TRANSPORT FINDINGS

Learning about the Transit Passenger Experience from Microblogging Posts

Egbe-Etu Etu, Ph.D.[1], Asha Weinstein Agrawal, Ph.D.[2], Jordan Larot[1], Chaitanya Tatipigari[3], Imokhai Theophilus Tenebe, Ph.D.[4]

[1] Marketing and Business Analytics, San José State University, [2] Mineta Transportation Institute, San José State University, [3] H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, [4] Booth School of Business, University of Chicago

---

Findings

Transit agencies often track delays and malfunctions but may overlook riders' real-time social experiences. To fill this gap, we analyzed 559 tweets from 2020 to 2023 that captured personal accounts of public transport use. Using thematic analysis, we found that 74% of posts criticized other passengers' behavior, far outweighing comments on transit ride quality, such as complaints about delays (25%). Notably, 86% of tweets expressed negative sentiment, though some highlighted exceptional service. These findings demonstrate that interpersonal dynamics are a key contributor to riders' experiences and illustrate the nature of the interactions that riders like and dislike.

1. QUESTIONS

This paper investigates how microblogging platforms—social media platforms where users post brief textual updates, such as Twitter—can be used to better understand public transit riders' experiences. Twitter, in particular, is widely used to share real-time feedback about transit systems. Systematic analysis of microblog posts has been found to offer broad insight into rider satisfaction (Etu et al. 2023; Qi, Costin, and Jia 2020; Méndez et al. 2019; Haghighi et al. 2018). However, many transit agencies monitor their social media feeds to identify incidents needing immediate attention or examples of posts describing passenger satisfaction, few agencies perform longitudinal, structured analyses to identify larger patterns (Das et al. 2022). We explored a set of Twitter posts to answer two research questions:

1. What topics do people write about in microblog posts related to public transit?

2. Can systematic analysis of microblog posts provide public transit operators with actionable insights that complement passenger surveys, the most common market research technique?
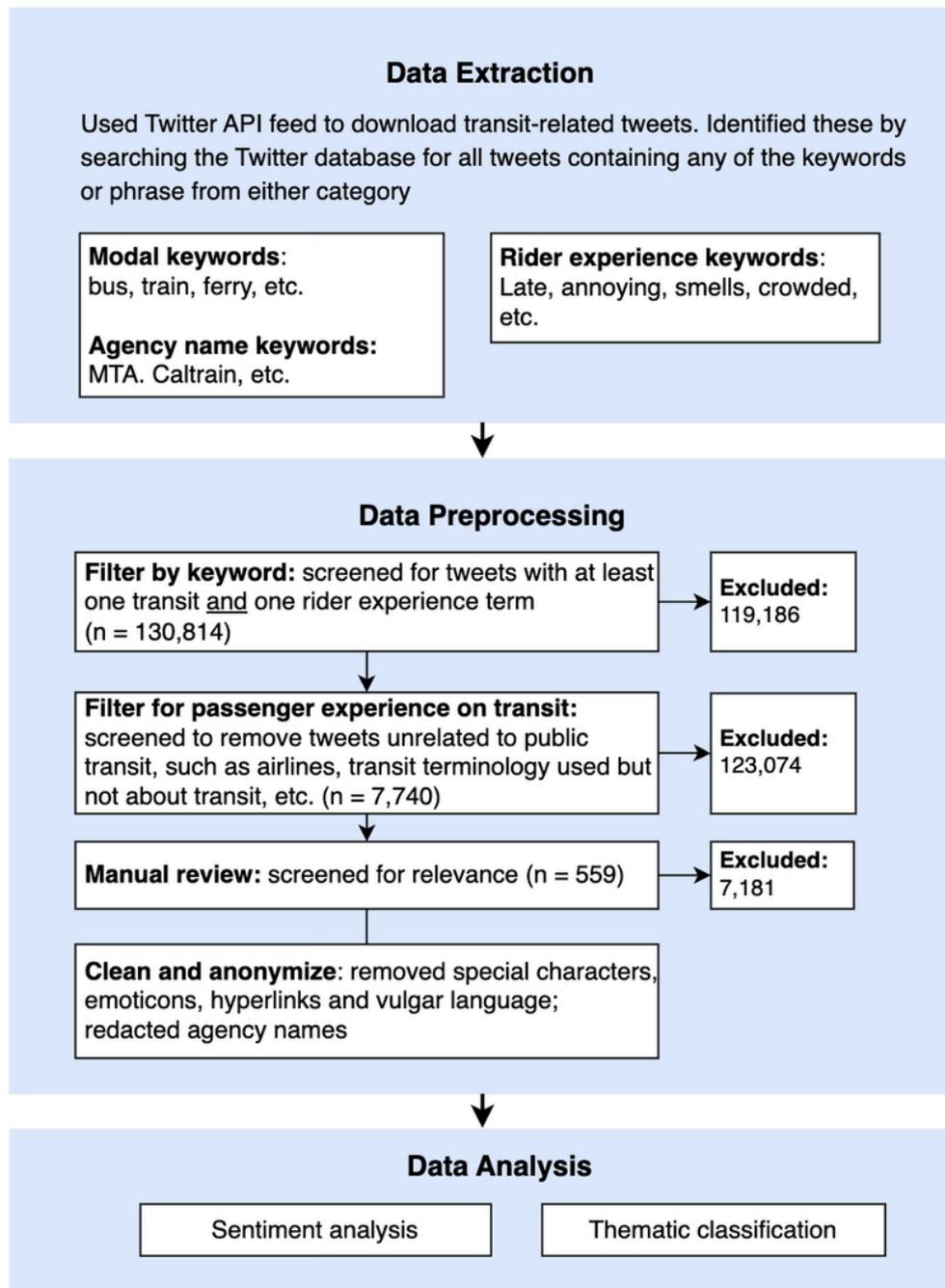
Data Extraction

