

Received November 4, 2020, accepted November 22, 2020, date of publication December 2, 2020, date of current version December 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3041762

# ST-Seq2Seq: A Spatio-Temporal Feature-Optimized Seq2Seq Model for Short-Term Vessel Trajectory Prediction

LAN YOU<sup>1</sup>, SIYU XIAO<sup>1</sup>, QINGXI PENG<sup>2</sup>, CHRISTOPHE CLARAMUNT<sup>3</sup>, XUEWEI HAN<sup>1</sup>, ZHENGYI GUAN<sup>1</sup>, AND JIAHE ZHANG<sup>1</sup>

<sup>1</sup>School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China

<sup>2</sup>School of Computer Science, Wuhan Donghu University, Wuhan 430212, China

<sup>3</sup>Naval Academy Research Institute, 29240 Brest, France

Corresponding author: Qingxi Peng (pqx@whu.edu.cn)

This work was supported in part by the Hubei Province Natural Science Foundation Item under Grant 2019CFB757, in part by the Key Project of Science and Technology Research Program of Hubei Provincial Education Department under Grant D20201006, and in part by the Open Fund of National Engineering Research Center for Water Transport Safety under Grant A2019011.

**ABSTRACT** Deep learning provides appropriate mechanisms to predict vessel trajectories for safer and efficient shipping, but still existing models are mainly oriented to longer-term prediction trends and do not fully support real time navigation needs. While most recent works have been largely exploiting Automatic Identification System (AIS), the complete semantics of these data haven't so far fully exploited. The research presented in this paper introduced an extended sequence-to-sequence model using AIS data. A Gated Recurrent Unit (GRU) network encodes historical spatio-temporal sequences as a context vector, which not only preserves the sequential relationships among trajectory locations, but also alleviates the gradient descent problem. The GRU network acts as a decoder, outputting target trajectory location sequences. Real AIS data from the Chongqing and Wuhan sections of the Yangzi River were selected as typical experimental areas for evaluation purposes. The proposed ST-Seq2Seq model has been tested against the LSTM-RNN and GRU-RNN baseline models for short term trajectory prediction experiments. A 10-minute historical trajectory sequence was used to predict the trajectory sequence for the next five minutes. Overall, the findings show that LSTM and GRU networks, while applying a recursive method to predict a sequence of continuous trajectory points, when the number of predicted trajectory points increases accuracy decreases. Conversely, the extended sequence-to-sequence model shows satisfactory stability on different ship channels.

**INDEX TERMS** Sequence-to-sequence (Seq2Seq), recurrent neural network, spatio-temporal, AIS, vessel trajectory prediction.

## I. INTRODUCTION

Over the last decades, maritime traffic has widely increased as a result of higher demand on global trade, likely causing channel congestions, collisions risks and environmental threats at sea [1]. Safety and security are then becoming crucial concerns of maritime navigation due to this worldwide exponential growth of maritime traffic [2], and where maritime surveillance data should be fully exploited to support higher levels of situational awareness. Consequently, the past few years have seen a rapid increase in research and development of information-oriented infrastructures and systems addressing many aspects of data integration, analysis,

visualization and diffusion of data related to movement at sea. Automatic Identification System (AIS) technology, as a navigation system, provides a vast amount of near-real time maritime positioning information [3], linking a unique ship identifier, the Maritime Mobile Service Identity (MMSI) to longitude and latitude, speed, course, and other contextual information [4]. Large-scale AIS data can be used to track most vessels, to mine extract ship navigation patterns. Time series can be for also applied to compare the patterns that emerge from different maritime trajectories [5]. This shipping knowledge can be exploited to predict ship trajectories for collision avoidance, trajectory monitoring, analysis and prediction [6].

Amongst many issues to deal with, trajectory prediction is a hot issue as this should be dealt with in real-time, high

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

accuracy and computational efficiency. Different attempts have been so far made from various domains to predict human, flight, and vehicle trajectories to mention a few examples [3], [7]–[9]. Trajectory data are a special kind of time series data combining the spatial and temporal dimensions [10]. Kalman filters have been commonly applied to waypoints for predicting ship trajectories. With the development of deep learning, methods such as recurrent neural networks (RNNs) and variational autoencoders have demonstrated effective generalization performance in trajectory prediction tasks [11].

The research introduced in this paper introduces an extended spatio-temporal feature optimized Seq2Seq Model whose objective is to predict short term vessel trajectories. The approach is based on incoming AIS data and the aim is to provide a prediction mechanism to mainly avoid vessel collisions. The main peculiarity of the approach is that it takes into account historical navigation data and contextual data to improve vessel navigation predictions while also considering the irregular sampling of AIS data. Overall the principles and contribution of our approach are four-fold:

- (1) Design of a Seq2Seq framework based on spatio-temporal data that automatically predicts a trajectory sequence. A Gated Recurrent Unit (GRU) is used as the neural unit, effectively alleviating the problem occurring when some valuable states are “forgotten”.

- (2) While most related works apply long-term prediction models, the peculiarity of our approach is that it is oriented towards short-term prediction. This is particularly significant for maritime navigation and ship collision warning and avoidance.

- (3) Trajectory tracking points are temporally processed as time intervals, and the position points are processed as relative positions in order to improve the prediction accuracy.

- (4) The demonstration of the relevance of the proposed approach is experimented on real datasets from the Wuhan and Chongqing waterways

The remainder of the paper is organized as follows. Section II briefly reviews current methods applied to vessel trajectory prediction. Section III provides the background of our modelling approach. The proposed method is detailed in Section IV while Section V develops the experiments. Finally, Section VI outlines the findings and draws a few perspectives for future work.

## II. RELATED WORK

Significant work has been done from the maritime engineering and scientific community on extracting valuable information from AIS data. This section mainly reviews AIS-based trajectory forecasting approaches. Current modeling and prediction vessel trajectory methods can be categorized into two broad classes according to their underlying implementation mechanisms: physical model based methods or learning model based methods [12].

