Predicting short-term travel demand in bike sharing system with neural networks

Yuanxuan Yang
School of Geography,
University of Leeds
Leeds, UK
gyyy@leeds.ac.uk

Alexis Comber
School of Geography,
University of Leeds
Leeds, UK
a.comber@leeds.ac.uk

Alison Heppenstall
School of Geography,
University of Leeds
Leeds, UK
A.J.Heppenstall@leeds.ac.uk

Andy Turner
School of Geography,
University of Leeds
Leeds, UK
A.G.D.Turner@leeds.ac.uk

Abstract

Understanding the dynamics and travel demand is curtail for the management of bike sharing schemes, also improving mobility in none-motorised traffic. This work compared the performance of different prediction methods on forecasting short-term bike travel demand using a case study. Results suggests that neural network obtained the best performance in terms of accuracy and robustness.

Keywords: travel demand, machine learning, neural network, bike sharing, traffic, urban dynamics

1 Introduction

Bike sharing schemes can contribute towards improving air quality and reducing congestion in cities as a part of a sustainable travel infrastructure (Shaheen et al., 2011). Its popularity has increased in the last few years globally mainly due to its advantages in cost and convenience over other forms of transport. Understanding the dynamics of bike travel demands in different areas is important for scheme efficiency, maintenance and bike fleet management.

Traditionally, activity-based models of travel demand have been used to understanding the mechanisms behind travel choices and their spatial context and to help predict travel demand in urban environments (Ortúzar et al., 2000; Rybarczyk & Wu, 2010). These methods normally focus on the supply side, using the attributes of individuals and local land use characteristics to forecast travel demand. They place less emphasis on understanding and predicting the highly dynamic interactions between different elements in urban systems. For example, gravity-based models of travel demand are perhaps the simplest and provide an effective representation of the spatial interaction between locations. Alternatively, utility-based destination choice models consider various behavioural factors driving underlying travel demand (Ortúzar et al., 2000; Rybarczyk & Wu, 2010). However, they normally characterize travel demands over low temporal granularities, despite the fact that traffic flow and demand is a highly dynamic process and changes constantly. For instance, Al-Ayyash et al (2016) modelled demand using time slices of one week. Therefore, activity-based models are not ideal for short-term prediction, which typically refers to time slices of no more than one hour (Ermagun & Levinson, 2016).

Transportation systems now generated large quantities of dynamic consumer data, such as bus tickets, bike rental, taxi transaction and GPS records (Birkin, 2018). They can be used for short-term travel demand forecasting using machine learning methods such as artificial neural network (Goves et al., 2016; Kumar, et al. 2013), Bayesian network (Wang et al.,2017) and Long-Short Term Memory network (deep learning) (Xu et al., 2017). Deep learning is a form of machine learning which has the potential to provide good short-term forecasts of traffic flows/demand by exploiting the dependencies in highly dimensional sets of explanatory variables. How well different temporal machine learning approaches support predictive forecasting of travel demand remains an unanswered question.

This work used deep neural network to predict short-term travel demand in bike sharing scheme and compared its performance against several other prediction methods.

2 Data

The data used in this analysis were traffic and meteorological data. Traffic data was from the New York bike sharing data (citi bike scheme) and each record include detailed travel information: trip start time, trip end time, start station, end station, user id, user gender, coordinates of start/end stations. The data records transactions at 827 docking stations. Meteorological data were obtained from openweather.com. It describes hourly weather information including rain, haze, scattered clouds, humidity, pressure, temperature and wind speed. Data from a continuous five-month period (2017/06/01-2017/10/31) were used.

3 Methods

3.1 Spatial data preprocessing

Bike docking stations were grouped into clusters based on their spatial proximity, to avoid predicting travel demand at each station (Chen et al., 2016; Li et al., 2015). The reasons are two-fold: (1) The dynamics of individual station is too chaotic to predict, because the status of any individual station may be impacted by the status of nearby stations (empty or full docks); (2) From the perspective of management, understanding the behaviour of a small group of stations is sufficient for bike rebalancing strategies and fleet management. Bike docking stations were firstly clustered into 36 groups using hierarchical clustering (Figure 1). The

number of 36 is an arbitrary choice since too many clusters are meaningless and chaotic, and too few fails to represent system dynamics in a large area, and to indicate spatial heterogeneity. The prediction models in this research were applied to the clusters rather than individual stations. 80% of data were randomly choose for model training, the remaining 20% were used for testing.
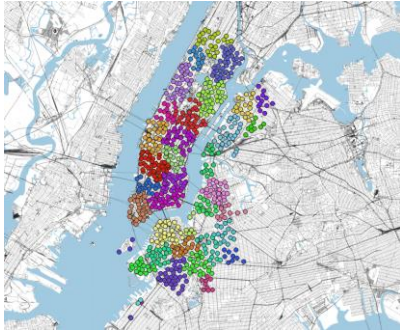

