A Hybrid Neural Network to Predict Short-term Passenger Flow at Bus Stops

Tahereh Arabghalizi
University of Pittsburgh
Pittsburgh, USA
tahereh.arabghalizi@pitt.edu

Xiaowei Jia
University of Pittsburgh
Pittsburgh, USA
xiaowei@pitt.edu

ABSTRACT

With the growth of human population and urbanization, the need for more efficient public transportation services has become crucial. Passenger flow forecasting is a key factor that can help transit operators with better planning and scheduling. While many studies have been focused on passenger flow forecasting in railways and Urban Rail Transit (URT), the prediction of passenger flow in bus transit systems remains largely unexplored. In this paper, we propose a novel hybrid model to predict short-term passenger flow at bus stops. This model extracts the spatial and temporal data patterns using a combination of Graph Convolutional model (GCN), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM). Our experimental study shows that our proposed model outperforms the baselines.

CCS CONCEPTS

• **Computing methodologies → Neural networks**.

KEYWORDS

Bus Passenger Flow, Short-term Prediction, Time Series, Neural Networks, Spatial-temporal Data

ACM Reference Format:
Tahereh Arabghalizi and Xiaowei Jia. 2021. A Hybrid Neural Network to Predict Short-term Passenger Flow at Bus Stops. In *DeepSpatial '21: 2nd ACM SIGKDD Workshop on Deep Learning for Spatiotemporal Data, August, 2021.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

There is an increasing stress on traffic systems in recent years as a result of the acceleration of urbanization and the growing human population. Intelligent passenger flow forecasting can provide timely information needed for making traffic planning decisions. In this paper, we focus on the short-term passenger flow forecasting at bus stops, which is critical for bus transit operators to optimize schedules in order to avoid crowdedness and provide passengers with better transit services.

While many studies have been conducted for passenger flow forecasting in railways and urban rail transit, the prediction of passenger flow for bus transit system remains largely unexplored. Bus transit system differs from URT in terms of variability over space and time, number of stops, number of vehicles, time schedules, and passenger flow that prevent existing techniques to be directly used in our problem.

In this work, we propose a hybrid model which extracts the spatial and temporal data patterns using a combination of Graph Convolutional model (GCN), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) for short-term passenger flow prediction at bus stops. In particular, we build separate model components for capturing spatial dependencies between consecutive bus stops from the inflow (number of people entering a bus stop) and outflow (number of people exiting a bus stop) data and we also use a graph structure to capture spatial correlations amongst multiple bus stops. An LSTM model is further applied to capture the temporal correlations of passenger flow. We evaluated the proposed method using the historical data collected from the Port Authority of Pittsburgh at a 15-min time granularity to predict the short-term passenger flow at 87 bus stops located on one of the busiest routes in Pittsburgh. Moreover, we did extensive experiments to compare the accuracy of our proposed model and a few baselines. Our results show that our model outperforms the baseline models.

2 RELATED WORK

The passenger flow prediction can be classified in two types: long-term and short-term. Long-term prediction refers to the forecasting of changes in the volume of passenger flow over a long period of time (e.g., a few months) and is usually used for public transit system planning while short-term prediction forecasts changes in the volume of passenger flow over a short time span (e.g., 15 minutes) and is mainly used in real-time scheduling [1]. Most researchers have focused on passenger flow prediction for transportation modes such as railways and urban rail transit, but they have rarely studied bus transit systems.

Most early work for short-term passenger flow prediction employed conventional mathematical methods [2, 3] such as linear models (e.g., Kalman filter) and time series analysis methods (e.g., Autoregressive Integrated Moving Average (ARIMA)) [4–6]. These models are limited in their ability to capture complex non-linear patterns from passenger flow over space (e.g., over consecutive stops) and time. In contrast, neural networks have been shown to outperform traditional empirical models because of their ability to capture spatio-temporal and topological information. Some studies have applied typical deep learning models [7, 8] or hybrid models such as SVM-LSTM [9], Back Propagation Neural Network and Markov Chain [10] and Principal Component Analysis and Back Propagation Neural Network [11] for passenger flow prediction.

