



PASSENGER FLOW PREDICTION FROM AFC DATA USING STATION MEMORIZING LSTM FOR METRO RAIL SYSTEMS

T.D. Sajanraj^{*}, *J. Mulerikkal*[†], *S. Raghavendra*[‡], *R. Vinith*[§], *V. Fabera*[¶]

Abstract: Metro rail systems are increasingly becoming relevant and inevitable in the context of rising demand for sustainable transportation methods. Metros are therefore going to have a consistently expanding user-base and hence user satisfaction will require meticulous planning. Usage forecast is clearly an integral component of metro planning as it enables forward looking and efficient allocation of resources leading to greater commuter satisfaction. An observation from studying the usage of Kochi Metro Rail Ltd. is that there is a consistently occurring temporal pattern in usage for every station. But the patterns differ from station to station. This hinders the search for a global model representing all stations. We propose a way to overcome this by using station memorizing Long Short-Term Memory (LSTM) which takes in stations in encoded form as input along with usage sequence of stations. This is observed to significantly improve the performance of the model. The proposed architecture with station parameter is compared with algorithms like SVR (support vector regression) and neural network implementation with the best architecture to testify the claim. The proposed model can predict the future flow with an error rate of 0.00127 MSE (mean squared error), which is better than the other models tested.

Key words: *metro rail transport, forecast, passenger flow, LSTM*

Received: October 27, 2020

DOI: 10.14311/NNW.2021.31.009

Revised and accepted: June 30, 2021

1. Introduction

Metros are increasingly becoming a part and parcel of modern transportation systems within highly populated and geographically large cities [22, 14]. It reduces

^{*}T.D. Sajanraj – Corresponding author; Department of Computer Science Christ Deemed to be University, Bengaluru, India and Rajagiri School of Engineering & Technology (Autonomous), Rajagiri Valley, Kakkanad, Kochi, India, E-mail: sajanraj.t.d@gmail.com

[†]Jaison Mulerikkal; Department of Computer Science, Rajagiri School of Engineering & Technology (Autonomous), Rajagiri Valley, Kakkanad, Kochi, India, E-mail: jaisonmpaul@gmail.com

[‡]S. Raghavendra; Department of Computer Science, Christ Deemed to be University, Bengaluru, India, E-mail: raghav.trg@gmail.com

[§]R. Vinith; Department of Computer Science, JyothiEngineering College, Cheruthuruthy, Thrissur, Kerala, India, E-mail: vinithr@jecc.ac.in

[¶]Vit Fabera; Head of the Department, Department of Applied Informatics in Transportation, Czech Technical University in Prague, Prague, Czech Republic

the pressure of traffic on roads and also offers an environmentally more sustainable alternative. More cities are progressing in this direction. But the smooth running of the systems is conditioned on the degree to which it seamlessly integrates with other modes of services. Therefore planning of metro services turns out to be a meticulous exercise meant to ensure that the system works in accordance with the demands of the commuters and also in sync with the other transportation services running in the city [10]. Hence the process of planning takes in data from metro services itself along with those generated by feeder services [13].

Demand forecast, using input data from the above sources, is the cornerstone step in system planning and management [4]. Inference of demand forecast is done by way of developing spatio-temporal analytical models that predict the usage patterns in the future [7]. An analytical system that makes the relevant inferences are instrumental tools in bringing out the latent patterns in the data. The importance of such systems are increasingly realised by transportation managers as well as researchers, as is evidenced by the number of recent publications in the area. This work proposes a sequential deep learning architecture using LSTM that effectively extrapolates the legacy usage pattern into the future. Empirical verification of the proposed system is performed using the local metro data available to us – Kochi metro, run by Kochi Metro Rail Limited (KMRL) [11]. More information about the KMRL is available at the Fig. 1. It is a growing metro rail service with a steadily increasing passenger count that touched 2 Cr in the year 2019, two years from the



Fig. 1 Route map of Kochi metro rail.

Source: <https://kochimetro.org/the-project/>

commencement of operations. Growing usage implies a proportional growth in the data produced by the system. Data, while being of anomalous information value, poses significant system design challenges when it comes to scalability. This work proposes an analytical model on a distributed scalable platform so that progressive data ingestion and corresponding periodic model update happens seamlessly. The total passenger flow in a specific period is represented in Fig. 2. The different

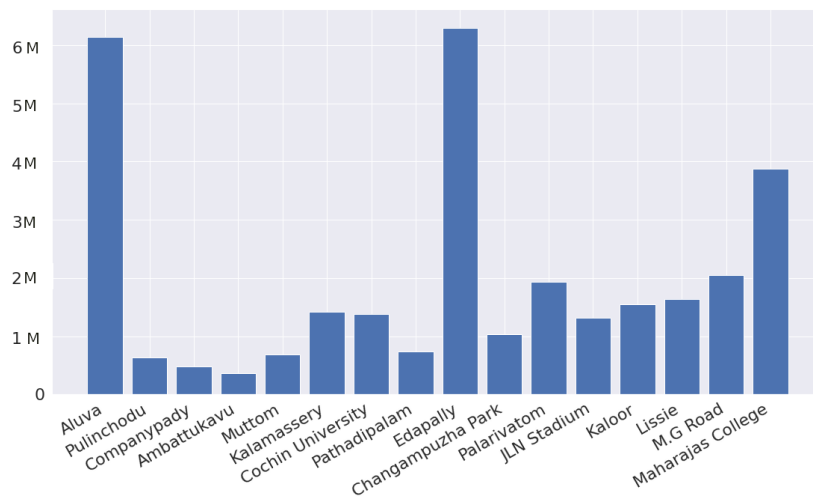


Fig. 2 Total passenger flow in station from 2017-06-19 to 2019-09-21.

metros use various technologies in their operations. The crucial component of the data generated is common. Hence the common platform that can absorb these salient aspects of the data can help us in developing global solutions applicable across many services.