### A. PHYSICAL MODEL-BASED METHODS

Early methods of vessel trajectory prediction rely on a physical model of the vessel movement and are mainly based on curvilinear models [13], [14], lateral models [4], [15] and ship model [16], [17]. Physical ship motion is represented using a conjunction of mathematical equations and laws that consider all possible influencing factors such as mass, size, inertia, and mass center. The accuracy of such methods relies on ideal or very precise representation of the environmental and state assumptions, which are difficult to attain in most real-world vessel trajectory prediction scenarios.

### B. LEARNING MODEL-BASED METHODS

Learning-based methods model ship motion from previous and current trajectories using historical and real-time AIS data, implicitly integrating all possible influencing factors. For instance, Kalman filter approaches use dynamic information from the vessel target, and remove noise to get a prediction of the next vessel target locations [18]. Siebert et al used EKF to track vessel trajectories [19], and allows for failure detection based on residual monitoring. In [12], an Extended Kalman Filter is proposed as an adaptive filter algorithm for the estimation of position, velocity, and acceleration that are in turn, used to predict maneuvers for ocean vessel trajectory. The advantage of this extended Kalman filter is that it is a well-studied classical method, so there are many successful applications that can be exploited. However, the extended Kalman filter is generally not globally optimal [7].

Deep learning is another popular approach for trajectory prediction due to its powerful ability to fit complex functions [20]. [21] [22] combined recurrent neural networks with latent variable modeling to address the peculiarities of AIS data streams: massive amount of streaming data, noisy data and irregular time sampling.

As detailed hereafter, we develop a seq2seq framework based on spatio-temporal data to predict vessel trajectory and demonstrate its relevance from experiments on a real AIS dataset on a regional scale. Most of the existing researches are long-term predictions, and there are few short-term predictions, and most of them are iterative predictions of trajectory points. The more iterations, the greater the error. The Seq2Seq model proposed in this paper solves the problem of short-term prediction of ship trajectory. At the same time, the model in this paper is a continuous trajectory point sequence prediction, which can reduce the error caused by iterative prediction.

## III. BACKGROUND KNOWLEDGE

Trajectory prediction is a difficult issue to deal with and different attempts have been made in various fields, including the prediction of human, animal, aircraft, vehicle and vessel trajectories. A trajectory can be roughly considered a spatial time series, where the spatial as well as the temporal

dimension are the main dimensions to represent plus the semantic one for some specific application-oriented cases. This section first briefly introduces the main principles behind the trajectory data model and the Seq2Seq model [23], [24].

### A. TRAJECTORY PATTERN

A trajectory represents the path tracked by an object moving in space over time. Trajectory data are usually collected by installing sensors on moving objects. A sensor periodically transmits location data of a considered object, such as in the case of AIS that periodically send ship location information. Based on different trajectory data mining techniques, trajectory patterns are generally divided into four types: common moving pattern, trajectory clustering, trajectory mining sequence pattern, and trajectory mining cycle pattern [18]. A trajectory mining sequential pattern represents a sequence of position points from a moving target according to some regular time intervals. This pattern is usually used for next location prediction for either trajectories in free or constrained spaces.

#### 1) TRAJECTORY SEQUENCE PATTERNS IN FREE SPACE

Trajectory sequence patterns in a free space are generally expressed as a spatio-temporal sequence (ST-sequence). A spatio-temporal sequence is denoted as a sequence of time and spatial-stamped tuples given as follows:

$$S = \langle (p_0, q_0, t_0), \dots, (p_k, q_k, t_k) \rangle \quad (1)$$

where  $t_i (i = 0 \dots k)$  represents a series of timestamps with  $\forall 0 \leq i < k, t_i < t_{i+1}$ , and  $(p_i, q_i)$  denotes a location.

#### 2) TRAJECTORY SEQUENCE PATTERNS IN ROAD GRIDS

When the location information in the track sequence mode needs to be represented by the road grid, first use the map matching algorithm to map each track to the road grid, and then the track is represented by a series of road segment (IDS).

### B. Seq2Seq MODEL

The sequence to sequence model (Seq2Seq) has been widely used in processing tasks of variable length input and output sequences, including speech recognition, machine translation and so on [25]–[27]. Its core idea is to map a variable length input sequence to variable length output sequence using cyclic neural network. Cho *et al.* introduced a neural network structure so-called RNN Encoder-Decoder [27]. Subsequently, the Google team put forward a Sequence-to-Sequence model similar to RNN Encoder-Decoder structure for machine translation [28]. Similar solutions have been then proposed in the literature, and the Seq2Seq model has also been generated, like building Emotional Conversation Systems Using Multi-task Seq2Seq Learning. The model consists of two cyclic neural networks, the structure is shown in Figure 1 [29], [30].

As shown in Figure 1, the Seq2Seq model consists of two parts: an encoder and a decoder. The encoder reads every

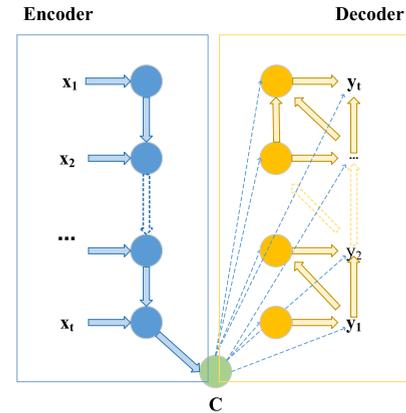


FIGURE 1. Seq2seq model.

symbol in the input sequence in sequence. When reading each symbol, the hidden state will change accordingly. Finally, a semantic vector  $C$  is formed as the input of the decoder. The decoder is also a  $\{y_1$ , which generates the output sequence  $y_2, y_t\}$ , which is different from the calculation of the hidden state in a common RNN under the condition of a given hidden state, and predicts the next symbol. The calculation of the hidden state in the decoder takes into account the semantic vector  $C$ , its formula is given as follows:

$$h_t = f(h_{t-1}, y_{t-1}, c) \quad (2)$$

$f$  is a nonlinear activation function, and  $f$  can be a tanh or sigmoid function.