Figure 1: Bike dock station clusters

3.2 Prediction methods

Four methods are compared: two baseline methods and two machine learning methods.
1) Baseline methods: The following methods create naïve prediction results, but provide a performance baseline for comparison with other models.
   - Last Hour (LH): This predicts the bike check-outs as a proxy for travel demand as the demand value of the last hour.
   - Historical Average (HA): This predicts the check-out number using the historical average in corresponding periods. E.g., for 1:00pm-2:00 pm on Tuesday, the corresponding periods are all of the historical time intervals from 1:00pm-2:00pm on weekdays.
2) Machine learning models:
   - GBRT (Gradient Boost Regression Tree): This method has been used in a number of recent traffic prediction models (e.g. Alajali et al., 2017) It is considered to be a strong prediction model, fits complex nonlinear relations and has the ability to handle different types of predictor variables.
   - (Neural Network): A Deep Neural Network is used in this work, constructed using the Keras and TensorFlow R packages. A feed-forward network structure was used, with 3 tensor layers and two drop out layers (dropout rates are 0.1 and 0.5 respectively) to overcome overfitting problems.

3.3 Feature selection

A number of temporal and meteorological features were selected for the GBRT and Neural Network approaches.
   - Temporal: day of month, day of week, holiday, weekday or weekend, hour of day

   - Meteorological: humidity, pressure, temperature, wind speed, condition description (e.g. heavy rain, cloudy, etc.)

Category information such as weather condition description were processed using one-hot encoding.

4 Preliminary results and analysis

Figure 2 shows the hourly travel demand (check-outs) of the 36 clusters in a week in 2017. The first two peaks - September 2nd and September 3rd - are at the weekend and their patterns are very different from the weekdays (Sep 5th – Sep 8th). Sep 4th is Labour day, a US Federal Holiday and its pattern is similar to a weekend. The repeated dynamics indicates dynamics in bike sharing schemes data have the potential to predict future demand from learning of historical temporal patterns.
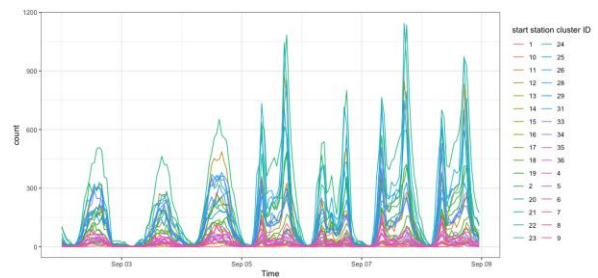

Figure 2: Travel demand (check-out amount) of different clusters in a week (Sep. 02 – seep 08, 2017)

The results of the different methods described in Section 3.2 are shown in Table 1. Firstly, we examined MAE (Mean Average Error) and MRE (Mean Relative Error) of the different models in all time periods. These two indicators help to understand the overall performance of each method. Secondly, short-term cycling activity is sensitive to local events or extreme weathers, as such conditions can lead to anomalous changes in hourly travel demand. Therefore, we identified periods of "raining" weather, and calculated the MAE and MRE separately, to measure model performance in these anomalous periods.

Table 1: Performance of different methods

Methods	All periods		Raining periods	
	MAE	MRE	MAE	MRE
LH	23.65	0.3378	18.59	0.4575
HA	14.60	0.2085	25.90	0.6371
GBRT	13.98	0.1996	16.75	0.4119
NN	10.90	0.1557	9.35	0.2301

Table 1 indicates that Neural Network results in the best predictions in both circumstance (all time periods and raining periods), with the lowest MAE and MRE. GBRT is slightly better than HA in all time period, but much more accurate (MRE 0.4119) in raining weather conditions than HA (MRE 0.6371). This is because GBRT included weather information

while HA and LH did not. LH is the worst over all periods, but it is better (0.4575) than HA (0.6371) in raining periods. Overall, the performance of NN is best among the four methods, especially in anomalous periods with rainfall.

5   Discussion and future work

This work compared the performance of different methods for predicting short-term travel demand in a bike sharing schemes. The results indicate that neural network shows the best performance in all period and in abnormal periods (i.e. those with rain). The results confirm the advance and robustness of neural network in traffic prediction. Future work will include comparing more methods, such as support vector machine, random forest etc., and will explore other neural network structures, for example, Long Short Term Memory (LSTM) networks, to determine the best NN structure for predicting short-term travel demand.