Used Twitter API feed to download transit-related tweets. Identified these by searching the Twitter database for all tweets containing any of the keywords or phrase from either category

Modal keywords:
bus, train, ferry, etc.

Agency name keywords:
MTA. Caltrain, etc.

Rider experience keywords:
Late, annoying, smells, crowded, etc.

Data Preprocessing

Filter by keyword: screened for tweets with at least one transit <u>and</u> one rider experience term (n = 130,814)

Excluded: 119,186

Filter for passenger experience on transit: screened to remove tweets unrelated to public transit, such as airlines, transit terminology used but not about transit, etc. (n = 7,740)

Excluded: 123,074

Manual review: screened for relevance (n = 559)

Excluded: 7,181

Clean and anonymize: removed special characters, emoticons, hyperlinks and vulgar language; redacted agency names

Data Analysis

Sentiment analysis

Thematic classification

Figure 1. Framework for extracting, preprocessing, and analyzing tweets about personal public transit experiences.

2. METHODS

[Figure 1](#) presents the methodology used to create the study dataset, a collection of tweets describing the writer's recent, personal experience using transit. We extracted from Twitter approximately 250,000 posts from January 2020 to March 2023 that contained keywords hypothesized to describe transit passenger experiences. The extracted tweets include metadata such as a tweet ID, timestamp, language, and—where available—latitude and longitude coordinates pinpointing the approximate location where the tweet was posted.

Data preprocessing included the following steps:

- Removing tweets that were not personal experiences using transit (e.g., news stories, COVID-19, airline travel)

- Assigning tweets to specific transit agencies using a combination of metadata and content-based identifiers found within the tweet text (e.g., agency names, route numbers, station names, and street references). Tweets without identifiable agency, route, or location markers were categorized as "Don't Know."

- Correcting obvious typos, such as misspelled words, and redacting transit agency names for anonymity.

With the final data set of 559 tweets, we conducted sentiment analysis and thematic classification. Further details, including specific filtering steps and exclusion counts, are available in the Supplementary Materials.

