Statistics Driven Suspicious Event Detection of Fishing Vessels Based on AIS Data

Harikumar Radhakrishnan, Saikat Bank, R Bharath and C P Ramanarayanan

February 16, 2023

Statistics Driven Suspicious Event Detection of Fishing Vessels based on AIS Data.

Hari Kumar Radhakrishnan[1], Saikat Bank[2], R Bharath[3] and Ramanarayanan C. P [4]

[1]Indian Navy, India,
[2,3] School of Computer Engineering and Mathematical Sciences,
Defence Institute of Advanced Technology, Pune, Maharashtra-411025, India,
saikatbank101@gmail.com, rambharath@diat.ac.in
[4]Defence Institute of Advanced Technology, Pune, Maharashtra-411025, India, vc@diat.ac.in

Abstract:

Fishing vessels have been widely used in contraband activities and are also highly vulnerable to accidents, malfunction of engines, etc. Fishing vessels are also reported in incidents by roaming in restricted border areas raising tensions across the nations. Timely monitoring and tracking of the fishing vessels will be needed such that it can improve the vigilance on the fishing vessel in contraband activities, provide rescue in case of accidents, malfunction of the vessel, or alarm in the case of sailing in restricted areas. In this paper, we propose an automated algorithm to detect any suspicious activity of the fishing vessel in real-time such that an alarm is issued to concerned authorities to take the necessary action. The algorithm is based on how frequently fishing vessels transmit the Automatic Identification System (AIS) data. Monitoring the fishing vessels all the time is not necessary and is also likely infeasible. Knowing the limitation, we propose a statistics-driven threshold, based on which we can reduce the instances for which we have to give attention to the fishing vessels.

Keywords: AIS data, Fishing vessels, Vessel type, marine traffic monitoring, statistics, box plots

1. Introduction:

Commercial Fishing is one of the most active marine activities associated with contraband activities, accidents, illegal fishing, etc, hence demands more vigilance. Fishing activities are done in imminence working conditions arise due to long working hours of the crew, operating in harsh weather conditions, etc. The accidents of fishing vessels constitute one third of the accidents according to the report of the transportation safety board of Canada [1]. According to the data, almost eighty-seven fishing vessels accident happens per year. From Figure 1, inference can made that fishing vessels are more prone to accidents compared to other vessels like Barge, Cargo, Ferry, etc. Also, The National Institute for Occupational Safety and Health (NIOSH) which maintains the Commercial Fishing Incident Database (CFID) of the United States reported that from year 2000 to year 2015, says that seven hundred and twenty-five fishermen died while fishing in the United States due to various accidents [1]. On the other side, illegal, unreported and unregulated fishing is also a major global problem. According to

