

Data-Driven Traffic Assignment: A Novel Approach for Learning Traffic Flow Patterns Using Graph Convolutional Neural Network

Rezaur Rahman^a and Samiul Hasan^{*/✉}

^a *Department of Civil, Environmental, and Construction Engineering*

University of Central Florida

Email: rezaur.rahman@knights.ucf.edu

^{*/✉} *Corresponding Author, Department of Civil, Environmental, and Construction Engineering*

University of Central Florida

Email: samiul.hasan@ucf.edu

ABSTRACT

We present a novel data-driven approach of learning traffic flow patterns of a transportation network given that many instances of origin to destination (OD) travel demand and link flows of the network are available. Instead of estimating traffic flow patterns assuming certain user behavior (e.g., user equilibrium or system optimal), here we explore the idea of learning those flow patterns directly from the data. To implement this idea, we have formulated the traditional traffic-assignment problem (from the field of transportation science) as a data-driven learning problem and developed a neural network-based framework known as Graph Convolutional Neural Network (GCNN) to solve it. The proposed framework represents the transportation network and OD demand in an efficient way and utilizes the diffusion process of multiple OD demands from nodes to links. We validate the solutions of the model against analytical solutions generated from running static user equilibrium-based traffic assignments over Sioux Falls and East Massachusetts networks. The validation results show that the implemented GCNN model can learn the flow patterns very well with less than 2% mean absolute difference between the actual and estimated link flows for both networks under varying congested conditions. When the training of the model is complete, it can instantly determine the traffic flows of a large-

[This is the pre-print version of the article: Rahman, R., & Hasan, S. (2023). Data-Driven Traffic Assignment: A Novel Approach for Learning Traffic Flow Patterns Using Graph Convolutional Neural Network. *Data Science for Transportation*, 5(2), 11, which has been published in final form at <https://doi.org/10.1007/s42421-023-00073-y>].

scale network. Hence this approach can overcome the challenges of deploying traffic assignment models over large-scale networks and open new directions of research in data-driven network modeling.

Keywords: Traffic assignment problem, Data-driven method, Deep learning, Graph convolutional neural network

1. INTRODUCTION

The traffic assignment problem (TAP) is one of the key components of transportation planning and operations. It is used to determine the traffic flow of each link of a transportation network for a given travel demand based on modeling the interactions among traveler route choices and the congestion that results from their travel over the network (Sheffi, 1985). Traditionally traffic assignment problems have been formulated as mathematical programs and solved based on user equilibrium (UE) principles (Wardrop, 1952). The UE solution relies on several assumptions on user behavior and knowledge such as: (i) drivers have perfect information and knowledge about the underlying network; (ii) drivers make rational choices when choosing a route; and (iii) all drivers are homogeneous (Kim et al., 2009). Although some of these assumptions may not hold in a real-world scenario, this approach has been providing the most reasonable solutions of the traffic assignment problem (Bar-Gera, 2002; Jafari et al., 2017; LeBlanc et al., 1975; Leurent et al., 2011; Mitradjieva and Lindberg, 2013; Tizghadam and Leon-garcia, 2007). One of the major issues with the static traffic assignment solution is that it assumes constant OD demand whereas traffic demand may significantly change over time, which makes it unsuitable to deploy for real world traffic scenario.

Static traffic assignment problem fails to capture traffic dynamics of a network, which motivated researchers to move towards dynamic traffic assignment methods. Dynamic traffic assignment (DTA) has become a state-of-the-art solution method to capture traffic network

dynamics in a more realistic way. Numerous formulation and solution approaches have been introduced ranging from dynamic programming to variational inequality to simulation-based approaches (Abdelfatah and Mahmassani, 2002; Ban et al., 2008; Barthélemy and Carletti, 2017; Ben-Akiva et al., 2012; Boyles et al., 2006; Foytik et al., 2017; Friesz et al., 1989; He et al., 2010; Janson, 1989; Ji et al., 2016; Jiang et al., 2011a, 2011b; Liu et al., 2007; Lo and Szeto, 2002; Mahmassani, 2001; Merchant and Nemhauser, 1978; Nie and Zhang, 2010; Peeta and Mahmassani, 1995; Peeta and Ziliaskopoulos, 2001; Primer, 2011; Ran et al., 1993; Shafiei et al., 2018; Ukkusuri et al., 2012; Waller et al., 2013; Watling and Hazelton, 2003). In a smaller model with less complexity, the DTA process works well in converging to a point of equilibrium. As the scale of a model grows in size, complexity, and congestion, the DTA procedure may become more difficult and require more computational time making it less suitable for a real-time deployment. Researchers are exploring to develop efficient solution methods for traffic assignment problem (TAP) which can be deployed for real-time transportation operation purposes.

Optimization of a transportation network largely depends on how accurately we capture the traffic dynamics of the network to forecast traffic in real-time. In particular, active traffic management requires a system that can estimate and forecast traffic in real-time. Over the last decade, traffic sensing technologies and emerging connected vehicles environment with added computational power has created an opportunity towards developing a data-driven approach as an alternative to traditional approaches of solving traffic assignment problems. Roadway sensors and devices provide us information on instantaneous travel time variation for different links indicating the congestion propagation throughout the network. Moreover, OD demand can be extracted from emerging data sources such as mobile phones, GPS observations, and social media posts. As such, we now have real-world O-D demand and traffic flow data available at large-scale over many days. However, one challenge is how to develop a data-

driven framework at a network scale to learn the relationship between OD demand and link flows.

To overcome this challenge, in this paper, we present a novel data-driven approach of learning traffic flow patterns of a transportation network given that many instances of OD demand and traffic flow of the network are available. Instead of estimating traffic flow patterns assuming certain user behavior (e.g., user equilibrium), here we explore the idea of learning those patterns from large-scale training data by developing a neural network architecture. In particular, we use the Graph Convolutional Neural Networks (GCNN) which generalize traditional neural networks to work on structured graphs (Kipf and Welling, 2016). GCNNs utilize an adjacency matrix or a Laplacian matrix to represent the structure of a graph. Recently, several studies have utilized the concept of graph convolution to represent traffic network as a generalized graph for traffic state prediction (Cui et al., 2018a; Li et al., 2018). Moreover, Graph Convolution Neural Network has emerged as a new approach to overcome the challenge of high dimensionality when predicting travel demand for a large-scale network (Kim et al., 2019; Lin et al., 2018).

However, this is the first time, we developed a GCNN model to learn the traffic flow pattern of a transportation network through learning the diffusion process of multiple OD demands from nodes to links. In this work, we seek to answer the question: *can a deep learning model learn the flow patterns of a network without relying on the assumptions of user behavior for assigning traffic in the network?* If positive, this will offer a learning-based traffic assignment solution, which means that the model can learn the interaction between demand and supply purely from data given that real-world data of O-D demand and traffic flows are available. Apart from assigning traffic in a purely data-driven way, this can open new directions in transportation network modeling research offering data-driven approaches for solving dynamic traffic assignment, network design, and many other problems with real-world

applications. Thus, this study makes several contributions:

- (i) It formulates the traffic assignment problem from the field of transportation science as a data-driven learning problem considering the underlying transportation network and available instances of OD demand and link flows.
- (ii) It develops a novel neural network architecture that can solve the learning problem to determine link flows based on OD demand and traffic flow data.
- (iii) It provides rigorously tested experimental evidence that such a neural network architecture can learn the user equilibrium traffic flow assignment of a transportation network only from data.

2. DATA DRIVEN METHODS IN NETWORK LEVEL TRAFFIC MODELING

With the emergence of big data and scalable data processing technologies data-driven modeling of complex system is gaining more attention. In transportation data-driven approaches have widely been applied to solve traffic prediction problems. These involve forecasting traffic speed (Cai et al., 2016; Ma et al., 2015; Polson and Sokolov, 2017; Rahman and Hasan, 2018), volume (Luo et al., 2019; Yasdi, 1999), and travel time (Billings and Jiann-Shiou, 2006; Chien and Kuchipudi, 2003; Elhenawy et al., 2014; Innamaa, 2005; Myung et al., 2011; Wu et al., 2004; Zhang and Haghani, 2015). Traditional data-driven methods such as Support vector regression (SVR), K-nearest Neighbor regression (KNN), Kalman Filtering (KF), Autoregressive Integrated Moving Average (ARIMA), Decision Tree, Artificial Neural Network (ANN), and Long Short Term Neural Network (LSTM) have been applied for short-term traffic prediction over only one or multiple segments of highways, but not at the scale of a network (Billings and Jiann-Shiou, 2006; Cai et al., 2016; Chien and Kuchipudi, 2003; Deshpande and Bajaj, 2016; Elhenawy et al., 2014; Innamaa, 2005; Lee, 2009; Luo et al., 2019; Ma et al., 2015; Myung et al., 2011; Polson and Sokolov, 2017; Rahman and Hasan, 2018; Wu et al., 2004; Yu et al., 2016; Zhang and Haghani, 2015). Moreover, they do not consider network

wide travel demand variations while predicting future traffic. As such, these approaches consider traffic prediction as a simple time series problem and predict the traffic state for a shorter time horizon (e.g., next 5 to 15 min). However, to estimate traffic for an entire network similar to the solution of a traffic assignment problem, we need to understand traffic evolution and congestion propagation for an entire road network rather than on a single road. Hence, researchers are investigating new data-driven methods to estimate network scale traffic which can be deployed for real-time transportation operation purposes.

Network level traffic prediction is more challenging due to the higher computational complexity required by the network topology. Such a problem requires (i) to capture the travel demand variation in the network, (ii) to capture the correlation of traffic in interconnected roads and mapping it in a spatial network (Polson and Sokolov, 2017), and (iii) to reflect drivers' route choice and associated congestion propagation inside the network. Convolutional Neural Network (CNN) methods are the initial attempt to model spatial and temporal correlation among the traffic states for network level traffic prediction. A few studies (Liu et al., 2017; Ma et al., 2017), have implemented the CNN and its variants such as CNN-LSTM model for network level traffic prediction. Although this model outperforms existing state-of-the-art data-driven models, in a CNN model, the traffic network is considered as an image and network information is extracted using a convolution filter, which is then fed into a fully connected layer or an LSTM layer to predict future traffic state. While traffic network is converted into an image, a certain amount of noise is included into the spatial relationships of traffic state causing erroneous results in prediction. Moreover, CNN cannot capture the long-term congestion propagation inside the network for long term traffic forecasting. Hence, it fails to capture the stochastic nature of traffic dynamics of a transportation network.

Recently, graph theory coupled with generalized neural network architecture has been utilized to model the dynamics between the structural properties and the functions of a network

(Atwood and Towsley, 2015; Defferrard et al., 2016; Li et al., 2018; Sanchez-Gonzalez et al., 2020; Zhang et al., 2019; Zhou et al., 2018). Thereby, solving problems such as modeling physical systems, learning molecular fingerprints, predicting protein interface, and classifying diseases, which require that a model learns from graph-based inputs. Several studies have applied graph convolution-based approaches to learn spatiotemporal features of a transportation network (Atwood and Towsley, 2015; Cui et al., 2020, 2018b, 2018a; Guo et al., 2021, 2020, 2019; Kim et al., 2019; Li et al., 2021, 2018; Lin et al., 2018; Peng et al., 2021, 2020; Tang et al., 2020; Yu et al., 2018; Zhao et al., 2020). One of the main limitations of these studies is that they focused on learning traffic representation at a network scale rather than modeling the interaction between travel demand and traffic flow diffusion process. In other words—without including demand in the modeling framework—these methods disregard the fundamental concept of network flow modeling. Consequently, they fail to explain the working mechanism and support the reliability of these methods in network wide traffic flow estimation.

In this study, we develop a data-driven method to model the interaction between network wide travel demand variations and links flows. Thereby, we formulate the traffic assignment problem as data-driven learning problem. To solve this problem, we proposed a novel deep learning architecture adapting the concept of graph convolution.

3. DATA-DRIVEN TRAFFIC ASSIGNMENT

In this section, first we formulate the data-driven traffic assignment problem. We then describe the Graph Convolutional Neural Network (GCNN) approach to model traffic flow propagation in the network.

