Use of Mobile Ticketing Data to Estimate an Origin-Destination Matrix for New York City Ferry Service

**Subrina Rahman**, City College of New York, srahman007@citymail.cuny.edu  
**James Wong**, New York City Economic Development Corporation, jwong@edc.nyc  
**Candace Brakewood (Presenting)**, City College of New York, cbrakewood@ccny.cuny.edu

**Introduction**

Urban ferry service plays an important role in the transit network in New York City. The New York City Economic Development Corporation manages the East River Ferry (ERF), a privately-operated ferry route connecting Manhattan, Brooklyn and Queens in New York City. In 2012, the ERF became the first urban transit service in the United States to launch a mobile application for fare payment (Tavilla, 2015). This mobile ticketing application, or “app”, allows passengers to buy tickets directly on their smartphone using a credit or debit card (Brakewood et al., 2014). Using the app, passengers can buy a ticket at any time and then activate the ticket before they are ready for travel on the ferry. Because riders can purchase tickets for multiple ferry routes, the app prompts users for their origin and destination in order to determine the ticket type. This process of purchasing and then activating a mobile ticket creates a record in a backend database, which provides a rich source of data about where and when passengers are traveling on the ferry service.

**Objective & Methodology**

The objective of this research is to assess if the backend data from mobile ticketing apps can be used to create origin-destination (OD) matrices, which are a fundamental input to the transit planning process. To do this, mobile ticketing activation data were compiled for the same period as a recent onboard survey for the ERF, and both datasets (mobile ticketing and survey data) were used to create seed OD matrices. The seed matrices were adjusted to meet aggregate boarding and alighting ridership counts using iterative proportional fitting (IPF), which is a technique that is commonly used for transit OD estimation (e.g., Mishalani et al., 2013). Then, the different OD matrices were compared using Euclidean distance.

**Results**

The results of the Euclidean distance calculations comparing OD matrices are shown in Figure 1. Euclidean distance was calculated for each time period (AM, PM, midday and weekend) in the following four ways: (1) between the seed matrix of the mobile ticketing data...
and the seed matrix of the onboard survey data; (2) between the seed and final IPF matrices for the mobile ticketing data only; (3) between the seed and final IPF matrices for the onboard survey data only; and (4) between the final IPF matrix of the mobile ticketing data and the final IPF matrix of the onboard survey data.

![Figure 1 Euclidean Distance Calculation for Four Time Periods](image)

As can be seen in Figure 1, the range of Euclidean distances is between 0.04 and 0.18. In general, the AM and PM peak periods have smaller distances compared to midday and weekend periods. The most important result is the comparison of IPF adjusted matrices for mobile and onboard survey data, which is at the bottom of Figure 1. This shows that the AM and PM peak distances are very small (0.048 for the AM and 0.039 for the PM), whereas the midday and weekend distances are larger (0.086 for the midday and 0.088 for the weekend). These results suggest that the origin-destination data from this mobile ticketing app most closely aligns with the onboard survey data during the peak periods.

Additionally, the onboard survey included a small number of questions pertaining to travel behavior. These questions showed that the majority of peak period passengers were traveling for commuting purposes (92% during the AM and 83% during the PM peak), and that most peak period passengers were regular riders (82% during the AM and 67% during the PM peak typically take the ferry at least 4 times per week).

Based on the results of the survey questions combined with the OD estimation results, it can be inferred that most mobile ticketing users during the AM and PM peak periods are likely commuters and/or regular ERF riders. Therefore, mobile ticketing apps are likely to provide the
most reliable travel behavior information during peak periods when travel patterns are more consistent compared to the midday or weekend. In summary, this research demonstrates that mobile ticketing apps have the potential to provide valuable data about where and when passengers travel, which could be used for transportation planning more broadly in future analyses.

References

Notes
1. The views and opinions expressed in this abstract are those of the authors and do not necessarily represent those of New York City Economic Development Corporation or The City of New York.