Research Article

Research on Real-Time Anomaly Detection of Fishing Vessels in a Marine Edge Computing Environment

Jie Huang ^{1,2}, Fengwei Zhu,² Zejun Huang,² Jian Wan,¹ and Yongjian Ren²

¹School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China

²School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China

Correspondence should be addressed to Jie Huang; huangjie@zust.edu.cn

Received 9 February 2021; Revised 14 April 2021; Accepted 24 April 2021; Published 5 May 2021

Academic Editor: Xiaoxian Yang

Copyright © 2021 Jie Huang 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.

Fishing vessel monitoring systems (VMSs) play an important role in ensuring the safety of fishing vessel operations. Traditional VMSs use a cloud centralized computing model, and the storage, processing, and visualization of all fishing vessel data are completed in the monitoring center. Due to the limitation of maritime communications, the data generated by fishing vessels cannot be fully utilized, and communication delays lead to inadequate warnings in cases of fishing vessel abnormalities. In this paper, we present a real-time anomaly detection model (RADM) for fishing vessels based on edge computing. The model runs in the edge layer, making full use of the information of moving edge nodes and nearby nodes, and combines a historical trajectory extraction detection model with an online anomaly detection model to detect anomalies. The detection model of historical trajectory extraction mines frequent patterns in historical trajectories through multifeature clustering and identifies trajectories that are different from the frequent patterns as anomalies. Online anomaly detection algorithms detect anomalous behavior in specific scenarios based on the spatiotemporal neighborhood similarity and reduce the impact of anomaly evolution. Experiments show that RADM was more effective than traditional methods in real-time anomaly detection of fishing vessels, which provides a new method for upgrading the technology of traditional VMS.

1. Introduction

At present, fishing vessel monitoring systems (VMSs) are widely used in fishing vessel safety management. In VMSs, the position and status information of each fishing vessel is collected and recorded by shipborne sensors at a certain time interval and then sent back to a monitoring center, and thus a series of spatiotemporal data points will form a trajectory data set. The trajectory data set is the core data of the VMS and has very important application prospects. For example, in the process of sailing, fishing vessels may face unpredictable and/or abnormal conditions. In cases of equipment failure (radar, positioning equipment, etc.), bad weather (typhoon, etc.), or even terrorist events (such as hijacking by pirates), the monitoring center can timely identify the abnormal conditions using the trajectory data and then extract the information necessary to take measures to maintain and guarantee the safety of the fishing vessel.

In recent years, many new technologies have been applied in VMS, including big data trajectory computing and visualization [1–3], VMS data mining [4–6], machine learning [7–9], and maritime IoT [10–13]. However, there are the following shortcomings in the anomaly detection of marine fishing vessels through a monitoring center using transmitted trajectory data:

- (1) Traditional VMSs use a centralized cloud computing model, and the storage, processing, and visualization of all fishing vessel data are completed in the monitoring center. The interaction between the VMS monitoring center and the fishing vessels is carried out through marine communications, but the bandwidth of marine communications is far behind that of land communication, which affects the real-time performance and accuracy of anomaly detection.
- (2) Traditional vessel anomaly detection methods extract frequent patterns in trajectories through spatial

location, sequence features, and abnormal behavior, and they detect abnormalities based on these frequent patterns. They require a large amount of data from all vessels. When the historical trajectory data set is small or the trajectory data in some areas are sparse, it is difficult to meet the requirements of detection accuracy, so it is not suitable for edge computing models.

To solve the above-stated problems, in this paper we propose a real-time anomaly detection model based on an edge computing framework. First, fishing vessel anomalies are detected according to the fishing vessel's behavior characteristics. Anomalies are mainly divided into spatial position anomalies and behavioral pattern anomalies. Secondly, the frequent patterns of fishing vessels are extracted through multifeature clustering, and a trajectory that deviates from a frequent cluster is recognized as an anomaly. When ignoring the time-varying evolution characteristics of trajectory flow data, a global feature model constructed using a sufficient number of historical trajectories has high anomaly detection accuracy. However, when considering the normal trajectory pattern based on the historical trajectory training set as the reference standard, new abnormal patterns and abnormal behaviors in specific scenes that may exist in the latest trajectory flow data cannot be detected accurately. Therefore, in this paper we use the communication function between the edge nodes to obtain the trajectory data of other nodes in the adjacent sea area in real-time, and then the spatiotemporal nearest neighbor similarities are combined to detect an online anomaly of the collected trajectory data set. Finally, the weighted detection results are calculated to determine whether an anomaly is present. That is, a high-precision and low-delay anomaly detection model is established in the edge layer through the cooperation between edge nodes, combined with an online anomaly detection algorithm and a historical trajectory extraction algorithm. Finally, the feasibility of the proposed model is verified experimentally.

Through this paper, we aim to make the following contributions to the state of the art:

- (1) To design a novel real-time anomaly detection model (RADM) for fishing vessels based on an edge computing framework to improve the real-time performance and accuracy of traditional VMSs.
- (2) To propose a fishing vessel trajectory anomaly detection algorithm based on multifeature clustering (VAD-MFC), which can not only combine the historical global trajectory feature model but also update the model incrementally at the edge layer.
- (3) To present a vessel anomaly detection algorithm based on spatiotemporal neighbor similarity (VAD-SNS), which can not only identify existing patterns but also detect new patterns and anomalous behavior in a specific scene. This model is suitable for running at edge nodes.

2. Related Work

Maritime anomaly detection has received much attention recently, based on the premise that if a trajectory does not appear frequently or in any cluster, then it may be abnormal.

The factors that affect the trajectory anomaly are not only reflected in the abnormal location but also hidden in the sequence of movements [14, 15]. In this context, we consider the behavior of a fishing boat to be abnormal if it is different from that of most fishing boats in the same sea area.

Generally, there are three types of anomalies, spatial position anomalies, sequence anomalies, and behavior anomalies [15]. Spatial position anomalies refer to anomalies based on vessel population density, that is, when the ship is located in a low-density area, this is considered to be an anomaly. The causes of an abnormal position may be that the vessel has entered a forbidden area or it is not moving in a specified area. A sequence anomaly refers to an exception based on a sequential pattern. Abnormal behavior means that the ship's behavior pattern is different from its nearest neighbor's trajectory with respect to some characteristic such as direction, speed, and so on. There are many reasons for abnormal behaviors; for example, some ships may not slow down after entering the port to improve efficiency.

At present, there is much research on trajectory anomaly detection [16]. According to the detection mechanism used, methods can be divided into four categories as follows: classification-based, distance-based, historical trajectory similarity-based, and grid-based.

