Research Article

Ship trajectory prediction via a transformer-based model by considering spatial-temporal dependency

Xinqiang Chen[1,2], Peiyang Wu[1], Yuzhen Wu[3], Loay Aboud[4], Octavian Postolache[5,6], Zichuang Wang[7]

[1]Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China.
[2]Chongqing Key Laboratory of Green Logistics Intelligent Technology, Chongqing Jiaotong University, Chongqing 400074, China.
[3]SPG Qingdao Port Group Co., Ltd., Shandong Port Group, CO. LTD., Qingdao 266011, Shandong, China.
[4]College of Maritime Transport & Technology CMTT, Arab Academy for Science & Technology and Maritime Transport AASTMT, Alexandria 1029, Egypt.
[5]Instituto de Telecomunicações (IT), Lisbon 1049-001, Portugal.
[6]Department of Information Science and Technology, Iscte - Instituto Universitário de Lisboa, Lisbon 1649-026, Portugal.
[7]Shanghai Ship and Shipping Research Institute Co., Ltd., Shanghai 201306, China.

Correspondence to: Prof. Octavian Postolache, Instituto de Telecomunicações (IT), North Tower, 10th Floor, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal. E-mail: octavian.adrian.postolache@iscte-iul.pt

Abstract

With the rapid development of global maritime trade, ship trajectory prediction plays an increasingly important role in maritime safety, efficiency optimization, and the development of green shipping. However, the complexity of the marine environment, multi-factor influences, and automatic identification system (AIS) data quality issues pose significant challenges to trajectory prediction. This study proposes a ship trajectory prediction model based on the Crossformer architecture comprising three core components: Dimension-Segment-Wise embedding, Two-Stage Attention layer, and Hierarchical Encoder-Decoder structure, which efficiently captures spatiotemporal dependencies in ship movement patterns. Through experiments on public AIS datasets, we validate the model using two navigation scenarios (complex turning and smooth sailing) and conducted comprehensive comparisons with traditional models such as gated recurrent unit (GRU), long short-term memory (LSTM), and temporal graph convolutional network (TGCN). Experimental results demonstrate that Crossformer significantly outperforms the comparative models across multiple evaluation metrics including average Euclidean distance error (ADE), mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE), reducing average error by over 60% in complex scenarios and up to 70% in smooth scenarios. For Case 1, Crossformer achieved the lowest

values across metrics with ADE of 2.35×10^{-2}, MSE of 7.00×10^{-4}, RMSE of 2.58×10^{-2}, and MAE of 2.35×10^{-2}, substantially outperforming GRU, LSTM, and TGCN models. For Case 2, Crossformer similarly excelled with an ADE of 1.64×10^{-2}, MSE of 4.00×10^{-4}, RMSE of 2.06×10^{-2}, and MAE of 1.64×10^{-2}. The model maintains low error levels in predicting both latitude and longitude dimensions, exhibiting excellent multi-dimensional prediction capability and robustness. This research not only provides a high-precision solution for ship trajectory prediction but also establishes an important technical foundation for intelligent ship scheduling, maritime traffic management, and navigation safety assurance.

Keywords: Ship trajectory prediction, Crossformer model, spatial-temporal dependency, smart ship

1. INTRODUCTION

In the context of rapidly developing international maritime shipping, ship trajectory prediction plays a significant role in enhancing maritime safety and efficiency. Ship trajectory prediction is the process of scientifically estimating future navigation routes, ship speeds, and arrival times based on comprehensive utilization of historical voyage data and real-time information. Accurate trajectory prediction can effectively mitigate maritime accidents, optimize port resource allocation, and reduce ship fuel consumption and pollutant emissions, thereby promoting the development of green shipping. However, current ship trajectory prediction still faces numerous challenges. First, adverse weather conditions, ocean current variations, and navigational channel factors make prediction difficult. Second, due to the influence of multiple aspects including voyage planning, operational decision-making, and economic considerations, the maritime industry faces greater risks. Simultaneously, while the extensive automatic identification system (AIS) provides rich navigational information, it also encounters issues such as signal loss, noise interference, and data incompleteness[1].

To date, there are two relatively common approaches to ship trajectory prediction: traditional machine learning methods and neural network-based methods[2]. Murray *et al.* first proposed the neighborhood heading distribution method, introducing probabilistic models such as Gaussian mixture models to handle the multi-modal nature of maritime traffic[3]. On this basis, researchers have increasingly turned to neural network-based methods, harnessing the strengths of deep learning to capture the complex, nonlinear patterns inherent in AIS data. For instance, Dalsnes *et al.* developed bilinear autoencoders and hybrid neural network models, significantly enhancing the accuracy and robustness of trajectory prediction[4]. These developments reflect the steady progress in ship trajectory prediction, with each approach contributing new insights while also revealing limitations that guide future advancements. Zhen *et al.* developed a method combining ship trajectory clustering and Naive Bayes classification to detect anomalous ship behavior in maritime surveillance systems, enhancing situational awareness in coastal waters[5]. Although these methods have the advantages of low computational complexity and ease of interpretation, the fundamental limitation is that they assume that the ship's motion is pre-designed, making it difficult to capture the complex nonlinear dynamics of the marine environment. Perera *et al.* proposed an extended Kalman filter method for estimating and predicting longitudinal ship trajectories in ocean navigation, implementing a curved motion model that successfully estimates ship position, velocity, and acceleration from noisy position measurements to enhance maritime safety and security systems[6]. However, this method is sensitive to noise and struggles with sudden maneuvers, limiting its effectiveness in complex navigation environments. Liu *et al.* introduced a support vector machine (SVM) prediction algorithm incorporating an adaptive chaotic differential evolution algorithm[7]. While this approach offers certain improvements in nonlinear modeling, its performance remains heavily dependent on the quality of feature engineering. Moreover, it faces challenges in capturing the inherent dependencies within high-dimensional spatio-temporal data.

The current traditional methods are not accurate enough for ship trajectory prediction, and now neural network-based methods are more advantageous. Suo *et al.* proposed a ship trajectory prediction framework utilizing gated recurrent unit (GRU) neural networks to process AIS data, establishing a long short-term memory (LSTM)-based ship trajectory prediction method with high prediction accuracy[8]. Borkowski proposed an algorithm that, through data fusion, considers measurements of a ship's current position from multiple dual autonomous devices to improve prediction reliability and accuracy. The algorithm employs artificial neural networks with adaptive training using data strings of varying lengths to predict trajectories of other ships[9]. Gan *et al.* introduced an innovative algorithm based on a multilayer perceptron (MLP) network with optimized hidden neurons, which adjusts network parameters through particle swarm optimization methods, significantly improving the accuracy of long-term ship speed prediction[10]. Park *et al.* proposed a novel ship trajectory prediction method that innovatively combines spectral clustering techniques with bidirectional LSTM networks (Bi-LSTM), adopting the longest common subsequence (LCSS) distance metric to quantify similarity between trajectories[11]. Zhao *et al.* introduced graph attention networks (GAT) and LSTM for ship trajectory prediction. The GAT-LSTM constructs a ship trajectory graph network based on dependencies between ship trajectory data, using GAT to extract spatial features of ship trajectory data and introducing LSTM to learn temporal features of ship trajectory data[12]. Johansen *et al.* proposed a conceptual framework for a ship collision avoidance system based on model predictive control. The system generates a finite set of alternative control strategies by dynamically adjusting two key parameters: first, the offset adjustment to the autopilot guidance heading angle, and second, the modification to propulsion commands. The core mechanism of the system relies on precise simulation predictions of obstacle positions and potential ship trajectories[13].

