Article

Predicting Ship Trajectory Based on Neural Networks Using AIS Data

Tamara A. Volkova [1,*], Yulia E. Balykina [2,*] and Alexander Bespalov [1,*]

[1] Department of Applied Mathematics, Admiral Makarov State University of Maritime and Inland Shipping, 5/7, Dvinskaya Str., Saint Petersburg 198035, Russia
[2] Faculty of Applied Mathematics and Control Processes, Saint Petersburg State University, 7-9, Universitetskaya Naberezhnaya, Saint Petersburg 199034, Russia
* Correspondence: volkovata@gumrf.ru (T.A.V.); j.balykina@spbu.ru (Y.E.B.); bespalov.aleksandr.p@gmail.com (A.B.)

Abstract: To create an autonomously moving vessel, it is necessary to know exactly how to determine the current coordinates of the vessel in the selected coordinate system, determine the actual trajectory of the vessel, estimate the motion trend to predict the current coordinates, and calculate the course correction to return to the line of the specified path. The navigational and hydrographic conditions of navigation on each section of the route determine the requirements for the accuracy of observations and the time spent on locating the vessel. The problem of predicting the trajectory of the vessel's motion in automatic mode is especially important for river vessels or river-sea vessels, predicting the trajectory of the route sections during the maneuvering of the vessel. At the moment, one of the most accurate ways of determining the coordinates of the vessel is by reading the satellite signal. However, when a vessel is near hydraulic structures, problems may arise connected with obtaining a satellite signal due to interference and, therefore, the error in measuring the coordinates of the vessel increases. The likelihood of collisions and various kinds of incidents increases. In such cases, it is possible to correct the trajectory of the movement using an autonomous navigation system. In this work, opportunities of the possible application of artificial neural networks to create such a corrective system using only the coordinates of the ship's position are discussed. It was found that this is possible on sections of the route where the ship does not maneuver.

Keywords: AIS Data; trajectory prediction; waterway transportation; neural networks; autonomous navigation

1. Introduction

Nowadays, to build the trajectory of the vessel and determine its position, it is a common approach to use data from satellite navigation systems and automatic identification systems (AIS), which allow obtaining coordinates with high accuracy and secondly, discreteness. However, despite the advances in the development of satellite navigation technologies in recent decades, their methods have a significant drawback: non-autonomy, which does not allow the use of a satellite system when a number of known circumstances occur. The main ones are the loss of communication or, under the conditions of maneuvering, too long processing the received data, which prevent it from making a decision in time. In this regard, the issues of studying autonomous methods for determining a ship's position are of great practical interest. The prediction of the trajectory of the vessel's movement is of practical importance for the development and creation of unmanned autonomous vessels.

Taking into account the inertia of the vessel, in order to react in time to possible obstacles, it is necessary to predict in advance the estimated nearest coordinates of the vessel. In this regard, it is necessary to take into account the capabilities of the ship's technical means, so that the time required for the autopilot to make a decision is sufficient enough to implement it. In this sense, predicting the trajectory of a vessel's movement based on

historical data, that is, using data received from other vessels, can only be used very approximately, since the trajectory of a vessel's movement depends quite strongly on specific external conditions. Therefore, the goal of this work was to analyze the application of neural network technologies that use knowledge of the previous coordinates of the vessel's trajectory to predict the estimated next coordinates of the vessel during river navigation.

Related Works

Various studies have examined different aspects of predicting vessel trajectories using AIS data, including behavior anomaly detection [1], waypoint detection [2], and the actual forecast of the vessel's trajectory [3–5]. Since artificial neural networks gained wide popularity, a number of studies have appeared that focused on the use of artificial intelligence in solving problems related to navigation. For example, there is a significant difference between the different patterns of vessel traffic, namely static navigation, normal navigation, and maneuvering. As the data that are collected during maneuvering are of the greatest interest, a separate issue is the task of separating them from the general data. In this regard, there are a number of studies that use artificial neural networks to classify AIS data types in relation to ship traffic [6].