A further important question is how can machine learning methods be used to analyse spatiotemporal graph structures and related information (centrality, flux etc) to forecast travel demand in urban areas. The idea of viewing individual's and collective travel behaviour as graph or network is important and not new. Numerous studies have used this to study the spatial behaviour of individuals and the dynamics in whole urban systems (Batty 2013). Travel demand, consisting of large quantity of origin-destination trips, can be viewed as a large-scale weighted directed graph. The related network-based analysis (graph theory) has potential to shed light on spatial interdependencies and interactions between location in a more comprehensive manner. The underlying graph properties are also likely to help improve the performance of travel demand. This is because, in general, network-based measures provide a quantifiable picture of interactions and interaction strengths between origins and destinations. Such information about connectivity between places is not fully captured by hourly depart and arrive data as analysed in this and many other studies.

While network-based characteristics of travel demand present a deeper understanding of interactions between places in a city, how to fully incorporate them in existing modelling approaches remains an open question. A direct and naive way might be to integrate additional graph structure information into the feature dataset. Data could then be sliced into short equal intervals in terms of temporal dimension, various spatiotemporal features and graph structure information are derived accordingly. Figure 3 shows how measures of degree, betweenness, closeness and PageRank of different station clusters change over a weekday in the bike sharing scheme. Degree (Figure 3 a) and closeness (Figure 3 c) show similar rhythms to travel demand in Figure 2, while betweenness (Figure 3 b) and PageRank (Figure 3 d) are more chaotic. Other graph structure properties such as modularity and connectivity will also be extracted used in machine learning models in future studies. Since various graph information implies the dynamics and interdependence in different part of the system, their potential in strength machine learning prediction models should not be ignored.
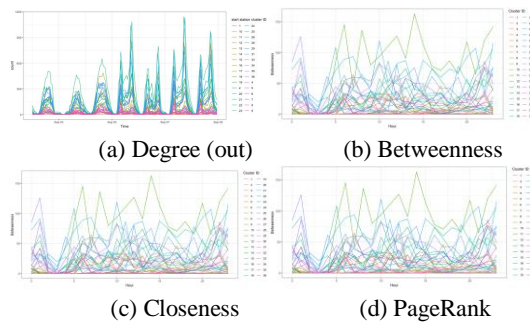

(a) Degree (out)   (b) Betweenness

(c) Closeness   (d) PageRank

Figure 3: Graph structure properties

6   References

Alajali, W., Wen, S., & Zhou, W. On-street car parking prediction in smart city: a multi-source data analysis in sensor-cloud environment. In *proceedings of SpaCCS*, Melbourne, 2017.

Al-Ayyash, Z., Abou-Zeid, M., & Kaysi, I.. Modeling the demand for a shared-ride taxi service: an application to an organization-based context. *Transport Policy*, 48: 169-182, 2016.

Batty, M. The New Science of Cities. *The MIT Press*, London, 2013.

Birkin M. Spatial data analytics of mobility with consumer data. *Journal of Transport Geography*, 04(12):1-9, 2018.

Chen, L. , Zhang, D. , Wang, L. , Yang, D. , Ma, X. , & Li, S. , et al. Dynamic cluster-based over-demand prediction in bike sharing systems. In *Proceedings of ACM International Joint* Conference on Pervasive & Ubiquitous Computing. Heidelberg, 2016.

Ermagun, A., & Levinson, D.. Spatiotemporal traffic forecasting: review and proposed directions. *Working Papers*, 1-29, 2016.

Goves, C., North, R., Johnston, R., & Fletcher, G.. Short term traffic prediction on the uk motorway network using neural networks ☆. *Transportation Research Procedia*, 13: 184-195, 2016.

Li, Y., Yu, Z., Zhang, H., & Lei, C.. Traffic prediction in a bike-sharing system. In *Proceedings of International Conference on Advances in Geographic Information Systems*, Seattle, 2015.

Kumar, K., Parida, M., & Katiyar, V. K. Short term traffic flow prediction for a non-urban highway using artificial neural network ☆. *Procedia - Social and Behavioral Sciences*, 104: 755-764, 2013.

Ortúzar J de D, Iacobelli A, Valeze C. Estimating demand for a cycle-way network. *Transportation Research Part A: Policy and Practice*. 34(5):353-373, 2010.

Rybarczyk, G., & Wu, C.. Bicycle facility planning using gis and multi-criteria decision analysis. *Applied Geography, 30*(2): 0-293, 2010.

Shaheen, S., Guzman, S., & Zhang, H..Bikesharing in Europe, the Americas, and Asia. *Transportation Research Record: Journal of the Transportation Research Board, 2143*: 159-167, 2010

Wang, Q., Sun, H., & Zhang, Q.. A Bayesian network model on the public bicycle choice behavior of residents: a case

study of Xi'an. *Mathematical Problems in Engineering*, 2017(4): 1-13, 2017.

Xu, J., Rahmatizadeh, R., Boloni, L., & Turgut, D..A Sequence Learning Model with Recurrent Neural Networks for Taxi Demand Prediction. In *Proceedings of IEEE Conference on Local Computer Network*s, Singapore, 2017.