ISSN 2616-5775 Vol. 9, Issue 8: 90-97, DOI: 10.25236/AJCIS.2026.090811

Simulation study on ship operation trajectory prediction based on Bayesian optimization

Li Weibo, Yang Dawei, Wang Zixuan

Naval Aviation University, Yantai, China
webblee2008@sina.com

Abstract: *Based on real-time position tracking recognition and trajectory prediction analysis of ships, dynamic control and route planning of maritime traffic are achieved, and a ship operation trajectory prediction method based on Bayesian optimization is proposed. A parameter model for ship trajectories is built based on original samples such as position, velocity, and heading. The point information of the ship's current position is used to constrain the maximum deviations of heading and velocity. A kernel density estimation model is used to summarize and infer the historical trajectory of the ship, revealing hidden relationships. In the Bayesian estimation model, historical ship trajectory data and trajectory features are used as inputs, and an adaptive dynamic weighting method is used for training as autocorrelation feature quantities for predicting different types of ship trajectories. A grid clustering method is used to cluster ship trajectories and predict the trajectories of selected categories of ships. The simulation results show that this method can effectively predict the safe navigation area of ships, with an average error of 2.8 kilometers after 4 hours of prediction. It does not require iterative weighting and bias, and is more suitable for shipping safety related application scenarios such as collision avoidance.*

Keywords: *Bayesian optimization; a ship; Track prediction; Grid clustering; Maritime transportation*

1. Introduction

With the development of intelligent maritime traffic control and transportation planning, research on predicting the trajectory of maritime vessels has attracted people's attention. Through predicting the operation trajectory of maritime vessels, it plays an important role in maritime traffic management and safety, route planning and resource scheduling, marine environmental protection and supervision. In areas or ports with dense maritime traffic and complex conditions, predicting the trajectory of vessels is an important research topic in the field of maritime traffic management[1]. In order to improve the navigation safety of vessels in complex and changing maritime environments, it is necessary to provide trajectory prediction and danger warning functions for intelligent navigation systems of vessels. However, the maritime navigation environment is prone to many unexpected events, and ship trajectory prediction requires consideration of various external conditions and influencing factors, such as weather, sea conditions, ship types, cargo volume, etc. This makes it difficult to accurately predict ship trajectories, especially in crowded port waters[2].

Traditional trajectory prediction methods often require the establishment of complex physical models and consideration of the interaction of multiple factors, resulting in large computational complexity and difficulty in accurate prediction[3]. Deep learning has powerful nonlinear modeling capabilities and adaptability, and has advantages in handling large-scale data and complex problems[4]. By using deep learning for ship trajectory prediction, it is possible to automatically extract features and establish a model by learning and training on a large amount of historical data, thereby obtaining more accurate prediction results[5]. Kim et al [6] decomposes ship motion into two directions: horizontal and vertical. In the horizontal direction, a Gaussian process model is used to model the uncertainty of lateral motion, and each traffic route corresponds to typical motion behavior. In the vertical direction, acceleration is used to estimate the uncertainty. By using the mean and variance to describe the lateral position and its uncertainty of the vessel, it is possible to estimate the predicted trajectory. The advantage of Gaussian process regression is its strong applicability and ease of understanding. The disadvantage is that it requires a large amount of computation, and the accuracy of the prediction results will greatly decrease over time. Last et al[7] proposed a method of fitting segmented polynomial features of ship trajectories using polynomial Kalman filters, which requires less storage space during

ISSN 2616-5775 Vol. 9, Issue 8: 90-97, DOI: 10.25236/AJCIS.2026.090811

the calculation process. The Kalman filter can estimate the system state under uncertain factors. It estimates unknown variables by combining joint distributions at different times, improving estimation and prediction accuracy. However, it has a significant impact on accuracy when the initial state is unknown.