The smart card data has lots of importance in holding the passenger flow information [15, 23, 21]. All those details about the passenger will be kept in a Central Automated Fare Collection (AFC) system. The study of AFC data reveals the flow patterns within a transport system. According to the transport authority, passenger flow information is very important for future planning. There are many forecasting strategies used from the beginning of transport planning. Some input and the methods are line fitting methods, moving average, simple linear regression, multiple linear regression, and support vector regression. The methods are more error-prone to the forecasting due to pattern change at each station. From the literature, the methods were observed to fail to fit the data from different stations, leading to an undesirable scenario of needing a model for every station. A more general, global model that can fit the pattern across different stations makes the analysis and prediction much easier and scalable. In order to achieve this, we have used a deep learning model, LSTM to learn the latent temporal relations in the sequence of passenger flow data.

The outcome of the paper is to study the result of the external features on passenger flow data and to improve the passenger prediction model by the addition of extra features to the current system. The overall organization of the article is

as follows. The paper describes the literature survey in Section 2. and Section 3 cover the dataset description and modelling of passenger flow system. Section 4 explain the formation of SM-LSTM followed by experimental result in Section 5.

2. Literature survey

Passenger flow pattern analysis is perhaps the most significant inference as far as metro planning is concerned. Both descriptive and predictive statistical models are relevant in terms of understanding the dynamics of the usage patterns. From earlier purely statistical models, attention has gradually drifted towards engineering solutions with the rise of machine learning. This migration was propelled by the availability of data generated by new generation transportation systems. A major proportion of the data is sourced from smart card usage. Referred to as Automated Fare Card, it is introduced as a means to enable easy transport and has a long history [18]. Utilization of the data to create facilitating services like identification of trip purpose [12], estimating the route choice [25] and evidence based planning in the transport medium show that this area is getting more and more attention of researches and solution developers.

Xiaolei Ma, Jiyu Zhang, Bowen Du, et.al [16] introduced convolutional bi-directional LSTM neural networks for network wide metro ridership prediction that leverages the power of deep learning algorithm along to predict the metro ridership. The spatial and temporal features are used in a parallel manner in two independent networks comprised of Convolutional Neural Network (CNN) and Bidirectional LSTM (Bi-LSTM) networks. The spatial features are derived from the metro ridership data after transforming it into images. CNN is used for extracting spatial features from the images and Bi-LSTM for handling time series data. The method used only 25-day data to train the network and claims that the deep architecture outperforms the other traditional methods.

Zhao, J., Qu, Q., et.al [24] proposed a method for spatio-temporal analysis of passenger travel patterns in massive smart card data. The pattern revealed by the analysis is utilized in policy making and marketing. Both spatial (origin-destination) and temporal (passenger-trips) patterns of individual are obtained from smart card data. Analysis is conducted using predefined rules for getting the regularity and anomaly from existing data. In Shenzhen metro, on an average of 2.8 million transactions happen/day. Which gives a clear idea about the growth of data within a metro system.

Z. Gan, T. Feng, et.al [6] examined the mobility pattern and travel distance in order to find the relationship between urban land usage and average travel distance (ATDs), taking Nanjing metro smart card data as a use case. Smart card data is one of the best sources of point to point travel information and travel pattern is easily decipherable from it. The analysis applied quantile regression to find out the urban land use considering different spatial scales from 0.8 km – 5 km. The findings are focused on the different time frames for a particular day, both weekdays and weekends. The analysis found that there exists some spatial correlation between different time frames. All these findings can be utilized for the metro planning. The analysis has been established only on the basis of in-flow and out-flow using quantile regression method.

Y. Qu, Y. Xiao, H. Liu [19] et al have done a study on crowd dynamics during boarding and alighting process in urban metro stations to analyze the commuter behavior pattern among users of Beijing metro rail. The investigation has targeted the videos from each metro station facing the boarding passenger to and from the train. The passenger entry and exit patterns are analyzed based on the time gap extracted from the video frame in each interval of time. Cameras kept above 3m height from the floor covers the entry to and exit from the train, of each passenger. The extracted data has been used for studying behavioral patterns, bottleneck capacity utilization, and time headway distribution. The analysis based on low-quality cameras and the future works aim for collecting high quality camera feeds for the analysis. The dwell time finding method is a rough approach and needs an integrated approach for better estimation.

The above works vindicate the usage of smart card as a mine of highly useful information whose proper utilization can take metro planning and management a lot more meaningful and effective. The study using deep recurrent neural network (RNN) presents, some descriptions about deep learning methods, auto-encoders, and Gated recurrent units (GRU) networks [17].

3. Passenger flow analysis

In passenger flow analysis, the exit pattern has a lot of significance in determining the behavior of the system. There exist many patterns with temporal relationships [20] among them. The source of passenger flow can be derived from the AFC [1] information, gathered at the metro station. In general, the collective database has a lot of common factors for mapping different relationships.

Each smart card transaction is time-stamped and has a unique transaction id for identify the passenger flow information. For the development of a generic platform, all the information needs to transformed into a common structure by the database administration. Since the inflow rate of data is very high a scalable and distributed platform for the development of analytical solutions is integral to any analytic solution. We have also developed a customized platform (JP-DAP) [9] for the purpose and all the programs are developed on it. The analytic platform (JP-DAP) is capable of handling the data at different scales though the implementation of a resource clustering application and distributed storage. The extraction of data is done through distributed computing engine with spark. Spark [3, 2] is designed to work with the advantage of the in-memory computing facility. After performing the extract-transform-load, the data is fed to the machine learning algorithms.