Zhou *et al.* proposed a trajectory prediction method integrating AIS data with back propagation (BP) neural networks. Based on the fundamental principles of BP neural networks, this method innovatively uses a ship's navigational behavior features at three consecutive time points as input variables and the behavioral features at the fourth time point as output variables, training the BP neural network through this pattern to achieve effective prediction of future ship navigation trajectories[14]. Gao *et al.* proposed a novel MP-LSTM method that combines the advantages of TPNet and LSTM, involving four components: AIS data preprocessing methods, solutions for target points and support points, and uncertainty analysis. This method demonstrates high prediction accuracy[15]. Graph neural network (GNN)-based methods are severely limited as they rely on predefined graph structures to capture static ship attributes, thus failing to consider dynamic interactions between ships[16]. Secondly, recurrent neural network (RNN)-based methods[17] show insufficient capability for modeling long-term temporal dependencies. Additionally, their complex recurrent architectures inhibit effective modeling of local (short-term) temporal dependencies, resulting in suboptimal inference efficiency. A novel gated spatio-temporal graph aggregation network (G-STGAN) has been proposed, comprising a ship spatial gated encoder (SSGE) that integrates graph convolutional networks (GCN) with transformer architecture to model dynamic and static spatial interactions. This approach builds on previous advances in graph-based modeling, which focus on spatial dependencies, and transformer architectures, known for capturing long-range temporal patterns. By bringing these techniques together, G-STGAN overcomes the limitations of traditional GNNs and RNNs, offering a more complete and nuanced representation of ship trajectories. Thereby enhancing predictive performance for ships, while also featuring a spatial-temporal gated encoder (STGE) that utilizes gated transformers (GT) and temporal convolution (TC) to effectively capture short-term and long-term temporal dependencies. Spatial and temporal features extracted from the SSGE and STGE modules are subsequently aggregated through temporal convolutional networks (TCN) to generate comprehensive trajectory predictions[18]. To advance intelligent maritime navigation, Jiang *et al.* proposed a spatio-temporal multi-graph fusion network (STMGF-Net), designed to model the complex spatio-temporal interactions among multiple ships using AIS data[19]. The model constructs multiple interaction graphs - such as motion,

risk, and attribute graphs - and integrates them through a multi-modal fusion framework. By incorporating squeeze-and-excitation TC modules, STMGF-Net improves both prediction accuracy and computational efficiency. Similarly, Wang *et al.* introduced an enhanced version of STMGF-Net, further refining the modeling of spatial and temporal dependencies among vessels and demonstrating superior trajectory prediction performance over classical and state-of-the-art GNN-based approaches[20]. Additionally, Syed and Ahmed developed a 1D CNN-LSTM framework tailored for continuous AIS data, treating vessel trajectories as multivariate time series. This architecture effectively captures spatial patterns and long-term temporal dependencies, maintaining robust performance even when dealing with overlapping trajectories or missing data. Experimental evaluations on AIS datasets confirm that this approach outperforms comparable neural network models in tracking accuracy[21].

2. METHOD

2.1. Data preprocessing

AIS is a ship dynamic information collection system based on wireless communication technology, widely applied in maritime traffic monitoring and ship management. During AIS data transmission, factors such as signal instability or channel congestion can cause anomalies in the data, including duplicate records and erroneous data[22]. These anomalous data, if left unprocessed, would affect model performance in experiments. Therefore, this research conducts data cleaning prior to experimentation. The data cleaning process comprises five steps: anomaly detection, information extraction, data interpolation, equal-interval processing, and data standardization[23].

Anomaly detection and information extraction were accomplished by establishing rational standards for key fields such as maritime mobile service identity (MMSI), BaseDateTime, latitude (LAT), longitude (LON), course over ground (COG), and speed over ground (SOG). Rows with MMSI values not containing 9 digits were deleted, as were rows lacking BaseDateTime, LAT, or LON values. Additionally, only rows with COG values between 0 and 360 and SOG values between 0 and 30 were retained. To address the requirements for trajectory correlation analysis between different ships, segmented cubic spline interpolation was applied to interpolate LAT and LON at equal time intervals[24], as given in

$$f(r) = \begin{cases} a_1 + b_1 r + c_1 r^2 + d_1 r^3 & r \in [r_0, r_1] \\ a_2 + b_2 r + c_2 r^2 + d_2 r^3 & r \in [r_1, r_2] \\ \vdots \\ a_n + b_n r + c_n r^2 + d_n r^3 & r \in [r_{n-1}, r_n] \end{cases} \tag{1}$$

where r_1, r_2, …, r_n denote the characteristic points of the ship trajectory; the trajectory segment of every two neighboring characteristic points is a section of cubic polynomial curve and satisfies the smoothness constraints that the function values are continuous (C^0-continuous) and the derivatives of the first and second orders are continuous (C^1, C^2-continuous) at the characteristic points. This method can effectively deal with the sampling problem of different time intervals during the ship navigation.

Finally, in order to eliminate the influence of different parameters on the data analysis, we normalize the data, and the normalized values are defined as

$$x_{\text{std}} = \frac{x_{\text{raw}} - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \tag{2}$$

where x_{std} is the normalized value, x_{raw} is the original observation, x_{max} and x_{min} represents the maximum and minimum values of the data.

2.2. Model structure

In this research, as shown in Figure 1, Crossformer is a Transformer model specifically designed for multivariate time series prediction[25]. The model has three key components: Dimension-Segment-Wise (DSW) Embedding, Two-Stage Attention (TSA) Layer, and Hierarchical Encoder-Decoder (HED) structure. The structure of the model is illustrated in Figure 1.

2.2.1. DSW embedding

As shown in Figure 2, traditional Transformer methods concatenate all variables at the same time step into a single vector, which is then linearly mapped to obtain embeddings. This approach only focuses on cross-temporal dependencies and fails to fully exploit the spatial correlations between different variables.

DSW Embedding is a new embedding method that processes the time series of each ship navigation variable (such as LAT and LON) by segmenting them. Each segment is transformed into a fixed-dimensional vector through a learnable linear mapping, while also incorporating the corresponding positional embedding $E^{(pos)}$ to preserve temporal position information. This generates a two-dimensional vector array that simultaneously carries information about both the ship's navigation time and geographical coordinates, with the specific process given in

$$S_{1:T} = \left\{ S_{i,j}^{(seg)} \mid 1 \leq i \leq \frac{T}{L}, 1 \leq j \leq D \right\} \tag{3}$$

$$S_{i,j}^{(seg)} = \left\{ S_{t,j} \mid (i-1) \times L < t \leq i \times L \right\} \tag{4}$$

where $S_{i,j}^{(seg)} \in \mathbb{R}$ represents the i-th time period in the preprocessed AIS data variable j, with each period standardized through spline interpolation to L time steps at fixed 1-minute intervals.

Each standardized AIS data segment is encoded through linear projection and positional embedding, which is given in

$$V_{i,j} = \Phi S_{i,j}^{(seg)} + \Psi_{i,j}^{(pos)} \tag{5}$$

This embedding approach is suitable for processing AIS data after anomaly detection, including cases with missing data, ensuring the quality of the data input to the model.