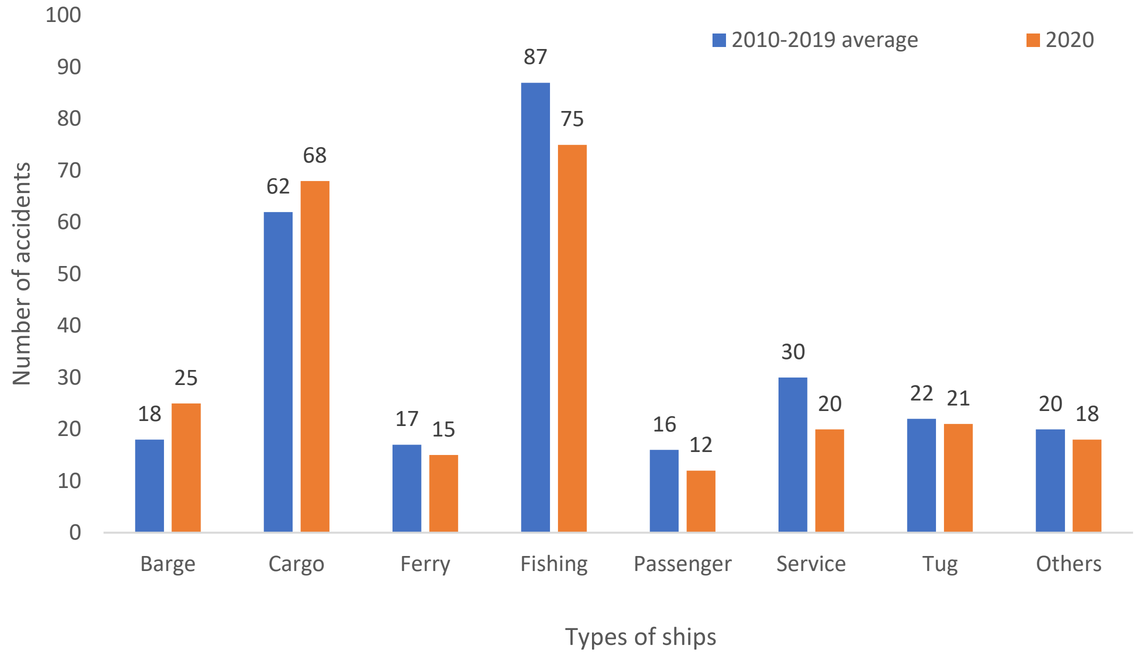


Figure 1: The number of accidents occurred from year 2010-2020 [1]

United Nations, Food and Agriculture Organization (FAO) reported a loss of 11-26 million tonnes of fish per year which has an estimated economic value of US$ 10-23 billion because of illegal, unreported and unregulated fishing activities [2].

To improve maritime monitoring, most ships are equipped with an Automatic Identification System (AIS) transponder, which continuously transmits the data about the ship to regional or national data centres [3]. The transmitted AIS data will be almost real-time, which will help to monitor and track the vessels [4]. According to International Maritime Regulations, all fishing vessels over fifteen meters should have onboard AIS transponders and maintain transmission of AIS data at all times, except where the international agreements, rules and standards provide for the protection of navigational information [5]. The AIS data mainly consists of seventeen features, namely Maritime Mobile Service Identity (MMSI), Base Date Time, Vessel type, Latitude, Longitude, Speed Over Ground (SOG), Course Over Ground (COG), Heading, Vessel Name, IMO (International Maritime Organization), Call Sign, Status, Length, Width, Draft, Cargo, Transceiver Class. Among them, MMSI, Vessel type, Vessel Name, Length, and Width are static data. MMSI (Maritime Mobile Service Identity) is the unique nine-digit integer id assigned to identify a vessel type; Vessel type assigns an integer-valued number to understand the type of vessel (for a fishing vessel, vessel type is 30, cargo ship 70, passenger ship 60 etc.) [6]. The field Base Date Time provides the date and time of receiving AIS data from a particular ship. Latitude, Longitude, SOG and COG will vary with time. Latitude and Longitude together give the position of the ship at the corresponding time stamp. SOG and COG give the velocity and direction respectively of the ship at the corresponding timestamp. With AIS data, maritime authorities can get the navigation status of a ship. The Centre for Maritime Research and Experimentation (CMRE) receives an average rate of 600 million AIS messages per month, and the rate is increasing day by day [7]. So, it is a very challenging task to keep real-time tracking of the vessels manually. Monitoring and tracking of all the ships even for a specific region of interest are highly complicated, seeing this as an opportunity, some fishing vessels are involved in contraband activities such as smuggling, trafficking, transporting weapons, drugs, etc. It is also reported that the fishing vessels ended up fishing in restricted and undesignated cross-border areas leading to inter-border tensions across the nations [8] [9].

Understanding the data timely is an important task to prevent any illegal activity or accident. If some ship is not sending the AIS data, then there is a high probability that the fishing vessel may involve in contraband activity, or the fishing vessel has met with an accident, malfunctioning equipment, etc. In all such scenarios, there is a timely need to pay more attention to that particular fishing vessel at that moment. In such methods, there is a need for an automated system to alert the control room such that the concerned authorities of the nearby base station can contact the ship in case of need or put more vigilance on the ship in case of contraband activity. As there are numerous fishing vessels sailing simultaneously, continuous monitoring and tracking the fishing vessels manually are highly infeasible, there is a need for automated algorithms to assist in detecting and monitoring the activities of the fishing vessels.

