Real-Time Riders

A First Look at User Interaction Data from the Back End of a Transit and Shared Mobility Smartphone App

Candace Brakewood, Niloofar Ghahramani, Jonathan Peters, Eunjin Kwak, and Jake Sion

A fundamental component of transit planning is understanding passenger travel patterns. However, traditional data sources used to study transit travel have some noteworthy drawbacks. For example, manual collection of travel surveys can be expensive, and data sets from automated fare collection systems often include only one transit system and do not capture multimodal trips (e.g., access and egress mode). New data sources from smartphone applications offer the opportunity to study transit travel patterns across multiple metropolitan regions and transit operators at little to no cost. Moreover, some smartphone applications integrate other shared mobility services, such as bikesharing, carsharing, and ride-hailing, which can provide a multimodal perspective not easily captured in traditional data sets. The objective of this research was to take a first look at an emerging data source: back-end data from user interactions with a smartphone application. The specific data set used in this paper was from a widely used smartphone application called Transit that provides real-time information about public transit and shared mobility services. Visualizations of individuals' interactions with the Transit app were created to demonstrate three unique aspects of this data set: the ability to capture multicity transit travel, the ability to capture multiagency transit travel, and the ability to capture multimodal travel, such as the use of bikeshare to access transit. This data set was then qualitatively compared with traditional transit data sources, including travel surveys and automated fare collection data. The findings suggest that the data set has potential advantages over traditional data sources and could help transit planners better understand how passengers travel.

A fundamental component of transit planning is understanding passenger travel patterns. However, traditional data sources used to study transit travel patterns have some noteworthy drawbacks. For example, manual collection of travel surveys can be time consuming and expensive; similarly, data from automated fare collection (AFC) systems, such as smart cards, often include only one transit agency and usually do not capture multimodal trips (e.g., access and egress mode). Rapid adoption of mobile phones has led to the creation of new data sources that can be used to study travel patterns, particularly smartphones that are GPS enabled or location aware (*1–3*). Smartphone applications (or apps) that focus on specific modes, such as those that provide real-time transit information, can capture data that may be harder to identify in larger mobile phone data sets (e.g., data from cellular network providers) because of difficulty in differentiating between modes, cell phone service issues in underground tunnels, and other challenges. The widespread use of apps to access information about public transit and new shared mobility services (e.g., carsharing, bikesharing, and ride-hailing) presents a unique opportunity to utilize the data generated from user interactions with these apps to study transit and shared mobility travel behavior. The objective of this paper is to take a first look at an emerging data source: back-end data from user interactions with a transit-related smartphone application. The specific data set used in this paper comes from a widely used app called Transit that provides real-time information about public transit and shared mobility services (*4*).

This paper proceeds as follows. First, a brief review of prior research is provided. Next, the smartphone application that is the focus of this paper is discussed in detail, including a description of the user interface, the data captured in various back-end tables from user interactions, and the specific data samples used in this paper. This discussion is followed by a visualization exercise that shows individuals' interactions with the Transit app to demonstrate three unique aspects of the data set. Next, this new data set is qualitatively compared with traditional sources of transit travel behavior data, including travel surveys and automated fare collection systems. Finally, conclusions and areas for future research are presented.

PRIOR RESEARCH

One of the most commonly used data sources to study transit travel behavior is surveys, which are often conducted as household travel surveys or in stations or on board transit vehicles (*5, 6*). Survey data have been extensively written about in the literature [e.g., *The Online Travel Survey Manual* (*7*)] and will not be reviewed in detail here.

AFC systems, particularly smart cards, have become commonplace in transit systems over the past two decades. Although smart card systems are installed for the purpose of revenue collection, they also provide a rich source of data about transit travel (*8, 9*). Passengers with contactless smart cards pay fares by tapping their cards at fare gates or on fare boxes. With each tap, a record is created that includes the date and time, fare type, route or station ID, a unique card ID number, and possibly other things (*9*). Pelletier et al. provided a literature review of the many uses of transit smart card data,

C. Brakewood and N. Ghahramani, Department of Civil Engineering, City College of New York, 160 Convent Avenue, New York, NY 10031. Current affiliation for C. Brakewood: Deparment of Civil and Environmental Engineering, University of Tennessee, Knoxville, 851 Neyland Drive, Knoxville, TN 37996-2313. J. Peters and E. Kwak, School of Business, College of Staten Island, 2800 Victory Boulevard, Staten Island, NY 10314. J. Sion, Transit, 5333 Avenue Casgrain, Suite 803, Montreal, Quebec H2T 1C2, Canada. Corresponding author: C. Brakewood, cbrakewo@utk.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2658, 2017, pp. 56–63.
http://dx.doi.org/10.3141/2658-07

including "strategic-level" analyses that relate to travel behavior and demand forecasting (*9*).