3.1 Problem Definition

In a transportation network, all nodes are connected, and each link is associated with information such as distance, speed limit, capacity etc. Here, we consider the transportation

network as a weighted directed graph $\mathcal{G}(v, \mathcal{E}, \mathbf{A}_w)$ where v denotes the set of nodes and \mathcal{E} denotes the set of links between nodes (i, j) . \mathbf{A}_w represents the connectivity between nodes as a weighted adjacency matrix, where weights are based on free flow travel time between any two nodes (i, j) , defined as follows:

$$\mathbf{A}_w(i, j) = \begin{cases} t_{i,j}^0 & \text{if } i \rightarrow j \\ t_{j,i}^0 & \text{if } j \rightarrow i \\ 0, & \text{if } i = j \end{cases} \quad (1)$$

where, t^0 denotes the free flow travel time between the origin and the destination nodes. The proposed data-driven formulation of the traffic assignment problem aims to *learn* the flow patterns of a transportation network based on the network structure and the instances available on origin to destination (OD) travel demand and link flows (Fig.1). Also, we have the information on network characteristics, such as location of each node with respect to other nodes, travel distance or free flow travel time between different nodes. From this information, we develop a data-driven method to estimate the link flows for given travel demands. During the estimation process, we also learn the traffic flow propagation from origin nodes toward destination nodes in a transportation network. Let, \mathbf{X} be the demand matrix for the transportation network \mathcal{G} , where each element of a row indicates the travel demand between origin node (i) and the destination node (j) .

The traffic assignment problem aims to learn a function $\mathcal{F}(\cdot)$ that maps m instances of OD demand matrix $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_2 \dots \dots \dots, \mathbf{X}_m)$ to m instances of flow $(F_1, F_2, F_3 \dots \dots \dots F_m)$, defined as follows,

$$\mathcal{F}([\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_2 \dots \dots \dots, \mathbf{X}_m]; \mathcal{G}(v, \mathcal{E}, \mathbf{A}_w)) = [F_1, F_2, F_3 \dots \dots \dots F_m] \quad (2)$$

where, \mathbf{A}_w indicates the weighted adjacency matrix, \mathcal{E} indicates the set of links of the network, and the vector F_m contains the link flows for each link of the network for a given OD demand (\mathbf{X}_m) . In this formulation, OD demands and network properties are input variables and link flows are the target variables.

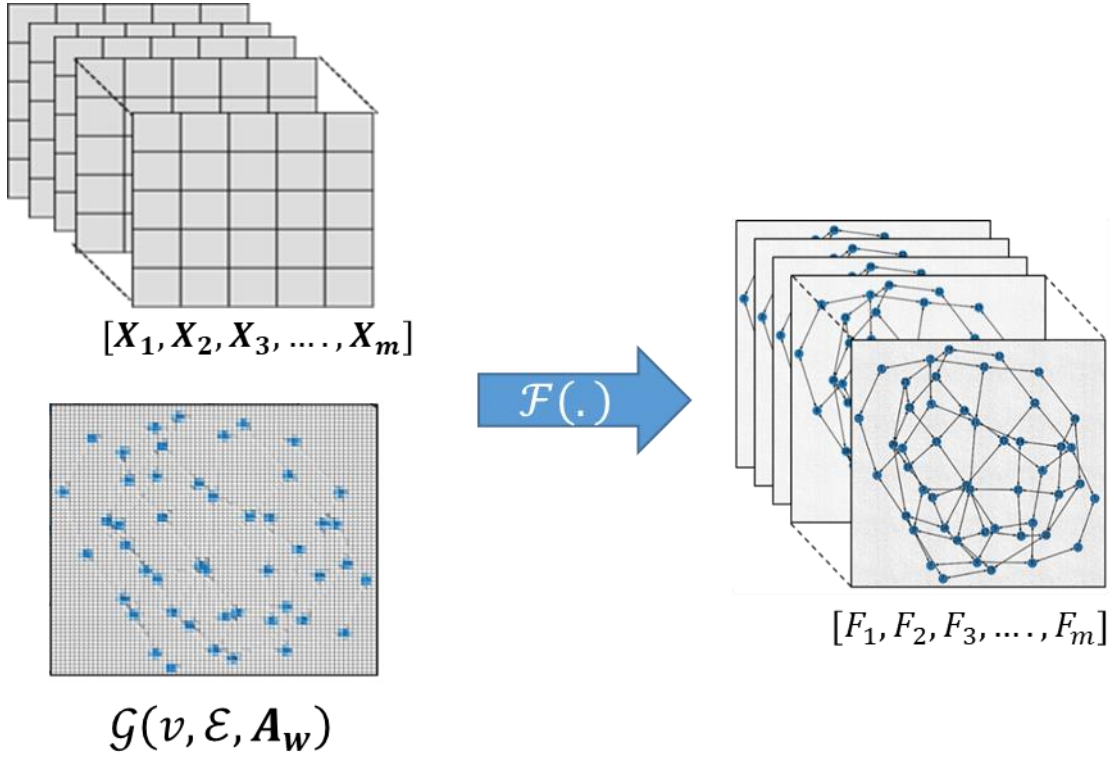


Fig. 1. A Schematic View of the Data-Driven Traffic Assignment Problem

Table 1 List of notations

Notation	Description
\mathcal{G}	Transportation network
v	Set of nodes in \mathcal{G} with size of $ v = N$
\mathcal{E}	Set of links in \mathcal{G} with size of $ \mathcal{E} = E$
$A_w \in \mathbb{R}^{N \times N}$	Weighted adjacency matrix of \mathcal{G} , defined by Equation (1)
$I \in \mathbb{R}^{N \times N}$	Identity matrix
$\bar{A}_w \in \mathbb{R}^{N \times N}$	Neighborhood matrix defined by Equation (7)
$D_w \in \mathbb{R}^{N \times N}$	Degree matrix of \mathcal{G} , a diagonal matrix where diagonal elements (i, i) indicate the number of links coming out from a node
$L_w \in \mathbb{R}^{N \times N}$	Laplacian matrix represents the structural properties of a network, defined by the equation (12)
$t_{i,j}^0$	Free flow travel time between nodes i and j
$X \in \mathbb{R}^{N \times N}$	OD demand matrix
$F \in \mathbb{R}^E$	Flow vector contains flows for each link of the network for a given OD demand X

Notation	Description
$\mathcal{P} \in \mathbf{R}^{N \times N}$	Routing matrix indicates the probability of diffusion of flow from node i to node j
\mathbf{g}_{θ}	Convolutional filter to learn the network features along with network function
$f(.)$	Activation function
$\Theta \in \mathbf{R}^{N \times N}$	Learnable parameters for the convolution filter
$\mathbf{W}_q, \mathbf{W}_F$	Learnable parameters for link flow estimation
\mathbf{q}	Flow distribution matrix indicating the propagation of flow from a given node to different links of the network
\mathbf{H}	Indicates the outputs from different layers of the proposed neural network architecture

All the bold letters denote a matrix

3.2 Graph Convolution Neural Network for Flow Pattern Learning

We develop a Graph Convolution based Neural Network (GCNN) architecture to assign traffic in the transportation network. We use graph convolution operation to learn the network properties and the flow diffusion process from origin nodes toward destination nodes. To estimate link flows, we model how this flow diffusion process is contributing to link flows. In other words, the model considers how flows are coming to a specific link from adjacent nodes while diffusing from origin nodes towards destination nodes. In the proposed GCNN model, a convolution filter is derived based on the network structure (position of the nodes and links) and the flow diffusion process inside the network. Hence, GCNN model captures changes in traffic states of a transportation network by modeling the flow diffusion from origin nodes toward destination nodes. As such, the GCNN model simultaneously learn the features (each node and link are embedded with valuable information) and function of the network.

Once the model learns the flow diffusion process, we can feed this information into a feed forward neural network to estimate the traffic flows at different links. In the following section, we describe the final architecture of the proposed deep neural network model to estimate traffic flows at different links.

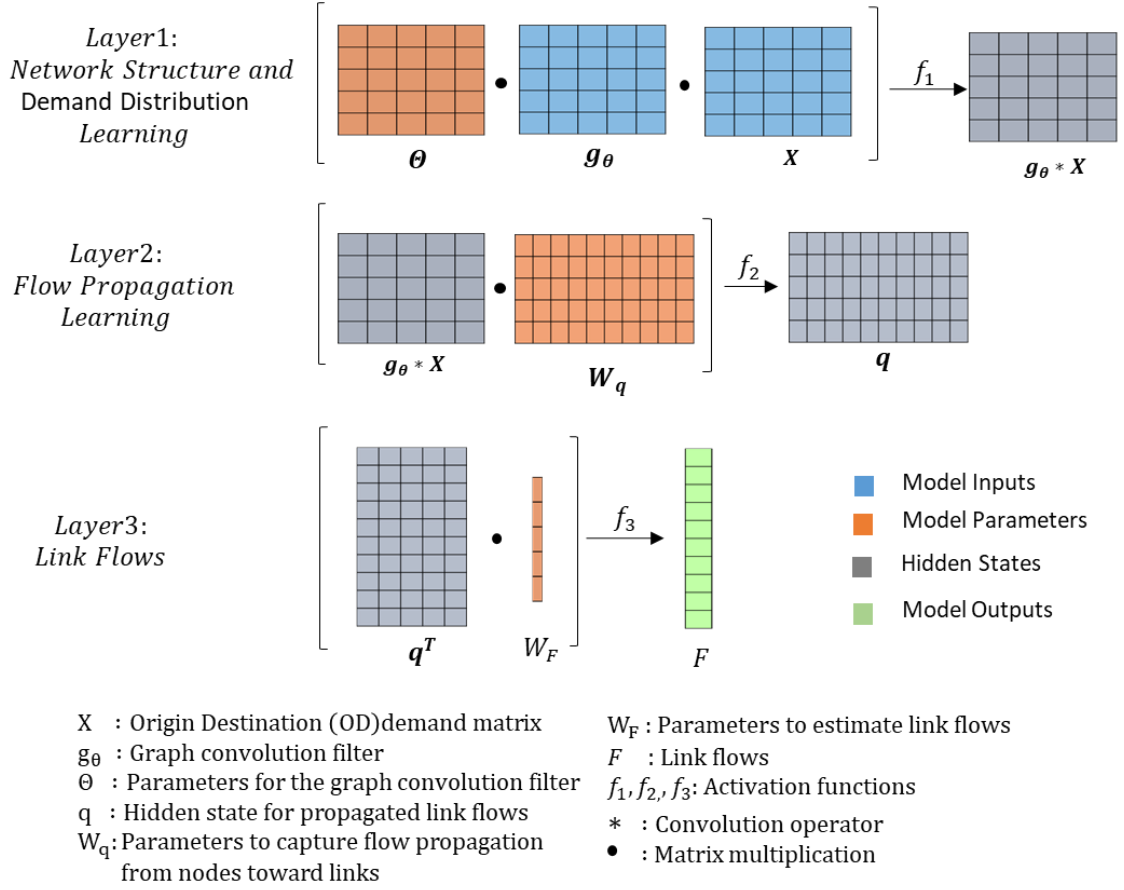


Fig. 2. GCNN architecture for traffic flow pattern learning

Fig. 2 shows all the layers and matrix operations in each layer for the proposed GCNN architecture. In the first layer, we define the graph convolutional operation to learn the network properties and demand distribution for different nodes. We derive the graph convolution filter from weighted adjacency matrix where weights are assigned based on the free flow travel time of a link (see equation 1). We perform the convolution operation between the OD demand (X) and graph convolution filter (g_θ). The convolutional filter (g_θ) represents the diffusion process of the traffic flows from the origin node towards destination nodes. While training the model, we estimate the parameters (θ) for this convolutional filter. We define the graph convolution layer as follows:

$$\mathbf{H}^1 = f_1(g_\theta * X) = f_1(\Theta g_\theta X) \quad (3)$$

here, $f_1(\cdot)$ denotes the nonlinear activation function for the convolution layer and $\mathbf{H}^1 \in R^{N \times N}$

indicates the output from the graph convolution layer (1st layer). From this layer, we obtain a convoluted demand matrix (\mathbf{H}^1) representing the flow diffusion process (from origin nodes towards destination nodes) inside a network.

The convoluted demand matrix is then fed into the 2nd layer of the GCNN, where we model traffic flow distribution from origin nodes towards adjacent links. In this layer, we create a simple neural network model with parameters W_q , which maps the convoluted demand matrix to a $N \times E$ dimensional space (same size as the link-node adjacency matrix) via matrix multiplication. In this way, the GCNN captures how flow diffusion process will assign traffic at different links of the network. We define the 2nd layer of the model as follows,

$$\mathbf{H}^2 = \mathbf{q} = f_2(W_q \mathbf{H}^1) \quad (4)$$

here, $f_2(\cdot)$ denotes the nonlinear activation function for the second layer and $\mathbf{H}^2 (= \mathbf{q}) \in R^{N \times E}$ indicates the output from the second layer representing distributed link flows from adjacent origin nodes (N) of the network. Which means, each row of the matrix \mathbf{q} , indicates the distributed link flows for all the links (E) associated with an origin node.

