P}(A_p \mid p \in B_{S_i})$ is estimated from the historical travel trajectories, we actually only obtain the estimated value (i.e., $\hat{\mathbb{P}}(A_p \mid p \in B_{S_i})$). It is a random variable and the corresponding variance $\text{Var}[\hat{\mathbb{P}}(\cdot)]$ reflects the estimation error. The estimation error depends on sample sizes (i.e., amount of normal day data) and passenger travel irregularity. The challenge of estimating $\text{Var}[\hat{\mathbb{P}}(\cdot)]$ is that this value is not available for passengers without historical information. Future studies can explore approximation techniques with reasonable distributional assumptions to calculate estimation errors. (2) Apply the model to different incident cases. According to Mo et al. (2022), the incident locations and the redundancy of surrounding public transit alternatives are influential in passenger mode choice behavior. Future studies may analyze more case studies and compare passengers' behavioral responses under different scenarios. (3) Analyze individual-level choices. In this study, we only output the aggregate level mode choice behavior (i.e., N_{S_i}). Future work may explicitly output $\mathbb{P}(p \in S_i)$, and analyze its relationship with passengers' characteristics (such as home location, fare card type, travel frequency, etc.).

These future studies can help improve the proposed method, and make the understanding of passenger responses more accurate. For example, with better quantification of estimation uncertainty, we can develop more robust or stochastic optimization methods for shuttle service design, headway design, path recommendations, flow control, etc.

Though there are extensive data in the AFC and AVL systems, machine learning methods do not fit into this study because of the lack of actual observed responses behaviors (i.e., lack of labels). In the future, if some passenger's actual response behavior can be observed (e.g., from self-report data, probe GPS data, or cell phone data), a supervised learning model may be trained to predict passengers' responses to incidents. The features may include passenger's spatial and temporal travel histories, incident information, and supply information. These features can be embedded with many advanced deep learning methods such as long short-term memory (LSTM) networks, convolutional neural networks (CNN), graph neural networks (GNN), etc.

CRediT authorship contribution statement

Baichuan Mo: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Haris N. Koutsopoulos:** Conceptualization, Formal analysis, Writing – review & editing, Supervision. **Jinhua Zhao:** Conceptualization, Formal analysis, Supervision, Project administration, Funding acquisition.

Acknowledgment

Appendix. Model formulations (continued)

A.1. Historical trip information + direct incident-related B_{S_i} (continued)

A.1.1. Inferring S_4 and S_{12}

S_4 and S_{12} are passengers who waited until the system recovered. Passengers in S_4 left and waited outside the blocked stations. Thus, they have at least one tap-in record before T_1, and another tap-in record at the blocked stations after T_2. We assume that passengers in S_{12} also waited outside the blocked stations (passengers usually take the train up to the blocked stations then start to wait).

We define $B_{S_{4,12}} = \{p : \exists k \in \{1, \ldots, K_p - 1\} \text{ s.t. } t_{p_k} \leq T_1, t_{p_{k+1}} \geq T_2, m_{p_k} = m_{p_{k+1}} = \text{rail}, o_{p_{k+1}} \in \mathcal{W}\}$, which means passengers with a rail tap-in trip before the incident and another rail tap-in trip after the system recovery, with the second tap-in station one of the blocked stations. As passengers who tap-in again at a blocked station may do so not because of the incident but as part of a normal routine, similar to Eq. (6),

$$\mathbb{E}[N_{S_4} + N_{S_{12}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{4,12}}\}} \cdot \mathbb{P}(A_p \mid p \in B_{S_{4,12}}). \tag{A.1}$$

$\mathbb{1}_{\{p \in B_{S_{4,12}}\}}$ is a constant. $\mathbb{P}(A_p \mid p \in B_{S_{4,12}})$ can be calculated the same way as Eqs. (7) and (8) by replacing B_{S_1} with $B_{S_{4,12}}$. The variance of $N_{S_4} + N_{S_{12}}$ is

$$\text{Var}[N_{S_4} + N_{S_{12}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{4,12}}\}} \cdot [\mathbb{P}(A_p \mid p \in B_{S_{4,12}}) - \mathbb{P}(A_p \mid p \in B_{S_{4,12}})^2] \tag{A.2}$$

A.2. Historical trip information + indirect incident-related B_{S_i} (continued)

A.2.1. Inferring S_{14} and S_{15}

Passengers in S_{14} and S_{15} have not entered the rail system when the incident happens. Therefore, they have no rail tap-in records before T_1.