3.1 Ridership modelling in metro system

Exploratory analysis is used to visualize and manually finds the trends in time series pattern. Beyond this point, the inference can be done by the time series prediction. Manual inference can reveal only the most apparent trends. there are typically patterns in the data that remain latent in direct examination. They can only be revealed through computational analysis using specialized methods.

The time series pattern has four basic properties:

- Seasonality: gives the variation in time series pattern within an interval.
- Trend: change in direction of time series pattern.
- Level: average value in time-series.
- Noise: unexpected/random variation in time-series.

3.2 Data description

Data is supplied by KMRL and it is available from their starting to the current phase. Dataset consists of passenger flow information from 2017 on-wards. The data contains information about 825 days from the start of the metro station. The format of the data with the relevant fields are given in the Tab. I. Among the table, a unique transaction ID is considered for extraction the passenger count.

Database Entry	Description / Contents
Stations	All working station information (till Feb 2019) Aluva, Pulinchodu, Companypady, Ambattukavu, Muttom, Kalamassery, Cochin University, Pathadipalam, EdapallyChangampuzha Park, Palarivattom, JLN Stadium, Kaloor, Lissie, M.G Road, Maharaja’s College
Equipment Type	Mode of Taking Tickets have done (3 Modes) EFO (Excess Fare Office) TOM (Ticket Office Machine) GATE (AFC Gates)
Equipment ID	Unique ID of each Machines
Fare Product	E-Purse, RJT (Return Journey Ticket), SJT (Single Journey Ticket), Free Exit Ticket, Paid Exit Ticket, Staff Card
Fare Media	EMV (using Kochi One Card) , QR (Normal Paper Ticket), RPT (RF-ID Paper Ticket)
Ticket Card Number	Unique Ticket ID Information
Transaction Type	Top-up, Issue , Adjustment, Entry, Exit ,Cancel
Transaction Time	YYY-MM-DD HH:MM:SS Format

Tab. I Automated Fare Collection (AFC) data format.

The transaction type is responsible for identifying the directional flows. The time division is obtained from transaction time in the form of date-time format.

By comparing the changes from the initial phase to the current phase the passenger flow trends and seasonality can be identified at the early stage of analysis. The changes have a significant variation each year. As observable in Fig. 3 which shows the monthly passenger count of the two busiest station, Aluva and Edappally. A slightly downward trend can be seen in 2017, but the passenger flow is reasonably consistent in 2018 and 2019. An interesting observation is that the maximum passenger entry reported on weekends.

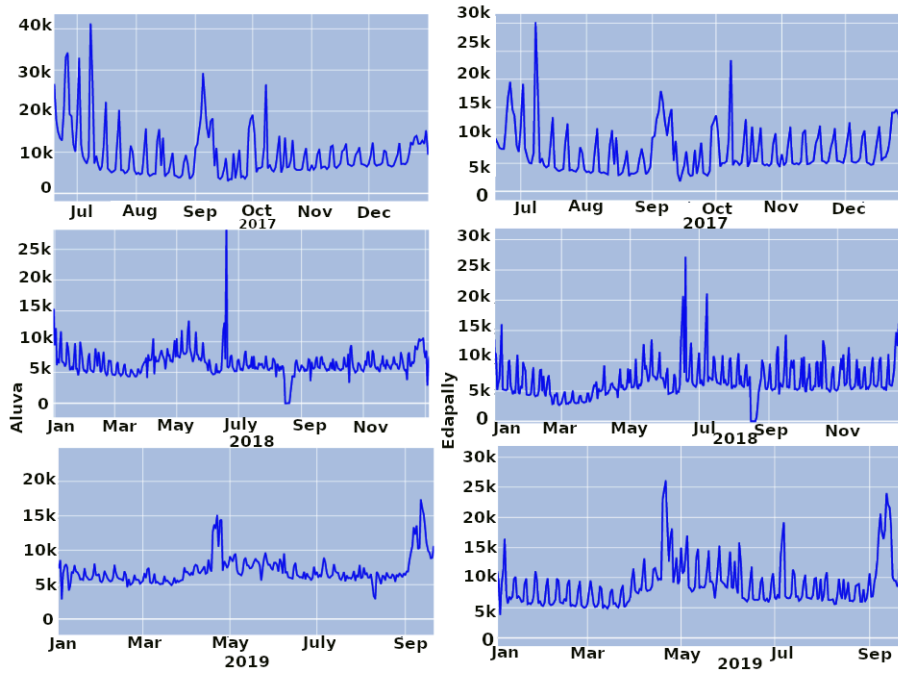


Fig. 3 Comparing 3 year data of station with maximum passenger entry.

3.3 Data pre-processing

Data needs to be pre-processed before being fed into the analytical system. Normal pre-processing such as scaling to $[0 - 1]$ range is done using min-max scalar. In the Eq. (1), the data vector is represented as X and data points are marked as X_i .

$$X_{\text{Scaled}} = \frac{X_i - \min(X)}{\max(X) - \min(X)} \quad (1)$$

There are extraneous factors which significantly impact travel patterns, such as cultural events, climatic condition etc. presently such factors are not considered and hence the anomalous patterns potentially caused them are smoothed out by replacing with the median of the time series.

3.4 Prediction of passenger flow using RNN

As can be seen in Fig. 4, the distribution changes from one station to another indicating the station is an important biasing parameter that needs to be accounted for modeling. Station memorizing LSTM proposed here takes station in encoded form as an input to predictive LSTM model.

4. Station memorizing LSTM

LSTM is a modern RNN, capable of memorizing long-range dependencies, and addressed the vital deficiency of RNNs in capturing them. The usage and efficiency of LSTM in passenger flow prediction is given in literature [8, 5]. LSTM differs from standard RNN units by the presence of special memory cell architecture.

In passenger flow prediction this memorizing layer has a lot of importance. The time sequence information along with the station encoding can be in-corporate into the input space of LSTM.