2.2.2. TSA layer

The TSA layer design takes into account the quasi-continuous nature of AIS data. In Figure 3, it includes two key stages:

The cross-time stage applies a multi-head self-attention mechanism to the dimensions (such as LAT and LON) in the AIS data, capturing the temporal patterns of ship movement, as given in

$$\widetilde{A}_{:,j}^t = \text{Norm}\left(A_{:,j} + \text{MHSA}^t(A_{:,j}, A_{:,j}, A_{:,j}) \right) \tag{6}$$

$$A^t = \text{Norm}\left(\widetilde{A}^t + \text{FFN}(\widetilde{A}^t) \right) \tag{7}$$
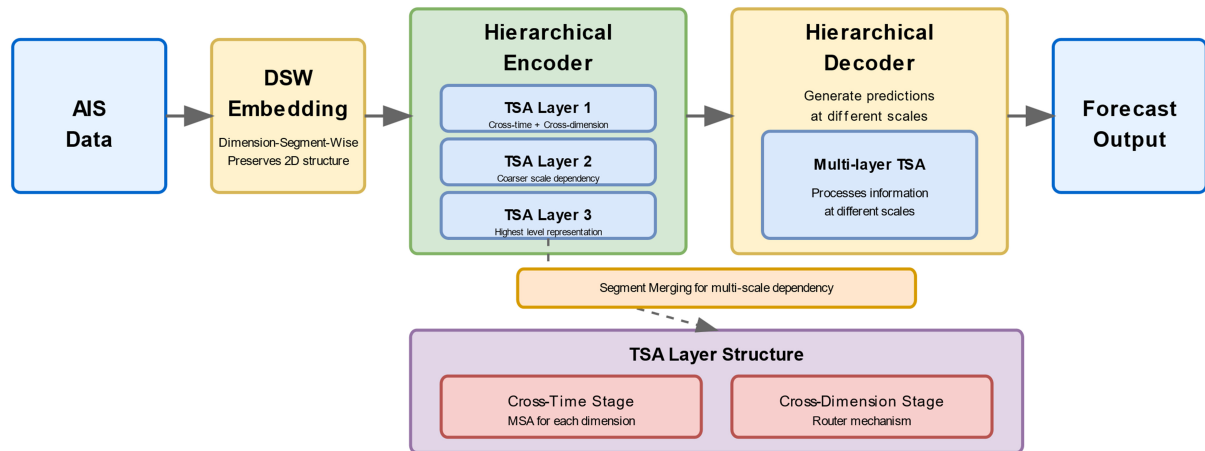
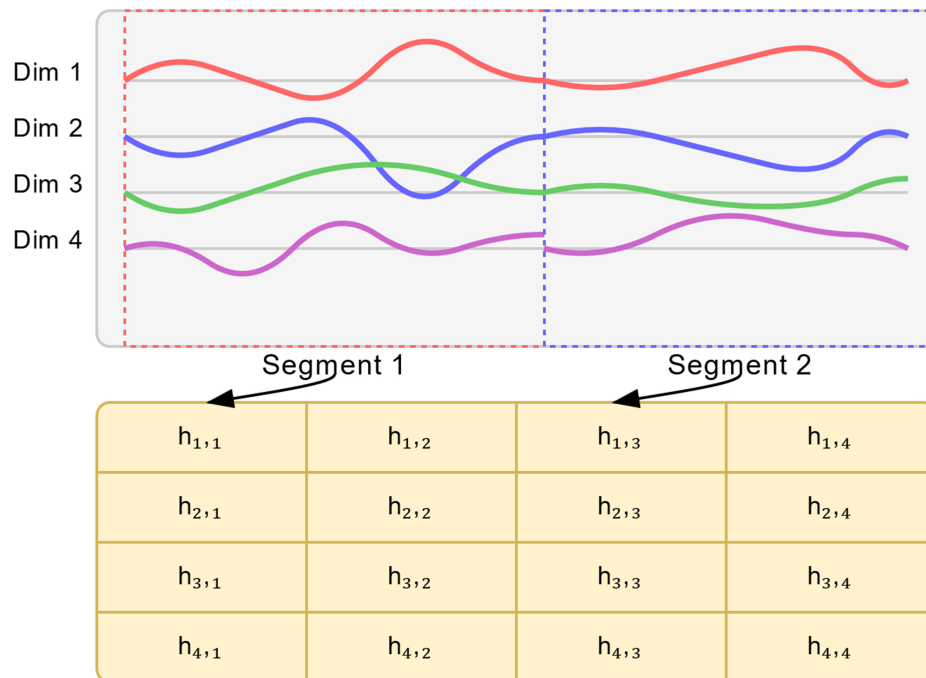**Figure 1.** Crossformer architecture diagram.

Figure 2. DSW embedding model. DSW: Dimension-Segment-Wise.

The other is the cross-dimension stage, which adopts a router mechanism that can efficiently establish connections between various dimensions of ship trajectory data, achieving information fusion among navigational parameters.
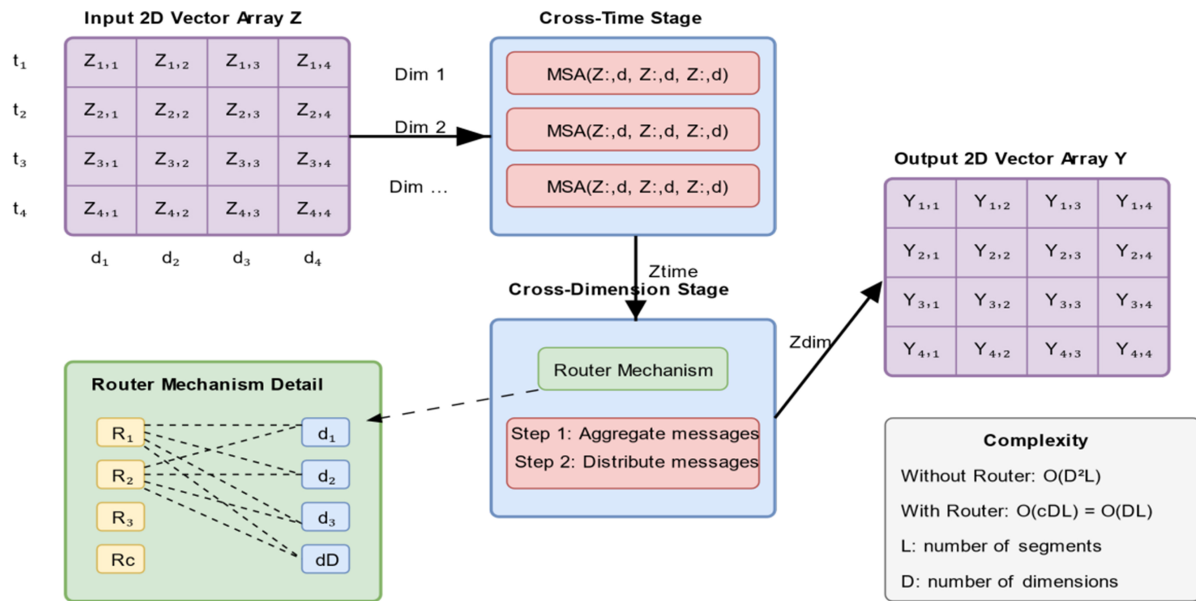
Figure 3. Structure and workflow of TSA. TSA: Two-Stage Attention.

$$C_{i,:} = \text{MHSA}_1^d(Q_{i,:}, A_{i,:}^t, A_{i,:}^t), 1 \leq i \leq L' \tag{8}$$

$$A_{i,:}^d = \text{MHSA}_2^d(A_{i,:}^t, C_{i,:}, C_{i,:}), 1 \leq i \leq L' \tag{9}$$

$$\widetilde{A}^d = \text{Norm}(A^t + A^d) \tag{10}$$

$$A^d = \text{Norm}\big(\widetilde{A}^d + \text{FFN}(\widetilde{A}^d)\big) \tag{11}$$

where $Q \in \mathbb{R}^{L \times c \times d_{model}}$ is c learnable router vector, and $C \in \mathbb{R}^{L \times c \times d_{model}}$ is the aggregated ship information. The router mechanism establishes information exchange between dimensions by setting a fixed number of learnable vectors as routers.