A ship operation trajectory prediction method is proposed based on Bayesian optimization to address the drawbacks of traditional methods. Firstly, a parameter model of ship trajectories is constructed based on information such as position, velocity, and heading as raw samples. Then, a grid clustering method is used to cluster the ship trajectories to predict the trajectory of the selected category. Finally, performance verification was conducted through simulation experiments, demonstrating the reliability and superiority of the proposed method in improving the ability to predict ship operation trajectories.

2. Parameter modeling and trajectory modeling

2.1 Parameter Model of Ship Trajectory

Firstly, a parameter model of ship trajectory is constructed using information such as position, velocity, and heading as raw samples, and the problem of predicting ship trajectory is described as a Markov process. For a given region of interest (ROI), which can be assumed to be a rectangular area described by two points, a state is assigned to each cell in the grid in order[7]. Starting from the first row, the cells are encoded from left to right and top to bottom. Based on the actual position of the ship, the ship is assigned to a specific cell, and the cell number is used as the state of the Markov process.

Assuming that the coordinates of the ship correspond to the state parameter $x_0, x_1, x_2, \ldots\ldots, x_n$ of the grid model, the encoding process and prediction problem are equivalent to finding the next state of the Markov process. The first-order Markov process can be described by the following equation:

$$P\left(X^{n+1} \mid X^n, X^{n-1}, X^{n-2}, \ldots\ldots, X^0\right) = P\left(X^{n+1} \mid X^n\right) \tag{1}$$

Where, X^n represents the state of the ship at time *n*, and the current state X^n of the ship corresponds to its current position coordinates. During the time migration process, these states are marked as the next state of the ship, that is X^{n+1}. After giving a timestamp of the ship's state, the probability of the ship moving from one cell to another can be obtained

$$\pi^1 = P^T \pi^0 \tag{2}$$

In the above formula, π^0 represents the segmented initial state vector of the ship's trajectory, and *PT* represents the transpose of the ship's state transition matrix after a timestamp. With n timestamps, we can obtain:

$$\pi^n = P^T \pi^{n-1} \tag{3}$$

The speed information of the ship is obtained by the odometer or GPS. Ship odometry includes two types: absolute odometry and relative odometry. When the main engine of the ship is stopped and the ship is docked, the ship speed is displayed as zero in the absolute odometry and GPS. There is a single ship trajectory data with a length of s and n records: $D(t) = \left(D_1(t), D_2(t), \cdots, D_n(t)\right), 1 \le t \le s$, there is a window with a length of, which slides the trajectory data by data values based on the velocity parameter to form a subsequence $(s - w + 1)$. The step size of each data interval segment is equal, and the window size is w. Because noisy data influence the sliding results, the average velocity within each interval segment is calculated during the sliding process to reduce the impact of noisy data on the results. After iteratively obtaining the state of each vessel, trajectory points in non anchored states are removed to form a trajectory dataset for all anchored vessels[8].

2.2. Modeling of Ship Operation Trajectory

K-means and network clustering algorithms are used to cluster the trajectory data of anchored ships to obtain the anchorage position and to model the operating trajectory. Based on the point information of the current position of the ship, the maximum deviation constraints of heading and velocity are

ISSN 2616-5775 Vol. 9, Issue 8: 90-97, DOI: 10.25236/AJCIS.2026.090811

applied[9]. The kernel density estimation model is used to summarize and infer the historical trajectory of the ship, revealing hidden relationships. The specific implementation steps of the algorithm are as follows:

(1) Determine the initial centroid, with K initial centroids;

(2)repeat;

(3) Assign each point to the nearest centroid to form K clusters of classes;

(4) Recalculate the centroid of each cluster;

(5) Until the centroid of each cluster remains unchanged.