More recently, data from GPS-enabled mobile phones have been utilized to study transit travel. Within this growing body of literature, one noteworthy segment pertains to smartphone-based household travel surveys that utilize prompted recall surveys on location-aware smartphones to collect travel behavior data [e.g., Cottrill et al. (*10*)]. Other studies have used crowdsourced GPS-traces from mobile phones to generate transit vehicle predictions (*11*) and map informal transit systems (*12*). A small number of studies have aimed to capture transit travel behavior by using mobile phone location data for the purpose of understanding passenger movements; one noteworthy example was recently published by Carrel et al., who developed a methodology for merging transit automatic vehicle location data with smartphone location data (*13*).

In summary, there is a growing body of literature that pertains to the use of mobile phone data to study transit travel. However, to the best of the authors' knowledge, previous studies have not utilized user interaction data from the back end of a smartphone application to study transit travel behavior. Therefore, this study aims to explore the potential uses of this emerging data source.

TRANSIT APP

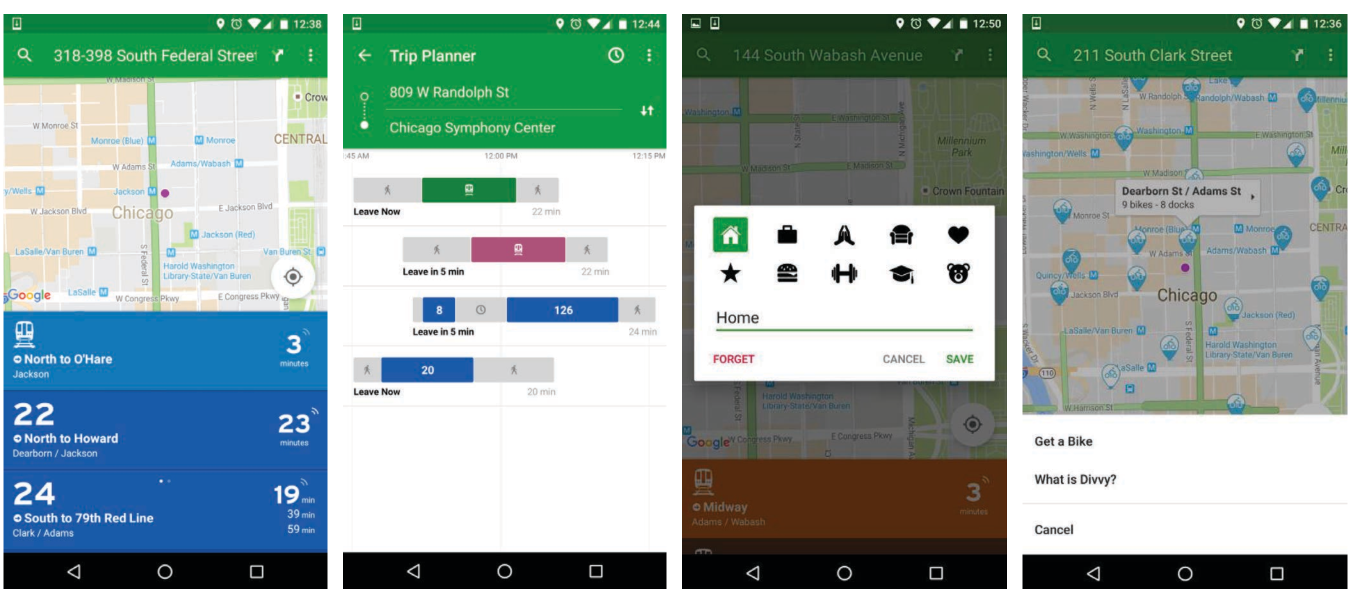
Transit App is a company based in Montreal, Canada, that has developed a freely available smartphone application known as "Transit" (*4*). In 2012, the company released the first version of its application for iPhone, and in the initial version, the app provided transit schedule information for Montreal, Toronto, and Quebec City, Canada. Since then, an Android version of the application has been launched, and the app has expanded to more than 125 cities in nine countries, including widespread coverage in the United States. Many additional features have also been incorporated into the app, including real-time transit information, transit trip planning, and multimodal support (including bikesharing, carsharing, and Uber). Transit app users can also store their favorite locations, such as home or work, in the app to facilitate quick finding of information that they commonly use. Figure 1 shows the Transit app Android interface displaying real-time transit information for nearby routes, trip planning, storing a home location, and bikesharing, respectively.

Whenever a user opens the Transit app, data about his or her interactions with the app are created and stored in a back-end database. Each user interaction is called a session and is identified by a unique ID number and time stamp. To provide relevant transportation information for nearby transit service, the application needs to identify the location of the device. Therefore, each time the application is opened, a session is generated, the device location is sent to the Transit app server, and this record is then stored in a back-end database. To protect the anonymity of users, neither names nor demographic information are requested or stored. More details about additional data fields that are stored are provided in the following section.

Back-End Data Tables

The back-end database generated by user interactions with the Transit app is divided into tables that capture data pertaining to the various functions within the app. The structure of the back end can change when there are application updates to include new features, and the 13 back-end data tables discussed in this paper and summarized in Table 1 represent a sample provided to the authors in 2016.