Neural networks are often used to study the trajectories of various types of ships in the oceans [7–9]. They allow one to study the possible routes of a ship's motion, and identify the most frequently used ones. There are various probabilistic approaches for solving this kind of problem as well [10]. There are studies that, while solving this problem, also touch upon the problem of using neural technologies to predict possible ship collisions [11]. It is believed that the identification and construction of optimal routes will allow the use of unmanned vessels, which will follow the given trajectories [12–14]. When using neural networks to predict the motion of the vessel and build the trajectory of its motion, neural networks are used either as a function for identifying the model of the vessel movement, or for the trajectory prediction, and the correction of their free parameters is performed in real time as the vessel is moving [15–17].

A detailed review of various approaches to predicting the trajectory of vessel movement or location is given in [18]. Both classical models, such as Constant Velocity Model (CVM), and modern approaches based on machine learning and statistical analysis are analyzed. Models for predicting the trajectory of a specific vessel based on neural networks using course of ground (COG) and speed over ground (SOG), attached with a corresponding future timestamp as input, are discussed in [19], and based on the obtained data the coordinates of the ship are restored.

2. Materials and Methods

Unlike the movement of sea vessels, which are faced with maneuvering in confined spaces mainly when passing straits and moving in port waters, river vessels are forced to constantly maneuver in confined spaces (e.g., the presence of a large number of possible obstacles, limited fairway, frequent course changes, a large number of oncoming and passing vessels). Moreover, the performance of satellite navigation systems is greatly influenced by the presence of possible sources of industrial interference, as well as a large number of hydraulic structures, located on the riverbanks. Therefore, AIS data are not error-free. Different sources of error that render a portion of these data noisy and inadequate for monitoring are discussed in detail in [20].

Since the aim of the research was to determine the possibility of predicting the position of the vessel at the next moment in time using only the previous coordinates of the vessel's movement for training the neural network, only these data were collected during the experiment. The real experiment was carried out specifically for collecting data in the conditions of river navigation, as the data in existing open databases often consist of information collected during sea navigation.

It was assumed that the problem of predicting the vessel trajectory could be solved on the basis of a focused time lagged feedforward network. For training the neural network

model, the data obtained during the experiment were used. The experiment was carried out on a ship with characteristics presented in Table 1:

Table 1. Vessel characteristics.

Feature	Value	Units
Maximum length	35.65	meters
Maximum width	5.80	meters
Side height	2.60	meters
Draft	1.51	meters
Displacement	157	tons
Main engine power	2×200	kW
Crew	10	man
Freeboard	1.096	meters
Endurance	6	days

There were two antennas placed on the roof of the ship's cabin at a distance of 90 cm along the center plane of the ship and 66 centimeters between the antennas. The data from the second antenna underwent a post-processing and were subsequently received as a reference. All vessel position data were recorded using the NMEA-0183 protocol.

NMEA is a format for transmitting data between ship instruments. It includes a message system for information exchange between GPS navigation receivers and navigation information consumers. All commands and messages are transmitted in ASCII text form. The last field of the message may contain the checksum of the current message, starting with the * separator. This is followed by an 8-bit checksum (exclusive OR) of all characters in the message, including spaces between the $ and * separators, not including the last one. The hexadecimal result is translated into two ASCII characters (0–9, A–F).

The GGA-GPS data, which were obtained during the experiment and used to train neural networks and verify the results, contained the following values:

- Greenwich Mean Time at the moment of determining the location.
- Latitude.
- North/South (N/S).
- Longitude.
- West/East (E/W).
- GPS signal quality indicator:
  0 = Positioning is not possible or is not correct;
  1 = GPS mode, normal accuracy, location possible;
  2 = Differential GPS mode, normal accuracy, location possible;
  3 = GPS precision mode, location possible.
- The number of satellites used (0–12, may differ from the number of visible ones).
- Horizontal Dilution of Precision (HDOP).
- Receiver antenna height above/below sea level.
- Unit of antenna's location height, meters.
- Geoid difference, i.e., the difference between the WGS-84 ellipsoid and the sea level (geoid), "-" = sea level is below the ellipsoid.
- Units of measurement, meters.
- GPS Differential Data Age—Time in seconds since the last SC104 type 1 or 9 update, filled with zeroes if the differential mode is not used.
- ID of the station transmitting differential corrections, ID, 0000-1023.
- Checksum.