2.2.3. HED architecture

As illustrated in Figure 4, the diagram demonstrates how Crossformer processes normalized AIS trajectory data through a hierarchical approach for ship trajectory prediction. Beginning with the processing of AIS data at the bottom layer, each ascension to a higher layer involves the fusion of adjacent temporal segments, enabling higher layers to capture coarser-grained temporal dependencies in ship movement. This architecture effectively models route planning. Predictions are generated at multiple scales and aggregated to form the final prediction:

$$S_{i,j}^{(pred),l} = \Gamma^l D_{i,j}^l \tag{12}$$

where $\Gamma^l \in \mathbb{R}^{L \times d}$ is the learnable projection matrix.

3. EXPERIMENTATION AND ANALYSIS

3.1. Data

In this study, we utilized AIS data publicly accessible from the MarineCadastre website, which provides free information on ship movements in U.S. waters from 2009 to 2023. The MarineCadastre database, maintained through collaboration between the Bureau of Ocean Energy Management (BOEM) and the
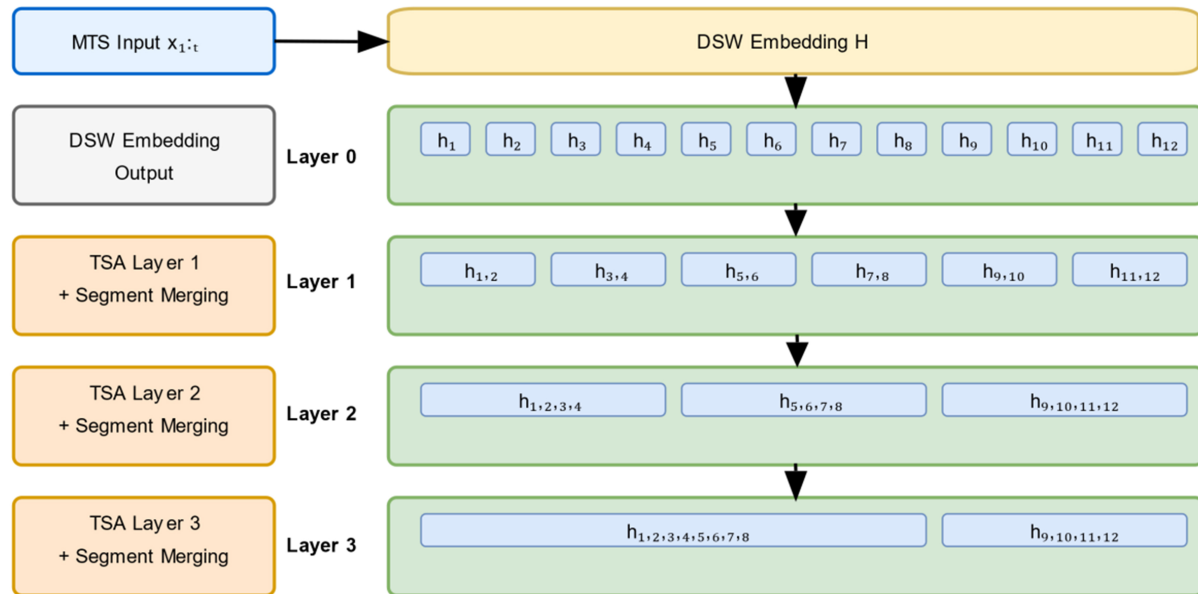
Figure 4. Hierarchical encoder structure model.

National Oceanic and Atmospheric Administration (NOAA). For this research, AIS data from March 2020 were selected, specifically focusing on the coastal waters of the Western Seaboard with LAT ranging from 33°N to 34°N, and LON ranging from 118°W to 120°W. In accordance with international coordinate standards, northern LATs and western LONs are represented as positive and negative values, respectively, maintaining consistency with navigational conventions.

3.2. Indicators for experimental evaluation

This study compares predicted AIS data with actual reference data using a statistically validated evaluation framework recognized in academic circles. The assessment employs average Euclidean distance error (ADE), final Euclidean distance error (FDE), mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE) as metrics to evaluate the accuracy of ship trajectory predictions. To evaluate the performance of the trajectory prediction models, several commonly used metrics are employed in this study. The ADE calculates the mean Euclidean distance between each predicted point and its corresponding ground truth across the entire trajectory. The FDE, on the other hand, measures the Euclidean distance between the predicted endpoint and the actual endpoint. Additionally, the MSE captures the average of the squared differences between predicted and true values, while the RMSE - as the square root of MSE - reflects the standard deviation of prediction errors. Finally, the MAE provides the average absolute difference between predictions and ground truth. In our study, these metrics collectively offer a comprehensive assessment of both the overall prediction quality and endpoint accuracy. ADE measures the average deviation between the entire predicted trajectory and the actual trajectory, while FDE quantifies the deviation between the predicted endpoint and the actual endpoint. MSE, RMSE, and MAE reflect the magnitude of prediction errors from different statistical perspectives. For all metrics, lower values indicate higher prediction accuracy of the ship trajectory model. These evaluation metrics provide a multi-dimensional assessment framework that enables objective comparison between different prediction models. The comprehensive evaluation method ensures that both overall trajectory similarity and specific location accuracy are appropriately quantified.

$$\text{ADE} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{(x_i^{\text{pred}} - x_i^{\text{gt}})^2 + (y_i^{\text{pred}} - y_i^{\text{gt}})^2} \tag{13}$$

$$\text{FDE} = \sqrt{(x_N^{\text{pred}} - x_N^{\text{gt}})^2 + (y_N^{\text{pred}} - y_N^{\text{gt}})^2} \tag{14}$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i^{\text{pred}} - y_i^{\text{gt}})^2 \tag{15}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i^{\text{pred}} - y_i^{\text{gt}})^2} \tag{16}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |y_i^{\text{pred}} - y_i^{\text{gt}}| \tag{17}$$

where N denotes the total number of predicted trajectory points. $(x_i^{\text{pred}}, y_i^{\text{pred}})$ and $(x_i^{\text{gt}}, y_i^{\text{gt}})$ represent the predicted and ground truth (reference) coordinates of the ship at the *i*-th time step, respectively. Similarly, y_i^{pred} and y_i^{gt} denote the predicted and ground truth values (e.g., LAT, LON) at the *i*-th time step. These statistical indicators serve as quantitative measures for evaluating the prediction accuracy of ship trajectories.

3.3. Experimental setup

In order to validate the performance of the proposed model in this study, we selected some existing ship trajectory prediction models as comparative models, including some established models, such as GRU[26,27], LSTM[28], and recent-developed temporal graph convolutional network (TGCN)[29].