We first consider passengers in S_{14}. Consider a $p \in \mathcal{P}^H$ with $T_1 < t_{p_1} < T_2$ and $m_{p_1} = \text{bus}$, which means the first trip for p during the incident period is bus. Define $B_{S_{14}} = \{p : T_1 < t_{p_1} < T_2, m_{p_1} = \text{bus}\}$. And define the event that p changed from rail to bus on the incident day as CB_p. The probability of CB_p for $p \in B_{S_{14}}$ can be calculated as

$$\mathbb{P}(CB_p \mid p \in B_{S_{14}}) = 1 - \frac{\# \text{ normal days } p\text{'s first trip in } [T_1, T_2] \text{ is bus}}{\# \text{ normal days } p \text{ has trips in } [T_1, T_2]} \qquad \forall p \in \mathcal{P}^H \tag{A.3}$$

Eq. (A.3) means the probability of CB_p equals 1 minus the frequency of using a bus on normal days. A high frequency of using a bus on normal days means using a bus is highly likely the typical behavior for p, instead of a change in the behavior. For $p \notin \mathcal{P}^H$, we can approximate the probability by

$$\mathbb{P}(CB_p \mid p \in B_{S_{14}}) = \frac{\sum_{p' \in \mathcal{P}^H \cap \mathcal{P}^F} \mathbb{P}(CB_{p'} \mid p \in B_{S_{14}})}{|\mathcal{P}^H \cap B_{S_{14}}|} \qquad \forall p \notin \mathcal{P}^H \tag{A.4}$$

However, passengers may change from rail to bus due to many reasons, not necessarily because of the incident. Similar to Eq. (19), we have

$$\mathbb{E}[N_{S_{14}}] = \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in S_{14}\}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{14}}\}} \cdot \mathbb{P}(CB_p, A_p \mid p \in B_{S_{14}}) \tag{A.5}$$

$$= N_{CB} - \tilde{N}_{CB}$$

where $N_{CB} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{14}}\}} \cdot \mathbb{P}(CB_p \mid p \in B_{S_{14}})$ is the expected number of passengers who change from rail to bus on the incident day. $\tilde{N}_{CB} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{14}}\}} \cdot \mathbb{P}(CB_p \mid p \in B_{S_{14}})\mathbb{P}((A_p)^c \mid CB_p, p \in B_{S_{14}})$ is the expected number of passengers who change from rail to bus but *not* because of the incident. It can be approximated by the number of passengers changing from rail to bus on normal days. Similar to Eq. (22),

$$\mathbb{E}[N_{S_{14}}] = N_{CB} - \tilde{N}_{CB} = N_{CB} - \frac{\sum_{j=1}^{M} \tilde{N}_{CB}^{(j)}}{M} \tag{A.6}$$

where $\tilde{N}_{CB}^{(j)}$ is the number of passengers changing from rail to bus on the jth normal day, calculated with the same method of calculating N_{CB} but using the jth normal day AFC data. Similar to Eqs. (23) and (24), we can calculate the variance of $N_{S_{14}}$ as:

$$\text{Var}[N_{S_{14}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{14}}\}} \cdot [\mathbb{P}(CB_p \mid p \in B_{S_{14}})\mathbb{P}(A_p \mid CB_p, p \in B_{S_{14}}) -$$

$$\mathbb{P}(CB_p \mid p \in B_{S_{14}})^2 \mathbb{P}(A_p \mid CB_p, p \in B_{S_{14}})^2] \tag{A.7}$$

where $\mathbb{P}(A_p \mid CB_p, p \in B_{S_{14}}) = (N_{CB} - \tilde{N}_{CB})/N_{CB}$ for all $p \in B_{S_{14}}$ (similar to Eq. (23)).

