Robust Federated Learning Approach for Travel Mode Identification from Non-IID GPS Trajectories

Yuanshao Zhu[1,2], Shuyu Zhang[1,2], Yi Liu[1,2], Dusit Niyato[3], James J.Q. Yu[1,2,*]
[1] *Department of Computer Science and Engineering, Southern University of Science and Technology*
[2]*Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation*
[3]*School of Computer Science and Engineering, Nanyang Technological University*
{*yasozhu,shuyuzhangcn*}*@gmail.com, 97liuyi@ieee.org, dniyato@ntu.edu.sg, yujq3@sustech.edu.cn*

***Abstract*—GPS trajectory is one of the most significant data sources in intelligent transportation systems (ITS). A simple application is to use these data sources to help companies or organizations identify users' travel behavior. However, since GPS trajectory is directly related to private data (e.g., location) of users, citizens are unwilling to share their private information with the third-party. How to identify travel modes while protecting the privacy of users is a significant issue. Fortunately, Federated Learning (FL) framework can achieve privacy-preserving deep learning by allowing users to keep GPS data locally instead of sharing data. In this paper, we propose a Roust Federated Learning-based Travel Mode Identification System to identify travel mode without compromising privacy. Specifically, we design an attention augmented model architectures and leverage robust FL to achieve privacy-preserving travel mode identification without accessing raw GPS data from the users. Compared to existing models, we are able to achieve more accurate identification results than the centralized model. Furthermore, considering the problem of non-Independent and Identically Distributed (non-IID) GPS data in the real-world, we develop a secure data sharing strategy to adjust the distribution of local data for each user, thereby the proposed model with non-IID data can achieve accuracy close to the distribution of IID data. Extensive experimental studies on a real-world dataset demonstrate that the proposed model can achieve accurate identification without compromising privacy and being robust to real-world non-IID data.**

***Keywords*-Travel mode identification, Federated learning, GPS trajectory, Deep learning, Convolutional neural network.**

I. Introduction

Travel mode identification is a critical component of the intelligent transportation system (ITS), mainly used for urban traffic modeling, management, and planning. An accurate and efficient travel mode identification system is essential for governments, companies, and institutes to better understand people's travel behavior and improve transportation modeling, planning, and operating [1]. Information about travel modes is usually collected through offline or online surveys, which are costly and inefficient due to incorrect or incomplete answers and low response rates [2]. In the last few decades, with the popularization of smart devices and advanced data collection approaches, especially the Global Positioning System (GPS), more travel information is recorded accurately and completely, i.e., time, location, and track. These trajectories provide a new idea for inferring the user's travel mode [3].

In the study of travel mode identification based on GPS trajectory data, the typical methods generally adopt the two-step paradigm as in [4]. In detail, organizations or companies need to collect large amounts of raw data at first. They then use feature engineering methods to clean and normalize the original data and fed it into the centralized classification models. Inspired by the above work, researchers used Convolutional Neural Networks (CNN) [2], Recurrent Neural Networks (RNN), or their variants [5] to achieve satisfactory results in this task. In addition, they are also working on using new models or attributes to identify travel modes more accurately [6], [7].

However, there are still research gaps in current GPS-based travel mode identification methods. On the one hand, current methods need to directly upload the original data to the company or organization for centralized training [8]. On the other hand, GPS track information is strongly connected to everyone's private data (e.g., location, travel time). Furthermore, with the increase in user privacy awareness and stricter government supervision, it is difficult for a company or organization to obtain the original GPS data directly [9]. Therefore, some researches have begun to focus on privacy protection and propose some secure models in travel mode identification. However, their models cannot protect privacy without losing accuracy [10].

To address this issue, we need to develop a model that protects privacy without compromising accuracy. Fortunately, federated learning (FL) paradigm-based deep learning model

This work was supported by the General Program of Guangdong Basic and Applied Basic Research Foundation No. 2019A1515011032, by the Department of Education of Guangdong Province, China No. SJJG201901, and by the Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation No. 2020B121201001.

James J.Q. Yu is the corresponding author.