The goal of joint training of two parts of Seq2Seq model is to maximize the conditional likelihood function,  $(x_n, y_n)$  is the sequence of the corresponding input and output..

$$\max_{\theta} \frac{1}{N} \sum_{n=1}^N \log p_{\theta}(y_n | x_n) \quad (3)$$

$$V = \{v_1, v_2, \dots, v_t\}$$

$\theta$  represents model parameters,  $N$  represents the number of samples in the training set,  $(x_n, y_n)$  is the sequence of the corresponding input and output.

### IV. THE EXTENDED SEQUENCE-TO-SEQUENCE MODEL FOR SHORT-TERM VESSEL TRAJECTORY PREDICTION

The ship AIS provides a large amount of near real-time water monitoring data. The main idea behind our approach is to design and apply a deep learning model to mine pre-processed tracking AIS data in order to effectively understand the behavior of the ship and improve the efficiency of water safety supervision. Specifically, if the predicted trajectories of two ships cross, there is indeed a risk of collision. The time granularity of such trajectory prediction tasks is generally within the range of 15 minutes. This is the main motivation behind our approach that considers short-term (5-15 minutes) trajectory sequences prediction. There are two complex factors that are challenging in the prediction task: irregular time sampling of the track data and accuracy of the track sequence prediction. In order to solve

these problems, the main features of our extended sequence-to-sequence model is to propose a short-term AIS trajectory sequence prediction model.

### A. AIS DATA PREPROCESSING

#### 1) TRAJECTORY SEGMENTATION OF RAW DATA

The way AIS broadcasts and receives information is automatic and continuous. An original AIS data stream is formatted according to the receiving time of the AIS message. Therefore, data preprocessing divides the original AIS data stream into ship trajectories. This specific process is done in two steps

1) Separate the trajectory data of different vessels. The MMSI number is the unique identifier of the ship, and the trajectory of different ships is separated by the MMSI number information of the AIS data;

2) Separate the trajectory data of the same vessel. In navigable waters, due to the large number of ships and the constraints of the AIS working mechanism, the network communication is blocked, For example, the data that should have been received 1 second ago was received after a delay of 1 second by the network. This 1 second is the time interval. Thus, the AIS cannot reserve or listen to the idle time slots, which causes the AIS information to be delayed, and the trajectory data of a given ship will appear larger interval. Overall, a continuous ship trajectory is separated based on the timestamp information of the AIS data.

#### 2) INPUT TRAJECTORY SEQUENCE STRUCTURE

The model input is a sequence of historical trajectories of a given ship, and the output is a sequence of predicted ship trajectories. A ship trajectory  $V$  is denoted as a time series of a state vector  $v_i$  at time  $t$ :

$$V = \{v_1, v_2, \dots, v_t\} \quad (4)$$

The state vector  $v_t$  at time  $t$  is the feature information separated from the AIS data. The trajectory features include position features, speed and heading, etc., expressed as follows:

$$v_t = [p_t, q_t, \bar{p}_t, \bar{q}_t, z_t]^T \quad (5)$$

where,  $p_t$  and  $q_t$  denote the relative longitude and relative latitude, respectively,  $\bar{p}_t$  denotes the Speed Over Ground (SOG),  $\bar{q}_t$  denotes the Course Over Ground (COG),  $z_t$  denotes the time interval. The temporal and spatial dimensions of a ship trajectory are qualitatively represented by time intervals and relative positions, respectively. This being a major difference with most current methods where ship positions are quantitatively timestamped and referenced by latitude and longitude coordinates. Therefore, in order to solve the problem of the irregular frequency sampling of AIS data we resample trajectory data using relative time and position features. This optimizes the data structure of the prediction model. For example,  $v_t = [121.070578, 31.756207, 12.9, 115.8, 0.01]$ , means this ship is located in 121.070578, 31.756207, and the COG is 12.9, the SOG is 115.8, the time interval is 0.01 second.

#### 3) NORMALIZATION

RNNs use the gradient descent method to solve the optimization problem during the training process. Therefore, the input data is normalized, and the input data is mapped into the range of  $[0, 1]$ , thereby speeding up the solution of the gradient descent and improving the model convergence rate.

The Min-Max Normalization method normalizes the ground heading, and the characteristics of the three tracking points for the ground speed and time interval.

$$X^* = \frac{X - \min}{\max - \min} \quad (6)$$

where  $\max$  denotes the maximum value found in the sample data,  $\min$  is the minimum value in the sample data, and  $X$  is the original data, where  $X^*$  is the normalized data. For example, the maximum speed is 21.4, the minimum speed is 0, and the average is 8.41.

### B. IMPROVED Seq2Seq TRAJECTORY SEQUENCE PREDICTION MODEL

The proposed Seq2Seq trajectory sequence prediction model combines  $t$  spatiotemporal sequence data and the Seq2Seq model. As shown in Figure 2, the ship AIS trajectory real-time prediction model based on Seq2Seq is mainly composed of two RNNs modules. The model design fully considers the influence of temporal and spatial features on the trajectory prediction accuracy (relative position, speed over ground, time interval, etc.).

The improved Seq2Seq trajectory sequence prediction model combines the characteristics of spatiotemporal sequence data and Seq2Seq model, as shown in Figure 2. The real-time Seq2Seq prediction model is mainly composed of two RNNs modules. This model can be regarded as a process of encoding historical ship AIS trajectory input data and obtaining trajectory features, and then decoding the features to help predict future ship AIS trajectories. The model goal is to estimate conditional probability  $p(y_1, \dots, y_m | x_1, \dots, x_n)$ , among them,  $x_1, \dots, x_n$  represent the input historical track data,  $y_1, \dots, y_m$  represent the output predicted trajectory sequence.