Trajectory anomaly detection methods based on classification are mainly divided into two stages: the training stage and the detection stage. Li et al. [17] proposed an anomaly detection algorithm based on the correlation between anomaly patterns and spatiotemporal attributes. The algorithm used *k*-means to cluster the motifs of subtrajectory-related attributes and construct a rule-based classifier. Trajectory anomaly detection methods based on classification can achieve high accuracy when provided with an accurate training set [18]. Besides, many abnormal behaviors are unknown and change with time, so studies about online anomalous trajectory detection have been proposed [19, 20]. However, it is expensive to obtain accurate labelled data for practical applications.

Distance-based anomaly detection methods aim to detect the deviation of a trajectory in the dataset through a distance model. Knorr et al. [21] first proposed algorithms to identify abnormal trajectories through global characteristics such as trajectory, velocity, direction, and distance. However, this method is used to detect anomalies of complete trajectories, so it is only suitable for detecting abnormal behaviors whose location and behavior characteristics are completely different from those of other trajectories. Lee et al. [22] proposed a two-stage anomaly detection algorithm based on a segmentation detection framework. Yu et al. [23] proposed a strategy based on spatiotemporal nearest neighbor similarity. Wang et al. [24] presented the difference-and-intersection set distance metric to evaluate the similarity between any two trajectories. These anomaly detection methods suffer from high complexity and poor detection accuracy. Distance-based anomaly detection methods usually only focus on outlier behaviors with respect to location but ignore outlier trajectories whose behavior characteristics deviate from the nearest neighbors in space

and time. Therefore, these methods are difficult to use in real-time trajectory anomaly detection.

Anomaly detection methods based on historical trajectory similarity formulate a global feature model by extracting frequent patterns from large-scale historical trajectory data and then identifying trajectory deviations from the model as anomalies. Liu et al. [25] proposed an anomaly detection algorithm for identifying spatiotemporal outliers, which could discover the causal relationship between outliers in time. Rong et al. [26] proposed a data mining approach for the probabilistic characterization of maritime traffic and anomaly detection. Belhadi et al. [27] proposed a deep learning algorithm that learns the different features of historical data to determine groups of trajectory outliers. Lei [15] constructed an anomaly detection framework based on the spatial and behavioral characteristics of trajectory data. In this class of methods, the accuracy of anomaly detection is affected by the number of historical tracks. When the historical data set is small or track data in some areas are sparse, the detection effect's accuracy is insufficient for practical applications.

Trajectory anomaly detection methods based on grid partition divide the specified area into grid cells of equal size to detect abnormal behavior and are widely used in the field of transportation. Pang et al. [28] proposed a pattern recognition method based on likelihood ratio test statistics, which can detect "persistent" and "new" outliers in trajectory data effectively. Uniform grids have been used to represent the trajectory space and detect abnormal behaviors of taxis in real-time [29, 30]. The detection accuracy of grid-based anomaly detection methods is poor for trajectories with nonnormal distribution. Besides, these algorithms are only suitable for anomaly detection using geospatial information.

However, the above methods are based on traditional central computing architectures, and they are not optimized for edge nodes, so they are difficult to apply directly in edge computing scenarios. Moreover, in the field of marine fishing vessel management, only a few researchers have studied edge computing models of VMSs and established edge computing frameworks [31, 32], but there is no research on trajectory anomaly detection on this basis. In this paper, we present a real-time anomaly detection model for fishing vessels based on edge computing, and our models are experimentally verified to be more effective than traditional methods.

3. Study Area

The scope of this study was the sea area between 120 and 130 east longitude and 25–35 north latitude, located near Zhoushan, Zhejiang Province, China. The trajectory data set of the fishing VMS used in the experiment is shown in Figure 1. The system used BeiDou Satellite Navigation and communication technology to realize real-time monitoring of fishing boats' status. Its functions include ship query, voyage statistics, trajectory review, alarm and rescue, and so on. From May 22, 2016, to November 8, 2018, 220 ships were tracked and recorded.

FIGURE 1: The study area.

Since the anomaly state of a fishing vessel is closely related to the sequence of its trajectory, in this study we focus only on the detection of adjacent trajectory segments at a certain moment when an anomalous state of the fishing vessel trajectory was detected. Given a fishing vessel trajectory segment $ST_o(m, n)$, denoted as $\{p_m, p_{m+1}, \dots, p_{n-1}, p_n\}$, we define the spatial and behavioral anomalies of a trajectory.

A *spatial* anomaly considers the correlation between the adjacent trajectory segments. The distance between the current trajectory segment and other segments is calculated using a trajectory distance metric, and then the current trajectory segment's proportion of the adjacent trajectory segments is calculated using the adjacent trajectory distance threshold. If the proportion is smaller than the anomaly threshold, then the current trajectory segment can be identified as a spatial anomaly.

A *behavioral* anomaly of a fishing vessel is an anomaly in the vessel's multidimensional trajectory features, such as instantaneous angular acceleration, average angular acceleration, instantaneous speed, average speed, and acceleration. Each component receives a trajectory anomaly score by calculating the proportion of the abnormal trajectory point to the total trajectory points of the trajectory segment. Then, the component anomaly scores are integrated to obtain the final behavior anomaly score. If the score exceeds a predefined anomaly threshold, then the current trajectory is identified as a behavioral anomaly.

The abovementioned trajectory anomaly definitions are mainly aimed determining abnormalities in behavior of maritime fishing vessels. The sailing behavior of fishing vessels at sea is very different from that of general freight and passenger vessels. The latter determine the launch and destination port information before departure. The route of a voyage is sailed in accordance with established trade routes. Based on this premise, although at sea there is no physical restriction as in the case of urban road network traffic, a normal vessel does not deviate too far from these established routes. When crews find that the vessel has deviated from the route, they correct the ship's course

accordingly. Therefore, because they need to follow the determined route network, freight and passenger vessels are much like land vehicles when sailing at sea. Anomaly detection for such vessels can rely on the route information, and both the anomaly definition and anomaly detection are relatively simple. Unlike other vessels, fishing vessels do not have a defined destination area when at sea, nor do they follow a specific route, so they exhibit considerable randomness in their navigation. The sailing behavior of fishing vessels is mainly divided into two types, namely, sailing and fishing. Therefore, in this paper we make special distinctions for behavioral anomalies, as defined above. Thus, the aspects of position and behavior are considered to obtain the degree of anomaly suspicion of the fishing vessel's current state, while fishing vessel behavioral anomaly detection specifically is further classified according to whether it pertains to fishing or sailing.

4. Methods

4.1. Framework of Real-Time Trajectory Anomaly Detection Using Edge Computing. The three-module RADM framework is shown in Figure 2 and consists of historical trajectory modeling, anomaly evaluation, and the edge computing application.