DeepSpatial '21, August, 2021, Virtual

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

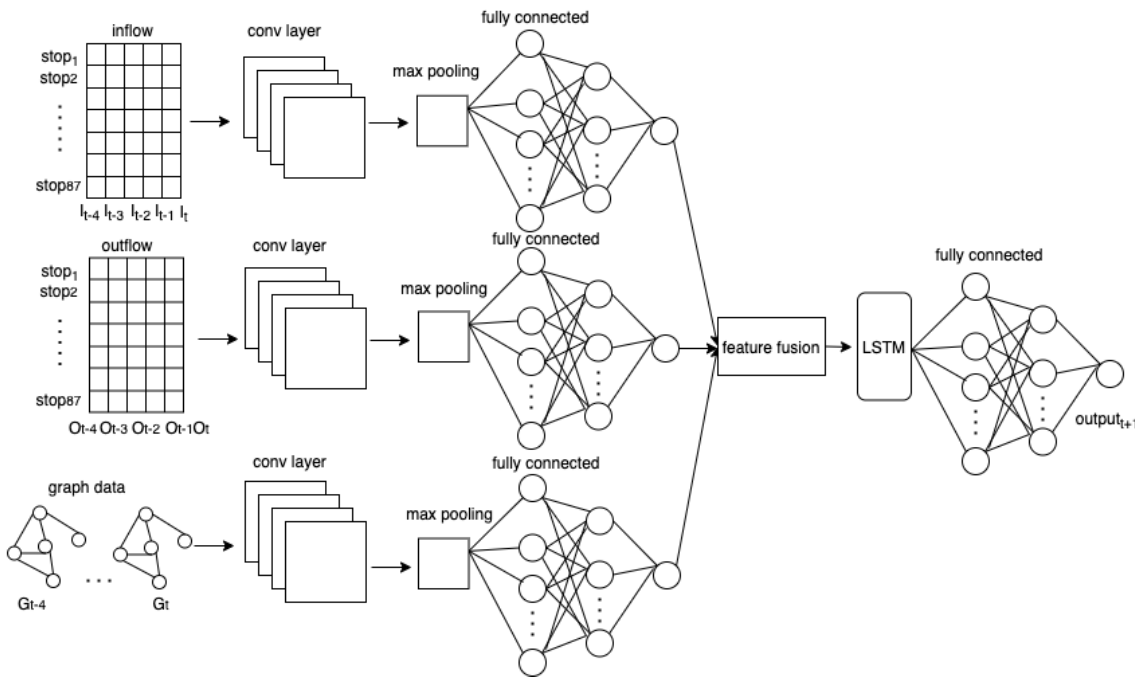

Figure 1: Model Architecture

However, most of these studies either have focused only on the railways and urban rail transit or they did not take the topological structure of the transit network into consideration.

As mentioned, a limited number of work has been done for short-term flow passenger prediction at bus stops. Wang et. al [12] proposed a hybrid model consisting of two single models of BP neural network and ARIMA time series model to predict the short-term passenger flow at bus stops. Liu et. al [13] presented a Stacked AutoEncoders-Deep Neural Network (SAE-DNN) model to predict the hourly passenger flow in bus rapid transit stations. However, topological structure of the bus transit network has been overlooked and their experiments are limited to only a few bus stops. Although, the CNN is widely used to model spatial dependence in traffic systems, it has limitations on traffic networks with complex topological structures. In recent years, the Graph Convolutional Network model [14] has been developed to deal with this problem and capture the structural feature of graph networks. To this end, we propose a novel hybrid model that combines the CNN, GCNN and LSTM models to capture both the spatial-temporal correlations and topological structure in the bus network and unlike the previous studies, we apply our model to predict passenger flow on 87 different bus stops in Pittsburgh.

The remaining of this paper is organized as follows. Section 3 presents our proposed model that predicts short-term passenger flow at bus stops. Section 4 demonstrates our experimental results conducted using a real-world dataset and section 5 summarizes this paper and future work.

3 PROPOSED MODEL

In this section, we introduce the components of our proposed model. This hybrid model detects the spatial and temporal dependencies using a combination of GCN, CNN, and LSTM models. The architecture of our proposed model is illustrated in Figure 1. We preprocess all input data including inflow, outflow and graph data to obtain the inputs for times $t - n$ (n is tuned and set to 4) to t and predict the output (passenger inflows) for time $t + 1$. The first and second process use the inflows and outflows respectively to capture the spatial features and the third process uses the graph data to extract the network topological information. Furthermore, the last process contains the LSTM model to capture the temporal correlations and predict the output data. In the following, we will provide more details about the model architecture.

3.1 Neural Network Methodologies

In this section, we shortly describe the neural network methodologies used in our proposed model and their configurations:

CNN: A convolutional layer with 32 filters and a 3×3 kernel size followed by a max-pooling layer are used to extract the spatial features of each input data.

