Understanding spatiotemporal mobility patterns using smart card data: a case study in Shenzhen

Yuanxuan Yang[*1], Alexis Comber[†1] and Alison Heppenstall[‡1]

[1]School of Geography, University of Leeds

Dec 26, 2016

Summary

This paper describes research analysing smart card data (16.3 million individual trip records) collected in Shenzhen, China over five weekdays to characterise commuters' mobility pattern and understand urban dynamics. A number of temporal and spatiotemporal features as well as regional characteristics were identified using approaches drawn from data mining and spatial analytics. Population segments were inferred and this research considered school or minor student travel flows and identified a number of unique mobility behaviours.

KEYWORDS: Geocomputation, Smart Card Data, Travel flow, Mobility Pattern, Metro System

1. Introduction

Exploring the citizen's mobility pattern and travel flow helps to understand the regional characteristics such as semantic functions and land use types and complex urban systems (Pan et al., 2013). Traditionally, household travel surveys are used to link individual's travel habits and socio-economic attributes. However, in many countries such as China, household travel surveys are conducted only every five or 10 years, in different cities, and requiring high cost and time resources.

Smart Card Data (SCD) of public transit systems records spatiotemporal data in a large quantity at individual level with low cost, providing the opportunity for new insight into individual travel flows and mobility behaviours, also revealing the spatial functions of cities. However, inferring the socio-economic background of card holders is still a challenge because the cards are anonymous. There have been few studies of SCD in China that have sought to identify the spatiotemporal pattern of certain social groups. Wang et al. (2016) and Long, et al. (2015) analysed university students (normally adults) and extreme transit commuters (above 15 years old) (Long, et al. 2015). Little research has been undertaken on the travel patterns of minor students.

Using data mining and spatial analytics, this study analysed two million individuals' travel flow and the spatiotemporal features in Shenzhen Metro System SCD, characterized the movement of citizens and the regional characteristics. Looking into the minor students' "digital footprint", this work also examined their particular travel flow in order to shed light on their mobility patterns.

2. Study Area and Data

Shenzhen, located immediately north of Hong Kong, is the southern mainland China's major financial and high-tech centre. With unprecedented urbanization and rapid urban transport development, Shenzhen built five metro lines by 2014. The metro lines also connect with the Hong Kong metro system at Futian Checkpoint Station.

SCD of Shenzhen Metro System collected from 9th June 2014 to 13th June 2014 was used in this work. The data include two million smart cards and related travelling information in the five working

[*] gyyy@leeds.ac.uk
[†] a.comber@leeds.ac.uk
[‡] a.j.heppenstall@leeds.ac.uk

days. The total number of trips is 16.3 million, with a mean of 3.26 million transit records per day. The data contains attributes describing smart card ID, trip price and discount price, station name, metro line number, train ID and tap-in/out status. Taking advantage of the above information, this work analysed the travel flow of two million individual travellers, identified as school students by their ticket pricing and a unique student discount (50%) for students in all primary school, secondary school, secondary technical school and technician training school. We identified 79 thousand students through their travelling records, thus giving insight into school students' daily mobility pattern inside metro.

3. Result and Initial Analysis

3.1. Temporal Mobility analysis

This research analysed travellers' trip frequency and temporal density, providing insight into their mobility pattern and revealing the regularity of these patterns.

In this work, each day was divided into 48 time intervals and the trip numbers in each interval in metro system were counted. Figure 1 shows the trip tap-in amount of all travellers and minor students during the weekdays, x-axis is the timeline by half hour and y-axis indicates the number of trips per 30 minutes. For all travellers, there are two peak hours in weekdays, one from 7:30 to 8:30 and another from 17:30 to 18:30. The students' morning peak is one hour earlier (from 6:30 to 7:30), this is because of the early school time (commencing at 7:30-8:00 in a different school). The number of all travellers' trips in the Friday evening is slightly higher than other days.

