A PREDICTION OF BIKE FLOW IN BIKE RENTING SYSTEMS WITH THE TENSOR MODEL AND DEEP LEARNING

Yerkezhan Seitbekova
Kazakh British Technical University, Almaty, Kazakhstan
Bakytzhan Assilbekov
Satbayev University, Almaty, Kazakhstan
Alibek Kuljabekov
Satbayev University, Almaty, Kazakhstan
Iskander Beisembetov
Satbayev University, Almaty, Kazakhstan

ABSTRACT

Rental bikes are popular in many urban areas to help people expand their mobility. It is important to make the rental bicycle usable and available to the general public at the appropriate time and place. Inevitably, providing the city with a steady supply of rental bicycles becomes a major concern. The most important aspect is the estimation of the number of bicycles required in each bicycle sharing station at any given hour. This paper gives an examination of human mobility as indicated by bicycle renting information of the bike sharing system. In this paper, we proposed a new approach for forecasting the bike inflow and outflow from one station to another during certain time slots. Our method analyses human mobility pattern by two steps: (1) Using Tuckers tensor decomposition to create a 3D tensor to model human mobility and extract latent temporal and spatial characteristics of various stations and time slots. (2) to use a Long-Short Term Memory

Neural Network to model the relationship between mobility patterns and the derived latent spatial and temporal features in order to predict bike flow between stations. The main contribution of this study that with the extracted latent characteristics through Tuckers factorization we improve the accuracy of prediction by 16% and decrease the amount of training data that used in prediction. Also, a root mean squared error of prediction is 1,5 bike.

We compare our model with baseline models as historical average, ARMA, the feed-forward neural network, and KNN. The proposed method showed the best results

1. INTRODUCTION

Public bike rental systems have become popular in recent years. People increase attention to bicycles because of their flexibility, low cost, and benefit for health. The list of cities, that provides public bike-sharing systems are still growing. A bike-sharing world guide can be found from the following link www.bikesharingmap.com.

For the ideal execution of such frameworks, there must be (a) the likelihood of finding a bike when the client needs to start the ride, (b) the likelihood of leaving the bike at the client's destination, and (c) the distribution of bike stations around the city.

To avoid overgrowing the structure, there are two different ways to solve these problems: firstly, inform the client in advance about where to get or leave bicycles, and improve the redistribution of bicycles from entire stations to empty ones.

In this study, we developed a solution to these problems by analysing cyclical mobility models, which we then used to predict hourly the number of bicycles available at stations by predicting inflow and outflow between stations. These predictions will be made by the current bike rental systems; increase user satisfaction with the system. Knowing the established patterns that follow people can lead to optimization of the bike sharing system, forcing the operator to predict in advance overcrowding or shortage of bicycles at precise stations and optimize their redistribution according to the situation.

The understanding of human mobility and find traveling patterns of passengers is crucial for public transportation systems, taxi providers, and bike renting companies because predicting it accurately can increase citizens living conditions and increase profit of transportation companies. The transportation-related works published in different fields, such as sociology, urban planning, computer science and other areas. In this section, some studies on human mobility discussed.

Many machine-learning algorithms were used to predict the demand for bicycles. Jia et.al (2019) used a Gaussian mixture model for the bike-sharing system clustering to group stations by migration trends, then apply the gradient boosting regression tree to predict renting traffic. Sathishkumar et.al. (2020) applied the data mining technique to predict hourly rental bike demand and found that Gradient Boosting Machine showed the highest and best result. Graph structured information was added to deep learning models in the study of Yang et.al. (2020) to short term forecasting of travel demand. Feng et.al. (2017) analysed the future availability of bicycles at cycle stations using instant analysis of a Markov chain model with continuous time and time-dependent metrics. A random forest model applied in work of Huang et.al. (2013) to predict the demand of bikes, and then authors applied the hub-firs-route-second bike repositioning technique to redistribute bikes.

Raviv et. al. (2013) analyses bike-sharing systems imbalances caused by various levels of attractiveness and generation of station-level trips. Systems and Lacker (2013) and Garcia-Palomares (2012) provides efficient bike redistribution strategies to reduce bike distribution imbalance. With a similar goal of introducing a more balanced system, other studies by Khatri (2015) modeled demand or developed models that optimize the location of stations. Wergin and Buehler (2018) have focused on the GPS analysis of casual cyclists' routes. Artificial neural networks have been widely used in several fields of transportation engineering. The work of Polson and Sokolov (2017) found the fact that the deep learning architecture can capture spatio-temporal effects and deep learning provides accurate short-term traffic flow predictions. The temporal relationship is significant in the task of a forecasting time series. LSTM is a time series prediction algorithm that designed to merge short-term and long-term time information with good forecasting performance.