Marzuki et al. proposed an abnormal fishing vessel behaviour based on vessel trajectories [10]. They employed unsupervised learning machine learning algorithms to extract the characteristics of trajectory patterns and after then they employed random forest and support vector machines to classify abnormal pattern. Castaldo et.al trained a behaviour/interaction model based on trajectory data of ships, and identified abnormalities using FishNET which is a CNN based model [11]. Harikumar et al. proposed a semi-supervised approach to label the data and classify the features of AIS data into normal and abnormal using a support vector machine classifier [12]. In this paper, we took a different approach for detecting the suspicious event detection based on how frequent the fishing vessel transmits the AIS data.

The rest of the paper is organized in the following way. Section 2 discusses the proposed approach used for detecting the abnormal activity of the fishing vessel. The AIS data used in the analysis and its characteristics are discussed in Section 3. The results of the proposed approach are discussed in Section 4, and Section 5 concludes the paper.

2. Statistics-Driven Approach for Suspicious Event Detection

In this paper, we propose an approach for identifying the instances where attention has to be given to the ship. By doing this, we eliminate the burden of surveillance of the vessels all over the course of travel.

Tracking the abnormal event detection of fishing vessels is challenging due to its peculiar characteristics compared to other vessels like cargo, tankers, ferry, etc. Some of the characteristics of the fishing vessel are: it follows very irregular trajectories hence it does not follow any fixed path. Also, the time intervals between transmissions of two consecutive AIS data samples are also irregular. Due to this irregular behaviour, there are more opportunities for fishing vessels to escape from the surveillance system and participate in contraband activity. In case of an accident of some ships, sending help may also be delayed. So, detecting anomalies based on how frequent it transmits the AIS data carries a lot of importance, and is also equally challenging due to the irregularity of time intervals between two consecutive AIS data. In this paper, we propose to identify the anomaly based on the time difference associated with two successive AIS data reception. To identify anomalies based on how frequently the vessel transmits the AIS data, we proposed the thresholds for the time difference between the AIS data transmission. We will use these threshold values of time difference to identify the anomaly. If the fishing vessel is not transmitting the AIS data before the pre-set threshold values, then the system can raise a warning alarm for concerned authorities of the base station demanding attention from the authorities to monitor the fishing vessel.