Finally, the distributed link flow matrix (\mathbf{q}) is transposed and fed into the output layer (see Fig. 2). Inside the transposed matrix ($\mathbf{q}^T \in R^{E \times N}$) each row indicates the distributed link flows for a given link from all the origin nodes (N). In the output layer, we assign a linear activation function ($f_3(\cdot)$) with N parameters, which aggregates the distributed link flows and outputs assigned traffic flow for a given link. We define the output layer as follows,

$$\mathbf{H}^3 = \mathbf{F} = f_3(W_F(\mathbf{H}^2)^T) = f_3(W_F \mathbf{q}^T) \quad (5)$$

Here, f_3 denotes the linear activation function and $\mathbf{H}^3 (= \mathbf{F}) \in R^E$ denotes the assigned traffic flows for all the links. From the output layer (\mathbf{H}^3), we obtain link flows (\mathbf{F}) for a given OD demand (\mathbf{X}). So, the mathematical formulation of the GCNN model to estimate the link flows can be generalized as follows,

$$\mathbf{F} = f_3((f_2(f_1(\Theta \mathbf{g}_\theta \mathbf{X}) \mathbf{W}_q)^T \mathbf{W}_F) \quad (6)$$

here, we select the hyperbolic tangent function $\tanh = \frac{e(x)-e(-x)}{e(x)+e(-x)}$ as the nonlinear activation function ($f_1(.) = f_2(.) = \tanh$) for the model.

In the following section, we describe the graph convolution operations and the parameters associated with graph convolution filter in details.

3.3 Modeling Traffic Flow using Diffusion Graph Convolution

In a transportation network, the traffic flow pattern changes in response to the changes in travel demand. We represent this relationship via flow diffusion process from origin nodes towards destination nodes. To capture the stochastic nature of traffic flow variation at a network level, we consider the flow diffusion process by a random walk (random movement into adjacent neighboring nodes) in the network, \mathcal{G} with restart probability $\alpha \in [0,1]$ and a state transition matrix $\bar{D}_w^{-1}\bar{A}_w$, where \bar{A}_w is the neighborhood matrix. In the neighborhood matrix, we add an identity matrix (I) with the adjacency matrix (A_w). By adding the identity matrix, we create a self-loop for each node; for a given diffusion step it will capture the traffic flows having same node as origin and destination; in other words, it captures that the traffic flow remains in the origin node rather than diffusing from origin node to destination nodes. The neighborhood matrix \bar{A}_w is defined as,

$$\bar{A}_w = A_w + I \quad (7)$$

The restart probability indicates the probability of starting of a random walk from node i . From the starting node such random walks take multiple steps (diffusion steps, K) to traverse the adjacent nodes until reaching the destination node j . After many time steps, such diffusion process converges to a stationary distribution $\boldsymbol{p} \in \mathbf{R}^{N \times N}$ where i th row of \boldsymbol{p} indicates the probability of flow diffusion from node i towards j . The stationary distribution of the diffusion process can be represented as a weighted combination of infinite random walks on the graph

(Teng, 2016) and be calculated in closed form,

$$\mathcal{P} = \sum_{k=0}^{\infty} \alpha(1 - \alpha)^k (\bar{\mathbf{D}}_w^{-1} \bar{\mathbf{A}}_w)^k \quad (8)$$

where k is the diffusion step. In practice, we can consider a finite number of diffusion steps and assign a trainable weight at each step. Similar diffusion processes have been adopted in previous research (Atwood and Towsley, 2015; Li et al., 2018). Based on that we can define the diffusion convolution over the network features (e.g., OD demand) \mathbf{X} as follows,

$$\mathbf{g}_{\theta} * \mathbf{X} = \sum_{k=0}^{K-1} \Theta_k (\bar{\mathbf{D}}_w^{-1} \bar{\mathbf{A}}_w)^k \mathbf{X} \quad (9)$$

If we consider a 2-step diffusion process the above equation becomes,

$$\mathbf{g}_{\theta} * \mathbf{X} = \Theta_0 \mathbf{X} + \Theta_1 (\bar{\mathbf{D}}_w^{-1} \bar{\mathbf{A}}_w) \mathbf{X} \quad (10)$$

where, Θ_0 and Θ_1 are the weights for each of the steps. While modeling the traffic flow pattern of transportation network, we can consider different number of diffusion steps to reach to a stationary distribution. However, as the network grows, this process becomes computationally expensive, since for a larger network the value for the diffusion step will be higher. Hence, such diffusion processes are only applicable for small scale networks, where flow diffusion occurs among the nearest neighbors (Li et al., 2018; Wang et al., 2020).

In our problem, we propose an alternative approach to represent the diffusion process. Instead of selecting a value of the diffusion step (K) and assigning parameter to each step, we assign parameters (Θ) to locally learn the stationary probability distributions (probability matrix). So, while the training the model we estimate the parameters (Θ) to learn the stationary probability distribution for the flow diffusion process. In other words, we obtain a routing matrix (Leon-Garcia and Tizghadam, 2009) which indicates the probability of diffusion of flow from node i to node j . The resulted diffusion convolution over the OD demand \mathbf{X} can be written as follows:

$$\mathbf{g}_{\theta} * \mathbf{X} = \Theta (\bar{\mathbf{D}}_w^{-1} \bar{\mathbf{A}}_w) \mathbf{X} \quad (11)$$

where, $\Theta \in \mathbf{R}^{N \times N}$ are the parameters of the convolution filter and $\bar{\mathbf{D}}_w^{-1} \bar{\mathbf{A}}_w$ represents the

transition probability matrix of the diffusion process. In the demand matrix \mathbf{X} , each row indicates travel demand from origin node i to destination node j . So, when we perform the matrix multiplication between the operator $\Theta(\bar{\mathbf{D}}_w^{-1}\bar{\mathbf{A}}_w)$ and \mathbf{X} , we obtain a convoluted feature matrix which captures the influence all OD pairs on link flows associated with origin node.

We can also model the diffusion process using a normalized Laplacian matrix. Laplacian matrix better represents the structural properties of a network: the diagonal elements indicate the number of links originating at a given node, while the other elements indicate the connection between the origin and destination nodes. We define the Laplacian matrix as follows,

$$\bar{\mathbf{L}}_w = \bar{\mathbf{D}}_w - \bar{\mathbf{A}}_w \quad (12)$$

Normalizing the Laplacian matrix with degree matrix,

$$\bar{\mathbf{D}}_w^{-1}\bar{\mathbf{L}}_w = \bar{\mathbf{D}}_w^{-1}(\bar{\mathbf{D}}_w - \bar{\mathbf{A}}_w) = \mathbf{I} - \bar{\mathbf{D}}_w^{-1}\bar{\mathbf{A}}_w \quad (13)$$

Now, the convolution over OD Demand matrix, \mathbf{X} can be written as follows,

$$\mathbf{g}_\theta * \mathbf{X} = \Theta(\mathbf{I} - \bar{\mathbf{D}}_w^{-1}\bar{\mathbf{A}}_w)\mathbf{X} \quad (14)$$

where, $\Theta \in \mathbf{R}^{N \times N}$ are the parameters of the convolution filter, in other words, the coefficient matrix of the diffusion equation. During the training of deep learning model, we learn these parameters, which capture the flow diffusion process inside the network.

In this study, we focus on a probabilistic approach to model the flow diffusion by estimating transition probability matrix in two ways: using a random walk on adjacency matrix (equation 11) and Laplacian matrix (equation 14). We also compare these approaches with spectral graph convolutional neural network (Kipf and Welling, 2016) to learn traffic flow patterns. In a spectral graph convolutional approach, the convolutional filter is estimated by decomposing the adjacency matrix into its eigenvalues to represent different properties of the graph such as strength of a node, shortest path distance etc. In Appendix A, we provide the details of the concepts of spectral graph convolution to learn traffic flow patterns of a transportation network.

4. DATA GENERATION

Although we propose this method for real-world traffic data, recent sensing technologies are not densely distributed yet to provide us data necessary to test this approach. Especially, for large networks the OD demand variations are not accessible to us, though such OD demand data exist due to the availability of mobile phone data (Alexander et al., 2015; Gundlegård et al., 2016). In addition, we seek to verify if the proposed approach works across different congested conditions and to measure the gaps between the actual traffic assignment solutions (i.e., analytical solutions) and the solutions obtained from the proposed neural network-based approach. Thus, to verify our approach, we generate synthetic traffic data based on user equilibrium (UE) solutions of static traffic assignments over two networks: Sioux Falls network (24 nodes and 76 links) and East Massachusetts Network (74 nodes and 258 links). It should be noted that our approach is not an alternative method to determine UE solutions. We use the UE based traffic assignment mainly to generate the data to verify our approach.

We obtained the OD demand and information on network characteristics from (Transportation Networks for Research Core Team, 2016). To generate multiple OD demand, we multiplied the OD demand matrix by random factors collected from a uniform distribution which varies 0.1 to 1.0. To test our approach in different scenarios, we consider three conditions: uncongested, moderately congested, and fully congested. We generate 5,000 OD matrices for each condition and solve both networks using the Frank Wolfe algorithm to obtain user equilibrium traffic flows. To represent the prevailing traffic condition, we estimate the flow over capacity ratio. We assume that, for uncongested condition the flow-capacity ratio remains less than 0.5, for moderate condition the flow-capacity ratio varies between 0.4 and 0.8, and for uncongested condition, most of the cases the flow-capacity ratio is greater than 1.0. Fig. 4 shows the traffic flow variations for different links of Sioux Falls Network.