Historical trajectory modeling extracts the characteristic of differences between normal and abnormal behavior and then clusters historical trajectories to find frequent patterns, which are considered normal behavior. The anomaly evaluation module firstly identifies trajectories that are different from the frequent patterns as anomalies, while the real-time trajectory is clustered incrementally to optimize the global feature model. Meanwhile, input trajectory data from nearby fishing vessels are used to detect new anomaly patterns and anomalous behavior in a specific scene. Finally, the RADM uses a combinational algorithm to calculate a comprehensive anomaly index. The edge computing module receives the output of RADM and applies it to various applications such as safety protection and navigation of the fishing vessel.

4.2. Fishing Vessel Trajectory Anomaly Detection Based on Multifeature Clustering. First, a global feature model is formulated based on the historical trajectory data set. Supervised and unsupervised learning algorithms are two commonly used modeling methods. The trajectories of the fishing vessels studied in this paper include two parts: (1) unlabeled data with a small number of abnormal trajectories and many normal trajectories and (2) a small number of labelled abnormal trajectories. Thus, it is more suitable to use an unsupervised learning algorithm. This study is based on an edge computing framework, that is, a framework where real-time anomaly detection of fishing vessels is needed in the edge layer [33, 34]. With the continuous inflow of trajectory data, the same trajectory may be indifferent to frequent patterns at different times because the global feature model changes in real-time. Therefore, it is necessary to update the model incrementally while identifying anomalies.

To solve the above problems, in this paper we propose a vessel anomaly detection algorithm based on multifeature

clustering (VAD-MFC). This algorithm extracts the characteristic quantity of obvious differences between normal and abnormal behavior and clusters historical trajectories according to these characteristic quantities to find the trajectory characteristic of normal behavior. Then, the minimum distance between the trajectory collected by the edge node in the current period and the normal behavior pattern is calculated, and this is considered the abnormality value of the current trajectory. Finally, the real-time trajectory is clustered incrementally to find a new trajectory characteristic and optimize the global feature model. It should be noted that the VAD-MFC algorithm is based on two premises: (1) the number of normal trajectories is much larger than that of abnormal trajectories and (2) there are obvious differences between the normal and abnormal behaviors of fishing vessels.

4.2.1. Frequent Area Formulation. Traditional grid-based clustering methods divide the monitored sea area into grids of equal size and then calculate the trajectory point density of each grid [35, 36]. If it is greater than a predefined threshold, then the area where the grid is located is marked as a frequent area. However, the above methods determine grid density by calculating the number of trajectory points in the grid, and thus the results of frequent regions may be lost when the object is moving too fast or the sampling interval is too large. This is demonstrated in Figure 3 [37, 38]. To avoid this kind of missing data problem, in this paper grid density is defined as the number of trajectory segments passing through using the method proposed in a prior study [15]. Each trajectory segment represents the line segment formed by adjacent trajectory points $TS_o(t) = \{p_{t-1}, p_t\}$.

4.2.2. Subtrajectory Division. The suspicion level of a fishing vessel's trajectory deviation is measured through the difference between the trajectory and the historical routes. Therefore, before anomaly detection, historical trajectories need to be clustered to obtain the historical trajectory mode. However, since the trajectory obtained in this paper is a complete trajectory with multiple stages and no markers, it is necessary to divide it to subtrajectories before modeling and ensure that only one behavior is assigned to each subtrajectory. If the historical trajectory TR contains the navigation stage subtrajectories TR_{steam} , the fishing stage subtrajectories TR_{fish} , and the anchoring stage subtrajectories TR_{stop} , then the results of the sun trajectory division can be expressed as follows:

$$TR = \{TR_1, TR_2, \dots, TR_n\}, \quad (1)$$

$$TR_i \in TR_{steam} \text{ or } TR_i \in TR_{stop} \text{ or } TR_i \in TR_{fish}. \quad (2)$$

In this study, the spatiotemporal distance model of trajectory points and the density-based spatial clustering of applications with noise (DBSCAN) clustering algorithm are used to classify the trajectory of fishing vessels. The basis for the classification is that the behavior of fishing vessels is continuous in time and localized in space. The algorithm first defines a spatiotemporal distance model based on the characteristics of fishing vessel behavior, and then DBSCAN clustering is applied

FIGURE 2: RADM framework.

FIGURE 3: Frequent regions are defined based on the number of trajectory segments passing through the grid.

on the trajectory points according to the model to form subtrajectory segments. For trajectory points P_i and P_j , the spatiotemporal distances are defined as follows:

(i) *Temporal Distance*. The distance of the trajectory points in time reflects their temporal correlation. Because fishing vessel behavior is continuous in time, the smaller the interval, the greater the probability of finding the same behavior. The distance definitions of PT_a and PT_b in time are given in the following equation:

$$T(a, b) = \frac{|\text{time}_a - \text{time}_b|}{\max(T) - \min(T)}. \quad (3)$$

(ii) *Spatial Distance*. The geographic distance of the trajectory points reflects their spatial correlation. Because the behavior of fishing vessels is spatially localized, that is, the

trajectory falls in the same area, then the probability of the trajectory being the outcome of the same behavior is greater. The spatial distance definitions of PT_a and PT_b are given in the following equation:

$$S(a, b) = \frac{\sqrt{(\text{lon}_a - \text{lon}_b)^2 + (\text{lat}_a - \text{lat}_b)^2}}{\max(S) - \min(S)}. \quad (4)$$

(iii) *Speed Distance*. Because the speed of fishing boats is different during navigation, fishing, and anchoring, the larger the speed difference is, the smaller the probability the corresponding behaviors being the same. The definition of the distance between PT_a and PT_b on speed is given in the following equation:

$$V(a, b) = \frac{|\text{speed}_a - \text{speed}_b|}{\max(V) - \min(V)}. \quad (5)$$

(iv) *Direction Distance*. Due to the continuity of behavior in time, the distance in the direction $\text{Dir}(a, b)$ is defined as the number of times the direction changes in a given interval.

When dividing the trajectories of fishing vessels, the four distances defined above should be considered. The distance between trajectories PT_a and PT_b is defined as

$$D(a, b) = W_T * T(a, b) + W_S * S(a, b) + W_V * V(a, b) + W_D * \text{DIR}(a, b). \quad (6)$$

4.2.3. Trajectory Spatiotemporal Distance Model. The abnormal behavior of fishing boats is continuous in time, which is represented by several abnormal trajectory points in the trajectory data. Therefore, we should not only consider the attributes of the trajectory points themselves but also pay attention to the differences between adjacent trajectory points. This paper describes how to extract the multiple eigenvalues of a trajectory; on this basis, a spatiotemporal distance model between trajectories is defined.