All methods in this study were implemented in the PyTorch framework in Python, with training conducted on a single NVIDIA RTX 4060 GPU. The dataset was partitioned into training, validation, and test sets. From Table 1, we determined the optimal experimental parameters, with comparative results shown in the table. As indicated in the table, when the initial learning rate was set to 0.001, the model yielded the minimum evaluation metrics. Consistent experimental parameters were maintained across all experiments during the training phase[30].
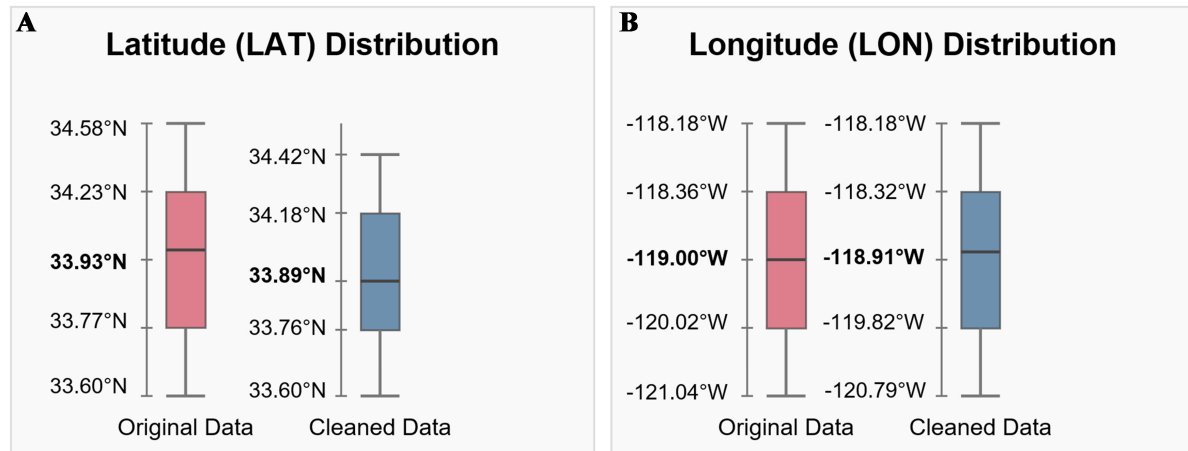
To provide clear illustrations of the model's effectiveness, we selected two representative ships from the dataset and presented their trajectories as examples in this paper. For convenience, these trajectories are referred to as Case 1 and Case 2. Case 1 represents a ship entering the port, with a trajectory containing multiple complex turning points, as illustrated. Case 2 depicts a ship departing the port, featuring a smooth trajectory during its progression, as shown. We observe the distribution of both ship trajectories, noting that the LAT and LON of the ships vary within a small range, indicating that the ships navigated back and forth within a confined area. However, several significant outliers are identified in the original ship trajectory data, necessitating preprocessing. The distribution diagrams presented in Figure 5 demonstrate the comparison between original data and cleaned data regarding LAT and LON parameters. This box plot visualization effectively illustrates the impact of the data preprocessing framework on spatial distribution characteristics.

For the LAT distribution shown in Figure 5A, the original data ranged approximately from 33.60°N to 34.25°N. After cleaning, the standard deviation decreased by 7.6%, and the data range contracted from 0.98° to 0.82°, representing a reduction of 16.2%. The box plot indicates that the cleaning process removed certain

Table 1. Comparison of different learning rates

Learning rate	1×10^{-3}	1×10^{-4}	5×10^{-4}	1×10^{-5}
MSE	4.66×10^{-3}	6.26×10^{-3}	1.897×10^{-2}	4.28×10^{-2}
MAE	4.81×10^{-2}	5.418×10^{-2}	9.969×10^{-2}	1.522×10^{-1}

MSE: Mean square error; MAE: mean absolute error.

Figure 5. Difference between before and after data preprocessing.

high-LAT outliers, resulting in a more concentrated data distribution. Similarly, the LON distribution shown in Figure 5B demonstrates that after cleaning, the standard deviation decreased by 9.2%, and the data range narrowed by 8.6%. The chart exhibits a more regular distribution of LON data post-cleaning, with extreme values effectively eliminated. Following data preprocessing, the dataset is partitioned into training, testing, and validation sets in a ratio of 7:1:2.

3.4. Comparison between Crossformer and other models

To more comprehensively demonstrate and compare the performance of different prediction models in ship trajectory prediction, this study selected historical trajectory data from Case 1 and applied various prediction models. Subsequently, the actual trajectory (blue) was presented alongside the prediction results from Crossformer, GRU, LSTM, and TGCN models in Figure 6 for comparison, as illustrated. To ensure a fair comparison, we configured the baseline models with comparable complexity. Both the GRU and LSTM models were set with a hidden dimension of 512 and three layers, while the TGCN model used 64 hidden units across two graph convolutional layers. All models followed the same data preprocessing steps, shared identical input and output dimensions, and were trained using consistent procedures. As evident from Figure 6, under identical parameters and input data conditions, significant differences exist among the models in predicting LAT-LON variations and ship heading changes.

We can clearly observe from Figure 6 that the LSTM model's prediction performance for Case 1 is relatively poor. Although this model can capture the overall trend of ship heading changes, substantial deviations persist between the predicted trajectory and the actual trajectory when the ship undergoes sharp turns. In contrast, the GRU and TGCN models demonstrate better prediction accuracy, both achieving closer alignment with the actual trajectory across most segments. However, when the ship executes substantial turns, the predicted trajectories from GRU and TGCN still exhibit a certain degree of deviation.
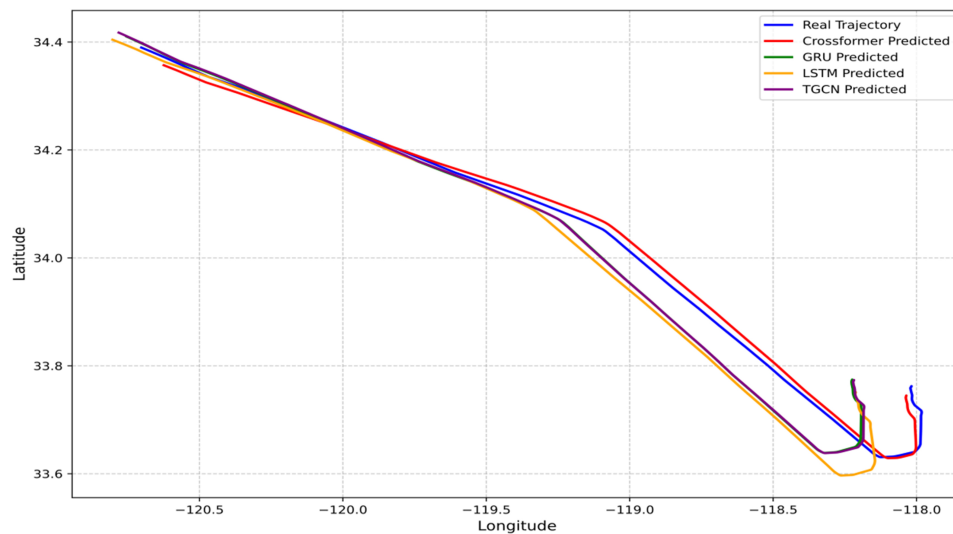
Figure 6. Multiple model results for predicting Case 1 trajectories.

The Crossformer model proposed in this study shows a high degree of congruence between its prediction results for Case 1 and the actual trajectory. It not only accurately captures the overall trend of ship heading changes but also maintains high prediction accuracy during turning phases. This superior performance is attributed to better integration of temporal and spatial factors during the modeling process, enabling Crossformer to ultimately demonstrate the highest prediction accuracy among all models.

As shown in Figure 6, the trajectory comparisons highlight key differences in how the models handle complex navigational scenarios. The large deviations in LSTM predictions, especially during sharp turns, are likely due to the vanishing gradient problem common in traditional RNNs. This limitation makes it difficult for them to capture the long-term spatial-temporal dependencies essential for maintaining trajectory continuity. In contrast, the Crossformer model performs noticeably better, largely thanks to its DSW embedding mechanism, which helps preserve the intricate relationships between LAT and LON during complex maneuvers. Its attention-based architecture allows the model to dynamically focus on relevant segments of past trajectories, particularly when sudden directional changes occur. These results are consistent with recent advances in transformer-based time series prediction, where attention mechanisms have repeatedly shown advantages over recurrent models in handling long-range dependencies.