The first back-end data file, "Locations," is the primary file in the Transit app back-end server that contains a user's interactions with the app. Every time a user opens the app, a session is created, and each session can be identified by a unique ID. The locations file records numerous items for each session, including the following: coordinates (latitude and longitude) of the user, time stamp, accuracy, speed (if the user is in a vehicle) and simulation



(a) (b) (c) (d)

FIGURE 1 Transit app screenshots: (*a*) nearby transit routes, (*b*) trip planning, (*c*) storing home location, and (*d*) bikesharing (Android Version 3.11.2).

TABLE 1 Summary of Tables from the Transit App Back End

Table No.	Table Name	Description of Contents
1	Locations	Includes location (latitude and longitude) for each time user opens app. Date, time, accuracy of location, and speed (e.g., if user is in vehicle) are recorded, and unique session ID is created each time user opens app.
2	Session Complete	Provides event-based view of user, including beginning location and ending location for each session of Transit app. (Session is loosely defined as when app is opened until it is closed, with some variation in actual timing and frequency caused by phone widgets on Android devices.)
3	Placemarks	Includes coordinates of home and work locations that users have stored in Transit app. These data are from optional function in Transit app that allows users to store places to which they often go (e.g., home, work) to easily access relevant transit information for that specific location.
4	Sharing System Actions	Provides information on booking of carshare, bikeshare, and other services. Contains information on location of station for bikeshare or location of vehicle for carshare systems. Provides information on booking and cancellation process for these systems.
5	Sharing System Purchase	Provides information on purchase of shared vehicle passes, which are primarily bikeshare passes. Variables include type of pass, time of request, number of passes selected, and cost of transaction.
6	Uber Request	Lists requests for service from Uber, which are then handed off to Uber's app for fulfillment. Contains time of request, location of user at time of request, type of service requested, and, in some cases, drop-off location.
7	Trips	Contains information about usage of trip-planning feature in Transit app, including starting and ending coordinates (latitude and longitude), date, and time stamps of trip-planning requests.
8	Nearby View	Contains information about routes presented to user in each session upon opening app. Information includes transit route, corresponding agency, number of taps for each route, and whether that route is designated as user's favorite transit route.
9	User Feed Session	Provides information on number of times Transit app is opened by user and number of transit agencies for which user has requested information.
10	Installed App	Reports on other installed apps on user's device that can affect Transit app functionality. For example, user who has Uber installed will be linked to Uber app when ordering ride, whereas user who does not will be sent to app store.
11	Feed Download	Provides overall summary of activity on Transit app by day.
12	Favorite	Provides information on user-designated favorites in terms of transit routes.
13	Device	Contains Transit app–specific ID number that is utilized in other tables to identify unique device. Also includes information on user-selected language, type of device, model of device, operating system used, version of Transit app software installed, and last date of app use.

(if the user has moved the GPS point in the Transit app map and searched for information in a location other than where he or she actually was).

The second and third files described in Table 1 are "Session Complete" and "Placemarks," respectively. The session complete table provides an event-based view of a user. This information includes the starting and ending location of each session in terms of user coordinates (latitude and longitude), starting and ending time stamps, number of records transmitted during the session, and whether the session was simulated or not. The placemarks table includes data about an optional feature to store frequently used places, such as home and work, which can be saved within the app to help users find relevant information quickly.

The next three files listed in Table 1 pertain to the shared mobility services that are integrated within the Transit app. The fourth table, "Sharing System Actions," provides information on the booking of carshare, bikeshare, and other services. The file contains information on the location of the station for bikeshare systems or the location of the vehicle for carshare systems, as well as the location of users when they are searching for sharing system information. Further, it provides information on the booking and cancellation process for these systems, which can be done through the Transit app directly. The fifth table, "Sharing System Purchase," provides information on the successful purchase of shared vehicle passes (primarily bikeshare). Variables include type of pass, time of request, number of passes, and the cost of the transaction. The sixth table, "Uber Request," lists requests for service from Uber. Once a user requests the service, the request is handed off to Uber through its smartphone application for fulfillment. This file contains the time of request, location of the user at time of request, type of service requested, and, in some cases, drop-off location. For additional information about Transit app Uber request data, see Davidson et al. (*14*).