When a fishing vessel's behavior is abnormal, its speed or direction can change dramatically. The direction of the fishing vessel is also unstable during the fishing phase. Therefore, to reduce the impact of the fishing stage on anomaly detection, in this study we use the average acceleration, the average angular acceleration, and the average velocity as three characteristics of behavior. In spatial locations, fishing vessels may appear in illegal locations due to acts such as invading restricted zones, and thus the frequency is selected in this paper. To sum up, the anomaly index of the trajectory $\text{TR} = \{P_1, P_2, \dots, P_n\}$ is defined in four ways as follows:

(i) *Frequency*. The frequency of a trajectory indicates the density of the area through which the current trajectory passes. A lower frequency indicates an inactive fishing area, and the probability of an anomaly is greater; conversely, it is assumed that the probability of an anomaly is relatively small. The specific definition is given in equation (7). Here, $\text{FR}(P_i)$ indicates whether trajectory P_i is in a frequent area, with 1 corresponding to frequent areas, and 0 indicating that it is away from a frequent area:

$$\text{FR} = \frac{\sum_{i=1}^n \text{FR}(P_i)}{n}. \quad (7)$$

(ii) *Velocity*. Velocity can be used to distinguish the behavior of fishing vessels, and it plays an important role in clustering normal behavior. Because the instantaneous velocity is random, it cannot reflect the general state of the current trajectory well, and thus, the average velocity is selected as the eigenvalue in this paper, as follows:

$$\text{AS} = \frac{\sum_{i=2}^n \sqrt{(\text{lon}_i - \text{lon}_{i-1})^2 + (\text{lat}_i - \text{lat}_{i-1})^2}}{\text{time}_n - \text{time}_1}. \quad (8)$$

(iii) *Acceleration*. Acceleration reflects the stability of the fishing vessel. The greater the acceleration, the more unstable the behavior of the fishing vessel is and the greater the probability of abnormality. In this paper, the average acceleration is used as the fishing vessel characteristic quantity, as follows:

$$\text{ADS} = \frac{1}{n-1} \sum_{i=2}^n \frac{|\text{speed}_i - \text{speed}_{i-1}|}{\text{time}_i - \text{time}_{i-1}}. \quad (9)$$

(iv) *Angular Acceleration*. Angular acceleration reflects the change in direction. Similarly to the above, the mean angular acceleration is used as the characteristic quantity of fishing vessels, as given in the following equations:

$$\text{ADA} = \frac{1}{n-1} \sum_{i=2}^n \frac{\text{dirDif}}{\text{time}_i - \text{time}_{i-1}}, \quad (10)$$

$$\text{dirDif} = \begin{cases} |\text{dir}_i - \text{dir}_{i-1}| & |\text{dir}_i - \text{dir}_{i-1}| \leq 180, \\ 360 - |\text{dir}_i - \text{dir}_{i-1}| & |\text{dir}_i - \text{dir}_{i-1}| > 180. \end{cases} \quad (11)$$

In this paper, the anomaly index of a fishing vessel is evaluated by calculating the difference between the current trajectory and the normal trajectory mode. Therefore, a spatiotemporal distance model between trajectories needs to be defined. Because the different data magnitudes of different eigenvalues can affect the calculation of distances, the data needs to first be normalized, as shown in equation (12). The spatiotemporal distances of the trajectories TR_i and TR_j are then calculated, as shown in equations (13) and (14):

$$C_i = \frac{C_i - \min(C)}{\max(C) - \min(C)}, \quad (12)$$

$$D(\text{TR}_i, \text{TR}_j) = \sqrt{\text{Dif}(\text{FR})^2 + \text{Dif}(\text{AS})^2 + \text{Dif}(\text{ADS})^2 + \text{Dif}(\text{ADA})^2}, \quad (13)$$

$$\text{Dif}(C) = |C_i - C_j|. \quad (14)$$

4.2.4. Anomaly Detection Process. When moving to the edge layer for the real-time anomaly detection of fishing vessels, the trajectory data continue to flow in as a “trajectory flow.” To improve the performance of anomaly detection, timeseg is used to represent a basic time unit because of the insufficient information contained in the individual trajectory points. A timeseg contains more than one trajectory point. In this study, anomaly detection was performed on the incoming trajectory data in the current timeseg. Comparing the difference between TR and normal trajectory patterns collected in the current timeseg, the higher the difference, the greater the probability of abnormal fishing vessel behavior. Therefore, before anomaly detection, normal behavior patterns need to be obtained through historical trajectory modeling, which is the basis of trajectory anomaly detection.

(i) *Modeling Historic Trajectories.* Historic trajectories are clustered based on selected feature values during the modeling phase to discover frequent patterns hidden in the trajectory data set, as detailed in Algorithm 1.

The VAD-MFC historical trajectory modeling is divided into four main steps, as follows:

First, Determination of the Frequent Areas. This study considered the spatial location characteristics of trajectories, that is, whether or not they occurred in frequent areas. Therefore, in this step the algorithm constructs frequent regions mainly from historical trajectories, which facilitates the extraction of frequencies. *Second, Division into Subtrajectories.* This stage mainly groups the continuous trajectory points with the same behavior to ensure that each subtrajectory contains only one behavior to optimize the clustering effect. This stage mainly uses the method mentioned in Section 4.2, that is, clustering trajectory points based on a predefined spatiotemporal distance model. The clustering algorithm used here was DBSCAN. *Third, Extraction of Subtrajectory Features.* After the subtrajectories are divided, the spatial location (frequency) and behavior (average velocity, average acceleration, and average angular acceleration) characteristics of all the subtrajectories are extracted. *Finally, Building a Normal Trajectory Model.* In this stage, the subtrajectories are clustered based on the spatiotemporal distance model between trajectories, and the frequent pattern of trajectories, that is, the normal trajectory pattern, is extracted. Because DBSCAN can process noise points effectively and cluster quickly, it was again used to cluster historical trajectories in this stage.

(ii) *Trajectory Anomaly Detection.* The abnormal index A_c of the trajectory data TR collected in the current timeseg is calculated based on the prebuilt normal trajectory model M . If A_c is greater than the anomaly threshold δ , then the current behavior of the fishing vessel is considered to be abnormal. Detailed steps are given in Algorithm 2.

Trajectory anomaly detection using VAD-MFC is mainly divided into four main steps, as follows:

First, extract the characteristics of the current trajectory flow TR. Features extracted include the average velocity AS,

average acceleration ADS, average angular acceleration ADA, and frequency FR. Second, update the normal trajectory model dynamically. The incremental DBSCAN algorithm is used to incrementally cluster the trajectory stream data TR collected in the current timeseg, so as to update the global feature model dynamically and improve the accuracy of anomaly detection. Third, calculate the distance between TR and all the normal trajectory modes. The calculation methods are given in equations (13) and (14). Finally, obtain the minimum difference between TR and the normal trajectory mode, that is, the anomaly index A_c .

4.3. Vessel Trajectory Anomaly Detection Based on Spatio-temporal Neighbor Similarity. Anomaly detection methods based on multifeature clustering uses the normal historical trajectory patterns as the reference standard. They can detect existing patterns better, but it is difficult to detect new anomaly patterns and anomalous behavior in a specific scene. To solve these problems, we present VAD-SNS.