GCNN: We applied GCNN to capture the topological dependencies in the bus network. Consider the graph as G = (V , E) where V is the set of vertices and E is the set of edges that represent the relationships between adjacent vertices, the GCNN function can be defined as follows[14]:

$$H^{l+1} = \sigma(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}H^l W^l) \tag{1}$$

where $\hat{A} = A + I$, $A \in R^{s,s}$ is the adjacent matrix where s is the number of bus stops, I is the identity matrix, $\hat{D}$ is the diagonal

node-degree matrix of $\hat{A}$, W is the weight matrix of the l^{th} layer, $H \in R^{s,t}$ is the feature matrix where s is the number of bus stops, t represents the number of historical time steps for each stop, and σ is an activation function like ReLU.
LSTM: we employed LSTM with 128 neurons to capture the temporal correlations. In other words, LSTM is used to capture the weights of different timesteps.

3.2 Inflow and Outflow Inputs

Inflow: historical inflow is the most important input for predicting the output inflow. The real-time inflow and outflow in each bus stop is obtained by the Automatic Passenger Counting (APC) system. The inflow time series is defined as a $S \times T$ matrix where S is the number of bus stops, T is the number of historical timesteps for each stop and each cell carries the total number of people getting on the busses that stop at the bus stop s at timestep t. Since the bus stops belong to each route are ordered geographically, the rows in the inflow matrix represent the adjacent bus stops. We use the time series for timesteps t - 4 to t (4 is tuned using trial and error) to predict the inflow for timestep t + 1. This data is the input to a 32-filter convolutional layer followed by a max pooling layer to reduce the computational cost by reducing the number of parameters. Then, the data is flattened and fully connected. The output data of this process is then the input into the feature-fusion section.
Outflow: the outflow process is identical to the inflow process except for the input data which is the passenger outflow instead of the inflow. We build a different time series matrix in which each cell carries the total number of people getting off the busses that stop at the bus stop s at timestep t.

3.3 Graph-based Input

As mentioned before, we use a GCNN model to capture the impact of the bus network topology. According to equation1, the graph input for GCNN is defined and built as follows:

$$Graph\ Input\ Data = \hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}(I_{s,t}) \tag{2}$$

where A is the adjacent matrix and $I_{s,t}$ is the inflow matrix. It is worth mentioning that to create the adjacent matrix we computed the distance between each two stops using their latitude and longitude and filled the matrix with the obtained values. The input data is then being processed similar to what described for the inflow and outflow process.

4 EXPERIMENTAL STUDY

4.1 Dataset

We received a one-year-long (2018-2019) historical data from the Port Authority of Pittsburgh, which consists of the data coming from the Automatic Vehicle Location (AVL) and Automatic Passenger Counting (APC) systems for around 200 routes and 7,000 bus stops in Pittsburgh. However, for the purpose of this study, we only used the data of the bus stops located on the route "61C" in March 2019. 61C is one of the busiest routes in Pittsburgh and there are 87 consecutive stops on the path of this route which are also shared by 32 other routes. The timeline used in our experiments is between 5:00 am to 12:00 am (19 hours in total) for 21 workdays in March 2019 (595,516 records). Each record contains stop ID, stop name, date, bus arrive and depart times, number of passengers boarding and alighting, latitude and longitude of each stop, etc. This data is processed to create the inflow and outflow time series according to the structure that explained in previous section. As mentioned before, the time granularity used in this study is 15 minutes (4 time slots per hour) which results in the inflow and outflow to be shaped as $87 \times (21 \times 19 \times 4)$ or $87(rows) \times 1596(columns)$ matrices. It should be noted that about 12% of the input data was missing so imputed with zero.

4.2 Model Configuration

We used data from the first 16 days of March 2019 (80%) to train and data from the last 5 days to test (20%). The validation split rate is set to 0.1 to calibrate the model. As mentioned earlier, we used the flow from the previous five timesteps to predict the next one. The convolutional layers are with 32 filters and 3×3 kernel size followed by fully connected layers with 87 (number of stops) neurons. After the feature fusion, the LSTM and final fully connected layers consist of 128 and 87 neurons respectively. We computed the training loss and validation loss for different number of epochs. both losses had a significant vibration for the first 100 epochs but after that both remained almost stable so we trained each model for 100 epochs in total.