978-1-7281-9074-7/20/$31.00

is a promising solution. The fundamental idea of federated learning is to keep user data locally instead of sharing raw data, and collaboratively train a global model [11]. This idea has been adopted in ITS for traffic flow prediction [12]. Nevertheless, their study did not use GPS data and did not research on travel mode identification. To close the research gaps in privacy protection of the existing travel mode identification methods, we propose a federated learning-based travel mode identification system to protect privacy. In the proposed system, we design a CNN model based on the attention mechanism [13], which is used for travel mode classification and attribute feature extraction. Besides, considering the non-independent and identically distributed (non-IID) travel modes of different users, we design a data sharing strategy that allows users to adjust their data distribution dynamically. The main contributions of this paper are summarized as follows:

- Unlike existing methods, we propose a privacy-preserving travel mode identification algorithm that combines federated learning and neural networks to identify travel modes accurately without compromising privacy.
- We propose a data sharing strategy that can effectively improve the robustness of the proposed model for non-Independent and Identically Distributed (non-IID) problem in federated learning.
- We propose an attention mechanism-based CNN model to identify travel modes. This model uses the attention mechanism to extract the fine-grained features of GPS data to achieve high-precision travel pattern classification.

The rest of this paper is organized as follows. Section II reviews relevant research on travel mode identification and privacy in this task. Section III presents the privacy-related issues and challenges of using federated learning for travel mode identification. Section IV details the methods used in this paper, including data pre-processing steps, federate learning algorithms based on data sharing, and the neural network model the proposed approach adopted. Section V analyzed relevant experimental results. Finally, we conclude the paper in Section VI.

II. Related Work

A. Travel Mode Identification

Travel mode identification is one of the most significant fields of ITS, which has attracted the attention of many researchers. The travel mode identification methods based on GPS data can be divided into two categories: machine learning-based and neural network-based methods.

Machine learning-based method: The machine learning-based approaches achieve travel mode identification by using machine learning algorithms (e.g., decision tree, support machine vector) to classify input data (e.g., GPS trajectory, location data) [4]. For example, Zheng *et al.* [4], [14] proposed a change-point-based segmentation algorithm with decision tree to identify travel mode by using GPS trajectory data. Specifically, they extracted the indicators from GPS trajectory data based on each road section's characteristics, including the mean and variance of speed and the expected value of speed. These indicators were used as input to the classification algorithm to infer travel modes. Inspired by the above work, many researchers have turned their attention to the use of machine learning-based methods to classify travel modes.

Neural network-based method: Due to the development of deep learning technology, some researchers began to focus on using neural networks to identify travel modes by extracting high-level data features. For example, Wang *et al.* [15] utilized a sparse autoencoder to transform point-level features to obtain high-dimensional features and used DNN model to learn these high-dimensional features for traffic mode identification. Although their work used neural networks to extract high-level data features, they cannot extract non-linear features in a fine-grained manner [4]. To solve this problem, Dabiri *et al.* [2] used a CNN network to extract non-linear features at a fine-grained level. Specifically, the authors performed data pre-processing to obtain data features by following [4], [14] and stack these feature vectors into a structure that can be accepted by the CNN model training. This method further improves the accuracy of classification, but cannot capture the time dependence of the GPS trajectory. Therefore, many researchers develop recurrent neural networks (RNN) and their variants (e.g., long short-term memory (LSTM)) for time-dependent capture. Jeyakumar *et al.* [7] and Liu *et al.* [16] utilized a convolutional bi-LSTM network to capture the time dependence of GPS trajectories to achieve better classification results. Furthermore, Yu *et al.* [17] combined the LSTM network with discreet Fourier transform (DFT) and discreet wavelet transform (DWT) to obtain both time-domain and frequency-domain attributes.

Inspired by the above work, feature extraction is an essential step in the task of travel mode identification directly affecting the accuracy of identification. To extract fine-grained features, we propose the attention augmented CNN model to extract features in a fine-grained manner [13]. The reason is that convolution operation is limited by its locality and lacks an understanding of the global context. Therefore, we combine the attention mechanism and the CNN model to utilize the attention mechanism's ability to capture the long time-series information.

B. Privacy Issues For Travel Mode Identification

In recent years, with the development of sensor technologies and data collection technologies, more and more data resources (GPS data, mobile signals, etc.) have provided significant support for the travel mode identification research. However, these data resources, especially GPS data, are

closely related to personal privacy (i.e., location). We cannot directly collect user GPS data due to privacy concerns, which poses a challenge to traditional methods for identifying travel modes by aggregating raw data. Therefore, researchers started to pay attention to protecting users' private data when researching ITS. For example, Zhou *et al.* [10] proposed a privacy-persevering scheme that allows the collection of total traffic flow data while preserving individual vehicles' privacy to monitor traffic flow. Hoh *et al.* [18] presented a system based on virtual trip lines and the corresponding cloaking technique, which can perform travel time estimates without knowing the actual geographic locations of trip lines. Although they have made many privacy protection efforts, these methods did not achieve great results and even lost the accuracy of the model.