The VAD-SNS algorithm is based on the similarity of fishing vessel behaviors within a certain spatiotemporal range (for example, the return of a fishing vessel during a storm), and thus, if there is a significant difference between the vessel trajectory in an area and the vessel trajectory in the adjacent area, the behavior is considered abnormal. The following describes in detail how to detect anomalies in fishing vessel behavior online using spatiotemporal near-neighbor similarity.

This study builds on a framework for edge computing from previous research [32], in which fishing vessels collect positioning information in real-time through an on-board terminal device during navigation. For the positioning information TR collected in the current timeseg T_c , it can be formally expressed as a series of trajectory points in chronological order, that is, $TR = \{(p_1, t_1), (p_2, t_2), \dots, (p_n, t_n)\}$, where p_i is the fishing vessel at time stamp t_i 's position coordinates and $T_c.start < t_i < t_{i+1} < T_c.end$. At the same time, the fishing boat sends position information to nearby nodes through edge computing nodes so that each edge node can obtain a neighbor trajectory set $TF_c = \{TR_0, TR_1, \dots, TR_n\}$, where TR_i represents the position information collected within the current timeseg of node i .

Because only the position information of fishing vessels in adjacent sea areas can be received passively through automatic identification systems (AIS), there may be missing data in the TF_c trajectories. To improve the accuracy of trajectory anomaly detection, the abnormal behavior of fishing vessels can be evaluated by synthesizing the abnormality degree of all trajectory segments in TR. That is, the segment set of adjacent trajectories for each trajectory segment is obtained based on the distance model between trajectories, and then the abnormality degree of the current trajectory segment is evaluated based on the behavioral difference characteristics in the local neighborhood. Finally, the abnormality index of the current trajectory is obtained by combining the abnormality degree of all the trajectory segments.

Input:
historical trajectory set $D = \{P_1, P_2, \dots, P_n\}$
Output:
normal trajectory model N
Begin
 (1) discover frequent regions
 (2) get subtrajectories D_s by dividing the training set D
 (3) **for each** $TR_i \in D_s$ **do**
 (4) extract trajectory features F
 (5) build normal trajectory model N
 (6) **return** N
End

ALGORITHM 1: Historical trajectory modeling.

Input:
normal trajectory model N
trajectory $TR = \{P_1, P_2, \dots, P_n\}$
Output:
abnormal index A_c
Begin:
 (1) trajectory feature extraction
 (2) update normal trajectory model by Incremental DBSCAN
 (3) **for each** $N_i \in N$ **do**
 (4) calculate trajectory distance
 (5) get the minimum trajectory difference A_c
 (6) **return** A_c
End

ALGORITHM 2: Trajectory anomaly detection.

Input:
trajectory $TR = \{tf_0, tf_1, \dots, tf_n\}$
Output:
abnormal index A_N
Begin
 (1) **for each** $tf_i \in TR$ **do**
 (2) Obtain a time-space neighbor trajectory segment set for tf_i
 (3) Calculate the local difference of tf_i
 (4) Calculate tf_i 's local anomaly factor
 (5) Calculate the local anomaly factors the for current timebin trajectory flow TR
 (6) Obtain the anomaly index A_N using time decay function
 (7) **return** A_N
End

ALGORITHM 3: VAD-SNS.

4.3.1. Neighbor Trajectory Segmentation. The VAD-SNS algorithm requires the spatiotemporal neighbor trajectory segment set first. The spatiotemporal local neighbors include temporally local neighbors and spatially local neighbors because each trajectory point in TF_c is collected for the current timeseg and all the trajectory segments belong to

temporal neighbors. For any trajectory segment tf_i , the spatial neighbor N_{T_c} representation of the segmented trajectory set with tf_i spacing is no greater than a given threshold d , as

$$N_{T_c} = \{tf_i \in TF_c | \text{dis}(tf_i, tf_j) \leq d\}. \quad (15)$$

FIGURE 4: Sliding window model based on time decay.

Here, $\text{dis}(\text{tf}_i, \text{tf}_j)$ represents the spatial distance of the trajectory segment between tf_i and tf_j . In this paper, the Hausdroff distance [39] is used, as shown in the following equations:

$$H(A, B) = \max(h(A, B), h(B, A)), \quad (16)$$

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|. \quad (17)$$

4.3.2. Local Anomaly Evaluation. The local outlier factor (LOF) is a density-based anomaly detection mechanism [40]. It is widely used to detect local neighborhood anomaly data because it considers the density of data samples relative to their spatial neighbors and does not require prior knowledge of the data distribution. Based on the LOF, in this method we use the local difference density (LDD) to evaluate the degree of behavioral difference between current trajectory segment tf_i and its neighbor trajectory segment set N_{T_c} , as

$$\text{LDD}_{T_c}(\text{tf}_i) = \frac{\sum_{\text{tf}_j \in N_{T_c}(\text{tf}_i)} \text{Diff}(\text{tf}_i, \text{tf}_j)}{|N_{T_c}(\text{tf}_i)|}. \quad (18)$$

Here, $|N_{T_c}(\text{tf}_i)|$ denotes the number of trajectory segments contained in $N_{T_c}(\text{tf}_i)$ and $\text{Diff}(\text{tf}_i, \text{tf}_j)$ denotes the degree of behavioral difference between trajectory segments and tf_j . In this paper, we mainly consider the characteristics of direction and speed, as

$$\text{Diff}(\text{tf}_i, \text{tf}_j) = W_S * \text{dis}(\text{Speed}) + W_D * \text{dis}(\text{Dir}). \quad (19)$$

When a fishing vessel is sailing, it is often found that there is a significant difference in the behavior of all the fishing vessels in the adjacent sea area. At this time, it will be considered that tf_i contains an exception that does not correspond to the facts. It can only be considered an anomaly if the behavior of other fishing vessels in the neighborhood is similar and there is a significant difference between the current fishing vessels. Therefore, based on the LOF, the local anomaly factor (LAF) is used to evaluate the degree of abnormality of each trajectory segment in its neighborhood [41]. Larger LAF values correspond to a greater the likelihood of abnormality of the trajectory segment, as

$$\text{LAF}_{T_c}(\text{tf}_i) = \frac{\sum_{\text{tf}_j \in N_{T_c}(\text{tf}_i)} \text{LDD}_{T_c}(\text{tf}_i) / \text{LDD}_{T_c}(\text{tf}_j)}{|N_{T_c}(\text{tf}_i)|}. \quad (20)$$

The LAF describes the degree of abnormality of trajectory segments within their local neighborhoods, while the TR collected in the current timeseg contains multiple trajectory segments. Thus, the abnormality of trajectory TR can be defined as

$$\text{LAF}_{T_c}(\text{TR}) = \frac{\sum_{\text{tf}_i \in \text{TR}} \text{LAF}_{T_c}(\text{tf}_i)}{|\text{TR}|}. \quad (21)$$

Given the time-varying evolution of the trajectory flow data, the same behavioral characteristics of fishing vessels may be abnormal within the current timeseg and but become normal in the next timeseg. Therefore, we should not only focus on the LAFs of the current timeseg but also consider the influence of the historical LAFs of the fishing vessel. Sliding window models are widely used in real-time data streaming because they can process persistent streaming data effectively while eliminating obsolete and invalid data. In this paper, we use a sliding window model based on time decay [42] to ensure data freshness.

