

## Research Article

# Anomalous Trajectory Detection Using Masked Autoregressive Flow Considering Route Choice Probability

Pengqian Cao,<sup>1,2</sup> Renxin Zhong ,<sup>1,2</sup> and Wei Huang <sup>1,2</sup>

<sup>1</sup>School of Intelligent Systems Engineering, Shenzhen Campus of Sun Yat-sen University, Shenzhen, Guangdong, China

<sup>2</sup>Guangdong Key Laboratory of Intelligent Transportation Systems, Guangzhou, Guangdong, China

Correspondence should be addressed to Wei Huang; [huangwei5@mail.sysu.edu.cn](mailto:huangwei5@mail.sysu.edu.cn)

Received 11 April 2022; Revised 27 September 2022; Accepted 29 September 2022; Published 18 October 2022

Academic Editor: Antonino Vitetta

Copyright © 2022 Pengqian Cao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Taxis play a critical role in public traffic systems, and they deliver myriad travelers with convenient service due to temporal-spatial availability. However, anomalous trajectories such as trip fraud often occur due to greedy drivers. In this study, we propose an anomalous trajectory detection method that incorporates Route Choice analysis into Masked Autoregressive Flow, named MAFRC-ATD. The MAFRC-ATD integrates data-driven and model-based methods. First, we divide the urban traffic network into small grids and represent subtrajectories with a sequence of grids. Second, based on the subtrajectories, we employ the MAFRC-ATD model to calculate the anomaly score of each trajectory. Third, according to the anomaly score, we can identify the anomalous trajectories and distinguish between intentionally and unintentionally anomalous. Finally, we evaluate our method with a real-world dataset in Porto, Portugal. The experiment demonstrates that the MAFRC-ATD can effectively discover anomalous trajectories and can identify the unintentional detours due to traffic congestion.

## 1. Introduction

Taxis provide convenient and available services for a multitude of urban passengers. However, passengers may be victims who are overcharged deliberately by greedy drivers. Two examples are as follows.

*Example 1.* Many people, mostly tourists, are victims of taxi-driving frauds committed by greedy taxi drivers who overcharge passengers by deliberately taking unnecessary detours [1]. In New York City, nearly 1.1 million dollars have been overcharged due to 21,819 taxicab drivers swindling money out of customers arbitrarily [2].

*Example 2.* Car hailing is a combination of the Internet and sharing economy, fundamentally changing the way people travel. However, car-hailing platforms lower employment requirements and increase effective supply, bringing safety and economic risks. The risk of overcharging, leaking personal information, or even crimes [3] could occur due to its

strong characteristics of a “stranger economy.” Thus, to avoid potential crises, scouring the potential anomalous behaviors from historical record data of cabbies is prohibitively crucial.

To tackle this issue, collecting the GPS traces and discovering the anomalous trajectories from the sheer volume of data is an effective approach. The rapid advances in big data and machine learning grant us an unprecedented opportunity to detect anomalous trajectories via the trace data of moving objects.

Anomalous trajectory detection has attracted extensive research attention. Recently, with the proliferation of deep learning and deep generative models (DGMs), anomalous trajectory detection methods based on deep learning have made remarkable progress in recent years. However, the existing methods are usually faced with the following challenges:

Spatiotemporal information for DGM: as a branch of deep learning, deep generative models are employed to identify anomalous trajectories by reconstruction error.

However, most deep generative models [4–7] are difficult to directly capture sequential information of trajectory data.

**Unintentional anomalous trajectory:** mostly, drivers may take unintentionally anomalous detours with their experience to avoid traffic congestion. The experience of drivers, particularly, is crucial for route recommendations when traffic congestion occurs. However, traditional anomalous trajectory detection methods seldom consider mining this unintentionally anomalous trajectory pattern.

We take an unintentional anomaly for instance. As shown in Figure 1, there are two trajectories denoted as  $t_1$  and  $t_2$ . Obviously,  $t_1$  is normal, whereas  $t_2$  seems to take a detour trace. However, the travel time of  $t_1$  is similar to the travel time of  $t_2$ , suggesting that the driver of  $t_2$  may take a detour to reduce the travel time if traffic congestion or a traffic incident occurs. In real life, this unintentional anomaly is critical to route recommendations for traffic congestion.

To mitigate the above problems, this study proposes a framework that integrates the route choice model in the masked autoregressive flow (MAF), which is one framework of DGMs. Our key insight is that anomalous trajectory detection can be transformed into a distribution estimation task. MAF can fully capture the sequential information of trajectory data by leveraging the autoregressive mechanism and ought to assign the low probability density at the anomalous trajectory (i.e., the likelihood). On the contrary, to capture the actual route choice behavior, the route choice model is applied to distinguish unintentional anomalies due to traffic congestion or traffic incident. When traffic incident or traffic congestion occurs, people will normally take other routes which may be longer in distance, to avoid congestion or incident. Henceforth, the route choice model can successfully mine the patterns of unintentionally anomalous trajectories through the choice probability of corresponding routes.

In a nutshell, this study proposes an anomalous trajectory detection method based on masked autoregressive flow considering route choice analysis, named MAFRC-ATD. First, we extract valid trajectories from the historical dataset and split the traffic network into equal-sized grids; hence, subtrajectories can be represented with a sequence of grids. Second, we employ the MAFRC-ATD model to calculate the anomaly score of each trajectory through the likelihood of subtrajectories and the choice probability of the corresponding route. Finally, according to the anomaly score, each trajectory can be determined whether it is anomalous and, if anomalous, which type of anomalous trajectory it belongs to.

Overall, the contributions of this study can be summarized as follows.

We propose an unsupervised framework based on normalizing flow, which takes advantage of the



FIGURE 1: An illustration of trips traversing from Porto cathedral to Clérigos church.

spatiotemporal information and autoregressive characteristics of trajectories.

To explore the unintentional anomalies, we incorporate the route choice model in the normalizing flow. As such, we develop a method MAFRC-ATD which integrates the data-driven and model-based method. In MAFRC-ATD, the route choice model is exploited to distinguish the unintentionally anomalous trajectories due to traffic congestion or traffic incident.

We evaluate the proposed method in a real-life dataset in Porto, Portugal. The results show that the MAFRC-ATD model can effectively discover anomalous trajectories and successfully discover unintentional trajectories due to traffic congestion.

This study is organized as follows. Section 2 briefly reviews the related work. Problem formulation and proposed methodology are introduced in Section 3. In Section 4, we evaluate our model with a real-life dataset and analyze the results. Section 5 concludes this study and outlines the future scope.

## 2. Related Work

In this section, we will give a brief literature review on anomalous trajectory detection, the probability density estimation with the flow-based model.

**2.1. Anomalous Trajectory Detection.** Trajectory anomaly detection methods can be divided into statistic-based, cluster-based, grid-based, classification-based, and deep learning methods.

**2.1.1. Statistic-Based Method.** With the help of the popular routes, Zhu et al. [8] proposed time-dependent popular routes-based trajectory outlier detection (outlier is also referred to as anomaly detection in this study) method, considering both spatial and temporal anomalies

simultaneously. Liu et al. [9] found that top drivers were better at finding the fastest route than ordinary drivers. However, these algorithms mainly focus on the geometric information of trajectories and neglect semantic information. Thus, Qin et al. [10] proposed the ProbDetect model, which transforms anomaly detection problems into a route choice problem considering semantic information such as trip time and fare, as well as traffic variability. In addition, Kong et al. [11] employed a statistical model to calculate the travel distribution of travel time, driving distance, and travel cost, making it possible to explore the reason for outlier trajectories.

*2.1.2. Cluster-Based Method.* Knorr and Ng [12, 13] and Knorr et al. [14] applied the distance-based algorithm to detect trajectory anomalies in multidimensional multiscale space. However, “few” anomalies may be averaged over the whole trajectories’ dataset. Based on subtrajectories, Lee et al. [15, 16] proposed TRAOD, which includes two phases: partitioning a trajectory into a set of line segments and then detecting outlying line segments. In addition, Bu et al. [17] and Yu et al. [18] built local clusters and neighbor-based trajectory outliers, respectively, to discover the outlier of trajectory streams. In contrast to the above distance-based mechanism, Wang et al. [19] proposed an anomalous trajectory detection method based on edit distance and hierarchical clustering, which was employed to determine the number of clusters.