To address this problem, we need to develop a model that protects privacy without compromising accuracy. Fortunately, a federated learning paradigm-based deep learning model is a promising solution. For example, Liu *et al.* [12] introduced the FL framework to tackle the traffic flow prediction problem and proposed a federated learning-based Gated Recurrent Unit neural network (FedGRU) algorithm to predict traffic flow without compromising privacy. Lu *et al.* [19] used the FL framework to design a two-stage scheme to solve the data leakage problem of multiple participants and multiple transmission channels in Vehicular Cyber-Physical Systems.

Although they used the FL framework to address privacy concerns, there are many real-world problems not considered. For example, in the task of travel mode identification, the common travel mode of various users has significant differences. In this paper, we not only use FL to protect privacy but also design a secure public data sharing strategy to solve the non-IID data distribution problem in the real world, which will be introduced in Section IV-B.

III. Problem Definition

In this paper, we use the term 'client' to define each user in the FL framework, and the term 'server' to define the entity that aggregates clients' parameters. Let $\mathcal{K} = \{k_1, k_2, \ldots, k_n\}$ and $\mathcal{D} = \{D_1, D_2, \ldots, D_n\}$ denote the set of client and client's dataset in travel mode identification. Each client participates in FL framework computation as an independent node. We aim to train a global model $\hat{y} = f(x)$ at the server, and the server sends this model to the client, which can identify each client's travel mode without compromising privacy, where x is the client's travel feature, and $\hat{y}$ is the predicted category. First, we define the privacy protection problem in travel mode identification as follows:

Definition 1. ***(Privacy Sensitivity** δ**):*** *Privacy sensitivity is defined to measure the degree of leakage of user privacy. Specifically, D denotes the user's local dataset and D_{access} represents the amount of data that the user needs to be accessed by a third party, so the privacy sensitivity δ can be formally defined as follows:*

$$\delta = \frac{||D_{access}||}{||D||}. \tag{1}$$

According to Definition 1, our goal is to classify the travel mode when δ is minimized. In this paper, we use the FL framework to achieve this goal. The federated travel mode identification can be formally defined as follows.

Definition 2. ***(Federated Travel Mode Identification):*** *Given the input samples $x_k \in D_k$ at the client k, the optimization objective function is defined as follows:*

$$\begin{cases} \min\limits_{\omega} F(\omega) = \arg\min\limits_{\omega} \sum\limits_{i=1, x_k \in D_k}^{K} f_k(x_k; \omega_i) \\ \delta \to 0 \end{cases}, \tag{2}$$

where $f_k(\cdot)$ is the local loss function at the client k and δ is the privacy sensitivity. Definition 2 indicates that users cannot optimize the common loss function by sharing data. Therefore, we need to convert Eq. (2) *into a distributed optimization problem, which is exactly what the FL framework can solve.*

In the context of travel mode identification, we will face the following challenges due to different data distribution and collection environments:

Challenge 1. ***(Statistical Heterogeneity [20]):*** *Clients' data may be generated by users using different vehicles, for example, some users often take buses, and other users often drive, which may result in the attributes and distribution of data between users are different. It violates the assumption of independent and identically distributed (IID) that commonly used in distributed training, which will lead to difficulties in modeling and optimization [21].*

Challenge 2. ***(System Heterogeneity [20]):*** *The clients participating in federated learning may use different operating systems (e.g., Android, IOS), heterogeneous devices (e.g., mobile phone, smartwatch), and different network conditions (4G, 5G, Wi-Fi, etc.) Additionally, there is no guarantee that every client is active and always connected to the server in each learning round. Such various characteristics of the system-level carry great challenges to the convergence of the global model.*

IV. Methodology

Existing approaches for travel mode identification follow the two-step framework proposed by Zheng *et al.* [14], [22]. In detail, the server calculated each GPS segment's motion features and fed them into a classification algorithm (e.g., CNN) for mode identification [2], [23], [5], [17]. However, these methods need to access the original GPS

data and follow the assumption of independent and identical distribution. In short, these methods did not consider the challenges of privacy protection and distributed optimization in travel mode identification. To solve these issues, we adopt a federated learning framework to address data privacy concerns and design a public data sharing strategy for non-IID data distribution where the proposed mechanism can significantly enhance the performance of the model on non-IID data.

In this section, we first present the data processing techniques. Then we elaborate on the federated averaging algorithm (FedAVG) and the proposed public data sharing strategy. Finally, briefly introduce the employed neural network model.

A. Processing of GPS Record

GPS raw data is the user's travel information collected by GPS collection devices within a period, i.e., longitude, latitude, and sampling timestamp. Therefore, we can divide the original GPS data with the same travel mode into a trip based on its timestamp. We then calculate multiple motion features according to the geographic coordinates and timestamp of each GPS point in the trip segment.