4.1 Layered architecture for sequential LSTM

The LSTM units contains:

- Forget Gate: Allows the cell state to forget memory using sigmoid function.
- Input Gate: Control input to the cell state, information using the tanh activation function.
- Output Gate: Deliver output to the next cell from the current state.

The LSTM network is designed to have a special kind of memory network as shown in Fig. 5. The cell state information is fed to the input gate and the process has 3 steps.

1. The sigmoid function regulate the values passing through the cell. x_t and h_t information were filtered in this phase.
2. The tanh function formulate all input into the range -1 to $+1$ and create an output vector.
3. The output from sigmoid and tanh is multiplied and added to form next cell state.

$$Gate : \begin{bmatrix} i_t = \sigma(W_{x_i} * x_t + W_{h_i} * h_{t-1} + b_i) \\ f_t = \sigma(W_{x_f} * x_t + W_{h_f} * h_{t-1} + b_f) \\ o_t = \sigma(W_{x_o} * x_t + W_{h_o} * h_{t-1} + b_o) \end{bmatrix} \quad (2)$$

$$Input Transform : [c' = \tanh(W_{x_c} * x_t + W_{h_c} * h_{t-1} + b_{c'})] \quad (3)$$

$$State Update : \begin{bmatrix} c_t = f_t \times c_{t-1} + i_t \times c' \\ h_t = o_t \times \tanh(c_t) \end{bmatrix} \quad (4)$$

The input gate governs what proportion of the present state goes into deciding the new state of the system. In the final state, the out O_t consists of a sigmoid

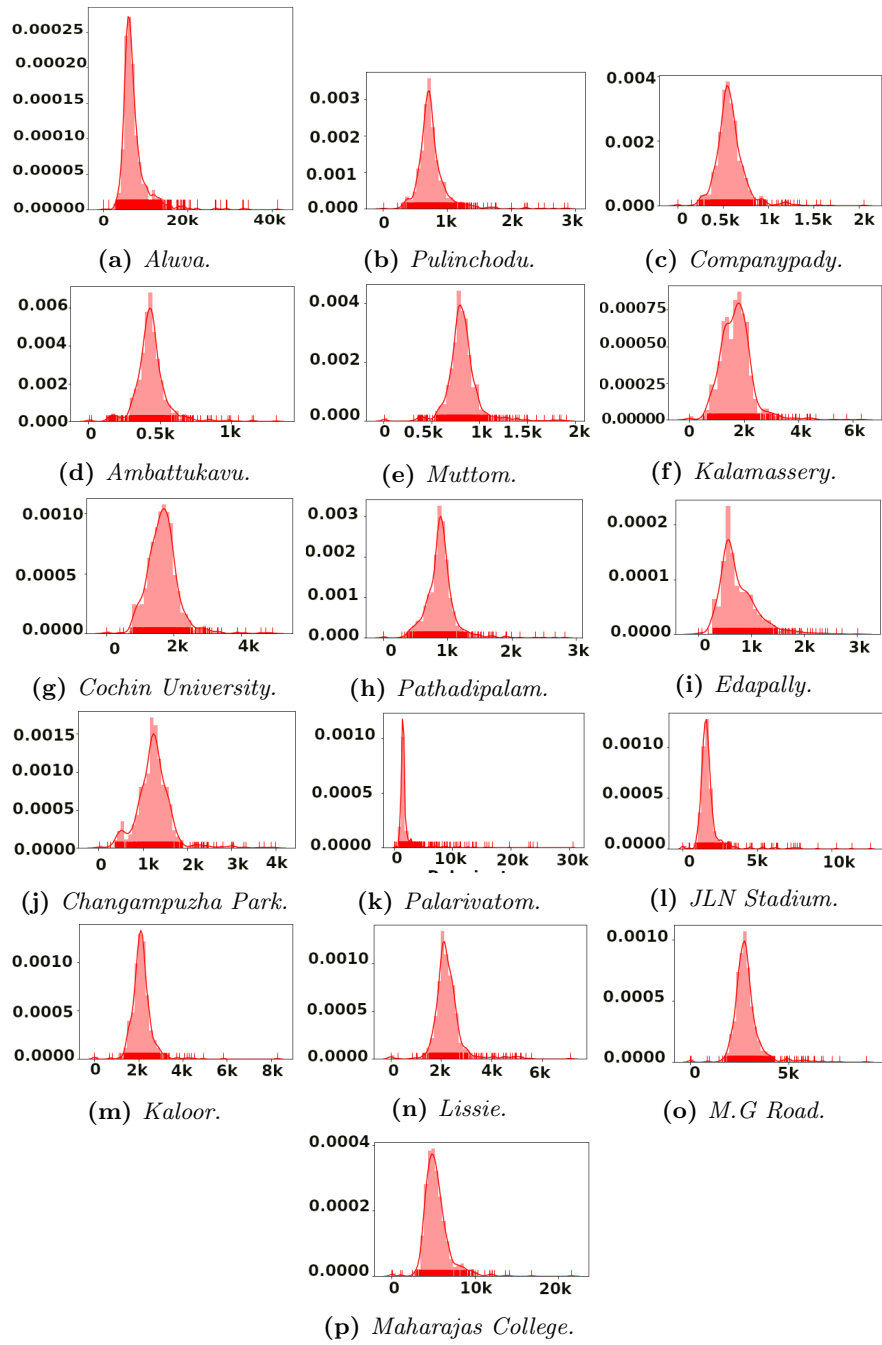


Fig. 4 Distribution of passenger flow observations from Kochi metro rail in Phase 1 with individual stations (passenger flow at x-axis and density at y-axis).

