IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 7 March 2023, accepted 20 March 2023, date of publication 23 March 2023, date of current version 28 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3260981

RESEARCH ARTICLE

GANs for Privacy-Aware Mobility Modeling

IVAN FONTANA, MARC LANGHEINRICH, (Member, IEEE), AND MARTIN GJORESKI
Faculty of Informatics, Università della Svizzera italiana, 6900 Lugano, Switzerland

Corresponding author: Martin Gjoreski (martin.gjoreski@usi.ch)

This work was supported by the Swiss National Science Foundation through Behavioral Analytics for Smart Environments (BASE) under Project 200021-182109.

ABSTRACT Human mobility modeling is crucial for many facets of our society, including disease transmission modeling and urban planning. The explosion of mobility data prompted the application of deep learning to human mobility. Along with the growth of research interest, there is also increasing privacy concern. This study first examines the cutting-edge approaches for trajectory generation, classification, and next-location prediction. Second, we propose a novel privacy-aware approach for predicting next-week trajectories. The approach is based on two modules, a Generative Adversarial Network used for generating synthetic trajectories and a deep learning model for user identification which safeguards privacy. These two modules are combined with a next-week trajectory predictor that uses privacy-aware synthetic data. The experiments on two real-life datasets show that the generator creates trajectories similar to the real ones yet different enough to safeguard privacy. The low user-recognition recognition accuracy of models trained on the generated data demonstrates privacy awareness. Statistical tests confirm no significant difference between the original and the generated trajectories. We further demonstrate the utility of the synthetic data by predicting week-ahead trajectories based on the synthetic trajectories. Our study shows how privacy and utility can be managed jointly using the proposed privacy-aware approach.

INDEX TERMS Deep learning, generative adversarial networks, location data, machine learning, mobility modeling, privacy.

I. INTRODUCTION

The urban population is rapidly increasing, and human mobility is becoming more complex and voluminous. These changes affect critical aspects of people's lives, such as the spread of viral diseases (COVID-19 pandemic) [1], [2], people's behavior in natural disasters [3], public and private transportation and the resulting traffic volumes [4], and people's well-being [5]. Moreover, migrations from rural to urban regions, such as those caused by natural disasters, climate change, and wars, impact crowd mobility between cities [6].

However, authorities are not defenseless in the face of these problems. The emergence of interconnected devices and platforms, e.g., mobile phones, the Internet of Things, and social media, produce up-to-date and precise mobility data at multiple temporal and spatial dimensions. Examples of mobility data include traces from GPS devices embedded in smartphones [7], [8], modern cars [9], records produced by phone-to-cellular network communication [10], and geo-tagged posts from social media platforms [11]. This stream of digital data drove a massive research output on many aspects of human mobility, such as trajectory data mining [12] and next-location (next-place) prediction [13]. It is critical in many applications, such as advertisement (e.g., travel recommendations and marketing), early warning of possible public hazards, and friend recommendations on social networks. The development of sophisticated Artificial Intelligence (AI) techniques and the availability of large amounts of mobility data provided researchers with the expansive potential to employ Deep Learning (DL) methodologies to address mobility-related problems.

Along with the growth of research interest, there is also increasing concern regarding privacy in handling these types of data, given the many methods capable of tracing back to the source of the data (user) even after data anonymization [14]. Simply eliminating unique identifiers from mobility traces does not preserve privacy. A recent example is Trajectory-User Linking (TUL), which allows one to identify

The associate editor coordinating the review of this manuscript and approving it for publication was Junggab Son.

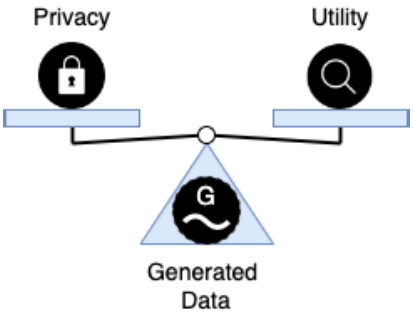


FIGURE 1. Privacy – Utility trade-off in generated data.

which (anonymous) trajectories belong to which user [15]. Thus, it is important to resolve these issues by developing technologies progressively focused on user privacy.

The main challenges in human mobility that we address with this work are twofold:

- **Privacy – Utility trade-off**. Privacy is always a concern when it comes to personal data modeling. However, the process of maximizing data privacy may destroy data utility. This trade-off between privacy and utility is well known [16], but is rarely evaluated with a joint approach. In our work, privacy refers to data anonymization, i.e., the source user cannot be identified from the corresponding mobility data. On the other hand, we define the utility as the amount of information in a dataset with which further modeling can be done. This study explores this trade-off to find the equilibrium where privacy and utility are balanced (Fig. 1).
- **The Impact of Data Sparsity**. Mobility data is already inherently sparse due to the collection methods, e.g., check-ins, on-demand sensing, and geotagging. Furthermore, compared to the often overwhelming number of places accessible, a user visits very few. This study includes one sparse and one dense dataset to analyze the influence of data sparsity.

We start by examining the cutting-edge approaches currently dealing with human mobility trajectory generation, classification, and next-location prediction. We propose a new privacy-aware approach for accurately predicting week-ahead trajectories based on the state-of-the-art (SOTA). Besides the accurate predictions, the approach generates realistic synthetic mobility data (compared to existing datasets) that are used to ensure user privacy (evaluated via the accuracy of a user-identification model).

While the proposed method can be applied for a variety of tasks in privacy-aware mobility modeling, here we explicitly list two examples. The COVID-19 contact-tracing apps raised citizens' privacy concerns in many countries [17], [18]. Having privacy-aware predictive models would benefit both the citizens (e.g., smartphone users by having better privacy conditions) and the authorities (e.g., better end-user compliance and thus better pandemics modeling). The second scenario where the proposed approach can be directly applied is privacy-aware geomarketing. Based on the privacy-aware week-ahead trajectory predictions, smartphone users can receive week-ahead recommendations regarding traffic, restaurants, and shopping places, enabling personalized services for the users and thus improved quality of life.

The main contributions presented in this work include:

- A unified implementation of two SOTA approaches for human mobility modeling, LSTM-TrajGan [16] and MARC [19], enabling usage of these two methods through a single processing pipeline.[1]
- Development of a novel approach for predicting variable-sized next-week trajectories in a privacy-aware manner. The approach is based on the two SOTA approaches, LSTM-TrajGan [16] used for generating synthetic trajectories, and MARC [19] used for user identification. The combination of these two approaches is extended with a next-week trajectory predictor, which provides accurate predictions using only the privacy-aware synthetic data. Unlike the related work, we take into account the two important aspects of mobility modeling, i.e., privacy and utility, through a unified approach.
- Evaluation of the proposed approach on two real-life datasets with respect to: (i) the quality of the synthetic data, including spatial, temporal, and semantic characteristics; (ii) user-privacy, i.e., whether users can be traced back from the synthetic data; (iii) data utility, i.e., whether accurate next-week trajectory predictors can be developed using the privacy-aware synthetic data.

The rest of the paper is structured as follows: In section two, we summarize the related work, including methods for generating data, user identification, and trajectory prediction. In section three, we present the two datasets used in the study, and in section four, we present the proposed approach. In section five, we present the experimental setup and experimental results. We finalize the paper with a discussion section followed by a conclusion section.

A. SUMMARY OF THE RELATED WORK

Human mobility modeling is a complex subject that includes data-processing pipelines and models (e.g., generative or predictive) that try to maximize data utility. Improving the accuracy of the mobility models is one task that many related studies tackle. User privacy is the second task that has gained attention recently. Maximizing data privacy may decrease the data utility and models' accuracy. Most of the related-work studies focus on only one of these components at a time (accuracy or privacy), solving specific tasks or improving existing techniques. Different from the related work, our work aims to include these two important aspects of mobility modeling, i.e., privacy and utility, into a joint approach. This is carried out by (i) using a GAN to generate a synthetic version of the original trajectories; (ii) using a TUL DL model to ensure

[1] https://github.com/fonticode/Privacy-aware-Mobility-Modeling

that the users can not be traced back from the synthetic data; and (iii), using a next-place prediction DL model over the synthetic data to ensure that the generated data is still useful for user mobility modeling.

