International Journal of Engineering in Computer Science

E-ISSN: 2663-3590
P-ISSN: 2663-3582
www.computersciencejournals.com/ijecs
IJECS 2024; 6(2): 96-100
Received: 07-05-2024
Accepted: 19-06-2024

Dr. T Srikanth
Associate Professor, Department of CSE, Malla Reddy Engineering College for Women, Autonomous, Hyderabad, Telangana, India

E Kavya
Student, Department of CSE, Malla Reddy Engineering College for Women, Autonomous, Hyderabad, Telangana, India

E Madhavi
Student, Department of CSE, Malla Reddy Engineering College for Women, Autonomous, Hyderabad, Telangana, India

K Shailaja
Student, Department of CSE, Malla Reddy Engineering College for Women, Autonomous, Hyderabad, Telangana, India

Corresponding Author:
Dr. T Srikanth
Associate Professor, Department of CSE, Malla Reddy Engineering College for Women, Autonomous, Hyderabad, Telangana, India

Metro systems: Predicting origin-destination passenger flow using deep learning algorithms

Dr. T Srikanth, E Kavya, E Madhavi and K Shailaja

DOI: https://doi.org/10.33545/26633582.2024.v6.i2b.129

Abstract
The efficiency and quality of metro services may be improved with accurate predictions of Origin-Destination (OD) passenger flow. Projecting OD in metro networks has received less attention than projecting incoming as well as outgoing flows for particular stations, according to existing research. The problems with OD flows stem from three main sources: 1) their complicated geographical correlations and high temporal dynamics; 2) their susceptibility to external influences; and 3) the fact that their data slices are sparse and incomplete. In this study, we present an AFFN that can a) learn the effects of external factors automatically and b) precisely represent the periodic variations in passenger flows by fusing spatial dependencies from various based on knowledge graphs as well as hidden correlations between stations. Improving the accuracy of OD predictions is our secondary goal in extending AFFN to multi-task AFFN, which allows us to handle sparse and incomplete OD matrices by also predicting the input and output of each station. Two authentic Chinese metro trip datasets, gathered in Nanjing and Xi'an, were the subjects of our comprehensive experiments. When compared to state-of-the-art baseline approaches and AFFN variations, our AFFN and multitasking AFFN perform better on several accuracy criteria. This proves that AFFN and its components are useful in OD prediction.

Keywords: Metro systems, origin-destination (OD) flow, deep learning algorithms

Introduction
When it comes to public transit in large cities, METRO is among the best and most often used options. The majority of cities' commuters opted for the metro as their primary mode of transportation. In the Big Apple, Hong Kong, and Tokyo, the percentage Revised on November 6, 2022, and approved on January 12, 2023, after manuscript received on March 21, 2022. Revision date: 8 May 2023; publishing date: 26 January 2023. This work was partially funded by the following sources: the National Key Development and Research Program of China (Grant 2019YFB2102200), the Natural Sciences and Engineering Foundation of China (Grants 62232004, 62102082, 61632008, 61902062, 61672154, 61972086, 61932007, 61806053, and 61807008), the Jiangsu Organic Science Foundation of China (Grants BK20210203, BK20180356, and BK20180369), the Graduate Training and Innovative technology Program of Jiangsu Province, the Republic of China (Grant KYCX19_0089), and the Key Laboratory about Computer Network and Data Integration (Southeastern University, Ministry of Education). This article was edited by Q. Zhang. (The contributions of Yohan Xu & Yan Lye were equal.) (Wu Weiwei is the listed author.) The authors of the paper are from the School for Computing Sciences and Engineering at Southeast University in Nanjing, China. You can reach them at the following email addresses: hangul@ seu.edu.cn, lvyanly@seu.edu.cn, guangweixiong@seu.edu.cn, showings.edu.cn, weiweiwu@seu.edu.cn, and jluo@seu.edu.cn. Email: chl@nwpu.edu.cn Hele Cui is associated with the School of Computer Science at Northwestern Polytechnical University in Xi'an, Shaanxi, China (710129). Even higher, between 80% and 90%, are metro passengers with the Digital Object Identifier 10.1109/TITS.2023.3239101. Rapid urbanization and population growth are putting a strain on metro systems, which must optimize service operations like elastic timetable scheduling and flexible skip-stop line planning in order to meet the high demand for dynamic travel. This can only be achieved with the help of accurate predictions of origin-destination (OD) passenger flows. Predicting the inflow and outflow of specific metro stations for the purposes of metro administration