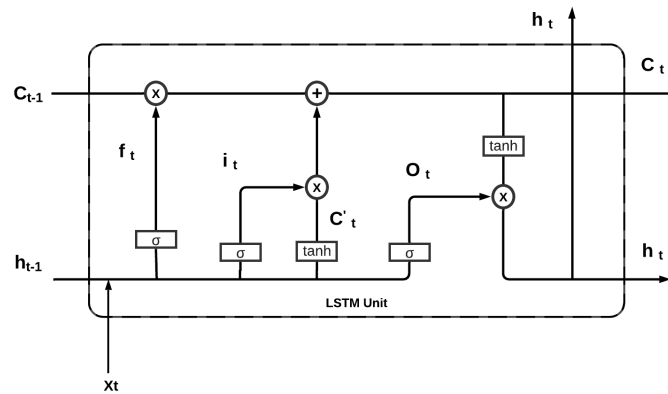


Fig. 5 LSTM unit.

filter fed by the external input and hidden state. The resultant value is multiplied with the output from \tanh to form new hidden state h_t to the next layer. The gated cell with sigmoid function is given in the Eq. (2). The input transform c' is generated using \tanh activation as in Eq. (3). The cell state update in Eq. (4) is the consolidated sum derived from input gate and forget gate. The final outcome from the LSTM unit has to output c_t and h_t .

The proposed system has two layers, one is LSTM and the other dense neural network layer. The input is fed to the LSTM sequential cells to derive patterns from the existing input and the neural computing unit works as a non-linear regressive unit to create the future forecast. The dense layer functionality is represented using Eq. (5). The intermediate output produced by the LSTM is fed to the prediction layer, which makes the future forecast. The intermediate representation produced by each 1st layer unit is a consolidated form of information flowing from the past, and hence richer than the original sequence for making the forecast. As shown in the station were distribution plot in Fig. 4, there are significant variations in the commuter behavior across stations. The global model should be able to decrease these shifts in distributions. As an improvement to the forecast, the statistical descriptors of the passenger flow patterns are given as input feed to the RNN. Therefore each station is encoded in terms of the parameters of their commuter flow distributions.

In the present work the median of the distribution is scaled to [0-1] is used as a station indicators. This enables the model to make predictions that are likely given the station parameters and renders it a generic character. The effect of these parameter is to regularize the prediction using a posterior station specific distributions.

$$O_{\text{final}} = f(W_i * X_i + B)$$

X_i : is the input matrix

W_i : weight matrix, B: bias

(5)

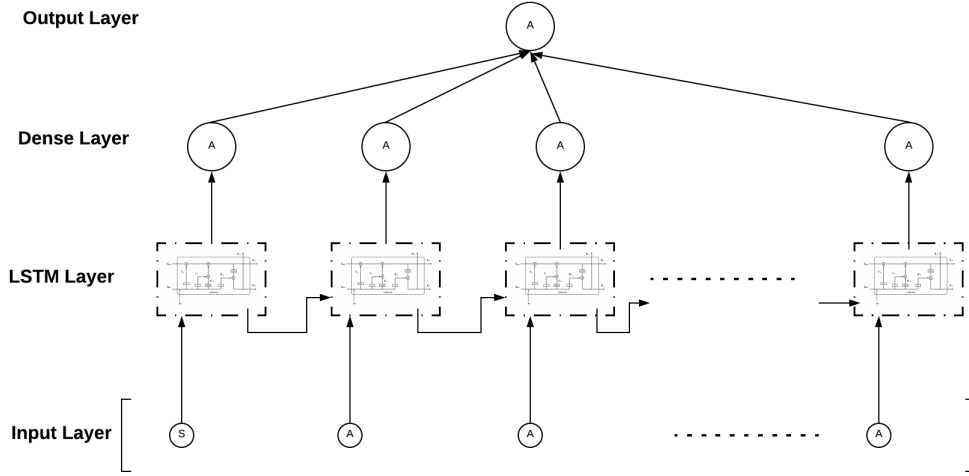


Fig. 6 LSTM layered architecture.

The relationship lies in data from historical mapping as well as the current status. All the relationship between the station and surrounding is reflected in passenger flow and the importance of station can be easily identified.

$$X = \begin{pmatrix} P_0 & D^1 & D^2 & \dots & D^n \\ P_0 & D^2 & D^3 & \dots & D^{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ P_{SN} & D^k & D^{k+1} & \dots & D^{k+n} \end{pmatrix} \quad (6)$$

The input to the system is $(n + 1)$ dimensional denoted by x and is given by $[P_i, D_1, D_2, \dots, D_n]$. The training data is represented by X which is $N \times (n + 1)$ and P is the station parameter. The input space consists station count values for continues days along with the station encoded feature. The station encoding feature is computed from the passenger flow count which gives median (the 50th percentile) as in Fig. 7. The final feature value is obtained from min-max scaling as in Eq. (1). The output is a scalar which is the normalized passenger flow prediction. The training output is denoted by y which is of N dimension.

Mean squared error is taken as the loss function. The configuration giving minimal loss is saved for the prediction of passenger flow. The learning uses Adagrad (adaptive gradient algorithm) as an optimizer because of the adaptive property of updating the parameter without manual triggering. The output uses the Relu activation function since the output can not be negative valued. Experiments were carried out on the analysis platform Fig. 3. The results are compared with the relevant existing model to prove performance improvement.