Zhao et al. (2015) proposed an LSTM model in which the two measurements straightforwardly shown to the spatial-temporal association with inspect the spatio-temporal connections in busy time gridlock transfer. The work of Yu et al. (2017) proposed a combined deep LSTM approach that uses deep LSTM to reenact normal traffic from case in exceptional conditions. Ma et al. (2015) used another deep LSTM on far off microwave sensor information to catch non-linear traffic elements.

2. METHODS

In this work, to predict the demand of bikes, first, we modelled human mobility patterns using tensor decomposition, and then this pattern was used to estimate the number of bikes to be taken and returned to stations. We compress all historical bike trip data to the 3-dimensional tensor to reduce the amount of input data for training model as shown in Fig.1. We use tensor to extract mobility patterns. Only this pattern is used to train model. Also, for training the model we use last month's trip information and weather condition information. This section describes the tensor model, a model for prediction, and input parameters to the model.

2.1 Tensor model

According to the study of Kolda and Bader (2009) tensor is an array with more than three-dimension. A higher order tensor decomposition used to compress a volume of data or to find some dependencies between data. Tensor decomposition is widely used in graphical analysis, numerical analysis, computer vision, data mining, neuroscience, etc. In this article, we propose to simulate the movement of a bicycle between different stations using a three-dimensional tensor $\mathcal{H} \in \mathbb{R}^{N \times N \times L}$, as shown in Figure 1.a. The first tensor dimension $\mathcal{H}$ means the identifiers of the source cycling stations, the second dimension means the N identifiers of the destination stations, and the third dimension means the L time intervals.

Each element of the tensor $\mathcal{H}\ (i, j, l)$ stores an average amount of trips from station i to station j over a period of time l. With this tensor model, we extract the latent spatial characteristics of each source station, destination station, and the latent temporal characteristics of each time slot using the Tucker decomposition.

The Tucker is a higher-order principal component analysis technique (PCA). In each dimension, it decomposes a tensor into a base tensor multiplied by a matrix. In our case, we decompose the tensor $\mathcal{H}$ into three matrices $o \in \mathbb{R}^{N\times P}$, $Sd \in \mathbb{R}^{Q\times N}$, $T \in \mathbb{R}^{L\times R}$ and the base tensor $G \in \mathbb{R}^{P\times Q\times R}$, as shown in Fig. 1. In terms of a mathematical formula, this relationship can be written as in equation (1):

$$\mathcal{H} \approx G \times {}_1So \times {}_2Sd \times {}_3T \tag{1}$$

The feature vector indicating the characteristics of origin station i is the row i of matrix So after tensor factorization. The feature vector indicating the characteristics of destination station j is the same, the row j of matrix Sd, Sd_j. T_k is a feature vector that indicates the quality of the time gap k. The degree of cooperation between different components of So, Sd, and T is specified by each element of the core tensor G.