This study also selects historical trajectory data from Case 2 for prediction and conducted comparative analysis using the proposed ship trajectory prediction models. The results illustrated in Figure 7 indicate that while different models can generally simulate the movement trajectory of Case 2 effectively, there are notable differences in their prediction accuracy. It is evident that the LSTM model underperforms in terms of prediction accuracy, showing significant deviation from the actual trajectory. Although this model can capture the general trend of trajectory changes, there remains considerable room for improvement regarding key point localization and local trajectory fitting. The TGCN model demonstrates good prediction performance in the initial stages but exhibits noticeable deviation in later prediction phases, suggesting that the model still requires further optimization for longer time spans or changes in data distribution. The GRU model performs relatively consistently overall, accurately depicting the ship's movement trends during most time periods; however, comparative analysis reveals that prediction effectiveness of GRU is slightly inferior to that of Crossformer. The Crossformer model proves most effective in Case 2 trajectory prediction, showing the highest degree of overlap between its predicted
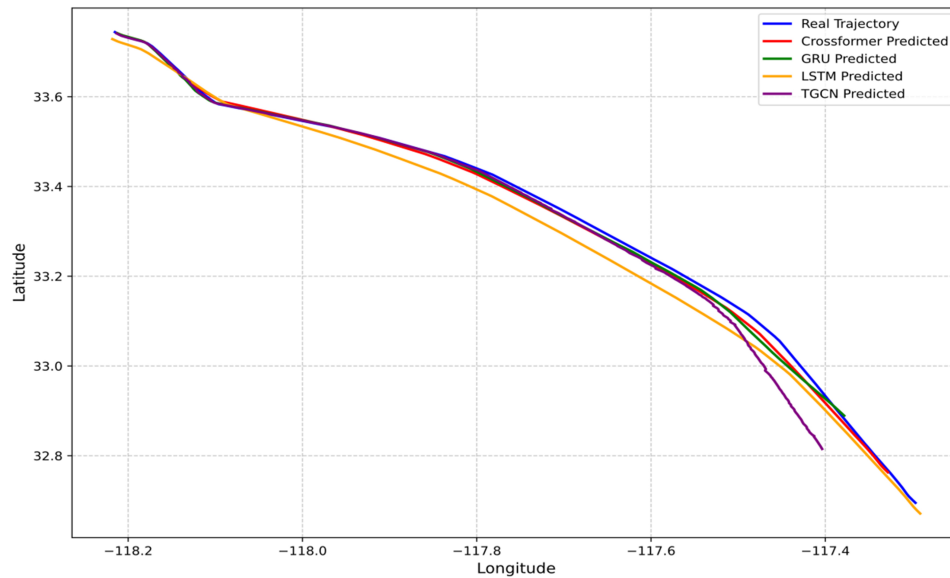
Figure 7. Multiple model results for predicting Case 2 trajectories.

trajectory and the actual trajectory. This indicates that the model can more thoroughly extract temporal and spatial information when processing relatively smooth ship tracks with less dramatic turns, thereby achieving more accurate prediction results.

Meanwhile, compared to the trajectory in Case 1, Case 2 does not exhibit significant directional changes throughout the entire prediction interval, with an overall smoother trajectory in Figure 7. All models are able to adequately reflect its navigational trend, with differences primarily manifested in prediction granularity and the ability to capture local inflection points. Based on the comparative analysis of trajectory prediction results for Case 1, it can be concluded that Crossformer provides more accurate prediction results when processing ship trajectory prediction (particularly in complex scenarios involving frequent heading changes). Crossformer maintained high prediction accuracy and stability in the Case 2 prediction task as well.

As shown in Table 2, the Crossformer model achieved the lowest values across the ADE, MSE, RMSE, and MAE metrics, with values of 2.35×10^{-2}, 7.00×10^{-4}, 2.58×10^{-2}, and 2.35×10^{-2}, respectively, indicating that this model significantly outperforms baseline models such as GRU, LSTM, and TGCN in terms of overall trajectory prediction accuracy. Notably, for the FDE metric, the GRU model exhibited the minimum error (9.20×10^{-3}), slightly lower than the Crossformer model (3.55×10^{-2}), suggesting that GRU possesses certain advantages in endpoint prediction accuracy for short-term forecasting. However, considering the overall prediction error, Crossformer demonstrated superior performance across multiple metrics, exhibiting stronger global predictive capability.

These results indicate that Crossformer can more precisely capture trajectory evolution patterns, reduce prediction errors, and particularly excel in overall trajectory accuracy control compared to traditional sequence modeling methods (such as GRU and LSTM) and GCN.

To more intuitively demonstrate the differences between models, this research introduced comparative diagrams of ADE and FDE for both LON and LAT, as shown in Figure 8. In Figure 8, the dimensional

Table 2. Predictive performance results of different models for Case 1

Statistical indicator	Proposed Crossformer	GRU	LSTM	TGCN
ADE	**2.35×10^{-2}**	9.38×10^{-2}	2.53×10^{-2}	1.00×10^{-1}
FDE	3.55×10^{-2}	**9.20×10^{-3}**	5.58×10^{-2}	5.32×10^{-2}
MSE	**7.00×10^{-4}**	9.10×10^{-3}	8.00×10^{-4}	1.02×10^{-2}
RMSE	**2.58×10^{-2}**	9.55×10^{-2}	2.85×10^{-2}	1.01×10^{-1}
MAE	**2.35×10^{-2}**	9.38×10^{-2}	2.53×10^{-2}	1.01×10^{-1}

The values labeled with bold fonts demonstrate the best results. GRU: Gated recurrent unit; LSTM: long short-term memory; TGCN: temporal graph convolutional network; ADE: average Euclidean distance error; FDE: final Euclidean distance error; MSE: mean square error; RMSE: root mean square error; MAE: mean absolute error.

Figure 8. Comparison of model results based on various indicators in Case 1.

analysis offers valuable insights into how the models perform across spatial coordinates. Notably, Crossformer achieves much lower ADE values in both LON (3.36×10^{-2}) and LAT (1.34×10^{-2}), reflecting a well-balanced predictive accuracy across these geographic dimensions. Such balanced performance is critical in maritime navigation, where safety depends on precise positioning in both directions. The considerable gap between Crossformer and baseline models - often exceeding a 60% reduction in error - highlights the effectiveness of its HED architecture in capturing the intrinsic correlation between LAT and LON during vessel movement. In contrast, traditional sequential models struggle particularly with LON predictions, revealing challenges in modeling the complex interplay between temporal ship dynamics and spatial coordinate changes - an issue that the proposed TSA mechanism successfully overcomes.

This performance is consistent with the aforementioned overall metric comparison [Table 2], leading to the conclusion that the Crossformer model demonstrates significant advantages in processing complex spatiotemporal data. Its dimensional-segmented embedding technique and TSA mechanism effectively extract data features at both temporal and spatial levels, markedly reducing prediction errors. In contrast,

while GRU, LSTM, and TGCN exhibit certain accuracy in local predictions, they remain somewhat inadequate in multi-dimensional error control, struggling to comprehensively accommodate prediction precision in both longitudinal and latitudinal directions. Through repeated verification and cross-comparison in the above experiments, we can conclude that the Crossformer model demonstrates superior stability and robustness in predictions across different dimensions.

The quantitative results highlight a notable improvement over existing ship trajectory prediction methods. Crossformer achieves over a 60% reduction in error compared to baseline models, marking a significant step forward in prediction accuracy with clear benefits for maritime safety. Its consistently low error rates across various metrics demonstrate a robustness that makes it well-suited for supporting real-time decision-making in autonomous navigation systems. When compared with recent approaches such as GAT-LSTM proposed by Zhao *et al.*[12] and MP-LSTM used by Gao *et al.*[15], Crossformer stands out by effectively combining dimension-wise processing with hierarchical temporal modeling. This advantage is especially evident given the challenging nature of the test cases - ranging from complex turning maneuvers in Case 1 to smoother navigation patterns in Case 2 - which together reflect the diverse conditions encountered in real-world maritime navigation.