To address the issue of data anomalies caused by abnormal ship speeds, we compare the speed and position relationships between the ship speed anomaly points and the two preceding and following points, calculate the density parameters for each record, retain the data points corresponding to high-density parameters, and use Euclidean distance as a standard to measure the similarity between data objects:

$$dist\left(a_i, b_i\right) = \sqrt{\sum_{d=1}^{D}\left(a_{i,d} - b_{j,d}\right)^2} \tag{4}$$

The cluster center of the single ship anchoring trajectory data obtained by calculation can be used as the anchoring position of the anchored ship. In clustering, each iteration requires recalculating the corresponding cluster center. The average value of all data objects in the corresponding cluster center is the updated cluster center of the cluster:

$$\text{Cent}=\frac{1}{\left|C_k\right|}\sum_{a_i \in C_k} a_i \tag{5}$$

The problem of abnormal trajectories of different ships with the same MMSI is usually directly eliminated to ensure data cleanliness. For the same MMSI ship, the trajectory in chronological order in the selected area, anchored ship, is defined as:

$$\text{recor}=\left(MMSI, l, T, t_s, t_e, p\right) \tag{6}$$

Where, l represents the length of the vessel (m); T represents the type of vessel; t_s represents the starting time of ship anchoring; t_e is the termination time of vessel anchorage; p represents the position of the anchor point ( ).

By processing the ship trajectory data mentioned above, traverse the trajectory data of each ship and calculate the average value of each data interval segment. Choosing the location with the highest probability can predict the future position of the ship:

$$\pi^{(t+T)} = P^T\left(\pi^{(t,(t-T))}\right) \tag{7}$$

The value of T in the above equation is 30 minutes, $\pi^{(t,(t-T))} \in \mathbb{R}^{n^2}$. During the process of ship anchoring, the characteristics of the ship's trajectory are obtained by using the backward and forward throwing methods to anchor, and the current and past position information of the ship is used to predict the future position. The future position coordinates of a ship can be considered as generated from a random distribution, whose parameters depend on the current and past positions[10].

3. Implementation of ship trajectory prediction

3.1 Track parameter estimation based on Bayesian estimation model

The ship trajectory prediction problem can be regarded as a time series prediction problem because the data is ordered in time. The goal of this task is to use the current and past position information of the vessel to predict future positions. Assuming that at time t, the vessel's position is x_t, the next

ISSN 2616-5775 Vol. 9, Issue 8: 90-97, DOI: 10.25236/AJCIS.2026.090811

position of the vessel can be represented as x_{t-1}, and the position of the preceding time window can correspond to $x_{t-1}, x_{t-2}, x_{t-3}$, etc. Therefore, the problem of ship trajectory prediction is transformed into a problem of finding the position x_t of the ship at time t:

$$x_t = (\text{Longitude}_t,\ \text{Latitude}_t) \tag{8}$$

So the next position of the ship can be represented as x_{t+1}, and the position of the preceding time window can be represented as $x_{t-1}, x_{t-2}, x_{t-3}$ In order to predict from historical data x_{t+1}, it is necessary to learn an *h* that:

$$x_{t+1} \approx h\left(x_t, x_{t-1}, \ldots, x_{t-k}\right) \tag{9}$$

Where k is the number of preceding timestamps. The input feature space consists of the current and previous positions of the ship, and the output vector consists of the next or future position of the ship. To obtain the final prediction method, the spherical distance between the actual position coordinates and the predicted position coordinates is calculated, which is defined as the prediction error of the trajectory:

$$e_i^{t+nT} = 2R\sin^{-1}\left(\sqrt{\sin^2\left(\frac{Lat_2 - Lat_1}{2}\right) + \cos\left(Lat_2\right)\cos\left(Lat_1\right)\sin^2\left(\frac{Lon_2 - Lon_1}{2}\right)}\right) \tag{10}$$

Based on Bayesian estimation, after the area is divided into grids, each cell is treated as a feature according to the position of the ship in each cell, and it is converted into a vector representation using one hot encoding. The previous position can be represented as a vector of length 16, and the next position can be represented as another vector of length 16, forming a feature vector of length 32. The adaptive dynamic weighting method is used to train and use it as an autocorrelation feature for predicting the trajectories of different types of ships, thereby achieving parameter estimation and prediction.