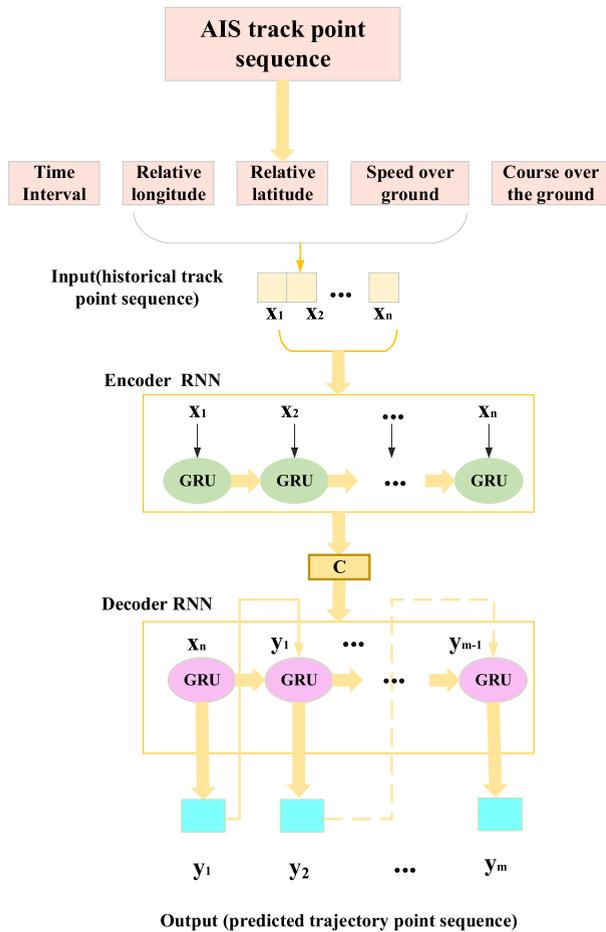
#### 1) ENCODER

The encoder part uses a cyclic neural network composed of a single-layer GRU unit to obtain the context feature information of the trajectory point sequence, and encode the ship trajectory input sequence (historical trajectory) into an abstract context vector  $C$ , the formula is as follows:

$$c = f(h_1, \dots, h_n) \quad (7)$$

Among them,  $h_1, \dots, h_n$  represent the hidden state value of each step in RNNs. For RNN, for a certain sequence, for time  $t$ , it is related to the state at the previous moment and the current input, that is  $h_t = f(h_{t-1}, x_t)$ ,

where  $f$  is a non-linear activation function, and the context vector  $c$  is directly assigned by the last hidden state  $h_n$ .



**FIGURE 2.** A Seq2Seq neural network model for vessel real-time trajectory prediction.

## 2) DECODER

The decoder part uses a recurrent neural network composed of a single-layer GRU unit. The traditional Seq2Seq is used for machine translation, and the initial input value in the decoder is the sequence start character  $\langle \text{BOS} \rangle$ . The model predicts the future trajectory from the current position, so  $x_n$  is used as the initial input value of the decoder part. It can be seen from Figure 2 that the decoder part of RNNs uses the context vector  $c$  as mentioned in formula (7), and the initial input value  $x_n$  to generate the predicted ship trajectory sequence. That is, the decoder uses the trajectory features extracted by the encoder and the current trajectory point position to predict the future ship trajectory. The calculation formula for the first hidden state  $h_1$  is as follows:

$$h_1 = f(c, x_n) \quad (8)$$

The sequence of ship predicted trajectory points is generated one by one following the time step of the encoder. The input of each RNN unit consists of the output of the hidden layer of the previous unit and the output of the previous unit. The expression is as follows:

$$y_t = f(h_{t-1}, y_{t-1}) \quad (9)$$

While  $f$  is our given activation function.

## 3) MODEL TRAINING

The Seq2Seq-based ship AIS trajectory real-time prediction model is a smooth interconnected model. The encoder and decoder parts of the model are all differentiable, that is, the trainable parameters in the model can be updated using the gradient descent method. This paper uses an adaptive learning rate in the optimization algorithm Adam to update the network parameters.

During the training process, the model is optimized by minimizing the root mean square error for each output. Adam algorithm is different from traditional stochastic gradient descent. Stochastic gradient descent maintains a single learning rate (that is, alpha) to update all weights, and the learning rate does not change during the training process. And Adam designs independent adaptive learning rates for different parameters by calculating the first-order moment estimation and the second-order moment estimation of the gradient. The author of the Adam algorithm described it as a set of advantages of two extended stochastic gradient descent:

(1) The adaptive gradient algorithm (AdaGrad) reserves a learning rate for each parameter to improve performance on sparse gradients (i.e. natural language and computer vision problems).

(2) Root Mean Square Propagation (RMSProp) adaptively preserves the learning rate for each parameter based on the mean of the nearest magnitude of the weight gradient. This means that the algorithm has excellent performance on non-steady state and online problems.

The root mean square error formula is as follows:

$$loss = \sqrt{\frac{\sum_{i=1}^n (\bar{y}_i - y_i)^2}{n}} \quad (10)$$

where  $\bar{y}$  represents the true value,  $y$  represents the predicted value, and  $n$  represents the number of samples.

## V. EXPERIMENTS

In this section, we present the evaluation of the benefits of the spatio-temporal Feature-optimized Seq2Seq Model using two real-world AIS datasets. This paper introduces real AIS data of the Yangtze River Channel as experimental data to compare and verify the validity of the model. There are differences in the channel between different areas of the Yangtze River Channel, as show in Fig3, Wuhan is located in the middle reaches of the Yangtze River, and the navigation channel is relatively straight, Chongqing is located in the upper reaches of the Yangtze River, with a curved navigation channel and many bayonet bay. Therefore, two data sets of Chongqing section and Wuhan section are used to conduct a confirmatory study on the accuracy and training speed of the model.