The sliding window model based on time decay is shown in Figure 4. The size of the sliding window is N . Each time a new *timeseg* datum is processed, the sliding window moves one timeseg to the right, thereby removing the LAF of the farthest *timeseg*. A time decay function reduces the influence of historical LAFs. For all the fishing vessels, the historical LAF obtained for the previous timeseg is multiplied by a forward decay time function, and then the LAFs produced by the current timeseg are accumulated to obtain the online anomaly index A_N of the fishing vessels. The value of the forward decay time function is the distance between the current timeseg and the leftmost timeseg of the sliding window, as defined in the following equation:

$$\begin{cases} A_N = \sum_{T_k=T_l}^{T_c} \frac{g(T_k - T_s) \text{LAF}_{T_k}(\text{TR})}{g(T_c - T_s)}, \\ g(n) = n^2. \end{cases} \quad (22)$$

Here, T_c is the current timebin, T_l is the leftmost timebin of the sliding window, and $g(n)$ is a monotonic nondecaying time function.

The VAD-SNS algorithm is given in Algorithm 3.

4.4. Real-Time Anomaly Detection. RADM is set at the edge layer to improve the real-time performance and accuracy of anomaly detection, as shown in Figure 4. It can detect the current timeseg trajectory data in real-time. If abnormal behavior is detected, then it will alert and send the results to the monitoring center.

RADM uses a combination of an online anomaly detection algorithm and a historical trajectory extraction algorithm. First, normal patterns hidden in the historical trajectory set are extracted based on multifeature clustering, and the distance from normal patterns is taken as the historical anomaly index A_C . Then, combined with the dynamic information of moving neighborhood nodes, the online anomaly index A_N is obtained based on spatiotemporal

FIGURE 5: Frequent areas.

FIGURE 6: The interpolated trajectory.

TABLE 1: AIS device dynamic information update time [43].

Update Time	Navigation Status
3 minutes	Anchoring
10 seconds	Speed between 0 and 14 knots, sailing
3.33 seconds	Speed between 0 and 14 knots, changing course
6 seconds	Speed between 14 and 23 knots, sailing
2 seconds	Speed between 14 and 23 knots, changing course
2 seconds	Speed >23 knots, sailing
2 seconds	Speed >23 knots, changing course

neighborhood similarity. Finally, a comprehensive anomaly index A is obtained by combining the weighted two anomaly detection algorithms:

$$A = W_N * A_N + W_C * A_C. \quad (23)$$

FIGURE 7: Feature clustering results.

Here, W_N represents the weight of the online anomaly index and W_C represents the weight of the historical anomaly index. Because the moving edge node mainly receives ship data from nearby sea areas through AIS, when a fishing vessel travels to sparse areas, data from too few trajectories may be received, which will affect the accuracy of the A_N index. At this time, the W_N value can be reduced appropriately, mainly through historical trajectory extraction for trajectory anomaly detection. Conversely, the W_N value can be increased.

5. Results and Discussion

5.1. Experimental Setup. The study sea area was divided into $0.01^\circ(\text{latitude}) \times 0.01^\circ(\text{longitude})$ grids for a total of 1000×1000 grid squares. We installed edge computing nodes on vessels using the Zhoushan fishing VMS [32]. Using fishing vessel trajectory data collected using the Zhoushan fishing VMS from May 2016 to November 2018, the frequent regions were constructed. Grids with trajectory densities greater than 100 were marked as frequent regions, and the sampling interval between the trajectory points was 30 seconds. The results are shown in Figure 5, where black represents frequent areas and white represents infrequent areas. The experiment runs in a Windows 10 system with an Intel Core i7 4.8 GHz processor and 32 GB of RAM.

To verify the effectiveness of the RADM, we conducted comparative experiments using VAD-MFC, VAD-MFC, and RADM on a selected data set consisting of 50 normal trajectories and eight abnormal trajectories. We used the anomaly index mentioned above and the commonly used receiver operating characteristic (ROC) curve as the evaluation standard of anomaly detection.

To simulate the edge computing environment better, we interpolated the real trajectory data. The sample interval after interpolation was 2 seconds, and the following six possible anomalies were randomly added: steady direction and large fluctuations in speed, stable speed and large fluctuations in direction, large fluctuations in speed and direction, speed greater than 15 knots, avoidance of frequent

FIGURE 8: Frequencies of normal and abnormal trajectories.

FIGURE 9: Results of anomaly detection with VAD-MFC.

FIGURE 10: The results of abnormal detection with the RADM.

areas, and abnormal behavior in a specific environment (e.g., fast transit through a port area).

The interpolated data are shown in Figure 6.

In a real environment, the number of normal trajectories is much larger than the number of abnormal trajectories. Therefore, based on the dataset's statistics, the ratio of normal and abnormal trajectories in the experimental setup was 42195:1005. Edge nodes broadcasted dynamic data according to the update time (as shown in Table 1) and the communication distance of the AIS device. To simplify the calculation model, the communication distance of the experiment was set to 5 nautical miles; that is, each edge node

could receive AIS information from anywhere within the surrounding 5 nautical miles. Finally, anomaly detection was performed on the received trajectory stream data.

5.2. Experimental Results. To ensure that each subtrajectory contains only independent behavior characteristics when building a normal trajectory model, VAD-MFC first partitions the trajectories according to the spatiotemporal distance model and then extracts the behavior characteristics of the subtrajectories and clusters them. Because of the large time span of the subtrajectories, the latter were divided into

FIGURE 11: ROC curves of VAD-MFC and VAD-SNS.

FIGURE 12: RADM anomaly detection under different weight indexes.

FIGURE 13: The anomaly index distribution of RADM.

FIGURE 14: The ROC curve of the RADM.

five-minute interval trajectory segments to produce a better clustering effect. The trajectory segments were then clustered. The clustering result is shown in Figure 7, where the x -axis represents the change of speed, the y -axis represents the change of direction, and the z -axis represents the average speed. As shown in the graph, the trajectory segments were grouped into three categories corresponding to the three normal behaviors of fishing vessels (green for navigation, orange for mooring, and blue for fishing). The clustered centroids of each class represent the frequent pattern of characteristics of the current behavior.