4.3 Baselines

We compare the accuracy of our proposed model with several baseline models as described in the following:
CNN: three CNN models are used to process the three input data (with 32 filters and kernel size 3×3) and no LSTM model is applied afterwards.
LSTM: three LSTM models are used to process the three input data (with 128 neurons) and no CNN model is applied.
No-graph: the graph process is deleted from our proposed model and the input data includes inflow and outflow.
Graph-only: only the graph process is kept in our proposed model and the input data does not include the inflow and outflow.
Inflow-only: we only keep the inflow process and ignore the outflow and graph-based inputs.

4.4 Loss Function and Evaluation Metrics

The Mean Squared Error (MSE) is used as the loss function. The optimizer is "Adam" with a learning rate of 0.001. We used two more metrics to evaluate the model accuracy: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE).

4.5 Experimental Results

We conduct two types of experiments depending on the input data we fit our models with:
Comparison of the accuracy of models when using the whole day data: in this experiment, we use the input data with the same structure that explained in section 4.1 to fit the models. The figures in Table 1 indicate the accuracy of our proposed model and the baselines in terms of RMSE and MAE. As you can see, our proposed model outperforms the baseline models up to 20% and 29% in terms of RMSE and MAE respectively. Moreover, Graph-only

Table 1: Comparison of the accuracy of models when using the whole day data

Model	RMSE	MAE
CNN	4.12	1.99
LSTM	3.42	1.62
No-graph	3.73	1.67
Graph-only	3.33	1.43
Inflow-only	3.92	1.76
Proposed Model	**3.28**	**1.42**

model performs better than the other baselines which proves that the topological information has a significant impact on the passenger flow prediction. The CNN has the lowest accuracy since it only captures the spatial features.

Furthermore, Figure 2 shows the comparison of the actual and predicted passenger flow at the stop 6th Ave at Smithfild. Predicted passenger flow is computed for the test set that consists of the data for the last five days of March 2019. As it can be seen, the predicted values are in line (most of the time) with the actual values both in peak periods and off-peak periods which approves the accuracy of our proposed model. As shown, each day has several peek times and when we look closely, it seems that the model did not predict very well during the time interval between 10 am to 3 pm for the first day of the week (03/25/2019). We checked the weather condition on that day and found out that it was a rainy day so we believe one of the reasons why the proposed model did not work well for this time interval is because we did not take the external factors like weather conditions into account.

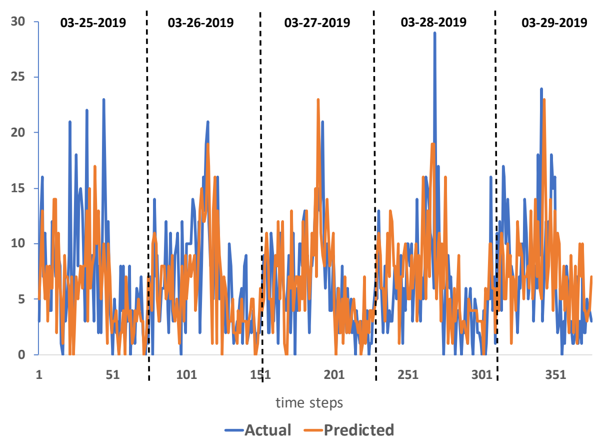

Figure 2: Comparison of actual and predicted passenger flows for the stop "6th at Smithfield" during the last five workdays of March 2019

Comparison of the accuracy of models for different time intervals: we designed this experiment to test the accuracy of the models for different time intervals. We divided the whole day into 4 time intervals including morning rush hour (5 am to 10 am), noon-afternoon (10 am to 3 pm), afternoon rush hour (3 pm to 7 pm]) and evening (7 pm to 12 am). We then built four separate inflow and outflow input matrices according to these time intervals. We trained and tested our proposed models and the baselines (with their previous configurations) using the new input data which led to building four separate models (e.g. CNN1 (morning rush), CNN2 (noon-afternoon), CNN3 (afternoon rush), CNN4 (evening)), one for each time interval. Figure 3 shows RMSE and MAE for values for these models. This figure indicates that our proposed model in all four time intervals outperforms their corresponding baselines by up to 28%. Besides, the accuracy of all models is lower during the afternoon rush hour compared to the other intervals which might be due to larger variability over that time period. In overall, we believe that there are stronger correlations between temporal features in shorter time intervals and having separate models for separate time periods of the day could provide us with more accurate predictions for those time intervals.