B. GENERATION OF SYNTHETIC TRAJECTORIES

Generative Model Definition [20]: *"A generative model for human mobility M is any algorithm capable of generating a collection of n synthetic trajectories* $T_M = \{T_{a_1}, \ldots, T_{a_n}\}$, *which describe the motions of n independent agents* $a_1, \ldots, a_n$ *over a given time period."*

Generative models aim to create synthetic trajectories representing real-life patterns in human mobility. Synthetic trajectories must mimic the collection of geographical, temporal, and semantic characteristics of mobility, that is, the distribution of coordinates and spatial distances, the temporal time spans, and the routinely predictable behavior. The model should experimentally represent the time and length of visits and the set of specific places with their meaning. The most difficult challenge is describing users' routines while considering the erratic tendency to explore new places. The synthetic data should represent the routine and predictable nature of human displacements [21], and the occurrence of visits to places far from one another [10]. Regarding geographical criteria, the synthetic trajectories should emulate both classes of users: returners and explorers [22]. The first type tends to travel shorter distances and visit the same set of places, while the latter is characterized by arbitrary behavior and longer distances.

The most recent generative models are based on DL generative frameworks such as GANs and Variational Auto Encoders (VAEs). Both frameworks deploy DL techniques to model the data distributions and generate novel synthetic trajectories based on those distributions. GANs and VAEs can capture several elements of mobility concurrently because of the adaptability of DL modules, whereas previous techniques can only capture certain attributes of trajectories (e.g., limited to spatial or spatiotemporal). Furthermore, DL models can learn subtle and non-linear correlations in data that standard techniques may miss. As a result, these techniques can produce synthetic data that is more authentic than typical models.

Two Stage GAN (TSG) [23]: The first stage extracts the geographical points to a grid (spatial tessellation) to be modeled with a GAN using Convolutional Neural Networks, to learn general patterns, particularly the start and end point of every trajectory. The second stage deploys an additional GAN, processing each square of the grid to extract the road information from the geographical map. The GAN uses two parallel LSTMs to generate the sequence of intermediate stay points of the journey between the start and end of the trajectory. The generative capabilities are assessed on the Porto taxi dataset [7] by calculating the Jensen-Shannon (JS) divergence between the distribution of trajectory lengths and the frequencies of the 50 most visited places.

MoveSim [24]: MoveSim is a model-free GAN framework that incorporates human mobility regularity domain knowledge. Unlike other GAN-based approaches, MoveSim pre-trains the generator and discriminator to accelerate the learning process. MoveSim was tested using two datasets, a private mobile phone dataset, and a publicly available dataset, GeoLife [25]. The synthetic trajectories are tested by computing the distance distributions, the radius of gyration, the daily frequency of visits, the G-rank, and the I-rank.

Sequential Variational Autoencoder (SVAE) [26]: SVAE introduced a framework that combines the capabilities of VAEs and LSTMs. VAE generates a latent space capturing relevant features of trajectories. LSTMs, on the other hand, are used to handle sequential input. The purpose of SVAE is trajectory reconstruction to solve the problem of incomplete mobility data (sparsity).

Ouyang GAN [27]: A trajectory generator based on Wasserstein GAN [28]. A CNN serves as the basis for both the generator and the discriminator. The evaluation uses the MDC dataset [8] with a 64×64 squared tessellation of Lausanne, Switzerland.

The latest methods for generating trajectories have introduced powerful techniques for high-quality simulations of human mobility. The quality of the generated trajectories has been evaluated using metrics such as distribution similarities [23], [24], distance metrics [26], road-network matching accuracy [23], and marginal-distribution metrics [27]. In addition to such evaluation metrics, our work introduces two additional evaluation steps: (i) we evaluate whether the users can be identified from the synthetic trajectories; (ii), we evaluate the quality of the synthetic trajectories by building next-week trajectory predictors using DL.

C. TRAJECTORY-USER LINKING (TUL)

Trajectory classification is a critical problem in mobility data mining that aims to predict the class labels of moving entities based on their trajectories. Examples include the prediction of the object's mean of transport (e.g., automobile, taxi, bus, pedestrian, bike), vessel type (e.g., cargo, fishing), the user who owns the trajectory, and so on. TUL is a specific type of classification problem.

TUL definition [29]: Let $T_{u_i} = (p_{1_i}, p_{2_i}, \ldots p_{n_i})$ *indicate a trajectory belonging to user* u_i. *An unlinked trajectory* $T_k = (p_{1_k}, p_{2_k}, \ldots p_{n_k})$ *is a trajectory for which, unlike the previous one, we do not know the ID of the user who originated. Supposed we now have a set of these unlinked trajectories* $\mathcal{T} = (T_1, T_2, \ldots, T_m)$ *produced by users* $\mathcal{U} = (U_1, U_2, \ldots, U_n)$ *where* $m \gg n$. *TUL is the task of mapping unlinked trajectories to users* $\mathcal{T} \rightarrow \mathcal{U}$.

Similar to other mobility modeling tasks, TUL utilizes the latest end-to-end DL techniques, such as RNNs, LSTMs, CNNs, and attention mechanisms. Some specific DL architectures for TUL are listed below.

Bi-TULER [14]: This study introduced the TUL task as a subset of trajectory classification in which the labels are

the users, i.e., the individuals who produced each trajectory. Bi-TULER is a bidirectional RNN, which learns embeddings of Points of Interest (POIs) based on the distributional hypothesis [30]. The embeddings are exploited to train an RNN model for classifying trajectories, and despite the fact that it can capture more complicated patterns of mobility data than earlier efforts, just the POI identifier is employed to train the model.

TULVAE [15]: TULVAE is a successor of Bi-TULER, proposed especially to address the TUL problem. POI embeddings, like Bi-TULER, are pre-learned and then fed into the model. Trajectories are modeled using a VAE architecture. TULVAE is primarily dependent on POIs visited by users and hence does not allow the spatial, temporal, and semantic properties of trajectories.

Our work utilizes the latest TUL methods as one component of the overall proposed methods. This component safeguards user privacy, ensuring that the user IDs cannot be traced back from the synthetic data.

D. NEXT-LOCATION PREDICTION

Human mobility prediction models attempt to estimate future locations, either individually or collectively. Individually, next-location predictors anticipate an individual's future locations based on previous observations. Predicting people's future locations is useful in a variety of applications, including monitoring public health [31], well-being [32], and traffic congestion [33], as well as improving travel recommendation, geomarketing, and link prediction in social network platforms [34]. The prediction of a user's next location may be difficult since it involves collecting spatiotemporal patterns characterizing its unique routine [35], as well as merging disparate data sources to simulate different factors driving human mobility.

Definition of Next-location prediction [20]): Next-location prediction is the prediction of the next location (stay point) that an individual will visit in the future based on their previous mobility data. Let u be a user, T_u their trajectory, and $p_t \in T_u$ u's present position, next-location prediction seeks to forecast u's next destination p_{t+1}.

CLNN [36]: Classification Learning Neural Network (CLNN) is based on spatiotemporal features, user features, and POIs from current and previous trajectories. The coordinates are processed using an LSTM. A dense representation contains the date, time, user attributes, and POIs. One fully-connected layer handles the dense representation, while another fully-connected layer processes previous trajectories. The two resulting outputs are combined with coordinates representation to form a unique output. The output is then processed to estimate the future location. The performance is tested by computing the mean Haversine distance on the Porto taxi dataset [7].