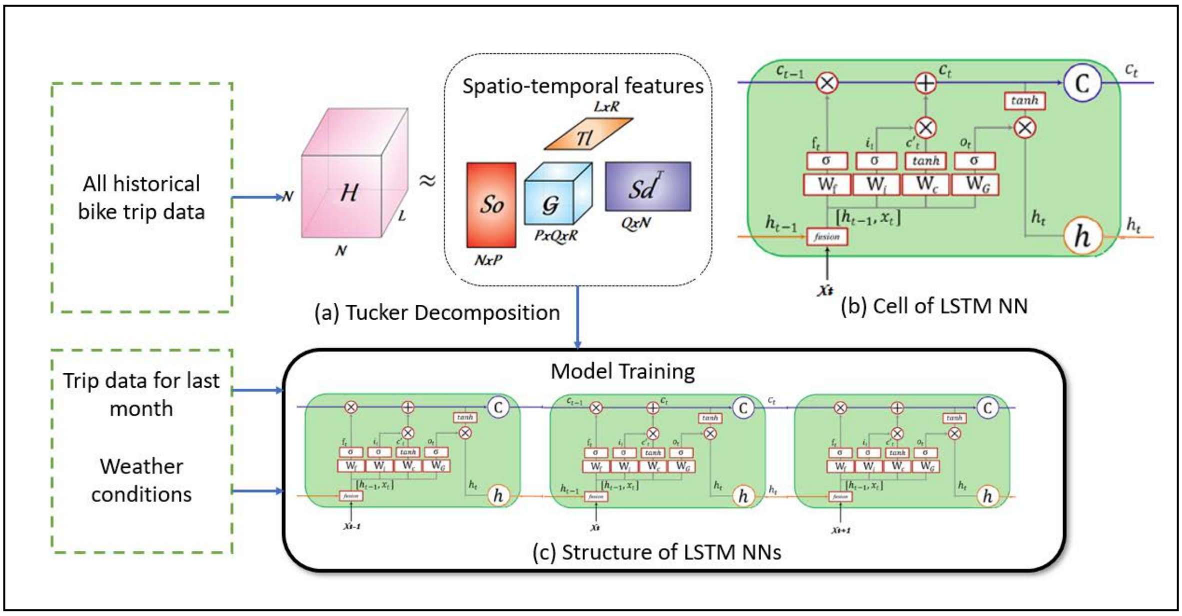

Fig. 1 – Illustrations of LSTM NN and its inputs.

The volumes of outflow $x_{oj:k}$ and inflow $x_{lj:k}$ of a station i at time gap k are under on their hidden spatial features So_i, Sd_j, latent temporal features T_k , and their previous values, $x_{oi:k-1}$, $x_{li:k-1}$ respectively. This is verified using the PC algorithm in the work of Guo and Karimi (2017). To see this dependency, we draw the values of the matrix T when P, Q, R is equal to 1 (Fig. 2). According to the graph, we have two maximum points at time slot (8-9) in morning rush hour, and at time slot (17-18) in evening rush hour. In addition, there is a minimum number of bike riders in the early morning. Likewise, the sum of all bikes taken from one station is high correlated with So matrix. Correlation coefficient is 0.806. It means that these values can be used to estimate the demand for bikes.

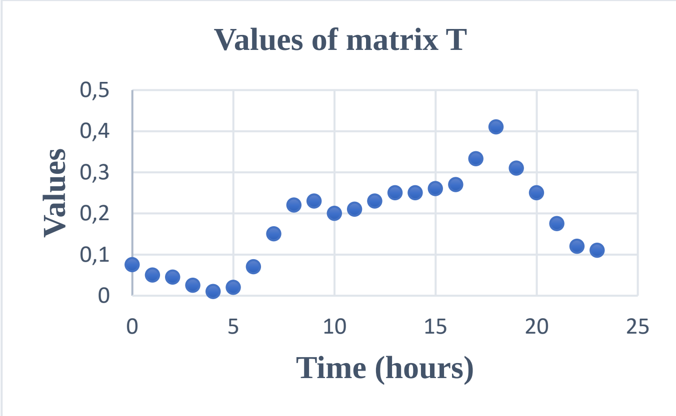

Fig. 2 – Values of matrix after tensor decomposition

2.2 Model Inputs

Anticipating the inflow and outflow of groups in each the bicycle sharing station is exceptionally testing, influenced by the accompanying three complex components:

Spatial conditions. The inflow of one station is influenced by outflows of another close-by stations. In like manner, the outflow of one station would influence the inflows of other stations.

Temporal conditions. The flow of groups in a station is influenced by ongoing time stretches, both all over. For example, a traffic flow happening at 8 am will influence that of 9 am.

Moreover, traffic conditions during morning times of heavy traffic might be comparable on successive workdays, rehashing like clockwork. Moreover, morning heavy traffic times can continually occur later when winter comes. At a time when temperatures constantly drop and the sun rises later, people rise later and later.

External conditions. Some external variables, such as climatic conditions and circumstances, can dramatically alter the flow of groups in different locations in a city. There spatial conditions and worldly conditions are separated from Tucker decomposition. The external impact is very troublesome. The traveler streams can be influenced by different outside elements, for example, climate and occasions. To investigate the impacts of these variables, Liu et al (2019) looked at traveler streams under various conditions. The precipitation information got for the examination records the evaluation of precipitation. Creator contrasted with a typical day, occasion, and end of the week have an obvious impact on the traveler streams. For example, on the off chance that a hefty downpour is seen on Monday, we will at that point contrast it and the information of past Monday to guarantee that they have similar qualities. To see is there some closeness in non-weekend days we look at them.

We found that the use of bicycles does not generally rely upon workday or end of the week data. At last, non-weekend day data, season of the day, and grade of precipitation is included as info boundaries for the demonstrating. Ashqar et.al. (2019) investigated climate conditions on bicycle checks and found that mugginess, season of-day, and temperature are noteworthy indicators of bicycle tallies. These boundaries were taken as an outer impact for forecast.

For metadata (for example, day of the week, hour of the day, and rainfall estimates), the insertion procedure is used to plan irreducible qualities in 3D vectors. It should be noted that the provisioning suite and the test suite are processed using the same bounds. At this point, the model itself can learn the data, which improves the accuracy of the predictions.

