A Data Protection Method for Short-Term Traffic Prediction with Applications to Network Active Mode Operations

Wen, Xiamei; Krishnakumari, Panchamy; Hoogendoorn, Serge

DOI
10.1109/ITSC57777.2023.10422693

Publication date
2023

Document Version
Final published version

Published in
2023 IEEE 26th International Conference on Intelligent Transportation Systems, ITSC 2023

2023 IEEE 26th International Conference on
Intelligent Transportation Systems (ITSC)
24-28 September 2023. Bilbao, Bizkaia, Spain

A Data Protection Method for Short-term Traffic Prediction with Applications to Network Active Mode Operations

Xiamei Wen[1], Panchamy Krishnakumari[1], Serge Hoogendoorn[1]

2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC) | 979-8-3503-9946-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ITSC57777.2023.10422693

Abstract— **Accurate prediction of active mode traffic is imperative for optimizing traffic operations in Intelligent Transportation Systems. However, existing data-driven approaches heavily rely on extensive datasets to achieve reliable traffic prediction. This dependence poses a challenge when it comes to data sharing, particularly when collecting information from multiple local clients, such as institutions, organizations, and mobile devices, and transmitting it to a central server for model training and application. To overcome this challenge and enhance data security, we introduce the FedASTGNN model for active mode traffic prediction. This approach combines the federated averaging (FedAvg) algorithm with an attention-based spatial-temporal graph neural network (ASTGNN) model. Subsequently, we conduct an evaluation to determine the performance gap between the centralized ASTGNN model and the proposed distributed FedASTGNN model. This evaluation takes into account the model's performance across different time aggregation intervals and prediction horizons. Moreover, considering the unique attributes and intricacies of active mode data, we create three scenarios to demonstrate the influence of diverse active mode data from different local clients (subnetworks) on the FedASTGNN model. The findings of our study illustrate that the FedASTGNN model effectively preserves the advantages of the ASTGNN model while ensuring data confidentiality in active mode traffic prediction. Furthermore, we observe that the performance of the FedASTGNN model is significantly affected by the varying degrees of imbalanced data distribution among subnetworks. The insights shed light on the potential and challenges presented by the FedASTGNN model as an efficient and secure solution for predicting active mode traffic in Intelligent Transportation Systems.**

Keywords: Traffic prediction, Data security, Federated learning, Graph neural network

I. INTRODUCTION

Intelligent Transportation Systems (ITS) have enabled the collection of vast amounts of valuable traffic data using sensors [1]. These data contain rich features that are crucial for accurate traffic prediction [2]. The precise prediction of traffic conditions holds the potential to greatly reduce travel time, emissions, and energy consumption, making it an area of substantial interest and research within the field of artificial intelligence [3]. Recently, Advanced tools are being developed, with a particular focus on leveraging deep learning techniques. These methods have proven to be highly effective in handling large amounts of traffic data and generating reliable and accurate traffic prediction results [4].However, the majority of current deep learning methods for traffic prediction that consider the spatial-temporal relationships of traffic data mainly concentrate on vehicular traffic modes and predicting bicycle-sharing demand. Predicting active-mode traffic operations, such as cycling and walking, is more challenging due to their complex features, which include irregular travel routes, varying speeds, and position instability.

Centralized traffic prediction models require the exchange of large amounts of traffic data, often owned by different (private or public) organizations or compagnies [5] to central server. However, these data can contain valuable information that may give a competitive edge to companies. Even within the same company, sharing such data across different departments can be challenging due to their sensitive nature. Also, the exchange of massive traffic data between different organizations or departments can result in a significant computational cost, and the lack of security safeguards during this process can pose a threat to the interests of the companies involved [6]. Additionally, integrating data can be challenging and may involve complex administrative procedures [7].