Flashback [37]: Flashback is based on RNNs and the idea of flashbacks, i.e., a strategy that predicts future POIs by searching similar temporal properties in trajectories in the training data, using trajectories with sparse semantics. Flashback also uses embedding to estimate individuals' preferences for visiting various POIs. The next POI visit is predicted by combining embedding and the recurrent outputs with a fully-connected layer. Flashback used ACC@k (e.g., acc@5) for evaluation on the Foursquare [38] and Gowalla [39] datasets.

VANext [40]: The historical and current trajectories in Variational Attention-Based Next Location (VANext) are encoded into latent features via two independent encoders to capture the similarity and periodicity of POIs. A CNN is used to learn from previous trajectories, while the embedded trajectories are modeled with GRUs. A variational attention mechanism process both resulting representations, which finds previous trajectories similar to the current to be compared for predicting the user's next POI. VANext is evaluated with ACC@k on Gowalla [39] and Foursquare [38].

DeepMove [41]: DeepMove uses attention and a recurrent neural network that predicts future location from long and sparse trajectories. Both past and present trajectory data are embedded using multi-modal embeddings, which generates latent representations of spatial, temporal and semantic features. An attention mechanism identifies mobility patterns from historical trajectories, while a GRU processes current trajectories.

The latest next-location predictors are based on DL, which enabled researchers to overcome major challenges by capturing temporal, geographical, and social-geographic patterns in data utilizing mechanisms such as RNNs, LSTMs, CNNs, and attention mechanisms [42]. Our work utilizes a DL-based predictor as one component of the overall proposed methods. This component ensures that the synthetic data generated is useful for downstream machine learning tasks, such as predicting future mobility trajectories.

II. DATASETS

To analyze the influence of data sparsity, our study included experiments on two different datasets, Breadcrumbs (a dense dataset) and Foursquare (a sparse dataset). Both datasets are publicly available. Breadcrumbs requires researchers to sign an agreement, whereas Foursquare is freely available. The details of the two datasets are presented in the following subsections.

A. BREADCRUMBS

Breadcrumbs [43] is a rich mobility dataset with POI annotation. In the spring of 2018, DOPLab and the Information Security and Privacy Lab organized a Data Collection Campaign at and around the University campus of Lausanne and EPFL. The campaign ran for three months, collecting multi-sensor data (GPS, GSM, WiFi, Bluetooth) on smartphones of 80 individuals. The main goal was to create a novel dataset that addresses common limitations on available mobility datasets. These limitations include:

- Lack of location data and related information captured from multiple sensors.
- Unavailability of location data at a high spatiotemporal granularity throughout the data collection.
- Lack of ground-truth information regarding participant POI.
- Unavailability of semantic information regarding POIs.

Participants had to keep the data-gathering application running for the duration of the whole experiment. Moreover, they had to fill out questionnaires to validate their data, with respect to the semantics of POIs and relationships with other participants, as well as provide additional information regarding demographic and lifestyles. Given the granularity, the wide set of semantic labels, and the key ground truth, this dataset is a perfect candidate for modeling experiments.

B. FOURSQUARE

Foursquare is a local search-and-discovery mobile app developed by Foursquare Labs Inc. The app offers local search and personalized recommendations of places near the users' location based on their previous browsing and check-in history. The dataset used in this work is extracted in the study by Yang et al. [38], provided by Petry et al. [19]. It was originally used for studying the spatial-temporal regularity of user activity in LBSNs (Location-Based Social Network). It contains check-ins in New York City collected from 12 April 2012 to 16 February 2013 (about ten months). Each check-in entry records GPS coordinates, associated timestamps, and venue categories. There are 193 users, 3,079 trajectories, and 66,962 trajectory points in total in the dataset.

III. APPROACH

Based on two SOTA approaches for human-mobility modeling, we developed a new privacy-aware approach for predicting week-ahead trajectories. To ensure privacy awareness, the developed approach combines the strengths of LSTM-TrajGAN [16] and Multi-Aspect Trajectory Classification (MARC) [19] by: (i) using LSTM-TrajGAN to train a GAN model for generating realistic synthetic data; and (ii), using MARC to ensure user privacy by monitoring whether the actual users can be identified from the synthetic data. The developed approach is capable of predicting week-ahead visits described via the day of the week, time of day, and categorical POIs. The approach is presented in Fig. 2. In the next subsections, we first present the data preprocessing steps, and then we present the details of each component.

A. DATA PRE-PROCESSING

The two datasets come with their different characteristics, features, and semantics. However, for an accurate analysis, a common ground has to be defined, i.e., a format that allows easy processing for modeling and a direct correspondence in features – ensuring consistency and variability in the evaluation. Furthermore, the crucial aspect we want to model is

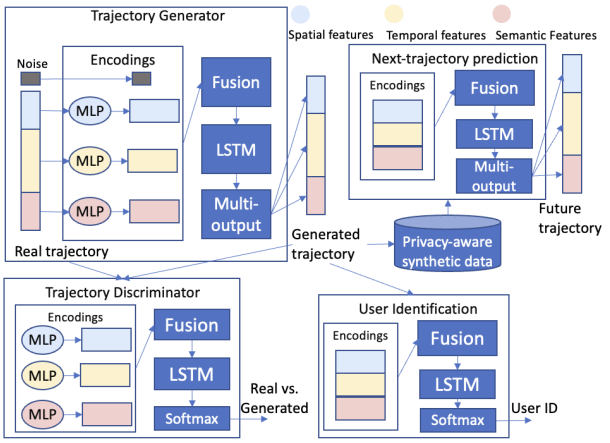


FIGURE 2. Workflow of the approach.

TABLE 1. Preprocessed dataset format.

Column	Feature	Type
uid	user ID	Integer
tid	trajectory ID	Integer
lat	latitude coordinate	Float
lon	longitude coordinate	Float
day	day of the week	Integer
hour	hour of the day	Integer
POI	point of interest	Integer or String

the trajectory. The definition of trajectory is defined by Petry et al. [19] with the following constraints and transformations:

- Check-ins belonging to broad categories are removed.
- Duplicated check-ins considering a 10 minutes threshold are also removed.
- Trajectories are grouped as a set of weekly check-ins from each user.
- Trajectories contains at least ten check-ins.
- Users have at least ten valid weekly trajectories.

The next step is to define the common characteristics of the datasets. The final formatted header of data is in the form **[uid, tid, lat, lon, day, hour, poi]** presented below in Table 1.

B. NEXT-WEEK TRAJECTORY PREDICTION

This component is responsible for evaluating the utility of synthetic data. Therefore it is a realistic application where the purpose is to predict the weekly trajectory with respect to temporal and semantic features (POIs), thus a POI recommendation system that provides awareness for the users' routines. The architecture is divided into four parts:

- **Encoding layer**, which encodes spatial, temporal, and semantic features into a representation based on their respective vocabulary size.
- **Fusion layer**, which combines all embeddings in fixed-size vectors.
- **LSTM layer**, which models the fused latent vectors sequentially and captures the temporal characteristics.
- **Prediction layer**, which provides the prediction of a possible future trajectory.

1) ENCODING LAYER

The Encoding layers are responsible for converting the data into a format that can be fed as input to the DL models. The interval data (e.g., date-time) has to be converted into two separate features composed of a single number representing the day of the week and the hour. The coordinates (latitude and longitude) are standardized using the centroid of all trajectories to obtain the deviations from the centroid. The nominal attributes such as the POI category are converted into a numerical representation.

Coordinate deviations are made such the model can learn the spatial deviation pattern between distinct trajectory points more effectively [44], they are treated as numerical input since the generation for these is set as a regression task.

Temporal and Semantic attributes are one-hot encoded based on the vocabulary sizes, encoding the features into high-dimensional binary vectors. For example, the feature `day` indicating the day of the week is represented by a seven-dimensional binary vector (e.g., Monday is encoded as [1, 0, 0, 0, 0, 0, 0]). Similarly, the feature `hour` is encoded in a 24-dimensional binary vector, and `POI` depends on the unique set of POIs available in each dataset.