*2.1.3. Grid-Based Method.* Ge et al. [20] proposed an algorithm Top-Eye, where the continuous data space was discretized into grids. Later, Ge et al. [21] applied travel route evidence and driving distance evidence to the trip fraud detection system, finding that the fraudulent drivers do earn more income than drivers without scams. Additionally, Kong et al. [22] proposed the LoTAD method to mine anomalous regions, where the center of each region and the region’s boundary were determined by the improved k-means method and the Voronoi graph while conforming to traffic conditions. Wang et al. [23] proposed a detection method based on the difference and intersection set (DIS) distance.

Inspired by the isolation mechanism [24] that the anomalous trajectories are susceptible to being isolated, iBAT [25] was proposed to identify the anomalous trajectories. Based on iBAT, Chen et al. [26] and Sun et al. [27] proposed the online model iBOAT and application, respectively.

*2.1.4. Classification Method.* Motion-alert algorithm [28, 29] leveraged the support vector machine (SVM) to learn the features of the extracted pattern motif, which are a sequence of motion features related to time and location. Similarly, Piciarelli and Foresti [30] employed a single-class SVM, and then, Piciarelli et al. [31] introduced the clustering method to improve the robustness.

*2.1.5. Deep Learning.* With the development of the neural network, some researchers have employed deep learning methods, such as recurrent neural networks (RNN) [32, 33], to discover anomalies. Generative models were widely employed to discover anomalous trajectories. GM-VSAE method [34] and IGMM-GAN method [4], for instance, can be applied to uncover outliers in the latent space. In addition, Dias et al. [35] proposed a method using an NF scheme, effectively discovering anomalies in contrast to the GMM method. However, it might be unreasonable to directly flatten multidimensional trajectory data into one dimension and unpractical to evaluate the model in bus data since buses might take bypass routes as well.

*2.2. Probability Density Estimation with the Flow-Based Model.* Many tasks in machine learning, including anomalous trajectory detection, can be formulated as learning and manipulating density distribution. The problem of modeling a probability distribution from a set of given samples is a major goal of probabilistic unsupervised learning or generative modeling [36]. A generative model ought to assign low likelihoods to samples from different distributions, making it attractive to anomaly detection methods.

Generative modeling has undergone tremendous advances in recent years, such as generative adversarial networks (GANs) [37] and variational auto-encoders (VAEs) [38]. Recently, NF methods have improved dramatically. Contrary to the GANs and VAEs, NF methods offer an exact density evaluation of probability density. NF was defined in Tabak and Vanden-Eijnden [39] and Tabak and Turner [40] and later popularized by Rezende and Mohamed [41] and by Dinh et al. [42]. NF methods can be divided into autoregressive flows, linear flows [43], residual flows [44–46], planar flow [41], Sylvester flow [47], and infinitesimal flows [48]. Among these flows, autoregressive flows remain the most popular.

Regardless of its wide application nowadays, autoregressive architecture was even overlooked in previous work [49] due to the huge cost of high dimensionality. To alleviate this problem, several approaches have been proposed, including recurrent mechanism, masked mechanism, and coupling layer. Recurrent mechanism [50] shares parameters by jointly using an RNN. The main drawback of the RNN-based mechanism is that the internal state of RNN is calculated sequentially, with higher computational complexity than the other two mechanisms. The masked mechanism [51] uses a feedforward neural network and removes appropriate connections by multiplying a binary matrix. However, masked mechanism MAF is efficient to evaluate, but slower to invert, indicating that MAF is appropriate to estimate density probability. Coupling layers [42], which splits the input into two disjoint partitions that the first part stays the same as the input and the second part undergoes an affine transformation, which are equally fast to evaluate the probability density and invert for sampling (sampling from the model and evaluating the model’s density are two invertible operations for normalizing the flow model).

**2.3. Route Choice Model.** Route choice modeling is essential to evaluate travelers' choice perception within a set of alternative routes, helping analyze and understand travelers' behavior [52]. Generally, a route choice model has two stages [53]: first is the choice set of possible alternative routes and second is the choice probability of a path, given the route choice set [54, 55]. Route choice models mainly include logit structures [56], GEV structures [57], and non-GEV structures [58].

### 3. Problem Formulation

Problem formulation is as follows.

**3.1. Problem Definition and Framework.** Problem definition and framework is described in the following sections.

**3.1.1. Problem Definition.** As stated by Agrawal and Agrawal [59], the goal of anomaly detection is to find data patterns that differ from what is expected. In our work, such unexpected patterns are related to trajectories sufficiently different from the majority which is considered to be normal, such as detour traces. The definitions of this anomalous trajectory detection problem are as follows.

**Definition 1** (trajectory). A raw trajectory  $T = \{q_1, q_2, \dots, q_n\}$  is a sequence of records, and each recorded  $q_i$  is represented by  $(\text{lon}_i, \text{lat}_i)$ , where  $(\text{lon}_i, \text{lat}_i)$  is a pair of latitude and longitude coordinates.  $q_1$  and  $q_n$  are the origin and destination of the trajectory  $T$ , respectively.

**Definition 2** (anomaly). Anomaly is data that are different from others in the dataset, making people assume that the outlier of these data is not caused by random factors, but by a completely different mechanism [60].

**Definition 3** (anomalous trajectory). Given an OD pair, we assumed that there are a set of trajectories whose pick-up locations and drop-off locations are the origin and destination respectively. Let  $D = \{T_1, T_2, \dots, T_M\}$  denote this set of trajectories, where the anomalous trajectory is significantly different from the majority in this OD pair.

**Definition 4** (unintentionally anomalous trajectory). The unintentionally anomalous trajectory is an anomalous trajectory in a spatiotemporal pattern, which is caused by travelers' unintentional detours to avoid traffic congestion.

**3.1.2. Overview of the Methodology.** To achieve the goal of anomalous trajectory detection, we propose the MAFRC-ATD model, and the methodology is shown in Figure 2.

In detail, we first deal with treating the raw trajectories in the following procedures: (1) removing the invalid trajectories, (2) discretizing the traffic network into equal-sized grids (that is grid mapping), (3) extracting all traces crossing the same OD cell pairs, and (4) partitioning trajectories into subtrajectories.

Based on the subtrajectories, we employ the MAFRC-ATD to discover anomalies. The anomaly score of each trajectory can be computed through the likelihood of subtrajectories and the choice probability of the corresponding route.

Finally, according to the anomaly score, each trajectory can be determined whether it is anomalous and, if anomalous, which type of anomalous trajectory it belongs to.

**3.2. Data Preprocessing.** As stated in Section 1, data preprocessing mainly includes two phases: grid mapping and partitioning trajectories into subtrajectories.

**3.2.1. Grid Mapping.** As shown in Figures 3(a) and 3(b), grid mapping is to split the city traffic network into small grids. In detail, for given  $N_{\text{lon}}$  and  $N_{\text{lat}}$ , the urban traffic network can be split into  $N_{\text{lon}} \times N_{\text{lat}}$  equal-sized grids. Consequently, each taxi trajectory  $T_k = \{q_{k,1}, q_{k,2}, \dots, q_{k,i}, \dots, q_{k,n}\}$  can be represented with a sequence of traversed grids  $T_k = \{t_1, t_2, \dots, t_i, \dots, t_n\}$ , where  $t_i = g(q_{k,i}, N_{\text{lon}}, N_{\text{lat}})$  and  $g$  is a grid mapping function.