For passengers in S_{15}, we define $B_{S_{15}} = \{p : T_1 < t_{p_1} < T_2, m_{p_1} = \text{rail}\}$, and denote the event that p changes tap-in station to o_{p_1} on the incident day as CS_p. Similar to Eq. (A.3), we have

$$\mathbb{P}(CS_p \mid p \in B_{S_{15}}) = 1 - \frac{\text{\# normal days that } p\text{'s first rail tap-in station in } [T_1, T_2] \text{ is } o_{p_1}}{\text{\# normal days that } p \text{ has rail trips in } [T_1, T_2]} \qquad \forall p \in \mathcal{P}^H \tag{A.8}$$

Analogue to the estimation of $\mathbb{E}[N_{S_{14}}]$, we have

$$\mathbb{E}[N_{S_{15}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{15}}\}} \cdot \mathbb{P}(CS_p, A_p \mid p \in B_{S_{15}}) \tag{A.9}$$

$$= N_{CS} - \tilde{N}_{CS}$$

where $N_{CS} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{15}}\}} \cdot \mathbb{P}(CS_p \mid p \in B_{S_{15}})$ is the expected number of passengers changing tap-in stations on the incident day. And $\tilde{N}_{CS} := \frac{\sum_{j=1}^{M} \tilde{N}_{CS}^{(j)}}{M}$, where $\tilde{N}_{CS}^{(j)}$ is the number of passengers changing tap-in stations on the jth normal day, which is calculated with the same method as calculating N_{CS} using the AFC data on the jth normal day. The variance of $N_{S_{15}}$ can be calculated the same way as in Eq. (A.7) by replacing $B_{S_{14}}$ and CB_p by $B_{S_{15}}$ and CS_p, respectively.

A.2.2. Inferring S_{17} and S_{18}

Passengers who decided to use other undetected modes or cancel trips after the incident (i.e., S_{17} and S_{18}) have no rail tap-in records between T_1 and T_e on the incident day. The inference is based on passengers who were supposed to have tap-in records in this period according to their behavior on normal days. Define $B_{S_{17,18}} = \{p : p \text{ has rail tap-in records within } [T_1, T_e] \text{ on any of the } M_p \text{ normal days, but not on the incident day}\}$. These are potential passengers who might change to other undetected modes or cancel trips on the incident day. Due to the nature of the AFC data, there is no direct way to differentiate these two groups. We assume that the probability of a passenger p using other undetected modes in this situation is β_p, that is, $\mathbb{P}(\mathbb{1}_{\{UMOS_p\}} = 1) = \alpha_{p,d}$, where $UMOS_p$ is the event that passenger p will other undetected modes when he/she is outside the system. The value of β_p can be obtained from previous survey-based studies (similar to α_p). Note that if we focus on the aggregate estimation of the passengers who do not use public transit (i.e., canceling trips + using other undetected modes), the value of β_p is not needed.

Consider a passenger $p \in B_{S_{17,18}}$. As in Section 3.3, we assume that if p has a high probability of having trips in $[T_1, T_e]$ on normal days, then the disappearance of the trip on the incident day is highly likely an atypical behavior (i.e., canceling the trip or switching to undetected modes). Define the event that p canceled the trip or switched to undetected modes on the incident day as $CTSM_p$. According to the assumption above and Eq. (17):

$$\mathbb{P}(CTSM_p \mid p \in B_{S_{17,18}}) = \frac{\text{\# normal days } p \text{ having rail trips in } [T_1, T_e]}{M_p} \qquad \forall p \in \mathcal{P}^H \tag{A.10}$$

However, as p may cancel the trip or switch to other undetected modes for other reasons, not necessarily due to the incident. We have $\mathbb{1}_{\{p \in S_{17}\}} = \mathbb{1}_{\{p \in B_{S_{17,18}}\}} \cdot \mathbb{1}_{\{CTSM_p \cap A_p \mid p \in B_{S_{17,18}}\}} \cdot \mathbb{1}_{\{UMOS_p\}}$. Since $\mathbb{1}_{\{UMOS_p\}}$ and $\mathbb{1}_{\{CTSM_p \cap A_p \mid p \in B_{S_{17,18}}\}}$ are independent, similar to Eqs. (19)–(21), we have

$$\mathbb{E}[N_{S_{17}}] = \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in S_{17}\}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{17,18}}\}} \cdot \beta_p \cdot \mathbb{P}(CTSM_p, A_p \mid p \in B_{S_{17,18}})$$