Following the features encoding procedure, all trajectory's spatial, temporal, and semantic attributes are stored in a multidimensional matrix, where the first dimension provides the index for each trajectory. Because the length of each trajectory (number of POIs) varies, we use padding to ensure that all trajectories have the same length as the longest trajectory in the dataset. Concretely, we employ zero pre-padding to pad empty trajectory points (points with all features set to zero) to each trajectory until all trajectories have the same maximum length. The fundamental reason behind this is that data of the same size may be used for batch processing when training the model, accelerating the training process. Moreover, pre-padding (up against post-padding) has been proven more efficient when modeling, especially with LSTMs [45]. These padding trajectory points will be masked (omitted) throughout the model training and inference stages and will not affect the model's weights. The padded data is then fed into feature-specific multi-layer perceptrons (MLPs), which learn vector representations for each feature type separately [46].

2) FEATURE FUSION, LSTM, AND PREDICTION

Following the embedding procedure, all vectors are concatenated and then fused into 100-dimensional vectors using a dense layer. This blends together spatial, temporal, and semantic attributes of each trajectory point to provide spatiotemporal information.

Next, the LSTM layer models the input with a many-to-many strategy, i.e., it accepts a sequence with certain time steps as input and constructs (generates) a sequence with the same time steps as output. Finally, the synthetic trajectories produced by the LSTM layer must be decoded (given the many-to-many approach). For the latitude and longitude variations decoding, a dense layer of two units with a hyperbolic tangent activation function (tanh) is applied. To reconstruct the one-hot representation of categorical features dense layer is also applied but with the same number of units as the vocabulary size using the softmax normalized exponential function. The embedding and decoding weight matrices are shared across all trajectory points. Each output is predicted independently as a multi-task prediction problem.

C. TRAJECTORY GENERATOR AND DISCRIMINATOR

The generative component is trained to produce synthetic trajectories, and the discriminative component is trained to discriminate between the synthetic trajectories and the real-life trajectories. The generator architecture is composed of: an input layer, an embedding layer, a feature fusion layer, an LSTM layer, and a regression/multi-classification output layer. The encoded real trajectories and random noise are input into the generator, which then embeds the trajectories using MLPs [47]. The spatial dimension, defined as a pair of latitude and longitude deviations, is embedded into a fixed 64-dimensional vector. Temporal (*day* and *hour*) and semantics (POI) features are also embedded in a fixed-dimensional vector. Following the embedding procedure, all the vectors and random noise are concatenated and fused into 100-dimensional vectors using a dense layer. The resulting embedded trajectory sequence is then modeled with a many-to-many strategy by an LSTM layer. In the last step, the synthetic trajectories generated by the LSTM layer are decoded to match the input format of the data.

The discriminator and the generator share a similar architecture. The main distinctions between them are:

- The discriminator only accepts trajectory data as input (no random noise required).
- The discriminator's LSTM layer employs a many-to-one strategy that accepts a sequence as input and produces one scalar as output.
- The scalar output of the discriminator is then passed into a one-unit dense layer with a sigmoid activation function to perform the binary classification (real/fake).

The balance between the two modules is achieved via a special loss function, TrajLoss similarity function [16]. Unlike the original GANs, which take as input just random numbers (noise), in our case, real trajectory data is required as input in addition. As a result, a special loss metric function is needed. TrajLoss measures the difference between the actual trajectory and the corresponding synthetic trajectory data in spatial, temporal, and semantic dimensions, and this loss function is deployed to train the generator. This is how the TrajLoss is defined:

$$\begin{aligned} TrajLoss(y^r, y^p, t^r, t^s) = \alpha BCE(y^r, y^p) + \beta MSE(t^r, t^s) \\ + \gamma CCE_t(t^r, t^s) + \delta CCE_s(t^r, t^s) \end{aligned} \tag{1}$$

In the equation, y^r and y^p are the ground truth label and the discriminator's prediction for the trajectory, respectively. t^r and t^s are the real and the matching generated trajectories.

BCE is the discriminator's original binary cross-entropy loss. *MSE* is the Mean Squared Error of deviations loss, CCE_t is the temporal categorical cross-entropy loss (applied over the modeled as categories), and CCE_s is the spatial categorical cross-entropy loss. All loss functions represent the mismatch between the real and synthetic trajectories through a different perspective. α, β, γ, and δ are the weights for these losses, which can be parameterized differently for different scenarios. The TrajLoss updates the generator weights during model training to improve the quality of the generated synthetic trajectories.

D. USER IDENTIFICATION

This module safeguards user privacy, ensuring that the user IDs cannot be traced back from the synthetic data. The module is based on MARC [19], which is a method that can be used to train machine-learning models for user identification based on trajectory data. The goal of the model is to minimize the categorical cross-entropy loss, which is provided by the following equation:

$$-\frac{1}{\mathcal{N}_{train}} \sum_{T \in \mathcal{T}_{train}} \sum_{L \in \mathcal{L}} 1_{T \in L} \cdot \log p[T \in L] \tag{2}$$

In the equation, $\mathcal{T}_{train}$ is the set of trajectories used to train the model, $\mathcal{N}_{train}$ is the number of training instances (trajectories), and $\mathcal{L}$ is the set of labels used to classify the trajectories. In other words, we seek to increase our model's likelihood of accurately predicting each trajectory's label (the user ID) T. Finally, we also employ dropout [48] and regularization techniques to prevent overfitting our model to the training data, which is an intrinsic issue in deep neural networks such as RNNs. Dropout layers are used throughout the model to ensure that units are discarded at random throughout the training phase. Furthermore, the LSTM units' weights and biases are regularized using L1 regularization.

IV. EXPERIMENTS

This section first presents the experimental setup, and then it presents the experimental results in three separate subsections: the first subsection presents results about the quality of the synthetic data by looking into the spatial, temporal, and semantic characteristics of the generated data. The second subsection presents the results on user privacy, and the third subsection presents the results on next-week trajectory prediction.

A. EXPERIMENTAL SETUP

The experimental setup is represented in Figure 3. Where the generator, the classifier (user identification), and the predictor (next-week trajectory prediction) are tested individually. The main purpose is to test the generated data in a practical way with the two use cases representing privacy and utility. All experiments are carried out using the two datasets that have already been prepared and divided into a train (two-thirds of the data of each user) and a test (one-third of the data of each user) set, colored in green and red, respectively.

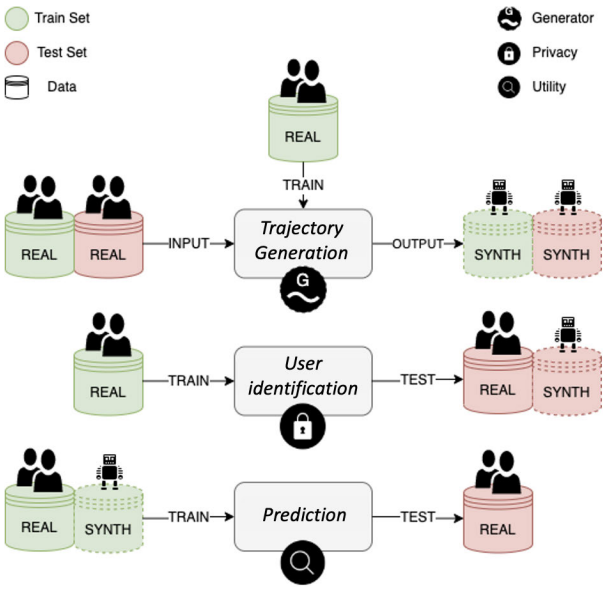


FIGURE 3. Experiments workflow.

The generator is trained according to the min-max game using the train set. Once convergence is reached, the train and test sets are passed through the trained model to generate their respective synthetic versions.

The model for user identification is responsible for assessing the privacy of the data. The model tries to identify users through their corresponding trajectories. For this purpose, the model is trained with the real training set and tested with both real and synthetic test sets. The evaluation with the real test set represents a baseline performance. The goal is to achieve lower accuracy with the synthetic evaluation showing that the synthetic dataset secures the users' privacy.