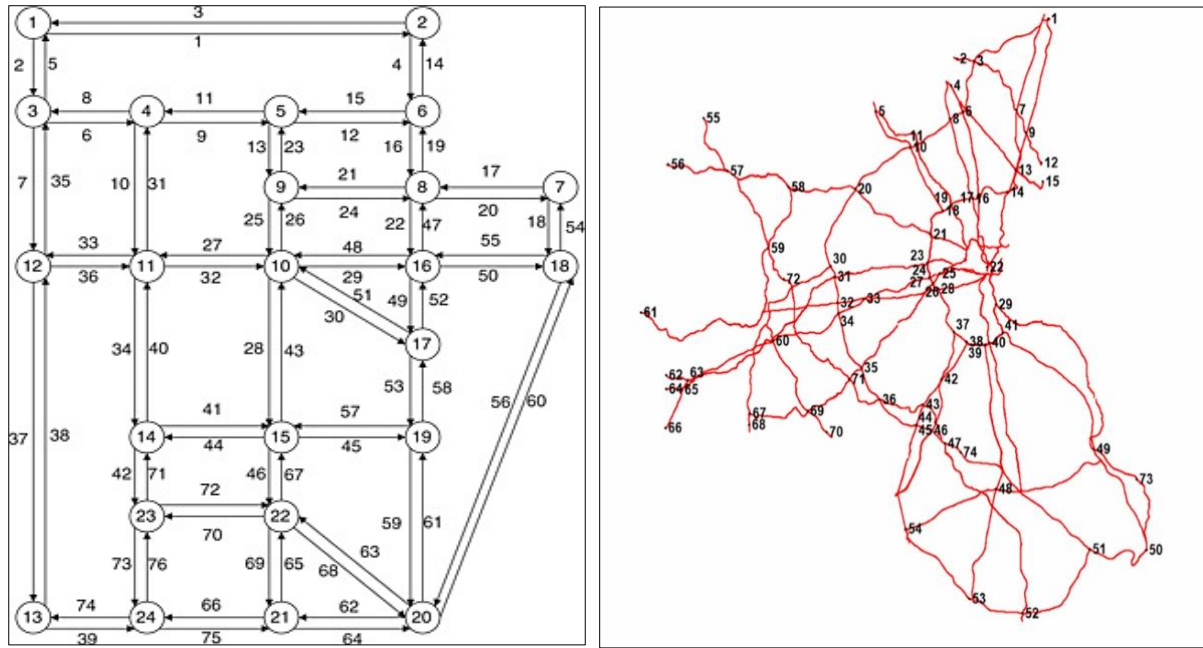
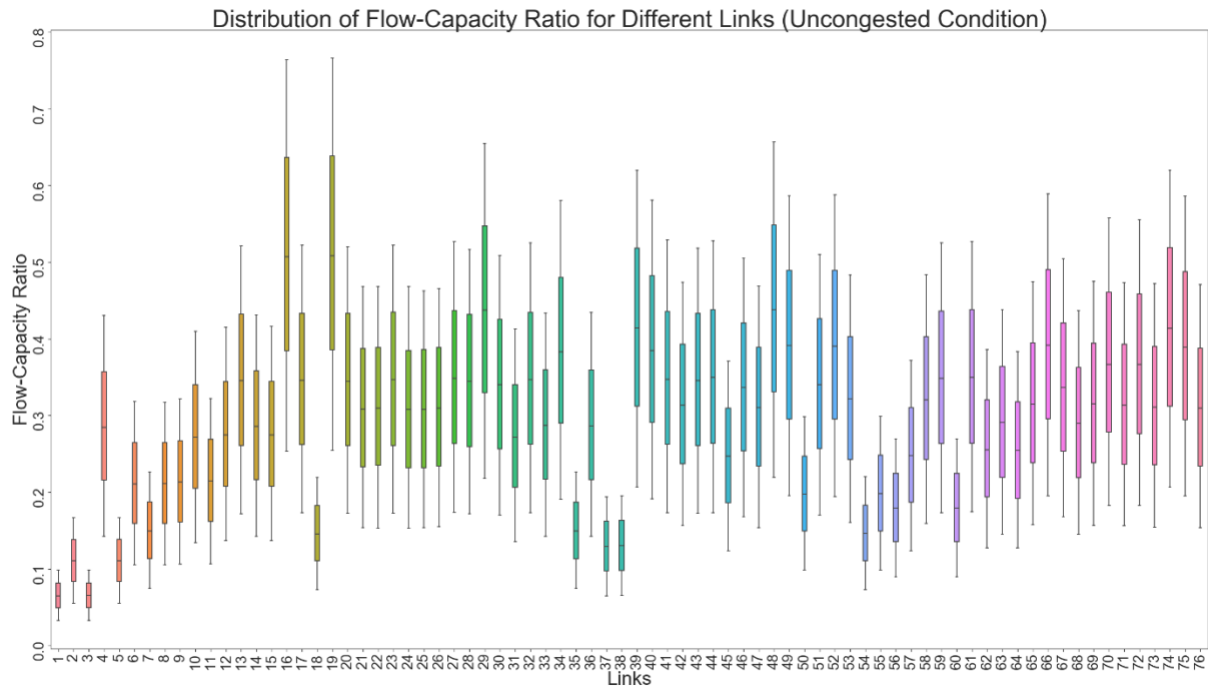
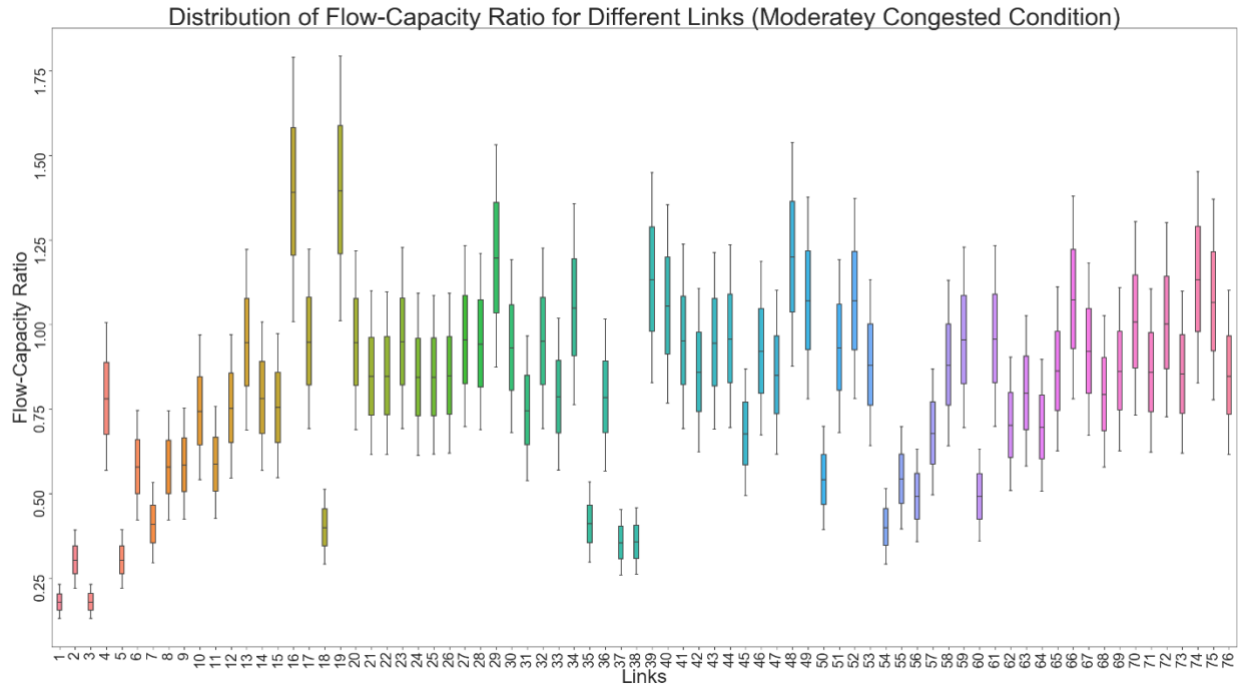


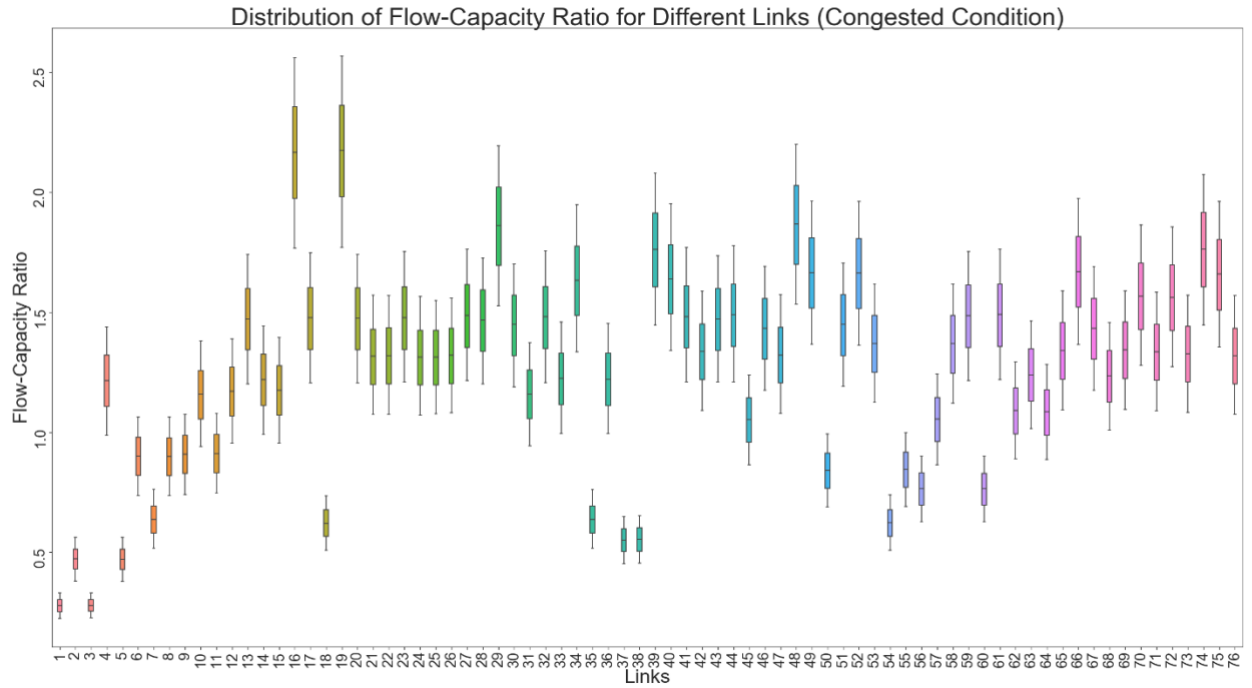
Fig. 3. Transportation Networks (a) Sioux Falls Network (Transportation Networks for Research Core Team, 2016) (b) East Massachusetts Network (US Census Bureau, 2015) (Huang and Kockelman, 2019)



(a) Uncongested Condition



(b) Moderately Congested Condition



(c) Congested Condition

Fig. 4. Illustrates the distribution of flow-capacity ratio for different traffic conditions

5. RESULTS AND DISCUSSION

We implemented all the models using PyTorch (“PyTorch,” 2016) library and train our model with dual NVIDIA Tesla V100 16GB PCIe GPU. Among the OD demand matrices, we use 70% ($n=3,500$) for training, 20% ($n=1,000$) for validation, and rest 10% ($n=500$) for testing the model. We train the model on training data and check the accuracy for the model on validation data set. Based on the validation accuracy, we tune the hyperparameters such as learning rate, types of activation functions (i.e., tanh, sigmoid etc.) and maximum number of iterations. We also check whether the model is overfitting or not. Once the final model is fixed, we test it on the test data set. We calculate Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) as performance measures to check the accuracy of the implemented model. Performance metrics are defined as,

$$RMSE = \sqrt{\frac{1}{m} \sum_{m=1}^m \frac{\sum_{\mathcal{E}=1}^E (F_m^{\mathcal{E}} - \hat{F}_m^{\mathcal{E}})^2}{E}} \quad (15)$$

$$MAE = \frac{1}{m} \sum_{m=1}^m \frac{\sum_{\mathcal{E}=1}^E |F_m^{\mathcal{E}} - \hat{F}_m^{\mathcal{E}}|}{E} \quad (16)$$

In Table 2, we report the performance of model on the test dataset. From the result, we find that both diffusion convolution and spectral convolution have similar accuracy, which is expected, since the spectral convolution operation is a special case of the diffusion convolution (Li et al., 2018). Based on performance metrics values, we can conclude that the proposed approaches are performing well to capture the flow diffusion. RMSE and MAE values provides aggregated information (average over all the outputs) on the performance of the models, hence, we also estimated R^2 score. As shown in Table 2, for each model the R^2 score is nearly 1, indicating the accuracy of the model to learn traffic flows in the network. We have also compared between actual and estimated link flows for a given OD demand. Fig. 4 and Fig. 5 show that for both Sioux Falls and East Massachusetts networks the difference between actual and estimated link flow is quite low.

Table 2 Model accuracy on the test datasets.

	Network	Flow Propagation Function	Minimum Flow	Maximum Flow	Mean Flow	MAE	RMSE	% Error Over Mean Flow	R ² Score
Uncongested Free Flow Condition	Sioux Falls	Random Walk	642.8	7001.4	2447.4	8.6	11.9	0.350	0.9999
		Laplacian Graph				8.5	11.8	0.346	0.9999
		Spectral Graph				8.6	11.9	0.353	0.9999
	East Massachusetts	Random Walk	0	8220.1	1762.0	25.5	44.6	1.444	0.9991
		Laplacian Graph				25.2	44.5	1.431	0.9991
		Spectral Graph				25.5	44.6	1.449	0.9991
Moderately Congested Condition	Sioux Falls	Random Walk	2562.9	16394.4	6704.5	23.2	31.7	0.345	0.9998
		Laplacian Graph				23.1	31.7	0.345	0.9998
		spectral Graph				23.4	32.1	0.350	0.9998
	East Massachusetts	Random Walk	0	10858.7	2337.9	33.6	59.5	1.436	0.9991
		Laplacian Graph				33.5	59.4	1.431	0.9991
		Spectral Graph				33.6	59.5	1.438	0.9991
Congested Condition	Sioux Falls	Random Walk	4489.6	23437.3	10408.7	36.0	48.4	0.346	0.9998
		Laplacian Graph				35.3	47.4	0.338	0.9998
		Spectral Graph				36.9	49.2	0.355	0.9999
	East Massachusetts	Random Walk	0	12174.8	2698.7	38.8	68.9	1.438	0.9991
		Laplacian Graph				38.7	68.8	1.439	0.9991
		Spectral Graph				39.3	69.1	1.4575	0.9991

From the results, we observe that: (i) a neural network can capture the traffic assignment of a network without any prior knowledge on user behavior; (ii) we can achieve a better accuracy with an appropriate representation of the physical process of flow propagation.

