Efficient Traffic Demand Forecasting Using A Meaningful Representation With Social Multiplex Networks and Community Detection

Eleftheria Karakitsou[1], Panagiotis Fafoutellis[2]*, Eleni I. Vlahogianni[3]

[1] Masters Student, School of Electrical and Computer Engineering

[2] Ph.D. Candidate, School of Civil Engineering

[3] Associate Professor, School of Civil Engineering

National Technical University of Athens, Greece

SHORT SUMMARY

In this paper, a meaningful representation of the road network using Multiplex Networks, as well as a novel feature selection framework that enhance the predictability of future traffic conditions of an entire network are proposed. Using data of traffic volumes and tickets' validation from the transportation network of Athens, we were able to develop prediction models that achieve very good performance but are also trained efficiently, do not introduce high complexity and, thus, are suitable for real-time operation. More specifically, the network's nodes (loop detectors and subway/metro stations) are organized as a multilayer graph, each layer representing an hour of the day. Nodes with similar structural properties are then classified in communities and are exploited as features to predict the future demand values of nodes belonging to the same community. The results imply the potential of the method to provide reliable and valid predictions.

Keywords: Multiplex network, community detection, traffic forecasting, road network representation

1. INTRODUCTION

Traffic forecasting is a vital part of modern transportation systems and its implications include traffic management, user information and travel time estimation (Vlahogianni et al., 2014). Deep Learning structures have been an obvious direction for researchers and practitioners during the last decades, due to the great development of technology and computing systems, as well as the availability of vast amounts of relevant data. The high accuracy of the predictions of the Deep Learning methods, compared to classic statistical modeling, is mentioned as their main advantage (Wang et al., 2019).

However, various researchers disagree with the overwhelming use of Deep Learning in traffic forecasting problems and suggest focusing on models that are more interpretable and actionable, i.e. can be exploited in real-world conditions. Indeed, Deep Learning models require massive amounts of data and their training process is time-consuming and require a lot of computing resources that are not often available, limiting the usability of Deep Learning models by policy-makers and authorities (Tedjopurnomo et al., 2020).

In addition, complex Deep Learning structures, such as Recurrent and Graph Convolutional Neural Networks, are difficult to calibrate as they include a large number of hyperparameters which also contribute to the higher computational cost (Boukerche & Wang, 2020). Most importantly, the complex topology of big cities' road networks cannot be efficiently processed using the representations that are most popular nowadays (Jiang & Luo, 2021) and does not favor the deployment of Deep Learning architectures (Do et al., 2019).

In contrast, researchers argue that a more lightweight model which would highlight the existing spatiotemporal correlations of the road network is more likely to be efficient for real-time prediction applications (Tedjopurnomo et al., 2020; Yin et al., 2021). Therefore, an effective and accurate but also efficient representation of the road network is considered equally important with the selection of the appropriate modeling technique. (Lee et al., 2021) and would increase the chance of the models being used in real-time applications (Do et al., 2019).

Taking into account the above, in this paper, we adopt the concept of Multiplex Networks from the research area of Social Network analysis to represent the transportation network of the city of Athens, Greece and develop an innovative framework of feature selection based on Community Detection on multilayer graphs. Moreover, we proceed to predict the future traffic volume of the entire road network and the number of validated tickets at metro stations using a simple Feed Forward Neural Network. As far as the authors are concerned, this is the first time that Multiplex Networks are exploited to enhance a traffic forecasting application.

The rest of this paper is organized as follows: In Section 2 the basic methodological concepts are presented, while Section 3 includes the implementation details and results. In Section 4 the main conclusions are discussed.