The remaining seven tables are not used in the following analysis but are briefly described here for completeness. The seventh table, "Trips," contains information about usage of the trip-planning feature in the Transit app, which provides A-to-B transit directions. This table includes the starting and ending coordinates (latitude and longitude), date, and time stamps of trip-planning requests. The eighth file, "Nearby View," contains information about the routes presented to a user in each session upon the opening of the app. This information includes the transit route and the corresponding transit agency, number of taps for each route, and whether that route is designated as the user's favorite transit route. The ninth table, "User Feed Session," provides information on the number of times the Transit app is opened by a user each day and the transit agencies for which the user requested information. The 10th file, "Installed App," reports on other apps installed on the user's device that can affect Transit app functionality. For example, a user who has Uber installed on his or her phone will be linked to the Uber app when he or she requests a ride, whereas a user who does not have the Uber app installed will be sent to the app store. The 11th file, "Feed Download," provides an overall summary of Transit app activity by day. The 12th table, "Favorite," provides information on user-designated favorites in terms of transit routes. The 13th table, "Device," creates a Transit app–specific ID number that is utilized in other tables to identify a unique device in the data archive. This table also includes the user's selected language, type of device, model of device, operating system used, version of the Transit app software installed, and last date of Transit app use.

Samples

Two data samples of Transit app data were provided to the research team for this exploratory analysis. The first sample included records from 1 month in 2014 for any user that opened the Transit app at least once in the New York City region. This sample included approximately 10.8 million records and contained five tables, two of which—the locations table and the placemarks table—were used for the visualization exercised presented below. The second sample of data was much larger (approximately 12 terabytes) and included 418 days of data from 2015 and 2016 for all geographic regions available in the Transit app. This larger data set included all 13 tables shown in Table 1; however, only the session complete and sharing systems tables were used in the visualization exercises.

VISUALIZATIONS

Visualizations of individuals' interactions with the app were created to demonstrate the potential uses of this unique data set. The three examples in this exercise are multicity transit travel, multiagency transit travel, and multimodal travel (e.g., from bikesharing to transit). For the purpose of this study, all Transit app user location data (latitude and longitude) were offset by a random number. Anonymizing the location data in this way ensured that user privacy was protected. Throughout this paper, whenever user locations are mentioned, they refer to the anonymized version of the data point.

Analysis 1. Multicity Transit Travel

Because the Transit app covers more than 125 metropolitan regions, this data set can be used to identify intercity travelers and understand how they use transit systems in metropolitan areas other than their home city (*15*). A simple example of this is shown in Figure 2. Data from the locations table of the first data sample (2014) are displayed to show how an individual traveler used the Transit app in Los Angeles, California; Houston, Texas; and New York City over the course of 5 days. Each colored circle in Figure 2 represents that individual's interactions with the Transit app on a specific day. On the first day (light blue), the app user is observed in the vicinity of the Ontario International Airport near Los Angeles, California. On the following 2 days, this individual uses the app in various areas of the Los Angeles metropolitan region. On the fourth day (pink), the app user has Transit app records in the vicinity of Houston Hobby Airport and New York LaGuardia Airport. These records imply that he or she has traveled from Los Angeles to New York and has transferred planes in Houston, even though there is no observation of this user in a Los Angeles airport on that day. On the fifth day, this user has Transit app records in the Bronx, which is his or her self-reported home location in the placemarks table, as well as records in Manhattan and New Jersey (not shown in Figure 2).

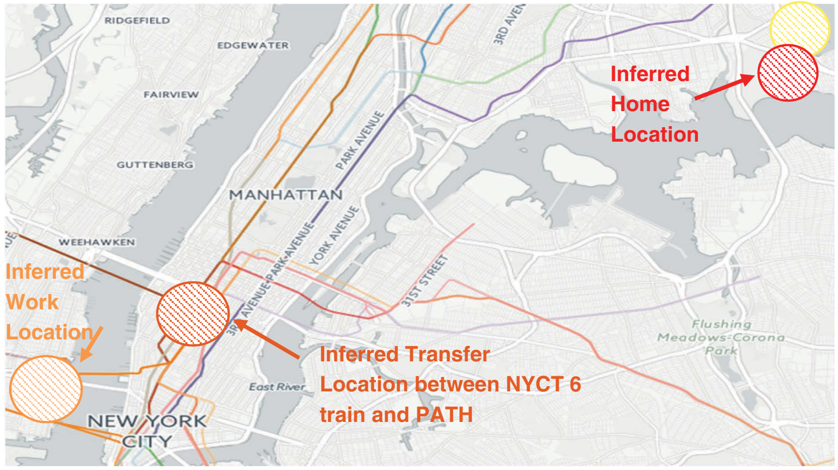
Analysis 2. Multiagency Transit Travel

Similar to the previous example, the Transit app data set enables the identification of users who search for transit information from multiple transit agencies operating within the same metropolitan region. Figure 3 uses the locations table from the first data sample (2014) to show an example of an individual's interactions with the Transit app. The figure shows two weekdays that are typical of this user and that likely represent a commuting pattern in the New York region. During the two weekdays shown in Figure 3, the app user is observed in the Bronx in the early morning (shown in yellow) and late evening (shown in red), which suggests that the Bronx is likely this user's home location (labeled "Inferred Home Location"). He or she is also observed in New Jersey during the daytime, which suggests that this location is likely his or her work location (labeled "Inferred Work Location"). In addition to inferred home and work locations, this app user's commute route presumably includes multiple transit operators. He or she appears to take the New York City Transit Number 6 train route from the Bronx to Manhattan and then transfer to the Port Authority of New York and New Jersey's PATH train from Manhattan to New Jersey. This user is observed on both weekdays in the vicinity of 23rd Street Station in midtown Manhattan, which is likely the transfer location between these two transit operators (labeled "Inferred Transfer Location").

