Differentially Private Traffic Flow Prediction Using Transformers: A Federated Approach

Sargam Gupta(✉) and Vicenç Torra

Department of Computing Science, Umeå University, Umeå, Sweden
{sgupta,vtorra}@cs.umu.se

Abstract. Accurate traffic flow prediction plays an important role in intelligent transportation management and reducing traffic congestion for smart cities. Existing traffic flow prediction techniques using deep learning, mostly LSTMs, have achieved enormous success based on the large traffic flow datasets collected by governments and different organizations. Nevertheless, a lot of these datasets contain sensitive attributes that may relate to users' private data. Hence, there is a need to develop an accurate traffic flow prediction mechanism that preserves users' privacy. To address this challenge, we propose a federated learning-based temporal fusion transformer framework for traffic flow prediction which is a distributed machine learning approach where all the model updates are aggregated through an aggregation algorithm rather than sharing and storing the raw data in one centralized location. The proposed framework trains the data locally on client devices using temporal fusion transformers and differential privacy. Experiments show that the proposed framework can guarantee accuracy in predicting traffic flow for both the short and long term.

Keywords: Federated Learning · Traffic Flow Prediction · Differential Privacy · Temporal Fusion Transformer · Time Series Data

1 Introduction

Urban transportation is a vital part of everyday life. Traffic congestion on roads is one of the major concerns in today's transportation system. Most of the people traveling on the road utilize their own observations for selecting an optimum time and path to commute. In the absence of accurate traffic flow predictions, this leads to longer commute times and delays. Hence, everybody requires a timely and accurate traffic flow prediction. Using accurate traffic flow prediction techniques, we can use historic traffic data to predict future road conditions that can be utilised in different location-based services.

This study was partially funded by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

S. Katsikas et al. (Eds.): ESORICS 2023 Workshops, LNCS 14398, pp. 260–271, 2024.
https://doi.org/10.1007/978-3-031-54204-6_15

Even after being such an essential part of Intelligent Transport Management, traffic flow prediction is difficult and quite challenging. First of all, the nature of the traffic data is spatiotemporal. The models predicting the traffic flow must capture both the time-series information and spatial features of the location. Secondly, the traffic flow in a particular region is highly dependent on many different external factors like which day of the week is it or is there any special event happening on a particular day. Hence, these factors need to be considered. Thirdly, most of the existing works can only do a short-term prediction for about the next 30 to 50 min which might not be enough time for the commuters to plan their journey route. Lastly and most importantly, a lot of the historic traffic data may contain some sensitive information about the vehicle which may reveal some private information Hence, it is very essential to build a framework for traffic flow prediction that gives accurate predictions and at the same time preserves sensitive information.

Most of the traffic data is collected by the sensors deployed on the road. This collected data is stored at a central location over which a traffic prediction model is trained. This intrudes on the privacy of the data collected and increases the prediction duration Hence, to address the above-mentioned issues, we propose a novel federated differentially private traffic flow prediction framework based on Temporal Fusion Transformers. The contributions of this paper are summarized as follows:

- We propose a novel privacy-preserving traffic flow prediction framework that integrates Federated Learning (FL), Differential Privacy (DP) and Temporal Fusion Transformers (TFTs). This framework gives accurate and timely predictions without actually sharing the raw data collected from the sensors.
- We incorporate various static information like the day of the week and the calendar holidays within the region which improve the accuracy of the prediction.
- We have included the long-term prediction of traffic flow in a region that was missing in the existing literature.

We use two evaluation metrics (Mean Squared Error and Mean Absolute Error) on a real-world dataset for the simulation of the proposed framework. From the obtained results we can clearly see that our proposed algorithm has a higher performance when compared with the existing works.