$$= N_{CTSM1} - \tilde{N}_{CTSM1} \tag{A.11}$$

where $N_{CTSM1} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{17,18}}\}} \cdot \beta_p \cdot \mathbb{P}(CTSM_p \mid p \in B_{S_{17,18}})$ is the expected number of passengers using other undetected modes on the incident day (not necessarily due to the incident). $\tilde{N}_{CTSM1} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{17,18}}\}} \cdot \beta_p \cdot \mathbb{P}(CTSM_p \mid p \in B_{S_{17,18}}) \cdot \mathbb{P}((A_p)^c \mid CTSM_p, p \in B_{S_{17,18}})$ is the expected number of passengers using other undetected modes and the reason is not the incident day, which can be approximated by the number of passengers using other undetected modes on normal days: $\tilde{N}_{CTSM1} = \frac{\sum_{j=1}^{M} \tilde{N}_{CTSM1}^{(j)}}{M}$, where $\tilde{N}_{CTSM1}^{(j)}$ is the expected number of passengers using other undetected modes on the jth normal day, which is calculated with the same method as calculating N_{CTSM1} but with the AFC data on the jth normal day.

And the variance of N_{17} is

$$\text{Var}[N_{S_{17}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{17,18}}\}} \cdot [\beta_p \cdot \mathbb{P}(CTSM_p \mid p \in B_{S_{17,18}}) \cdot \mathbb{P}(A_p \mid CTSM_p, p \in B_{S_{17,18}}) - $$

$$\beta_p^2 \cdot \mathbb{P}(CTSM_p \mid p \in B_{S_{17,18}})^2 \cdot \mathbb{P}(A_p \mid CTSM_p, p \in B_{S_{17,18}})^2] \tag{A.12}$$

where

$$\mathbb{P}(A_p \mid CTSM_p, p \in B_{S_{17,18}}) = \frac{N_{CTSM} - \tilde{N}_{CTSM}}{N_{CTSM}}. \qquad \forall p \in B_{S_{17,18}} \tag{A.13}$$

$N_{CTSM}, \tilde{N}_{CTSM}$ are calculated the same way as $N_{CTSM1}, \tilde{N}_{CTSM1}$ by replacing β_p to 1.

Similarly, for passengers in S_{18}, $\mathbb{E}[N_{S_{18}}]$ and $\text{Var}[N_{S_{18}}]$ are calculated the same way as $\mathbb{E}[N_{S_{17}}]$ and $\text{Var}[N_{S_{17}}]$, respectively, by replacing β_p to $1 - \beta_p$.

A.2.3. Inferring S_{19}

Passengers in S_{19} are those who continued to use their original routes but delayed their departure times. In this study, we define "delay departure time" as departing $2\sigma_p$ later than μ_p, where μ_p is the mean departure time of p's first rail trip in the analysis period on the normal days, and σ_p is the corresponding standard deviation. μ_p and σ_p can be calculated using the tap-in times of previous rail trips at station o_{p_1} on normal days. We define $B_{S_{19}} = \{p : t_{p_1} \geq T_2, m_{p_1} = \text{rail}, t_{p_1} > \mu_p + 2\sigma_p\}$, which is the set of passengers who delayed their departure times and departed after T_2 (i.e., after system recovery). However, as passengers may delay departure time for different reasons, not necessarily because of the incidents, similar to Eq. (19), we have

$$\mathbb{E}[N_{S_{19}}] = \sum_{p \in \mathcal{P}} \mathbb{E}[\mathbb{1}_{\{p \in S_{19}\}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{19}}\}} \cdot \mathbb{P}(A_p \mid p \in B_{S_{19}}) \tag{A.14}$$
$$= \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{19}}\}}[1 - \mathbb{P}((A_p)^c \mid p \in B_{S_{19}})]$$
$$= N_{DD} - \tilde{N}_{DD}$$

where $N_{DD} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{19}}\}}$ is the expected number of passengers who delayed departure time on the incident day. And $\tilde{N}_{DD} := \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{19}}\}} \mathbb{P}((A_p)^c \mid p \in B_{S_{16}})$ is the expected number of passengers who delayed departure time but *not* because of the incident, which can be approximated by the number of passengers delaying departure time on normal days. Therefore, similar to Eq. (22), we have