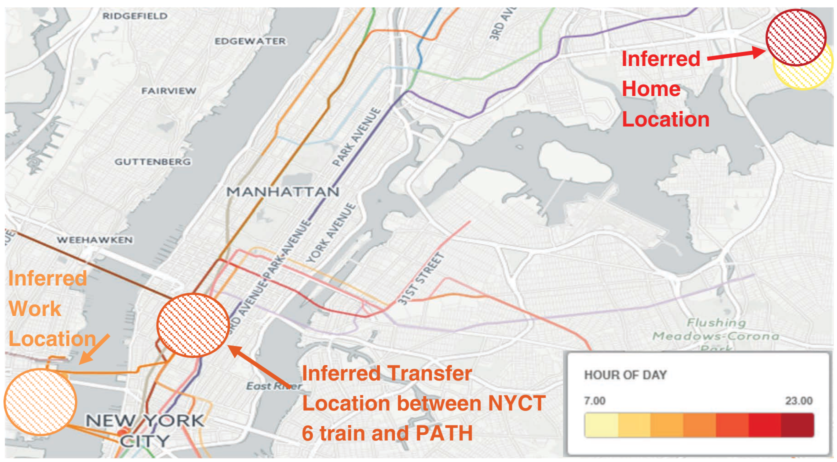
Analysis 3. Multimodal Travel

The third visualization demonstrates how the Transit app data set can be used to identify multimodal travel. Bikesharing is one of the shared mobility services available in the Transit app, and in

FIGURE 2 Example of an individual's multicity transit travel: (*a*) first day to third day, (*b*) fourth day, and (*c*) fourth and fifth days.



(a)



(b)

FIGURE 3 Example of an individual's multiagency transit travel: (*a*) first typical weekday and (*b*) second typical weekday.

some cities, such as Chicago, this app can be used to purchase bikeshare passes and unlock bicycles. Figure 4 uses bikeshare data from the sharing system actions table combined with the session complete table from the second (larger) sample of Transit app data to identify a multimodal traveler in Chicago on a single day in 2016. The individual unlocks a Divvy bikeshare bicycle in the morning (labeled "Divvy Request in the Morning"), which is likely near his or her home location (labeled "Inferred Home Location"). Shortly thereafter, this individual is observed in the vicinity of the Chicago Transit Authority's Blue Line at Irving Park Station, where he or she has probably transferred from bikeshare to rail (shown in yellow). Next, this user is observed during the daytime in the vicinity of Cumberland Station, which is also on the Blue Line and is likely near his or her work location (labeled "Inferred Work Location"). In the early evening, this user is observed again unlocking a Divvy bikeshare bicycle in a location close to Irving Park Station (labeled "Divvy Request in the Afternoon"). This is nearly the same location in which the user was observed transferring from bikeshare to rail in the morning, implying that he or she is transferring from rail to bikeshare and probably continuing his or her trip home via bicycle. Last, in the late evening, this individual is again observed near his or her inferred home location in Irving Park (shown in dark red).

COMPARISON OF TRANSIT APP DATA WITH OTHER SOURCES

This section briefly compares two commonly used sources of data on transit travel behavior—surveys and AFC systems—with Transit app data. These three data sets are compared on five dimensions:

1. Geographic coverage,
2. Institutional coverage,
3. Mode,
4. Timescale, and
5. Sample size.

The results are shown in Table 2 and discussed in the following paragraphs.

The first dimension, geographic coverage, refers to the geographic area covered by each data set. Surveys are typically conducted within