The remaining part of this paper is organized as follows. Section 2 reviews some basic concepts needed in understanding the paper. In particular, we discuss FL, DP, and TFTs. Section 3 describes, in brief, some existing literature on traffic flow prediction. Our suggested prediction mechanism is described in Sect. 4. Section 5 describes the dataset and simulation settings for experiments. In Sect. 6, we provide and discuss the results. Section 7 gives the conclusion and future directions.

2 Preliminaries

2.1 Federated Learning

The concept of FL was proposed by Google [5] in 2016 which allows building a collaborative model from distributed data without actually sharing and storing it at a centralized location thereby preserving the privacy and security of the data.

Assuming M clients $\{C_1, C_2, \ldots, C_M\}$, Yang *et al.* [18] define federated learning as a process of constructing collaborative model M_{Fed} with accuracy A_{Fed} such that

$$|A_{Fed} - A_{Cen}| < \delta \tag{1}$$

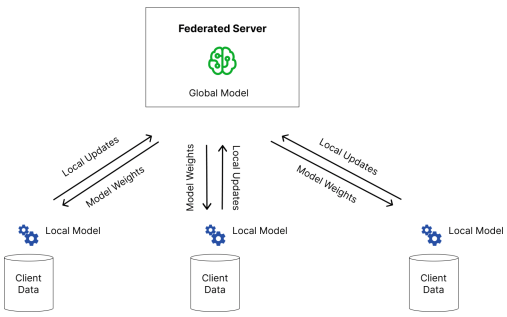

Fig. 1. Federated Learning

where A_{Cen} is the accuracy of centralized machine learning on the centralized dataset $D = D_1 \cup D_2 \cup \cdots \cup D_M$ and δ is a non-negative real number if Eq. (1) holds. The FL algorithm is said to have δ-accuracy loss [18]. There are mainly two categories of FL - one, where data at clients have the same features but different samples, called horizontal FL, and second, called vertical FL, where clients have different feature spaces. In our proposed solution, we work with horizontal FL.

In each communication round, the server transmits the global model parameters to the selected clients. These clients perform the local model training on their own individual dataset and send their updated parameters to the server which then aggregates the differences to the global model. This communication stops when convergence is achieved. The system architecture of FL is represented in Fig. 1.

2.2 Differential Privacy

Differential Privacy [2] is considered as the de facto standard of privacy by most researchers in the field of privacy. DP can provide strong privacy guarantees if the selected values of ε and δ are good. The formal definition of differential privacy is given as follows. An algorithm M is said to be (ε, δ) - differentially private if

$$P(M(D \in S) \leq e^{\in} P(M(D') \in S) + \delta \tag{2}$$

where D and D' are neighbouring datasets and S is an arbitrary subset of outputs of M. ε is the privacy budget and δ is the relaxation term. A smaller value of ε enforces stronger privacy.

2.3 Temporal Fusion Transformers

Transformers are the state of art deep learning models that were proposed recently in 2017 [16]. They use the self-attention mechanism for different types of tasks. Though they were originally proposed for natural language processing models but several different versions of transformers have become popular over the time. The major advantage of using the transformer models over the traditionally used LSTMs [4] in sequential processing tasks is that they require much less training time due to parallelization. One such type of transformer is the TFT.

TFTs were proposed by Lim et al. [7] in 2020. This model is specifically designed for interpreting and predicting the time-series data. It has several novel architectures that have improved the prediction performance for time series considerably. The TFTs consider different types of inputs like static inputs which could be the never-changing information like the ids of sensors, known inputs which are known even after the input time like a day of the week and holidays, and observed values.

The main modules of the TFT architecture are:

- Gating module: This module helps in filtering out the not-so-necessary complex details in the model formed and hence reducing the complexity of the trained model.
- Variable Selection Network: This module, being true to its name, is used for the feature selection mechanism.
- Static information encoder: This encodes the static information in the problem considered for prediction.
- LSTM encoder-decoder layers: Since our main input is sequential in nature, hence it is worthwhile to consider using LSTM layers to process the temporal information well.
- SeqtoSeq layer and Multi-head attention module: They are used for capturing the short-term and long-term dependencies in the data respectively.