5. Experimental results

The experiments were done with the data obtained from Kochi metro rail system. The dataset contains station information of the system along with passenger flow

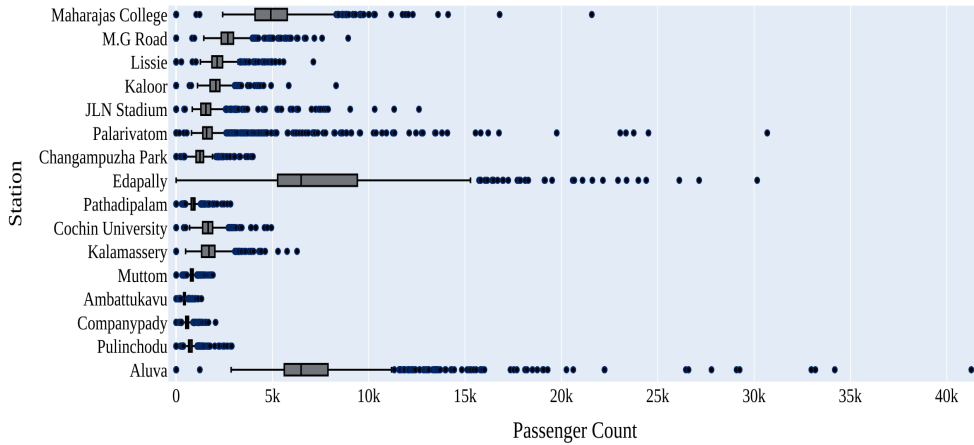


Fig. 7 Boxplot with station passenger flow.

count is as described in Tab. I. The dataset obtained from KMRL covers a time span of 824 days involving 21 stations. 16 stations were chosen for training and validation of the model as the rest 5 were operational only for 5 months. The station information is available from the metro AFC system in the raw format. The transactions available in the database is processed through big data tools in-order to extract the input for the forecast. The AFC data is processed within a 6 node cluster using hive and spark tools. The database store the information and the transformation is handled by spark dataframe which uses in-memory computing to accelerate processing.

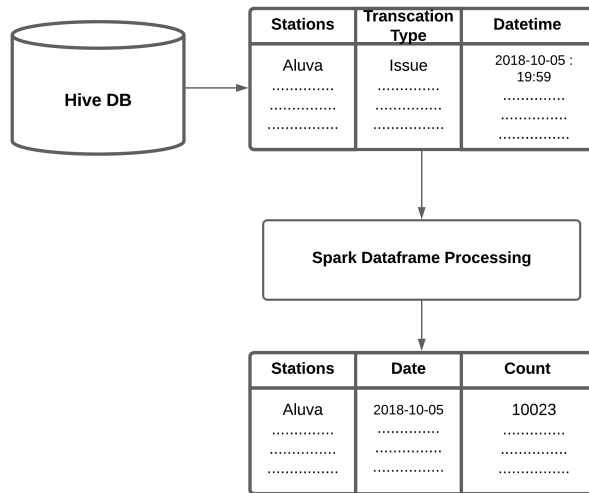
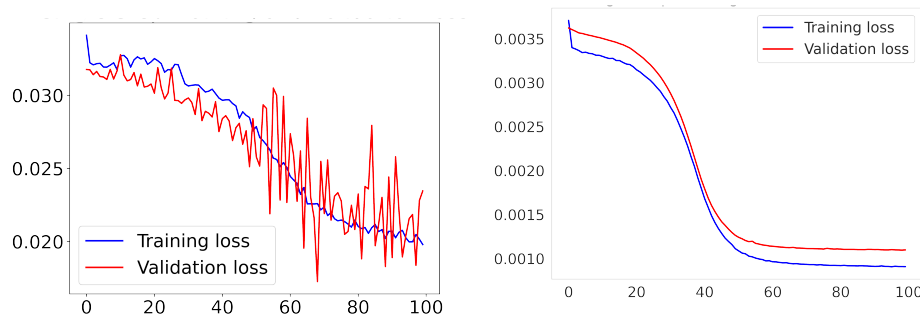


Fig. 8 Data transformation.

Feature Map	Method	MSE	RMSE
Passenger Count	SVR	0.0542	0.2320
	ANN	0.0038	0.0621
	LSTM	0.0281	0.0530
Passenger Count, Staion Encoding	SVR	0.0523	0.228
	ANN	0.0032	0.0565
	LSTM	0.00127	0.0356

Tab. II The experimental result shows the error ratio of algorithms compared, MSE denotes mean squared error and RMSE denotes root mean square error.

The test data included the passenger flow count containing weekdays and weekends. Passenger flow entry/exit at each station is recorded in AFC system while user swipe the ticket at gates. Here, the system only consider the entry passenger count at each station. The test data-set contains data of 2017 (from June), 2018 and 2019 (two month data) in form of database transaction. A sliding window of 7 days is used for generating the training and validation sets of passenger flow sequence. Thus the dataset contains a total of 16×818 instances of flow sequences. To incorporate the station information, median of the passenger flow of each station is appended to the flow sequence to obtain the final input to the model as is shown in Eq. (6). The train-test split ratio is chosen to be 0.7. The learning rate is adjusted along with the value of momentum in order to attain the minimum error rate. In the initial stage the loss plot shows damping values. After increasing both learning rate and momentum the optimal value is obtained as shown in Fig. 9. The batch size of the learning model is also dependent on this loss value. The initial phase the model takes only 100 batch for training in Fig. 9a. In Fig. 9b, the batch size is increased to 1000 which will decrease the loss to minimum level in the first epoch itself. The test result of proposed on two stations JLN Stadium and M.G Road is given in the Fig. 10.



(a) learning rate: 0.0010 and momentum: 0.020.

(b) learning rate: 0.015 and momentum: 0.030.

Fig. 9 Training and validation plot up-to 100 epochs (epochs at x-axis and loss at y-axis).