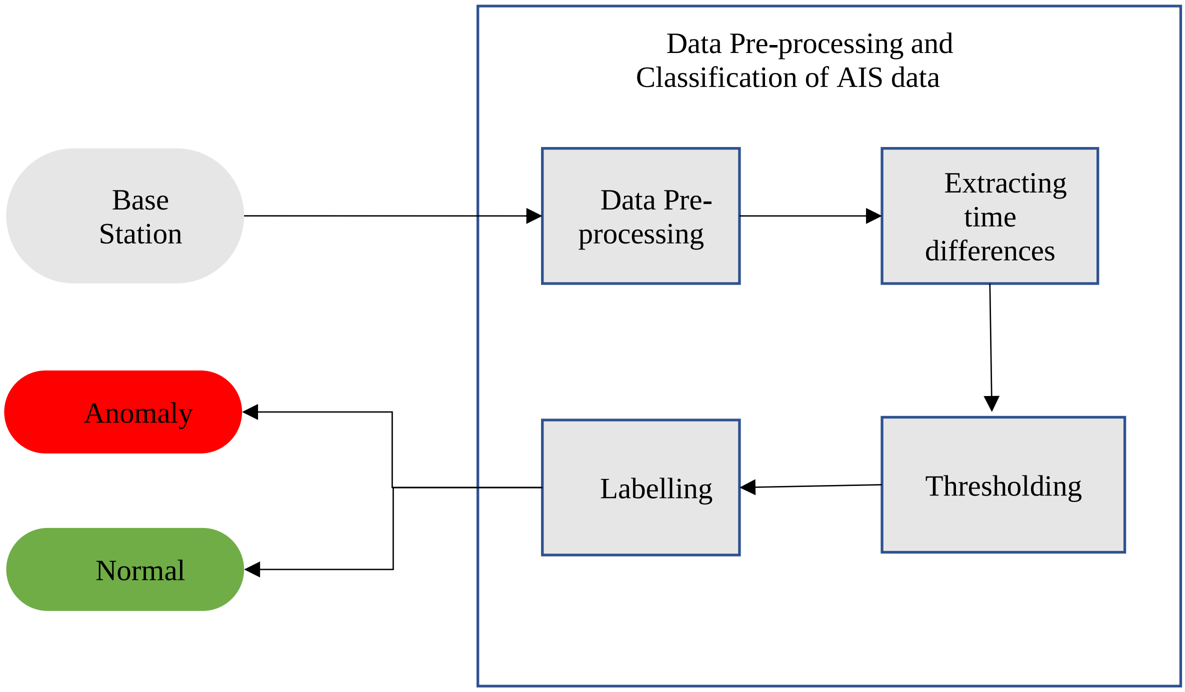


Figure 2: The basic building blocks of the proposed algorithm in detecting the suspicious activity of the fishing vessels

The basic building blocks of the proposed algorithm is shown in Figure 2. The base station receives the AIS data from different vessels simultaneously. Since, we are interested only in fishing vessels from the database, the fishing vessels got separated based on the MMSI number. The data collected are not timely aggregated, and hence the data will be arranged sequentially based on time for further processing.

3. Description of AIS Data used for the analysis

In this work, we used AIS maritime data downloaded from https://marinecadastre.gov/ [13]. The analysis is carried on fifteen days of data, recorded from 1/1/2022 to 15/1/2022. This AIS data consisted of 25948 unique MMSI Ids and 84 unique vessel types. Vessel type 30 represents the fishing vessels. Based on the vessel type, we extracted the data corresponding only to fishing vessels. The time difference between the transmission of two consecutive AIS data is very irregular. If the fishing vessel is not transmitting the AIS data frequently, then there is a chance that the ship may be met with an accident, or there is a malfunction of the AIS transponder, or it is involved in the contraband activity. So, a study of the time difference of fishing vessel AIS data is required to find some threshold which can differentiate between the normal and abnormal activity of fishing vessels. If we choose the threshold very high, then we will miss some anomaly activities, and if choose the threshold very low then the number of outliers will be very high which is not expected. So, some optimal point is required which can address both problems. To compute the threshold, we conducted statistical data analysis on AIS data which will provide some threshold to detect the anomaly. Since we are interested in studying how frequent the fishing vessels transmit the AIS data, will take only MMSI, Base Date Time (Time Stamp), Latitude, Longitude, and Vessel type column from the AIS data to extract the time differences. These time differences have units in seconds. Then we extracted all the data points which are corresponding to vessel type 30 from AIS database. The data is extracted based on MMSI Id. Among all MMSI IDs, we have taken those MMSI IDs which has at least 1000 AIS samples. By doing this, we got a total of 190 unique fishing vessels on which we have done our experiments. After extracting the data, we transformed Base Date Time columns into two new columns for date and time and then sorted first based on the date and then based on time to arrange the data sequentially. Then we found the time difference between two consecutive times. After sequentially arranging the data, the statistical distribution of the time difference of AIS data of the MMSI is analysed. The flowchart of the proposed algorithm is shown in Figure 3.