The predictor represents a real-world use case for determining the utility of the synthetic data. The goal is to predict the user's next-week trajectory given the current one. The predictor is trained with real and synthetic train sets but only tested with the real test set. This is because we want to see if it is possible to have a good performance using only synthetic training data instead of real training data. A good quality predictor trained with synthetic data should have similar predictive power as a predictor trained with real data.

B. QUALITY OF SYNTHETIC TRAJECTORIES

1) SPATIAL DISTRIBUTIONS

The spatial distributions for both real and generated data are presented in Figure 4. Where the x and y axes are latitude and longitude coordinates, respectively. These are grouped in bins for visualization convenience and colored to differentiate the distributions of real (blue color) from fake (synthetic in orange color). In both cases, it can be seen how the distribution of the synthetic data mirrors the real one, even generating disconnected and distant points from the base distribution. On the Breadcrumbs dataset, the generator tries to reach more distant and scattered points in the generation. On the

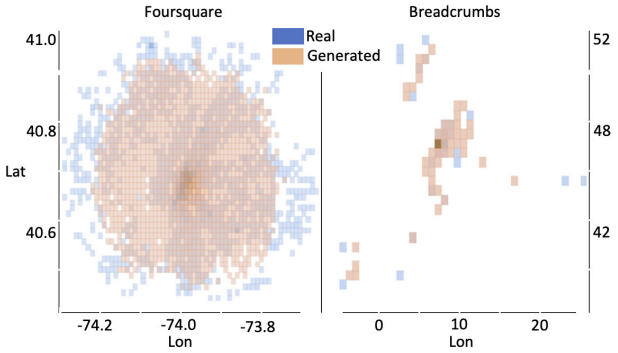


FIGURE 4. Spatial distribution (longitude and latitude coordinates) of the real data vs. the generated data for the two datasets (Foursquare and Breadcrumbs).

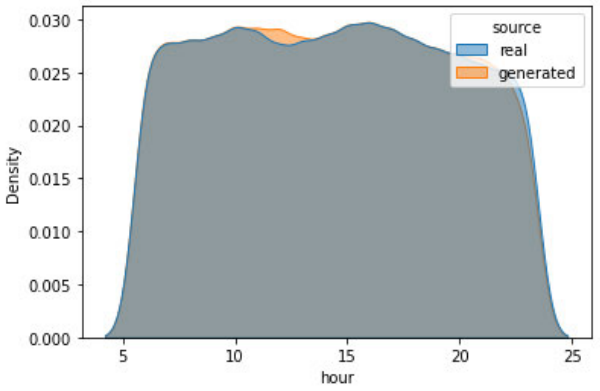


FIGURE 5. Temporal distribution (hour of the day) of the real data vs. the generated data for the Breadcrumbs dataset.

Foursquare dataset, this is less pronounced because of how the data is distributed and the resolution of the bins. Nonetheless, it happens on the edges of the distribution "area".

2) TEMPORAL CHARACTERISTICS

Figure 5 presents the distribution of the feature *hour* of the day for the real and for the generated data. We visualized only the results for the Breadcrumbs because they were more challenging compared to the Foursquare dataset. For the Foursquare dataset, the two distributions matched even better than the presented results. In Figure 5, the x-axis represents the *hour* of the day (from 0 to 24), and the y-axis represents the density of the distribution. The figure shows that the two distributions generally overlap (the gray area). There is a slight difference between 10h and 22h. The figure also shows that the low frequency between 00:00 and 05:00 does not prevent the model's excellent performance.

Figure 6 presents the distribution of the feature *day* of the week. The x-axis numerically indicates the day of the week, starting from 0 (Monday) to 6 (Sunday), and the y-axis represents the density of the frequency distribution of the categorical attribute. From the figure, it can be seen that there is an almost perfect match between the two distributions.

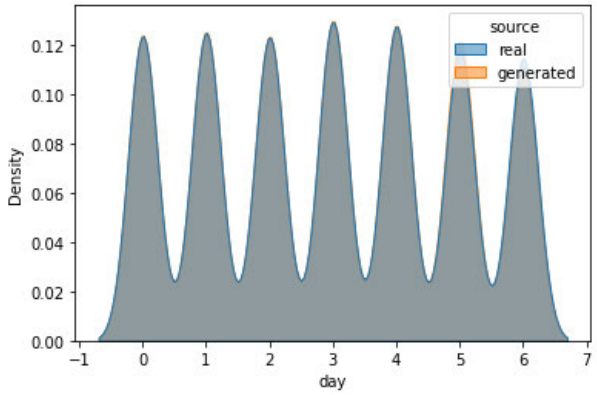


FIGURE 6. Temporal distribution (day of the week) of the real data vs. the generated data for the Breadcrumbs dataset.

TABLE 2. Pearson's correlation coefficient (PCC) of Temporal attributes (real vs. generated).

Data	Breadcrumbs		Foursquare	
Feature	day	hour	day	hour
Mean PPC (Std.Dev.)	.99 (.01)	.98 (.01)	.99 (.02)	1.0 (.01)

TABLE 3. Chi-square p-values of categorical attribute frequencies.

Data	Breadcrumbs	Foursquare
Chi-square (p-value)	0.35	0.99

Table 2 presents the Person's correlation coefficient for the two features *day* of the week and *hour* of the day. The results confirm the match between the real and the generated data distributions for the two features.

3) SEMANTIC CHARACTERISTICS

We performed statistical analysis using the chi-square test to quantify the similarities between the distributions of the POI semantic category. The null hypothesis is that the two distributions are similar, i.e., the frequencies of the POI categories are the same for both the real and the generated data. The alternative hypothesis is that the two distributions are different. The results of the tests are presented in Table 3. The p-values are above 0.05, thus the null hypothesis cannot be rejected, i.e., we cannot state the two distributions are different. Furthermore, it can be seen that the Breadcrumbs dataset is more problematic than the Foursquare dataset.

Figure 7 presents an example of the visit frequency for each POI semantic category in the Foursquare dataset. The bars represent the visit frequency, and the lines indicate the kernel density estimation (KDE). The frequency is represented on a logarithmic scale. The figure shows that the only problematic category with more prominent differences between the real and the generated data is the category *Event*. This is probably because the visit frequency for that place was the smallest compared to the other POIs.

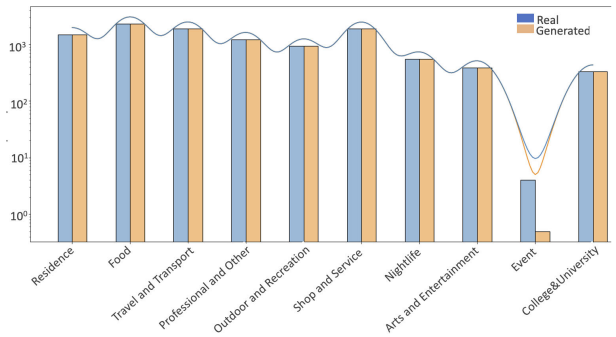


FIGURE 7. Temporal distribution (day of the week) of the real data vs. the generated data for the Breadcrumbs dataset.

C. USER PRIVACY

Table 4 presents the results in TUL accuracy. In these experiments, higher accuracy represents lower user privacy, i.e., the user can be identified from the corresponding trajectories more accurately. We present the accuracy@K score, which shows whether the actual user was part of the top-k predicted users, ranked by the likelihood scores that the TUL model outputs. We tested two versions of the model. One that uses longitude and latitude coordinates as input (*coord.*), and another without the coordinates, i.e., the *no_coord.* model uses only trajectories represented via the date-time features and the POI category.