Federated learning, originally proposed by Google in 2016, is an emerging method that enables the training of high-quality global models using multiple devices without the need to share their local data [8]. The field of traffic prediction has seen the emergence of federated learning as a promising framework that addresses the challenge of protecting traffic raw data from different organizations during the training of deep learning models [9]. To investigate the opportunities and challenges associated with federated learning in active mode traffic prediction, we integrate a federated learning framework with an attention-based spatial-temporal graph neural network model called FedASTGNN for active mode traffic prediction, and evaluate the performance gap between the centralized ASTGNN model and the proposed distributed FedASTGNN model. Furthermore, we examine the impact of data heterogeneity among different subnetworks on the FedASTGNN model through the analysis of three distinct scenarios.

*This work is supported by China Scholarship Council

[1]Xiamei Wen is with the Department of Transport and Planning, Delft University of Technology, Stevinweg 1, 2628CN, Delft, the Netherlands X.Wen@tudelft.nl

[2]Panchamy Krishnakumari is with the Department of Transport and Planning, Delft University of Technology, Stevinweg 1, 2628CN, Delft, the Netherlands P.K.Krishnakumari@tudelft.nl

[2]Serge Hoogendoorn is with the Department of Transport and Planning, Delft University of Technology, Stevinweg 1, 2628CN, Delft, the Netherlands S.P.Hoogendoorn@tudelft.nl

II. LITERATURE REVIEW

The increasing number of vehicles and the diverse modes of transportation have led to common challenges in transportation systems, including traffic congestion, delays, and air pollution [10]. Accurate traffic predictions play a crucial role in the analysis of traffic conditions, as they are essential for effective traffic management and control strategies. For instance, based on the prediction results, traffic operators can provide guidance to drivers, suggesting suitable routes to take before they embark on their journey [11]. In this context, various advanced approaches have been developed to achieve accurate and reliable traffic predictions [12]. In addition to motor vehicle prediction, there has been an increasing focus on active modes, such as bicycle-sharing traffic prediction, within the field of Intelligent Transportation Systems (ITS). This attention is driven by the recognition of the importance of active modes in addressing issues related to the first-last mile connectivity, traffic-related air pollution and traffic congestion [13]. However, the current body of research in the field has primarily emphasized the prediction of bicycle sharing mode while neglecting the prediction of private bicycle mode.

In the era of big data, data collection, sharing, and analytics have become commonplace. These practices provide researchers with the opportunity to extract valuable insights from large-scale datasets. However, the process of handling data from diverse devices can be time-consuming and complex. Moreover, sharing data with others can potentially compromise the competitiveness of companies and raise concerns about data security [14] [15]. For example, Zhang et al. [16] realized that data sharing is vulnerable to illegal attacks and can lead to data being lost and destroyed; To ensure data sharing safely in the application of Industrial Internet of Things (IIoT) technology, they proposed a data sharing model based on privacy protection. However, the aforementioned solutions focus on protecting data during the transfer process, while the data itself still needs to be accessed and leave its local storage.

Distributed learning, an emerging paradigm in machine learning, allows for efficient training of a high-quality centralized model by leveraging decentralized training data [7]. However, it should be noted that distributed learning is typically designed for data that follows an independent and identically distributed (IID) pattern. Allowing communication among devices may introduce potential security risks to the data stored on these devices [17]. Federated learning (FL) is one of the distributed learning frameworks introduced by McMahan et al. [18] in 2016 with raw data stored locally and not transferred to any device, which provides an effective data protection mechanism [19]. However, in the FL framework, datasets across different clients may follow different distributions. While McMahan et al. [18] claimed that their proposed Federated Averaging (FedAvg) algorithm can achieve acceptable accuracy even on non-IID data, several studies have found that the accuracy of FedAvg may deteriorate when dealing with non-IID data. Zhao et al. [20] observed a significant drop in accuracy (over 55%) in federated learning when training neural networks using highly imbalanced non-IID local data. Furthermore, it has been observed by several researchers that the imbalanced data distribution among clients can also impact the performance of the model [21]. Existing studies have primarily focused on the imbalance data distribution within the federated learning framework. However, they often overlook the influence of specific situations or contexts on the performance of the models. In particular, in the context of active mode prediction.