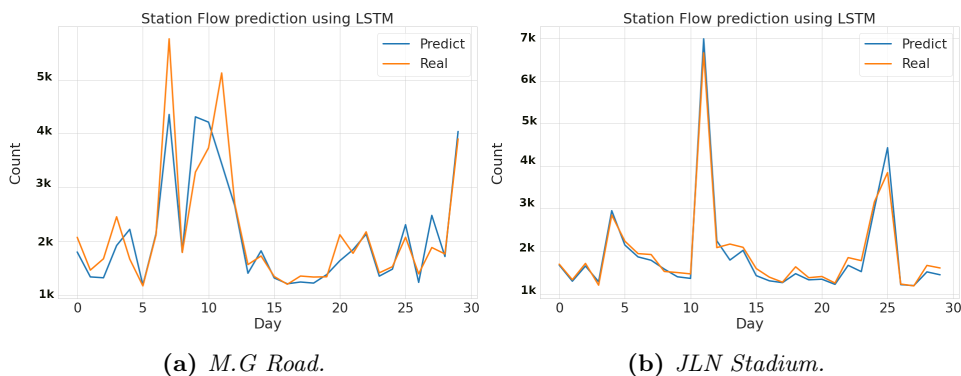


Fig. 10 Testing result of Station Memorizing LSTM.

The forecast is compared with those of SVR and various neural networks. The experimental results are shown in Tab. II. SVR with the best kernel produced an error of 0.0530. The meta parameters of ANN obtained using a cross-validation procedure and the best configuration produce an error of 0.0032.

To evaluate the influence of station information in the prediction power of the system, experiments were conducted with and without station encoding. The observations, as shown in Tab. II clearly indicates that station encoding reduces the error and hence a critical factor in making the forecast. It also shows that without station information, a universal model that performs optimally is not feasible.

The proposed model produced a mean squared error 0.00127 which is better than the other models compared. The tested result of passenger flow prediction is shown in Fig. 10. The passenger flow is represented as a count on y-axis and the x-axis shows the passenger flow in date wise manner.

6. Conclusion

Passenger flow prediction is an important problem in transport management. Development of a global model for an entire metro system is therefore very important in effective metro management. A hold-back in developing such a system is spatial variation in usage pattern distribution which is evident in the station wise distribution plot in Fig. 4. We propose a modified approach for overcoming the problem by incorporating station information in encoded form into the system. This enables the system to memorize the stations and make conditional predictions based on the station information and legacy usage patterns.

Two major observations that this work gave are:

- RNN models are capable of modelling usage patterns better than conventional regression based methods. This indirectly implies that there is significant information latent in the temporal dependencies.
- Station characteristics also contain information vital for modelling these temporal sequences. This is evident from the improvement that station encoding provides to the normal LSTM model.

This work encodes every station as the median of the commuter usage distribution. The system infers the probable distribution of a station from the median value provided and this lowers the mean squared error to 0.00127. It is intuitively expected that better statistical descriptors of station-wise usage distribution can lead to further improvement in forecast accuracy. Another major future direction to extend the work is to factor in extraneous influences like an event, presence of commercial or other establishments etc. during the modelling of usage sequence.

Acknowledgement

This research is supported by Interdisciplinary Cyber Physical Systems Division of Department of Science and Technology (DST), Government of India (Project ID: DST/ICPS/CPS Individual/2018/1091) under the Principal Investigator, Fr. Dr. Jaison Paul Mulerikkal CMI, Vice Principal & Professor, Computer Science, Rajagiri School of Engineering & Technology (Autonomous), Rajagiri Valley, Kakkanad, Kochi, India. The authors also wish to thank Kochi Metro Rail Limited for sharing their data with us for this project under a mutually agreed MoU.

References

- [1] AMPELAS A. Automatic fare collection. In: *IEEE Conference on Intelligent Transportation Systems, Proceedings C*, 2001, doi: [10.1109/itsc.2001.948827](https://doi.org/10.1109/itsc.2001.948827).
- [2] APACHE SOFTWARE FOUNDATION. *Apache Spark*. 2021. Available also from: <https://spark.apache.org/>.
- [3] ARMBRUST M., XIN R.S., LIAN C., HUAI Y., LIU D., BRADLEY J.K., MENG X., KAFTAN T., FRANKLIN M.J., GHODSI A., ZAHARIA M. Spark SQL: Relational data processing in spark. In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 2015, doi: [10.1145/2723372.2742797](https://doi.org/10.1145/2723372.2742797). ISBN 9781450327589.
- [4] ATA A., KHAN M., ABBAS S., AHMAD G., FATIMA A. Modelling smart road traffic congestion control system using machine learning techniques. *Neural Network World*. 2020, pp. 99–110, doi: [10.14311/NNW.2019.29.008](https://doi.org/10.14311/NNW.2019.29.008).
- [5] DU B., PENG H., WANG S., BHUIYAN M.Z.A., WANG L., GONG Q., LIU L., LI J. Deep irregular convolutional residual LSTM for urban traffic passenger flows prediction. *IEEE Transactions on Intelligent Transportation Systems*. 2019, 21(3), pp. 972–985, doi: [10.1109/TITS.2019.2900481](https://doi.org/10.1109/TITS.2019.2900481).
- [6] GAN Z., FENG T., WU Y., YANG M., TIMMERMANS H. Station-based average travel distance and its relationship with urban form and land use: An analysis of smart card data in Nanjing City, China. *Transport Policy*. 2019, doi: [10.1016/j.tranpol.2019.05.003](https://doi.org/10.1016/j.tranpol.2019.05.003). ISSN 1879310X.
- [7] GAN Z., YANG M., FENG T., TIMMERMANS H. Understanding urban mobility patterns from a spatiotemporal perspective: daily ridership profiles of metro stations. *Transportation*. 2020, 47(1), pp. 315–336, doi: [10.1007/s11116-018-9885-4](https://doi.org/10.1007/s11116-018-9885-4).
- [8] HAN Y., WANG S., REN Y., WANG C., GAO P., CHEN G. Predicting station-level short-term passenger flow in a citywide metro network using spatiotemporal graph convolutional neural networks. *International Journal of Geo-Information*. 2019, 8(6), pp. 243, doi: [10.3390/ijgi8060243](https://doi.org/10.3390/ijgi8060243).