3. RESULTS AND DISCUSSION

The methodology given in this paper is trained and tested on public data available on web site https://www.citibikenyc.com/system-data. There can be found information about all CitiBike New York users' trips and annual monthly reports from May 2013. The data used in work contain records from 1st January 2017 to 31st December 2017 and January 2018. We use all data for 2017 to build 3D tensor and used data of January 2018 to test and train the model. From January 2018, we chose the last 10 days as the testing set and the left samples as the training set. In this case, the time period for traffic aggregation for forecasting is 1 hour. It's worth noting that if the chosen time interval is too short, the forecast would be incorrect and meaningless. Furthermore, short-term flows are often trivial, making the prediction approach difficult to use. We used the Min-Max normalization technique to scale the passenger traffic data in the range [-1, 1] for both the training and test sets. During estimation, the normalized predicted values are scaled and compared to the actual performance.

The results of the proposed method are compared with the following baseline models:

Historical average: We use the average historical supervision at the same time gap, the average of the past week's bike flow at the same time gap is set as the prediction outcome.
ARMA: Autoregressive moving-average model - is the mathematical model used for the forecasting and analysis of stationary time series data in statistics. The ARMA model combines two simpler time series models - the autoregressive (AR) model and the moving average (MA) model. The ARMA(p, q) model, where p and q are numbers showing the order of the model, a time series {Xt} is generating by next process:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j}, \quad (2)$$

Where c is a constant, ε_t is white noise, that is, a sequence of identically and independent normally distributed random variables, with zero mean, and $\alpha_1, \ldots, \alpha_p$ and $\beta_1, \ldots, \beta_q$ – autoregressive coefficients and moving average coefficients.

This model can be explained as a linear multiple regression model, in which the previous values of the dependent variable itself go as illustrative variables and moving averages of white noise elements go as the regression remaining.

FFNN: The dynamic nonlinear relationship between different variables can be captured by a feedforward neural network. We use the passenger flow in recent tree adjacent time intervals to reflect temporal dependencies in FFNN. (i.e. $[x_{t-1}, x_{t-2}, x_{t-3}]$).

KNN: K-nearest neighbours' algorithm, k-NN - a metric algorithm for automatic classification of items or regression. In our case, the object is given an average value based on the k objects closest to it, whose values are already known, using the regression method. In this article, we use a value of 4 for k..

The study's efficiency metrics are as follows: Root Mean Square Error (RMSE), , Mean Absolute Error (MAE), Symmetric Mean Absolute Percent Error (SMAPE), and Mean Relative Error (MRE).

The proposed methodology is assessed using RMSE, SMAPE, MRE, and MAE. The outcomes are shown in Table 1.

Performance metrics	HA	ARMA	FFNN	KNN	LSTM	LSTM ST
SMAPE	92.5	88.	48.1	46.4	46.1	44.2
RMSE	3.3	3.1	2.5	2.4	2.4	1.5
MAE	2.1	1.9	1.5	1.5	1.5	1.2
MRE	72.8	66.8	65.7	61.9	59.2	47.6

Table 1- Performance comparison

Different variation models were checked and compared to the baseline outcome to assess the forecast accuracy of the models in terms of RMSE, SMAPE, MRE, and MAE. Table 1 shows that using only the average bike flow over the previous time period (Historical average) results in a lower forecast.. Autoregressive models show better result than just calculating average. The accuracy remarkably improved when we apply FFNN, KNN and LSTM NN models. If we compare results of FFNN, KNN and LSTM NN, LSTM NN is better than other models, because LSTM NN model best suits for time series forecasting problem.

Models FFNN, KNN, LSTM accuracy approximately the same, the error in the number of bikes is about 2 or 3. But after applying the Tucker decomposition result error decreased to 1 or 2 bikes. That is good enough to help a bike managements system to improve the availability of bikes and docks in stations. The Tensor model increased forecasting accuracy for RMSE metric to 15 % from 2.43 to 1.53

4.CONCLUSION

We proposed a method for forecasting spatial-temporal bike mobility patterns, namely the inflow and outflow of bicycles from one station to another during a time gap. Our approach consists of two stages: (1) using a 3D tensor to model human mobility and recovering hidden spatial and temporal characteristics such as origin and destination stations and timeframes through tensor factorization; and (2) determining a connection between mobility patterns and recovering hidden features using a Long-Short Term Memory Neural Network for human mobility forecasting. We conduct a study of bicycle trips in New York City in order to validate the proposed technique. The results showed that the recover hidden features effectively identify attributes of timespans and spatio - temporal features with a strong correlation coefficient with bicycle sharing station inflow and outflow. The proposed method for extracting hidden characteristics can be applied to existing models to increase precision (MAE error is reduced by 16 percent).