III. Methodology

Transportation networks are typically depicted as graphs, and for traffic prediction, graph neural networks (GNNs) have emerged as cutting-edge deep learning models. In contrast to earlier machine learning models like convolutional neural networks (CNNs), GNNs excel at capturing the non-Euclidean graph structure, making them well-suited for modeling irregular network architectures and effectively learning traffic network information in transportation systems. GNNs are mainly categorized into Recurrent Graph Neural Networks (RecGNNs), Convolutional Graph Neural Networks (ConvGNNs), Graph Autoencoders (GAEs), and Spatial–Temporal Graph Neural Networks (STGNNs) [22]. STGNNs are widely applied to traffic prediction due to their ability to simultaneously consider spatial and temporal correlations [23], which is essential to get a well-performed prediction model.

In this study, we construct our traffic networks as an undirected graph $\mathcal{G}_t = (\mu, x_t, \varepsilon, A)$, where μ is the set of nodes representing sensors; x_t represents the traffic flow or density at time t, The historical traffic data of the k time steps is $x = (x_{t-k+1}, x_{t-k+2}, ..., x_t)$, which would be used to predict the following T time steps $\hat{x} = (\hat{x}_{t+1}, \hat{x}_{t+2}, ..., \hat{x}_{t+T})$; The set of edges is denoted as ε; and $A \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ is the adjacency matrix of $\mathcal{G}$ recorded as a binary matrix, which means that if $\forall \mu_i, \mu_j$ connect to each other $A[i,j] = 1$, otherwise $A[i,j] = 0$; N is the number of nodes. To ensure data security, we apply the FL framework with the traffic prediction model to predict traffic flow and density. The entire transportation network is represented as a global network $\mathcal{G}$. The global network is manually divided into sub-networks $\tilde{G} = \{\tilde{G_1}, \tilde{G_2}, ..., \tilde{G_s}\}$ based on the size of the graph data of each sensor, where s is the number of sub-networks or clients. It is assumed that the sub-networks exhibit distinct structures and non-overlapping data.

A. Traffic Prediction Method under data security Constraints

In this study, we integrated the federated averaging algorithm (FedAvg) [18] with ASTGNN [24] to explore the performance of the traffic prediction method without compromising data security.

1) The FedAvg algorithm offers a promising approach for local clients to collectively learn a global machine learning model, denoted as M_r, without the need to exchange their raw data. Here, r represents the communication round. In FedAvg, the training process unfolds as follows.

Step 1 Global model initialization: The central server determines the hyperparameters of the global model, including training strategies like the communication round and learning rate. Additionally, the central server selects the required number of clients for the ongoing communication round and broadcasts the initialized global model, denoted as M_0, to the chosen clients. $\tilde{G} = \{\tilde{G}_1, \tilde{G}_2, ..., \tilde{G}_s\}$. Our study encompasses all the clients we have.

Step 2 Client model training and parameter update: The clients engage in training the global model, denoted as M_r, using their local data, resulting in the update of respective parameters w_r and s. Subsequently, the clients upload these updated parameters to the central server.

Step 3 Global model parameter aggregation and update: The central server gathers the parameters $w = w_{r,1}, w_{r,2}, ..., w_{r,s}$ from the clients and aggregates them using the following formulation. The model is then updated based on the aggregated results.

$$w_{r+1} = \sum_{s=1}^{S} \frac{n_s}{n} w_{r+1}^{s} \tag{1}$$

Where S represents the total number of selected clients, while s denotes the label of an individual client. w_{r+1} corresponds to the updated parameter of round r. Additionally, n_s stands for the number of client-specific data, n represents the total number of training data, and w_{r+1}^{s} signifies the updated parameter of client s.