3. FINDINGS

Almost all the posts mentioned bus and rail systems as the mode of transit (49% and 46%, respectively). The tweets originated from 36 states plus the District of Columbia. A few states had significant numbers of tweets: 33% from New York, 16% from California, 10% from Illinois, and 7% each from Massachusetts and Pennsylvania (Figure 2A). Fifty-eight percent of tweets named a specific transit agency; 16 agencies were named (Figure 2B). Concerning timing, tweets were most frequent in January, February, and March; on weekdays (compared to weekends); and between 9 pm and 1 am (see Figure 3). These patterns align with existing literature showing that tweet content can vary geographically across U.S. cities (Das and Zubaidi 2023).
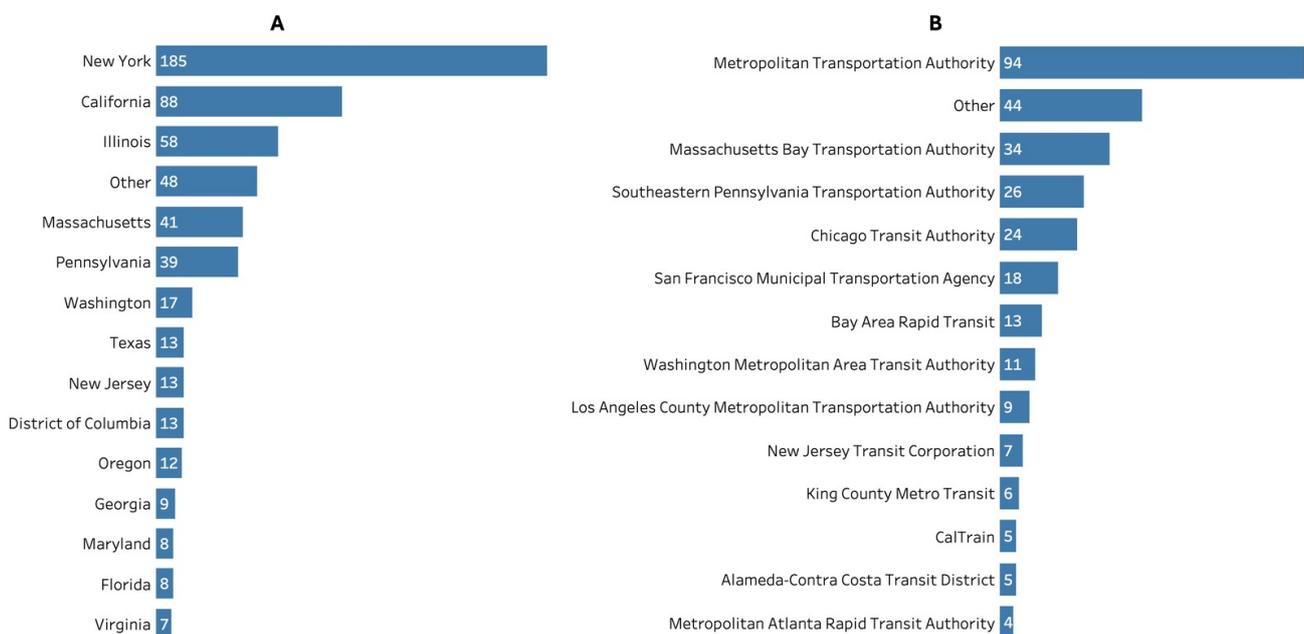