$$\mathbb{E}[N_{S_{16}}] = N_{DD} - \tilde{N}_{DD} = N_{DD} - \frac{\sum_{j=1}^{M} \tilde{N}_{DD}^{(j)}}{M} \tag{A.15}$$

where $\tilde{N}_{DD}^{(j)}$ is the number of passengers delaying departure time on jth normal day, calculated with the same method of calculating N_{DD} but using the jth normal day AFC data. Similar to Eqs. (23) and (24), we can calculate the variance of $N_{S_{19}}$ as:

$$\text{Var}[N_{S_{19}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{19}}\}}[\mathbb{P}(A_p \mid p \in B_{S_{19}}) - \mathbb{P}(A_p \mid p \in B_{S_{19}})^2] \tag{A.16}$$

where $\mathbb{P}(A_p \mid p \in B_{S_{16}}) = (N_{DD} - \tilde{N}_{DD})/N_{DD}$ as per Eq. (23).

A.3. Subsequent trip information only (continued)

A.3.1. Inferring S_{16}

Passengers in S_{16} are those who did not change tap-in stations, but chose to transfer halfway to avoid the blocked stations. We assume that passengers who make decisions after the incident are informed of the service interruption. Hence, if they decided to still use rail between T_1 and T_2, the possible situations for them are (1) changing tap-in station (S_{15}), (2) choosing alternative routes by transferring (S_{16}), and (3) not affected. Let $B_{S_{16}} = \{p : T_1 < t_{p_1} < T_2, m_{p_1} = \text{rail}\}$, which means passengers with a rail trip during the incident time. We notice that the third possibility can be excluded if we find that a passenger's original path is blocked. Therefore, for all passengers in $\mathcal{P}^F$, we first infer their destinations based on the next non-transfer trip after $(t_{p_1}, o_{p_1}, m_{p_1})$ (see Section 3.2). Given an inferred destination $d \in D_{p_1}$, denote the event that p's original path is blocked but a transfer option is available as $RBTA_p(d)$. From the above analysis, we know that all passengers in $B_{S_{16}}$ and with $\mathbb{1}_{\{RBTA_p(d)\}} = 1$ can only be in S_{15} and S_{16}. Define $N_{B_{S_{16}} \cap RBTA} := \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_1}} \mathbb{1}_{\{p \in B_{S_{16}}\}} \cdot \mathbb{1}_{\{RBTA_p(d)\}} \cdot \mathbb{1}_{\{\tilde{d}_{p_1} = d\}}$, which is the number of passengers with a rail trip during the incident and the original route blocked. Therefore, the mean of $N_{S_{16}}$ can be calculated as:

$$\mathbb{E}[N_{S_{16}}] = \mathbb{E}[N_{B_{S_{16}} \cap RBTA} - N_{S_{15}}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_1}} \mathbb{1}_{\{p \in B_{S_{16}}\}} \cdot \mathbb{1}_{\{RBTA_p(d)\}} \cdot \mathbb{P}(\tilde{d}_{p_1} = d) - \mathbb{E}[N_{S_{15}}] \tag{A.17}$$

where $\mathbb{E}[N_{S_{15}}]$ is estimated as in Appendix A.2.1. To calculate $\text{Var}[N_{S_{16}}]$, we notice that the covariance between $N_{B_{S_{16}} \cap RBTA}$ and $N_{S_{15}}$ is zero:

$$\text{Cov}[N_{15}, N_{B_{S_{16}} \cap RBTA}]$$
$$= \text{Cov}\left[\sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{15}}\}} \cdot \mathbb{1}_{\{CS_p, A_p \mid p \in B_{S_{15}}\}}, \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_1}} \mathbb{1}_{\{p \in B_{S_{16}}\}} \cdot \mathbb{1}_{\{RBTA_p(d)\}} \cdot \mathbb{1}_{\{\tilde{d}_{p_1} = d\}}\right]$$
$$= 0 \tag{A.18}$$

This is based on the observation that $\text{Cov}[\mathbb{1}_{\{CS_p, A_p \mid p \in B_{S_{15}}\}}, \mathbb{1}_{\{\tilde{d}_{p'_1} = d\}}] = 0$ for all $p, p' \in \mathcal{P}$ (even if $p = p'$, this still holds because the derivation of destination relies on future information while the derivation of atypical behavior relies on historical information). Hence, the variance of $N_{S_{16}}$ can be estimated as.