and emergency response has been the primary emphasis of most previous research. The amount of metro journeys between each pair of origin and destination stations has only been predicted by a small number of studies. The total number of taxi journeys from each origin location to the goal region is a well-studied example of OD prediction in the context of ride-hailing services. However, these methods won't work on the metro since there aren't many connections connecting the stations, unlike dense arterial roads where distances may be reasonably estimated using Euclidean distances. This is why we're interested in learning more about sparse metro network OD flow prediction. The following facts make OD prediction regarding a citywide subway system problematic. There are intricate spatial linkages and high temporal dynamics. OD traffic in subway systems is quite variable, particularly during rush hours. It just takes a little period for the amount of OD excursions to fluctuate substantially. Two points may have comparable temporal OD flows in the spatial dimension because to their close proximity, common urban functions in the surrounding area, or other unobservable shared qualities. Accurately and concurrently capturing these intricate spatial and temporal connections is crucial. 2) Recurring patterns and outside influences. It is clear that OD flow follows regular patterns that repeat every few days or weeks. At the same time, it is susceptible to outside influences that could disrupt regularity, such as holidays and weather. The current body of research treats external variables and periodic patterns as separate entities, but it does not account for how the former influences the latterThirdly, OD matrices that are sparse and partial. perhaps, a metro ride will take a considerable amount of time, perhaps 30 minutes or more. Due to the fact that we do not know the passengers' final destinations until they tap out at the final station, the Synchronous OD matrix does not include incomplete trips. In addition, sparsity is a common feature of OD matrices. While the majority of OD pairings have few travels between them, just a small number of origin-destination point pairs cover the vast majority of OD excursions. It is challenging to make an accurate forecast with such a sparse and partial data. Adaptive Feature Fusion Networks (AFFNs) were suggested as a solution to these problems; they would 1) fuse spatial relationships between stations using various pieces of information and even hidden correlations, and 2) fuse periodic patterns using the auto-learned influence of external variables. For the purpose of encoding spatial relationships between stations, we provide Enhanced Multi-Graph The process of convolution GRU (EMGC-GRU), which makes use of an attention-based network for hidden connections in addition to several knowledge-based graphs. Each GRU layer includes graph convolutions to capture dynamics over time. After that, EMGC-GRU using a gating unit incorporates periodic OD flow into real-time prediction by learning attention weights from external sources. We augment AFFN to perform multiple tasks AFFN in order to anticipate the incoming and outgoing data from every sensor as a separate sub-task, therefore addressing the sparsity and incompleteness of OD matrices. Because IO matrices can be more comprehensive, dense, and closely connected with OD prediction, IO prediction is much simpler than OD prediction. As a result, OD prediction accuracy is enhanced by sharing the IO forecast network. Basically, what we've done is:

- An EMGC-GRU is a kind of Extended Multi-Graph The process of convolution gated recurrent unit that can learn hidden attention-based correlations among stations in GRUs as well as capture spatial correlations that are preset in various knowledgebase graphs.
- A suggested attention module that is based on external variables aims to enhance prediction accuracy by integrating periodic stream of information with emphasis weights gained from external factors.
- Improve the accuracy of OD predictions even further using an asymmetric multitasking Adaptive Feature Integration Network (AFFN) that uses a common IO encoder and an external factor-based attention to forecast both IO and OD flows.
- AFFN and its core components are successful in forecasting OD flow, as shown by analyses on three real-world data sets where our AFFN and multitasking AFFN beat state-of-the-art baseline approaches and AFFN variations in terms of prediction errors. Here is how the remainder of the article is structured: Prior research is covered in Section II, while essential ideas are presented and prediction issues are defined in Section III. The adapting feature fusion network is introduced in Section IV, and its application to multi-task AFFN is discussed in Section V. Section VI presents the evaluation, and Section serves as the paper's conclusion.

Related work

"The current state of urban rail transit in China and its plans for the future: An analysis of national policies and strategies from 2016 to 2020"

Global interest has been piqued by China's massive and quick urban rail transportation expansion in the last several years. Future planning, policymaking, financial investment, and service enhancement pertaining to regional urban transportation might benefit from the findings and analysis presented in the progress report. This report provides an overview of infrastructure statistics. This paper thoroughly examined the urban rail transit system in Mainland China, looking at its spatial service protection, passenger service performance, and operational efficiency. It also analyzed its development features, including creating scales as well as multi-type downtown rail transit modes, using data from the China Urban Mass Transit Annual Report 2008-2015. We provide some patterns and suggestions on how China's urban rail transport system may grow in the future.