An example of the received data is presented in Figure 1.

```
$PNVGDBG,085228.20,1,6.8,38,0.0016576,-0.0076*62
$PNVGIMU,085228.20,-4.60,3.86,320.70,0.07,-0.22,-0.03*7B
$PNVGIRQ,28272,28272,0,2828960,2828960,0,2828960,2828960,0*59
$PNVGCPU,38774,33463,0,2,0,17617,9177,0,0*67
$PNVGKF0,332,7,436,86,407,701,101,1128,60,24,26*1F
$PNVGKF1,228,4,139,57,463,539,60,481,74,24,13*18
$GPRMC,085228.40,A,5952.0525254,N,03058.2179220,E,3.99,312.71,270918,,,R*45
$GPGGA,085228.40,5952.0525254,N,03058.2179220,E,4,15,0.8,31.883,M,15.787,M,0.4,0000*7A
$GPGSV,5,1,18,7,4,308,40,8,41,296,46,10,56,180,51,13,15,23,41*77
$GPGSV,5,2,18,15,30,61,51,16,22,216,46,18,6,268,38,20,67,117,49*72
```

Figure 1. Sample data received from the first antenna.

To obtain images correlated with geographic maps, the algorithm for transferring data from the NMEA protocol to the WGS84 system (World Geodetic System 1984) was used [21]:

$$flTemp = Value \ (NMEA \ 0183)$$

$$nTemp = flTemp/100.0;$$

$$nTemp = nTemp - (nTemp\% \ 1);$$

$$flMin = flTemp - 100.0 * nTemp;$$

$$Lo = nTemp + flMin/60.0;$$

where /—division, %—remainder of division, *—multiplication, Lo—value in WGS84.

3. Results and Discussion

The experiment was conducted on a segment of inland waterways of the Neva-Ladoga region. The resulting vessel path is shown in Figure 2. It should be noted that the results of all antennas are very close to each other.
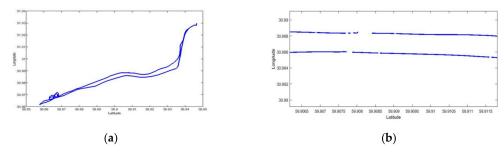

(**a**)   (**b**)

Figure 2. Total path of the ship (**a**), the part of the path near and under the bridge (**b**).

During the experiment, the case of a ship passing under a bridge was considered separately. Since the passage of the vessel under the bridge is quite difficult due to the small distance between the bridge supports, the flow of the river, etc., in order to reduce the risk of a dangerous situation, it is necessary to have alternative methods of determining the exact location of the vessel.

One can see (Figures 3–5) how the received signals behave while passing the bridge. It is clear that the signal is either interrupted, or does not reflect in any way the real trajectory of the vessel. This fact is most evident for the processed signal when the system is trying to restore the trajectory.

In this regard, the task was set to assess the possibilities of using neural networks to solve such problems, in particular, to determine whether it is possible to build a neural network that will restore the trajectory of a ship using the previous AIS data. Neural networks with various activation functions were studied.

(**a**)   (**b**)

Figure 3. Vessel trajectory, obtained from the first antenna (**a**), vessel trajectory under the bridge, obtained from the first antenna (**b**).
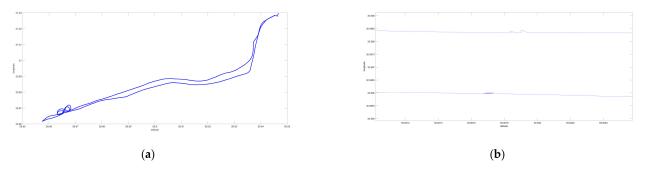

(**a**)   (**b**)

Figure 4. Vessel trajectory, obtained from the second antenna (**a**), vessel trajectory under the bridge, obtained from the second antenna (**b**).