$$\text{Var}[N_{S_{16}}] = \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_1}} \mathbb{1}_{\{p \in B_{S_{16}}\}} \cdot \mathbb{1}_{\{RBTA_p(d)\}} \cdot [\mathbb{P}(\tilde{d}_{p_1} = d) - \mathbb{P}(\tilde{d}_{p_1} = d)^2] + \text{Var}[N_{S_{15}}] \tag{A.19}$$

A.4. Other

Passengers in S_6 and S_{13} are those who are not affected by the incident. They are inferred based on the results of other groups, which do not belong to any formulation cases and thus are described separately in this section.

A.4.1. Inferring S_6 and S_{13}

Passengers in S_6 are those who were not affected by the incident even though they were in the rail system while the incident happened. According to the diagram in Fig. 2, we can infer N_6 as all passengers in the rail system subtracting other subgroups of passengers given the mutually exclusive definition. Define $B_{S_6} = \{p : \exists k \in \{1, \ldots, K_p\} \text{ s.t. } t_{p_k} < T_1, \ m_{p_k} = \text{rail}\}$, which means all passengers who might be in the rail system when the incident happened. Therefore, we have

$$\mathbb{E}[N_{S_6}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_6}\}} - \sum_{i=1}^{5} \mathbb{E}[N_{S_i}] - \sum_{i=7}^{12} \mathbb{E}[N_{S_i}] \tag{A.20}$$

Note that $\mathbb{E}[N_{S_3} + N_{S_{10}}]$, $\mathbb{E}[N_{S_4} + N_{S_{12}}]$, and $\mathbb{E}[N_{S_5} + N_{S_{11}}]$ are calculated as a whole (see Sections 3.4, 3.3, and Appendix A.1.1).

The calculation of variance needs to consider the possible correlation among N_{S_i}. First of all, B_{S_1}, B_{S_2}, B_{S_8}, B_{S_9}, and $B_{S_{4,12}}$ do not intersect with other B_{S_i}'s, which implies N_{S_1}, N_{S_2}, N_{S_8}, N_{S_9}, and $N_{S_4} + N_{S_{12}}$ are independent and they are also independent of other N_{S_i}'s (because the behavior of different passengers is assumed to be independent). As shown in Section 3.4, the inference of $N_{S_5} + N_{S_{11}}$ uses the historical trip while the inference of $N_{S_3} + N_{S_{10}}$ and N_{S_7} relies on the information of subsequent trips (after the incident). Hence, $N_{S_5} + N_{S_{11}}$ is independent of $N_{S_3} + N_{S_{10}}$ and N_{S_7}. Then, the variance of N_{S_6} can be calculated as

$$\text{Var}[N_{S_6}] = \sum_{i=1}^{2} \text{Var}[N_{S_i}] + \sum_{i=8}^{9} \text{Var}[N_{S_i}] + \text{Var}[N_{S_4} + N_{S_{12}}] + \text{Var}[N_{S_5} + N_{S_{11}}] + \text{Var}[N_{S_3} + N_{S_{10}} + N_{S_7}] \tag{A.21}$$

Note that the variance of $N_{S_3} + N_{S_{10}} + N_{S_7}$ can be calculated as a whole according to Section 3.4:

$$\begin{aligned}
&\text{Var}[N_{S_3} + N_{S_{10}} + N_{S_7}] \\
&= \sum_{p \in \mathcal{P}} \sum_{d \in D_{p_K}} \mathbb{1}_{\{p \in B_{S_{3,7,10}}\}} \cdot [\mathbb{1}_{\{RBTA_p(d)\}} + \mathbb{1}_{\{RBTN_p(d)\}}] \cdot [\mathbb{P}(\tilde{d}_{p_K} = d) - \mathbb{P}(\tilde{d}_{p_K} = d)^2] + \text{Var}[N_{S_5} + N_{S_{11}}]
\end{aligned} \tag{A.22}$$