Let $\mathcal{R} = \{R_1, R_2, \ldots, R_M\}$ represent the GPS record in the segment with length M. Each GPS record is represented by a triple $R_i = \langle \mathrm{lat}_i, \mathrm{long}_i, \mathrm{t}_i \rangle$ as the latitude (lat_i) and longitude (long_i) of the device's location at the time of t_i. For two consecutive records R_i, R_{i+1}, Vincenty formula [24] can be used to calculate the relative distance:

$$RD_i = \text{Vincenty}\left(\mathrm{lat}_i, \mathrm{long}_i; \mathrm{lat}_{i+1}, \mathrm{long}_{i+1}\right). \quad (3)$$

Denoting the time interval between R_i and R_{i+1} as Δt_i, based on the relative distance (RD_i), we can calculate the speed (S_i), acceleration (A_i) and jerk (J_i) of the R_i location. These motion features can be calculated by using the following equations:

$$S_i = \frac{RD_i}{\Delta t_i}, \quad 1 \le i \le M, \quad S_M = S_{M-1}, \quad (4)$$

$$A_i = \frac{S_{i+1} - S_i}{\Delta t_i}, \quad 1 \le i \le M, \quad A_M = 0, \quad (5)$$

$$J_i = \frac{A_{i+1} - A_i}{\Delta t_i}, \quad 1 \le i \le M, \quad J_M = 0. \quad (6)$$

In addition, some researchers have also introduced attributes (e.g., bearing rate, DFT, DWT) to boost the model [2], [17]. However, we use only the four attributes mentioned above, namely relative distance, speed, acceleration, and jerk rate. For each segment, we extract the motion feature, arrange it according to the time sequence, and fix the feature vector's length. Then we stack these vectors into a 4-channel tensor, which is used as the input of the proposed travel mode identification model.

Algorithm 1 Federated Averaging (FedAVG) Algorithm with Data Sharing

Input:
Client set is $\mathcal{K} = \{k_1, k_2, \ldots, k_n\}$, clients' dataset are $\mathcal{D} = \{D_1, D_2, \ldots, D_n\}$, server's public dataset is $\mathcal{D}_g$, the sharing rate is α, the local mini-batch size is B, the number of local epochs is E, the fraction of clients on each round is C, and η is the learning rate.

Output:
Parameter w.

for each client $k_i \in \mathcal{K}$ **do**
  Send global model w_0 to k_i
  Pre-train the model of k_i using D_i
  Sample from $\mathcal{D}_g$ with probability α to get D_s and send it to k_i
  Use the client model to select required data from D_s, and combine with D_i to obtain new client dataset D_i^*
end for
for each round $t = 1, 2, \ldots, T$ **do**
  $m \leftarrow \max(C \cdot n, 1)$
  $\mathcal{K}_t \leftarrow$ (random set of m clients)
  for each client $k \in K_t$ in parallel **do**
    Initialize $w_{t+1}^k \leftarrow \mathrm{ClientUpdate}\ (k, w_t)$
    $w_{t+1} \leftarrow \frac{1}{|\mathcal{K}_t|} \sum_{k \in \mathcal{K}_t} w_t^k$
  end for
end for
ClientUpdate (k, w)
$\mathcal{B} \leftarrow$ (split D_k^* into batches of size B)
for each local epoch i from 1 to E **do**
  for batch $b \in \mathcal{B}$ **do**
    $w \leftarrow w - \eta \nabla \ell(w; b)$
  end for
end for
return w to server.

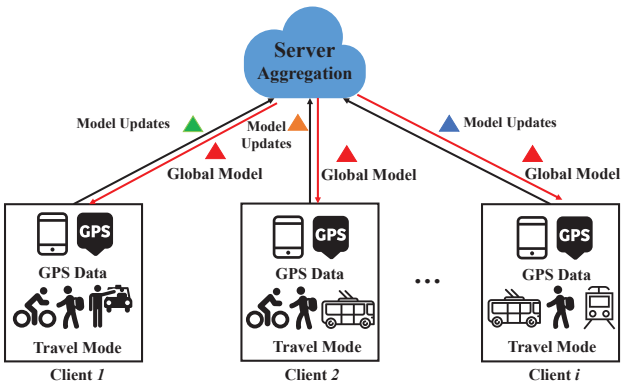

Figure 1. Federated learning-based travel mode identification architecture.