3.2. Ship trajectory prediction and performance indicators

In trajectory prediction, continuous input values such as x_{t-1} and x_t are used as feature spaces, which also include factors such as vessel length, destination, and vessel type. We will use the next position as the output marker to transform the classification problem from the previous section into a regression problem. In addition, the position coordinates of two timestamps were considered. The reason for using vessel types as features is that vessels of the same type typically exhibit similar behavioral patterns. This article uses data from two different types of ships, cargo ships and oil tankers. These vectors are used as inputs to the Kernel Density Estimation (KDE) algorithm. Specifically, the heat map generated by this method can be described as

$$P_{KDE}(D) = \frac{1}{n}\sum_{i=1}^{n} K\left(D - D_i\right) \tag{11}$$

In the above equation, based on trajectory data D, the i-th trajectory rate density estimate can be calculated using kernel function K. The KDE heatmap is further randomly cropped to train with the goal of minimizing the mean square error (MSE) of the reconstruction error. The corresponding loss function can be expressed as:

$$CAE_{MSE} = \frac{1}{n * m} \| P_i - \hat{P}_i \|_2^2 \tag{12}$$

In order to evaluate the performance of different methods, this article used three metrics, namely Average Displacement Error (ADE), Final Displacement Error (FDE), and Interval Accuracy (IAcc). All indicators are provided in kilometers. Among them, the ADE index represents the average L2 distance from the real trajectory to the horizontal trajectory at all time steps. Specific description:

$$ADE = \frac{\sum_{i=1}^{N}\sum_{t=0}^{t=T_{\text{pred?}}} \| \hat{Y}_i^t - Y_i^t \|_2}{NT_{pred}} \tag{13}$$

ISSN 2616-5775 Vol. 9, Issue 8: 90-97, DOI: 10.25236/AJCIS.2026.090811

Where N is the total number of ships, and T_{pred} is the predicted number of time steps, *FDE* is the L2 distance on the last length, expressed as the distance between Yi and i, expressed as:

$$FDE = \frac{\sum_{i=1}^{N} || \hat{Y}_i^{T_{\text{pred?}}} - Y_i^{T_{\text{pred?}}} ||_2}{N} \tag{14}$$

In addition, *IAcc* refers to the percentage of time step errors in the predicted trajectory that fall within the 2-standard range of the actual trajectory. This ratio is used to measure the prediction accuracy of the model

$$IAcc = \frac{\sum_{i=1}^{N} \sum_{t=0}^{t=T_{pred}} acc_i^t}{NT_{pred}} \tag{15}$$

Where, acc_i^t represents the acceptance rate of the trajectory planner considering the uncertainty of ship operation, and thus, a prediction model and indicator performance for ship operation trajectory are constructed.

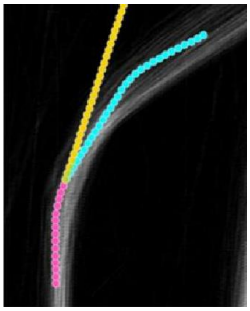
4. Simulation experiment

In the simulation experiment, the generator model was sampled 20 times to generate a set of 20 candidate trajectories. In addition to the mkvGAN network based on Markov model probability distribution input described in this article, this article also measured the GAN model that directly accepts the ship's operating trajectory as input (hereinafter referred to as the general GAN model) to compare the effect of obtaining trajectory distribution characteristics through Markov trajectory distribution. For GAN and LSTM models, this paper uses average trajectories to calculate FDE and ADE. Comparing the above two GAN networks, the final results are shown in Table 1, where FDE and ADE indicators are measured in kilometers.

Table 1 Statistical Table of Prediction Performance of Multiple Prediction Models

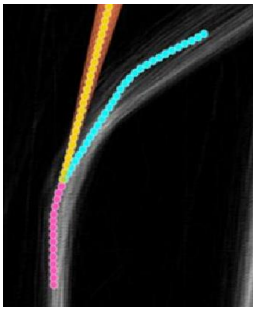
Comparative indicators	Markov model	LSTM model	GAN	Bayesia
ADE	0.489	0.498	0.359	0.370
FDE	1.361	1.153	1.399	1.303
IACC	0.916	0.767	0.891	0.621