$N_{S_{13}}$ can be inferred in a similar way as the total number of potentially affected passengers outside the system minus the number of passengers in other groups. It is worth noting that the potentially affected passengers include those who do not have tap-in records on the incident day (e.g., $B_{S_{17,18}}$). Define $B_{S_{13}} = \{p : t_{p_1} \geq T_1\} \cup B_{S_{17,18}}$ as the set of passengers outside the system who were potentially affected. Then,

$$\mathbb{E}[N_{S_{13}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{13}}\}} - \sum_{i=14}^{19} \mathbb{E}[N_{S_i}] \tag{A.23}$$

The variance of $N_{S_{13}}$ also needs to consider the correlations. Notice that $B_{S_{14}}$ and $B_{S_{19}}$ do not intersect with other B_{S_i}'s. So, $N_{S_{14}}$ and $N_{S_{19}}$ are independent of other N_{S_i}'s. According to Appendix A.2.2, the variance of $N_{S_{17}} + N_{S_{18}}$ can be estimated as a whole:

$$\begin{aligned}
\text{Var}[N_{S_{17}} + N_{S_{18}}] = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{p \in B_{S_{17,18}}\}} \cdot [\mathbb{P}(CTSM_p \mid p \in B_{S_{17,18}}) \cdot \mathbb{P}(A_p \mid CTSM_p, p \in B_{S_{17,18}}) - \\
\mathbb{P}(CTSM_p \mid p \in B_{S_{17,18}})^2 \cdot \mathbb{P}(A_p \mid CTSM_p, p \in B_{S_{17,18}})^2]
\end{aligned} \tag{A.24}$$

And from Appendix A.3.1, $N_{S_{15}}$ and $N_{S_{16}}$ are independent, and independent of $N_{S_{17}} + N_{S_{18}}$ since $B_{S_{17,18}}$ does not intersect with $B_{S_{15}}$ or $B_{S_{16}}$. Therefore, The variance of $N_{S_{13}}$ can be estimated as:

$$\text{Var}[N_{S_{13}}] = \sum_{i=14}^{16} \text{Var}[N_{S_i}] + \text{Var}[N_{S_{17}} + N_{S_{18}}] + \text{Var}[N_{S_{19}}] \tag{A.25}$$

References

Adelé, S., Tréfond-Alexandre, S., Dionisio, C., Hoyau, P.-A., 2019. Exploring the behavior of suburban train users in the event of disruptions. Transp. Res. F 65, 344–362.

Bai, Y., Kattan, L., 2014. Modeling riders' behavioral responses to real-time information at light rail transit stations. Transp. Res. Rec. 2412 (1), 82–92.

Barry, J.J., Freimer, R., Slavin, H., 2009. Use of entry-only automatic fare collection data to estimate linked transit trips in new york city. Transp. Res. Rec. 2112 (1), 53–61.

Barry, J.J., Newhouser, R., Rahbee, A., Sayeda, S., 2002. Origin and destination estimation in New York City with automated fare system data. Transp. Res. Rec. 1817 (1), 183–187.

Ben-Akiva, M.E., Lerman, S.R., Lerman, S.R., 1985. Discrete Choice Analysis: Theory and Application To Travel Demand, Vol. 9. MIT Press.

Cheng, Z., Trépanier, M., Sun, L., 2020. Probabilistic model for destination inference and travel pattern mining from smart card data. Transportation 1–19.

Cox, A., Prager, F., Rose, A., 2011. Transportation security and the role of resilience: A foundation for operational metrics. Transp. Policy 18 (2), 307–317.

CTA, 2019. CTA monthly ridership report for september 2019. URL: https://www.transitchicago.com/assets/1/6/Monthly_Ridership_2019-09.pdf.

Derrible, S., Kennedy, C., 2010. The complexity and robustness of metro networks. Physica A 389 (17), 3678–3691.

Dimitrov, S.D., Ceder, A.A., 2016. A method of examining the structure and topological properties of public-transport networks. Physica A 451, 373–387.

Gordon, J.B., Koutsopoulos, H.N., Wilson, N.H., Attanucci, J.P., 2013. Automated inference of linked transit journeys in London using fare-transaction and vehicle location data. Transp. Res. Rec. 2343 (1), 17–24.

Goulet-Langlois, G., Koutsopoulos, H.N., Zhao, Z., Zhao, J., 2017. Measuring regularity of individual travel patterns. IEEE Trans. Intell. Transp. Syst. 19 (5), 1583–1592.