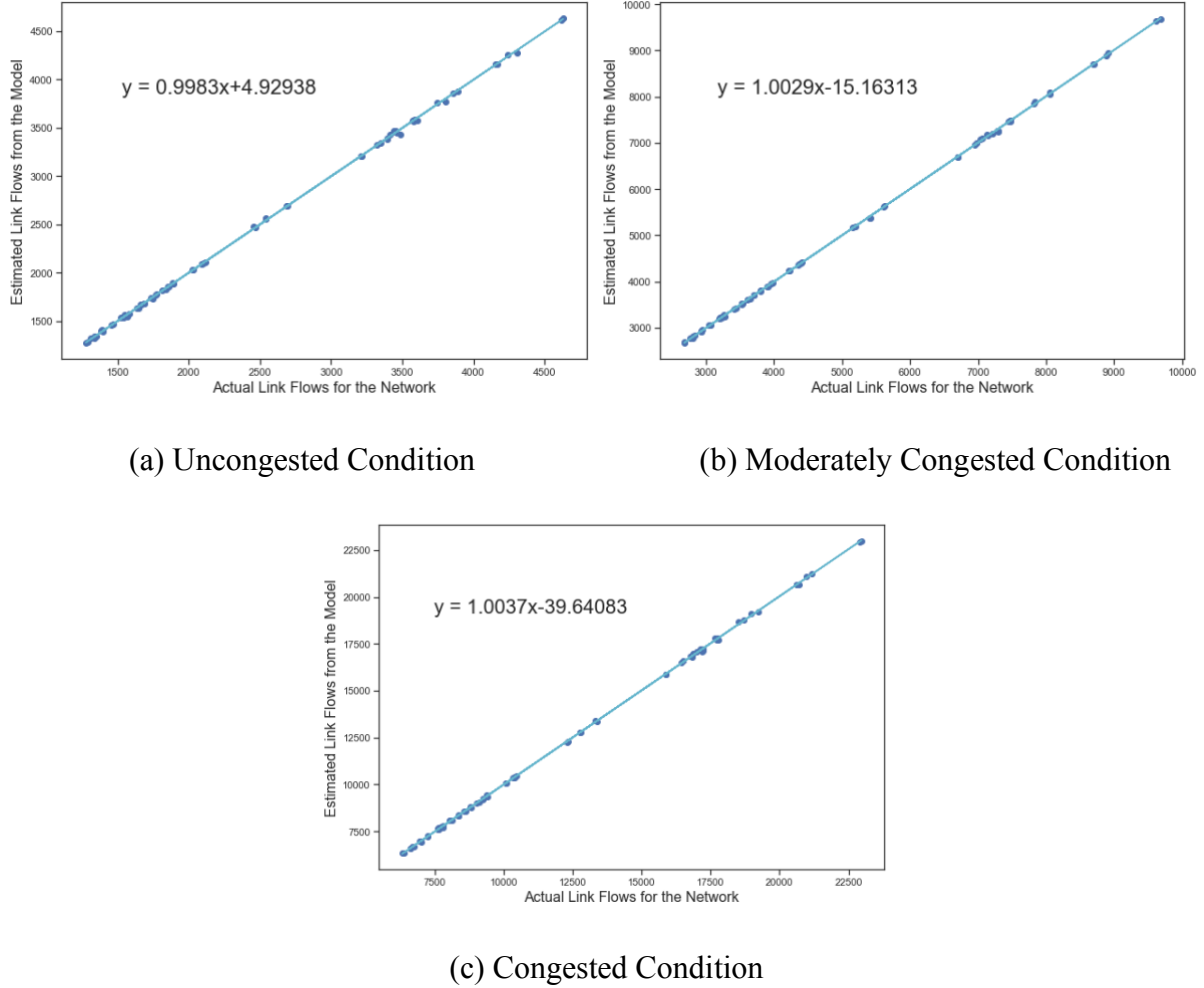
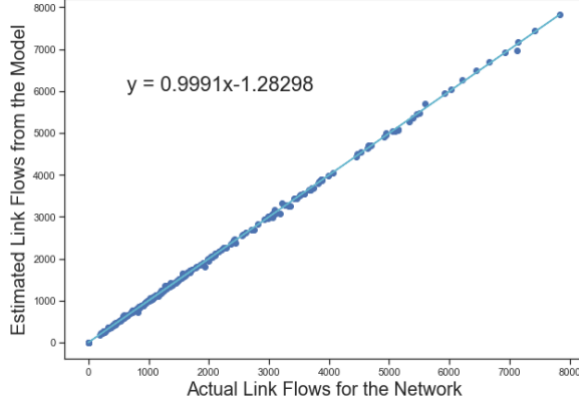
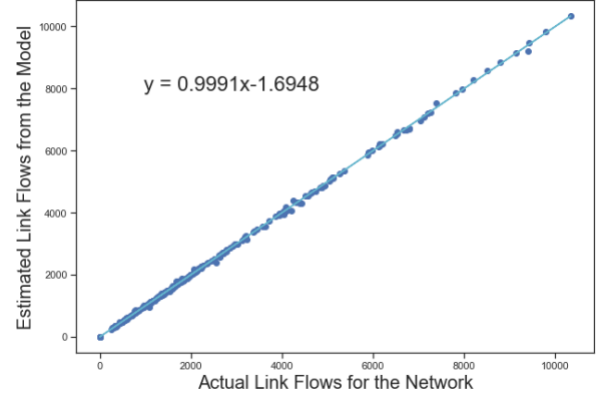


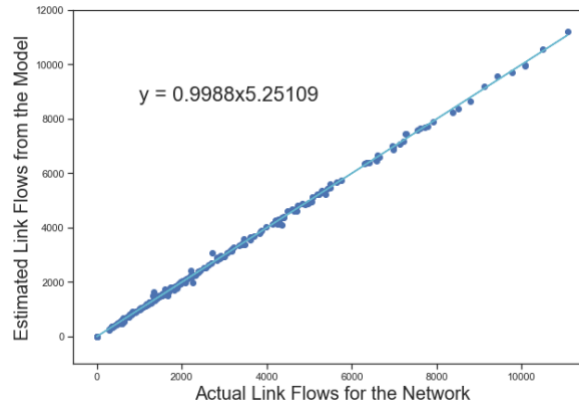
Fig. 4. Comparison between actual and estimated link flows for a given OD demand in Sioux Falls Network



(a) Uncongested Condition



(b) Moderately Congested Condition



(c) Congested Condition

Fig. 5. Comparison between actual and estimated link flows for a given OD demand East Massachusetts network

5.1 Result interpretation

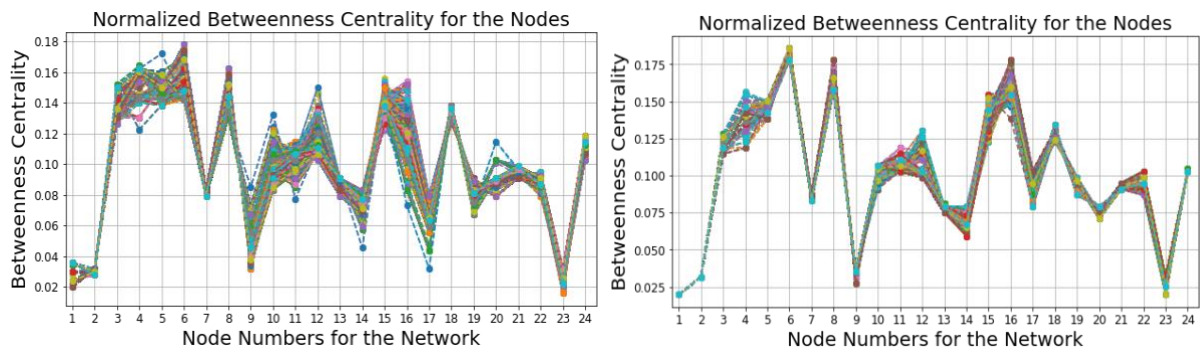
To understand how well the model has learned the flow propagation for the given networks, we perform network topology analysis for the Sioux Falls network based on the betweenness centrality of the nodes and correlate them with the variations in trained parameters. Betweenness centrality of a node indicates the fraction of the total number of shortest paths passing through that node (Brandes, 2001), which means a node with a higher value of betweenness centrality will have a higher number of shortest paths incidence on that particular node.

For a given OD demand matrix, to estimate betweenness centrality values we need to

find the shortest paths for all pairs of origin-destination nodes at user equilibrium. However, the shortest path for a pair of origin and destination nodes depends on link travel times. In our analytical solutions (i.e., using Frank-Wolfe algorithm), we use BPR travel time function (see equation 17) to update link travel times. Based on the updated travel time, we find the shortest paths for assigning traffic to the network. These steps continue iteratively prior to reaching an user equilibrium solution, when travel time for all the used paths remain same for a given O-D pair. Hence, we use the same function to estimate user equilibrium travel time ($t_{i,j}$) for all the links of the network.

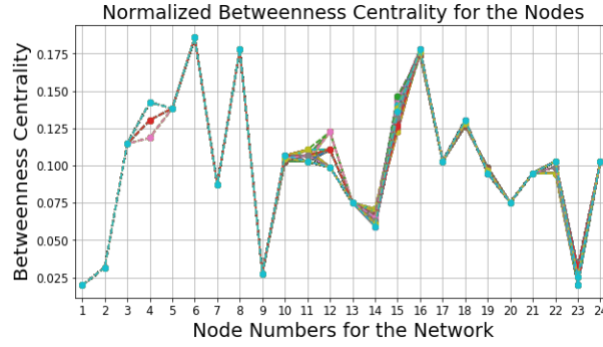
$$t_{i,j} = t_{i,j}^0 \left(1 + 0.15 \left(\frac{f_{i,j}}{c_{i,j}} \right)^4 \right) \quad (17)$$

where $\frac{f_{i,j}}{c_{i,j}}$ indicates the flow-capacity ratio for a given link. Based on the estimated links' travel time from equilibrium link flows we find the shortest paths and estimate the betweenness centrality for all the nodes. We apply this approach on all the training OD demand samples. Fig. 6 shows the distribution of betweenness centrality values for each node of the Sioux Falls Network across different traffic conditions. From the figures we find that except nodes 1, 2, 9, and 23, all the nodes have higher betweenness centrality values. Nodes 6, 8, 12, 15, 16, and 18 are the most critical for the Sioux Falls network. Moreover, we find that for congested condition (i.e., higher demand), variations of betweenness centrality values are higher compared to moderately and uncongested conditions.



(a) Congested Condition

(b) Moderately congested Condition



(c) Uncongested Condition

Fig. 6. Variations of Betweenness Centrality for different nodes at different traffic condition
(Sioux Falls network)

From the BPR function, we find that, if flow capacity ratio is less than 1.0, there will be no significant change in travel time with link flow variations or demand variations. However, when the flow capacity ratio is greater than 1.0, travel time will vary significantly with demand variations (i.e., link flow variations). Consequently, the shortest paths will change abruptly (i.e., shortest path are not stable) leading to significant variations in betweenness centrality for all the nodes. In our case, for congested condition, the flow capacity ratio mostly varies from 1.0 to 2.0, thereby we observe a significant variation of betweenness centrality for different nodes. Whereas for moderately congested and uncongested conditions, since the flow capacity ratio is mostly less than 1.0 so we, do not see significant variations in Betweenness Centrality.

In the proposed model, we assign parameters (\mathbf{W}_q) to learn the flow propagation from nodes into adjacent neighboring links. We assume that the weight parameters associated with critical nodes will be higher and will vary significantly due to the changes in betweenness centrality of nodes. In other words, the weight parameter associated with a node is likely to be correlated with the betweenness centrality value of the node.

In Fig. 7, we plot the weight distributions for node-link flow propagation inside Sioux Falls network at different traffic conditions. Since for uncongested and moderately congested

conditions, the variations of betweenness centrality for different nodes are similar, the distributions of weight parameter (W_q) are also similar. In both cases, the critical nodes 6, 8, 15, and 16 have high positive weights (Fig. 7 (a) & (b)). For congested condition, the variation of betweenness centrality over different nodes are higher, thereby the weight parameters vary significantly compared to uncongested and moderately conditions (Fig. 7(c)). In a congested condition, the model cannot identify the critical nodes from all nodes to pass the traffic efficiently. This could be a possible reason that the model does not give higher positive weights for critical nodes similar to uncongested and moderately congested conditions.

5.2 Stability of the Solution

We train the model using mean squared error as the loss function. At each iteration, the model estimates the mean squared error for the estimated flows ($\hat{F}_m^\mathcal{E}$) and the actual flows ($F_m^\mathcal{E}$) of the network. Afterward, the gradient of the loss function is backpropagated to adjust the weights to reduce loss function value. The loss function can be defined as,

$$L_m = Loss(F_m^\mathcal{E}, \hat{F}_m^\mathcal{E}) \quad (18)$$

where, $Loss(.)$ is the function to estimate the error between the actual ($F_m^\mathcal{E}$) and estimated values ($\hat{F}_m^\mathcal{E}$) and \mathcal{E} denotes the set of links for the network. In this study, we estimate mean square error (MSE) as a loss function.