From the table, it can be seen that the mkvGAN model proposed in this article outperforms other prediction models in both ADE and FDE aspects. At the same time, the GAN model that directly generates predictions from ship trajectories is worse than the Markov model and LSTM model, proving that historical trajectory features are of great help to prediction results. Markov models and LSTM models have good performance in trajectory prediction accuracy, therefore they have good IAcc. Considering that large vessels generally need to maintain a safe distance of 0.5 nautical miles from other vessels, the FDE obtained by the mkvGAN model is 0.677, equivalent to 0.36 nautical miles, which can ensure the safe distance of vessel operation. This makes the mkvGAN model more suitable for collision avoidance through trajectory prediction compared to other models, which has significant implications for navigation safety. The predictive performance of four models for short-term turning of ships is shown in Figure 1.

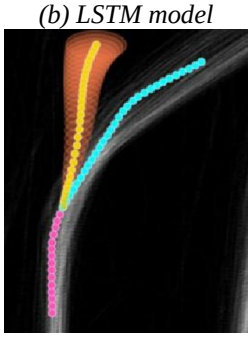


(a) Markov model

ISSN 2616-5775 Vol. 9, Issue 8: 90-97, DOI: 10.25236/AJCIS.2026.090811



(b) LSTM model



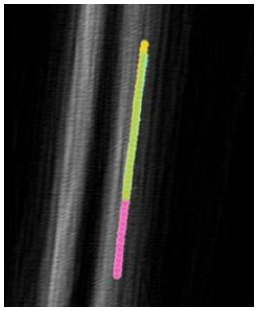
(c) GAN



(d) Bayesian optimization model

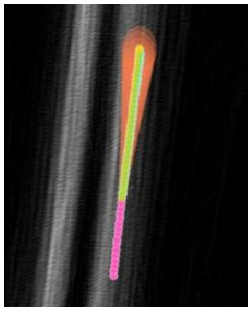
Figure 1: The predictive performance of four models for short-term turning of ships during operation

The pink dots in Figure 1 represent the historical trajectory of the input, the blue data dots represent the true ground trajectory, the yellow dots represent the average predicted value of the given method, and the orange ellipsoid represents the 95% confidence interval of the given time step. From the above figure, it can be seen that by combining the use of Markov transition probability, Bayesian optimization models can better predict the short-term position of ships and achieve collision avoidance. The simulation is shown in Figure 2.

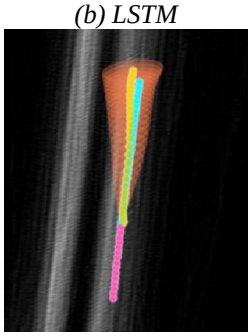


(a) Markov model

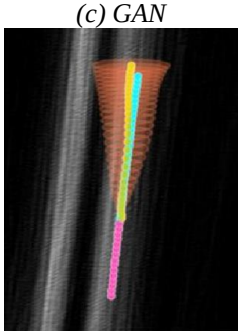
ISSN 2616-5775 Vol. 9, Issue 8: 90-97, DOI: 10.25236/AJCIS.2026.090811



(b) LSTM



(c) GAN



(d) Bayesian optimization model

Figure 2 Prediction effect of straight ship trajectory

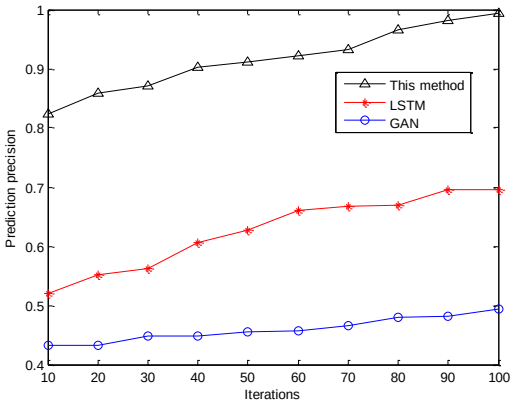


Figure 3 Predicted Position Coordinate Feature Error of Ship Operation Trajectory