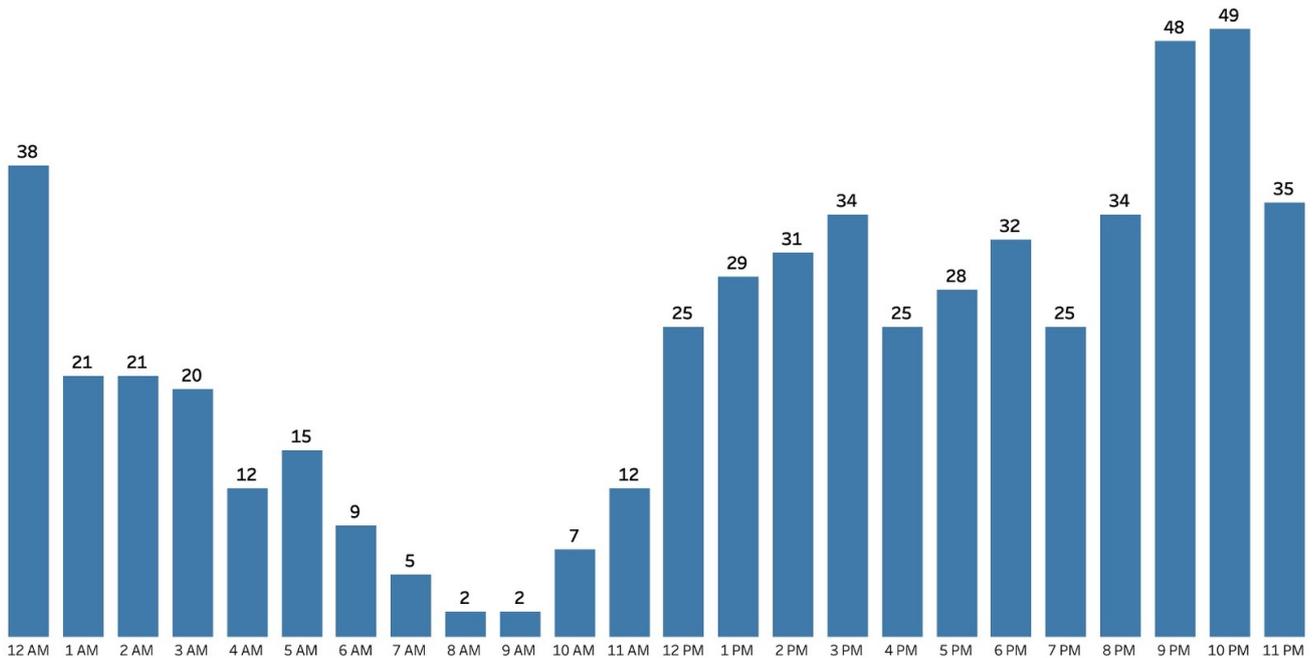Figure 2. Number of tweets by (a) U.S. state and (b) transit agency

Figure 3. Number of tweets per hour

Most tweets described people whom the writer encountered while using transit (74%): passengers (43%) and/or staff from the transit operator or local law enforcement (37%). Only 25% of tweets commented about characteristics of the transit service unrelated to actions by a specific person, such as comments about schedule disruptions. Almost all tweets were purely negative (86%).

Tweets about Other Passengers

The most frequent subject of the tweets was complaints about the behaviors of other passengers (43%), and all but nine of these posts were negative (Table 1). Common topics were noise or smells from other passengers, and passengers acting in a rude, obnoxious, or harassing manner (18% of all tweets). The handful of positive tweets commented approvingly on the entertainment provided by watching other passengers.

Tweets about Transit Operations and Law Enforcement Personnel

About a third of the tweets commented on the behaviors of specific transit operations or law enforcement personnel, most commonly the former (Table 2). The most common complaints centered on what the tweet writers considered rude or harassing behavior as staff tried to enforce agency rules. A much smaller number of tweets complained about unsafe driving. Some writers complained that police or transit operators did not intervene in unsafe situations, while other tweets complained that staff were too aggressive about enforcing rules, such as yelling at passengers who did not pay fares or wear