**3.2.2. Partitioning Trajectories.** As discussed by Lee et al. [15], the difference between anomalies and normal trajectories might be averaged out over the whole trajectory. To alleviate this problem, trajectories are segmented with subtrajectories. A trajectory  $T_k = \{t_1, t_2, \dots, t_i, \dots, t_n\}$  can be partitioned with equal  $d$ -length sized subtrajectories as follows:

$$\Omega(T_k): \{s_1, s_2, \dots, s_t, \dots, s_l\}, \quad (1)$$

where  $\Omega(T_k)$  is a set of continuous subtrajectories and  $s_l$  can be completed as last  $d$ -length GPS points. Thus, we can obtain the set of subtrajectories  $\Omega(T)$  regarding this OD.

**3.3. Masked Autoregressive Flow.** As stated in Section 1, the anomalous trajectory detection problem can be formulated as the probability density estimation problem. NF ought to assign low likelihoods to data from different distributions, making it possible to discover anomalies with low-likelihood predictions. MAF is applied to detect anomalous trajectories as follows.

NF pushes an initial density (such as normal distribution) through a series of transformations to produce a richer distribution with more expressive power than the constituent components of these transformations [51]. Assuming that  $s$  is a  $d$ -length subtrajectory and  $u$  is the  $d$ -dimensional random variable of the base distribution  $p_u(u)$ , then  $s$  can be represented through the transformation  $F$ :

$$s = F(u), \quad (2)$$

where  $u \sim p_u(u)$ , and the transformation  $F$  must be differentiable and invertible. Then, the density of  $s$ , i.e.,  $p(s)$ , can be calculated as follows:

$$p(s) = p_u(F^{-1}(s)) |\det J_{F^{-1}}(s)|, \quad (3)$$

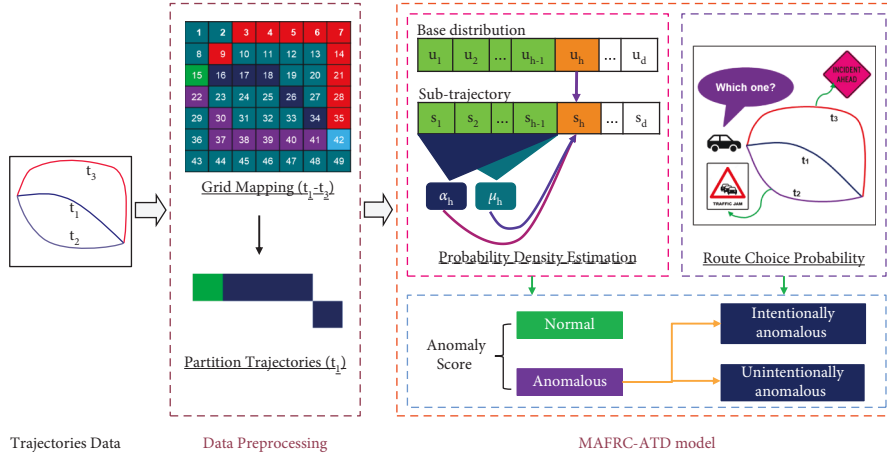
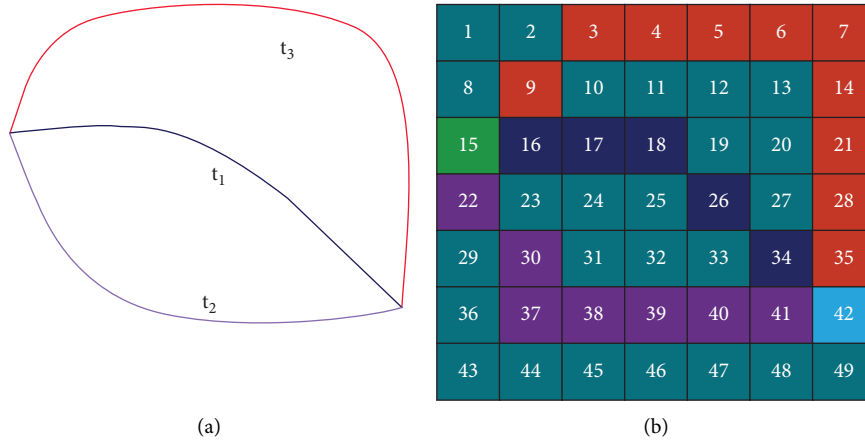


FIGURE 2: The overall architecture of MAFRC-ATD.


 FIGURE 3: An illustration of grid mapping.  $t_1$  to  $t_3$  are shown in Figure 2 (a) An illustration of  $t_1$  to  $t_3$ . (b) Grid mapping.

where the inverse of differentiable transformation  $F^{-1}$  and Jacobian determinant  $|\det J_{F^{-1}}(s)|$  can be implemented by MAF.

Since the spatiotemporal autoregressive character of  $s$ ,  $p(s)$  can be decomposed into a product of conditional densities using the chain rule of probability:

$$p(s) = \prod_{h=1}^d p(s_h | s_1, s_2, \dots, s_{h-1}) = \prod_{h=1}^d p(s_h | s_{1:h-1}). \quad (4)$$

Assume that conditional distributions  $p(s_h | s_{1:h-1})$  are parameterized as single Gaussian:

$$p(s_h | s_{1:h-1}) = N(s_h | \mu_h, (\exp \alpha_h)^2), \quad (5)$$

where  $\mu_h = f_{\mu_h}(s_{1:h-1})$  and  $\alpha_h = f_{\alpha_h}(s_{1:h-1})$ , for  $h = 1, 2, \dots, d$ . Specifically,  $f_{\mu_h}$  and  $f_{\alpha_h}$  are unconstrained scalar functions that compute the mean and log standard deviation given all previous variables, and both two scalar functions are estimated by Germain et al. [61]. The  $h$ th

dimension of subtrajectory, i.e.,  $s_h$ , can be sampled from the base distribution:

$$s_h = u_h \exp \alpha_h + \mu_h, \quad (6)$$

where  $u_h \sim N(0, 1)$ . Thus, the invertible transformation  $F$  is easy to access as follows:

$$u_h = (s_h - \mu_h) \exp(-\alpha_h), \quad (7)$$

where the absolute determinant of the Jacobian matrix is

$$|\det J_{F^{-1}}(s)| = \left| \det \left( \frac{\partial F^{-1}}{\partial s} \right) \right| = \exp \left( - \sum_h \alpha_h \right). \quad (8)$$

Therefore,  $p(s)$  can be successfully evaluated through (3), where the inverse transformation and its Jacobian determinant can be calculated by (6) and (8)

After calculating the likelihood of subtrajectories, the likelihood of the whole trajectory can be calculated as follows. For any trajectory  $T_k$ , the likelihood of the

corresponding subtrajectory  $p(s_t)$  can be obtained via the aforementioned process, and the likelihood of trajectory  $T_k$  can be defined as the mean value of the  $p(s_t)$ , i.e.,

$$p(T_k) = \frac{1}{l} \sum_{t=1}^l (p(s_t)), \quad (9)$$

where  $t = 1, 2, \dots, l$ .

**3.4. Route Choice Analysis.** As mentioned above, we have detected anomalous trajectory with the MAF method, mostly using geometrical information and sequential information. However, in the real world, anomalous trajectories may occur due to lots of factors, such as travel time and traffic conditions. As described in Section 1, detours can be unintentionally generated by drivers whose intentions are to avoid traffic congestion to reduce trip time during peak hours. Apparently, it is extremely crucial to explore this route choice motivation behind anomalous behaviors. However, modeling the motivation behind the choice behavior could be problematic due to the complicated traffic conditions of the road network. The impact of traffic control regulation at an intersection on the Network Macroscopic Fundamental Diagram (NMFD) [62, 63], for example, could also influence the aggregate traffic conditions of the road network. In this study, following Qin et al. [10], we adopt a random utility model based on a multivariate Gaussian model to explore the motivation behind the random choice behaviors.