B. Federated Learning with Data Sharing Strategy

In the case of conventional centralized methods, since all

users' data can be accessed, privacy-related issues are not considered. On the other hand, in federated learning the each user's private data will not be shared, and the living styles of each user are different, which makes the data inconsistent with the expectation of independent and identical distribution. To address this problem, we proposed a data sharing-based FedAVG algorithm, as shown in Algorithm 1. This method will not violate the setting of FL and compensate for the drawbacks of uneven data distribution and lack of data in the client.

*1) **FedAVG Algorithm**:* The FedAVG algorithm can effectively reduce overhead communication between the server and the client. As shown in Fig. 1, we suppose there are n clients in FL framework, and the private data of each customer k is D_k. The main steps of FedAVG algorithm are as follows:

1) **Step1:** For the round of training t, with a percentage $C \in [0, 1]$, randomly select $C \cdot n$ clients to participate in training. The selected clients initialize the local model according to the model parameter w^t broadcast by the server.
2) **Step2:** Each client k trains E epochs with local data. For each epoch, client conducts gradient optimization by batch-size B. The goal of client k is to find optimal parameters w_k^* minimize its local loss function $J_k(w)$, that is:

$$w_k^* = \arg\min_{w_k} J_k(w_k). \tag{7}$$

3) **Step3:** The server aggregates the parameters of all clients participating in the training through the secure parameters aggregation method and updates the global model. At last, the server sends parameters back to the client for the next training round. The server wishes to minimize the global loss function:

$$J(w_g) = \frac{1}{n} \sum_{k=1}^{n} J_k(w_k). \tag{8}$$

Repeat the above steps until the global loss function converges.

*2) **Data Sharing Strategy**:* Since the data between users exists in the form of non-IID, this will make it difficult for the federated global model to achieve high-precision results and convergence. We need to find an efficient method to alleviate the non-IID distribution problem of data. We have the key insight that we use data sharing strategy to adjust user data distribution, thereby mitigating this problem.

Let $L = \{l_1, l_2, \ldots, l_c\}$ as the number of travel modes for each category, where c represents the total number of categories. We can define the non-IID data distribution characteristics as:

$$\tau = \frac{\max(L) - \min(L)}{\sum L}. \tag{9}$$

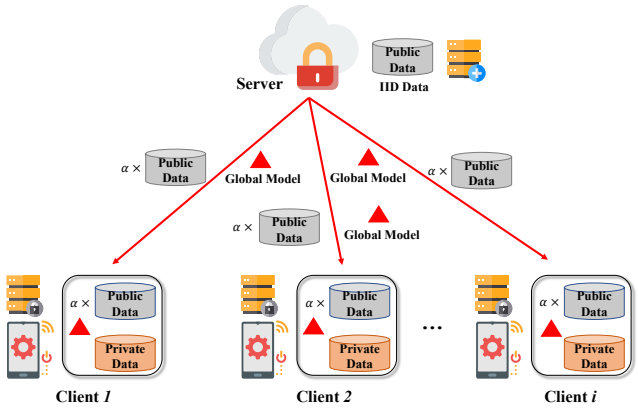

Figure 2. Overview of the federated learning with data sharing strategy.

The goal of the data sharing strategy is to make $\tau \to 0$, That is, $l_1 \approx l_2 \approx \cdots \approx l_c$. As shown in Fig. 2, the concrete implementation of this strategy involves the following three stages:

1) **Server preparation:** Before conducting the FL training task, the server uses the public dataset to pre-train a global model. At the same time, the server samples with random probability α to obtain a subset D_s, where $D_s \subset \mathcal{D}_g$, in which each category is uniformly distributed.
2) **Transmission of data and parameter:** The server sends the pre-trained global model parameters and D_s to each client participating in the FL framework by a secure mechanism.
3) **Client processing:** After the client receives the model parameters and the subset D_s, it selects the missing or fewer data from D_s according to its own private data distribution and obtains $D_c \subset D_s$. Then the client gets a new dataset D_k^*, where, $D_k^* = D_k \bigcup D_c$. All subsequent training steps in FL use the new dataset D_k^*.

C. Attention Augmented CNN Model

Based on the efficiency of CNN for image feature extraction, Dabiri *et al.* [2] constructed the raw GPS trajectory as a multi-channel tensor by data processing method, which can be used for CNN training. Such a CNN-based method achieved great success in travel mode identification. However, the existing CNN architectures [2], [23] for identifying travel modes use too many convolutional layers, which increase the computational burden and do not improve the accuracy of the model. Furthermore, CNN suffers from the limitation of its local receptive field and loses much global information.