### A. DATASETS

We use real AIS data of the Yangtze River Channel. The dataset consists of a set of streaming data samples containing warp-separated tuples. Each tuple is sent by one vessel,

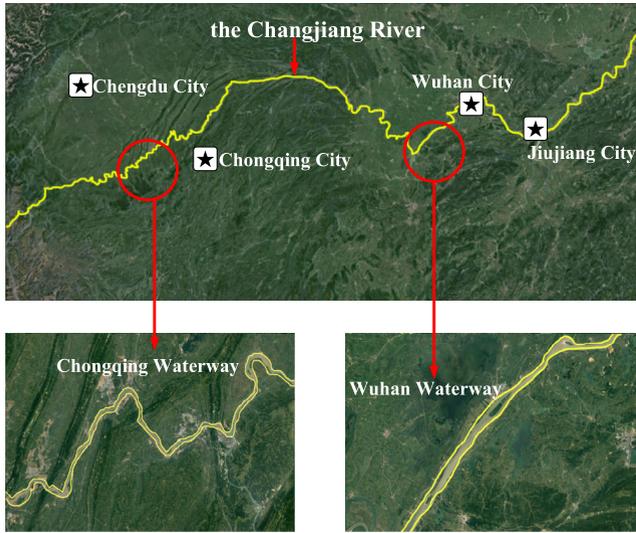


FIGURE 3. Chongqing and Wuhan waterway.

TABLE 1. Experiments data details.

| Dataset            | Wuhan Waterway  | Chongqing Waterway  |
|--------------------|---|---|
| Time Span          | 6/1/2017- 6/10/2017   | 6/1/2017- 6/10/2017                                       |
| Time interval      | 30s   | 30s   |
| Longitude Span     | 114.05°- 114.5°   | 106.00°- 106.61°  |
| Latitude Span      | 30.22°- 30.70°  | 29.00°- 29.62°  |
| Trajectory feature | point timestamp, relative longitude, relative latitude, SOG,COG | timestamp, relative longitude, relative latitude, SOG,COG |

containing its contemporary behavior under AIS specifications including MMSI, latitude, longitude, speed, course, heading, time stamp, departure port and so on. There are differences in waterway between different regions of the Yangtze River Channel, for example, Wuhan is located in the middle reaches of the Yangtze River, and the waterway is relatively straight, The ship travels more smoothly here, with less changes in direction. Chongqing is located in the upper reaches of the Yangtze River and the waterway has sharp bends and more bayonet sections, as show in Fig3, the current here is turbulent and the direction of the ship changes frequently. Make model predictions in different geographic environments, so that you can better detect the applicability of the model. We evaluate our model on Wuhan waterway datasets and Chongqing waterway datasets as show in Table 1. Each dataset detailed as follows.

1) WUHAN WATERWAY DATASET

Trajectory data is the ship AIS data in Wuhan waterway from 1st Jun.2017 - 10th Jun. 2017. The geographical range is 114.05° to 114.5° east longitude and 30.22° to 30.70° north latitude and the selected features are time, longitude, latitude, ground speed, and heading.

2) CHONGQING WATERWAY DATASET

Trajectory data is the ship AIS data in Chongqing waterway from 1st Jun.2017 - 10th Jun. 2017. The geographical range is 106.00° to 106.61° east longitude and 29.00° to 29.62° north latitude and the selected features are time, longitude, latitude, ground speed, and heading.

AIS training data segment as show in Table 2, For example, the data of number 1 indicates that at 18:00:10, the ship with MMSI number 268166 is located at the location of 114.2705067 longitude and 30.53765 latitude. At this time, the speed over the ground is 2.2 and the course over the ground is 206.2.

TABLE 2. AIS training data segment.

| No. | Time     | MMSI   | SOG | Lon         | Lat         | COG   |
|-----|----------|--------|-----|-------------|-------------|-------|
| 1   | 18:00:10 | 268166 | 2.2 | 114.2705067 | 30.53765    | 206.2 |
| 2   | 18:00:50 | 268166 | 2.2 | 114.2701867 | 30.53727    | 211.6 |
| 3   | 18:01:39 | 268166 | 2.2 | 114.2698667 | 30.53686333 | 207.6 |
| 4   | 18:02:00 | 268166 | 2.3 | 114.2697467 | 30.53668667 | 210.9 |
| 5   | 18:02:13 | 268166 | 2.2 | 114.2696667 | 30.53657333 | 211.7 |

B. HYPERPARAMETERS

The python library, Keras, was used to build our models. In this paper, a stochastic objective function optimization algorithm (Adam) was used to update the parameters during training. The Adam algorithm dynamically adjusts the learning rate for each parameter according to the first-order moment estimation and the second-order moment estimation of the gradient of each parameter according to the loss function. This algorithm is also based on the gradient descent method, but the learning step size of each iteration parameter has a certain range, and the large gradient will not lead to a large learning step size, so the parameter values are relatively stable. The learning rate was set to 0.001 and the decay set to 0.0, indicating the attenuation of the learning rate after each parameter update. In addition, in order to prevent over-fitting, a dropout mechanism was used in the experiment and its value was set to 0.1. Since the experimental training data set was relatively large, the batch-size used in this experiment was 100. We selected 85% of the training data for training each model and the remaining 15% was chosen as the validation set, used to early stop our training algorithm for each model, based on the best validation score.

C. EXPERIMENT DESIGN AND ANALYSIS

Seq2Seq-based real-time ship AIS trajectory prediction model was compared with typical recurrent neural network model in Wuhan straight and Chongqing curved waterways to verify the validity of the model.

1) COMPARISON OF ST-Seq2Seq WITH LSTM-RNN, GRU-RNN

The proposed model and two recurrent neural network models are compared on Wuhan and Chongqing waterway datasets to verify the effectiveness of the model.

(1) LSTM-RNNs, A prediction model for ship trajectories based on LSTM was adapted from the literature using a sequence of historical trajectory points as the input of the network and output a trajectory point at a time. It cannot predict the sequence of trajectory points at multiple consecutive moments at one time, and can only predict multiple times in a recursive manner.

(2) GRU-RNNs, A recurrent neural network composed of neurons for the GRU is used. Compared with LSTM, the convergence time and model training time of GRU are shorter. Like LSTM, GRU cannot predict multiple consecutive trajectory point sequences at one time, and can only predict multiple times recursively.

(3) ST-Seq2Seq, This paper presents a ship AIS trajectory prediction model based on Seq2Seq. The model uses the RNNs composed of single-layer GRU units to memorize a historical trajectory point sequence and then uses the RNNs composed of single-layer GRU units to decode the historical information to predict a trajectory point sequence in the future.