Assumed there are a set of trajectories  $D = \{T_1, T_2, \dots, T_M\}$  for each pair of OD,  $\mathbb{C} = \{C_1, C_2, \dots, C_m\}$  is a set of choice routes that all distinct trajectories pass through in this OD. For any route  $C_j$  in the route choice set  $\mathbb{C}$ ,  $L_{C_j}$  and  $T_{C_j}$  are its travel length and trip time, and the generalized cost is

$$\mathcal{G}_{C_j} = L_{C_j} + cT_{C_j}, \quad (10)$$

where  $c$  is the velocity equivalent coefficient, which can be interpreted as the trip length value for a unit of the trip time. The choice probability of the route  $C_j$  can be defined as a vectorized form:

$$p_{C_j} = P(\mathbf{Z} \leq 0), \quad (11)$$

where  $\mathbf{Z}$  is the difference of generalized cost, i.e.,  $\mathbf{Z} = (\mathcal{G}_{C_j} - \mathcal{G}_{C_1}, \dots, \mathcal{G}_{C_j} - \mathcal{G}_{C_m})^T$ , and  $\mathbf{Z}$  and  $0$  are both  $R^{(m-1) \times 1}$  vectors. Thus,  $p_{C_j}$  can be calculated by the integral of the probability density function for  $\mathbf{Z}$ :

$$p_{C_j} = \int_{-\infty}^0 \cdots \int_{-\infty}^0 \frac{\exp(-(1/2)(\mathbf{Z} - \mu_{\mathbf{Z}})^T \Sigma^{-1} (\mathbf{Z} - \mu_{\mathbf{Z}}))}{\sqrt{(2\pi)^{m-1} |\Sigma_{\mathbf{Z}}|}} d\mathbf{Z}, \quad (12)$$

where  $\mu_{\mathbf{Z}}$  and  $\Sigma_{\mathbf{Z}}$  are the mean vector and covariance matrix of  $\mathbf{Z}$ . To calculate  $\mu_{\mathbf{Z}}$  and  $\Sigma_{\mathbf{Z}}$ ,  $\mathbf{Z}$  are further decomposed into as follows:

$$\mathbf{Z} = \mathbf{A}\mathbf{U}^T, \quad (13)$$

where  $\mathbf{A} = \begin{bmatrix} -1 & \cdots & 1 & \cdots & 0 \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & -1 \end{bmatrix}_{(m-1) \times m}$  and

$$\mathbf{U} = [\mathcal{G}_{C_1}, \mathcal{G}_{C_2}, \dots, \mathcal{G}_{C_m}]. \text{ Thus,} \quad (14)$$

$$\begin{cases} \mu_{\mathbf{Z}} = \mathbf{A}\mu_{\mathbf{U}}, \\ \Sigma_{\mathbf{Z}} = \mathbf{A}\Sigma_{\mathbf{U}}\mathbf{A}^T, \end{cases}$$

where  $\mu_{\mathbf{Z}}$  and  $\Sigma_{\mathbf{Z}}$  can be derived by  $\mu_{\mathbf{U}}$  and  $\Sigma_{\mathbf{U}}$ . Obviously,

$$\mu_{\mathbf{U}} = (\mu_{\mathcal{G}_{C_1}}, \mu_{\mathcal{G}_{C_2}}, \dots, \mu_{\mathcal{G}_{C_m}}), \quad (15)$$

and  $\Sigma_{\mathbf{U}}$  is

$$\Sigma_{\mathbf{U}} = \begin{bmatrix} \text{cov}(\mathcal{G}_{C_1}, \mathcal{G}_{C_1}) & \text{cov}(\mathcal{G}_{C_1}, \mathcal{G}_{C_2}) & \cdots & \text{cov}(\mathcal{G}_{C_1}, \mathcal{G}_{C_m}) \\ \text{cov}(\mathcal{G}_{C_2}, \mathcal{G}_{C_1}) & \text{cov}(\mathcal{G}_{C_2}, \mathcal{G}_{C_2}) & \cdots & \text{cov}(\mathcal{G}_{C_2}, \mathcal{G}_{C_m}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(\mathcal{G}_{C_m}, \mathcal{G}_{C_1}) & \text{cov}(\mathcal{G}_{C_m}, \mathcal{G}_{C_2}) & \cdots & \text{cov}(\mathcal{G}_{C_m}, \mathcal{G}_{C_m}) \end{bmatrix}. \quad (16)$$

Furthermore,  $\Sigma_{\mathbf{U}}$  is given by

$$\Sigma_{\mathbf{U}} = \begin{bmatrix} \sigma_{C_1}^2 & \sigma_{C_1 C_2}^2 & \cdots & \sigma_{C_1 C_m}^2 \\ \sigma_{C_2 C_1}^2 & \sigma_{C_2}^2 & \cdots & \sigma_{C_2 C_m}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{C_m C_1}^2 & \sigma_{C_m C_2}^2 & \cdots & \sigma_{C_m}^2 \end{bmatrix}, \quad (17)$$

where  $\sigma_{C_i C_j}^2$  is the variance of  $C_i C_j$  which reflects the overlapped portion of two routes  $C_i$  and  $C_j$ , and  $C_i$  and  $C_j$  are arbitrary routes in route set  $\mathbb{C}$ , for  $i, j = 1, 2, \dots, m$ .

In this way, the choice probability of the route  $C_j$  can be obtained by equation (12) by substituting equations (15) and (17) into equation (14).

**3.5. Anomaly Score.** As described in Section 1 and Section 3.1, the MAFRC-ATD model integrates the data-driven learning and route choice model. Thereby, the anomaly score is designed to combine the two aspects as follows.

The likelihood of a given trajectory and the route choice probability of the given trajectory can be calculated in Sections 3.3 and 3.4. For a given OD pair, we assumed that  $C_j$  is the corresponding route of trajectory  $T_k$ ,  $p(T_k)$  and  $p_{C_j}$  are the likelihood of  $T_k$  and choice probability of route  $C_j$  respectively, and an anomaly score  $E(T_k)$  is defined as

$$E(T_k) = -\left(p_{C_j} - \frac{1}{m}\xi\right) \cdot \text{sigmoid}(\log p(T_k)) + 1, \quad (18)$$

where  $m$  is the number of all routes for this OD pair and  $\xi$  is the coefficient of unintentionally anomalous. In (18),  $\text{sigmoid}(\cdot)$  is a sigmoid function, i.e.,  $\text{sigmoid}(x) = (1/1 + e^{-x})$ , and is monotonically increasing

with outputs in the range 0 to 1. Obviously,  $p_{C_j}$  and  $p(T_k)$  of normal trajectories are higher than anomalous trajectories.

Assuming that threshold  $\delta$  is applied to decide whether the trajectory is anomalous, there are three cases.

When  $E(T_k) \geq 1$ , i.e.,  $p_{C_j} - (1/m)\xi \leq 0$ ,  $T_k$  is likely to be an intentionally anomalous trajectory

When  $E(T_k) < 1$  and  $E(T_k) \geq \delta$ ,  $T_k$  can be categorized as unintentionally anomalous trajectory

When  $E(T_k) > 0$  and  $E(T_k) < \delta$ ,  $T_k$  can be regarded as a normal trajectory

## 4. Case Study

In this section, we will evaluate our MAFRC-ATD model using the real-world dataset and present the results as well as discussions of experiments.

### 4.1. Experimental Setup

**4.1.1. Dataset Preprocessing.** Our experiments are conducted on the real-world taxi trajectory dataset provided by Kaggle ECML-PKDD 2015 competition (<https://www.kaggle.com/c/pkdd-15-predict-taxi-service-trajectory-i/data>). In detail, the dataset is generated from 442 taxis running in the city of Porto from 01/07/2013 to 30/06/2014. There are 1710670 raw trajectories in total and each taxi reports its location every 15 seconds.

Out of approximately 1.7 million trajectories from the original dataset, we remove the trajectories with less than 20 GPS. After filtering, we split the city traffic network space into  $100m \times 100m$  grids, i.e.,  $g_l = g_w = 100m$ . Consequently, GPS points can be represented with grids and each trajectory can be denoted by a sequence of grids. Subsequently, we extract 4 OD pairs, denoted as  $OD_1$  to  $OD_4$ , respectively. In this study, we segment a trajectory into subtrajectories with 20 GPS points, that is,  $d = 20$ . Given the fact that there is no labeled dataset available, we choose to manually label anomalies in this paper.