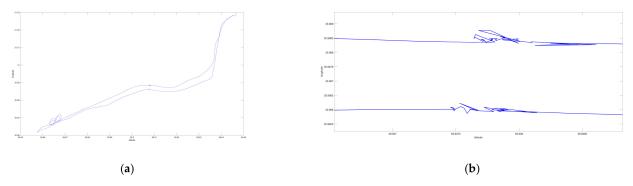(**a**)   (**b**)

Figure 5. Vessel trajectory (**a**); trajectory under the bridge, the result of post-processing the signal (**b**).

To train the neural network, an array of 1000 lines containing latitude and longitude was taken (see Table 2 for data sample).

Table 2. Input data sample.

Latitude	Longitude
59.9466765683333	31.0296530033333
59.9466765750000	31.0296530366667
59.9466765833333	31.0296530600000
59.9466765916667	31.0296530766667
59.9466765900000	31.0296531200000
59.9466765883333	31.0296531333333
59.9466765733333	31.0296531333333

There was 10% of the initial set left for the validation process. Of the remaining 90%, 70% was taken as a training set, and 20% was left for testing.

The constructed neural network consisted of two layers, the first one contained neurons with sigmoidal activation functions, and the other one had one neuron with identical activation functions. The network, taking the vector of the delayed values of the ship's coordinates in the geographic system as an input, predicts its value one step ahead. The network is trained in real time based on a sample containing a certain number of previous trajectory points.

The Levenberg-Marquardt algorithm, which implements the idea of gradient descent, was used as a training algorithm. The mean square of the course error was chosen as a measure of the training accuracy. This algorithm most adequately reproduces the dynamics of the vessel's movement. Its advantages, in comparison with other methods of conjugate gradients, are high speed calculations and ensuring convergence.

The first of the constructed networks included 200 input neurons. The hyperbolic tangent was used as a neuron activation function (1).
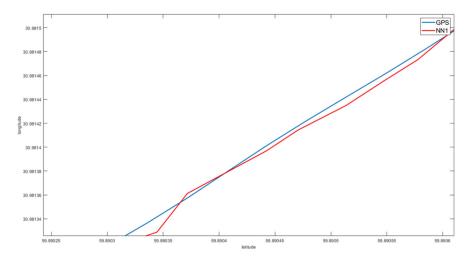
$$a = \frac{2}{(1 + \exp(-2 \cdot n))} - 1 \tag{1}$$

The network with 200 neurons in the hidden layer was trained in 50 iterations, and the validation showed the difference in latitude by 10^{-6}, which is an insignificant discrepancy from the true data, since the numbers after 6 decimal places are not significant (Table 3). Resulting average error on the straight sections of the vessel path was 0.02° An increase in the number of neurons did not lead to a decrease in errors, but increased the calculation time.

Table 3. Results of the first neural network.

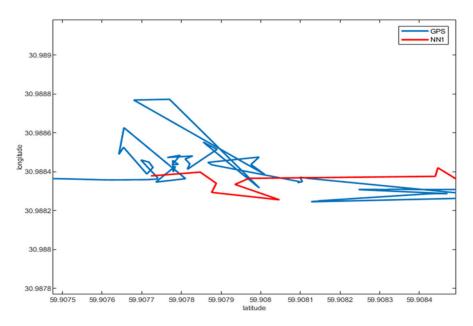|	Input Data	Output Data	Reference Data
Latitude	59.8665262166667	59.866514809161487	59.8665157383333
Longitude	30.970706225	30.969370218195859	30.9706893433333

The results of the first neural network performance are presented in Figures 6 and 7.

Figure 6. Comparison of the trajectory predicted by the neural network with the reference trajectory on the straight section of the path.

The second of the options for building and training a neural network was a network with the following structure: the first layer was the input data for the neural network and it did not participate in the calculation; the second layer was a hidden layer; and the third one processed the data received from the second layer and returned the desired value. For the convenience of working with the neural network, the input data were transformed as follows: the GPS coordinate values were divided by 100 to bring them to the interval (0; 1).