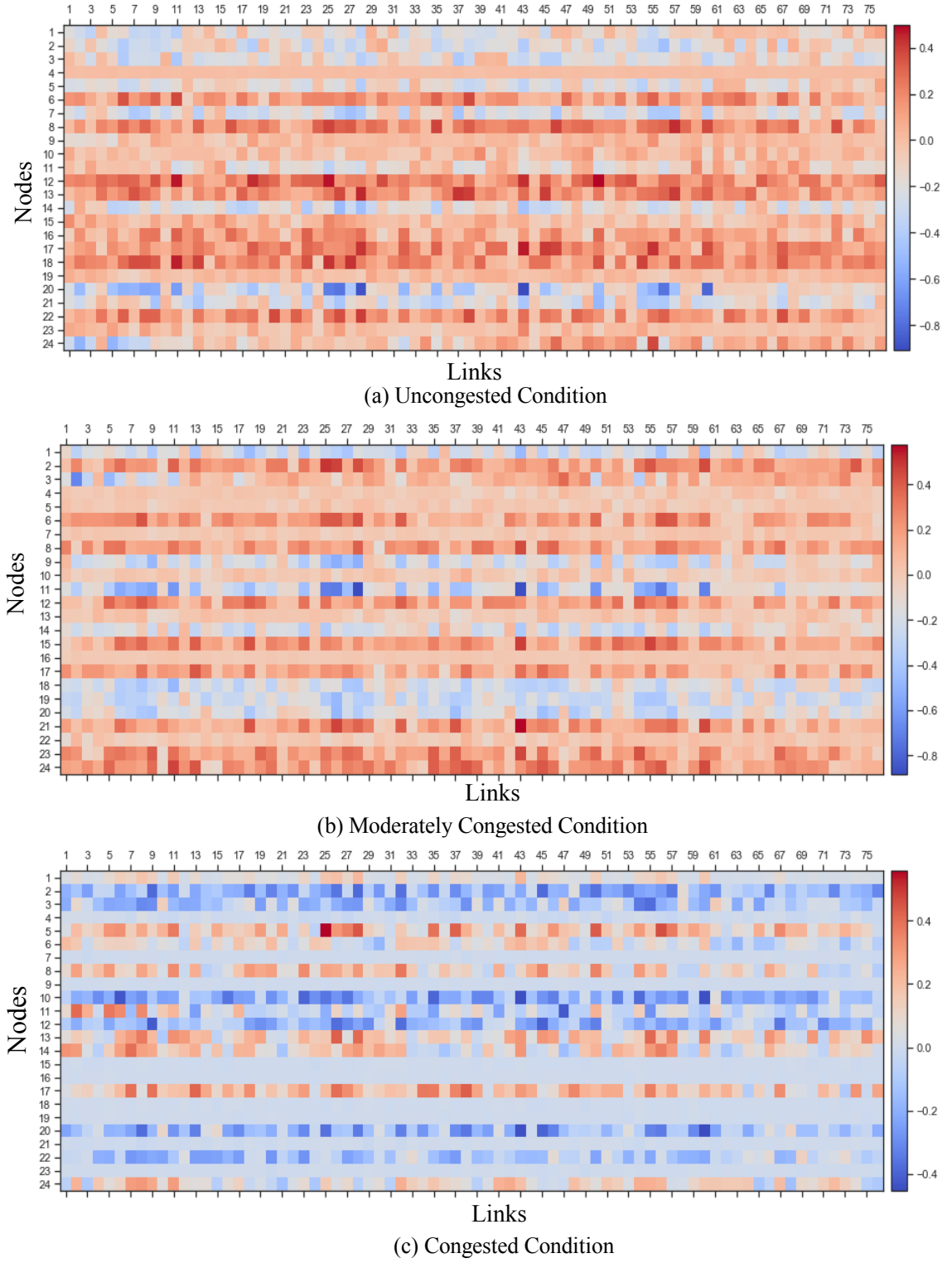
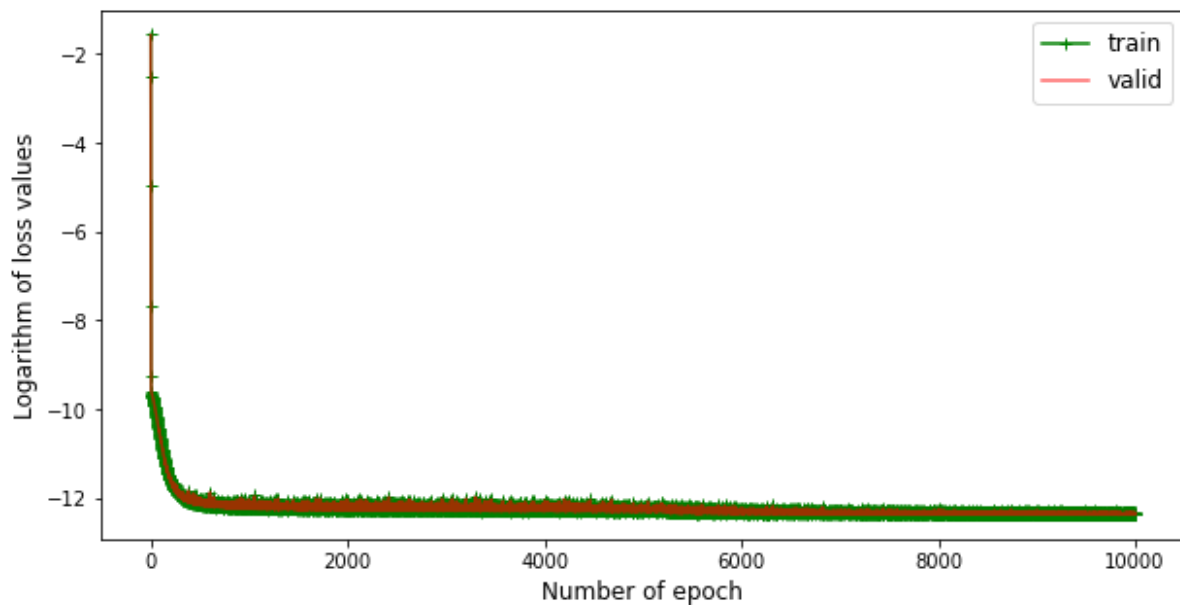


Fig. 7. Distribution of weight (w_q) over different traffic conditions (Sioux Falls network)

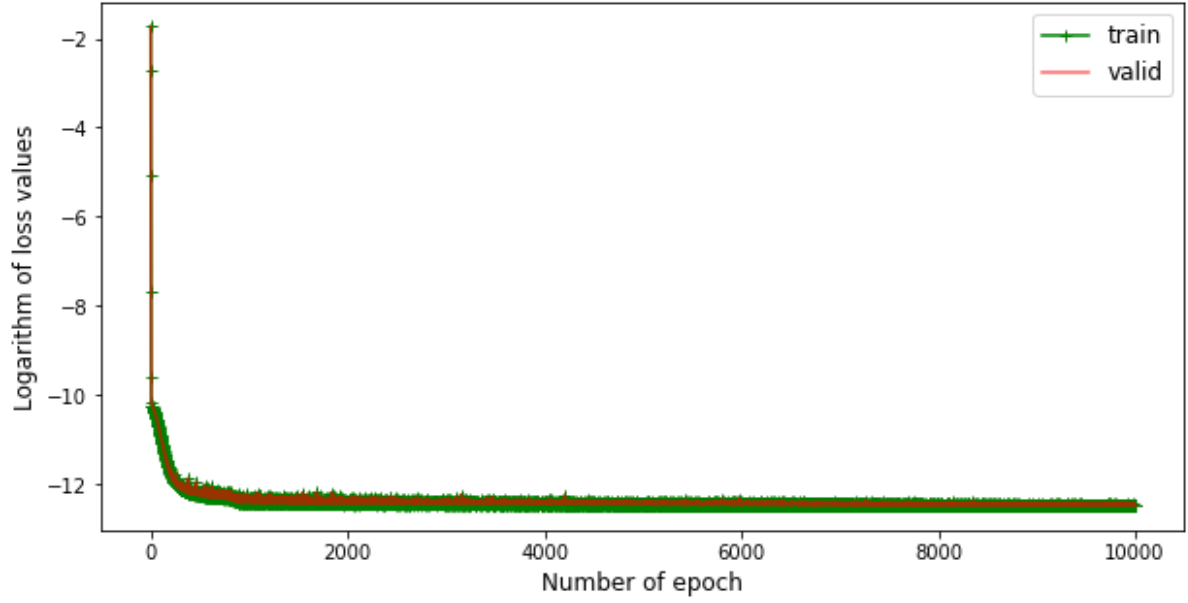
To check the stability of solution, we observed the training and test loss values for the model (Fig. 8). We train each model for 10,000 iterations to check variation of train and

validation loss values. We find that it takes about 2000 iterations for the model to converge to a stable solution, after that there are merely any variations in loss values. Moreover, after 10,000 iterations the loss function value for the validation data gradually start increasing. From 2,000 to 10,000 iterations the difference in MSE values varies from 4 to 9, indicating a stable solution with minimal variance. We experimented with two different optimizers to train the model: root mean square propagation (RMSProp) and adaptive moment estimation (ADAM) optimizer. Among these two, RMSProp takes less iterations (~ 1500 iterations) to converge (i.e. similar train and validation error) compared to ADAM optimizer (~ 2000 iterations). However, ADAM optimizer gives more stable solutions, which means MSE values for train and validation data almost remain same after convergence (i.e., after 2000 iterations). Whereas, for RMSProp optimizer we observe slight variations in MSE values for both train and validation data samples even after convergence (i.e., after 1500 iterations).

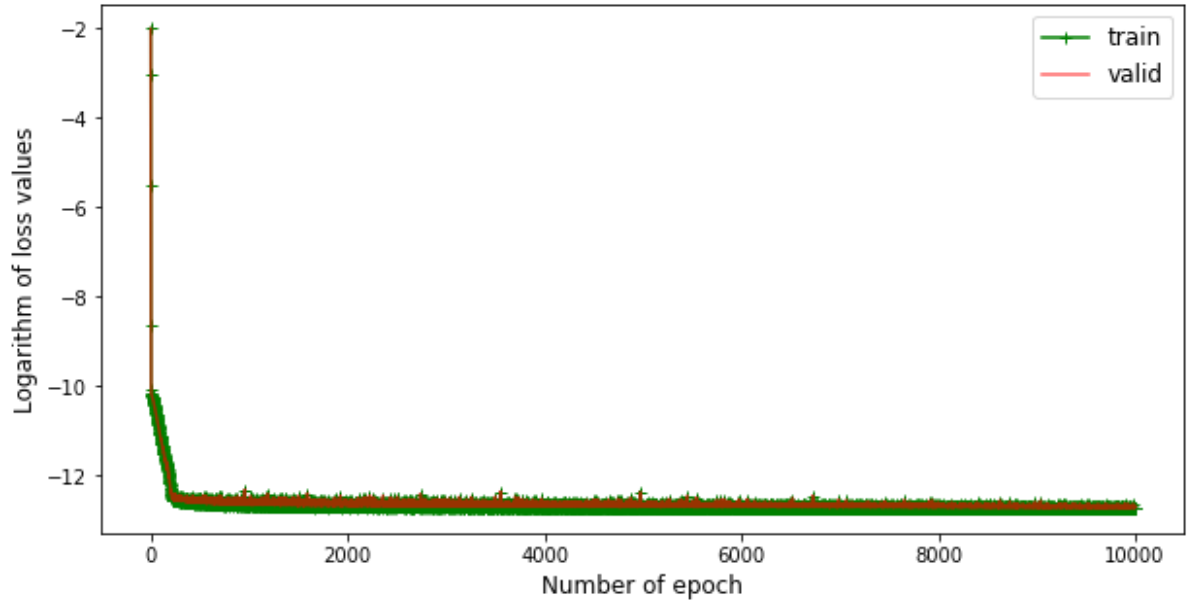
We also check the computation time required to train the models. It takes about 19 minutes to train the models on Sioux Falls network for 10,000 iterations, while for East Massachusetts network it takes 30 minutes. So, our approach performs reasonably well to estimate network level traffic flows with less computation time.



(a) Congested Condition



(b) Moderately Congested Condition



(c) Uncongested Condition

Fig. 8. Loss Function Values for Laplacian Spectral Graph (Sioux Falls network)

6. CONCLUSIONS

In this study, we present a data-driven formulation of traffic assignment problem based on

learning traffic flow patterns of a transportation network from origin destination (OD) travel demand variations. Adapting graph convolution approach, we develop a deep learning architecture to solve this traffic assignment problem by capturing the diffusion of OD demand inside the network. To efficiently represent the diffusion process of multiple OD demands from nodes to neighboring links, we customize the traditional graph convolutional neural network and introduce the concept of learning-based assignment (i.e., routing matrix) of OD demands to generate link flows. Finally, we provide experimental evidence on the validity of the approach by training the model to learn the user equilibrium traffic flows for Sioux Falls and East Massachusetts networks. The experiment results show that the implemented GCNN model can capture the user equilibrium traffic flow of the network very well with less than 2% mean absolute difference between the actual and estimated link flows under varying congested conditions. Moreover, when the training of the model is complete, it can instantaneously determine the traffic flows of a large-scale network. Hence this approach can overcome the challenges of deploying traffic assignment models over large-scale networks. Furthermore, this method is completely data-driven without requiring any assumption on user behavior. Thus, it will improve the reliability and stability of traffic assignment solutions.