3 Related Work

In this section, we discuss the most relevant research in the field of traffic flow prediction. In the initial studies related to traffic flow prediction, most researchers used traditional machine learning algorithms to solve the time-series problem. Gary et al. [1] proposed a K-Nearest Neighbour approach for short-time traffic flow prediction. Another Bayesian network-based [14] approach was proposed by Sun et al. which took into consideration the adjacent road links to analyze the traffic better. A few more machine learning-based approaches were proposed based on Support Vector Machines (SVMs) [6] and Autoregressive Integrated Moving average (ARIMA) [13] but none of them was accurate enough. This could be because

of the complicated relationship between features, volume, and uncertainty of the traffic flow data. Hence, researchers started exploring some deep-learning techniques for time-series prediction [11]. Since, the time-series data is sequential in nature, hence, researchers found that Recurrent Neural Networks (RNNs) could be good at capturing the temporal features in traffic flow prediction. Ma et el. [9] proposed a bidirectional LSTM to capture the time features better while Fu et al. RNNs indeed performed better than traditional machine learning algorithms. Another set of researchers used Convolutional Neural Networks (CNNs) and demonstrated good ability to capture the features in the field of computer vision. Zhang et al. [21] used CNNs for predicting urban traffic flow and captured the correlations of traffic with each road in a city. Some researchers tried combining properties of both CNNs and RNNs like Zhang et al. [20] which used the ST-ResNet, to collectively forecast the inflow and outflow of crowds in a city. Xia et al. [17] proposed a distributed WND-LSTM model in MapReduce that can predict traffic flow for distributed traffic networks. All these models could predict the traffic flow with a decent accuracy but all of them were centralized models and hence did not take data privacy into consideration. Since data privacy is a major concern, hence it is very important to find out an alternative to these models. FL being an emerging field attracted a lot of researchers' eyes. Liu et al. [8] proposed a federated learning-based highway traffic prediction using GRUs. FL, though, is more secure when compared to the centralized approach but it is still not enough. To ensure more privacy in FL approaches Yang et al. [19] proposed privacy-preserving Additive Homomorphic Encryption (AHE) in FL. AHE is a good way of securing the FL environment but it is very computationally intensive and slow and hence cannot be used in timely traffic flow prediction. To overcome these Qi et al. [12] proposed a blockchain-based federated learning approach combined with GRUs and Tang et al. [15] proposed a differential privacy-based federated learning approach with LSTMs for short-term prediction. Though LSTM models are reasonably accurate but training them is difficult as they have a larger number of parameters and cannot parallelize the task. Hence, we propose transformer-based models to predict short-term as well as long-term traffic flows in conjugation with privacy protection.

4 Differentially Private Federated Traffic Flow Prediction Using Temporal Fusion Transformers

In this section, we propose two variants of a new federated traffic flow prediction framework. The two variants differ only in the noise added to them.

Suppose we have m different sensors located in different parts of the city. Each sensor collects the traffic flow data D_i from its region. Each client's data D_i is not shared with anyone and is only used by the client for training their model. Each sensor constructs a TFT model on the dataset D_i. Then, the model updates are sent to the aggregating server where we use the FedAVG [10] algorithm to aggregate these parameters. These aggregated parameters form the global model. This global model is then sent back to the clients. This results in learning from each other's datsets without actually knowing the data.

4.1 Client-Side Training

Figure 2 presents the client-side training steps. Each client collects road traffic data every hour. When feeding this input to the TFT, we segregate the inputs into different types. The first type is the static input values which is the detector id and road number. Then, we input the known inputs. These values are the one that we know even after the prediction time like the hour of the day, day of the week, month, holiday or not etc. Lastly, we input the observed value. This is the value that we want our model to predict after training. In our case, this observed value is the number of vehicles on the road at any particular time. The data format file is created and we set the look back historical window and the prediction length of our model. We then set the hyperparameters which include the number of LSTM layers, dropout rate, minibatch size and number attention heads. Lastly, we train our TFT model and obtain the model parameters.