Subsequently, the data were returned to their original state. To train the neural network, 24,000 datasets were used with 10,000 iterations for each dataset.

Figure 7. Comparison of the trajectory predicted by the neural network with the reference trajectory along the path with data loss and distortion.

Function (2) was used as an activation function.

$$\sigma(x) = \frac{1}{(1 + \exp(-tx))} \tag{2}$$

In this case, t is the weight of the neuron, which is automatically selected during training of the neural network. The peculiarity of neurons with such a transfer characteristic is that they amplify strong signals much less than weak ones, since the areas of strong signals correspond to the flat areas of the activation function.
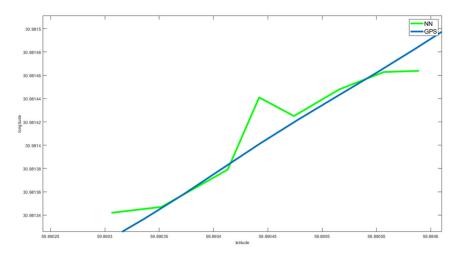
It should be noted that functions (1) and (2) differ only in the range of values, however, this difference significantly changes the possible input and output values of the neuron, which subsequently affects the learning rate and accuracy of the neural network.

For the input data, 10 pairs of latitude and longitude values were used, from which the values of the next coordinate were obtained. Data in Table 4 show that the accuracy of the data produced by the neural network differs for the worse, since in this case the difference in latitude from the true data is 10^{-5}.
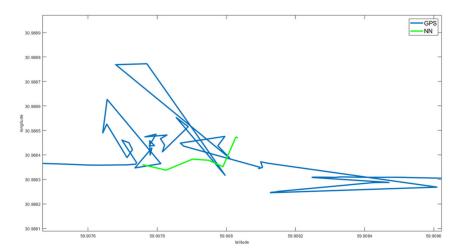
Table 4. Results of the second neural network.

Input Data		Output Data	
59.8902644083333	30.9812980466667		
59.8902644083333	30.9812980466667		
59.8902644083333	30.9812980466667		
59.8903000583333	30.9812980466667		
59.8903000583333	30.9813171183333		
59.8903000583333	30.9813171183333	59.8903520000000	30.9813470000000
59.8903000583333	30.9813171183333		
59.8903000583333	30.9813171183333		
59.8903360133333	30.9813171183333		
59.8903060000000	30.9813420000000		

The results of the second neural network are presented in Figures 8 and 9.

Figure 8. Comparison of the trajectory predicted by the neural network with the reference trajectory on the straight section of the path.
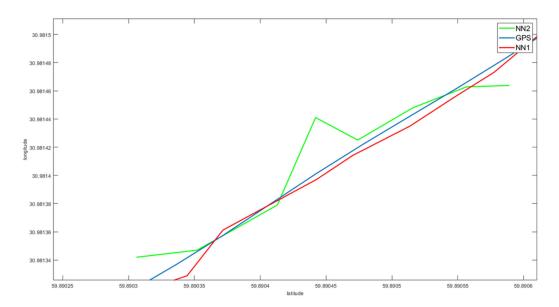

Figure 9. Comparison of the trajectory predicted by the neural network with the reference trajectory along the path with data loss and distortion.

While assessing the performance of the second neural network, the following peculiarities appear:
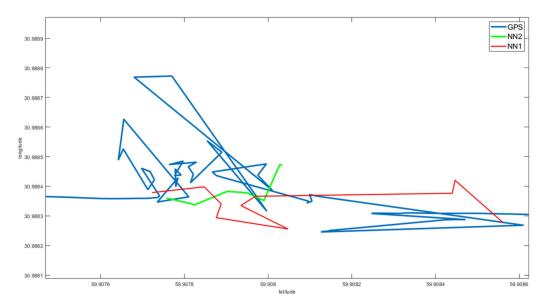
- the considered neural network can predict the trajectory of a vessel's motion on straight sections; however, due to the accuracy of such a prediction, it is impossible to guarantee the safety of the vessel's motion (Figure 8);
- in areas with GPS data loss or distortion, the first few iterations of the neural network show a good result, which can be seen in Figure 9;
- the developed neural network is not suitable for long-term use due to the lack of stable prediction quality.

