The University of Chicago
MACS 30200 Research Paper Literature Review:
An Analysis of the Chicago Taxi Industry:
How can Traditional Taxi Industry Survive
in the Age of the Sharing Economy?

Zhang, Dongping *

April 24, 2017

1 Background

Uber launched in Chicago on September 22, 2011 and Lyft entered the Chicago market on May 11, 2013. A few years after those two ride-sharing company started to operate in the Chicago area, the city of Chicago reached an agreement with Uber and Lyft on November 25, 2015, allowing them to pick-up and drop-off passengers at O'Hare and Midway airports.

Even worse to taxi drivers, on January 1, 2016, taxi fare increased of about 15 percent in response to an increase in taxation of the Chicago area. Those two events together have significantly increased local competition for the taxi drivers after January 1, 2016.

2 Research Question

After Uber and Lyft started to operate in Chicago airports and after Chicago taxi fare increased of about 15 percent in 2016, how should Chicago taxi drivers operate to maximize their daily profit under significantly increased market competition?

*The University of Chicago, MA in Computational Social Science, CNetID: 12144965, dpzhang@uchicago.edu.

3 Literature Review

Cities have varying structures and layouts determined by their unique urban plannings. In order to identify optimal operating pattern and driving strategy of local taxi drivers, it is intuitive to first identify top-performing taxi drivers in the Chicago area by drivers' average daily income using taxi medallion ID.

After successful identifications of productive drivers, more refined research inquiries could be formulated, such as: Are those productive taxi drivers performing consistently over years? How do these successful drivers optimize over the bounded resources of space and time? Are these drivers profiting by deliberately choosing routes that are more costly for the customer, as every out-of-towner fears when stepping into a taxi in an unfamiliar city?

3.1 Identifying Top-performing Taxi Drivers

Liu et al. (2010) systematically studied the driving strategy of the top-performing cab drivers in one major city of China. They systematically composed a distribution of taxi drivers' earning profile of the city using a k-mean clustering algorithm so as to identify top-performing taxi drivers in certain time interval of a day. After identifying those productive drivers, they used two novel statistics composed by Turner (2009), which are ratio of real path length over shortest path length (RRSL) and ratio of real path travel time over shortest path travel time (RRST), to quantify those productive taxi drivers' operation patterns during corresponding time interval.

3.2 Spatial Flow Analysis

Ashbrook et al. (2003) used GPS data and created a predictive model of user's future movements based on user's past movements. They presented a system that would automatically cluster GPS data of individuals, taken over an extended period of time, into meaningful locations at multiple scales. These locations are then incorporated into a Markov model that can be consulted for use with a variety of applications in both single-user and collaborative scenarios. Although the location of taxi drivers would be largely determined by taxi demands in each spatial unit of the city at different temporal interval as well as the drop-off locations of previous passengers, Ashbrook et al.

(2003)'s predictive model could be incorporated when conducting a spatial flow analysis to answer questions such as would taxi drivers choose to drive in consistent routes when they pick-up and drop-off passengers from certain census tracts?

3.3 Pick-up Strategy

Another novel method to represent passenger-finding strategies for taxi drivers, or Taxi Driver Mobility Intelligence (TDMI), is by using time-location-strategy triplets according to time and locations of taxi drivers (Li et al., 2011). By representing the passenger-finding strategies in a Time-Location-Strategy feature triplet and constructing a train/test dataset containing both top-performance and ordinary-performance taxi features, Li et al. (2011) adopted a powerful feature selection tool, L1-Norm SVM, to select the most salient feature patterns determining the taxi performance.

Chang et al. (2009) recognized an issue that in urban area, the demand for taxis was not always matched up with the supply. They composed a four-step process: data filtering, clustering, semantic annotation, and hotness calculation. By mining historical data, they aimed to predict demand distribution with respect to contexts of time, weather and taxi locations.

3.4 Cost-saving Strategy

In addition to Li et al. (2011)'s work, Ge et al. (2010) also proposed recommendation systems for taxi drivers to optimize the cost of finding passenger by reducing gas emission and fuel consumption. Ge et al. (2010) proposed a concept called "Green Knowledge", which was defined as energy-efficient transportation patterns used as guidance for reducing inefficiencies in energy consumption of the transportation sectors. Ge et al. (2010)'s approach is to save energy consumption of taxi drivers and reducing the number of idle hours and dead miles (the driving miles when there is no passengers in the car), which is unarguably important cost-reducing strategy for Chicago taxi drivers.

3.5 Spatial Visualization

Visualization of trips and routes is particularly important and helpful when conducting temporal-spatial flow analysis. Ferreira et al. (2013) recognized the challenges of studying taxi trips data due to reasons such as data complexity, difficulty of specifying exploratory queries and performing comparative analyses, as well as the typical large size of the datasets. They proposed a new model in their paper that would allow researchers to visually query taxi trips, and to conduct origin-destination queries, which had enabled researchers to study taxi mobility across the city.

References

Ashbrook, D., & Starner, T. (2003). Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous computing, 7*(5), 275-286.

Chang, H. W., Tai, Y. C., & Hsu, J. Y. J. (2009). Context-aware taxi demand hotspots prediction. *International Journal of Business Intelligence and Data Mining, 5*(1), 3-18.

Ferreira, N., Poco, J., Vo, H. T., Freire, J., & Silva, C. T. (2013). Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. *IEEE Transactions on Visualization and Computer Graphics, 19*(12), 2149-2158.

Ge, Y., Xiong, H., Tuzhilin, A., Xiao, K., Gruteser, M., & Pazzani, M. (2010, July). An energy-efficient mobile recommender system. *In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 899-908). ACM.

Liu, L., Andris, C., & Ratti, C. (2010). Uncovering cabdrivers? behavior patterns from their digital traces. *Computers, Environment and Urban Systems, 34*(6), 541-548.

Liao, L., Patterson, D. J., Fox, D., & Kautz, H. (2006). Building personal maps from GPS data. *Annals of the New York Academy of Sciences, 1093*(1), 249-265.

Li, B., Zhang, D., Sun, L., Chen, C., Li, S., Qi, G., & Yang, Q. (2011, March). Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset. In Pervasive Computing and Communications Workshops (PERCOM Workshops), *2011 IEEE International Conference* on(pp. 63-68). IEEE.

Liu, L., Andris, C., Biderman, A., & Ratti, C. (2009). Uncovering taxi driver?s mobility intelligence through his trace. *IEEE Pervasive Computing*, 160, 1-17.

Liu, L., Andris, C., & Ratti, C. (2010). Uncovering cabdrivers? behavior patterns from their digital traces. *Computers, Environment and Urban Systems, 34*(6), 541-548.

Turner, A. (2009, September). The role of angularity in route choice. *In International Conference on Spatial Information Theory* (pp. 489-504). Springer Berlin Heidelberg.

Yuan, J., Zheng, Y., Zhang, L., Xie, X., & Sun, G. (2011, September). Where to find my next passenger. *In Proceedings of the 13th international conference on Ubiquitous computing* (pp. 109-118). ACM.