These three models were tested on the Wuhan and Chongqing datasets. A 10-minute historical trajectory sequence (20 trajectory points) is input to predict the trajectory sequence in the next five minutes (10 trajectory points). The experimental results are shown in Table 3.

TABLE 3. Performance comparison of different models.

| Dataset            | Model   | index | Root mean square error | Training time |
|--------------------|---------|-------|------------------------|---------------|
| Wuhan waterway     | LSTM    | RMSE  | 0.01607                | 459s          |
|                    | GRU     | RMSE  | 0.01918                | 374s          |
|                    | Seq2Seq | RMSE  | <b>0.00386</b>         | <b>167</b>    |
| Chongqing waterway | LSTM    | RMSE  | 0.01229                | 114s          |
|                    | GRU     | RMSE  | 0.01536                | 94s           |
|                    | Seq2Seq | RMSE  | <b>0.00849</b>         | <b>65s</b>    |

From the experimental results shown in Table 3, it can be seen that the Seq2Seq model achieved a reasonable short-term trajectory sequence prediction on both the Chongqing section and the Wuhan section. The LSTM and GRU models predicted the continuous sequence of trajectory points in sliding window mode. That is, every time the output of the next time  $t$  is predicted, the entire window is moved backward by one time. The new time window sequence is used to predict the output of the next time  $t + 1$ , recursively until the complete prediction sequence is output. The results of predicting the sequence of 10 consecutive track points are shown in Figure 4. When predicting the first and second trajectory points, the error of the LSTM and GRU models was smaller than the Seq2Seq model. When the predicted trajectory points were greater than two, the prediction error in

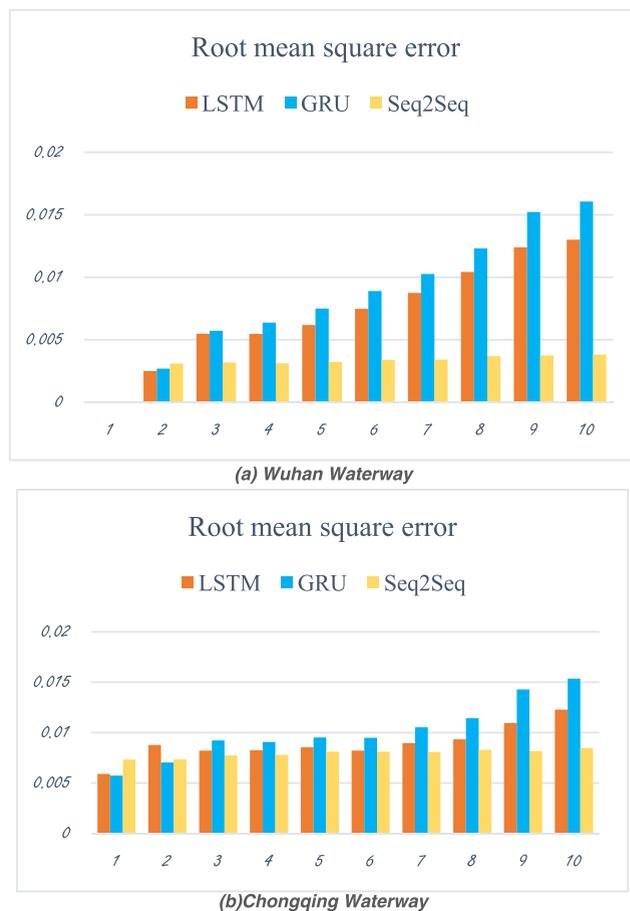


FIGURE 4. Comparison of loss of three models.

the Seq2Seq model was significantly smaller than the former. This is mainly because the sliding window LSTM and GRU models will accumulate prediction error of the last moment as the window moved back, the error gradually increased. Especially when the number of predicted trajectory points was eight the error between the two greatly increased. When the Seq2Seq model was used to predict a continuous sequence of trajectory points, the error changed more smoothly. Since the basic LSTM and GRU networks use a recursive method to predict a sequence of continuous trajectory points, when the number of predicted trajectory points increased, the prediction effect also decreased. As shown in Figure 4, the three models on the Wuhan segment and Chongqing segment data sets predict the variation of the error in the sequence of increasing trajectory points.

2) COMPARISON OF DIFFERENT STRUCTURE Seq2Seq MODELS

The encoder and decoder in the Seq2Seq are composed of recurrent neural networks; therefore, when predicting a ship’s trajectory, the structure of the recurrent neural network in the decoder and encoder will affect the prediction effect of the trajectory. This experiment compares the ship trajectory predictions from differently structure Seq2Seq models. The input sequence is a 10-minute

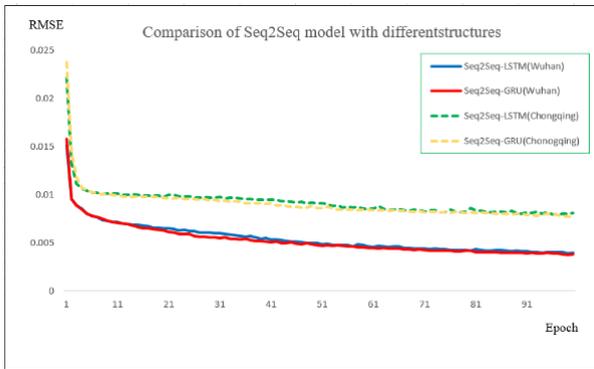


FIGURE 5. Comparison of Seq2Seq model with different structures.

historical trajectory sequence, and the output sequence is a 5-minute predicted trajectory sequence.

(1) The encoder is a recurrent neural network composed of a single-layer LSTM unit, and the decoder is a recurrent neural network composed of a single-layer LSTM unit.

(2) The encoder is a recurrent neural network composed of a single-layer GRU unit, and the decoder is a recurrent neural network composed of a single-layer GRU unit. The experimental results are shown in Table 4.