For the Breadcrumbs dataset, the model can recognize users from their trajectories with an accuracy of 99.37%. On the other hand, the respective test on synthetic data is less than 6%. In the version without coordinates, we see that the accuracy is lower. Nevertheless, the respective synthetic version shows an even lower performance. For the Foursquare dataset, the model that uses all the features reflects the same performance in identifying users. Although between 93.96% and 21.32% the difference is smaller, there is still a high rate of privacy protection. In the version of the model without coordinates, the performance with the synthetic data is the same as with the real data. For this case, including coordinates in the classification is important for performance. Nevertheless, these experiments also confirm how easily a user can be recognized from mobility trajectories.

These results indicate that the latest machine learning methods for user identification (e.g., MARC) can identify the users from the training data with 99% accuracy. However, if we use the synthetic data instead of the original data to train the predictive models (and the attackers get access to the training data), the user-identification accuracy drops to 6% on the Breadcrumbs dataset and 21% on the 4square dataset.

D. NEXT-WEEK TRAJECTORY PREDICTION

The evaluation results for these experiments are presented in Table 5. Higher accuracy represents better quality data, i.e., better models can be created from the data. "Real" and "Generated" data refer to the dataset on which the model was trained. Multi-task prediction is performed on the temporal

TABLE 4. User identification results for models with latitude and longitude coordinates as input (coord.) and models without input coordinates (no_coord.).

Data	ACC	coord. (Real)	coord. (Gen.)	no_coord. (Real)	no_coord. (Gen.)
Bread.	@1	**0.99**	0.06	0.12	0.07
	@5	**1.0**	0.14	0.28	0.18
	@10	**1.0**	0.20	0.41	0.32
4sq.	@1	**0.94**	0.21	0.32	0.32
	@5	**0.98**	0.48	0.61	0.61
	@10	**0.99**	0.62	0.71	0.72

TABLE 5. Accuracy for next-week trajectory prediction.

Data	Feature	LSTM (Real)	LSTM (Generated)	Dummy (Most frequent)
Bread.	day	0.30	**0.33**	0.18
	hour	**0.24**	0.23	0.08
	POI	0.45	0.59	**0.63**
4sq.	day	0.89	0.89	0.31
	hour	0.89	**0.90**	0.20
	POI	0.91	0.91	0.61

(*day*, *hour*) and categorical(POI) features. Also, these results are compared with a "Dummy" predictor, which makes a prediction based on the most frequent occurrences. To predict trajectory T_{i+1} the model uses the most frequent corresponding feature value in T_i.

For the Breadcrumbs dataset, the decrease is from 24% to 23% for feature *hour*, an improvement from 30% to 33% for the feature *hour*, and an improvement from 45% to 59%feature POI. The result of the Dummy model is the best in the case of POI. This is probably due to the complexity of modeling the Breadcrumbs dataset, given its density of check-ins. In contrast, for the Foursquare dataset, the difference between the models is even smaller. In this case, also the baseline results achieved by the Dummy predictor are far exceeded.

V. DISCUSSION

A. GENERATION OF SYNTHETIC TRAJECTORIES

The generator of synthetic trajectories demonstrated high performance in modeling spatial, temporal, and semantic feature characteristics. The spatial characteristics were the most challenging to capture. Nevertheless, the distribution of longitude and altitude coordinates depicted in Figure 4 showed that the generator can synthesize points far detached from the centroid. For the temporal features, the correlations between the real trajectories and the synthetic generated for both Breadcrumbs and Foursquare datasets were above 0.97 (see Table 2). The semantic characteristic (POI type) was the most complex feature to synthesize; nonetheless, the generator considered even those with a low frequency. The Chi-square statistical tests for the categorical features indicated that there is no statistically significant difference between the POIs of the original trajectories and the synthetic generated. The p-values of Breadcrumbs and Foursquare, 0.35 and 0.99, respectively, indicate that the hypothesis of

the synthetic trajectories being similar cannot be rejected. The model is very complex and includes a multitude of parameters that still have the potential to be fine-tuned.

The learning rate of the generator was one of the most important parameters. Depending on this parameter, the model could either make a complete copy of the real data observing only the most general patterns or instead, it could encode more fine-grained habits, generating realistic synthetic trajectories rich in information that can be used in practice.

Also, the importance of data pre-processing, especially the coordinate preparation technique, should be emphasized. These are standardized to amplify the difference between two points and thus more easily learned by the model. Since the difference between two sets of relatively close coordinates is only visible after the decimal point.

B. USER PRIVACY–TUL

The experiments performed on user identification (see Table 4) demonstrated the ability of the generator to preserve privacy. The user-identification accuracy for the real data is 99% for the Breadcrumbs dataset and 93% for the Foursquare dataset. The synthetic trajectories decreased the accuracy to 5.9% and 21% for Breadcrumbs and Foursquare datasets, respectively. We suspect that the high density of the Breadcrumbs dataset is the reason for the higher difference in accuracy between real and synthetic trajectories. These experiments also show that using the coordinates as input is the most important source for user identification.

C. NEXT-WEEK TRAJECTORY PREDICTION

The novel privacy-aware approach for predicting future trajectories achieved high accuracy with both datasets in their respective real and synthetic versions(Table 5). This demonstrated the effectiveness of the proposed approach, i.e., the approach can achieve high utility while maintaining a low probability of user identification. For the Breadcrumbs dataset, the accuracy for predicting the POI feature was 45% in the original trajectories and the same for the synthetic data. The same applies to the Foursquare dataset, with 91% accuracy on the POI feature.

The major difference between the Foursquare and Breadcrumbs datasets is the number of trajectories per user and their length. The foursquare dataset, despite its sparsity, provided a bigger number of trajectories per user, therefore more examples to learn from.

D. COMPUTING TIME

We implemented the method using TensorFlow and Python, and we ran all experiments on the Google Colab platform, which provides access to a Tesla T4 GPU (16 GB RAM.) To estimate the computational cost, we measured the training time and the inference time for each of the three components of the method. The results are presented in Table 6, where we can see that the most computationally demanding step is the training process of the GAN model, which requires close to one hour for the smaller dataset (Breadcrumbs), and close to two hours and a half for the larger dataset (Foursquare). The training time of the MARC and the next-week trajectory predictors is less than a minute. The inference time for all three components is between 0.14 and 1.3 seconds.

TABLE 6. Computing time (in seconds) required during training and inference for the three components (GAN, MARC, and next-week trajectory predictor) of the proposed method.

	GAN		MARC		Traj. Predictor	
	Train	Infer.	Train	Infer.	Train	Infer.
Bread.	3900	1	41	0.1	36	0.14
4sq.	9180	1.3	71	0.4	63	0.2

E. GAN STABILITY

The stability of the trajectory generator in our proposed method is an important component that needs to be considered, given that GANs utilize complex training of several modules (e.g., a generator and a discriminator). Nevertheless, our experiments showed that the generator is quite stable. One evidence of stability is the low standard deviation presented in Table 2, i.e., the Parsons's correlation coefficient was above 0.95, signifying high-quality trajectories in all experiments. Furthermore, in the experiments on next-week trajectory prediction (Table 5) we used two different datasets and three different modalities, two of which are numerical (*day* and *hour*), and one is categorical (POI). Thus, the GAN models in our work were evaluated in six different combinations (three modalities, two datasets).

To further evaluate the stability of the GAN, and the impact of the training data on the overall proposed method, we performed experiments where we varied the size of the training set used to train the GAN model. More specifically, we trained GAN models using n% of the training data of each participant, where n was set to 30, 50, and 100. We then generated synthetic trajectories using the three variations of the GAN models, and we subsequently trained three next-week trajectory predictors. The next-week predictors were then evaluated on the real test set. The test data was kept constant across these experiments to allow for a fair comparison. The results of these experiments are presented in Table 7. From the table, it can be seen that the accuracy for the feature *day* is stable for both datasets (Breadcrumbs and Foursquare), regardless of the size of the training set. The accuracy for the feature *hour* is stable for the Breadcrumbs dataset, but it decreases on the Foursquare dataset when the models are trained with smaller training datasets (e.g., 50% or 30% of the overall training data). The accuracy for the feature POI is stable for the Foursquare dataset. It decreases by 14 percentage points for the Breadcrumbs dataset when the models are trained with 50% or with 30% of the training dataset.