2. METHODOLOGY

Multiplex Networks are a kind of multilayer graphs that are most often used in Social Network analysis. Each node represents an actor and each edge a relationship between the actors. Each layer represents a different type of relationship of the same nodes/actors (Bródka et al., 2018). Multiplex Networks are used to compare the interactions between the actors n different layers. A formal definition is the following:

Definition 2.1 (Multiplex network): A multiplex network is a tuple (A,L,V,E) where A is a set of actors, L is a set of layers, $V \subseteq A \times L$ and $E \subseteq V \times V$ where $\forall (\alpha_1, l_1, \alpha_2, l_2) \in E : l_1 = l_2$. (Magnani, Rossi, et al., 2021)

Thus, we are referring to a set of graphs that are at various layers. The presented definition allows some of the nodes not to be present in some layers. In some cases, when the term multiplex is used, it is assumed that all nodes are present in all layers. To avoid confusion, in this paper we explicitly talk about this kind of graph which is defined as a node-aligned multiplex network (Bródka et al., 2018).

A common network mining task is the identification of communities, which is a subgroup of actors, also known as clusters or cohesive groups and can be applied for multiplex networks. Multiplex network nodes are in the same community when they have similarities and tend to share common properties. Therefore, revealing the community structure in a multiplex graph can provide a better understanding of the overall functioning of the network (Magnani, Hanteer, et al., 2021).

Using R programming language and the multinet package (Magnani, Rossi, et al., 2021), the above procedure was implemented and the graph was ready to be defined as the input to the “glouvain” function which uses a community detection algorithm described by (Mucha et al., 2010). The goal is to find community structures across layers, where vertices in different layers can belong to the same or a different community despite corresponding to the same actor. This community detection algorithm is based on modularity optimization; that is, it tries to find an assignment of the vertices to communities so that the corresponding value of modularity is as high as possible. Multiplex modularity is a quality metric function that takes higher values if most of the edges are between vertices in the same community and if vertices corresponding to the same actors are also often in the same community. Modularity is defined as:

$$Q_m = \frac{1}{2\mu}\sum_{i\ jsr}\left[\left(a_{ijs} - \frac{k_{is}k_{js}}{2m_s}\right)\delta(s,r) + \omega\delta(i,j)\right]\delta(\gamma_{is},\gamma_{jr}) \quad (1)$$

where i, j are actors, s, r are layers, α_{ijs} is 1 if i, j are adjacent on layer s, k_{is} is the degree of actor i on layer s, μ is the number of pairs of vertices either adjacent on α layer or corresponding to the same actor, m_s is the number of edges in layer s, γ_{is} is the community to which actor i on layer s is assigned to, δ is the Kronecker delta, and ω is a weight; when the same actor belongs to the same community on two different layers, then Q_m is increased by ω.

Next, we need to set the omega parameter, which takes values from 0 to 1, and study its impact. Setting higher values of omega will result in communities that will span on multiple layers and will consist of the same actors, since this way the value of modularity increases. On the other hand, with omega set to 0, having the same actors on different layers in the same community does not contribute to modularity (Magnani, Rossi, et al., 2021).

The purpose of community detection on our data is to include different layers and actors in the same community, in order to extract useful information. So, it was considered the omega value to be 0.001, as communities will span multiple layers and will not have the same actors on different layers in the same community. From the different groups that were created, since the nodes that belong to the same community have some similarities and common characteristics, it is deduced that the number of passing vehicles or validated tickets of one area of Athens a specific hour may have an effect or relation on another area at the same or different hour. This is a useful feature and is used later to build a forecasting model.

3. IMPLEMENTATION AND RESULTS

The Dataset

The data used in this work were collected from an open-source database that was developed by the Greek Government and the Region of Attica that includes the hourly traffic volume passing from more than 400 loop detectors and the hourly number of passengers embarking on the subway (metro) train at each of the 63 stations. The data are available for download for academic purposes at https://www.data.gov.gr/search/?topic=transport. For the purposes of this paper, we used data of 10 months (January to October of 2021) from 113 of the most significant loop detectors that are located around the center of the city of Athens, as well as the demand of all metro stations. Figure 1 shows the exact location of the loop detectors and the stations exploited.