Figures 10 and 11 show the comparison of the results of the neural networks' performance with respect to each other and the trajectory obtained from the reference coordinates.

Despite the fact that the second network obviously gives a better result than the first one, it is still far from the true trajectory of the ship. In the case when the ship begins to make any maneuvers (e.g., turn), or enters the interference zone, the available AIS data are insufficient to train neural networks that predict the position or trajectory of the vessel. Therefore, additional research is needed to adapt neural network technologies in relation to the physical characteristics of each specific vessel, as well as swimming conditions, characterized by a change in the density of water and air, the influence of shallow water, and other similar effects.

Figure 10. Comparison of the trajectories predicted by the considered neural networks with the reference trajectory on the straight section of the path.

Figure 11. Comparison of the trajectories predicted by considered neural networks with the reference trajectory along the path with data loss and distortion.

4. Conclusions

The problem of determining the position of the vessel in real time is very important for various kinds of applications, in particular, for the autopilot system development. The use of existing algorithms based on AIS data does not allow assessing the position of the vessel in the presence of interference that affects the receipt of AIS signal. One of the ways to solve this issue is the development of an autonomous system that predicts the trajectory of the vessel's movement, which will be a source of additional information for timely decision-making in case of AIS signal distortion. In this paper, the possibility of using neural network technologies is being considered as one of the options for predicting the trajectory of a ship. Neural networks with different activation functions and different structures were built and trained, using experimental data of the real movement of the vessel in the fairway of the Neva River. The goal was to assess whether it was possible to

obtain reliable results using only the coordinates of the ship's movement obtained from AIS for neural network training.

Simulation modeling has shown that when the vessel moves along a straight trajectory, the use of neural networks gives quite satisfactory results. The main problems appear when the ship starts to maneuver. With this type of movement, the results shown by trained neural networks are not satisfactory, since the discrepancy between the results of their work and real data is very large. In fact, it is possible to predict the trajectory only one coordinate ahead, which in the conditions of ship maneuvering does not permit using this technology to create autopilots.

Thus, additional research is required to identify which parameters should be added to train neural networks to predict the trajectory during ship maneuvering more accurately. Further research will be devoted to these issues.