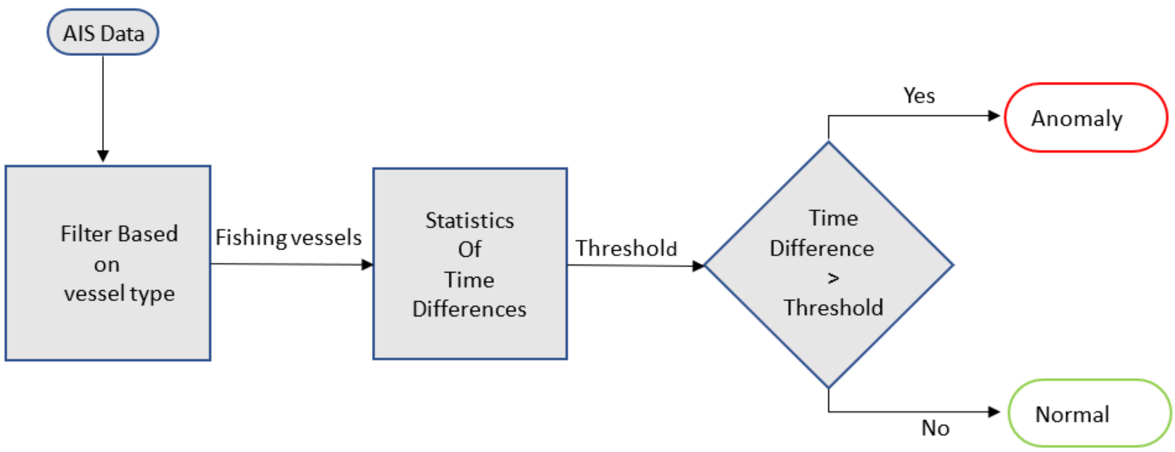


Figure 3: The flowchart of the proposed algorithm.

4. Results

The statistical distribution of the time difference of the statistics for the ten MMSI IDs is shown in Table 1. From there we can see ships are irregular although some vessels are sharing the same kind of statistics. The statistical distribution of the time difference of data across all 190 MMSI IDs is shown in Table. 2

In Table 1, 50 percentile values indicate the median of the time differences of the corresponding MMSI IDs. Similarly, 75 percentile and 90 percentile values of the time differences of the corresponding MMSI IDs. Upper Bound (UB) is the statistical quantity Q3 + (1.5*IQR) where IQR = Q3 – Q1, Q3 is the 3rd quartile and Q1 is the 1st quartile. As the unit of the time difference is seconds, the same thing follows by all percentiles and UB. If we take 50, 75, and 90 percentile values or UB as a threshold then the percentage of outliers is shown in the Outlier column. If we keep the threshold around 90 %, we can infer that 10% of the instances are marked as anomalies. If we keep the threshold with UB, then the percentage of outliers lies between 4.8% to 20.3%. Inferring from Table. 1, we can also infer that the time difference between the transmission of AIS data is irregular.

Table 1: Distribution of Time difference in transmission of AIS data for first 10 MMSI IDs

MMSI	**Total No. Of Data samples**	**Percentile Values (Seconds)**			**Upper Bound (UB) (Seconds)**	**Percentage of samples labelled as Outlier**			
		50	**75**	**90**		**50**	**75**	**90**	**UB**
1	3693	181	360	539	630	40.3	18.2	9.5	4.8
2	5748	90	120	180	168	43.6	21.0	8.8	13.5
3	1357	90	91	120	94	25.2	16.0	9.5	15.9
4	2854	119	180	360	315	49.8	24.8	9.4	12.1
5	18667	70	70	71	71.5	22.1	22.1	4.9	4.9
6	1462	70	80	100	96.5	47.2	20.3	9.3	11.2
7	1945	180	181	184	316	26.9	15.2	9.9	9.4
8	5127	180	182	361	186.5	42.7	22.5	7.2	20.3
9	3433	180	180	181	181.5	23.1	23.0	6.7	6.7
10	2476	178	181	360	317.5	47.7	23.2	8.6	13.6

Based on the inference from Table. 1, to obtain the appropriate threshold which work for every ship, we computed the distribution of percentile values for all the fishing vessels. The distribution of the percentile values is shown in Table. 2.

We have chosen six suitable numbers from Table 2. Among them, 360 seconds and 361 seconds are very close to each other we discarded 360 seconds and computed the percentage of anomalies labelled by the proposed algorithm for all the thresholds such as 182.75, 317.5, 361, 541 and 630 seconds. The proposed algorithm marked the abnormalities from 12.03 to 2.12% of the samples, respectively. The outlier percentages when we checked for all 351 fishing vessels and obtained results are shown in Table. 3.

The percentage of samples labelled as anomaly or outlier by the proposed algorithm across each 10 MMSI of Table. 1 is shown in Table. 3. The outlier percentages have been computed for all 351 fishing vessels.