Table 3 presents comparative results of performance evaluation metrics for different models on Case 2. As shown in Table 3, the Crossformer model achieved superior results across all evaluation metrics, particularly in overall trajectory error (ADE, MAE) and MSE, where the error reduction magnitude reached over 70%. Although in terms of the FDE metric, Crossformer's values were comparable to those of GRU and TGCN models, its advantages were more pronounced when compared to the LSTM model.

As shown in Figure 9, after comparing the MSE for both LON and LAT dimensions, it is evident that the Crossformer model achieved extremely small error values in both metrics, significantly outperforming comparative models such as GRU, LSTM, and TGCN. In terms of LON prediction, Crossformer's MSE is merely 4×10^{-4}, although the LSTM model exhibited an MSE of 1×10^{-4}; however, regarding LAT prediction, LSTM model performed notably worse than the Crossformer model. Crossformer maintained a low MSE level (6×10^{-4}), further demonstrating its advantages in capturing spatiotemporal dynamic characteristics of ship trajectories. Synthesizing the various comparative methods discussed previously, we can readily conclude that the model proposed in this paper outperforms other comparative models and can achieve high-precision ship trajectory prediction.
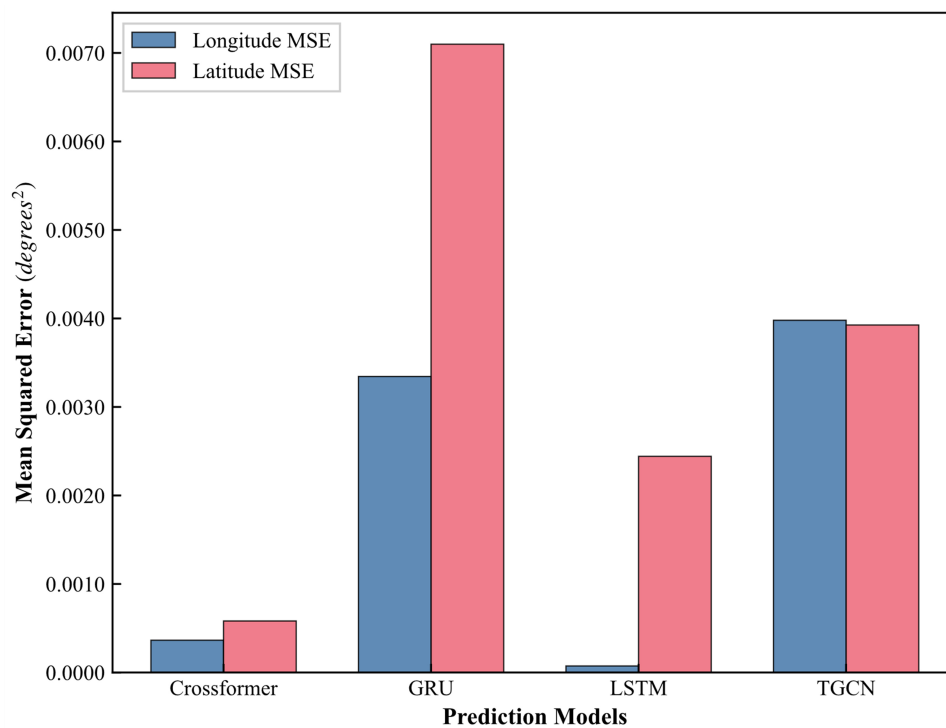
4. CONCLUSION

This research proposes a ship trajectory prediction method based on the Crossformer model. The Crossformer model features a triple architectural framework. It employs DSW embedding technology to segment ship AIS data, enabling it to simultaneously capture spatial features of LON and LAT variations; the TSA mechanism establishes data correlations across temporal and dimensional planes, effectively extracting complex dependencies in ship movement; the HED structure achieves precise control of navigational situations across various temporal spans through multi-scale modeling. Experimental results demonstrate that the Crossformer model exhibits significantly superior performance in ship trajectory prediction tasks. With ADE of 2.35×10^{-2}, MSE of 7.00×10^{-4}, RMSE of 2.58×10^{-2}, and MAE of 2.35×10^{-2} in Case 1, and even better performance with ADE of 1.64×10^{-2}, MSE of 4.00×10^{-4}, RMSE of 2.06×10^{-2}, and MAE of 1.64×10^{-2} in Case 2. Compared to traditional models such as GRU, LSTM, and TGCN, the Crossformer model reduced average prediction errors by over 60% in Case 1 and up to 70% in Case 2, demonstrating its superior capability in capturing spatiotemporal dependencies in ship movement patterns.

Table 3. Predictive performance results of different models for Case 2

Statistical indicator	Proposed Crossformer	GRU	LSTM	TGCN
ADE	**1.64 × 10^{-2}**	6.05 × 10^{-2}	9.59 × 10^{-2}	5.65 × 10^{-2}
FDE	**2.50 × 10^{-3}**	2.50 × 10^{-3}	4.44 × 10^{-2}	2.60 × 10^{-3}
MSE	**4.00 × 10^{-4}**	4.80 × 10^{-3}	9.40 × 10^{-3}	4.90 × 10^{-3}
RMSE	**2.06 × 10^{-2}**	6.95 × 10^{-2}	9.70 × 10^{-2}	6.24 × 10^{-2}
MAE	**1.64 × 10^{-2}**	6.05 × 10^{-2}	9.59 × 10^{-2}	5.65 × 10^{-2}

The values labeled with bold fonts demonstrate the best results. GRU: Gated recurrent unit; LSTM: long short-term memory; TGCN: temporal graph convolutional network; ADE: average Euclidean distance error; FDE: final Euclidean distance error; MSE: mean square error; RMSE: root mean square error; MAE: mean absolute error.

Figure 9. MSE comparison across models for Case 2 trajectory prediction. MSE: Mean square error.

To enhance model performance, the following directions could be further explored. The current model primarily focuses on single-ship trajectory prediction and has not fully considered interactive influences between ships or predictive capabilities under special scenarios such as extreme weather conditions and complex channels. Furthermore, the model's real-time computational efficiency and operational resource consumption require further optimization to meet deployment requirements in actual maritime monitoring systems. Simultaneously, more powerful attention mechanisms could be introduced based on the existing architecture to further improve long-term trajectory prediction capabilities; additionally, optimization combinations of network layer structures could be improved, such as adjusting encoder-decoder layer counts or introducing specific functional layers to enhance the model's expressive capacity. The Crossformer could also be combined with advanced technologies to further enhance its adaptability in dynamically complex environments; extending its application to multi-ship collaborative prediction holds promise for bringing more innovative solutions to smart port and intelligent shipping domains.

DECLARATIONS

Authors' contributions
Writing - review and editing, supervision, formal analysis, conceptualization: Chen, X.
Writing - original draft, visualization, validation, methodology, investigation: Wu, P.
Formal analysis, conceptualization: Wu, Y.
Writing - review, data curation, conceptualization: Aboud, L.
Writing - review and editing, performed data acquisition: Postolache, O.
Supervision, methodology, conceptualization: Wang, Z.

Availability of data and materials
The AIS data used in this study are publicly available from the MarineCadastre website (https://marinecadastre.gov/ais/). Specifically, the dataset corresponding to March 2020, covering the coastal waters of the Western Seaboard (33°N-34°N, 118°W-120°W), was utilized in this research. These data can be freely accessed and downloaded by any user following the terms of use specified on the MarineCadastre website.