Additionally, inspired by Liu et al. [34], to alleviate the problem of unbalanced learning in anomaly detection, we select several detour paths from which anomalous trajectories are generated. The detour paths, called generated routes, traverse the real-world links and are implemented in ArcGIS and Google maps. Among the generated routes, generated anomalous trajectories can be generated through sampling from the GPS points of generated routes. The generated anomalous trajectories are injected into the original dataset for “data argument.”

As stated above, the related information of those OD pairs is shown in Table 1, where O and D are represented with (longitude and latitude), and each OD pair can be shown in Figure 4. Out of all testing trajectories, the coral color lines represent the original historical dataset in the testing dataset, whereas other colors represent the generated routes from which the generated anomalous trajectories are sampled.

**4.1.2. Baseline.** In this study, we will evaluate our model compared with the following methods.

iBAT [25]: iBAT exploits an isolation mechanism that anomalous trajectories are susceptible to, i.e., anomalies are easy to be separated from the majority of the trajectories since anomalies are “few and different.” It is more likely to be anomalous for a higher anomaly score.

ATDC [23]: ATDC proposes a distance metric, namely, the DIS distance, to evaluate the similarity between any two trajectories. Based on the DIS distance, ATDC can further calculate the anomalous score of a trajectory and then determines if it is an anomalous trajectory.

**4.1.3. Evaluation Criteria.** In this study, we employ the area under the receiver operating characteristic curve (AU-ROC) as the evaluation metric. To evaluate the performance of the anomaly detection method, detection rate (DR) and false alarm rate (FAR) (in this study, we refer to true positive rate (TPR) and false positive rate (FPR) as detection rate and false alarm rate respectively) are two important metrics. As shown in the confusion matrix Table 2, trajectories are labeled either as anomalous (i.e., positive) or normal (i.e., negative).

A trajectory will be classified as one of our possible outcomes: (1) true positive (TP): an anomalous trajectory is correctly predicted as anomalous; (2) false positive (FP): a trajectory is predicted as anomalous while it is normal; (3) false negative (FN): a trajectory is predicted as normal while it is anomalous; (4) true negative (TN): a normal trajectory is correctly predicted as normal. The DR measures the rate of correctly predicted anomalous trajectories among all real anomalies, i.e.,

$$DR = \frac{TP}{TP + FN}. \quad (19)$$

The FAR is defined as the proportion of normal trajectories that are predicted as anomalous, i.e.,

$$FAR = \frac{FP}{FP + TN}. \quad (20)$$

The better anomaly detection method will have a higher DR and a lower FAR. The ROC curve shows the detection rate (Y-axis) against the false alarm rate (X-axis) at various thresholds, which depicts the tradeoff between false alarms and accurate classifications. AU-ROC value is the area under the ROC curve, and the anomaly detection method is of better quality for higher AU-ROC.

## 4.2. Results

**4.2.1. Performance Evaluation on Anomalous Trajectory Detection.** We evaluate the results on OD pair with the above baselines. Taking the instance of  $OD_2$ , the results of anomalous trajectory detection in  $OD_2$  are shown in Figure 5. As shown in the figure, the MAFRC-ATD model shows the highest detection rate under a given false alarm

TABLE 1: The information on OD pairs.

|                            | $OD_1$             | $OD_2$             | $OD_3$             | $OD_4$             |
|----------------------------|--------------------|--------------------|--------------------|--------------------|
| O                          | (-8.6399, 41.1598) | (-8.6109, 41.1457) | (-8.6141, 41.1412) | (-8.5856, 41.1485) |
| D                          | (-8.6542, 41.1806) | (-8.6098, 41.1605) | (-8.6497, 41.1544) | (-8.6068, 41.1457) |
| Historical trajectories    | 996                | 946                | 687                | 559                |
| Training dataset           | 897                | 852                | 619                | 504                |
| Testing dataset            |                    |                    |                    |                    |
| Historical testing dataset | 99                 | 94                 | 68                 | 55                 |
| Generated trajectories     | 12                 | 9                  | 6                  | 9                  |
| Anomalous trajectories     | 15                 | 10                 | 10                 | 10                 |
| Anomaly ratio (%)          | 13.51              | 9.71               | 13.51              | 15.63              |

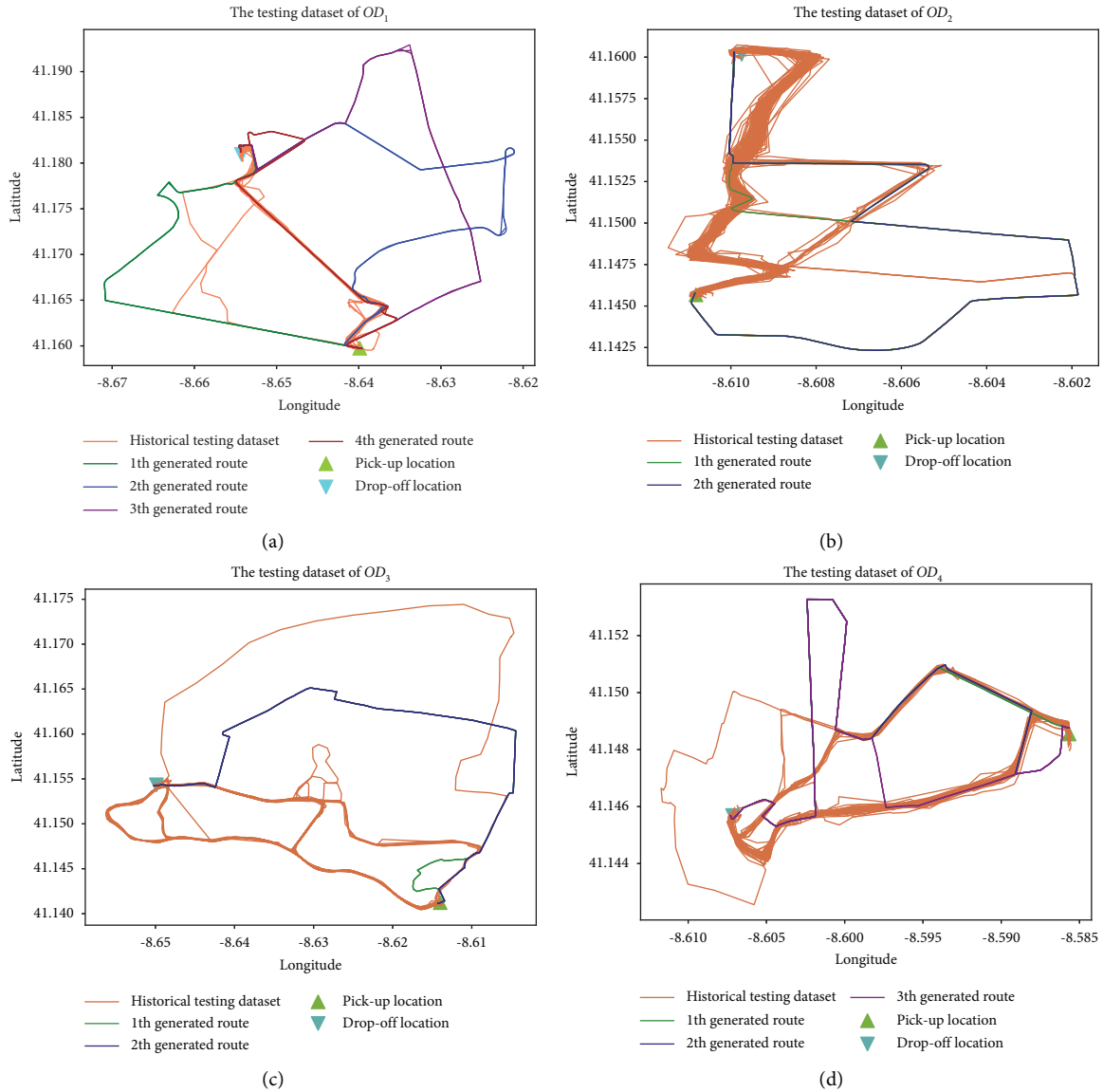


FIGURE 4: The testing dataset of each OD pair, where the lawn green color upper triangle represents the pick-up location and the cyan color lower triangle represents the drop-off location. (a)  $OD_1$ , (b)  $OD_2$ , (c)  $OD_3$ , and (d)  $OD_4$ .