Optimizing urban rail transport schedules with two objectives in mind: time-dependent passenger volume

Concerns about the environment and society are on the rise, making energy conservation a difficult issue for urban rail transportation systems. Customer demand for each station exhibits time-variant features, which are often disregarded in the current work on this issue. In this study, we build a bi-objective schedule optimization model to reduce energy usage and overall passenger waiting time using real-world data on schedule-dependent smart-card automated fare collection. The model says that in an oversaturated state, the entire waiting time for passengers is equal to the train's capacity, and that the pure energy used equals the sum of the propulsion energy consumption minus the regeneration energy during a certain time period. Using actual data from the Beijing Yichang metro line, numerical demonstrations are carried out. When compared to the existing schedule, the findings show that the created model can effectively enhance passenger service while reducing energy usage.

"An energy-efficient model for rescheduling metro train schedules that takes ATO profiles with dynamic passenger flow into account"

Because of the high density and frequency of metro traffic, when an unanticipated disruption happens under overcrowded circumstances, the running of trains could be disrupted. Because of delays in service and trains' limited capacity, many people may end themselves waiting on platforms. In this paper, we present a mixed integer program (MIP) framework for a metro train schedule rescheduling problem. The goal is to optimize all three metrics simultaneously: total train delay, number of stuck passengers, and energy consumption. To achieve this, we use integer variables as choosing signs for ATO profiles that were already present in on-board the Australian Taxation Office systems provided by metro signal suppliers. By adding together the tractive use of energy and the regenerated energy, we get the total energy consumption, which takes into account the mass of everyone in the vehicle. After that, we solve the suggested model using the commercial optimization tool CPLEX, which quickly finds solutions with trade-offs. To confirm the efficacy of the suggested approach, three numerical experiments are conducted using operational data from the actual world.

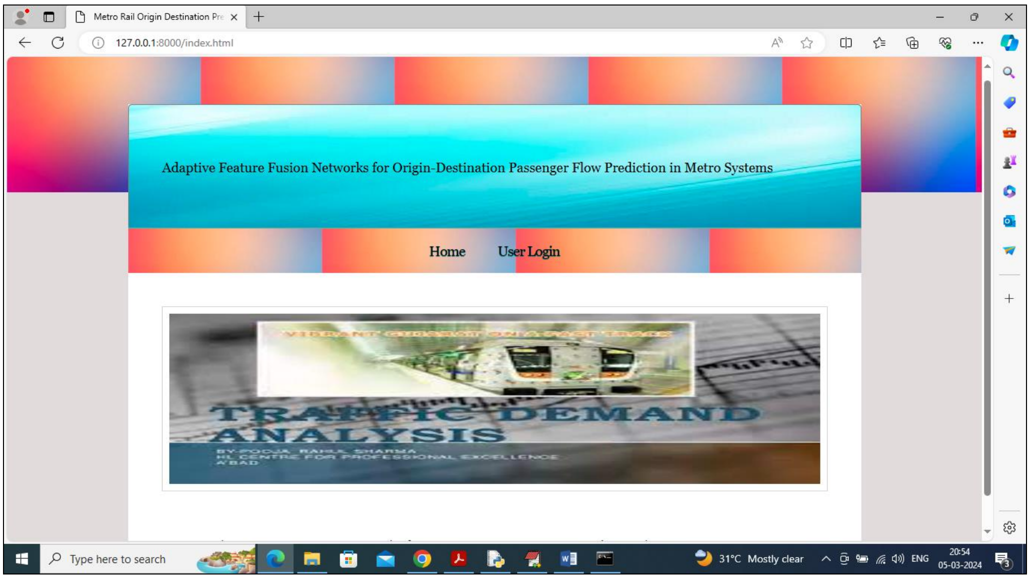
"Cleaning up metro line operations with smart card information"

A low-cost method to improve the effectiveness of metro operations and the travel experience for passengers is skip-stop operation. Optimal operation of the skip-stop strategy for bidirectional subway lines with the goal of minimizing average passenger journey time is proposed in this work. The suggested Flexible Skip-Stop plan (FSSS) is an alternative to the traditional "A/B" plan that is better able to handle passenger demand that varies both in space and time. The next step is to create a method that uses genetic algorithms (GAs) to find the best answer as quickly as possible. We use smart card data to derive time-dependent passenger demand and undertake a case study based on a real-world bidirectional subway line in Shenzhen, China. Findings show that efficient skip-stop operation may shorten average passenger journey times, and this scheme has the potential to save energy and operating costs for transit agencies. When a certain amount of people miss their train (because of a skip operation), it's important to figure out what happens next. Even when the majority of riders from the missed OD pairs are bewildered and unable to board the correct train, the results reveal that FSSS consistently surpasses the all-stop system.