Table 1. Themes of tweets about the behavior of other passengers, with examples

Topic		Example
Negative Sentiment	Harassing statements directed at passengers of transit operations staff	**A:** "Shout out to the creepy old guy hitting on me on the train today and making my ride very uncomfortable." **B:** "Yelled at an unmasked antivaxxer ranter on the subway earlier today when they started targeting me with what he thought were racist slurs. "Maha Laxmi"" **C:** "I've had the most ridiculous day of my life today and to top it off this girl on the bus won't stfu. She's being obnoxious and disrespectful to the bus driver for no reason. I'm ready to fight someone and it might just be her." **D:** "@[agency] This guy tried to get on the bus through a closed door and when he got on he started yelling at the driver calling him names. He's now making a false report full of lies to [agency, line number, bus number]. The driver is a model employee."
	Non-verbal harassment	**E:** "A man on the train walked by me fully licking his lips while staring at my legs and I'm so pissed I can't literally curse someone with my eyes. I've been angry, sad and disgusted about it all day." **F:** "Literally just peacefully sitting on the train and this man points his phone at me and takes a picture with a very creepy grin"
	Passengers creating smells or noise	**G:** "To the woman in a fur coat and reeking of the Macy's cosmetic department, was it really necessary to push past me getting onto the [agency] when the platform was not crowded?" **H:** "Normally I'd be concerned about making this train reek of bud, but luckily there is a man here who smells so fucking bad I don't have to worry about that! It is now 5:07 AM. I am still on the train. Kill me." **I:** "To the gentleman sitting behind me on the bus who is speaking loudly on the phone. I hate you." **J:** "This lady is on the train with a big ass LOUD ass speaker losing her s– to rock music. Not for cash either. She just in the damn zone."
	Fighting, assault, drug use, or other threatening behavior by passengers	**K:** "Good morning to everyone except the homeless man who yelled in my face and spit on me on the train this morning :)" **L:** "Im ok, I was mostly just shaken up, no injuries, and I still think [police chief] is wrong. Anyhow, a homeless individual pushing a giant wheeled dumpster onto the train decided that I was in the way and physically shoved me and my wheelchair off of the train. I was in pain and scared." **M:** "Nothing starts your day like a crazed man screaming at everyone in the subway car that they stole his stuff and that he's 'going to have the last laugh.' Cops are here now. Good morning everyone." **N:** "The mental health crisis in the [agency name] is real. Second time this year I've had someone in the car targeting riders. In February it was a person who spat on and threatened to kill me when he saw my Pride bracelet. Today it was a person crying and grabbing people asking for help." **O:** "Taking the trolley to [location]. So far one passenger ejected for actively smoking meth. Another passenger who tried to pick a fight (not with me I don't think) given final warning. Lively crowd." **P:** "Observed on the @[agency] today. 1) Man atop the [agency] entrance on [street intersection] with his pants down, ass cheeks out gleefully enjoying a hit off the ole crack pipe. 2) man in the connecting [line name] tunnel shamelessly urinating. Care free. @[mayor name] what's good?"
Positive Sentiment	Entertaining	**A:** "Shout out to the street performers on [agency], we need u those days when our phones [are] not charged and bored as hell." **B:** "Shout out to my fellow subway passenger currently making his own cocktail (looks like a variation of a Bloody Mary) on the train. You're a braver man than me."

masks. Although most tweets about personnel were negative, more than 10% of all tweets praised transit operations staff for helping passengers, having cheerful personalities, skilled driving, or enforcing rules appropriately.

Tweets about the Service Characteristics

Twenty-five percent of tweets commented on some aspect of the transit service *in general* (as opposed to tweets about specific staff), virtually all with a complaint (Table 3). The most common topic was service not running according to the schedule (15% of all tweets). Many tweets also described

Table 2. Themes of tweets about transit operations or law enforcement personnel, with examples