2) FedASTGNN algorithm: The FedASTGNN model is an amalgamation of the federated averaging (FedAvg) algorithm [18] and the ASTGNN model. The ASTGNN model, initially introduced by Guo et al. [24], is designed for traffic prediction and effectively captures the temporal dynamics of traffic data while incorporating spatial correlations through the use of self-attention mechanisms and graph convolution networks. The training procedures of FedASTGNN are outlined in **Figure 1**. Initially, the central server aggregates the sub-networks from the clients, disregarding their connections, to construct the global ASTGNN model. This means that the global network does not include any connection information between the sub-networks. Subsequently, the central server broadcasts the initialized global ASTGNN model to the clients. The local clients then train the global ASTGNN model using their respective networks and data, obtaining updated parameters that are subsequently uploaded to the central server. The central server employs the FedAvg algorithm to aggregate the updated parameters from the local clients, thereby updating the global ASTGNN model. The latest global model is then transmitted back to the local clients for the next communication round, and this process continues until convergence is achieved.

B. Performance Comparison in Different Scenarios

1) In order to evaluate the performance of the FedASTGNN model, we compare it with the ASTGNN model [24] using balanced graph data distribution among local sub-networks. Considering that different time aggregation intervals of raw data and prediction horizons can impact prediction accuracy, we conduct experiments using aggregation intervals of 5 minutes, 10 minutes, 15 minutes, and 20

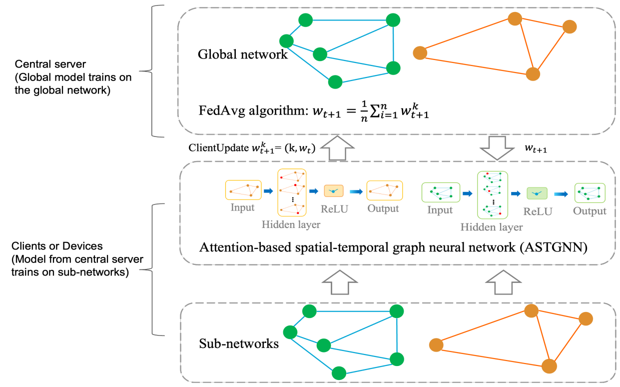

Fig. 1: FedASTGNN communication process

minutes, respectively. The prediction horizons are set from 5 minutes to 60 minutes. By examining various error metrics, we assess the predictive capabilities of both models.

2) To investigate the impact of active mode data on traffic prediction using the FedASTGNN model, we consider three types of graph structures with varying data distributions. Structure1 consists of two sub-networks with approximately balanced amounts of data. Structure2 and Structure3 also have two sub-networks, but they exhibit imbalanced data distributions. Notably, Structure3 exhibits a more significant data imbalance between its sub-networks compared to Structure2.

C. Accuracy Evaluation

This paper assesses the prediction results of various scenarios using metrics including Mean Square Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). The formulas for calculating these metrics are presented below:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{4}$$

Where $\hat{y}_i$ and y_i represent the predicted value and ground truth data of the value i. n is the number of sample values.

IV. Data Description and Modeling Configuration

The bicycle dataset utilized in this paper was obtained from sensors installed on the campus of TU Delft (https://micd.tudelftcampus.nl/projects/outdoor-mobility-digital-twin/). The dataset used in this paper includes geographic information obtained from sensors. It contains recorded active mode flow and density data at one-minute intervals as captured by the sensors.