As CNN cannot capture global correlation, in this paper, we introduce attention augmented convolution to extract raw GPS trajectory's global features. The reason is that attention augmented convolution has a powerful ability to

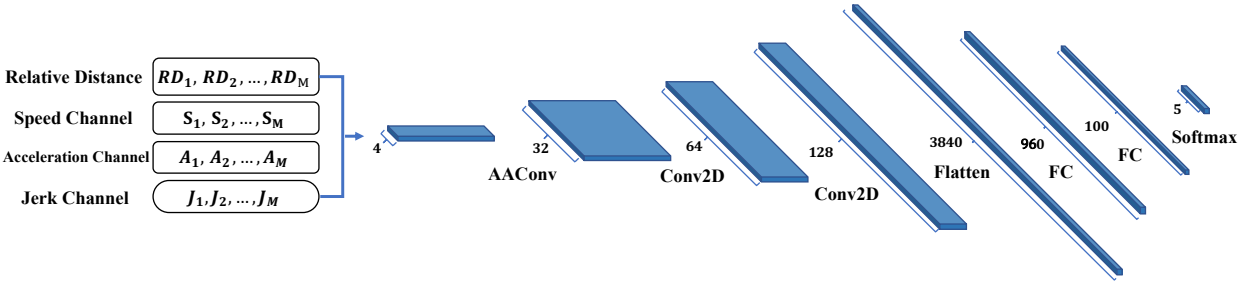

Figure 3. Structure of attention augmented CNN model.

capture long-range interactions [13], which is suitable for obtaining GPS motion trajectory information. To concrete, we simplified the existing CNN architecture and removed many unnecessary convolutional layers. Considering that the local receptive field of the convolutional layer cannot capture the global correlation, we combine the convolutional neural network with the self-attention mechanism and use the attention augmented convolutional layer as the first layer. More specifically, this model uses three convolutional layers to extract features, where the first layer is an augmented convolutional layer, and the next two layers are the traditional convolutional layer. The number of channels in each convolutional layer is 32, 64, 128, respectively. After obtaining the extracted high-dimensional features, we expand them and add two fully connected layers with 960 and 100 neurons. The last layer uses 5 neurons with the softmax activation function for classification. The proposed architecture of this model is shown in Fig. 3.

V. Experiment

To fully evaluate our proposed model's performance on travel mode identification, we applied the model on a real-world dataset to conduct a comprehensive study. In this section, we first investigated the performance of the proposed FL-based model and the previous model. Then under the FL framework, we compared the accuracy of the standard FedAVG algorithm and our data sharing method under different data distributions. Finally, we analyzed the effect of various parameters set on the model.

A. DataSet and Model Parameter Configurations

1) ***Dataset****:* In this work, we employ the read-world data from the Geolife GPS trajectory dataset [4] for investigation, which contains the travel trajectories of 182 users about five years. We adopt 69 users' raw trajectory data with five travel modes and apply the processing method mentioned in Section IV-A to regulate the GPS records. This work mainly considered five real travel modes for identification, i.e., walking, biking, bus, driving, and train. We first split 8% of all data into the public dataset, and for the remaining data, use 80% as the training set and 20% as the testing set.

2) ***Model Parameter Configurations****:* We distribute the training data to 50 clients, i.e., $K = 50$. Each client has at most two travel modes to simulate the distribution of non-IID data in the real world as much as possible. By default, we set $C = 0.2$, sharing rate $\alpha = 50\%$, local batch size $B = 30$, and the learning rate $\eta = 0.0005$. Each client uses Adam optimizer to train the proposed model locally for E ($E = 10$) epochs. All simulations are developed in Python and Pytorch [25]. All case studies are conducted on a computing server with an nVidia GeForce RTX2080 Ti GPU and Intel(R) Xeon(R) Sliver CPU.

B. Performance of the Proposed Framework

In this section, we evaluate the performance of the proposed model by comparing it in an identical configuration with the previous centralized models, i.e., CNN [2], image-based DNN [6], Graph-based decision tree (DT) [14], k-nearest neighborhood (KNN), support vector machine (SVM), and random forest (RF) [26]. Note that all baseline algorithms are applied to the same dataset. As shown in Table I, the proposed model can identify travel modes more accurately than the previous centralized models. Specifically, the accuracy of the proposed model is 1.3% higher than that of the centralized CNN model. The reason is that the proposed model uses attention augmented CNN, which has better global spatial feature extraction capabilities in travel mode identification tasks. Furthermore, the federated learning-based model is more accurate than the centralized methods. This is because the federated learning based-model inherits the fine-grained feature extraction capabilities of the centralized attention augmented CNN model. Besides, in the federated learning framework, the server needs to make an accuracy trade-off between each local and global models, which results in losing accuracy.