These results indicate that the quality of the synthetic trajectories generated by the GAN model depends on the type of the dataset, i.e., Foursquare is bigger but sparser,

TABLE 7. Accuracy for next-week trajectory prediction for varying training data sizes (100%, 50%, and 30%).

Data	Feature	Size of training data 100%	50%	30%
Bread.	day	0.33	0.29	0.29
	hour	0.23	0.20	0.19
	POI	0.59	0.45	0.45
4sq.	day	0.89	0.89	0.89
	hour	0.90	0.11	0.01
	POI	0.91	0.90	0.89

and Breadcrumbs is a smaller but denser dataset. The quality also depends on the feature type (*day*, *hour*, or POI in our experiments). And, as expected, the quality depends also on the size of the training data, i.e., models trained with the overall training data produce better synthetic trajectories compared to the models trained with 50% or 30% of the training dataset.

F. FUTURE WORK AND LIMITATIONS

The generator requires significant computing power. It is also constrained by the number of POIs, which increases the number of parameters exponentially by influencing the size of the encoding layers. A common human mobility dataset with tens of thousands of one-hot encoded POIs can occupy as much as 10 GBs of working memory. A future direction would be to augment the approach to work with multi-hot encoding or other types of dense representations. Furthermore, the generator outputs synthetic trajectories with the same length. This may be solved using input and output sequences that have a dynamic length, similar to generative models that generate sentences with a dynamic length [49].

In this work, we used the user-identification task for testing privacy, but other assessments could also be used to evaluate privacy from other perspectives.

Regarding the hyperparameters of the TrajLoss function, a detailed analysis of their influence on the GAN models has been presented by Rao et al. [16]. For example, they have investigated TrajLoss with and without spatial, temporal, and categorical components. However, our user-identification experiments showed that the spatial input, i.e., POI coordinates, is the most important input for user identification from the generated data (see Table 4). Consequently, the weight of the spatial component of the TrajLoss function (the parameter δ), may be an interesting component to analyze, providing a trade-off between the quality of the generated trajectories and user privacy.

VI. CONCLUSION

We presented a novel approach for predicting next-week trajectories in a privacy-aware manner. The approach is based on two SOTA approaches, LSTM-TrajGan [16] used for generating synthetic trajectories, and MARC [19] used for user identification. The combination of these two approaches was augmented with a next-week trajectory predictor, which provides accurate predictions using only the privacy-aware synthetic data. All components of the approach work with spatial, temporal, and semantic properties that characterize multiple-aspect trajectories and employ multi-feature embedding layers to represent these heterogeneous dimensions.

The experiments show how privacy and utility are not two opposing dimensions but instead can be combined with the new approach, creating a new synthetic dataset based on the original one that preserves both utility and privacy. The utility was demonstrated to be high, given the ability to predict week-ahead trajectories, having learned only from synthetic trajectories. Additionally, user privacy was preserved, given that the user-identification model that used synthetic trajectories performed much worse than the model that used real trajectories.

The generator was able to create trajectories that are similar to the real ones but difficult to trace back to the original users. This was demonstrated by the low user-recognition recognition accuracy and by the accurate next-week trajectory prediction. The statistical descriptors, including Pearson's correlation coefficient and Chi-square statistical tests, confirmed that there is no significant difference between the original and the generated trajectories.

The results per dataset show that the problem of sparsity is present but not necessarily the primary issue. Sparsity refers to the fact that trajectories in the datasets are often composed of considerably distant points, distant from a spatial, temporal, and semantic perspective. Depending on the application, the distribution of coordinates, the number of different POIs, and their frequency may be the main factors influencing the results. Furthermore, the size of the data might have a bigger influence on the quality of the models than the sparsity of the data, i.e., models built on smaller and dense datasets (Breadcrumbs in our case) may perform worse than models built on large and sparse datasets (e.g., Foursquare in our case). Nevertheless, this relation should be investigated with a wider variety of datasets.

REFERENCES

[1] M. U. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, and D. M. Pigott, "The effect of human mobility and control measures on the COVID-19 epidemic in China," *Science*, vol. 368, no. 6490, pp. 493–497, 2020.

[2] N. Oliver, B. Lepri, H. Sterly, R. Lambiotte, S. Deletaille, M. De Nadai, and E. Letouzé, "Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle," *Sci. Adv.*, vol. 6, no. 23, Jun. 2020, Art. no. eabc0764.

[3] Y. Wang and J. E. Taylor, "Coupling sentiment and human mobility in natural disasters: A Twitter-based study of the 2014 South Napa Earthquake," *Natural Hazards*, vol. 92, no. 2, pp. 907–925, 2018.

[4] A. Rossi, G. Barlacchi, M. Bianchini, and B. Lepri, "Modelling taxi drivers' behaviour for the next destination prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 2980–2989, Jul. 2020.

[5] A. Mehrotra, R. Hendley, and M. Musolesi, "Towards multi-modal anticipatory monitoring of depressive states through the analysis of human-smartphone interaction," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput., Adjunct*, Sep. 2016, pp. 1132–1138.

[6] C. L. Gray and V. Mueller, "Natural disasters and population mobility in Bangladesh," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 16, pp. 6000–6005, Apr. 2012.

[7] L. Moreira-Matias, J. Gama, M. Ferreira, J. Mendes-Moreira, and L. Damas, "Predicting taxi–passenger demand using streaming data," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1393–1402, Sep. 2013.

[8] J. K. Laurila, D. Gatica-Perez, I. Aad, O. Bornet, T.-M.-T. Do, O. Dousse, J. Eberle, and M. Miettinen, "The mobile data challenge: Big data for mobile computing research," Tech. Rep., 2012.

[9] L. Pappalardo, S. Rinzivillo, Z. Qu, D. Pedreschi, and F. Giannotti, "Understanding the patterns of car travel," *Eur. Phys. J. Special Topics*, vol. 215, pp. 61–73, Jan. 2013.

[10] M. C. González, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.

[11] R. Jurdak, K. Zhao, J. Liu, M. AbouJaoude, M. Cameron, and D. Newth, "Understanding human mobility from Twitter," *PLoS ONE*, vol. 10, no. 7, Jul. 2015, Art. no. e0131469.

[12] S. Wang, Z. Bao, J. S. Culpepper, and G. Cong, "A survey on trajectory data management, analytics, and learning," *ACM Comput. Surv.*, vol. 54, no. 2, pp. 1–36, Mar. 2022.

[13] X. Zheng, J. Han, and A. Sun, "A survey of location prediction on Twitter," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 9, pp. 1652–1671, Sep. 2018.

[14] Q. Gao, F. Zhou, K. Zhang, G. Trajcevski, X. Luo, and F. Zhang, "Identifying human mobility via trajectory embeddings," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, Aug. 2017, pp. 1689–1695.

[15] F. Zhou, Q. Gao, G. Trajcevski, K. Zhang, T. Zhong, and F. Zhang, "Trajectory-user linking via variational AutoEncoder," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 3212–3218.

[16] J. Rao, S. Gao, Y. Kang, and Q. Huang, "LSTM-TrajGAN: A deep learning approach to trajectory privacy protection," 2020, *arXiv:2006.10521*.

[17] L. Bradford, M. Aboy, and K. Liddell, "COVID-19 contact tracing apps: A stress test for privacy, the GDPR, and data protection regimes," *J. Law Biosci.*, vol. 7, no. 1, Jul. 2020, Art. no. lsaa034.

[18] M. Gjoreski, M. Laporte, and M. Langheinrich, "Toward privacy-aware federated analytics of cohorts for smart mobility," *Frontiers Comput. Sci.*, vol. 4, p. 88, Jul. 2022.