Author Contributions: Conceptualization, T.A.V., Y.E.B. and A.B.; methodology, T.A.V.; Y.E.B. and A.B.; validation, T.A.V., Y.E.B. and A.B.; writing—original draft preparation, T.A.V., Y.E.B. and A.B.; writing—review and editing, T.A.V., Y.E.B. and A.B.; supervision, T.A.V., Y.E.B. and A.B. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sidibé, A.; Shu, G. Study of Automatic Anomalous Behaviour Detection Techniques for Maritime Vessels. *J. Navig.* **2017**, *70*, 847–858. [CrossRef]
2. Dobrkovic, A.; Iacob, M.-E.; Hillegersberg, J. Using machine learning for unsupervised maritime waypoint discovery from streaming AIS data. In Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business, Graz, Austria, 21–22 October 2015; pp. 1–8. [CrossRef]
3. Mao, S.; Tu, E.; Zhang, G.; Rachmawati, L.; Rajabally, E.; Huang, G.-B. An Automatic Identification System (AIS) Database for Maritime Trajectory Prediction and Data Mining. *arXiv* **2018**, arXiv:1607.03306. Available online: https://arxiv.org/abs/1607.03306 (accessed on 26 December 2020).
4. Alizadeh, D.; Alesheikh, A.; Sharif, M. Prediction of vessels locations and maritime traffic using similarity measurement of trajectory. *Ann. GIS* **2020**. [CrossRef]
5. Alizadeh, D.; Alesheikh, A.; Sharif, M. Vessel Trajectory Prediction Using Historical Automatic Identification System Data. *J. Navig.* **2021**, *74*, 156–174. [CrossRef]
6. Chen, X.; Liu, Y.; Achuthan, K.; Zhang, X. A ship movement classification based on Automatic Identification System (AIS) data using Convolutional Neural Network. *Ocean Eng.* **2020**, *218*, 108182. [CrossRef]
7. De Vries, G.K.D.; Van Someren, M. Machine learning for vessel trajectories using compression, alignments and domain knowledge. *Expert Syst. Appl.* **2012**, *39*, 13426–13439. [CrossRef]
8. Murray, B.; Perera, L.P. A dual linear autoencoder approach for vessel trajectory prediction using historical AIS data. *Ocean Eng.* **2020**, *209*, 107478. [CrossRef]
9. Soldi, G.; Gaglione, D.; Forti, N.; Di Simone, A.; Daffinà, F.C.; Bottini, G.; Quattrociocchi, D.; Millefiori, L.M.; Braca, P.; Carniel, S.; et al. Space-based Global Maritime Surveillance. Part II: Artificial Intelligence and Data Fusion Techniques. *arXiv* **2020**, arXiv:2011.11338.
10. Rong, H.; Teixeira, A.P.; Guedes Soares, C. Data mining approach to shipping route characterization and anomaly detection based on AIS data. *Ocean Eng.* **2020**, *198*, 106936. [CrossRef]
11. Ahn, J.-H.; Rhee, K.; You, Y. A study on the collision avoidance of a ship using neural networks and fuzzy logic. *Appl. Ocean Res.* **2012**, *37*, 162–173. [CrossRef]
12. Gao, M.; Shi, G.-Y. Ship-handling behavior pattern recognition using AIS sub-trajectory clustering analysis based on the T-SNE and spectral clustering algorithms. *Ocean Eng.* **2020**, *205*, 106919. [CrossRef]
13. Xu, H.; Rong, H.; Guedes Soares, C. Use of AIS data for guidance and control of path-following autonomous vessels. *Ocean Eng.* **2019**, *194*, 106635. [CrossRef]

14. Tu, E.; Zhang, G.; Mao, S.; Rachmawati, L.; Huang, G. Modeling Historical AIS Data for Vessel Path Prediction: A Comprehensive Treatment. *arXiv* **2020**, arXiv:2001.01592.

15. Borkowski, P. The Ship Movement Trajectory Prediction Algorithm Using Navigational Data Fusion. *Sensors* **2017**, *17*, 1432. [CrossRef]

16. Tang, H.; Yin, Y.; Shen, H. A model for vessel trajectory prediction based on long short-term memory neural network. *J. Mar. Eng. Technol.* **2019**, 18. [CrossRef]

17. Zheng, H.; Negenborn, R.R.; Lodewijks, G. Trajectory tracking of autonomous vessels using model predictive control. *IFAC Proc. Vol.* **2014**, *47*, 8812–8818. [CrossRef]

18. Xiao, Z.; Fu, X.; Zhang, L.; Goh, R.S.M. Traffic Pattern Mining and Forecasting Technologies in Maritime Traffic Service Networks: A Comprehensive Survey. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 1796–1825. [CrossRef]

19. Xiao, Z.; Fu, X.; Zhang, L.; Zhang, W.; Liu, R.W.; Liu, Z.; Goh, R.S. Big Data Driven Vessel Trajectory and Navigating State Prediction With Adaptive Learning, Motion Modeling and Particle Filtering Techniques. *IEEE Trans. Intell. Transp. Syst.* **2020**. [CrossRef]

20. Patroumpas, K.; Alevizos, E.; Artikis, A.; Vodas, M.; Pelekis, N.; Theodoridis, Y. Online event recognition from moving vessel trajectories. *Geoinformatica* **2017**, *21*, 389–427. [CrossRef]

21. GeoStar Navigation. GEOS Module. NMEA Communications Protocol Ver. 3.0. Available online: http://geostar-navi.com/files/docs/geos3/geos_nmea_protocol_v3_0_eng.pdf (accessed on 25 December 2020).