Topic		Example
Negative Sentiment	Rude or harassing behavior	A: "The driver is 25 min late for the switch, and started harassing and threatening riders, saying he'll put the bus out of service when we asked why we've been standing for 25 min." B: "i really hate [agency] drivers so much, there was no reason for you to humiliate a man who can't afford a pass, you could've quietly told him he can't ride rather than say you do this all the time sir, no like he already admitted out loud he has no job and no place to sleep" C: "Ah yes. [agency] let's yell at passengers in wheel chairs and with strollers to make them get off of your train that has been shocking taken out of service. Get f—." D: "Two transit cops are on my [agency] bus yelling at kids to tag on with their [electronic fare] cards. What a job." E: "Bus driver called me fat ass several times. For no reason. Threatening me and continued to do it in front of all these people at [location] outbound. At 3:49 PM now. I'll continue to #boycott"
	Not enforcing rules or laws	F: "Bro f— [police]. This man literally followed me from the station to the train car I was taking and said he's going to hunt me down. They literally stood there and let this man harass me. Thank God people were here to say something, keep me safe. Ugh." G:"[Agency] zerotolerance harassment policy working well eh? Getting yelled a faggot and a sex doll mannequin from across the bus and bus driver didn't say anything or ask him to get off. This does not make me feel safe using [agency] in the future." H: "#[agency and station names] approx 2:26pm. [Agency] police and 2 red shirt helpers near entrance. They just ignored 6 people pushing thru the gates to not pay and none had a mask on."
	Unsafe driving	I: "I'm legit scared. This driver is driving so fast. Like it's his personal vehicle. I can't wait to get to my stop at the mall. I don't feel safe." J: "[Agency] please tell the driver of bus [route number] to stop slamming on the brakes over and over again. It's rush hour, the bus is crowded. It's extremely dangerous and unnecessary."
Positive Sentiment	Helping people	A: "@[agency] A good morning like to send 2 shout outs first awesome bus driver on [route] helped a disabled woman off the bus and always nice to ride with nice lady blue line operator who always say a nice saying" B: "Shout out to @[agency]: Today I got on a different bus than usual and the driver recognized me. He took a second to make sure I knew I was getting on not my usual bus in case it was a mistake. Made my day." C: "@[agency] The automated reminder is broken on my bus, so our driver is shouting out the upcoming stops, and all of that stop's related transfer points, completely from memory. Give this man a raise!!"
	Cheerful personality	D: "I love watching the smiles of the other @[agency] riders when the operator of the [route] starts talking to us #thebest" E: "In related news, the driver on [@ agency name, Route number, bus number] is an absolute ray of sunshine. He has enthusiastically greeted every new rider, thanked and wished well each departure, and is presently whistling a tune."
	Skilled driving	F: "It's terrifying when cars turn right in front of a bus! Just experienced this on the [route and location]. An SUV nearly met its death, but our awesome @[agency] driver had quick reflexes and amazing patience. He didn't even yell at the driver #reckless" G: "Hey @[agency] shout out to the bus driver on [route]. She navigated a VERY tough and crowded [intersection] and was quite funny while doing so! Kudos!"
	Enforcing rules	H: "I'm grateful for the driver of an [route] bus that refused to allow a mask-less aggressive customer to board the bus. Thank you for enforcing the rules and keeping us safe! #MaskUp #[city]" I: "Shout out to the @[agency] conductor who just made about a dozen [team name] fans mask up on the train ride home."
	Not enforcing rules	J: "I said it nice lemme take the bus...Sun is shining my makeup is looking like BUTTER the bus approaches, I get on and the driver gets a good look at this beaut and puts his hand over the metro slot..."

smells, dirty transit environments (4%), and feeling unsafe from crime (4%). Other issues that came up repeatedly, if less often, were crowding, missing or inaudible stop announcements, and uncomfortable temperatures.

Summary of Findings

This study demonstrates how microblog posts can offer insights into the passenger experience that complement what agencies learn from traditional passenger surveys:

1. Microblog posts provide data on passengers' *good* experiences, information that is particularly hard to elicit in surveys that ask for overall ratings and/or require recall of past events.

2. Microblog posts typically describe specific events that people are experiencing, providing qualitative insights into the specifics of passenger experiences. By contrast, a passenger survey can confirm how many passengers feel unsafe or safe, but only the microblog tweets illustrate the events that may create fear.

3. Because microblog posts appear continuously, they can be used to assess changes over time or reactions to a specific event, such as a temporary service disruption. By contrast, agencies only survey passengers intermittently, often only every few years.

The analysis of our dataset of tweets shows that (perceived) anti-social behavior from passengers and, to a lesser extent, from frontline staff is a key complaint among transit passengers. Further, the positive tweets highlight the vital role that staff can play in shaping a welcoming transit environment. These findings suggest that agencies may wish to (1) put a higher priority on identifying and implementing processes to address anti-social passenger behaviors, (2) offer better support to the frontline employees who must manage challenging passenger behaviors with professionalism and care, and (3) formally reward excellent service.