A. Dataset Pre-processing

For our experiments, we utilized a dataset collected by seven sensors to conduct the prediction task. The sampling data used in our study spans from March 1 to March 31, 2021. It includes data collected for 12 hours each day, specifically from 8:00 to 20:00. The cumulative flow for each sensor in March is as follows: 164184, 122485, 48981, 15217, 8786, 21988 and 27058, respectively. In the first scenario of the Results and Discussion section, the bicycle flow data is aggregated into 5-minute, 10-minute, 15-minute, and 20-minute intervals. This is done to determine the optimal aggregation intervals in the dataset and assess the prediction performance with different prediction horizons. In this paper, the prediction is performed for different time steps: 1 step, 2 steps, and 3 steps. The prediction horizons are determined by multiplying the time step by the data aggregation interval (time step $\times$ aggregation interval). In the second scenario, we construct three network structures based on the dataset size of each sensor. Structure1 consists of two sub-networks: the first sub-network is composed of three sensors (TUD_SBX01, TUD_SPX04, TUD_SPXS06_IO) with a total of 206,459 trips, and the second sub-network comprises four sensors (TUD_SBX02, TUD_SBX03, TUD_SBX05, TUD_SBX06) with a total of 202,240 trips. The second network structure is divided into two sub-networks: one includes three sensors (TUD_SBX01, TUD_SBX03, TUD_SPXS06_IO) with 240,223 trips, and the other includes four sensors (TUD_SBX02, TUD_SPX04, TUD_SBX05, TUD_SBX06) with 168,476 trips. The last network structure also consists of two sub-networks: the first sub-network includes three sensors (TUD_SBX01, TUD_SBX02, TUD_SBX03) with 335,650 trips, and the second sub-network includes four sensors (TUD_SPX04, TUD_SBX05, TUD_SBX06, TUD_SPXS06_IO) with 73,049 trips.

B. Parameters Setting

All experiments in this paper were conducted on an Apple M1 Pro computing server. The dataset was split into 60% for training, 20% for validation, and the remaining 20% for testing. The FedASTGNN model was configured with a model dimension of 64, 3 encoder and decoder layers, a convolution kernel size of 3, and 8 attention heads. The hyperparameters such as learning rate, batch size, and training epochs were determined based on the model's performance on the validation set.

For flow prediction with different aggregation intervals and prediction horizons using balanced data distribution between sub-networks, the batch size ranged from 500 to 560, the learning rate was set to 0.0001, and the number of communication rounds was 10. In the case of structure2 and structure3, the batch sizes were set to 520 and 480, respectively, both with a learning rate of 0.0001, and a fixed number of communication rounds of 10.

For ASTGNN, we utilized the model dimension settings provided by the original authors and their original code. However, we adjusted the learning rate, batch size, and training epochs to optimize the model's performance on our specific dataset.

V. Result and Discussion

A. Performance Comparison of the ASTGNN and FedASTGNN model

Table I shows the comparison of active mode flow prediction results for FedASTGNN and ASTGNN models across various time aggregation intervals and prediction horizons. Specifically, we employ different time aggregation intervals (5-minute, 10-minute, 15-minute, and 20-minute) to predict traffic flow. These intervals are paired with prediction horizons of 1 step (5-minute, 10-minute, 15-minute, and 20-minute, respectively), 2 steps (10-minute, 20-minute, 30-minute, and 40-minute, respectively), and 3 steps (15-minute, 30-minute, 45-minute, and 60-minute, respectively). The findings indicate a decline in the performance of both models as the prediction horizons increase.

For instance, in the case of a 5-minute aggregation interval, the ASTGNN model demonstrates an increase of 0.44 in Mean Squared Error (MSE), 0.08 in Mean Absolute Error (MAE), and 0.14 in Root Mean Squared Error (RMSE) when transitioning from a one-step prediction horizon to a three-step prediction horizon. Likewise, in the FedASTGNN model, we observe an upward trend in MSE, MAE, and RMSE as the prediction horizons increase. Specifically, the MAE increases from 0.95 to 1.03, 0.86 to 1.02, 0.86 to 1.09, and 0.86 to 1.15 for the 5-minute, 10-minute, 15-minute, and 20-minute aggregation intervals, respectively.