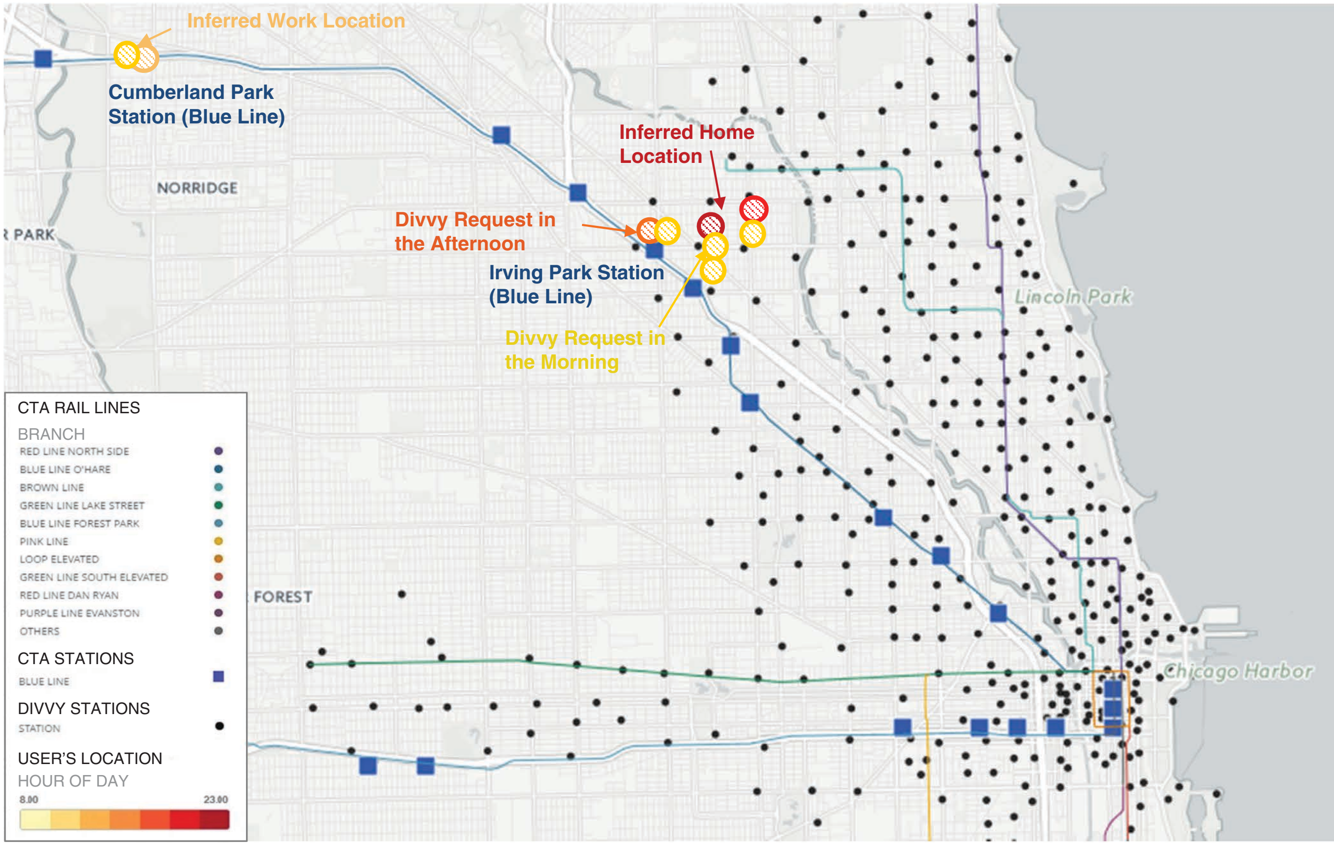


FIGURE 4 Example of an individual's multimodal travel on a typical weekday.

a single metropolitan area by the local transit agency or metropolitan planning organization, with a notable exception being the American Community Survey collected by the Census Bureau (*16*). Similarly, AFC systems are typically not compatible across different metropolitan areas (*17*); subsequently, the data generated by these systems typically encompass transit travel records for a single region. Because the Transit app includes more than 125 metropolitan regions in nine countries, it collects user interaction data for many different metropolitan regions, and these data can then be used to understand multicity travel patterns, as demonstrated in Figure 2.

The second dimension, institutional coverage, refers to the transit agencies within a single metropolitan region for which the data set is collected. Surveys conducted by the local transit agency typically include questions that pertain to only one transit provider, whereas surveys conducted by the metropolitan planning organization are more likely to consider numerous operators in the region. AFC systems usually include only one transit agency (*15*), although a small number of regions allow for interoperability between multiple local operators. The Transit app integrates many transit agencies that have open data (e.g., General Transit Feed Specification, or GTFS, schedule information or real-time vehicle data), and this integration has made user interaction data available for numerous transit agencies in the same region. The visualization in Figure 3, in which an individual likely transferred between New York City Transit and PATH, is an example of such integration of data from multiple agencies.

The third dimension, mode, refers to the mode of transportation for which data are collected. Surveys conducted by the local transit agency typically focus on transit travel and occasionally include a limited number of questions about access and egress mode to transit stations and stops; surveys conducted by metropolitan planning organizations are more likely to capture multiple modes, including

TABLE 2 Comparison of Transit App Data with Traditional Transit Data Sources

Dimension	Travel Survey Data	AFC Data	Transit App Data
Geographic coverage	Single region	Single region	Multiregion
Institutional or agency coverage	Single- or multiagency	Single agency[a]	Multiagency
Mode	Transit or other	Transit	Transit, ride-hailing, bikesharing, carsharing
Timescale	Cross sectional	Continuous (when transit system open)	Continuous
Sample size	Small	Large	Large

[a]A small number of AFC systems allow for use on multiple transit operators in the same region.

automobile, transit, and nonmotorized modes. AFC systems typically include only data pertaining to transit travel, because data are captured when transit fares are paid. The Transit app integrates information about numerous shared mobility modes, including bikesharing, carsharing, and ride-hailing (Uber), and allows users to purchase passes and, in some cities, to utilize shared mobility vehicles such as Divvy bicycles in Chicago. The multimodal nature of the Transit App data set is shown in Figure 4.