Table 2: Statistical distribution of percentile of percentile values across MMSI

Statistics	Percentile values (Seconds)			Upper Bound (UB) = Q3 + (1.5 * IQR) (Seconds)
	50	75	90	
Min	62	66	70	71.50
25%	88	91.25	180	133.75
50%	179	181	240	186.50
75%	180	182.75	361	317.50
90%	181	360	541	630
Max	360	900	2412	1980

Table 3: The percentage of samples labelled as anomalies for different thresholds

Percentage of anomalies					
Thresholds	**182.75**	**317.5**	**361**	**541**	**630**
Outlier Percentage	12.0	8.7	4.4	2.4	2.1

The map-based ship movement plots shown in Figure 4 for three thresholds will make more sense of the anomalies. The number of anomalies is changing with respect to different thresholds can be seen in the Figure 4. Like, when threshold is 182.75, the number of outliers is higher than the threshold 630.

The normal and anomaly points labelled by the proposed algorithm for different thresholds are shown in Figure 4. Red points correspond to anomalies and the blue points correspond to normal points.

Table 4: Percentage of anomalies corresponding to various thresholds

MMSI	Total No. of Data Samples	Outlier % if threshold is				
		182.75	317.5	361	541	630
1	3693	35.5	33.0	13.9	5.8	4.8
2	5748	7.1	4.6	3.1	1.8	1.5
3	1357	4.7	3.1	2.1	1.4	1.4
4	2854	22.2	12.1	9.1	5.5	4.6
5	18667	0.13	0.1	0.1	0.1	0.1
6	1462	2.7	0.9	0.8	0.6	0.5
7	1945	11.2	9.4	4.6	2.9	2.8
8	5127	22.5	19.9	7.2	2.5	2.2
9	3433	6.6	5.6	1.7	0.7	0.7
10	2476	20.7	3.6	7.4	3.8	3.4

4. Conclusion

In this paper, we proposed a statistical-based approach for detecting the normal and abnormal instances of fishing vessels in their voyage. All-time manual tracking of the fishing vessels on their voyage is a tedious task, and hence the proposed approach will reduce the number of instances to a finite number where attention has to be given. From the statistical nature of the time difference in the transmission of the AIS data, a selection of the lower threshold will result in a greater number of anomalies, and having the higher threshold will result in a lower number of anomalies. The threshold should be selected in such a way that, there will be a lesser number of outliers and simultaneously, we should not miss out on the suspicious events.

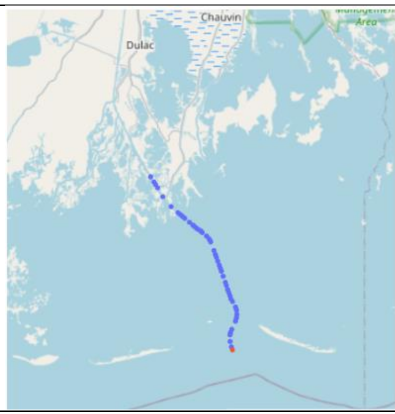
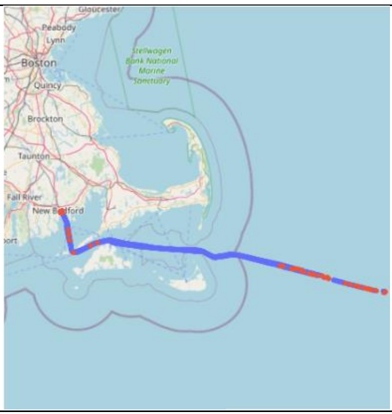
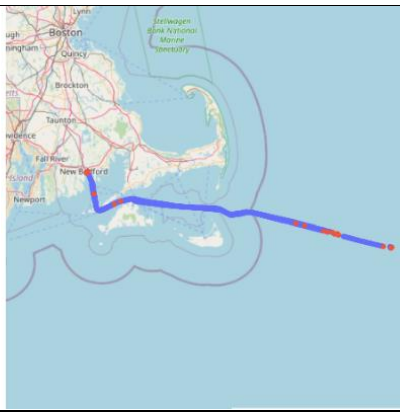
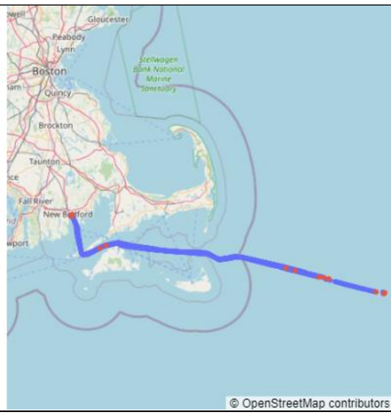
MMSI	Thresholds (Seconds)		
	182.75	361	630
1			
2			