By picking up from past historical bike rent data and past weather information, the proposed LSTM model with Tucker decomposition results can foresee the interest in bikes at a particular time. Considering the estimate, we can make the proposal for bike associations about how to scatter the bikes expressly to each station to satisfy the need of customers similarly as saving a silly cost of keeping bikes. The use of the proposed model will be a useful answer for both the bike renting organizations and the bike riders

ACKNOWLEDGEMENTS

This work has done under the project № AP05134776«Location Analytics Techniques for Prediction of Mobility Patterns» of the Ministry of Education and Sciences of the Republic of Kazakhstan.

REFERENCES

ASHQAR, H. I., ELHENAWY, M., & RAKHA, H. A. (2019). Modeling bike counts in a bike-sharing system considering the effect of weather conditions. Case studies on transport policy, 7(2), 261-268.

FENG, C., HILLSTON, J., & REIJSBERGEN, D. (2017). Moment-based availability prediction for bike-sharing systems. Performance Evaluation, 117, 58-74.

GARCÍA-PALOMARES, J. C., GUTIÉRREZ, J., & LATORRE, M. (2012). Optimizing the location of stations in bike-sharing programs: A GIS approach. Applied Geography, 35(1–2), 235–246.

GUO Q. L., KARIMI H. A. A novel methodology for prediction of spatial-temporal activities using latent features //Computers, Environment and Urban Systems. – 2017. – T. 62. – C. 74-85.

HUANG, D., CHEN, X., LIU, Z., LYU, C., WANG, S., & CHEN, X. (2020). A static bike repositioning model in a hub-and-spoke network framework. Transportation Research Part E: Logistics and Transportation Review, 141, 102031.

JIA, W., TAN, Y., LIU, L., LI, J., ZHANG, H., & ZHAO, K. (2019). Hierarchical prediction based on two-level Gaussian mixture model clustering for bike-sharing system. Knowledge-Based Systems, 178, 84-97.

KHATRI, R. (2015). Modeling route choice of Utilitarian Bikeshare users from GPS Data.University of Tennessee. Retrieved from http://trace.tennessee.edu/cgi/viewcontent.cgi?article=4837&context=utk_gradthes

KOLDA, T. G., & BADER, B. W. (2009). Tensor decompositions and applications. SIAM Review, 51, 455–500.

LIU Y., LIU Z., JIA R. DEEPPF: A deep learning based architecture for metro passenger flow prediction //Transportation Research Part C: Emerging Technologies. – 2019. – T. 101. – C. 1

POLSON, N.G., SOKOLOV, V.O., 2017. Deep learning for short-term traffic flow prediction. Transportation Research Part C: Emerging Technologies. 79, 1–17.

R. YU , Y. LI , C. SHAHABI , U. DEMIRYUREK , Y. LIU , Deep learning: a generic approach for extreme condition traffic forecasting, in: Proceedings of the SIAM International Conference on Data Mining, 2017, pp. 777–785 .

RAVIV, T., TZUR, M., & FORMA, I. A. (2013). Static repositioning in a bike-sharing system: Models and solution approaches. EURO Journal on Transportation and Logistics, 2(3), 187–229. Retrieved from http://link.springer.com/10.1007/s13676-012-0017-6

SATHISHKUMAR, V. E., PARK, J., & CHO, Y. (2020). Using data mining techniques for bike sharing demand prediction in metropolitan city. Computer Communications, 153, 353-366.

SYSTEMS, D. T. & LACKNER, B. (2013). Modeling demand for bicycle sharing systems Neighboring stations as a source for demand and a reason for structural breaks. TRB, 1–19.

WERGIN, J., & BUEHLER, R. 2018. Where do bikeshare bikes actually Go? An analysis of capital bikeshare trips using GPS data. Transportation Research Record, (January), 2662, 12– 21.

X. MA , Z. TAO , Y. WANG , ET AL. , Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, Transportation Research Part C: Emerging Technologies. 54 (2015) 187–197 .

YANG, Y., HEPPENSTALL, A., TURNER, A., & COMBER, A. (2020). Using graph structural information about flows to enhance short-term demand prediction in bike-sharing systems. Computers, Environment and Urban Systems, 83, 101521.

ZHAO, J., WANG, J., DENG, W., 2015. Exploring Bikesharing travel time and trip chain by gender and day of the week. Transportation Research Part C: Emerging Technologies 58, 251–264.