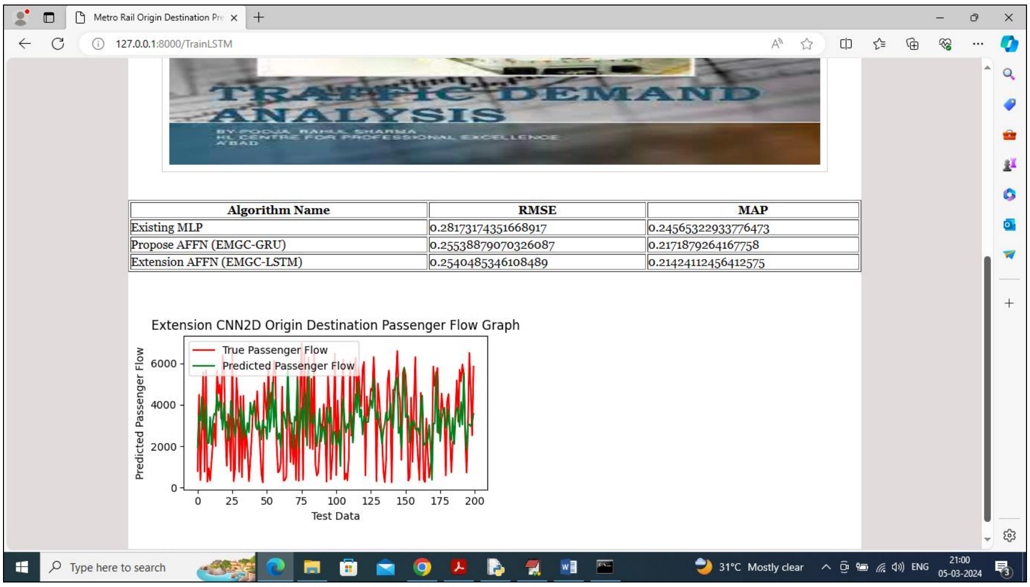
Methodology

1) Sign in as user: A user with the administrative login credentials may access the system.
2) Retrieve and Analyze Dataset: The user may normalize and divide the dataset into a train and test set using this module after uploading it to the program.
3) Current MLP set: Our current technique for passenger flow prediction may be executed with the help of this module. Then, using both the actual and anticipated numbers, we can compute RMSE and MAP.
4) AFFN (EMGC-GRU): This module allows us to anticipate passenger flow using a convolution graph driven GRU algorithm. We can then measure RMSE and MAP utilizing the predicted and actual values.
5) Supplemental AFFN (EMGC-LSTM): In order to forecast the flow of passengers, this module allows us to execute a convolution graph-based LSTM algorithm. We can then compute RMSE & MAP using the actual and anticipated values.
6) Data visualization: will show both the actual and anticipated passenger flow data, and then create a graph comparing the methods based on RMSE and MAP.