The fourth dimension, timescale, refers to the period in which data are generated. Most travel surveys are conducted at a single point in time and therefore provide a cross-sectional snapshot of transit travel behavior. Although panel surveys may be conducted at multiple points in time, this is infrequent in practice, primarily because of cost constraints. AFC data are collected whenever the transit system is in operation, which means this data source is (nearly) continuous in time. The Transit app functions 24 hours a day, 7 days a week, and therefore collects user interaction data continuously. The continuous nature of this data set could advantageously allow for future analyses that would examine travel behavior during events that are difficult to capture in cross-sectional data sets (e.g., extreme weather events).

The fifth dimension, sample size, refers to the quantity of data collected. Travel surveys typically sample only a small portion of the population of interest. Despite the relatively small sample size, however, the methods used to conduct surveys usually aim to be representative of the entire population of interest. AFC systems generate vast quantities of data, and, depending on the level of AFC adoption by riders in a region, they can represent all or nearly all transit riders. The Transit app also generates vast quantities of data because many riders use the app on a daily basis (*18*); however, the sample for which Transit app data are generated depends on the app's adoption and utilization levels, which could be biased (e.g., toward younger, more technology-friendly riders) as compared with the overall population. Therefore, future research is needed to understand the potential biases of this new data set.

CONCLUSIONS AND FUTURE RESEARCH

This research took a first step toward examining an emerging data source: back-end data from user interactions with a smartphone application. The specific data set used in this paper is from a widely used smartphone application called Transit that provides real-time information about public transit and shared mobility services. A visualization exercise was conducted to demonstrate three unique aspects of the Transit app data set: the ability to study multicity transit travel, the ability to study multiagency transit travel, and the ability to study multimodal travel, such as taking bikeshare to access transit. These three aspects of the Transit app data set have the potential to provide unique travel behavior information that is not easily captured in traditional transit data sets such as travel surveys and AFC data.

Many areas for future research emerged from this study, and three specific areas are briefly described in the following paragraphs. First, this exploratory analysis highlighted some unique aspects of the Transit app data set with simple visualizations. An important next step in this research is to develop algorithms to identify this type of travel behavior (i.e., multicity, multiagency, and multimodal transit travel) in the larger data set or for desired subsets of the data set (e.g., for a single metropolitan area). While the visualizations are useful as a proof of concept, the development of algorithms is critical to using the larger data set for planning purposes. Data-mining techniques could be used for this purpose; however, challenges associated with manipulating this big data set may arise.

Second, once algorithms have been developed for the larger data set, potential biases of this data source should be explored. These data could be biased in comparison with data for the overall transit-riding population—for example, toward younger, more technology-friendly riders. Comparing Transit app data with traditional data sources (i.e., travel surveys and AFC data) could identify systematic biases, and then methods to correct for these biases could be developed.

Third, an important area for future consideration is the arrangements by which transit agencies and planners can access data from this and other smartphone apps. The company that provides the specific app used in this study has already begun sharing its data with a limited number of transit agencies. For example, Transit recently announced a partnership with the Massachusetts Bay Transportation Authority in which the agency promotes Transit as its preferred app. As part of the agreement, the app developers provide data to the transit agency to be used for planning purposes (*19*). This example suggests that there are innovative and interesting data-sharing policies that can facilitate utilization of this type of data in the future.

In summary, after algorithms for expanding these analyses have been developed and corrected for systematic biases, the Transit app data set is likely to be extremely valuable to transit planners, operators, and managers and has the potential to transform the understanding of public transit and shared mobility travel behavior.

ACKNOWLEDGMENTS

This research was funded in part by a 2015 City University of New York (CUNY) Collaborative Incentive Research Grant and a University Transportation Research Center faculty-initiated research grant. The authors are grateful to the Transit app for sharing data with the research team. They acknowledge CUNY Graduate Center student Adam Davidson for his contributions to an early draft of this paper and thank Liv Haselbach and Josephine Kressner for reviewing an early draft of this manuscript. Finally, the authors thank the Transportation Research Board for the opportunity to participate in the Chan Wui and Yunyin Rising Star Workshop for Early Career Professionals.