Two small student travel peaks emerge at lunch time due to returning home to eat. Therefore, the first one is for going home and the second one is the time that they go back to school. The lunch time peak of students on Friday is higher than other weekdays, which is similar to that of Singapore metro travellers (Zhong, et al. 2015). Moreover, students tend to have much more trips in the afternoon and evening on Friday: 7381 student trips, nearly twice the number of other weekdays. This may indicate that Chinese school students' travel pattern is different to other countries (Huang, et al. 2014).

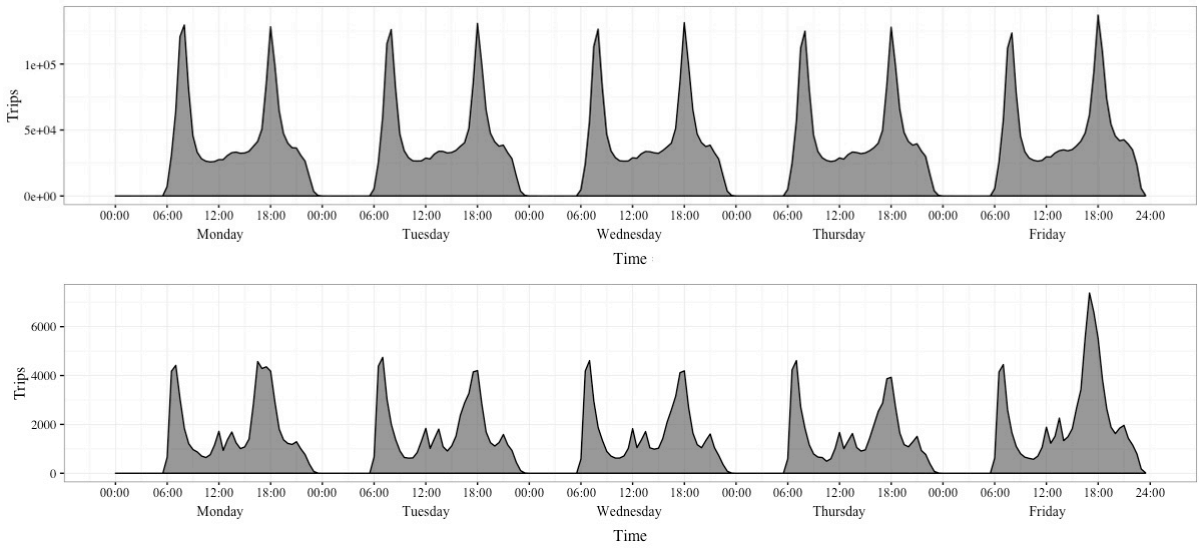
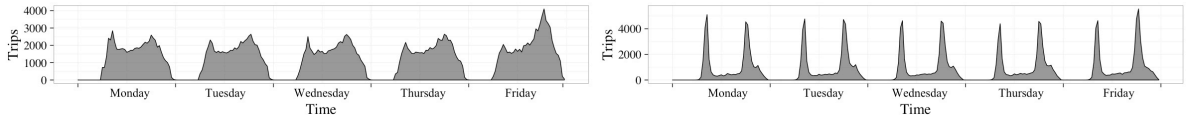


Figure 1. Number of All Travellers'(Top) and Students'(Bottom) Trips During the Five Weekdays (Half Hour*48*5)

This work also compared the temporal trip density of metro stations in different typical area. The temporal characteristics of stations related to scheduled activities, such as going to work or school, are much significant; while stations more related to free activities, have smoother density curve over time. For example, the trip density curve of Luohu station (next to Shenzhen railway station) is much smoother than that of Shenzhen university station (Figure 2).