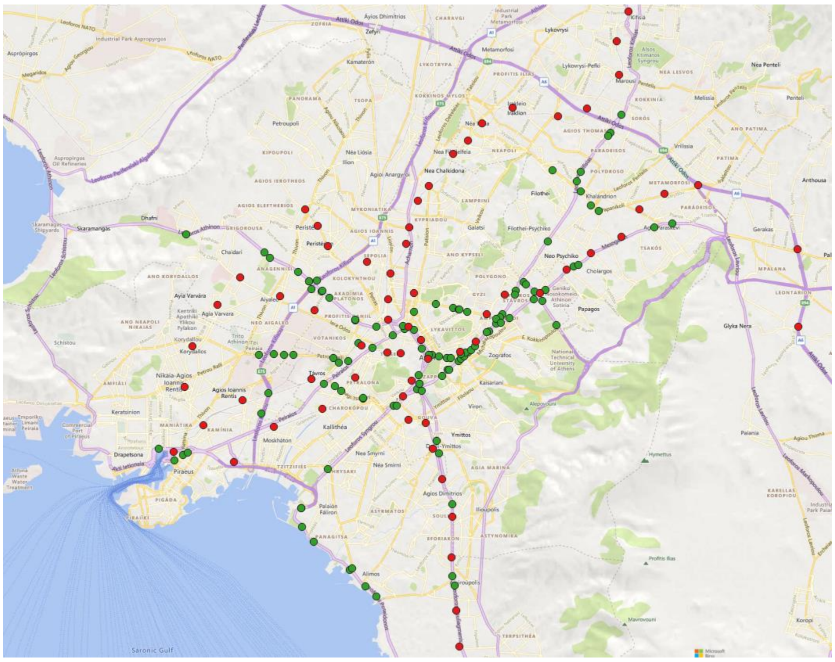


Figure 1: Geographic distribution of loop detectors (green) and metro stations (red)

Multiplex Network Representation and Community Detection

As was already mentioned, we decided to represent the transportation network of Athens (detectors and metro stations) as a multiplex graph, in order to detect communities and obtain valuable insights into the spatial and temporal relations that define it. The construction of the graph was based on the idea that each layer would correspond to each hour of the day, so the graph consists of 24 layers. Each layer consists of 176 nodes, each of which represents a loop detector or a metro station. Next, to define the edges of each of the layers, we got the time series of the demand of each node at the corresponding time of the day (e.g. every day at 9 a.m.) for a period of three months and calculated the mutual information between the time series of all nodes for the same hour. If the mutual information between two nodes is higher than 0.7, which indicates a significant correlation, an edge is created between them. Of course, we scaled the 2 types of data in order to place them in the same structure, as they refer to different measurements. That way we model the spatial correlation between the nodes. It should be reminded that the existence of an edge between two nodes is very important when estimating the modularity metric and detecting communities of nodes. By applying this methodology, we constructed each layer of the multiplex graph to which we can apply a community detection algorithm. We would prefer each community to contain different nodes at each time of the day (i.e. layer) and each instance of a node not to necessarily belong to the same community, in order to capture the temporal relations as well.

By applying the community detection algorithm described above, we end up with 7 communities which are presented in Figure 2. Different colors correspond to different communities. Nodes belonging to the same community have similar structural properties. As a significant amount of data was used to create the multilayer graph's edges, the results of the community detection are generalizable.