According to Figure 2, the Bayesian optimization model and LSTM model have more accurate predictive performance compared to the other two generative network-based models. The reason is that the KDE probability distribution generated through Markov transition probability adds more disturbance to trajectory prediction, so traditional trajectory prediction methods are a better solution when used for straight line prediction or long-term trajectory prediction. The error is compared with that of a model that contains only two timestamped positions, and the accuracy of the predicted position coordinate feature of the ship's trajectory is tested, as shown in Figure 3. The prediction error is shown in Table 2. Analysis of the simulation results in Figure 3 shows that the method proposed in this paper has a relatively small error in predicting ship trajectories, with an average error of 2.8

ISSN 2616-5775 Vol. 9, Issue 8: 90-97, DOI: 10.25236/AJCIS.2026.090811

kilometers after 4 hours of prediction. It does not require iterative weighting and bias, and is more suitable for shipping safety related application scenarios such as collision avoidance.

Table 2 Comparison of Prediction Errors

Iteration number	This method	LSTM model	GAN model
10	0.096	0.242	0.210
20	0.174	0.265	0.460
30	0.067	0.365	0.494
40	0.168	0.249	0.326
50	0.119	0.280	0.410
60	0.072	0.356	0.395

5. Conclusions

In this article, a ship operation trajectory prediction method based on Bayesian optimization is proposed. K-means and network clustering algorithms are used to cluster the trajectory data of anchored ships to obtain the anchorage position and to model the running trajectory. Based on the point information of the current position of the ship, the maximum deviation constraints of heading and speed are applied. The kernel density estimation model is used to summarize and infer the historical trajectory of the ship. Two different types of ship data, cargo ships and oil tankers, are used as inputs to the kernel density estimation algorithm (KDE). Grid clustering method is used to cluster the ship trajectory and predict the trajectory of the selected category of ships. Research has shown that the method proposed in this article has lower errors in predicting ship trajectories, and can better predict the short-term position of ships, achieving collision avoidance without the need for iterative weighting and bias. It is more suitable for applications related to shipping safety, such as collision avoidance, and can accurately predict the distribution of ship trajectories that traditional trajectory prediction methods cannot directly predict.

References

[1] JIANG Lingling. Research on Ship Movement Trajectory Tracking at Sea Assisted by 3D Computer Navigation [J]. Ship Science and Technology,2025,47(06):154-157.
[2] Wang Ziming, Chen Shunhuai, Zheng Long. Dynamic collision avoidance trajectory planning for unmanned ships based on improved PRM algorithm and event triggering mechanism [J]. Ship mechanics,2025,29(08):1239-1249.
[3] Jin Xiaodi. Finite time learning control for ship trajectory tracking [J]. Ship engineering, 2025,47(07):107-111.
[4] Zhu Mengwei. Application of VR Virtual Reality Technology in Ship Navigation Trajectory [J]. Ship Science and Technology,2019,41(22):31-33.
[5] Capobianco S, Millefiori L M, Forti N, et al. Deep learning methods for vessel trajectory prediction based on recurrent neural networks[J]. IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(6): 4329-4346.
[6] Kim K I, Lee K M. Deep learning-based caution area traffic prediction with automatic identification system sensor data[J]. Sensors, 2018, 18(9): 3172.
[7] Last P, Hering-Bertram M, Linsen L. Interactive history-based vessel movement prediction[J]. IEEE Intelligent Systems, 2019, 34(6): 3-13.
[8] Volkova T A, Balykina Y E, Bespalov A. Predicting ship trajectory based on neural networks using AIS data[J]. Journal of Marine Science and Engineering, 2021, 9(3): 254.
[9] Chen X, Ling J, Yang Y, et al. Ship trajectory reconstruction from AIS sensory data via data quality control and prediction[J]. Mathematical Problems in Engineering, 2020, 2020: 1-9.
[10] Jin Hao. Ship trajectory prediction based on multi-scale LSTM Attention [J]. Ship electrical technology,2025,45(07):76-81.