(a) Luohu Station (b) Shenzhen University Station

Figure 2. Trip Density of Different Stations

3.2. Spatiotemporal analysis

This research studied the individual's flow by identifying and linking the origin and destination of each trip and then analysed the time, money cost and distance of trips. The OD(Origins-Destinations) matrix of travels also helps to identify where people live and where they work/study. For example, the origins and destinations of students' morning ride (from 6:30 to 10:00 am) indicate the location of their home and schools. Figure 3 and Figure 4 show the distribution and number of tap-in and tap-outs of the students' morning ride to school. The radius of the metro station dot is proportionate to its normalized number of tap-in/out. The tap-in occurred most in Buji, Tai'an, Hubei stations, and students' favourite morning ride destinations are Dajuyuan, Lianhuaxi and Yijing stations. There are also many tap-out records in Futian checkpoint station which is connected to Hong Kong metro system, indicating that some school students in Shenzhen choose to go to Hong Kong for school.

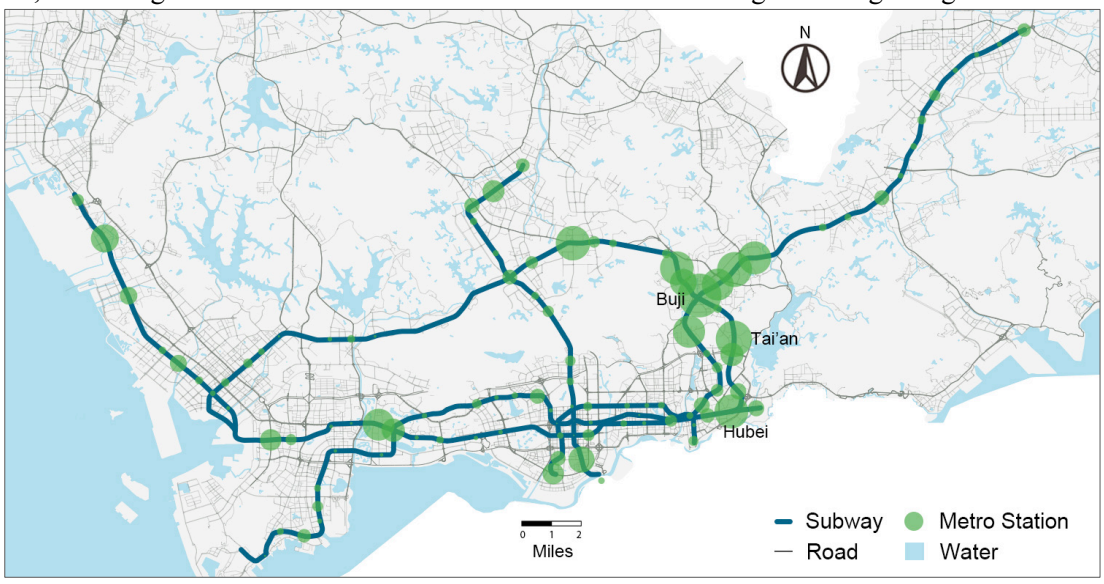


Figure 3. Distribution and Number of Tap-in of Students' Morning Ride in Metro System