To extend this framework towards a data-driven dynamic traffic assignment method, we need to consider the representation of the physical process of flow propagation to account travel time variations. In addition, existing data sources are not widely available to researchers to infer travel demand at a higher spatiotemporal resolution. Alternative data augmentation approach can be explored to prepare the travel demand data. Future research should consider these issues to develop a framework that can solve the dynamic traffic assignment problem and prepare demand data that can be fed into such frameworks.

A data-driven network modeling approach is warranted in the era of big data. Ubiquitous use of mobile phones, availability of GPS based vehicle trajectory data, emerging

connected and automated vehicle data, and so on will give us the opportunities to model travel demand and traffic flows at a high spatio-temporal resolution. The proposed deep learning architecture for solving a traffic assignment learning problem is an initial step towards exploiting such high-fidelity data for data-driven network modeling research.

ACKNOWLEDGMENT

This study was supported by the U.S. National Science Foundation through the grant CMMI #1917019. However, the authors are solely responsible for the facts and accuracy of the information presented in the paper.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

- Abdelfatah, A.S., Mahmassani, H.S., 2002. A simulation-based signal optimization algorithm within a dynamic traffic assignment framework 428–433. <https://doi.org/10.1109/itsc.2001.948695>
- Alexander, L., Jiang, S., Murga, M., González, M.C., 2015. Origin-destination trips by purpose and time of day inferred from mobile phone data. *Transp Res Part C Emerg Technol* 58, 240–250. <https://doi.org/10.1016/j.trc.2015.02.018>
- Atwood, J., Towsley, D., 2015. Diffusion-Convolutional Neural Networks. *Adv Neural Inf Process Syst* 2001–2009.
- Ban, X. (Jeff), Liu, H.X., Ferris, M.C., Ran, B., 2008. A link-node complementarity model and solution algorithm for dynamic user equilibria with exact flow propagations. *Transportation Research Part B: Methodological* 42, 823–842. <https://doi.org/10.1016/j.trb.2008.01.006>
- Bar-Gera, H., 2002. Origin-based algorithm for the traffic assignment problem. *Transportation Science* 36, 398–417. <https://doi.org/10.1287/trsc.36.4.398.549>
- Barthélemy, J., Carletti, T., 2017. A dynamic behavioural traffic assignment model with strategic agents. *Transp Res Part C Emerg Technol* 85, 23–46. <https://doi.org/10.1016/j.trc.2017.09.004>
- Ben-Akiva, M.E., Gao, S., Wei, Z., Wen, Y., 2012. A dynamic traffic assignment model for highly congested urban networks. *Transp Res Part C Emerg Technol* 24, 62–82. <https://doi.org/10.1016/j.trc.2012.02.006>
- Billings, D., Jiann-Shiou, Y., 2006. Application of the ARIMA Models to Urban Roadway Travel Time Prediction-A Case Study. *Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference on* 2529–2534.
- Boyles, S., Ukkusuri, S. V., Waller, S.T., Kockelman, K.M., 2006. A comparison of static and dynamic traffic assignment under tolls: A study of the dallas-fort worth network. *85th Annual Meeting of ...* 14.
- Brandes, U., 2001. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology* 25, 163–177. <https://doi.org/10.1080/0022250X.2001.9990249>
- Cai, P., Wang, Y., Lu, G., Chen, P., Ding, C., Sun, J., 2016. A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. *Transp Res Part C Emerg Technol* 62,

- 21–34. <https://doi.org/10.1016/j.trc.2015.11.002>
- Chien, S.I.-J., Kuchipudi, C.M., 2003. Dynamic Travel Time Prediction with Real-Time and Historic Data. *J Transp Eng* 129, 608–616. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2003\)129:6\(608\)](https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(608))
- Cui, Z., Henrickson, K., Ke, R., Wang, Y., 2018a. Traffic Graph Convolutional Recurrent Neural Network: A Deep Learning Framework for Network-Scale Traffic Learning and Forecasting. *IEEE Transactions on Intelligent Transportation Systems* 1–12. <https://doi.org/10.1109/TITS.2019.2950416>
- Cui, Z., Ke, R., Pu, Z., Wang, Y., 2018b. Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction. *International Workshop on Urban Computing (UrbComp)* 2017.
- Cui, Z., Lin, L., Pu, Z., Wang, Y., 2020. Graph Markov network for traffic forecasting with missing data. *Transp Res Part C Emerg Technol* 117, 102671. <https://doi.org/10.1016/j.trc.2020.102671>
- Defferrard, M., Bresson, X., Vandergheynst, P., 2016. Convolutional neural networks on graphs with fast localized spectral filtering. *Adv Neural Inf Process Syst* 3844–3852.
- Deshpande, M., Bajaj, P.R., 2016. Performance analysis of support vector machine for traffic flow prediction. *2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)* 126–129. <https://doi.org/10.1109/ICGTSPICC.2016.7955283>
- Elhenawy, M., Chen, H., Rakha, H.A., 2014. Dynamic travel time prediction using data clustering and genetic programming. *Transp Res Part C Emerg Technol* 42, 82–98. <https://doi.org/10.1016/j.trc.2014.02.016>
- Foytik, P., Jordan, C., Robinson, R.M., 2017. Exploring simulation based dynamic traffic assignment with large scale microscopic traffic simulation model.
- Friesz, T.L., Luque, J., Tobin, R.L., Wie, B.W., 1989. Dynamic Network Traffic Assignment Considered as a Continuous Time Optimal Control Problem. *Oper Res* 37, 893–901. <https://doi.org/10.2307/171471>
- Gundlegård, D., Rydergren, C., Breyer, N., Rajna, B., 2016. Travel demand estimation and network assignment based on cellular network data. *Comput Commun* 95, 29–42. <https://doi.org/10.1016/j.comcom.2016.04.015>
- Guo, K., Hu, Y., Qian, Z., Liu, H., Zhang, K., Sun, Y., Gao, J., Yin, B., 2021. Optimized Graph Convolution Recurrent Neural Network for Traffic Prediction. *IEEE Transactions on Intelligent Transportation Systems* 22, 1138–1149. <https://doi.org/10.1109/TITS.2019.2963722>
- Guo, K., Hu, Y., Qian, Z., Sun, Y., Gao, J., Yin, B., 2020. Dynamic Graph Convolution Network for Traffic Forecasting Based on Latent Network of Laplace Matrix Estimation. *IEEE Transactions on Intelligent Transportation Systems* 23, 1009–1018. <https://doi.org/10.1109/tits.2020.3019497>
- Guo, S., Lin, Y., Feng, N., Song, C., Wan, H., 2019. Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 922–929. <https://doi.org/10.1609/aaai.v33i01.3301922>
- Hammond, D.K., Vandergheynst, P., Gribonval, R., 2011. Wavelets on graphs via spectral graph theory. *Appl Comput Harmon Anal* 30, 129–150. <https://doi.org/10.1016/j.acha.2010.04.005>
- He, X., Guo, X., Liu, H.X., 2010. A link-based day-to-day traffic assignment model. *Transportation Research Part B* 44, 597–608. <https://doi.org/10.1016/j.trb.2009.10.001>
- Huang, Y., Kockelman, K.M., 2019. Electric vehicle charging station locations: Elastic demand, station congestion, and network equilibrium. *Transp Res D Transp Environ* 1–16. <https://doi.org/10.1016/j.trd.2019.11.008>
- Innamaa, S., 2005. Short-term prediction of travel time using neural networks on an interurban highway. *Transportation (Amst)* 32, 649–669. <https://doi.org/10.1007/s11116-005-0219-y>
- Jafari, E., Pandey, V., Boyles, S.D., 2017. A decomposition approach to the static traffic assignment problem. *Transportation Research Part B: Methodological* 105, 270–296. <https://doi.org/10.1016/j.trb.2017.09.011>
- Janson, B.N., 1989. Dynamic Traffic Assignment for Urban Road Networks. *Transportation Research Part B-Methodological* 25, 143–161.
- Ji, X., Shao, C., Wang, B., 2016. Stochastic Dynamic Traffic Assignment Model under Emergent Incidents. *Procedia Eng* 137, 620–629. <https://doi.org/10.1016/j.proeng.2016.01.299>
- Jiang, Y., Wong, S.C., Ho, H.W., Zhang, P., Liu, R., Sumalee, A., 2011a. A dynamic traffic assignment

- model for a continuum transportation system. *Transportation Research Part B: Methodological* 45, 343–363. <https://doi.org/10.1016/j.trb.2010.07.003>
- Jiang, Y., Wong, S.C., Ho, H.W., Zhang, P., Liu, R., Sumalee, A., 2011b. A dynamic traffic assignment model for a continuum transportation system. *Transportation Research Part B: Methodological* 45, 343–363. <https://doi.org/10.1016/j.trb.2010.07.003>
- Kim, H., Oh, J.S., Jayakrishnan, R., 2009. Effects of user equilibrium assumptions on network traffic pattern. *KSCE Journal of Civil Engineering* 13, 117–127. <https://doi.org/10.1007/s12205-009-0117-5>
- Kim, T.S., Lee, W.K., Sohn, S.Y., 2019. Graph convolutional network approach applied to predict hourly bike-sharing demands considering spatial, temporal, and global effects. *PLoS One* 14, e0220782. <https://doi.org/10.1371/journal.pone.0220782>
- Kipf, T.N., Welling, M., 2016. Semi-Supervised Classification with Graph Convolutional Networks. 5th International Conference on Learning Representations, ICLR 2017- Conference Track Proceedings.
- LeBlanc, L.J., Morlok, E.K., Pierskalla, W.P., 1975. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research* 9, 309–318. [https://doi.org/10.1016/0041-1647\(75\)90030-1](https://doi.org/10.1016/0041-1647(75)90030-1)
- Lee, Y.L.Y., 2009. Freeway travel time forecast using artificial neural networks with cluster method. 2009 12th International Conference on Information Fusion 1331–1338.
- Leon-Garcia, A., Tizghadam, A., 2009. A graph theoretical approach to traffic engineering and network control problem. 21st International Teletraffic Congress, ITC 21: Traffic and Performance Issues in Networks of the Future - Final Programme.
- Leurent, F., Chandakas, E., Poulhès, A., 2011. User and service equilibrium in a structural model of traffic assignment to a transit network. *Procedia Soc Behav Sci* 20, 495–505. <https://doi.org/10.1016/j.sbspro.2011.08.056>
- Li, G., Knoop, V.L., van Lint, H., 2021. Multistep traffic forecasting by dynamic graph convolution: Interpretations of real-time spatial correlations. *Transp Res Part C Emerg Technol* 128, 103185. <https://doi.org/10.1016/j.trc.2021.103185>
- Li, Y., Yu, R., Shahabi, C., Liu, Y., 2018. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. 6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings 1–16.
- Lin, L., He, Z., Peeta, S., 2018. Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach 97, 258–276. <https://doi.org/10.1016/j.trc.2018.10.011>
- Liu, H.X., Ban, X., Ran, B., Mirchandani, P., 2007. Analytical Dynamic Traffic Assignment Model with Probabilistic Travel Times and Perceptions. *Transportation Research Record: Journal of the Transportation Research Board* 1783, 125–133. <https://doi.org/10.3141/1783-16>
- Liu, Y., Zheng, H., Feng, X., Chen, Z., 2017. Short-term traffic flow prediction with Conv-LSTM. 2017 9th International Conference on Wireless Communications and Signal Processing, WCSP 2017 - Proceedings. <https://doi.org/10.1109/WCSP.2017.8171119>
- Lo, H.K., Szeto, W.Y., 2002. A cell-based variational inequality formulation of the dynamic user optimal assignment problem. *Transportation Research Part B: Methodological* 36, 421–443. [https://doi.org/10.1016/S0191-2615\(01\)00011-X](https://doi.org/10.1016/S0191-2615(01)00011-X)
- Luo, X., Li, D., Yang, Y., Zhang, S., 2019. Spatiotemporal Traffic Flow Prediction with KNN and LSTM 2019.
- Ma, X., Dai, Z., He, Z., Ma, J., Wang, Yong, Wang, Yunpeng, 2017. Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction. *Sensors* 17, 818. <https://doi.org/10.3390/s17040818>
- Ma, X., Tao, Z., Wang, Yinhai, Yu, H., Wang, Yunpeng, 2015. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transp Res Part C Emerg Technol* 54, 187–197. <https://doi.org/10.1016/j.trc.2015.03.014>
- Mahmassani, HaniS., 2001. Dynamic Network Traffic Assignment and Simulation Methodology for Advanced System Management Applications. *Netw Spat Econ* 1, 267–292. <https://doi.org/10.1023/A:1012831808926>
- Merchant, D.K., Nemhauser, G.L., 1978. A Model and an Algorithm for the Dynamic Traffic