Figure 8 compares the frequencies of the normal and abnormal trajectories. Both normal and abnormal trajectories were mostly located in frequent areas, but there were also low frequencies (such as the No. 6 abnormal trajectory). Because the number of abnormal trajectories marked in this experiment was small and the abnormal behavior features were sparse, fishing vessels away from frequent areas may have demonstrated abnormal behavior. Frequency was used as a spatial location feature of the trajectory in the following experiments; that is, fishing boats that were far away from the frequent areas may have been identified as anomalies.

Figures 9 and 10 show the abnormal detection results of VAD-MFC and RADM, respectively. The abscissa represents different trajectory sequences, and the ordinate represents the anomaly index of the current trajectory. As shown, VAD-MFC could not detect individual abnormal trajectories, and the detection effect of RADM was better than VAD-MFC.

To evaluate the effectiveness of RADM in the edge computing framework [32], in this section we compare the

anomaly detection accuracy of VAD-MFC, VAD-SNS, and RADM. Compared with other anomaly detection indicators, the ROC curve has greater tolerance to the imbalance of positive and negative samples, and thus we use the ROC curve as the evaluation standard for anomaly detection [44].

Figure 11 shows the ROC curves of VAD-MFC and VAD-SNS. Both VAD-MFC and VAD-SNS had a certain accuracy of anomaly detection, in which the area-under-curve (AUC) value of VAD-MFC reached 0.79, and the AUC value of VAD-SNS reached 0.8, which is much higher than the 0.5 obtained through random classification. When the false positive rate of VAD-MFC was 0, the true positive rate was close to 0.6; thus, VAD-MFC was better than VAD-SNS in the scenario of normal operation without false alarms. The abscissa of the equal error rate point in the VAD-SNS algorithm was smaller than that in VAD-MFC, and thus VAD-SNS was better than VAD-MFC when the probability of detecting normal and abnormal samples was equal.

The RADM combines the advantages of VAD-SNS and VAD-MFC. The results were normalized and weighted, and Figure 12 shows the accuracy of the RADM anomaly detection under different weight indexes, where the abscissa represents the weight of VAD-SNS and the ordinate represents the AUC value of RADM under the current weight. The AUC value of RADM increased with the increase of the VAD-SNS algorithm weight and then gradually decreased. The AUC value reached its maximum when the weight index was 0.29.

Figures 13 and 14 show the RADM anomaly detection results with a weighted index of 0.29. Figure 13 shows the distribution of the anomaly index, in which the red "X" represents the anomaly index of the abnormal trajectory, and the blue dot represents the anomaly index of the normal trajectory. Although the anomaly index of the normal trajectory was occasionally larger, the anomaly index of the abnormal trajectory was generally higher than that of the normal one, and the overall detection effect was better.

Figure 14 shows a comparison of the ROC curves of VAD-MFC, VAD-SNS, and RADM. The blue curve shows RADM with a weighted index of 0.29, and its AUC value was 0.92, which was much higher than the 0.79 of VAD-MFC and the 0.8 of VAD-SNS. Moreover, the abscissa of the equal error rate point of the model was also smaller than that of VAD-MFC and VAD-SNS. Therefore, RADM had higher anomaly detection accuracy.

6. Conclusions

In this study, we designed and implemented a real-time anomaly detection model for fishing vessel trajectories based on an edge computing paradigm. First, a trajectory anomaly detection algorithm based on multifeature clustering was designed according to the behavioral characteristics of fishing vessels. Then, aiming at the problem that the algorithm could not effectively detect new abnormal patterns that may have existed in the trajectory flow data and its weak real-time performance, the RADM was proposed by introducing an edge computing framework and combining it with the spatiotemporal near-neighbor

similarity. Finally, experiments were carried out on real data sets. The experimental results show that RADM had high detection accuracy in detecting anomalies of continuous trajectory flow. RADM can provide a trajectory anomaly detection service in the edge layer without interaction with a terrestrial monitoring center; thus, it can improve the real-time performance of trajectory anomaly detection without the limitations of marine communications.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (no. 61972358) and the Zhejiang Province Key Research and Development Project (no. 2017C03024).