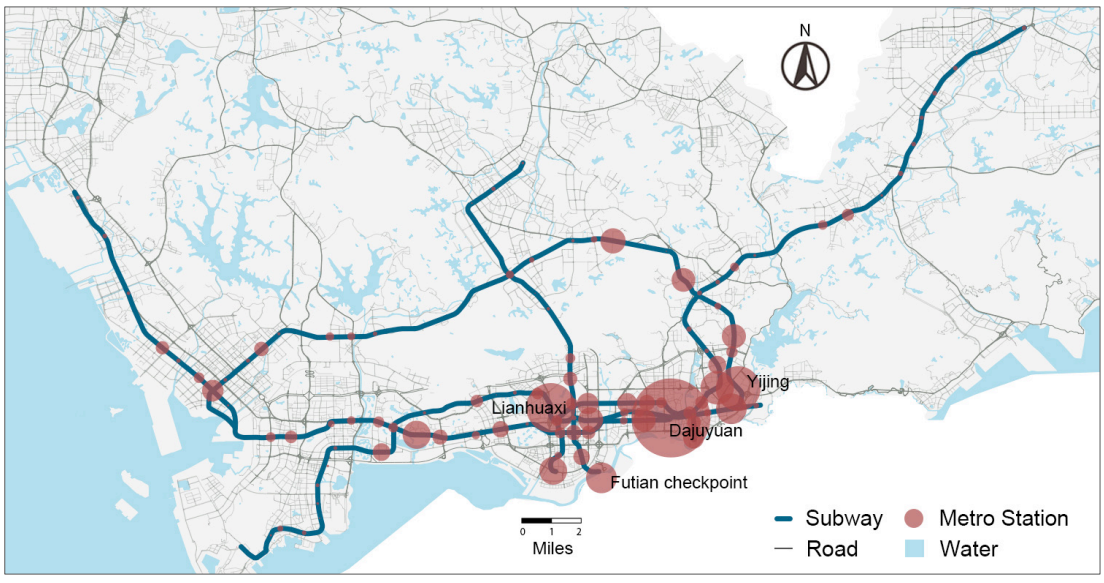


Figure 4. Distribution and Number of Tap-out of Students' Morning Ride in Metro System

4. Discussion and Outlook

Cities can be regarded as interactions that flow across networks (Batty & Cheshire, 2011). Using the SCD collected in Shenzhen Metro system, this work studied the temporal, spatiotemporal mobility pattern of citizen and minor students, compared the travellers and trip density of metro stations in different typical area, and continues to reveal the related regional characteristics.
Future work will link metro SCD with bus SCD, bus GPS and Taxi GPS data to gain a better understanding of the complex urban system and the regional characteristics. Associated with social network(Weibo) check-in POI data, it is also possible to shed light on attractiveness of places as well as the relationship between transit and urban land use.

5. Acknowledgements

We would like to thank University of Leeds and Chinese Scholarship Council for funding the project.

6. Biography

Professor **Lex Comber** holds a Chair in Spatial Data Analytics at the School of Geography. Lex is a leading international researcher in many areas of spatial science and geocomputation, with publications in accessibility, facility location optimisation, graph and network theory, spatial data uncertainty, citizen science, land use / land cover and remote sensing.

Professor **Alison Heppenstall** holds a Chair in Geocomputation. Her interests are in the application of artificial intelligent solutions for geographical problems, with applied research experience in the development and linkage of novel methodologies for a variety of socio-economic applications including education planning/management, retail analysis and crime. A particular focus of her work is in individual based modelling, in particular the development and application of agent-based modelling and microsimulation to geographical systems.

Mr. **Yuanxuan Yang** is currently studying (first year) for a Ph.D. at the School of Geography, University of Leeds. His interests are in big data, spatial microsimulation and land use/land cover.

Reference

Batty, M., & Cheshire, J. (2011). Cities as flows, cities of flows. *Environment and Planning B: Planning and Design*, *38*(2), 195-196.

Huang, X., & Tan, J. (2014). Understanding spatio-temporal mobility patterns for seniors, child/student and adult using smart card data. *International Archives of the Photogrammetry Remote Sensing & S, XL-1*, 167-172.

Long, Y., Liu, X., Zhou, J., & Chai, Y. (2015). Early birds, night owls, and tireless/recurring itinerants: an exploratory analysis of extreme transit behaviors in Beijing, china. *Computer Science, 57*, 223–232.

Pan, G., Qi, G., Wu, Z., Zhang, D., & Li, S. (2013). Land-use classification using taxi GPS traces. *IEEE Transactions on Intelligent Transportation Systems*, *14*(1), 113-123.

Wang, M., Zhou, J., Long, Y., & Chen, F. (2016). Outside the ivory tower: visualizing university students' top transit-trip destinations and popular corridors. *Regional Studies, Regional Science*, *3*(1), 202-206.

Zhong, C., Manley, E., Arisona, S. M., Batty, M., & Schmitt, G. (2015). Measuring variability of mobility patterns from multiday smart-card data. *Journal of Computational Science, 9,125-130*.