[19] L. M. Petry, C. L. D. Silva, A. Esuli, C. Renso, and V. Bogorny, "MARC: A robust method for multiple-aspect trajectory classification via space, time, and semantic embeddings," *Int. J. Geograph. Inf. Sci.*, vol. 2020, pp. 1–23, Jan. 2020.

[20] M. Luca, G. Barlacchi, B. Lepri, and L. Pappalardo, "A survey on deep learning for human mobility," *ACM Comput. Surv.*, vol. 55, no. 1, pp. 1–44, Jan. 2023.

[21] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[22] L. Pappalardo, F. Simini, S. Rinzivillo, D. Pedreschi, F. Giannotti, and A.-L. Barabási, "Returners and explorers dichotomy in human mobility," *Nature Commun.*, vol. 6, no. 1, pp. 1–8, Sep. 2015.

[23] X. Wang, X. Liu, Z. Lu, and H. Yang, "Large scale GPS trajectory generation using map based on two stage GAN," *J. Data Sci.*, vol. 19, no. 1, pp. 126–141, 2021.

[24] J. Feng, Z. Yang, F. Xu, H. Yu, M. Wang, and Y. Li, "Learning to simulate human mobility," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 3426–3433.

[25] Y. Zheng, X. Xie, and W. Y. Ma, "GeoLife: A collaborative social networking service among user, location and trajectory," *IEEE Data Eng. Bull.*, vol. 33, no. 2, pp. 32–39, Jan. 2010.

[26] D. Huang, X. Song, Z. Fan, R. Jiang, R. Shibasaki, Y. Zhang, H. Wang, and Y. Kato, "A variational autoencoder based generative model of urban human mobility," in *Proc. IEEE Conf. Multimedia Inf. Process. Retr. (MIPR)*, Mar. 2019, pp. 425–430.

[27] K. Ouyang, R. Shokri, D. S. Rosenblum, and W. Yang, "A non-parametric generative model for human trajectories," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 3812–3817.

[28] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–7.

[29] G. Wang, D. Liao, and J. Li, "Complete user mobility via user and trajectory embeddings," *IEEE Access*, vol. 6, pp. 72125–72136, 2018.

[30] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013, *arXiv:1301.3781*.

[31] L. Canzian and M. Musolesi, "Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Sep. 2015, pp. 1293–1304.

[32] L. Pappalardo, M. Vanhoof, L. Gabrielli, Z. Smoreda, D. Pedreschi, and F. Giannotti, "An analytical framework to nowcast well-being using mobile phone data," *Int. J. Data Sci. Anal.*, vol. 2, pp. 75–92, Dec. 2016.

[33] Y. Shi, H. Feng, X. Geng, X. Tang, and Y. Wang, "A survey of hybrid deep learning methods for traffic flow prediction," in *Proc. 3rd Int. Conf. Adv. Image Process.*, Nov. 2019, pp. 133–138.

[34] R. Wu, G. Luo, J. Shao, L. Tian, and C. Peng, "Location prediction on trajectory data: A review," *Big Data Mining Anal.*, vol. 1, no. 2, pp. 108–127, 2018.

[35] H. Barbosa, M. Barthelemy, G. Ghoshal, C. R. James, M. Lenormand, T. Louail, R. Menezes, J. J. Ramasco, F. Simini, and M. Tomasini, "Human mobility: Models and applications," *Phys. Rep.*, vol. 734, pp. 1–74, Mar. 2018.

[36] J. Tang, J. Liang, T. Yu, Y. Xiong, and G. Zeng, "Trip destination prediction based on a deep integration network by fusing multiple features from taxi trajectories," *IET Intell. Transp. Syst.*, vol. 15, no. 9, pp. 1131–1141, Sep. 2021.

[37] D. Yang, B. Fankhauser, P. Rosso, and P. Cudre-Mauroux, "Location prediction over sparse user mobility traces using RNNs: Flashback in hidden states!" in *Proc. 29th Int. Joint Conf. Artif. Intell.*, Jul. 2020, pp. 2184–2190.

[38] D. Yang, D. Zhang, V. W. Zheng, and Z. Yu, "Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs," *IEEE Trans. Syst. Man, Cybern., Syst.*, vol. 45, no. 1, pp. 129–142, Jan. 2015.

[39] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2011, pp. 1082–1090.

[40] Q. Gao, F. Zhou, G. Trajcevski, K. Zhang, T. Zhong, and F. Zhang, "Predicting human mobility via variational attention," in *Proc. World Wide Web Conf.*, May 2019, pp. 2750–2756.

[41] J. Feng, Y. Li, C. Zhang, F. Sun, F. Meng, A. Guo, and D. Jin, "DeepMove: Predicting human mobility with attentional recurrent networks," in *Proc. World Wide Web Conf.*, 2018, pp. 1459–1468.

[42] C. E. J. Ezequiel, M. Gjoreski, and M. Langheinrich, "Federated learning for privacy-aware human mobility modeling," *Frontiers Artif. Intell.*, vol. 5, Jun. 2022, Art. no. 867046.

[43] A. Moro, V. Kulkarni, P.-A. Ghiringhelli, B. Chapuis, K. Huguenin, and B. Garbinato, "Breadcrumbs: A rich mobility dataset with point-of-interest annotations," in *Proc. 27th ACM SIGSPATIAL Int. Conf. Adv. Geograp. Inf. Syst.*, New York, NY, USA, Nov. 2019, pp. 508–511.

[44] G. Mai, K. Janowicz, B. Yan, R. Zhu, L. Cai, and N. Lao, "Multi-scale representation learning for spatial feature distributions using grid cells," 2020, *arXiv:2003.00824*.

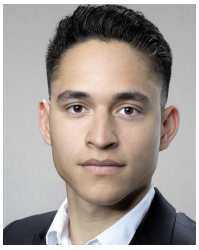
[45] M. Dwarampudi and N. V. S. Reddy, "Effects of padding on LSTMs and CNNs," 2019, *arXiv:1903.07288*.

[46] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

[47] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[48] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.

[49] T. B. Brown et al., "Language models are few-shot learners," in *Proc. Adv. Neural Inf. Process. Sys.*, vol. 33, 2020, pp. 1877–1901.



IVAN FONTANA received the B.Sc. degree in microtechnology and medical technology from the Bern University of Applied Sciences, Biel/Bienne, Switzerland, in 2019, and the M.Sc. degree in artificial intelligence from the Faculty of Informatics, Università della Svizzera italiana, Lugano, Switzerland, in 2022.

He is an Occupying Forces Sergeant of the Swiss Army. He is currently with xFarm Technologies, Manno, Switzerland, as a Computer Vision Engineer.

MARC LANGHEINRICH (Member, IEEE) leads the Research Group on ubiquitous computing with the Faculty of Informatics, Università della Svizzera italiana (USI) and is an expert in addressing privacy in pervasive and mobile computing. He is also a Dean of the Faculty of Informatics, USI. His articles on addressing privacy in smart environments have together been cited 1000 times. He is a member of the Steering Committee of UbiComp, the PerDis Symposium, and the IoT Conference. He was serving as the General Chair or the Program Chair for many major conferences in the area of pervasive and ubiquitous computing (e.g., UbiComp, Pervasive, PerCom, and the IoT Conference). He is the Editor-in-Chief of the *IEEE Pervasive Computing* magazine.

MARTIN GJORESKI received the Ph.D. degree in computer science from Jožef Stefan Postgraduate School, Ljubljana, Slovenia, in 2020. From 2014 to 2020, he was a Research Assistant with the Department of Intelligent Systems, Jožef Stefan Institute. Since 2021, he has been a Postdoctoral Researcher with the Faculty of Informatics, Università della Svizzera italiana, Switzerland. His research interest includes the development of machine-learning methods for monitoring human behavior. Jointly with several research teams, he has won five machine-learning competitions, from 2018 to 2020, including the Sussex-Huawei Locomotion Challenge, in 2018 and 2019. In 2021, he was also awarded the National Award "Jožef Stefan Golden Emblem," signifying an outstanding Ph.D. thesis in Slovenia.