Jin, J.G., Teo, K.M., Odoni, A.R., 2016. Optimizing bus bridging services in response to disruptions of urban transit rail networks. Transp. Sci. 50 (3), 790–804.

Leng, N., De Martinis, V., Corman, F., 2018. Agent-based simulation approach for disruption management in rail schedule. In: Conference on Advanced Systems in Public Transport and TransitData (CASPT 2018).

Lin, T., Shalaby, A., Miller, E., 2016. Transit user behaviour in response to service disruption: State of knowledge. In: Canadian Transportation Research Forum 51st Annual Conference-North American Transport Challenges in an Era of Change//Les DÉFis Des Transports En AmÉRique Du Nord à Une Aire de Changement Toronto, Ontario.

Lin, T., Srikukenthiran, S., Miller, E., Shalaby, A., 2018. Subway user behaviour when affected by incidents in toronto (SUBWAIT) survey—A joint revealed preference and stated preference survey with a trip planner tool. Can. J. Civil Eng. 45 (8), 623–633.

Liu, T., Ma, Z., Koutsopoulos, H.N., 2021. Unplanned disruption analysis in urban railway systems using smart card data. Urban rail transit 1–14.

López, F.A., Páez, A., Carrasco, J.A., Ruminot, N.A., 2017. Vulnerability of nodes under controlled network topology and flow autocorrelation conditions. J. Transp. Geogr. 59, 77–87.

Mo, B., von Franque, M.Y., Koutsopoulosc, H.N., Attanuccid, J., Zhao, J., 2022. Impact of unplanned service disruptions on urban public transit systems. arXiv preprint arXiv:2201.01229.

Mo, B., Ma, Z., Koutsopoulos, H.N., Zhao, J., 2020. Capacity-constrained network performance model for urban rail systems. Transp. Res. Rec. 2674 (5), 59–69.

Mo, B., Ma, Z., Koutsopoulos, H.N., Zhao, J., 2021. Calibrating path choices and train capacities for urban rail transit simulation models using smart card and train movement data. J. Adv. Transp. 2021.

Murray-Tuite, P., Wernstedt, K., Yin, W., 2014. Behavioral shifts after a fatal rapid transit accident: A multinomial logit model. Transp. Res. F 24, 218–230.

Nielsen, L.K., Kroon, L., Maróti, G., 2012. A rolling horizon approach for disruption management of railway rolling stock. European J. Oper. Res. 220 (2), 496–509.

Peng, D., Chao, L., Zhili, L., 2009. Walking time modeling on transfer pedestrians in subway passages. J. Transp. Syst. Eng. Inf. Technol. 9 (4), 103–109.

Rahimi, E., Shamshiripour, A., Shabanpour, R., Mohammadian, A., Auld, J., 2020. Analysis of transit users' response behavior in case of unplanned service disruptions. Transp. Res. Rec. 0361198120911921.

Silva, R., Kang, S.M., Airoldi, E.M., 2015. Predicting traffic volumes and estimating the effects of shocks in massive transportation systems. Proc. Natl. Acad. Sci. 112 (18), 5643–5648.

Sun, H., Wu, J., Wu, L., Yan, X., Gao, Z., 2016. Estimating the influence of common disruptions on urban rail transit networks. Transp. Res. A 94, 62–75.

Teng, J., Liu, W.-R., 2015. Development of a behavior-based passenger flow assignment model for urban rail transit in section interruption circumstance. Urban rail transit 1 (1), 35–46.

van der Hurk, E., 2015. Passengers, information, and disruptions. EPS-2015-345-LIS.

Veelenturf, L.P., Kidd, M.P., Cacchiani, V., Kroon, L.G., Toth, P., 2016. A railway timetable rescheduling approach for handling large-scale disruptions. Transp. Sci. 50 (3), 841–862.

Zhang, X., Miller-Hooks, E., Denny, K., 2015. Assessing the role of network topology in transportation network resilience. J. Transp. Geogr. 46, 35–45.

Zhao, J., Rahbee, A., Wilson, N.H., 2007. Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. Comput.-Aided Civ. Infrastruct. Eng. 22 (5), 376–387.

Zhu, Y., Koutsopoulos, H.N., Wilson, N.H., 2017. A probabilistic Passenger-to-Train Assignment Model based on automated data. Transp. Res. B 104, 522–542.