References

- [1] Z. Xiao, X. Fu, L. Zhang et al., "Big data driven vessel trajectory and navigating state prediction with adaptive learning, motion modeling and particle filtering techniques," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, pp. 1–14, 2020.
- [2] Y. Huang, Y. Li, Z. Zhang, and R. W. Liu, "GPU-accelerated compression and visualization of large-scale vessel trajectories in maritime IoT industries," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 10794–10812, 2020.
- [3] M. I. Marzuki, P. Gaspar, R. Garelo, V. Kerbaol, and R. Fablet, "Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories," *IEEE Journal of Oceanic Engineering*, vol. 43, no. 3, pp. 689–699, 2018.
- [4] H. Huang, F. Hong, Z. Wu et al., "FVRD: fishing vessel relations discovery through VMS trace analysis," in *Proceedings of the 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, pp. 1–4, Kobe, Japan, May 2018.
- [5] Sunarmo, A. Affandi, and S. Sumpeno, "Clustering spatial temporal distribution of fishing vessel based LON VMS data using K-means," in *Proceedings of the 2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, pp. 1–6, Yogyakarta, Indonesia, November 2020.
- [6] F. Hong, Y. Zhen, Y. Zong et al., "Fishing area separation and relation discovery through VMS trace analysis," in *Proceedings of the OCEANS 2017*, pp. 1–6, Aberdeen, UK, June 2017.
- [7] M. Terayama, C. Nishizaki, and T. Okazaki, "Fundamental study of traffic prediction in Tokyo Bay using machine learning," in *Proceedings of the 2020 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pp. 536–541, Chiang Mai, Thailand, September 2020.
- [8] X. Liu, W. He, J. Xie, and X. Chu, "Predicting the trajectories of vessels using machine learning," in *Proceedings of the 2020 5th International Conference on Control, Robotics and Cybernetics (CRC)*, pp. 66–70, Wuhan, China, October 2020.
- [9] H. Susanto and G. Wibisono, "Marine vessel telemetry data processing using machine learning," in *Proceedings of the 2019 6th International Conference on Electrical Engineering*, pp. 128–135, Computer Science and Informatics (EECSI), Istanbul, Turkey, September 2019.
- [10] Z. He, F. Yang, Z. Li, K. Liu, and N. Xiong, "Mining channel water depth information from IoT-based big automated identification system data for safe waterway navigation," *IEEE Access*, vol. 6, pp. 75598–75608, 2018.
- [11] M. S. Meera and S. N. Rao, "Comparative analysis of iot protocols for a marine IoT system," in *Proceedings of the 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 2049–2053, Bangalore, India, September 2018.
- [12] R. W. Liu, J. Nie, S. Garg, Z. Xiong, Y. Zhang, and M. S. Hossain, "Data-driven trajectory quality improvement for promoting intelligent vessel traffic services in 6G-enabled maritime IoT systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5374–5385, 2021.
- [13] P. Lousã, F. Moreira, and L. M. Silveira, "Micro-VMS: a VMS mobile unit for artisanal fishing vessels," in *Proceedings of the 2018 16th International Conference on Intelligent Transportation Systems Telecommunications (ITST)*, pp. 1–7, Xi'an, China, October 2018.
- [14] C. Iphar, C. Ray, and A. Napoli, "Data integrity assessment for maritime anomaly detection," *Expert Systems with Applications*, vol. 147, Article ID 113219, 2020.
- [15] P.-R. Lei, "A framework for anomaly detection in maritime trajectory behavior," *Knowledge and Information Systems*, vol. 47, no. 1, pp. 189–214, 2016.
- [16] M. Gupta, J. Gao, C. C. Aggarwal, and J. Han, "Outlier detection for temporal data: a survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 9, pp. 2250–2267, 2014.
- [17] 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; Proceedings; Society for Industrial and Applied Mathematics*, pp. 273–284, Philadelphia, PA, USA, October 2007.
- [18] D. A. Herrero, D. S. Pedroche, J. G. Herrero, and J. M. M. López, "AIS trajectory classification based on IMM data," in *Proceedings of the 2019 22th International Conference on Information Fusion (FUSION)*, pp. 1–8, Ottawa, Canada, July 2019.
- [19] Y. Han, R. Tse, and M. Campbell, "Pedestrian motion model using non-parametric trajectory clustering and discrete transition points," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2614–2621, 2019.
- [20] 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, Dallas, TX, USA, April 2020.
- [21] 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.
- [22] J. 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, Cancun, Mexico, April 2008.
- [23] 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, Association for Computing Machinery, New York, NY, USA, August 2014.
- [24] 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.
 - [25] W. Liu, Y. Zheng, S. Chawla, J. Yuan, and X. Xing, “Discovering spatio-temporal causal interactions in traffic data streams,” in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1010–1018, Association for Computing Machinery, New York, NY, USA, August 2011.
 - [26] H. Rong, A. P. Teixeira, and C. Guedes Soares, “Data mining approach to shipping route characterization and anomaly detection based on AIS data,” *Ocean Engineering*, vol. 198, Article ID 106936, 2020.
 - [27] A. Belhadi, Y. Djenouri, D. Djenouri, T. Michalak, and J. C.-W. Lin, “Deep learning versus traditional solutions for group trajectory outliers,” *IEEE Transactions on Cybernetics*, vol. 26, pp. 1–12, 2020.
 - [28] L. X. Pang, S. Chawla, W. Liu, and Y. Zheng, “On detection of emerging anomalous traffic patterns using GPS data,” *Data & Knowledge Engineering*, vol. 87, pp. 357–373, 2013.
 - [29] 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.
 - [30] N. Taniarza and S. Akbar, “Anomalous trajectory detection from taxi GPS traces using combination of IBAT and DTW,” in *Proceedings of the 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI)*, pp. 1–5, Langkawi, Malaysia, November 2017.
 - [31] J. C. Ferreira and A. L. Martins, “Edge computing approach for vessel monitoring system,” *Energies*, vol. 12, no. 16, p. 3087, 2019.
 - [32] J. Huang, J. Wan, J. Yu, F. Zhu, and Y. Ren, “Edge computing-based adaptable trajectory transmission policy for vessels monitoring systems of marine fishery,” *IEEE Access*, vol. 8, pp. 50684–50695, 2020.
 - [33] H. Gao, W. Huang, and Y. Duan, “The cloud-edge-based dynamic reconfiguration to service workflow for mobile ecommerce environments,” *ACM Transactions on Internet Technology*, vol. 21, no. 1, pp. 1–23, 2021.
 - [34] Y. Yin, Z. Cao, Y. Xu, H. Gao, R. Li, and Z. Mai, “QoS prediction for service recommendation with features learning in mobile edge computing environment,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 4, pp. 1136–1145, 2020.
 - [35] P. Lei, T. Shen, W. Peng, and I. Su, “Exploring spatial-temporal trajectory model for location prediction,” in *Proceedings of the 2011 IEEE 12th International Conference on Mobile Data Management*, pp. 58–67, Lulea, Sweden, June 2011.
 - [36] A. Y. Xue, R. Zhang, Y. Zheng, X. Xie, J. Huang, and Z. Xu, “Destination prediction by sub-trajectory synthesis and privacy protection against such prediction,” in *Proceedings of the 2013 IEEE 29th International Conference on Data Engineering (ICDE)*, pp. 254–265, Brisbane, Australia, April 2013.
 - [37] H. Jeung, H. T. Shen, and X. Zhou, “Mining trajectory patterns using hidden markov models,” in *Proceedings of the Data Warehousing and Knowledge Discovery*, I. Y. Song, J. Eder, and T. M. Nguyen, Eds., pp. 470–480, Springer, Berlin, Heidelberg, 2007.
 - [38] Y. Yin, Q. Huang, H. Gao, and Y. Xu, “Personalized APIs recommendation with cognitive knowledge mining for industrial systems,” *IEEE Transactions on Industrial Informatics*, vol. 1, no. 1, 2020.
 - [39] A. V. Arutyunov, S. A. Vartapetov, and S. E. Zhukovskiy, “Some properties and applications of the hausdorff distance,” *Journal of Optimization Theory and Applications*, vol. 171, no. 2, pp. 527–535, 2016.
 - [40] B. Tu, C. Zhou, W. Kuang, L. Guo, and X. Ou, “Hyperspectral imagery noisy label detection by spectral angle local outlier factor,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 9, pp. 1417–1421, 2018.
 - [41] J. Mao, T. Wang, C. Jin, and A. Zhou, “Feature grouping-based outlier detection upon streaming trajectories,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 12, pp. 2696–2709, 2017.
 - [42] G. Cormode, V. Shkapenyuk, D. Srivastava, and B. Xu, “Forward decay: a practical time decay model for streaming systems,” in *Proceedings of the 2009 IEEE 25th International Conference on Data Engineering*, pp. 138–149, Shanghai, China, March 2009.
 - [43] U. Dahana and R. O. S. Gurning, “Maritime aerial surveillance: integration manual identification system to automatic identification system,” *IOP Conference Series: Earth and Environmental Science*, vol. 557, Article ID 12014, 2020.
 - [44] H. Gao, C. Liu, Y. Li, and X. Yang, “V2VR: reliable hybrid-network-oriented V2V data transmission and routing considering RSUS and connectivity probability,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, pp. 1–14, 2020.