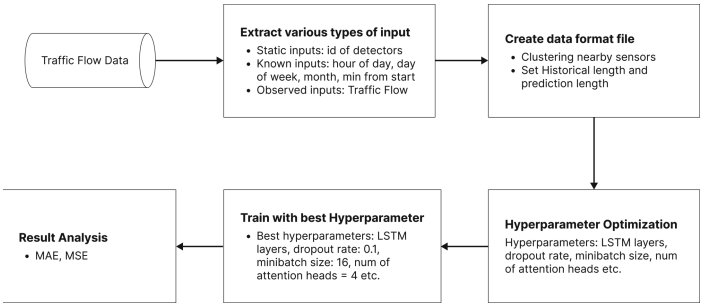

Fig. 2. Client side architecture

4.2 Model Perturbation

When the TFT model is trained on the client, then to ensure the security of the sensitive information we apply ε differential privacy on the client-side trained TFT model. More concretely, we supply the Gaussian mechanism and add noise to the parameters. In our experiments, we add the noise with varying values of epsilon from 0.1 to 0.9 and see how it impacts the global model's prediction accuracy.

4.3 Aggregation Algorithm

Since, aggregation is the key component of this framework, we use the FedAVG algorithm [10] for secure parameter aggregation. It is one of the simplest yet effective and very popular aggregation algorithms. Every iteration of the algorithm starts with initializing a global model to all the clients. The clients train on that model with their own local datasets and obtain a new updated model. The updates in the model parameters of these updated local models are then sent to the global server. The global server aggregates these updates by performing

a weighted average of their values. This forms a new global model. Again this updated global model is sent back to the clients for local training. This process continues iteratively until it reaches convergence. Once, the converged model is ready we use this model to predict on the data to evaluate its performance.

Table 1. Prediction results of the proposed model without DP

	Time Step	MSE	MAE
Centralized	24	0.0117	0.0567
	72	0.0134	0.0562
	720	0.0178	0.0713
2 Clients	24	0.0240	0.0931
	72	0.0435	0.1098
	720	0.0466	0.1536
4 Clients	24	0.0495	0.1589
	72	0.0639	0.1795
	720	0.0726	0.1689
6 Clients	24	0.0515	0.1525
	72	0.0552	0.1541
	720	0.0680	0.1646
8 Clients	24	0.0442	0.1461
	72	0.0556	0.1608
	720	0.0721	0.1667

5 Dataset and Experimental Settings

We are using the real-world public dataset collected by Caltrans Performance Measurement System (PeMS) (http://pems.dot.ca.gov) in California. This dataset contains the traffic flow information from the San Francisco Bay area freeways. The data is collected from 862 different sensors located on the highway system. The data is available for two years from 2015 to 2016 with a reading of traffic on roads after every hour. We used three months of data from January 2015 to March 2015 for training the TFT model and predicting the values for the following one day, three days and one month. We have used the Darts TFT [3] python library for implementing the TFT code. For simulating the federated settings we clustered the sensors located in a nearby region in proximity to each other into a single client. The values for the different TFT hyperparameters for the experiments were set as input chunk length as 64, output chunk length as 8, hidden size as 64, LSTM layers as 1, num attention heads as 4, dropout as 0.1, batch size as 16 and epochs=3. The proposed algorithm is simulated for 2, 4, 6 and 8 clients. We also apply DP on the client-side models with different ε values (0.1,0.5 and 0.9) to show how it impacts the prediction

results. For the evaluation of our results, we used Mean Squared Error and Mean Absolute Error.