Furthermore, it is evident that the 15-minute aggregation interval yields the lowest prediction errors among the four aggregation intervals when considering the one-step prediction horizon for both the ASTGNN and FedASTGNN models. When examining the ASTGNN model with a two-step prediction horizon, we observe a decrease in Mean Squared Error (MSE) from 2.41 for the 5-minute aggregation interval to 2.25 for the 10-minute aggregation interval, however, the MSE subsequently rises to 2.65 for the 20-minute aggregation interval. Additionally, when considering a two-step prediction horizon, we find that the dataset with 10-minute aggregation intervals exhibits the lowest prediction errors. Hence, the 10-minute aggregation interval proves to be the most suitable choice for the two-step prediction horizon in this flow dataset. As for the three-step prediction horizon, the 5-minute aggregation interval strikes a balance by capturing the traffic flow data fluctuations adequately while maintaining a high level of accuracy.

The findings presented in **Table I** also demonstrate that the ASTGNN model exhibits lower errors compared to the FedASTGNN model. This suggests that ASTGNN, being a centralized model with access to all the information in the graph network, outperforms the FedASTGNN model, where data owners train the model locally. However, the difference in prediction errors between the ASTGNN and FedASTGNN models for the 5-minute aggregation interval is not significant, ranging from 0.87% to 3.72%. Moreover, when considering the computational cost of each FedASTGNN

TABLE I: The Comparison of Traffic Flow Prediction Accuracy

Aggregation Interval (mins)	Prediction Horizon (mins)	ASTGNN			FedASTGNN		
		MSE	MAE	RMSE	MSE	MAE	RMSE
5	5	2.19	0.94	1.48	2.27	0.95	1.51
	10	2.41	0.98	1.55	2.48	0.99	1.57
	15	2.63	1.02	1.62	2.73	1.03	1.65
10	10	1.84	0.85	1.36	1.88	0.86	1.37
	20	2.25	0.93	1.50	2.33	0.94	1.53
	30	2.66	1.01	1.63	2.80	1.02	1.67
15	15	1.70	0.83	1.30	1.86	0.86	1.36
	30	2.36	0.95	1.54	2.54	0.98	1.59
	45	3.03	1.05	1.74	3.33	1.09	1.83
20	20	1.77	0.84	1.33	1.93	0.86	1.39
	40	2.65	0.99	1.63	2.89	1.01	1.70
	60	3.58	1.11	1.89	3.91	1.15	1.98

TABLE II: The Prediction Accuracy of Structures with Different Data Distribution

Structures	Sub-network1	Sub-network2	Error		
			MSE	MAE	RMSE
Structure1	206459	202240	1.86	0.86	1.36
Structure2	240223	168476	1.87	0.86	1.37
Structure3	335650	73049	1.98	0.88	1.41

communication round compared to ASTGNN, the training time consumption is nearly identical. Therefore, FedASTGNN, which shares a similar main structure with ASTGNN, can provide a trade-off between prediction accuracy and data security without compromising significantly on the accuracy of predictions.

B. Performance Comparison of FedASTGNN with Imbalanced Networks

Within this section, we propose three network structures that exhibit varying degrees of imbalanced data distributions among sub-networks for the purpose of traffic flow prediction using the FedASTGNN model. The prediction errors are shown in **Table II**. The performance metrics (MSE, MAE, and RMSE) for three different network structures are as follows:

Structure1: This structure exhibits approximately balanced data distribution between sub-networks, with one sub-network having 206,459 trips and the other sub-network having 202,240 trips. The corresponding MSE, MAE, and RMSE values are 1.86, 0.86, and 1.36, respectively.

Structure2: In this structure, the data distribution between the two sub-networks is slightly unbalanced, but not significantly. The MSE, MAE, and RMSE for this structure are 1.87, 0.86, and 1.37, respectively, which are slightly larger than those of the structure with a balanced data distribution.

Structure3: The number of trips in Subnetwork1 is more than four times the number of trips in Subnetwork2 in this structure, resulting in an imbalanced data distribution. With this configuration, the MSE, MAE, and RMSE values for Structure3 are 1.98, 0.88, and 1.41, respectively.