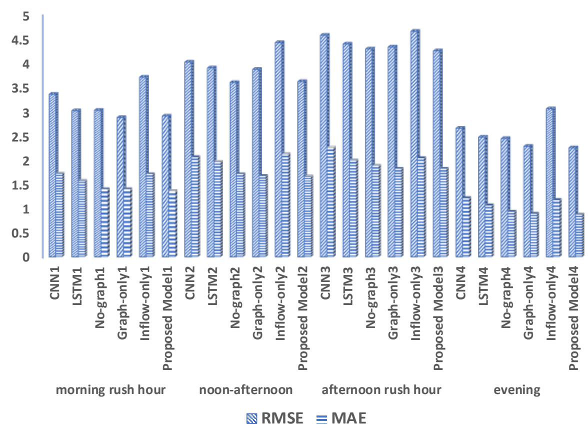

Figure 3: Comparison of the accuracy of models for different time intervals

5 CONCLUSION AND FUTURE WORK

In this study, we proposed a hybrid model which is the combination of GCNN, CNN and LSTM to predict short-term passenger flow at bus stops. We used the historical data from Port Authority of Pittsburgh as our dataset and did extensive experiments to compare the accuracy of our proposed model and several baselines. Our results showed that our model is more accurate than those baselines. As the future work, we intend to use other input features such as weather conditions that could affect the passenger flow. We also want to consider other time granularity e.g. 10 min or 30 min and compare the results with the current outcomes. We would also like to run more experiments in bigger scale by using data from bigger bus networks.

ACKNOWLEDGMENT

This work is part of the PittSmartLiving project which is supported by NSF award CNS-1739413, Pitt Momentum Award and USGS grant G21AC10207.

REFERENCES

[1] Y. Mo and Y. Su, "Neural networks based real-time transit passenger volume prediction," in *2009 2nd International Conference on Power Electronics and Intelligent Transportation System (PEITS)*, vol. 2. IEEE, 2009, pp. 303–306.

[2] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Short-term traffic forecasting: Overview of objectives and methods," *Transport reviews*, vol. 24, no. 5, pp. 533–557, 2004.

[3] H. Zhai, L. Cui, Y. Nie, X. Xu, and W. Zhang, "A comprehensive comparative analysis of the basic theory of the short term bus passenger flow prediction," *Symmetry*, vol. 10, no. 9, p. 369, 2018.

[4] J. Zhang, D. Shen, L. Tu, F. Zhang, C. Xu, Y. Wang, C. Tian, X. Li, B. Huang, and Z. Li, "A real-time passenger flow estimation and prediction method for urban bus transit systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 11, pp. 3168–3178, 2017.

[5] C.-h. Zhang, R. Song, and Y. Sun, "Kalman filter-based short-term passenger flow forecasting on bus stop," *Journal of transportation systems engineering and information technology*, vol. 11, no. 4, pp. 154–159, 2011.

[6] M. Gong, X. Fei, Z. H. Wang, and Y. J. Qiu, "Sequential framework for short-term passenger flow prediction at bus stop," *Transportation Research Record*, vol. 2417, no. 1, pp. 58–66, 2014.

[7] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.

[8] Y. Liu, Z. Liu, and R. Jia, "Deeppf: A deep learning based architecture for metro passenger flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 18–34, 2019.

[9] J. Guo, Z. Xie, Y. Qin, L. Jia, and Y. Wang, "Short-term abnormal passenger flow prediction based on the fusion of svr and lstm," *IEEE Access*, vol. 7, pp. 42 946–42 955, 2019.

[10] Z.-S. Xiao, B.-H. Mao, and T. Zhang, "Integrated predicting model for daily passenger volume of rail transit station based on neural network and markov chain," in *2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*. IEEE, 2018, pp. 578–583.

[11] S. Zhang, Z. Liu, F. Shen, S. Wang, and X. Yang, "A prediction model of buses passenger flow based on neural networks," in *Journal of Physics: Conference Series*, vol. 1656, no. 1. IOP Publishing, 2020, p. 012002.

[12] Z. Wang, C. Yang, and C. Zang, "Short-term passenger flow prediction on bus stop based on hybrid model," in *2017 2nd International Conference on Electrical, Control and Automation Engineering (ECAE 2017)*. Atlantis Press, 2017, pp. 343–347.

[13] L. Liu and R.-C. Chen, "A novel passenger flow prediction model using deep learning methods," *Transportation Research Part C: Emerging Technologies*, vol. 84, pp. 74–91, 2017.

[14] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.