However, centralized methods need to upload GPS data to the server, which cannot meet our privacy protection requirements. With the federated learning framework, each user's private data will not be shared. In short, the model using federated learning can make privacy sensitivity $\delta = 0$. Table I shows that the federated learning-based mode's accuracy is close to the centralized model, which fully

Table I
PERFORMANCE COMPARISON OF TRAVEL MODE IDENTIFICATION APPROACHES

Approach	Accuracy (%)
Proposed Model (Federated)	**85.1**
Proposed Model (Centralized)	**86.7**
CNN [2]	84.8
Image-based DNN [6]	67.9
RF [26]	82.0
Graph-based DT [14]	76.2
KNN [2]	63.5
SVM [2]	65.4

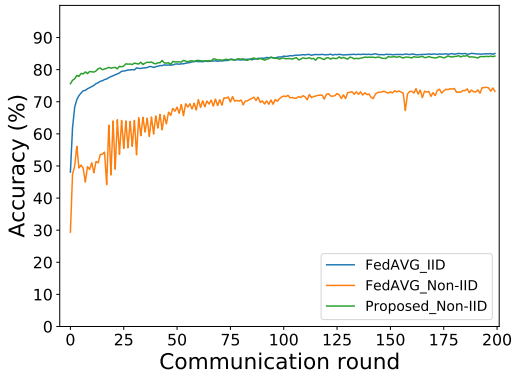

Figure 4. Performance comparison between FedAVG with IID data, FedAVG with Non-IID data and our sharing strategy with Non-IID data.

demonstrates that the proposed model can perform accurate travel mode identification without compromising privacy.

C. Performance Comparison under Non-IID Data

In this section, we will explore the impact of non-IID data distribution on the proposed model. As shown in Fig. 4, we observe that the FedAVG algorithm performs poorly on non-IID data and cannot converge quickly. This is because non-IID data distribution increases the parameter differences between various clients, which poses challenges to parameter aggregation and optimization. However, there is no significant difference in the performance of the proposed model under non-IID and IID data distribution, and it can greatly outperform the traditional FedAVG algorithm. It is shown that the proposed data sharing strategy can effectively compensate for the shortcomings of non-IID data distribution. In addition, the proposed model uses a pre-train method on the server to achieve converge faster.

D. Performance Comparison with Different Client Numbers and Sharing Rates

In this section, we investigate the impact of different client numbers and data sharing rates on the performance of the proposed model. As is shown in Fig. 5 (a), we can find that the increase in the number of clients has an adverse impact on the model's performance. The reason is

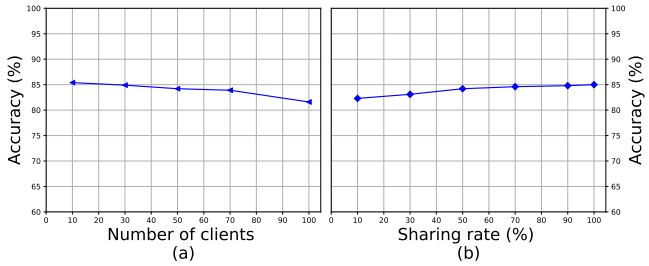

Figure 5. The performance of our model with (a) different numbers of client numbers and (b) different sharing rate.

that the increase in the number of clients will increase the diversity of data and more differences in parameters making the parameter aggregation more difficult. Moreover, we select $\alpha = \{10\%, 30\%, 50\%, 70\%, 90\%, 100\%\}$ for different sharing rates simulation. The experimental results are shown in Fig. 5 (b), it can be observed that an increase in α can improve the accuracy of the model, but when α increases to a certain level, the improvement effect is not significant. This is because the data sharing strategy allows clients to select only the missing data rather than using all the received data. On the contrary, too much data sharing would increase the communication overhead, so set $\alpha = 50\%$ is a good choice.

VI. CONCLUSION

In this paper, we propose a roust federated learning-based travel mode identification system that can accurately infer the user's travel mode without compromising privacy. Experimental results show that the proposed model's performance is better than the advanced centralized CNN model. Considering that the distribution of GPS data in real life usually exists in the form of non-IID, we propose a secure public data sharing strategy to alleviate this issue. Specifically, we adjust the distribution of GPS data by sharing public data to improve the model performance on non-IID data. In addition, we design an attention augmented mechanism-based CNN model to extract the features of GPS data in a fine-grained manner, whose performance exceeds the current advanced models. We conduct extensive experiments to demonstrate that the proposed model can perform accurate identification under the premise of protecting user privacy and achieve excellent results for non-IID data distribution.