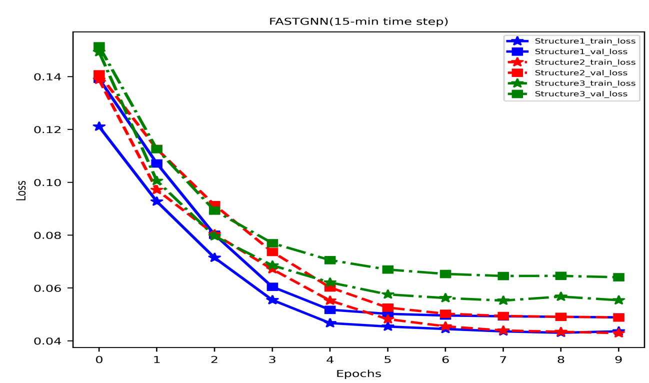

Fig. 2: Loss of Traffic flow prediction

These results indicate that Structure1 with a balanced data distribution yields the lowest prediction errors, while Structure3 with an imbalanced data distribution leads to slightly higher errors. Furthermore, depicted in **Figure 2**, the loss convergence speed of Structure1, which possesses a balanced data distribution, is considerably faster compared to Structure2 and Structure3. This highlights the influence of graph data distribution among local sub-networks on the performance of FedASTGNN in the situation of active mode traffic prediction. The presence of an unbalanced graph data distribution negatively impacts the performance of FedASTGNN, leading to a degradation in its active mode traffic predictive capabilities.

VI. Conclusion

In this study, we examine the performance of the FedASTGNN model for active mode traffic prediction while adhering to data sharing constraints. Additionally, we investigate the impact of imbalanced data distribution among sub-networks on the FedASTGNN model. This research is crucial for the development of highly effective prediction models that can be regularly updated in real-time using extensive traffic data while requiring minimal data storage space.

After conducting the analysis mentioned above, it has been determined that the FedASTGNN model shows great promise for predicting active mode traffic. However, there are certain challenges associated with the characteristics of active mode data, specifically the heterogeneous data distribution among local organizations, as well as the need to ensure data security. Moving forward, our focus will be on developing a federated learning approach to address these challenges in active mode traffic prediction. This method will take into account the varying data distributions among local organizations while enhancing the scalability and robustness of the model. By doing so, we aim to create an effective solution that can handle the complexities of active mode traffic prediction while maintaining data security.

ACKNOWLEDGMENT

I would like to express my gratitude for the support provided by the China Scholarship Council, which has played a vital role in facilitating this research.

References

[1] Lv, Y., Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, Traffic Flow Prediction With Big Data: A Deep Learning Approach. IEEE Transactions on Intelligent Transportation Systems, Vol. 16, No. 2, 2015, pp. 865–873.

[2] Liu, Z., Z. Li, K. Wu, and M. Li, Urban Traffic Prediction from Mobility Data Using Deep Learning. IEEE Network, Vol. 32, No. 4, 2018, pp. 40–46.

[3] Yao, H., X. Tang, H. Wei, G. Zheng, and Z. Li, Revisiting spatial-temporal similarity: A deep learning framework for traffic prediction. In Proceedings of the AAAI conference on artificial intelligence, 2019, Vol. 33, pp. 5668–5675.

[4] Meena, G., D. Sharma, and M. Mahrishi, Traffic Prediction for Intelligent Transportation System using Machine Learning. In 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), 2020, pp. 145–148.

[5] Lu, Y., X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, Blockchain Empowered Asynchronous Federated Learning for Secure Data Sharing in Internet of Vehicles. IEEE Transactions on Vehicular Technology, Vol. 69, No. 4, 2020, pp. 4298–4311.

[6] Lee, H. and J. Kim, Trends in Blockchain and Federated Learning for Data Sharing in Distributed Platforms. In 2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN), 2021, pp. 430–433.

[7] Yang, Q., Y. Liu, T. Chen, and Y. Tong, Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), Vol. 10, No. 2, 2019, pp. 1–19.

[8] Mammen, P. M., Federated Learning: Opportunities and Chal- lenges, 2021.