rate, indicating that it can effectively detect the anomalies. Similarly, the MAFRC-ATD model possessed of the lowest false alarm for a given detection rate, indicating that it has less potential for mistaken detection. For the overall curve, the ROC curve of the MAFRC-ATD model covers other the

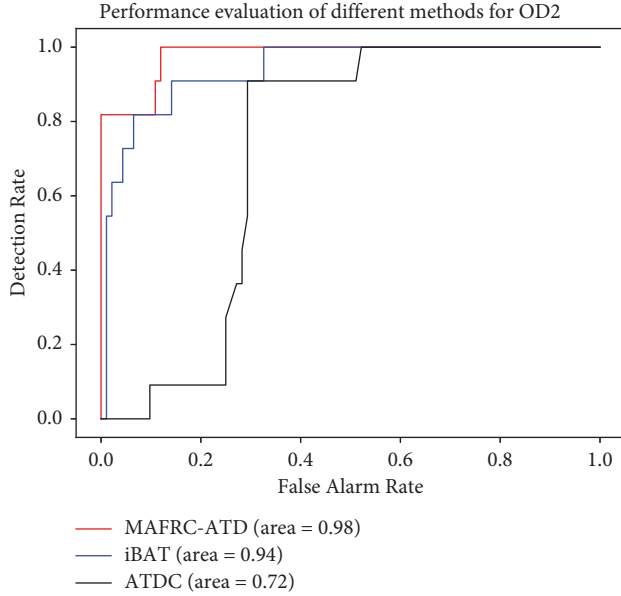
other two curves, suggesting that the MAFRC-ATD method performs better and has expressive power on anomalous trajectory detection.

To explore the influence of anomalous trajectory detection on different OD pairs instead of a given OD pair, we



TABLE 2: The confusion matrix of the binary classification problem.

| Prediction result | Real value |        |
|-------------------|------------|--------|
|                   | Anomalous  | Normal |
| Anomalous         | TP         | FP     |
| Normal            | FN         | TN     |

FIGURE 5: The performance evaluation on anomalous trajectories detection regarding  $OD_2$ .

conduct experiments on 4 OD pairs. In detail, we train the model with the training dataset of each OD pair and then evaluate the testing dataset, respectively. OD pairs and the final results are shown in Table 3. The results indicate that MAFRC-ATD achieves the best results over the 4 OD pairs and has the potential power to be well-suited for different situations.

The experiments indicate that the MAFRC-ATD model could effectively capture the autoregressive characteristics of trajectories and model the probability density of sub-trajectories. However, the reason for the undesirable results of iBAT and ATDC might be that those methods are based on distance or isolation mechanisms, mainly focusing on geometric information of the trajectory and neglecting the sequential information of the trajectory. Obviously, iBAT is well adapted to identify various detours, but may not be suitable for situations where anomalies are intertwined with normal ones ( $OD_4$ ). ATDC performs a little terribly when existing local anomalous trajectories (such as  $OD_2$  and  $OD_3$ ); it applies to intertwined situations.

**4.2.2. Unintentionally Anomalous Trajectory Analysis.** In addition to the performance of anomalous trajectories detection, discovering unintentionally anomalous trajectories is also essential. Route choice analysis is employed to discover the unintentional anomalies as follows.

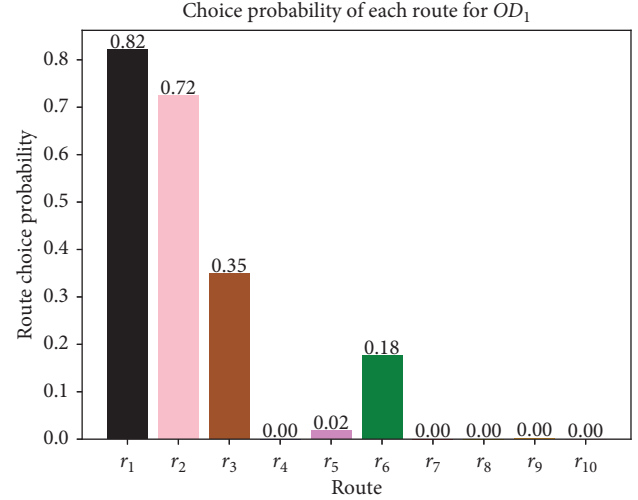
FIGURE 6: The route choice probability of each route in  $OD_1$ .

TABLE 3: Performance evaluation on anomalous trajectory detection in contrast to baseline.

| Methods   | OD     |        |        |        |
|-----------|--------|--------|--------|--------|
|           | $OD_1$ | $OD_2$ | $OD_3$ | $OD_4$ |
| MAFRC-ATD | 1.00   | 0.98   | 1.00   | 1.00   |
| iBAT      | 0.92   | 0.94   | 0.92   | 0.62   |
| ATDC      | 0.82   | 0.72   | 0.83   | 0.93   |

Taking  $OD_1$ , for instance, there are 10 routes for this OD pair, denoted as  $r_1$  to  $r_{10}$ , respectively, i.e.,  $m = 10$  and  $\mathbb{C} = \{r_j, \text{for } j = 1, 2, \dots, 10\}$ . Trajectory set as well as link information for each route regarding  $OD_1$  can be shown in Table 4. Among all paths,  $r_1$  contains the most amount of trajectories. For the sake of convenience, “(..)” in  $r_1$  represents the trajectories of  $OD_1$  except what are contained in  $r_2$ - $r_{10}$ . ?? shows each route in the  $OD_1$ .

In this study, we set velocity equivalent coefficient  $c$  as 8.91 m/s according to the mean speed value of historical trips. For any given link in  $OD_1$ , the generalized cost of each route is calculated as the mean value of the generalized costs of all trajectories corresponding to this route.

When  $\delta$  sets to 0.02 and  $\xi$  sets to 1.5, experiments show that the trajectories corresponding to  $\mathbb{C}_n = \{C_1, C_2, C_3\}$  are normal, whereas trajectories to  $\mathbb{C}_{an} = \{C_j, \text{for } j = 4, 5, \dots, 10\}$  are anomalous. Among all anomalous trajectories, the trajectory corresponding to  $\mathbb{C}_{un} = \{r_6\}$ , i.e., 82nd trajectory, is unintentionally anomalous, whereas trajectories corresponding to  $\mathbb{C}_{un} = \mathbb{C}_{an}/r_6$  are intentionally anomalous. Simply we can find that the 82nd trajectory happened on the evening peak hours, Tuesday. The choice probability of the 82nd trajectory is large enough, i.e., greater than  $1/m\xi$ , to make the anomaly score less than 1 despite greater than  $\delta$ . Henceforth, we have sufficient confidence that the 82nd trajectory (i.e., corresponding to  $r_6$ ) has the strong potential to be chosen when taking this trip. This result indicates that the 82nd trajectory could be

TABLE 4: The trajectory index set and link information regarding  $OD_1$ .

| Route | Trajectory index set | Link set                               |
|-------|----------------------|--|
| 1     | [...]                | [1, 2, 3, 4, 5, 6, 7, 8, 9]            |
| 2     | [12]                 | [1, 10, 18, 11, 4, 5, 6, 7, 8, 9]      |
| 3     | [7, 38, 83]          | [1, 2, 12, 13, 6, 7, 8, 9]             |
| 4     | [54]                 | [1, 14, 15, 16, 17, 7, 8, 9]           |
| 5     | [81]                 | [1, 10, 19, 1, 2, 3, 4, 5, 6, 7, 8, 9] |
| 6     | [82]                 | [1, 14, 20, 21, 22, 7, 8, 9]           |
| 7     | [99, 100, 101]       | [1, 14, 15, 23, 24, 22, 7, 8, 9]       |
| 8     | [102, 103, 104]      | [1, 2, 3, 4, 25, 26, 27, 28, 7, 8, 9]  |
| 9     | [105, 106, 107]      | [1, 10, 29, 27, 28, 8, 9]              |
| 10    | [108, 109, 110]      | [1, 10, 18, 11, 4, 5, 6, 7, 28, 30, 9] |

unintentionally generated by drivers whose intentions are to avoid traffic congestion to reduce trip time during peak hours. The results agree well with the route choice probability result, as shown in Figure 6.