Results and Discussion

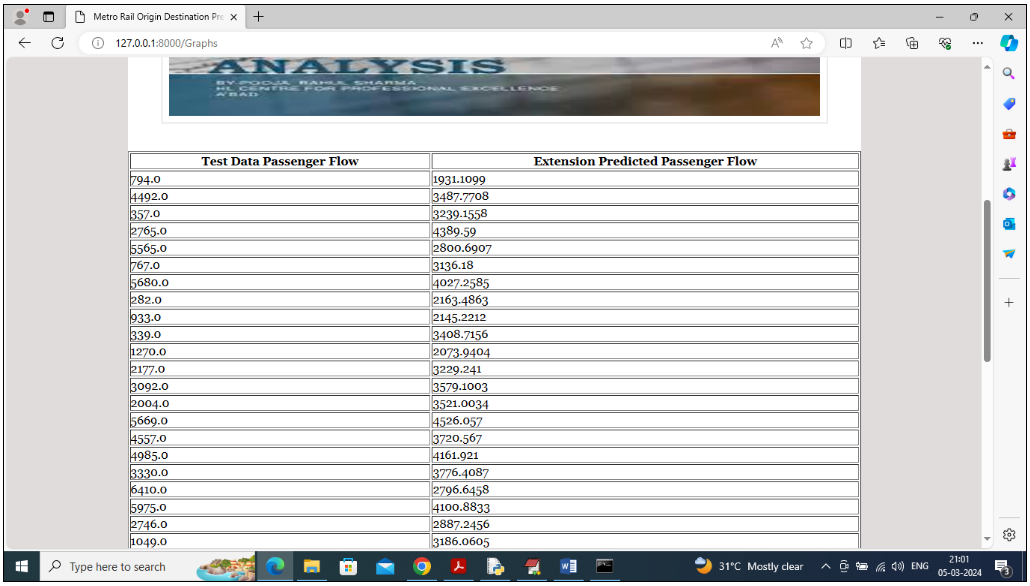


Select "User Login" from the results up top.



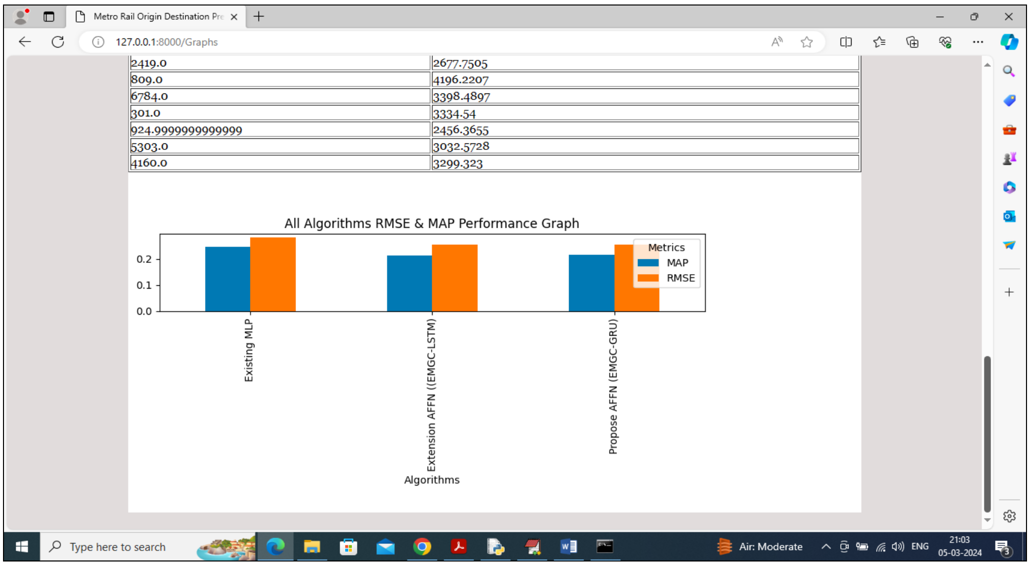
Algorithm Name	RMSE	MAP
Existing MLP	0.28173174351668917	0.24565322933776473
Propose AFFN (EMGC-GRU)	0.25538879070326087	0.2171879264167758
Extension AFFN (EMGC-LSTM)	0.2540485346108489	0.21424112456412575

Compared to the other two algorithms, AFNN with LSTM produced lower MAP and RMSE scores and a narrower separation between the green and red lines in the results shown above.



Test Data Passenger Flow	Extension Predicted Passenger Flow
794.0	1931.1099
4492.0	3487.7708
357.0	3239.1558
2765.0	4389.59
5565.0	2800.6907
767.0	3136.18
5680.0	4027.2585
282.0	2163.4863
933.0	2145.2212
339.0	3408.7156
1270.0	2073.9404
2177.0	3229.241
3092.0	3579.1003
2004.0	3521.0034
5669.0	4526.057
4557.0	3720.567
4985.0	4161.921
3330.0	3776.4087
6410.0	2796.6458
5975.0	4100.8833
2746.0	2887.2456
1049.0	3186.0605

You can see that some of your predictions are rather near by comparing the two columns of the results: true flow of passengers and extension passenger flow. To see the graph below, scroll down to the top of the page.



2419.0	2677.7505
809.0	4196.2207
6784.0	3398.4897
301.0	3334.54
924.9999999999999	2456.3655
5303.0	3032.5728
4160.0	3299.323

The x-axis in the following graph shows the names of the algorithms, while the y-axis shows the MAP & RMSE values in various color bars. Since extension achieved lower MAP & RMSE values than all of the methods, it is clear that extension is the superior choice.