$$MSE = 1/M \sum_{i}^{M} (actual_i - predicted_i)^2 \tag{3}$$

$$MAE = 1/M \sum_{i}^{M} |actual_i - predicted_i| \tag{4}$$

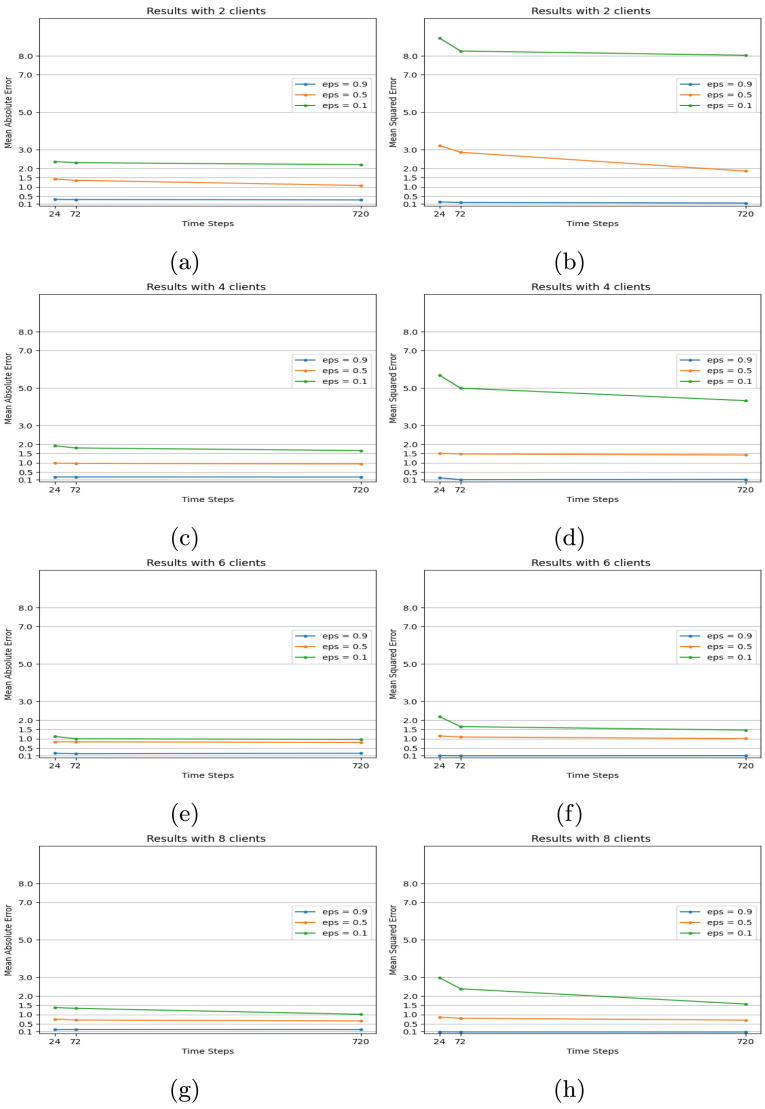

Fig. 3. Comparison of epsilon values for DP

6 Results and Analysis

The results of the experiments described are presented below.

Table 2. Prediction results of the proposed model with DP

Clients	Time Step	Eps	MSE	MAE
2 Clients	24	0.9	0.1808	0.3321
	72	0.9	0.1746	0.3298
	720	0.9	0.1706	0.3103
	24	0.5	3.2065	1.4260
	72	0.5	2.8518	1.3553
	720	0.5	1.8467	1.0783
	24	0.1	8.9417	2.3549
	72	0.1	8.2561	2.2964
	720	0.1	8.2084	2.1916
4 Clients	24	0.9	0.1027	0.2458
	72	0.9	0.0990	0.2440
	720	0.9	0.1140	0.2436
	24	0.5	1.5119	0.9814
	72	0.5	1.4715	0.9699
	720	0.5	1.4168	0.9445
	24	0.1	5.6733	1.9114
	72	0.1	4.9860	1.7923
	720	0.1	4.3179	1.6547
6 Clients	24	0.9	0.0898	0.2245
	72	0.9	0.0905	0.2242
	720	0.9	0.0980	0.2249
	24	0.5	1.1510	0.8403
	72	0.5	1.0912	0.8333
	720	0.5	1.0094	0.8054
	24	0.1	2.1788	1.1221
	72	0.1	1.6476	0.9995
	720	0.1	1.4634	0.9617
8 Clients	24	0.9	0.0768	0.1975
	72	0.9	0.0715	0.2056
	720	0.9	0.0679	0.2019
	24	0.5	0.8602	0.7552
	72	0.5	0.8055	0.7119
	720	0.5	0.7056	0.6606
	24	0.1	2.9688	1.3788
	72	0.1	2.3772	1.3330
	720	0.1	1.5631	1.0145