## 5. Conclusion

Anomaly trajectory detection is crucial for preserving the legitimate interests of passengers and discovering the pattern of anomalous driving behaviors. In this study, we propose an unsupervised detection framework, namely, MAFRC-ATD, to detect anomalous trajectories. Compared to traditional methods, MAFRC-ATD can carry out the following:

- (1) Taking the spatiotemporal information of trajectory data into consideration, the MAFRC-ATD model exploits the autoregressive character of trajectories and effectively captures the spatiotemporal information of trajectories, making it possible to assign lower likelihoods to anomalous trajectories
- (2) Incorporating the route choice model to discover the unintentional anomalous due to traffic congestion, route choice analysis is employed to further explore the motivation behind random choice behaviors, such as traffic congestion

Moreover, multi-od experiments are also conducted to evaluate the generalization properties of the MAFRC-ATD model. Experiments on the real-world dataset demonstrate the effectiveness of our methods over baseline methods. Furthermore, testing the proposed method with more real-world datasets is crucial to evaluate the applicability and robustness of the method, which is left for our future study.

In the future, we will extend our model online with time-series methods, such as the self-attention mechanism. In addition, the ensembling-based NF model could perform better to alleviate the imbalance learning problem.

## Data Availability

Our experiments are conducted on the real-world taxi trajectory dataset provided by Kaggle ECML-PKDD 2015 competition (<https://www.kaggle.com/c/pkdd-15-predict-taxi-service-trajectory-i/data>).

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This study was supported by the Shenzhen Science and Technology Program (Grant no. 2021A29).

## References

- [1] V. Dobnik, "NY agency: 600-plus Cabbies Repeatedly Overcharged," 2010, [https://archive.boston.com/news/nation/articles/2010/05/15/officials\\_fraction\\_of\\_nyc\\_cab\\_drivers\\_overcharge/](https://archive.boston.com/news/nation/articles/2010/05/15/officials_fraction_of_nyc_cab_drivers_overcharge/).
- [2] T. Namako, "Taxi Passengers Taken for a Ride to the Tune of \$ 1.1 Million: Officials," 2010, <https://nypost.com/2010/05/14/taxi-passengers-taken-for-a-ride-to-the-tune-of-1-1-million-officials/>.
- [3] A. Denney and R. Fenton, "Creepy Uber Driver Who Groped Sleeping Passenger Gets 3-year Sentence," 2019, <https://nypost.com/2019/06/24/creepy-uber-driver-who-groped-sleeping-passenger-gets-3-year-sentence/>.
- [4] D. Smolyak, K. Gray, S. Badirli, and G. Mohler, "Coupled igmm-gans with applications to anomaly detection in human mobility data," *ACM Transactions on Spatial Algorithms and Systems (TSAS)*, vol. 6, no. 4, pp. 1–14, 2020.
- [5] M. Sabuhi, M. Zhou, C.-P. Bezemer, and P. Musilek, "Applications of generative adversarial networks in anomaly detection," *A Systematic Literature Review. IEEE Access*, vol. 9, 2021.
- [6] G. Zhu, H. Zhao, H. Liu, and H. Sun, "A novel lstm-gan algorithm for time series anomaly detection," in *Proceedings of the 2019 Prognostics and System Health Management Conference (PHM-Qingdao)*, pp. 1–6, IEEE, Qingdao, China, October 2019.
- [7] J.-H. Oh, J. Y. Hong, and J.-G. Baek, "Oversampling method using outlier detectable generative adversarial network," *Expert Systems with Applications*, vol. 133, pp. 1–8, 2019.
- [8] J. Zhu, W. Jiang, A. Liu, G. Liu, and L. Zhao, "Time-dependent popular routes based trajectory outlier detection," in *International Conference on Web Information Systems Engineering*, pp. 16–30, Springer, 2015.
- [9] L. Liu, C. Andris, and C. Ratti, "Uncovering cabdrivers' behavior patterns from their digital traces," *Computers, Environment and Urban Systems*, vol. 34, no. 6, pp. 541–548, 2010.
- [10] G. Qin, Z. Huang, Y. Xiang, and J. Sun, "Probdetect: a choice probability-based taxi trip anomaly detection model considering traffic variability," *Transportation Research Part C: Emerging Technologies*, vol. 98, pp. 221–238, 2019.
- [11] X. Kong, B. Zhu, G. Shen et al., "Spatial-temporal-cost combination based taxi driving fraud detection for collaborative internet of vehicles," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3426–3436, 2022.
- [12] E. M. Knorr and R. T. Ng, "Algorithms for mining distance-based outliers in large datasets," *Proceedings of the 24th International Conference on Very Large Data Bases*, vol. 98pp. 392–403, Citeseer, 1998.
- [13] E. M. Knorr and R. T. Ng, "Finding intensional knowledge of distance-based outliers," *VLDB*, vol. 99, pp. 211–222, 1999.