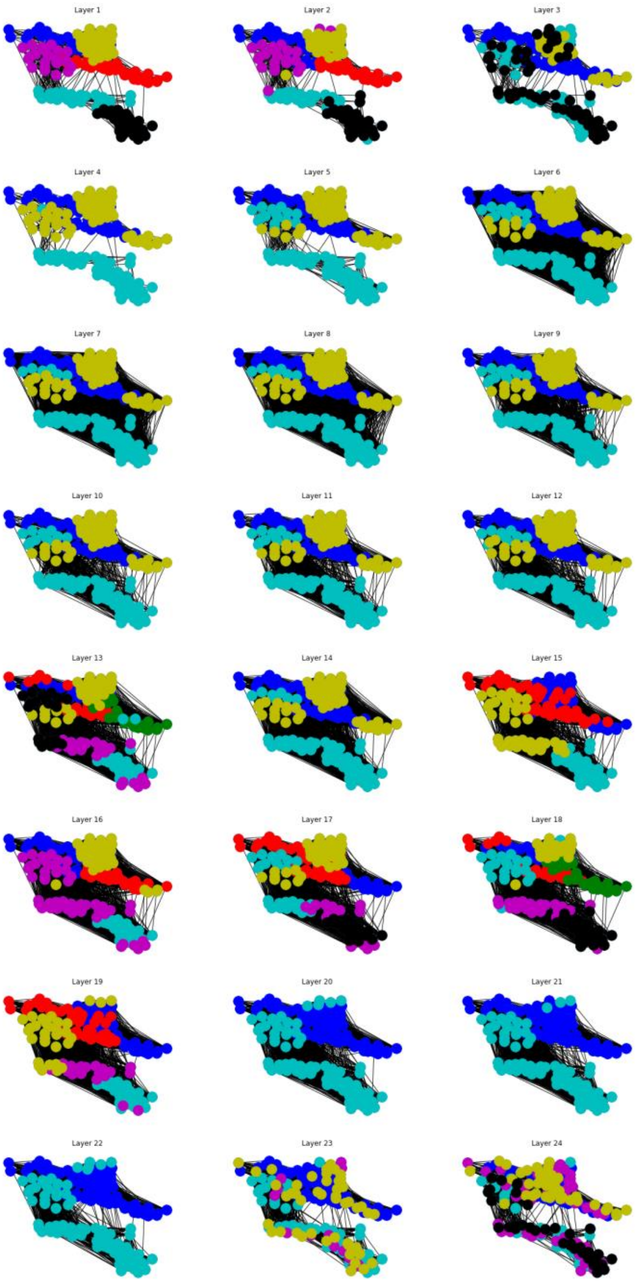


Figure 2: Detected communities

Model Development and Feature Selection

The overarching goal of this paper is the exploitation of a meaningful transport network representation, such as the multiplex network, and the community detection concept to enhance the predictability of traffic conditions of a road network, using a relatively simple model. After detecting the 7 communities presented earlier, a model for each node and for each time of the day can be developed to predict the demand at the corresponding detector or station, using as features the latest values of demand of the rest of the nodes that belong to the same community. For example, the demand at node i at 9:00 can be predicted from the demands of nodes j at 7:00 and k at 22:00 (of the previous day) if these three nodes belong to the same community.

So, the number of features of each model is equal to the size of the community to which the requested pair (device, hour) belongs and the value of each feature corresponds to the measurements of the previous 24 hours. Therefore the developed models would have a large number of features since the communities that were built have a size from 53 to over 1000. Thus, we used the PCA method in order to reduce dimensionality. The new dimension was chosen for each model to be set to 20 as it was observed that thus the cumulative explained variance was over 95%.

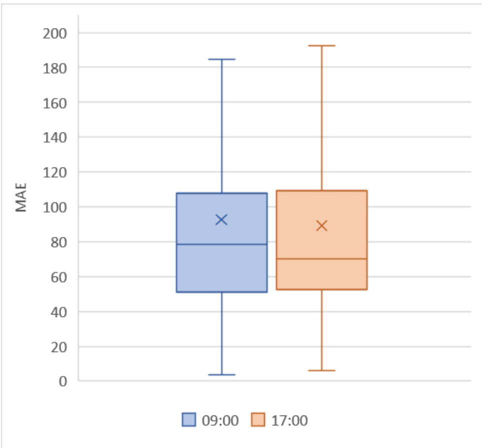
Then, for each device-time pair, a Neural Network was developed with an input size consisting of 20 features, as it emerged from PCA, with input layer size 32, 3 hidden layers size 64 and with the rectified linear unit (ReLU) being the activation function. Since the data covers a period of 303 days (10 months), so is the size of the input to each model. The train-test data split was random in which the test set consisted of 33% of the records. When adjusting the hyperparameters of the model, the validation set consisted of 20% of the training set, the number of epochs was 500 and the batch size was 32. With this method, we trained 4224 neural networks, which are the number of device-hour pairs.