- Assignment Problems. *Transportation Science* 12, 183–199. <https://doi.org/10.1287/trsc.12.3.183>
- Mitradjieva, M., Lindberg, P.O., 2013. The Stiff Is Moving—Conjugate Direction Frank-Wolfe Methods with Applications to Traffic Assignment *. *Transportation Science* 47, 280–293. <https://doi.org/10.1287/trsc.1120.0409>
- Myung, J., Kim, D.-K., Kho, S.-Y., Park, C.-H., 2011. Travel Time Prediction Using k Nearest Neighbor Method with Combined Data from Vehicle Detector System and Automatic Toll Collection System. *Transportation Research Record: Journal of the Transportation Research Board* 2256, 51–59. <https://doi.org/10.3141/2256-07>
- Nie, Y.M., Zhang, H.M., 2010. Solving the dynamic user optimal assignment problem considering queue spillback. *Netw Spat Econ* 10, 49–71. <https://doi.org/10.1007/s11067-007-9022-y>
- Peeta, S., Mahmassani, H.S., 1995. System optimal and user equilibrium time-dependent traffic assignment in congested networks. <https://doi.org/10.1007/BF02031941>
- Peeta, S., Ziliaskopoulos, A.K., 2001. Foundations of Dynamic Traffic Assignment : The Past , the Present and the Future. *Network and Spatial Economics* 1, 233–265.
- Peng, H., Du, B., Liu, Mingsheng, Liu, Mingzhe, Ji, S., Wang, S., Zhang, X., He, L., 2021. Dynamic graph convolutional network for long-term traffic flow prediction with reinforcement learning. *Inf Sci (N Y)* 578, 401–416. <https://doi.org/10.1016/j.ins.2021.07.007>
- Peng, H., Wang, H., Du, B., Bhuiyan, M.Z.A., Ma, H., Liu, J., Wang, L., Yang, Z., Du, L., Wang, S., Yu, P.S., 2020. Spatial temporal incidence dynamic graph neural networks for traffic flow forecasting. *Inf Sci (N Y)* 521, 277–290. <https://doi.org/10.1016/j.ins.2020.01.043>
- Polson, N.G., Sokolov, V.O., 2017. Deep learning for short-term traffic flow prediction. *Transp Res Part C Emerg Technol* 79, 1–17. <https://doi.org/10.1016/j.trc.2017.02.024>
- Primer, a, 2011. Dynamic Traffic Assignment. *Transportation Network Modeling Committee* 1–39. <https://doi.org/10.1016/j.trd.2016.06.003>
- PyTorch [WWW Document], 2016. URL <https://pytorch.org/>
- Rahman, R., Hasan, S., 2018. Short-Term Traffic Speed Prediction for Freeways During Hurricane Evacuation : A Deep Learning Approach 1291–1296. <https://doi.org/10.1109/ITSC.2018.8569443>
- Ran, B.I.N., Boyce, D.E., Leblanc, L.J., 1993. A New Class of Instantaneous Dynamic User-Optimal Traffic Assignment Models Author (s): Bin Ran , David E . Boyce and Larry J . Leblanc and Dynamic Models in Transportation (Jan . - Feb . , 1993), pp . 192-202 Published by : INFORMS Stable URL : <http://41>, 192–202.
- Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., Battaglia, P.W., 2020. Learning to simulate complex physics with graph networks. 37th International Conference on Machine Learning, ICML 2020 PartF16814, 8428–8437.
- Shafiei, S., Gu, Z., Saberi, M., 2018. Calibration and validation of a simulation-based dynamic traffic assignment model for a large-scale congested network. *Simul Model Pract Theory* 86, 169–186. <https://doi.org/10.1016/j.simpat.2018.04.006>
- Sheffi, Y., 1985. Urban transportation networks, PRENTICE-HALL, Inc. Prentice-Hall, Inc., Englewood Cliffs. [https://doi.org/10.1016/0191-2607\(86\)90023-3](https://doi.org/10.1016/0191-2607(86)90023-3)
- Tang, C., Sun, J., Sun, Y., Peng, M., Gan, N., 2020. A General traffic flow prediction approach based on spatial-Temporal graph attention. *IEEE Access* 8, 153731–153741. <https://doi.org/10.1109/ACCESS.2020.3018452>
- Teng, S.H., 2016. Scalable algorithms for data and network analysis. *Foundations and Trends in Theoretical Computer Science* 12, 1–274. <https://doi.org/10.1561/04000000051>
- Tizghadam, A., Leon-garcia, A., 2007. A Robust Routing Plan to Optimize Throughput in Core Networks 117–128.
- Transportation Networks for Research Core Team, 2016. Transportation Networks for Research [WWW Document]. 2016. URL <https://github.com/bstabler/TransportationNetworks> (accessed 8.7.18).
- Ukkusuri, S. V., Han, L., Doan, K., 2012. Dynamic user equilibrium with a path based cell transmission model for general traffic networks. *Transportation Research Part B: Methodological* 46, 1657–1684. <https://doi.org/10.1016/j.trb.2012.07.010>
- Waller, S.T., Fajardo, D., Duell, M., Dixit, V., 2013. Linear Programming Formulation for Strategic Dynamic Traffic Assignment. *Netw Spat Econ* 13, 427–443. <https://doi.org/10.1007/s11067-013-9187-5>
- Wang, H.W., Peng, Z.R., Wang, D., Meng, Y., Wu, T., Sun, W., Lu, Q.C., 2020. Evaluation and

- prediction of transportation resilience under extreme weather events: A diffusion graph convolutional approach. *Transp Res Part C Emerg Technol* 115, 102619. <https://doi.org/10.1016/j.trc.2020.102619>
- Wardrop, J.G., 1952. Some Theoretical Aspects of Road Traffic Research.
- Watling, D., Hazelton, M.L., 2003. The dynamics and equilibria of day-to-day assignment models. *Netw Spat Econ* 349–370. <https://doi.org/10.1023/A:1025398302560>
- Wu, C.H., Ho, J.M., Lee, D.T., 2004. Travel-time prediction with support vector regression. *IEEE Transactions on Intelligent Transportation Systems* 5, 276–281. <https://doi.org/10.1109/TITS.2004.837813>
- Yasdi, R., 1999. Prediction of Road Traffic using a Neural Network Approach. *Neural Comput Appl* 8, 135–142. <https://doi.org/10.1007/s005210050015>
- Yu, B., Song, X., Guan, F., Yang, Z., Yao, B., 2016. k-Nearest Neighbor Model for Multiple-Time-Step Prediction of Short-Term Traffic Condition. *J Transp Eng* 142, 4016018. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000816](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000816)
- Yu, B., Yin, H., Zhu, Z., 2018. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting, in: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, California*, pp. 3634–3640. <https://doi.org/10.24963/ijcai.2018/505>
- Zhang, S., Tong, H., Xu, J., Maciejewski, R., 2019. Graph convolutional networks: a comprehensive review. *Comput Soc Netw* 6. <https://doi.org/10.1186/s40649-019-0069-y>
- Zhang, Y., Haghani, A., 2015. A gradient boosting method to improve travel time prediction. *Transp Res Part C Emerg Technol* 58, 308–324. <https://doi.org/10.1016/j.trc.2015.02.019>
- Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H., 2020. T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Transactions on Intelligent Transportation Systems* 21, 3848–3858. <https://doi.org/10.1109/TITS.2019.2935152>
- Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M., 2018. Graph Neural Networks: A Review of Methods and Applications 1–20.

Appendix A. Modeling Traffic Flows using Spectral Graph Convolution

In spectral graph convolution, a spectral convolutional filter is used learn traffic flow patterns inside a transportation network in repose to travel demand variations. The spectral filter is derived from spectrum of the Laplacian matrix, which consist of eigenvalues of the Laplacian matrix. So, to construct the spectrum, we must calculate the eigenvalues of the Laplacian matrix. For a symmetric graph, we can compute the eigenvalues using eigen decomposition of the Laplacian matrix. In this problem, we consider the transportation network as a symmetric directed graph; same number of links getting out and getting inside a node, which means the in degree and out degree matrix of the graph is similar. Thus, the Laplacian matrix of this graph is diagonalizable as follows using eigen decomposition,

$$L_w = U\Lambda U^T \quad (19)$$

where, Λ is a diagonal matrix with eigenvalues, $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_N$ and U indicates the eigen vectors, $u_0, u_1, u_2, \dots, u_N$. Eigen values represent characteristics of transportation network in terms of strength of a particular node based on its position, distance between adjacent nodes, dimension of the network. The spectral graph convolution filter can be defined as follows,

$$g_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k \Lambda^k \quad (20)$$

where, θ is the parameter for the convolution filter shared by all the nodes of the network and K is the size of the convolution filter. Now the spectral graph convolution over the graph signal (X) is defined as follows,

$$\begin{aligned} g_\theta * X &= g_\theta(L_w)X = g_\theta(U\Lambda U^T)X = U g_\theta(\Lambda) U^T X = \sum_{k=0}^{K-1} \theta_k U \Lambda^k U^T X \\ g_\theta * X &= \sum_{k=0}^{K-1} \theta_k L_w^k X \end{aligned} \quad (21)$$

According to spectral graph theory, the shortest path distance i.e. minimum number of links

connecting nodes i and j is longer than K , such that $L^K(i, j) = 0$ (Hammond et al., 2011). Consequently, for a given pair of origin (i) and destination (j) nodes, a spectral graph filter of size K has access to all the nodes on the shortest path of the graph. It means that the spectral graph convolution filter of size K captures flow propagation through each node on the shortest path. So, the spectral graph convolution operation can model the interdependency between a link and its i th order adjacent nodes on the shortest paths, given that $0 \leq i \leq K$.

The computational complexity of calculating L_w^k is high due to K times multiplication of L_w . A way to overcome this challenge is to approximate the spectral filter g_θ with Chebyshev polynomials up to $(K - 1)$ th order (Hammond et al., 2011). Defferrard et al. (Defferrard et al., 2016) applied this approach to build a K -localized ChebNet, where the convolution is defined as,

$$g_\theta * X \approx \sum_{k=0}^{K-1} \theta_k T_k(\bar{L})X \quad (22)$$

in which, $\bar{L} = 2L_{sym}/\lambda_{max} - I$. \bar{L} represents a scaling of graph Laplacian that maps the eigenvalues from $[0, \lambda_{max}]$ to $[-1, 1]$. L_{sym} is defined as symmetric normalization of the Laplacian matrix $D_w^{-1/2}L_w D_w^{-1/2}$. T_k and θ denote the Chebyshev polynomials and Chebyshev coefficients. The Chebyshev polynomials are defined recursively by $T_k(\bar{L}) = 2\bar{L}T_{k-1}(\bar{L}) - T_{k-2}(\bar{L})$ with $T_0(\bar{L}) = 1$ and $T_1(\bar{L}) = \bar{L}$. These are the basis of Chebyshev polynomials. Kipf and Welling (Kipf and Welling, 2016) simplified this model by approximating the largest eigenvalue λ_{max} of \bar{L} as 2. In this way, the convolution becomes,

$$g_\theta * X = \theta_0 X - \theta_1 D_w^{-1/2} A_w D_w^{-1/2} X \quad (23)$$

where, Chebyshev coefficient, $\theta = \theta_0 = -\theta_1$. All the detail about the assumptions and their implications of Chebyshev polynomial can be found in (Hammond et al., 2011). Now the simplified graph convolution can be written as follows,

$$g_\theta * X = \theta(I + D_w^{-1/2} A_w D_w^{-1/2})X \quad (24)$$

Since $I + D_w^{-1/2} A_w D_w^{-1/2}$ has eigenvalues in the range $[0, 2]$, it may lead to exploding or

vanishing gradients when used in a deep neural network model. To alleviate this problem, Kipf et al. (Kipf and Welling, 2016) use a renormalization trick by replacing the term $\mathbf{I} + \mathbf{D}_w^{-1/2} \mathbf{A}_w \mathbf{D}_w^{-1/2}$ with $\bar{\mathbf{D}}_w^{-1/2} \bar{\mathbf{A}}_w \bar{\mathbf{D}}_w^{-1/2}$, with $\bar{\mathbf{A}}_w = \mathbf{A}_w + \mathbf{I}$, similar to adding a self-loop. Now, we can simplify the spectral graph convolution as follows,

$$\mathbf{g}_\theta * \mathbf{X} = \boldsymbol{\Theta} \left(\bar{\mathbf{D}}_w^{-1/2} \bar{\mathbf{A}}_w \bar{\mathbf{D}}_w^{-1/2} \right) \mathbf{X} \quad (25)$$

here, $\boldsymbol{\Theta} \in \mathbf{R}^{N \times N}$ indicates the parameters of the convolution filter to be learnt during training process. From Equation 21, we can observe that spectral graph convolution is a special case of diffusion convolution (Li et al., 2018), only difference is that in spectral convolution we symmetrically normalized the adjacency matrix.