- [14] E. M. Knorr, R. T. Ng, and V. Tucakov, "Distance-based outliers: algorithms and applications," *The VLDB Journal—The International Journal on Very Large Data Bases*, vol. 8, no. 3-4, pp. 237–253, 2000.
- [15] J.-G. Lee, J. Han, and X. Li, "Trajectory outlier detection: a partition-and-detect framework," in *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*, pp. 140–149, IEEE, Cancun, Mexico, April 2008.
- [16] J.-G. Lee, J. Han, and K.-Y. Whang, "Trajectory clustering: a partition-and-group framework," in *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, pp. 593–604, ACM, Beijing, China, June 2007.
- [17] Y. Bu, L. Chen, A. W.-C. Fu, and D. Liu, "Efficient anomaly monitoring over moving object trajectory streams," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 159–168, ACM, Paris, France, June 2009.
- [18] Y. Yu, L. Cao, E. A. Rundensteiner, and Q. Wang, "Detecting moving object outliers in massive-scale trajectory streams," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 422–431, ACM, New York, NY, USA, August 2014.
- [19] Y. Wang, K. Qin, Y. Chen, and P. Zhao, "Detecting anomalous trajectories and behavior patterns using hierarchical clustering from taxi gps data," *ISPRS International Journal of Geo-Information*, vol. 7, no. 1, p. 25, 2018.
- [20] Y. Ge, H. Xiong, Z.-h. Zhou, H. Ozdemir, J. Yu, and K. C. Lee, "Top-eye: top-k evolving trajectory outlier detection," in *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, pp. 1733–1736, ACM, Toronto, ON, Canada, October 2010.
- [21] Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou, "A taxi driving fraud detection system," in *Proceedings of the 2011 IEEE 11th International Conference on Data Mining*, pp. 181–190, IEEE, Vancouver, BC, Canada, December 2011.
- [22] X. Kong, X. Song, F. Xia, H. Guo, J. Wang, and A. Tolba, "Lotad: long-term traffic anomaly detection based on crowdsourced bus trajectory data," *World Wide Web*, vol. 21, no. 3, pp. 825–847, 2018.
- [23] J. Wang, Y. Yuan, T. Ni et al., "Anomalous trajectory detection and classification based on difference and intersection set distance," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 2487–2500, 2020.
- [24] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining*, pp. 413–422, IEEE, Pisa, Italy, December 2008.
- [25] D. Zhang, N. Li, Z.-H. Zhou, C. Chen, L. Sun, and S. Li, "ibat: detecting anomalous taxi trajectories from gps traces," in *Proceedings of the 13th International Conference on Ubiquitous Computing*, pp. 99–108, ACM, Beijing, China, September 2011.
- [26] C. Chen, D. Zhang, P. S. Castro et al., "iboat: isolation-based online anomalous trajectory detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 806–818, 2013.
- [27] L. Sun, D. Zhang, C. Chen, P. S. Castro, S. Li, and Z. Wang, "Real time anomalous trajectory detection and analysis," *Mobile Networks and Applications*, vol. 18, no. 3, pp. 341–356, 2013.
- [28] X. Li, J. Han, and S. Kim, "Motion-alert: automatic anomaly detection in massive moving objects," in *International Conference on Intelligence and Security Informatics*, pp. 166–177, Springer, Berlin, Heidelberg, 2006.
- [29] X. Li, J. Han, S. Kim, and H. Gonzalez, "Roam: rule-and motif-based anomaly detection in massive moving object data sets," in *Proceedings of the 2007 SIAM International Conference on Data Mining*, pp. 273–284, SIAM, Minneapolis, Minnesota, USA, 2007.
- [30] C. Piciarelli and G. L. Foresti, "Anomalous trajectory detection using support vector machines," in *Proceedings of the 2007 IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 153–158, IEEE, Washington, DC, USA, January 2007.
- [31] C. Piciarelli, C. Micheloni, and G. L. Foresti, "Anomalous Trajectory Patterns Detection," in *Proceedings of the 2008 19th International Conference on Pattern Recognition*, pp. 1–4, IEEE, Tampa, FL, USA.
- [32] L. Song, R. Wang, D. Xiao, X. Han, Y. Cai, and C. Shi, "Anomalous trajectory detection using recurrent neural network," in *International Conference on Advanced Data Mining and Applications*, pp. 263–277, Springer, Berlin, Heidelberg, 2018.
- [33] S. E. Huang, Y. Feng, and H. X. Liu, "A data-driven method for falsified vehicle trajectory identification by anomaly detection," *Transportation Research Part C: Emerging Technologies*, vol. 128, Article ID 103196, 2021.
- [34] Y. Liu, K. Zhao, G. Cong, and Z. Bao, "Online anomalous trajectory detection with deep generative sequence modeling," in *Proceedings of the 2020 IEEE 36th International Conference on Data Engineering (ICDE)*, pp. 949–960, IEEE, Dallas, TX, USA, April 2020.
- [35] M. L. Dias, C. L. C. Mattos, T. L. da Silva, J. A. F. de Macedo, and W. C. Silva, "Anomaly detection in trajectory data with normalizing flows," in *Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, Glasgow, UK, July 2020.
- [36] I. Kobayev, S. J. Prince, and M. A. Brubaker, "Normalizing flows: an introduction and review of current methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 11, pp. 3964–3979, 2021.
- [37] I. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," *Advances in Neural Information Processing Systems*, vol. 27, 2014.
- [38] D. P. Kingma and M. Welling, "Auto-encoding Variational Bayes," 2013, <https://arxiv.org/abs/1910.00292>.
- [39] E. G. Tabak and E. Vanden-Eijnden, "Density estimation by dual ascent of the log-likelihood," *Communications in Mathematical Sciences*, vol. 8, no. 1, pp. 217–233, 2010.
- [40] E. G. Tabak and C. V. Turner, "A family of nonparametric density estimation algorithms," *Communications on Pure and Applied Mathematics*, vol. 66, no. 2, pp. 145–164, 2013.
- [41] D. Rezende and S. Mohamed, "Variational inference with normalizing flows," in *Proceedings of the International Conference on Machine Learning*, pp. 1530–1538, PMLR, Baltimore, Maryland, USA, June 2015.
- [42] L. Dinh, D. Krueger, and Y. Bengio, "Nice: non-linear independent components estimation," 2014, <https://arxiv.org/abs/1410.8516>.
- [43] D. P. Kingma and P. Dhariwal, "Glow: generative flow with invertible 1x1 convolutions," *Advances in Neural Information Processing Systems*, vol. 31, 2018.
- [44] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, IEEE, Las Vegas, Nevada, USA, May 2016.

- [45] R. T. Chen, J. Behrmann, D. Duvenaud, and J.-H. Jacobsen, "Residual flows for invertible generative modeling," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [46] J. Behrmann, W. Grathwohl, R. T. Chen, D. Duvenaud, and J.-H. Jacobsen, "Invertible residual networks," in *Proceedings of the International Conference on Machine Learning*, pp. 573–582, PMLR, Beach, California, USA, 2019.
- [47] R. v. d. Berg, L. Hasenclever, J. M. Tomczak, and M. Welling, "Sylvester Normalizing Flows for Variational Inference," 2018, <https://arxiv.org/abs/1803.05649>.
- [48] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, "Neural Ordinary Differential Equations," *Advances in neural information processing systems*, vol. 31, 2018.
- [49] S. Chen and R. Gopinath, "Gaussianization," *Advances in Neural Information Processing Systems*, vol. 13, pp. 423–429, 2000.
- [50] J. Oliva, A. Dubey, M. Zaheer et al., "Transformation autoregressive networks," in *Proceedings of the International Conference on Machine Learning*, pp. 3898–3907, PMLR, Paris, France, 2018.
- [51] G. Papamakarios, T. Pavlakou, and I. Murray, "Masked autoregressive flow for density estimation," *Advances in Neural Information Processing Systems*, vol. 30, pp. 2338–2447, 2017.
- [52] C. G. Prato, "Route choice modeling: past, present and future research directions," *Journal of choice modelling*, vol. 2, no. 1, pp. 65–100, 2009.
- [53] S. Bekhor, M. E. Ben-Akiva, and M. S. Ramming, "Evaluation of choice set generation algorithms for route choice models," *Annals of Operations Research*, vol. 144, no. 1, pp. 235–247, 2006.
- [54] A. I. Croce, G. Musolino, C. Rindone, and A. Vitetta, "Route and path choices of freight vehicles: a case study with floating car data," *Sustainability*, vol. 12, no. 20, p. 8557, 2020.
- [55] G. Musolino, C. Rindone, and A. Vitetta, "A modelling framework to simulate paths and routes choices of freight vehicles in sub-urban areas," in *Proceedings of the 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pp. 1–6, IEEE, Heraklion, Greece, June 2021.
- [56] M. Ben-Akiva and M. Bierlaire, "Discrete choice methods and their applications to short term travel decisions," in *Handbook of Transportation Science*, pp. 5–33, Springer, Berlin, Heidelberg, 1999.
- [57] S. Bekhor and J. N. Prashker, "Stochastic user equilibrium formulation for generalized nested logit model," *Transportation Research Record*, vol. 1752, no. 1, pp. 84–90, 2001.
- [58] C. F. Daganzo and Y. Sheffi, "On stochastic models of traffic assignment," *Transportation Science*, vol. 11, no. 3, pp. 253–274, 1977.
- [59] S. Agrawal and J. Agrawal, "Survey on anomaly detection using data mining techniques," *Procedia Computer Science*, vol. 60, pp. 708–713, 2015.
- [60] D. M. Hawkins, "Identification of outliers," *London: Chapman and Hall*, Springer, vol. 11, , 1980.
- [61] M. Germain, K. Gregor, I. Murray, and H. Larochelle, "Made: masked autoencoder for distribution estimation," in *Proceedings of the International Conference on Machine Learning*, pp. 881–889, PMLR, Paris, France, 2015.
- [62] B. Alonso, Á. Ibeas, G. Musolino, C. Rindone, and A. Vitetta, "Effects of traffic control regulation on network macroscopic fundamental diagram: a statistical analysis of real data," *Transportation Research Part A: Policy and Practice*, vol. 126, pp. 136–151, 2019.
- [63] B. Alonso, Á. I. Pòrtilla, G. Musolino, C. Rindone, and A. Vitetta, "Network fundamental diagram (nfd) and traffic signal control: first empirical evidences from the city of santander," *Transportation Research Procedia*, vol. 27, pp. 27–34, 2017.