Conclusion

For a citywide metro system, we suggested an Adapted Feature Fusion Networks (AFFN) for forecasting the movement of passengers from their starting points to their final destinations. We started by creating an improved multi-graph convolution-gated recurrent section (EMGC-GRU) that combines the predefined correlations modeled by various knowledge-based graphs with the auto-learned attention-based undetected relationships between stations in GRUs. This allowed us to fully capture the multifaceted relationships between space and time in OD flows. Next, we incorporate the periodic data flow with external variables to create an attention module that is based on external factors. This module will correctly capture the periodic pattern. Additionally, we suggested an unbalanced multi-task framework for predicting IO and OD flows simultaneously in order to increase the accuracy of our predictions. On two actual metro trip datasets, our suggested approaches achieve better performance than current spatial-temporal prediction strategies with respect to different prediction errors. Here are some things we can work on for future projects: 1) creating a multi-step model for forecasting from the one-step one, 2) using data from video surveillance cameras as well as other sensors to predict more fine-grained flow of passengers (such as where people go and how long they wait at stations), 3) testing our model in more complicated metro systems (like those with circular lines or multiple shared tracks), and 4) incorporating non-metro trips (like those in a taxi) to improve prediction accuracy.

References

1. Bao X. Urban rail transit present situation and future development trends in China: Overall analysis based on national policies and strategic plans in 2016-2020. Urban Rail Transit. 2018;4(1):1-12.
2. Sun H, Wu J, Ma H, Yang X, Gao Z. A bi-objective timetable optimization model for urban rail transit based on the time-dependent passenger volume. IEEE Trans Intell Transp Syst. 2018;20(2):604-615.
3. Hou Z, Dong H, Gao S, Nicholson G, Chen L, Roberts C. Energy-saving metro train timetable rescheduling model considering ATO profiles and dynamic passenger flow. IEEE Trans Intell Transp Syst. 2019;20(7):2774-2785.
4. Zhang P, Sun Z, Liu X. Optimized skip-stop metro line operation using smart card data. J Adv Transp. 2017;2017:1-17.
5. Mei Y, Gu W, Cassidy M, Fan W. Planning skip-stop transit service under heterogeneous demands. Transp Res B Methodol. 2021;150:503-523.
6. Liu Y, Liu Z, Jia R. DEPP: A deep learning based architecture for metro passenger flow prediction. Transp Res C Emerg Technol. 2019;101:18-34.
7. Ye J, Zhao J, Ye K, Xu C. Multi-Signet: A graph convolution based spatial-temporal framework for subway passenger flow forecasting. In: Proceedings of the International Joint Conference on Neural Networks; Glasgow, UK; c2020. p. 1-8.
8. Lu Y, Ding H, Ji S, Sze N, He Z. Dual attentive graph neural network for metro passenger flow prediction. Neural Comput Appl. 2021;33:13417-13431.
9. Gao A, Zheng L, Wang Z, Luo X, Xin C, Luo Y. Attention based short-term metro passenger flow prediction. In: Proceedings of the International Conference on Knowledge Science, Engineering and Management; New York, NY, USA: Springer; c2021. p. 598-609.
10. Zhang J, Chen F, Cui Z, Guo Y, Zhu Y. Deep learning architecture for short-term passenger flow forecasting in urban rail transit. IEEE Trans Intell Transp Syst. 2021;22(11):7004-7014.
11. Botte M, Di Salvo C, Carports C, Montello B, Dracaena L. Defining economic and environmental feasibility thresholds in the case of rail signalling systems based on satellite technology. In: Proceedings of the IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC); Florence, Italy; c2016 Jun. p. 1-5.
12. Dracaena L, Botte M. A passenger-oriented optimization model for implementing energy-saving strategies in railway contexts. Energies. 2018;11(11):2946.
13. Dracaena L, Gallo M, Montello B, Placido A. The definition of a model framework for managing rail systems in the case of breakdowns. In: Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC); The Hague, Netherlands; c2013 Oct. p. 1059-64.
14. Dracaena L, Placido A, Botte M, Montello B. A methodological approach for managing rail disruptions with different perspectives. Int J Math Models Methods Appl Sci. 2016;10:80-86.
15. Toque F, Come E, El Maharshi MK, Okello L. Forecasting dynamic public transport origin-destination matrices with long-short term memory recurrent neural networks. In: Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC); Rio de Janeiro, Brazil; c2016 Nov. p. 1071-1076.
16. Liu L, Chen J, Wu H, Zhen J, Li G, Lin L. Physical-virtual collaboration modelling for intra- and inter-station metro ridership prediction. IEEE Trans Intell Transp Syst. 2020;23(4):3377-3391.