Results

For sake of brevity, the results of the models for the entire transport network (all detectors and stations) at 9:00 and 17:00 will be presented. The models were evaluated by the calculation of the mean absolute error (MAE) and the mean absolute percentage error (MAPE) of the test set. As it is obvious in Table 1, the models achieve a satisfying overall accuracy, which justifies the use of the proposed methodology. The overall MAE and MAPE values refer to the average error of the predictions concerning each node. The distributions of the error metrics are also presented in Figure 3.

Table 1: Average prediction error metrics

Time	**9:00**	**17:00**
Overall MAE	92.5	88.1
Overall MAPE	0.27	0.19



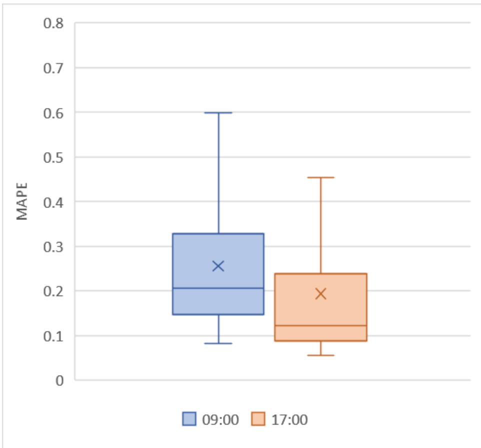
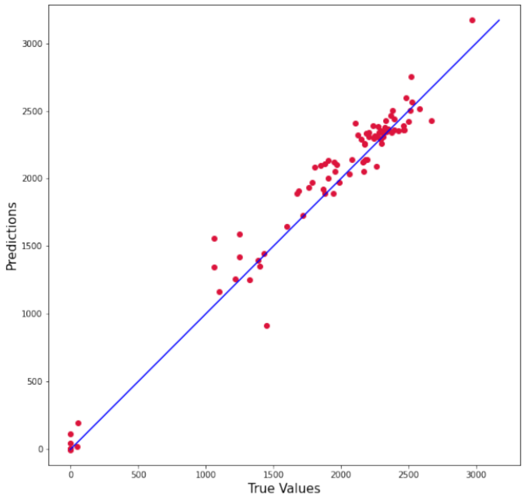
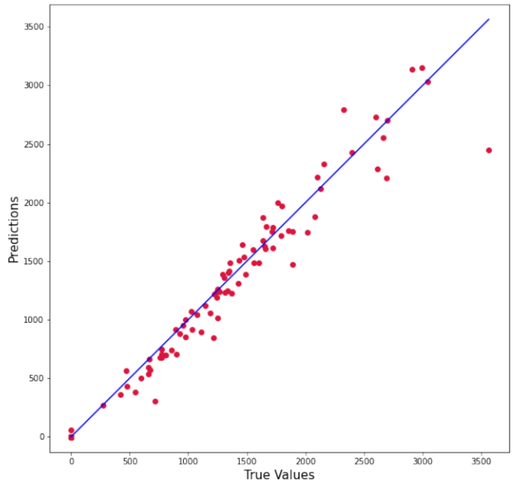


Figure 3: Error distribution on different nodes

It is evident that the values of the error metrics have a relatively high variation, which is reasonable, due to the high number of different models developed for the entire network. However, the majority of the nodes' demand is predicted with very good accuracy (MAPE 50% quartile ≤ 0.2 for both cases) and, interestingly, 25% of the nodes are predicted with a MAPE lower than 10%. Moreover, in Figure 4 the scatter plots of the predicted and actual values of two indicative nodes (one loop detector and a metro station) are presented. One may observe that the points are close to the x=y line and no systematic error can be detected.