- [9] J. M., S. T., V. R.D.M.D.K., B. A. Jp-dap: An intelligent data analytics platform for metro rail transport systems. *IEEE Transactions on Intelligent Transportation Systems*. 2021, pp. 1–11, doi: [10.1109/TITS.2021.3091542](https://doi.org/10.1109/TITS.2021.3091542). ISSN 15249050.
- [10] KAMENICKÝ D., BRANDEJSKÝ T., FÁBERA V., GOPAK V. RAILWAY TRAFFIC OPTIMIZATION. *Acta Polytechnica CTU Proceedings*. 2017, 11, pp. 31–34, doi: [10.14311/APP.2017.11.0031](https://doi.org/10.14311/APP.2017.11.0031).
- [11] KMRL. *Kochi Metro Rail Ltd. Kerala, India* [online]. 2021 [viewed 2019-11-11]. Available from: <https://kochimetro.org>.
- [12] LEE S.G., HICKMAN M. Trip purpose inference using automated fare collection data. *Public Transport*. 2014, 6(1-2), pp. 1–20, doi: [10.1007/s12469-013-0077-5](https://doi.org/10.1007/s12469-013-0077-5). ISSN 16137159.
- [13] LEFFLER D., BURGHOUT W., CATS O., JENELIUS E. Distribution of passenger costs in fixed versus flexible station-based feeder services. *Transportation Research Procedia*. 2020, 47, pp. 179–186, doi: [10.1016/j.trpro.2020.03.077](https://doi.org/10.1016/j.trpro.2020.03.077).
- [14] LIU Y., CHENG T. Understanding public transit patterns with open geodemographics to facilitate public transport planning. *Transportmetrica A: Transport Science*. 2020, 16(1), pp. 76–103, doi: [10.1080/23249935.2018.1493549](https://doi.org/10.1080/23249935.2018.1493549). ISSN 2324-9935.
- [15] LU J., REN G., XU L. Analysis of subway station distribution capacity based on automatic fare collection data of nanjing metro. *Journal of Transportation Engineering, Part A: Systems*. 2020, 146(2), pp. 04019067, doi: [10.1061/JTEPBS.0000304](https://doi.org/10.1061/JTEPBS.0000304).
- [16] MA X., ZHANG J., DU B., DING C., SUN L. Parallel Architecture of Convolutional Bi-Directional LSTM Neural Networks for Network-Wide Metro Ridership Prediction. *IEEE Transactions on Intelligent Transportation Systems*. 2019, doi: [10.1109/TITS.2018.2867042](https://doi.org/10.1109/TITS.2018.2867042). ISSN 15249050.
- [17] MEHRALIAN S Teshnehlab M N.B. Traffic data analysis using deep Elman and gated recurrent auto-encoder. *Neural Network World*. 2020, pp. 347–363, doi: [10.14311/NNW.2020.30.023](https://doi.org/10.14311/NNW.2020.30.023).
- [18] PARK J.Y., KIM D.J., LIM Y. Use of smart card data to define public transit use in Seoul, South Korea. *Transportation Research Record*. 2008, doi: [10.3141/2063-01](https://doi.org/10.3141/2063-01). ISSN 03611981.
- [19] QU Y., XIAO Y., LIU H., YIN H., WU J., QU Q., LI D., TANG T. Analyzing crowd dynamic characteristics of boarding and alighting process in urban metro stations. *Physica A: Statistical Mechanics and its Applications*. 2019, doi: [10.1016/j.physa.2019.121075](https://doi.org/10.1016/j.physa.2019.121075). ISSN 03784371.
- [20] TAO S., ROHDE D., CORCORAN J. Examining the spatial-temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap. *Journal of Transport Geography*. 2014, doi: [10.1016/j.jtrangeo.2014.08.006](https://doi.org/10.1016/j.jtrangeo.2014.08.006). ISSN 09666923.
- [21] TIAN X., ZHENG B., WANG Y., HUANG H.-T., HUNG C.-C. *TRIPDECODER: Study Travel Time Attributes and Route Preferences of Metro Systems from Smart Card Data*. 2020. Available also from: <https://arxiv.org/abs/2005.01492>.
- [22] VAN RYSWYK K., ANASTASOPOLOS A.T., EVANS G., SUN L., SABALIAUSKAS K., KULKA R., WALLACE L., WEICHENTHAL S. Metro Commuter Exposures to Particulate Air Pollution and PM 2.5 -Associated Elements in Three Canadian Cities: The Urban Transportation Exposure Study. *Environmental Science & Technology*. 2017, 51(10), pp. 5713–5720, doi: [10.1021/acs.est.6b05775](https://doi.org/10.1021/acs.est.6b05775). ISSN 0013-936X.

- [23] VERMA T., SIRENKO M., KORNECKI I., CUNNINGHAM S., ARAÚJO N.A. *Extracting Spatiotemporal Demand for Public Transit from Mobility Data*. 2020. Available also from: <https://arxiv.org/abs/2006.03351>.
- [24] ZHAO J., QU Q., ZHANG F., XU C., LIU S. Spatio-Temporal Analysis of Passenger Travel Patterns in Massive Smart Card Data. *IEEE Transactions on Intelligent Transportation Systems*. 2017, doi: [10.1109/TITS.2017.2679179](https://doi.org/10.1109/TITS.2017.2679179). ISSN 15249050.
- [25] ZHAO J., ZHANG F., TU L., XU C., SHEN D., TIAN C., LI X.Y., LI Z. Estimation of Passenger Route Choice Pattern Using Smart Card Data for Complex Metro Systems. *IEEE Transactions on Intelligent Transportation Systems*. 2017, doi: [10.1109/TITS.2016.2587864](https://doi.org/10.1109/TITS.2016.2587864). ISSN 15249050.