[9] Manias, D. M. and A. Shami, Making a Case for Federated Learning in the Internet of Vehicles and Intelligent Transporta- tion Systems. IEEE Network, Vol. 35, No. 3, 2021, pp.88–94.

[10] George, S. and A. K. Santra, Traffic prediction using multi- faceted techniques: a survey: Wireless Personal Communica- tions, Vol. 115, No. 2, 2020, pp. 1047–1106.

[11] Guo, K., Y. Hu, Z. Qian, H. Liu, K. Zhang, Y. Sun, J. Gao, and B. Yin, Optimized Graph Convolution Recurrent Neural Network for Traffic Prediction. IEEE Transactions on Intelligent Transportation Systems, Vol. 22, No. 2, 2021, pp. 1138–1149.

[12] Zhao, L., Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li, T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. IEEE Transactions on Intelligent Trans- portation Systems, Vol. 21, No. 9, 2020, pp. 3848–3858.

[13] Zhang, S.-M., X. Su, X.-h. Jiang, M.-l. Chen, and T.-Y. Wu, A Traffic Prediction Method of Bicycle-sharing based on Long and Short term Memory Network. J. Netw. Intell., Vol. 4, No. 2, 2019, pp. 17–29.

[14] Nguyen, D. C., M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, Federated learning for internet of things: A comprehensive survey. IEEE Communications Surveys Tu- torials, Vol. 23, No. 3, 2021, pp. 1622–1658.

[15] Dhawan, S. and R. Gupta, Analysis of various data security techniques of steganography: A survey. Information Security Journal: A Global Perspective, Vol. 30, No. 2, 2021, pp. 63–87.

[16] Zhang, Q., Y. Li, R. Wang, L. Liu, Y.-a. Tan, and J. Hu, Data security sharing model based on privacy protection for blockchain-enabled industrial Internet of Things. International Journal of Intelligent Systems, Vol. 36, No. 1, 2021, pp. 94–111.

[17] Verbraeken, J., M. Wolting, J. Katzy, J. Kloppenburg, T. Ver- belen, and J. S. Rellermeyer, A survey on distributed machine learning. Acm computing surveys (csur), Vol. 53, No. 2, 2020, pp. 1–33.

[18] McMahan, B., E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, Communication efficient learning of deep networks from decentralized data. In Artificial intelligence and statistics, PMLR, 2017, pp. 1273–1282.

[19] Zhang, T., L. Gao, C. He, M. Zhang, B. Krishnamachari, and A. S. Avestimehr, Federated Learning for the Internet of Things: Applications, Challenges, and Opportunities. IEEE Internet of Things Magazine, Vol. 5, No. 1, 2022, pp. 24–29.

[20] Zhao, Y., M. Li, L. Lai, N. Suda, D. Civin, and V. Chan- dra, Federated learning with non-iid data. arXiv preprint arXiv:1806.00582, 2018.

[21] Duan, M., D. Liu, X. Chen, Y. Tan, J. Ren, L. Qiao, and L. Liang, Astraea: Self-Balancing Federated Learning for Improving Classification Accuracy of Mobile Deep Learning Applications. In 2019 IEEE 37th International Conference on Computer Design (ICCD), 2019, pp. 246–254.

[22] Wu, Z., S. Pan, F. Chen, G. Long, C. Zhang, and P.S. Yu, A Comprehensive Survey on Graph Neural Networks. IEEE Transactions on Neural Networks and Learning Systems, Vol. 32, No. 1, 2021, pp. 4–24.

[23] Yu, B., H. Yin, and Z. Zhu, Spatio-temporal graph con- volutional networks: A deep learning framework for traffic forecasting. arXiv preprint arXiv:1709.04875, 2017.

[24] Guo, S., Y. Lin, H. Wan, X. Li, and G. Cong, Learning Dynamics and Heterogeneity of Spatial-Temporal Graph Data for Traffic Forecasting. IEEE Transactions on Knowledge and Data Engineering, 2021, pp. 1–1.