To the best of our knowledge, this is the pioneering work on travel mode identification with federated learning. In the future, we plan to combine the LSTM and federated learning framework to better capture long-term dependencies in GPS trajectory and further improve identification performance.

REFERENCES

[1] A. C. Prelipcean, G. Gidofalvi, and Y. O. Susilo, "Transportation mode detection – an in-depth review of applicability and

reliability," *Transport Reviews*, vol. 37, no. 4, pp. 442–464, 2017.

[2] S. Dabiri and K. Heaslip, "Inferring transportation modes from gps trajectories using a convolutional neural network," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 360–371, 2018.

[3] L. Wu, B. Yang, and P. Jing, "Travel mode detection based on gps raw data collected by smartphones: a systematic review of the existing methodologies," *Information-an International Interdisciplinary Journal*, vol. 7, no. 4, pp. 67–86, 2016.

[4] Y. Zheng, L. Liu, L. Wang, and X. Xie, "Learning transportation mode from raw GPS data for geographic applications on the web," in *Proceedings of the 17th International Conference on World Wide Web*. New York, NY, USA: Association for Computing Machinery, 2008, pp. 247–256.

[5] J. J. Yu, "Travel mode identification with gps trajectories using wavelet transform and deep learning," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2019.

[6] Y. Endo, H. Toda, K. Nishida, and A. Kawanobe, "Deep feature extraction from trajectories for transportation mode estimation," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2016, pp. 54–66.

[7] J. V. Jeyakumar, E. S. Lee, Z. Xia, S. S. Sandha, N. Tausik, and M. Srivastava, "Deep convolutional bidirectional lstm based transportation mode recognition," in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, 2018, pp. 1606–1615.

[8] J. Kang, Z. Xiong, D. Niyato, Y. Zou, Y. Zhang, and M. Guizani, "Reliable federated learning for mobile networks," *IEEE Wireless Communications*, vol. 27, no. 2, pp. 72–80, 2020.

[9] W. Shen, B. Yin, Y. Cheng, X. Cao, and Q. Li, "Privacy-preserving mobile crowd sensing for big data applications," in *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.

[10] Y. Zhou, Z. Mo, Q. Xiao, S. Chen, and Y. Yin, "Privacy-preserving transportation traffic measurement in intelligent cyber-physical road systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3749–3759, 2015.

[11] B. McMahan, 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*, 2017, pp. 1273–1282.

[12] Y. Liu, J. J. Q. Yu, J. Kang, D. Niyato, and S. Zhang, "Privacy-preserving traffic flow prediction: A federated learning approach," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7751–7763, 2020.

[13] I. Bello, B. Zoph, Q. Le, A. Vaswani, and J. Shlens, "Attention augmented convolutional networks," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 3285–3294.

[14] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, "Understanding mobility based on gps data," in *Proceedings of the 10th international conference on Ubiquitous computing*, 2008, pp. 312–321.

[15] H. Wang, G. Liu, J. Duan, and L. Zhang, "Detecting transportation modes using deep neural network," *IEICE TRANSACTIONS on Information and Systems*, vol. 100, no. 5, pp. 1132–1135, 2017.

[16] H. Liu and I. Lee, "End-to-end trajectory transportation mode classification using bi-lstm recurrent neural network," in *2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, 2017, pp. 1–5.

[17] J. J. Yu, "Semi-supervised deep ensemble learning for travel mode identification," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 120–135, 2020.

[18] B. Hoh, M. Gruteser, R. Herring, J. Ban, D. Work, J.-C. Herrera, A. M. Bayen, M. Annavaram, and Q. Jacobson, "Virtual trip lines for distributed privacy-preserving traffic monitoring," in *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services*. New York, NY, USA: Association for Computing Machinery, 2008, pp. 15–28.

[19] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Federated learning for data privacy preservation in vehicular cyber-physical systems," *IEEE Network*, vol. 34, no. 3, pp. 50–56, 2020.

[20] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020.

[21] J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 700–10 714, 2019.

[22] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Towards mobile intelligence: Learning from gps history data for collaborative recommendation," *Artificial Intelligence*, vol. 184–185, pp. 17–37, 2012.

[23] S. Dabiri, C. Lu, K. Heaslip, and C. K. Reddy, "Semi-supervised deep learning approach for transportation mode identification using gps trajectory data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 5, pp. 1010–1023, 2020.

[24] T. Vincenty, "Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations," *Survey Review*, vol. 23, no. 176, pp. 88–93, 1975.

[25] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *NIPS-W*, 2017.

[26] B. Wang, L. Gao, and Z. Juan, "Travel mode detection using gps data and socioeconomic attributes based on a random forest classifier," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1547–1558, 2017.