In Table 1, we compared the results of our proposed framework with the centralized Model. It shows the MSE and MAE values of the model by varying the number of clients. It can be seen from the results that our federated framework performs quite well and the obtained values are comparable to the centralized approach. Though the error increases slightly with the increase in the number of clients yet we consider that remains within reasonably good limits. We also compare our proposed work with FedGRU [8]. The MAE and MSE values in their work are 7.96 and 101.49 respectively for the same dataset. We can clearly see that these are quite high when compared with the values of our approach.

In Table 2, we share the values of the MAE and MSE when varying the value of ε for adjusting the privacy budget. In order to evaluate our results we can report from the literature, the MAE and MSE values of FedLSTM with Differential Privacy [15]. They are 7.65 and 100.47 respectively which are very high when compared to our framework's results with DP. Hence, our proposed model performs better than other baselines and existing works in the literature. In Fig. 3, we have plotted the values of MSE and MAE to measure the effectiveness of our model after adding noise. We have considered three values of ε, in our experiments: 0.1,0.5 and 0.9. Please note that the lower the value of ε, more

is the noise added. We can see from Table 2 that with the highest ε value, the noise added is less, thus the error values are low and vice versa. We can also observe that with the increase in the number of clients, the values of MAE and MSE show a reducing trend which makes our proposed framework suitable to be used in FL settings with large number of clients. Also, when comparing our error values with existing FedLSTM with DP [15], our values are smaller. Therefore, our approach is better.

7 Conclusion and Future Works

This paper presents a novel federated traffic flow prediction framework based on temporal fusion transformers and differential privacy which can make timely and accurate long-term as well as short-term predictions. The proposed federated framework is privacy-preserving as it does not promote any data sharing, is resistant to membership inference attacks, linkage attacks, and backdoor attacks and also satisfies differential privacy guarantees. This work is compared with some existing works and centralized models on the PEMS Dataset. Our results are comparable to the centralized ML algorithms yet preserve the privacy of the client's data. In the future, we would like to investigate more about the impact of different experiment settings on the proposed framework. We will also consider taking into account the spatial and weather information into account while traffic flow prediction.

References

1. Davis, G.A., Nihan, N.L.: Nonparametric regression and short-term freeway traffic forecasting. J. Transp. Eng. **117**(2), 178–188 (1991). https://doi.org/10.1061/(ASCE)0733-947X(1991)117:2(178)
2. Dwork, C., McSherry, F., Nissim, K., Smith, A.: Calibrating noise to sensitivity in private data analysis. In: Halevi, S., Rabin, T. (eds.) Theory of Cryptography, pp. 265–284. Springer, Berlin Heidelberg, Berlin, Heidelberg (2006). https://doi.org/10.1007/11681878_14
3. Herzen, J., et al.: Darts: User-friendly modern machine learning for time series. J. Mach. Learn. Res. **23**(124), 1–6 (2022). http://jmlr.org/papers/v23/21-1177.html
4. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735
5. Konecny, J., McMahan, H.B., Ramage, D., Richtárik, P.: Federated optimization: distributed machine learning for on-device intelligence (2016). https://arxiv.org/abs/1610.02527
6. Li, C., Xu, P.: Application on traffic flow prediction of machine learning in intelligent transportation. Neural Comput. Appl. **33**(2), 613–624 (2020). https://doi.org/10.1007/s00521-020-05002-6
7. Lim, B., Arik, S.O., Loeff, N., Pfister, T.: Temporal fusion transformers for interpretable multi-horizon time series forecasting. Int. J. Forecast. **37**, 1748–1764 (2020)