Table 3. Themes of tweets about service characteristics, with examples

Topic		Example
Negative Sentiments	Service not running on time	A: "I hate [agency] detour. I have been waiting for the bus for about 10 mins and I kept checking the Transit app. Without any warning, the bus changed the route and I missed it and I'm late. I hate it"
	Smelly or dirty environment	B: "one of the train cars smelled like ass so i went to a different train car and it smells like weed" C: "HEY I AM CURRENTLY STARING AT VOMIT ON A [agency] TRAIN HOW IS YOUR DAY" D: @[agency] clean this fucking [line name] up! This is ridiculous the filth is unreal! I don't even want the bottom of my shoes to touch the floor, much less put my ass in a seat. COVID is real, hepatitis is real, germs are real! Do better by your riders this is some bullshit"
	Concern about crime	E: "Shout out to the [line] stop for once again being randomly more sketchy at 4:45pm on a Wednesday than midnight on a Saturday."
	Malfunctioning equipment	F: "[Agency] Please have station train announcements about this! Not all riders follow you on Twitter. So far no announcements on my [train number] that it won't be running tomorrow." G: "[agency] please fix the escalator or put a sign up that it is broken, it has a HUGE 2 step gap and VERY dangerous as 5 people almost fell through the gap at [station name]!!" H: "@[agency] not one of ticket dispenser at #[line] was accepting credit cards just now. It's bad enough that you don't allow refills on a 10 trip card but now you forced me to pay full fare because you can't maintain the machines. You just stole from me"
	Crowding	I: "@[agency] What in the F— is going on with the [line name]? I don't think having over 100 people in one car is appropriate nor safe. Elderly people and children should not be fighting over seats"
Positive Sentiments		J: "Shout out to bus [#] for being warm and cozy this evening! @[agency]" K: "I never would have believed that a crowded [agency, vehicle type] would put a smile on my face! #letsgo #[city] #WFHsucks"

There are several limitations that should be considered when analyzing microblog posts:

1. The posts do not reflect the experiences of all transit riders, since most riders do not tweet about their transit experience.

2. Microblog posts are not always accurate. Posters may exaggerate, outright lie, or simply misinterpret the situation they observe. For example, the people described in posts as "homeless" may not be unhoused.

3. Choosing different keywords to extract tweets might have revealed significant new content about themes. For example, we did not include search terms related to cost, and possibly including those terms would have revealed riders' perceptions of fare affordability.

Data Availability

The dataset can be accessed via https://github.com/etujnr/Transit-Tweet-Data

---

REFERENCES

Das, S., N. S. Trisha, I. N. Sener, and M. Walk. 2022. *Uses of Social Media in Public Transportation*. Washington, D.C.: The National Academies Press. https://doi.org/10.17226/26451.

Das, S., and H. A. Zubaidi. 2023. "City Transit Rider Tweets: Understanding Sentiments and Politeness." *Journal of Urban Technology* 30 (1): 111–26. https://doi.org/10.1080/10630732.2021.1903288.

Etu, E. E., A. W. Agrawal, I. Tenebe, J. Larot, and D. M. N. Nguyen. 2023. "Misconduct on Public Transit: An Exploratory Analysis Using the Comments Formerly Known as Tweets." Mineta Transportation Institute. https://transweb.sjsu.edu/research/2317-Misconduct-Public-Transit-Safety-Tweets.

Haghighi, N. N., X. C. Liu, R. Wei, W. Li, and H. Shao. 2018. "Using Twitter Data for Transit Performance Assessment: A Framework for Evaluating Transit Riders' Opinions about Quality of Service." *Public Transport* 10:363–77. https://doi.org/10.1007/s12469-018-0184-4.

Méndez, J. T., H. Lobel, D. Parra, and J. C. Herrera. 2019. "Using Twitter to Infer User Satisfaction with Public Transport: The Case of Santiago, Chile." *IEEE Access* 7:60255–63. https://doi.org/10.1109/ACCESS.2019.2915107.

Qi, B., A. Costin, and M. Jia. 2020. "A Framework with Efficient Extraction and Analysis of Twitter Data for Evaluating Public Opinions on Transportation Services." *Travel Behaviour and Society* 21:10–23. https://doi.org/10.1016/j.tbs.2020.05.005.

SUPPLEMENTARY MATERIALS

Supplemental Information

Download: https://findingspress.org/article/140966-learning-about-the-transit-passenger-experience-from-microblogging-posts/attachment/289517.pdf