TABLE 4. Performance comparison of different structures of Seq2Seq.

| Dataset            | encoder | decoder | Root mean square error | Training time |
|--------------------|---------|---------|------------------------|---------------|
| Wuhan waterway     | LSTM    | LSTM    | 0.003971               | 1067s         |
|                    | GRU     | GRU     | 0.003823               | 997s          |
| Chongqing waterway | LSTM    | LSTM    | 0.008079               | 304s          |
|                    | GRU     | GRU     | 0.007765               | 261s          |

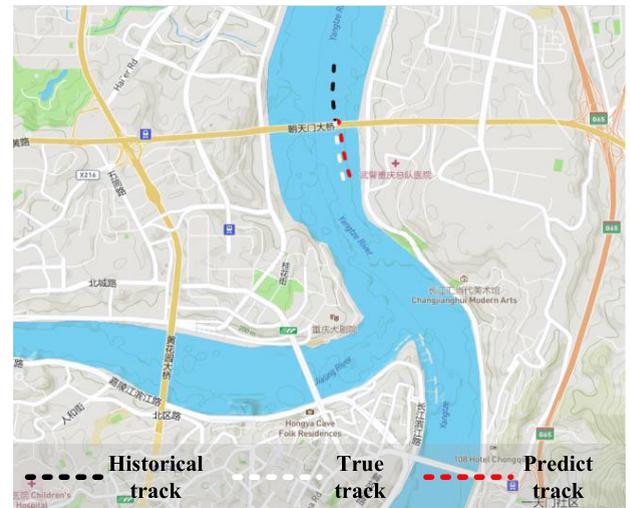
From the experimental results shown in Table 4, it can be seen that when the encoder and decoder are both GRU networks, the trajectory prediction effect is better. In addition, in the same environment, when the encoder and decoder are GRU networks, the model takes less time to complete an iteration. This is because the GRU structure is simpler than that of LSTM, and GRU has one less cell state than LSTM. Figure 5 shows the comparison of the training process of the two structures of the model on the Wuhan segment training set and the Chongqing segment training set when predicting the ship trajectory sequence.

It can be seen from Figure 5 that the prediction model for ship trajectories based on Seq2Seq can effectively predict the ship trajectory over a short period in the Wuhan and Chongqing segments. Where the encoder and decoder both use GRU networks, convergence was faster. Because the Wuhan waterway is smoother than the Chongqing waterway, the forecast error is small. Based on the comparison of Table 4 and Figure 5, it can be seen that when the encoder and decoder are GRU networks, the prediction effect is more reasonable and the training time is shorter.

As shown in Figure 6, when using the Seq2Seq model to predict the sequence of trajectory points, the error changes



(a) Wuhan Waterway



(b) Chongqing Waterway

FIGURE 6. Forecast result.

less as the number of predicted trajectory points increases. Moreover, the prediction result of Wuhan segment is slightly higher than that of Chongqing segment.

VI. CONCLUSION AND FUTURE

In this paper, we propose a seq2seq framework model based on spatio-temporal data, which can automatically predict the trajectory sequence, and the model can be used for short-term prediction, which can improve the timeliness of ship collision warning. We evaluate our model on two types of waterways in Wuhan and Chongqing, outperforming three baseline methods, confirming that our model is more applicable to ship trajectory prediction. In the current ship trajectory predictions, the time range of the predicted trajectory tends to be large, so the prediction result cannot be applied to collision avoidance, while some models that can predict the ship trajectory in a short time cannot predict multiple times at one time. The trajectory point can only be predicted multiple times in a recursive manner, so that the error between

the predicted trajectory point and the real trajectory point is continuously increased. The model proposed in this paper can encode the historical trajectory by the encoder, and effectively extracts historical trajectory features. The decoder decodes the feature and applies it to the prediction of all trajectory points, reducing error when predicting a trajectory.

The current model is not effective in predicting future trajectories over five minutes. Therefore, the focus of future work will be on improving the model, the encoder and decoder might use different network structures or the network deepened to predict the ship trajectories for a longer time and thus providing an early warning to avoid ship collisions. As it is difficult to combine LSTM and GRU methods with our ST-Seq2Seq method but this might be considered in our further work.