REFERENCES

1. Herrera, J.C., D.B. Work, R. Herring, X.J. Ban, Q. Jacobson, and A.M. Bayen. Evaluation of Traffic Data Obtained Via GPS-Enabled Mobile Phones: The Mobile Century Field Experiment. *Transportation Research Part C: Emerging Technologies,* Vol. 18, No. 4, 2010, pp. 568–583. https://doi.org/10.1016/j.trc.2009.10.006.
2. Jiang, S., G.A. Fiore, Y. Yang, J. Ferreira, Jr., E. Frazzoli, and M.C. González. A Review of Urban Computing for Mobile Phone Traces: Current Methods, Challenges and Opportunities. In *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing,* Association for Computing Machinery, New York, 2013. https://doi.org/10.1145/2505821.2505828.
3. Wolf, J., W. Bachman, M. Oliveira, J. Auld, A.K. Mohammadian, P. Vovsha, and J. Zmud. *NCHRP Report 775: Applying GPS Data to Understand Travel Behavior, Vol. II. Guidelines.* Transportation Research Board of the National Academies, Washington, D.C., 2014. https://doi.org/10.17226/23436.
4. Transit App. http://transitapp.com/. Accessed Oct. 11, 2016.
5. Schaller, B. *TCRP Synthesis of Transit Practice 63: On-Board and Intercept Transit Survey Techniques.* Transportation Research Board of the National Academies, Washington, D.C., 2005. https://doi.org/10.17226/13866.

6. Pratt, J., M. Lee-Gosselin, and S. Burbridge. Chapter 3: Options for Travel Surveys. In *The Online Travel Survey Manual: A Dynamic Document for Transportation Professionals* (TRB Travel Survey Methods Committee, ed.). http://www.travelsurveymanual.org/Chapter-3-1.html. Accessed Oct. 13, 2016.
7. *The Online Travel Survey Manual: A Dynamic Document for Transportation Professionals.* TRB Travel Survey Methods Committee. http://www.travelsurveymanual.org/. Accessed Oct. 13, 2016.
8. Bagchi, M., and P.R. White. The Potential of Public Transport Smart Card Data. *Transport Policy,* Vol. 12, No. 5, 2005, pp. 464–474. https://doi.org/10.1016/j.tranpol.2005.06.008.
9. Pelletier, M.P., M. Trépanier, and C. Morency. Smart Card Data Use in Public Transit: A Literature Review. *Transportation Research Part C: Emerging Technologies,* Vol. 19, No. 4, 2011, pp. 557–568. https://doi.org/10.1016/j.trc.2010.12.003.
10. Cottrill, C.D., F.C. Pereira, F. Zhao, I.F. Dias, H.B. Lim, M. Ben-Akiva, and P.C. Zegras. Future Mobility Survey: Experience in Developing a Smartphone-Based Travel Survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2354, 2013, pp. 59–67. https://doi.org/10.3141/2354-07.
11. Zimmerman, J., A. Tomasic, C. Garrod, D. Yoo, C. Hiruncharoenvate, R. Aziz, and A. Steinfeld. Field Trial of Tiramisu: Crowd-Sourcing Bus Arrival Times to Spur Co-Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Association for Computing Machinery, New York, 2011, pp. 1677–1686. https://doi.org/10.1145/1978942.1979187.
12. Williams, S., A. White, P. Waiganjo, D. Orwa, and J. Klopp. The Digital Matatu Project: Using Cell Phones to Create an Open Source Data for Nairobi's Semi-Formal Bus System. *Journal of Transport Geography,* Vol. 49, 2015, pp. 39–51. https://doi.org/10.1016/j.jtrangeo.2015.10.005.
13. Carrel, A., P.S. Lau, R.G. Mishalani, R. Sengupta, and J.L. Walker. Quantifying Transit Travel Experiences from the Users' Perspective with High-Resolution Smartphone and Vehicle Location Data: Methodologies, Validation, and Example Analyses. *Transportation Research Part C: Emerging Technologies,* Vol. 58, 2015, pp. 224–239. https://doi.org/10.1016/j.trc.2015.03.021.
14. Davidson, A., J. Peters, and C. Brakewood. Interactive Travel Modes: Uber, Transit and Mobility in New York City. Presented at 96th Annual Meeting of the Transportation Research Board, Washington DC, 2017.
15. Ghahramani, N., C. Brakewood, and J. Peters. An Exploratory Analysis of Intercity Travel Patterns Using Backend Data from a Transit Smartphone Application. Presented at 96th Annual Meeting of the Transportation Research Board, Washington D.C., 2017.
16. U.S. Census Bureau. American Community Survey. http://www.census.gov/programs-surveys/acs/. Accessed Oct. 11, 2016.
17. Acumen Building Enterprise, Inc. *TCRP Report 115: Smartcard Interoperability Issues for the Transit Industry.* Transportation Research Board of the National Academies, Washington, D.C., 2006. https://doi.org/10.17226/14012.
18. Ghahramani, N., and C. Brakewood. Trends in Mobile Transit Information Utilization: An Exploratory Analysis of Transit App in New York City. *Journal of Public Transportation,* Vol. 19, No. 3, 2016, pp. 139–160. https://doi.org/10.5038/2375-0901.19.3.9.
19. MBTA Selects "Best Transit App" Winner. *Metro Magazine,* Sept. 8, 2016. http://www.metro-magazine.com/management-operations/news/715324/mbta-selects-best-transit-app-winner. Accessed January 31, 2017.

The Ad Hoc Committee for the Chan Wui and Yunyin Rising Star Workshop peer-reviewed this paper.