Conflicts of interest
Chen, X. is an Editorial Board Member of *Intelligence & Robotics* and serves as the Guest Editor for the Special Issue titled "*Intelligent, Safe, and Green Shipping-oriented Maritime Data Exploitation and Knowledge Discovery*". Chen, X. was not involved in any part of the editorial process, including the selection of reviewers, handling of the manuscript, or the final decision-making. Wang, Z. is affiliated with Shanghai Ship and Shipping Research Institute Co., Ltd. The other authors declare no conflicts of interest.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Copyright

REFERENCES

1. Pallotta, G.; Vespe, M.; Bryan, K. Vessel pattern knowledge discovery from AIS data: a framework for anomaly detection and route prediction. *Entropy* **2013**, *15*, 2218-45. DOI
2. Short, D.; Lei, T.; Luo, C.; Carruth, D. W.; Bi, Z. A bio-inspired algorithm in image-based path planning and localization using visual features and maps. *Intell. Robot.* **2023**, *3*, 222-41. DOI
3. Murray, B.; Perera, L. P. A dual linear autoencoder approach for vessel trajectory prediction using historical AIS data. *Ocean. Eng.* **2020**, *209*, 107478. DOI
4. Dalsnes, B. R.; Hexeberg, S.; Flåten, A. L.; Eriksen, B. O. H.; Brekke, E. F. The neighbor course distribution method with gaussian mixture models for AIS-based vessel trajectory prediction. In *2018 21st International Conference on Information Fusion (FUSION)*, Cambridge, UK. Jul 10-13, 2018. IEEE; 2018. pp. 580-7. DOI
5. Zhen, R.; Jin, Y.; Hu, Q.; Shao, Z.; Nikitakos, N. Maritime anomaly detection within coastal waters based on vessel trajectory clustering and Naïve Bayes Classifier. *J. Navigation.* **2017**, *70*, 648-70. DOI
6. Perera, L. P.; Soares, C. G. Ocean vessel trajectory estimation and prediction based on extended Kalman filter. In *The Second*

International Conference on Adaptive and Self-Adaptive Systems and Applications. 2010. https://personales.upv.es/thinkmind/dl/conferences/adaptive/adaptive_2010/adaptive_2010_1_30_40029.pdf. (accessed 24 Jun 2025)

7.   Liu, J.; Shi, G.; Zhu, K. Vessel trajectory prediction model based on AIS sensor data and adaptive chaos differential evolution support vector regression (ACDE-SVR). *Appl. Sci.* **2019**, *9*, 2983. DOI

8.   Suo, Y.; Chen, W.; Claramunt, C.; Yang, S. A ship trajectory prediction framework based on a recurrent neural network. *Sensors* **2020**, *20*, 5133. DOI PubMed PMC

9.   Borkowski, P. The ship movement trajectory prediction algorithm using navigational data fusion. *Sensors* **2017**, *17*, 1432. DOI PubMed PMC

10.   Gan, S.; Liang, S.; Li, K.; Deng, J.; Cheng, T. Long-term ship speed prediction for intelligent traffic signaling. *IEEE. Trans. Intell. Transport. Syst.* **2017**, *18*, 82-91. DOI

11.   Park, J.; Jeong, J.; Park, Y. Ship trajectory prediction based on Bi-LSTM using spectral-clustered AIS data. *JMSE.* **2021**, *9*, 1037. DOI

12.   Zhao, J.; Yan, Z.; Zhou, Z.; Chen, X.; Wu, B.; Wang, S. A ship trajectory prediction method based on GAT and LSTM. *Ocean. Eng.* **2023**, *289*, 116159. DOI

13.   Johansen, T. A.; Perez, T.; Cristofaro, A. Ship collision avoidance and COLREGS compliance using simulation-based control behavior selection with predictive hazard assessment. *IEEE. Trans. Intell. Transport. Syst.* **2016**, *17*, 3407-22. DOI

14.   Zhou, H.; Chen, Y.; Zhang, S. Ship trajectory prediction based on BP neural network. *J. Artif. Intell.* **2019**, *1*, 29-36. DOI

15.   Gao, D.; Zhu, Y.; Zhang, J.; He, Y.; Yan, K.; Yan, B. A novel MP-LSTM method for ship trajectory prediction based on AIS data. *Ocean. Eng.* **2021**, *228*, 108956. DOI

16.   Bai, S.; Kolter, J. Z.; Koltun, V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv* 2018, arXiv:1803.01271. https://doi.org/10.48550/arXiv.1803.01271. (accessed 24 Jun 2025)

17.   Alizadeh, D.; Alesheikh, A. A.; Sharif, M. Vessel trajectory prediction using historical automatic identification system data. *J. Navigation.* **2021**, *74*, 156-74. DOI

18.   Zhang, X.; Liu, J.; Gong, P.; Chen, C.; Han, B.; Wu, Z. Trajectory prediction of seagoing ships in dynamic traffic scenes via a gated spatio-temporal graph aggregation network. *Ocean. Eng.* **2023**, *287*, 115886. DOI

19.   Jiang, J.; Zuo, Y.; Xiao, Y.; Zhang, W.; Li, T. STMGF-Net: a spatiotemporal multi-graph fusion network for vessel trajectory forecasting in intelligent maritime navigation. *IEEE. Trans. Intell. Transport. Syst.* **2024**, *25*, 21367-79. DOI

20.   Wang, S.; Li, Y.; Xing, H.; Zhang, Z. Vessel trajectory prediction based on spatio-temporal graph convolutional network for complex and crowded sea areas. *Ocean. Eng.* **2024**, *298*, 117232. DOI

21.   Syed, M. A. B.; Ahmed, I. A CNN-LSTM architecture for marine vessel track association using automatic identification system (AIS) data. *Sensors* **2023**, *23*, 6400. DOI PubMed PMC

22.   Chen, X.; Ling, J.; Yang, Y.; et al. Ship trajectory reconstruction from AIS sensory data via data quality control and prediction. *Math. Probl. Eng.* **2020**, *2020*, 1-9. DOI

23.   Zhao, J.; Yan, Z.; Chen, X.; Han, B.; Wu, S.; Ke, R. k-GCN-LSTM: a k-hop graph convolutional network and long-short-term memory for ship speed prediction. *Physica. A.* **2022**, *606*, 128107. DOI

24.   Zhang, D.; Li, J.; Wu, Q.; Liu, X.; Chu, X.; He, W. Enhance the AIS data availability by screening and interpolation. *In. 2017. 4th. International. Conference. on. Transportation. Information. and. Safety. (ICTIS).* 2017. pp. 981-6. DOI

25.   Zhang, Y.; Yan, J. Crossformer: transformer utilizing cross-dimension dependency for multivariate time series forecasting. In *The eleventh international conference on learning representations*. 2023. https://openreview.net/forum?id=vSVLM2j9eie. (accessed 24 Jun 2025).

26.   Cho, K.; van Merrienboer, B.; Bahdanau, D.; Bengio, Y. On the properties of neural machine translation: encoder-decoder approaches. *arXiv* 2014, arXiv:1409.1259. https://doi.org/10.48550/arXiv.1409.1259. (accessed 24 Jun 2025)

27.   Cho, K.; van Merrienboer, B.; Gulcehre, C.; et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* 2014, arXiv:1406.1078. https://doi.org/10.48550/arXiv.1406.1078. (accessed 24 Jun 2025)

28.   Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural. Comput.* **1997**, *9*, 1735-80. DOI

29.   Zhao, L.; Song, Y.; Zhang, C.; et al. T-GCN: a temporal graph convolutional network for traffic prediction. *IEEE. Trans. Intell. Transport. Syst.* **2020**, *21*, 3848-58. DOI

30.   Feroze, W.; Shahid, M.; Cheng, S.; et al. Enhancing temporal commonsense understanding using disentangled attention-based method with a hybrid data framework. *Intell. Robot.* **2025**, *5*, 228-47. DOI