(a) Detector of Vasilissis Sofias street



(b) Evangelismos metro station

Figure 4: Scatter plots of predicted vs actual hourly traffic volume.

4. CONCLUSIONS

In this paper, a novel representation of a road network is proposed along with a feature selection framework that are adopted from the research are of Social Networks analysis. The representation of the road network using a multilayer graph highlights the spatial and temporal relations of its nodes and combined with the community detection framework, lead to accurate and efficient models, as the results indicate.

Our future research will focus on using data of higher temporal resolution, which would allow shorter-term predictions, as well as the highlighting of causation between the independent and dependent variables.

ACKNOWLEDGMENT

This research is part of TANGENT project, which has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 955273. TANGENT partners are DEUSTO (coordinator) AIMSUN, NTUA, IMEC, CEFRIEL, Ruprecht, ID4CAR, Rennes, A-to-Be, Carris, TfGM, Panteia and Polis.

REFERENCES

Boukerche, A., & Wang, J. (2020). Machine Learning-based traffic prediction models for Intelligent Transportation Systems. *Computer Networks*, *181*(August), 107530.

https://doi.org/10.1016/j.comnet.2020.107530

Bródka, P., Chmiel, A., Magnani, M., & Ragozini, G. (2018). Quantifying layer similarity in multiplex networks: A systematic study. *Royal Society Open Science*, *5*(8). https://doi.org/10.1098/rsos.171747

Do, L. N. N., Taherifar, N., & Vu, H. L. (2019). Survey of neural network-based models for short-term traffic state prediction. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(1), 1–24. https://doi.org/10.1002/widm.1285

Jiang, W., & Luo, J. (2021). *Big Data for Traffic Estimation and Prediction: A Survey of Data and Tools*. http://arxiv.org/abs/2103.11824

Lee, K., Eo, M., Jung, E., Yoon, Y., & Rhee, W. (2021). Short-Term Traffic Prediction with Deep Neural Networks: A Survey. *IEEE Access*, *9*, 54739–54756. https://doi.org/10.1109/ACCESS.2021.3071174

Magnani, M., Hanteer, O., Interdonato, R., Rossi, L., & Tagarelli, A. (2021). Community Detection in Multiplex Networks. *ACM Computing Surveys*, *54*(3). https://doi.org/10.1145/3444688

Magnani, M., Rossi, L., & Vega, D. (2021). Analysis of multiplex social networks with r. *Journal of Statistical Software*, *98*(8). https://doi.org/10.18637/jss.v098.i08

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J. P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science*, *328*(5980), 876–878. https://doi.org/10.1126/science.1184819

Tedjopurnomo, D. A., Bao, Z., Zheng, B., Choudhury, F., & Qin, A. K. (2020). A Survey on Modern Deep Neural Network for Traffic Prediction: Trends, Methods and Challenges. *IEEE Transactions on Knowledge and Data Engineering*, *14*(8), 1–1. https://doi.org/10.1109/tkde.2020.3001195

Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2014). Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*, *43*, 3–19. https://doi.org/10.1016/j.trc.2014.01.005

Wang, Y., Zhang, D., Liu, Y., Dai, B., & Lee, L. H. (2019). Enhancing transportation systems via deep learning: A survey. *Transportation Research Part C: Emerging Technologies*, *99*(December), 144–163. https://doi.org/10.1016/j.trc.2018.12.004

Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H., & Yin, B. (2021). Deep Learning on Traffic Prediction: Methods, Analysis and Future Directions. *IEEE Transactions on Intelligent Transportation Systems*, 1–15. https://doi.org/10.1109/TITS.2021.3054840