Figure 4: Anomalies detected by the proposed algorithm for different thresholds. Blue dots represent the normal samples and the red dot represents the instances labelled as the anomaly by the proposed algorithm.

References

[1] "Commercial Fishing Safety: National Overview," [Online]. Available: https://www.cdc.gov/niosh/topics/fishing/nationaloverview.html.

[2] "International Day for the Fight against Illegal, Unreported and Unregulated Fishing," [Online]. Available: https://www.un.org/en/observances/end-illegal-fishing-day.

[3] T. Zhang, S. Zhao, B. Cheng, and J. Chen, "Detection of AIS Closing Behavior and MMSI Spoofing Behavior of Ships Based on Spatiotemporal Data," *Remote Sensing,* vol. 12, no. 4, p. 702, 2020.

[4] "Marine Traffic - the most popular online service for vessel tracking | AIS Marine Traffic," [Online]. Available: https://www.marinetraffic.com/en/ais/home/.

[5] "Automatic Identification Systems (AIS)", International Maritime Organization," [Online]. Available: https://www.imo.org/en/OurWork/Safety/Pages/AIS.aspx.

[6] "AIS Ship Types," [Online]. Available: https://api.vtexplorer.com/docs/ref-aistypes.html.

[7] Cimino, G.; Ancieri, G.; Horn, S.; Bryan, K, "Sensor Data Management to Achieve Information Superiority in Maritime Situational Awareness," CMRE Formal Report, NATO Unclassified, Brussels, Belgium, 2013.

[8] Fiorini, M.; Capata, A.; Bloisi, D.D, "AIS data visualization for Maritime Spatial Planning (MSP)," *International Journal of e-Navigation and Maritime Economy,* vol. 5, pp. 45-60, 2016.

[9] Goverment of Canda, "Illegal, Unreported and Unregulated (IUU) Fishing," [Online]. Available: https://www.dfo-mpo.gc.ca/international/isu-iuu-eng.htm.

[10] Marzuki, M.I.; Gaspar, P.; Garello, R.; Kerbaol, V.; Fablet, R., "Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories," *IEEE Journal of Oceanic Engineering,* vol. 43, pp. 689-699, 2017.

[11] Castaldo, F.; Palmieri, F.A.; Regazzoni, C.S., "Bayesian analysis of behaviors and interactions for situation awareness in transportation systems," *IEEE Transactions on Intelligent Transportation Systems,* vol. 17, pp. 313-322, 2015.

[12] Radhakrishnan, Hari Kumar, C. P. Ramanarayanan, and R. Bharath., "Machine Learning Based Automated Process for Predicting the Anomaly in AIS Data," in *International Conference on Data Management, Analytics & Innovation, Springer*, Singapore, 2023.

[13] "marinecadastre.gov," [Online]. Available: https://marinecadastre.gov/.

[14] "Marine transportation occurrences in 2020," [Online]. Available: https://www.bst-tsb.gc.ca/eng/stats/marine/2020/ssem-ssmo-2020.html.

[15] S. Arasteh, M. A. Tayebi, Z. Zohrevand, U. Glässer, A. Y. Shahir, P. Saeedi, and H. Wehn, "Fishing Vessels Activity Detection from Longitudinal AIS Data (Industrial Paper)," in *28th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '20)*, New York, NY, USA, 2020.

[16] Jiang X, Silver DL, Hu B, Souza EN, Matwin S., "Fishing activity detection from ais data using autoencoders," in *InCanadian Conference on Artificial Intelligence, Springer*, Cham, 2016.