8. Liu, Y., Zhang, S., Zhang, C., Yu, J.J.: FedGRU: privacy-preserving traffic flow prediction via federated learning. In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), pp. 1–6 (2020). https://doi.org/10.1109/ITSC45102.2020.9294453
9. Ma, C., Dai, G., Zhou, J.: Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM BILSTM method. IEEE Trans. Intell. Transp. Syst. **23**(6), 5615–5624 (2022). https://doi.org/10.1109/TITS.2021.3055258
10. McMahan, B., Moore, E., Ramage, D., Hampson, S., Arcas, B.A.y.: Communication-Efficient Learning of Deep Networks from Decentralized Data. In: Singh, A., Zhu, J. (eds.) Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, vol. 54, pp. 1273–1282. PMLR (2017)
11. Miglani, A., Kumar, N.: Deep learning models for traffic flow prediction in autonomous vehicles: a review, solutions, and challenges. Veh. Commun. **20**, 100184 (2019). https://doi.org/10.1016/j.vehcom.2019.100184. https://www.sciencedirect.com/science/article/pii/S2214209619302311
12. Qi, Y., Hossain, M.S., Nie, J., Li, X.: Privacy-preserving blockchain-based federated learning for traffic flow prediction. Futur. Gener. Comput. Syst. **117**, 328–337 (2021). https://doi.org/10.1016/j.future.2020.12.003
13. Shekhar, S., Williams, B.: Adaptive seasonal time series models for forecasting short-term traffic flow. Transp. Res. Rec. **2024**, 116–125 (2008). https://doi.org/10.3141/2024-14
14. Sun, S., Zhang, C., Yu, G.: A bayesian network approach to traffic flow forecasting. IEEE Trans. Intell. Transp. Syst. **7**(1), 124–132 (2006). https://doi.org/10.1109/TITS.2006.869623
15. Tang, H., Xue, N., Wang, G.: Differentially private decentralized traffic flow prediction approach based on federated learning. In: Proceedings of the 2022 10th International Conference on Information Technology: IoT and Smart City, pp. 280–285. ICIT 2022, Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3582197.3582244
16. Vaswani, A., et al.: Attention is all you need. In: 31st Conference on Neural Information Processing Systems (NIPS 2017) (2017)
17. Xia, D., et al.: A distributed WND-LSTM model on mapreduce for short-term traffic flow prediction. Neural Comput. Appl. **33**(7), 2393–2410 (2020). https://doi.org/10.1007/s00521-020-05076-2
18. Yang, Q., Fan, L., Yu, H. (eds.): Federated Learning. LNCS (LNAI), vol. 12500. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-63076-8
19. Yang, Q., Liu, Y., Chen, T., Tong, Y.: Federated machine learning: concept and applications. ACM Trans. Intell. Syst. Technol. **10**(2), 1–19 (2019). https://doi.org/10.1145/3298981
20. Zhang, J., Zheng, Y., Qi, D.: Deep spatio-temporal residual networks for citywide crowd flows prediction. In: AAAI 2017: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, pp. 1655–1661 (2017). https://doi.org/10.1609/aaai.v31i1.10735. https://ojs.aaai.org/index.php/AAAI/article/view/10735
21. Zhang, J., Zheng, Y., Qi, D., Li, R., Yi, X.: DNN-based prediction model for spatio-temporal data. In: Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. SIGSPACIAL 2016, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2996913.2997016, https://doi.org/10.1145/2996913.2997016