## REFERENCES

- [1] G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 12, pp. 2218–2245, Jun. 2013.
- [2] B. R. Dalsnes, S. Hexeberg, A. L. Flaten, B.-O.-H. Eriksen, and E. F. Brekke, "The neighbor course distribution method with Gaussian mixture models for AIS-based vessel trajectory prediction," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Cambridge, U.K., Jul. 2018, pp. 580–587.
- [3] L. P. Perera, P. Oliveira, and C. G. Soares, "Maritime traffic monitoring based on vessel detection, tracking, state estimation, and trajectory prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1188–1200, Sep. 2012.
- [4] P. Borkowski, "The ship movement trajectory prediction algorithm using navigational data fusion," *Sensors*, vol. 17, no. 6, p. 1432, Jun. 2017.
- [5] 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 Things J.*, early access, Oct. 5, 2020, doi: [10.1109/JIOT.2020.3028743](https://doi.org/10.1109/JIOT.2020.3028743).
- [6] A. Harati-Mokhtari, A. Wall, P. Brooks, and J. Wang, "Automatic identification system (AIS): Data reliability and human error implications," *J. Navigat.*, vol. 60, no. 3, pp. 373–389, Sep. 2007.
- [7] E. Tu, G. Zhang, L. Rachmawati, E. Rajabally, and G.-B. Huang, "Exploiting AIS data for intelligent maritime navigation: A comprehensive survey from data to methodology," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1559–1582, May 2018.
- [8] D. Caveney, "Numerical integration for future vehicle path prediction," in *Proc. ACC*, New York, NY, USA, 2007, pp. 3906–3912.
- [9] H. Li, J. Liu, Z. Yang, R. W. Liu, K. Wu, and Y. Wan, "Adaptively constrained dynamic time warping for time series classification and clustering," *Inf. Sci.*, vol. 534, pp. 97–116, Sep. 2020, doi: [10.1016/j.ins.2020.04.009](https://doi.org/10.1016/j.ins.2020.04.009).
- [10] E. Semerdjiev and L. Mihaylova, "Variable-and fixed-structure augmented interacting multiple model algorithms for manoeuvring ship tracking based on new ship models," *Int. J. Appl. Math.*, vol. 10, no. 3, pp. 591–604, Apr. 2000.
- [11] F. Tusell, "Kalman filtering in R," *J. Stat. Softw.*, vol. 39, no. 2, pp. 1–27, 2011.
- [12] L. P. Perera and C. G. Soares, "Ocean vessel trajectory estimation and prediction based on extended Kalman filter," in *Proc. 2nd SASO*, Budapest, Hungary, 2010, pp. 14–20.
- [13] Q. Bo and Y. B. Chen, "Prediction model of ship trajectory based on LSTM," *Comput. Sci.*, vol. 45, no. S2, pp. 136–141, 2018.
- [14] D.-D. Nguyen, C. Le Van, and M. I. Ali, "Vessel trajectory prediction using sequence-to-sequence models over spatial grid," in *Proc. 12th ACM Int. Conf. Distrib. Event-Based Syst.*, Hamilton, New Zealand, Jun. 2018, pp. 258–261.
- [15] S. J. Meng *et al.*, "Drift trajectory model of the unpowered vessel on the sea and its application in the drift simulation of the sanchi oil tanker," *Oceanologia et Limnologia Sinica*, unpublished, doi: [10.11693/hyhz20180200037](https://doi.org/10.11693/hyhz20180200037).
- [16] B. Murray and L. P. Perera, "A data-driven approach to vessel trajectory prediction for safe autonomous ship operations," in *Proc. 13th Int. Conf. Digit. Inf. Manage. (ICDIM)*, Berlin, Germany, Sep. 2018, pp. 240–247.
- [17] Z. Yan, "Traj-ARIMA: A spatial-time series model for network-constrained trajectory," in *Proc. 3rd ACM IWCTS*, San Jose, CA, USA, 2010, pp. 11–16.
- [18] Y. Zheng, "Trajectory data mining: An overview," *ACM Trans. Intell. Syst. Technol.*, vol. 6, no. 5, p. 29, May 2015.
- [19] G. Siegert, P. Banys, C. S. Martinez, and F. Heymann, "EKF based trajectory tracking and integrity monitoring of AIS data," in *Proc. IEEE/ION Position, Location Navigat. Symp. (PLANS)*, Savannah, GA, USA, Apr. 2016, pp. 887–897.
- [20] 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 Things J.*, vol. 7, no. 11, pp. 10794–10812, Nov. 2020, doi: [10.1109/JIOT.2020.2989398](https://doi.org/10.1109/JIOT.2020.2989398).
- [21] N. Deo and M. M. Trivedi, "Convolutional social pooling for vehicle trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Salt Lake City, UT, USA, Jun. 2018, pp. 1468–1476.
- [22] N. Deo and M. M. Trivedi, "Multi-modal trajectory prediction of surrounding vehicles with maneuver based LSTMs," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Changshu, China, Jun. 2018, pp. 1179–1184.
- [23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [24] L. Liu, D. Malak, and M. Medard, "Guesswork for inference in machine translation with Seq2Seq model," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Visby, Sweden, Aug. 2019, pp. 1–5.
- [25] Y. Zhang, D. Li, Y. Wang, Y. Fang, and W. Xiao, "Abstract text summarization with a convolutional Seq2seq model," *Appl. Sci.*, vol. 9, no. 8, p. 1665, Apr. 2019.
- [26] Y. Zhang and W. Xiao, "Keyphrase generation based on deep Seq2seq model," *IEEE Access*, vol. 6, pp. 46047–46057, 2018.
- [27] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*. [Online]. Available: <https://arxiv.org/abs/1406.1078>
- [28] A. Jain, A. R. Zamir, S. Savarese, and A. Saxena, "Structural-RNN: Deep learning on spatio-temporal graphs," in *Proc. IEEE CVPR*, Las Vegas, NV, USA, Jun. 2016, pp. 5308–5317.
- [29] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [30] R. Zhang, Z. Wang, and D. Mai, "Building emotional conversation systems using multi-task Seq2Seq learning," in *Proc. NLPCC*, Beijing, China, 2017, pp. 612–621.



LAN YOU was born in Wuhan, China, in 1978. She received the Ph.D. degree from LIESMARS, Wuhan University, in 2015.

From 2015 to 2017, she worked as a Research Associate with The Chinese University of Hong Kong. She is currently an Associate Professor with the Computer Science and Information Technology Faculty, Hubei University. She has published widely in the field with more than 20 publications. Her research interests include spatio-temporal data mining, knowledge graph, and virtual geographical environment.



SIYU XIAO is currently pursuing the master's degree in computer science with Hubei University. Her current research interests include knowledge mapping and spatio-temporal data.



**QINGXI PENG** was born in Wuhan, China, in 1974. He received the Ph.D. degree in computer software and theory from Wuhan University, in 2018.

He is currently a Professor with the Computer School, Wuhan Donghu University. He is also specialized in the design and implementation of large-scale distributed web extraction. He is the author of one book, more than 50 articles, and seven inventions. His research interests include machine learning, natural language processing, and information retrieval. He is also a member of China Computer Federation.



**ZHENGYI GUAN** is currently pursuing the B.Eng. degree in data science and big data technology with Hubei University. His current research interests include knowledge mapping and trajectory analysis.



**CHRISTOPHE CLARAMUNT** is currently a Professor of computer science and the Chair of the Naval Academy Research Institute, France. His research interests include the theoretical and pluri-disciplinary aspects of geographical information science and their applications to urban, maritime, and environmental systems. He has widely published in the domain of GIS and serves in the editorial boards of several international GIS journals and major GIS conferences, including acting as an Associate Editor of the *International Journal of Geographical Information Science*.

acting as an Associate Editor of the *International Journal of Geographical Information Science*.



**XUEWEI HAN** received the M.Eng. degree in computer technology from Hubei University, in 2020.

Her current research interests include spatio-temporal data and intelligent information systems.



**JIAHE ZHANG** is currently pursuing the B.Eng. degree in computer science and technology with Hubei University. His current research interests include knowledge mapping and neural networks.

...