International Journal of Advances in Data and Information Systems
Vol. 6, No. 1, April 2025, pp. 107~118
ISSN: 2721-3056, DOI: 10.59395/ijadis.v6i1.1367

Grid-Based Ship Density Analysis and Anomaly Detection for Ship Movements Monitoring at Tanjung Priok Port

Muhammad Ramadhan Ikhsan[1], Panca Dewi Pamungkasari[1], Babag Purbantoro[2], Ira Diana Sholihati[1], Frenda Farahdinna[1], Josaphat Tetuko Sri Sumantyo[3,4], Damy Matheus Heezen[5]
[1]Information System, Faculty of Communication and Information Technology, National University, Indonesia
[2]Research Center for Geoinformatics, National Research and Innovation Agency (BRIN), Indonesia
[3]Center for Environmental Remote Sensing, Chiba University, Chiba, Japan
[4]Department of Electrical Engineering, Universitas Sebelas Maret, Indonesia
[5]Changemakers Network B.V., Netherlands

Article Info

Article history:

Received Mar 02, 2025
Revised Apr 11, 2025
Accepted Apr 30, 2025

Keywords:

Maritime
Ship
Port
AIS
Anomaly

ABSTRACT

Indonesia, as a maritime country, depends on ports to support inter-island transport and a smooth regional economy. So, the awareness of knowing the marine status with various platforms is needed. This research distinguishes itself from several previous studies on ship movement detection by concentrating specifically on anomalies in ship movement within areas of high traffic density. This research proposes to find out the ship density area using the grid technique and identify the anomalies that have occurred, as information on ship movements at Tanjung Priok Port. Anomaly detection is done by looking for it through visualization, where AIS data is converted into a form of visualization using the Python language. The results obtained two pieces of information, namely that the areas with the highest density are around the harbor, docks, and ship lanes. Then, two types of anomalies were detected, namely large ships with dangerous cargo speeding in dense areas and ships that behave differently compared to other ships with the same status.

Corresponding Author:

Panca Dewi Pamungkasari,
Information System, Faculty of Communication and Information Technology,
National University,
Sawo Manila Road, Pejaten, Pasar Minggu, South Jakarta, Special Capital Region of Jakarta 12520.
Email: panca.dewi79@gmail.com

1. INTRODUCTION

Indonesia is a maritime country, according to information from the Coordinating Ministry for Maritime Affairs, covering a total area of 5,2 million km², which includes a land area of 1,92 million km² and a water area of 3,1 million km² [1]. Indonesia has the largest archipelago in the world, with more than 17,000 islands. Nonetheless, this vast expanse of sea poses many problems, including maritime security. Maritime security is fundamental to maintaining Indonesia's sovereignty and prosperity as a maritime nation [2]. Therefore, strict surveillance and strong maritime regulations are essential to protect marine resources and ensure optimal utilization of maritime wealth for the national interest. In addition to the defense component, maritime security relates to shipping safety and the protection of the marine environment. Indonesia is a major trade route traveled by many foreign ships [1], therefore, it is important to ensure the waterways remain safe from accidents or illegal activities. Maritime security is a shared responsibility between governmental authorities and coastal residents, including those involved in fishing and other marine activities, as a form of awareness of the importance of protecting the oceans. Maritime security represents a primary objective for Indonesia, considering that a significant portion of its territory consists of water and it shares borders with several other countries [2]. The waters of Indonesia serve as a significant international trade gateway, highlighting the imperative for rigorous surveillance methodologies, enhanced monitoring technologies, and cooperative efforts with neighboring countries to uphold maritime stability. Through maintaining its aquatic environments, Indonesia is actively protecting its natural resources while also ensuring the integrity of its sovereignty and the sustainability of the marine sector, which serves as a fundamental component of the national economy.

As a nation characterized by its maritime activities, ports play a crucial role in the transportation network of the archipelago. Understanding the condition of marine environments through diverse platforms is essential. The analysis of ship movement holds significant importance across multiple domains, particularly concerning maritime security, the management of marine traffic, and the protection of marine environments. The detection of anomalies in ship movement is crucial for maritime security, as it aids in identifying potential illicit activities, including smuggling, piracy, or violations of territorial boundaries.

Numerous previous studies have been conducted that are intricately linked to the significance of maritime monitoring [3], [4], [5], [6]. Automatic Identification System (AIS) was introduced in maritime to improve the security of sea traffic [3], [4]. AIS data has been utilized for the purpose of mapping maritime activities [7] and evaluating their effects on global environmental conditions [8]. AIS messages are sent as broadcasts to nearby ships and provide information about the ship's ID, position, speed, and course of the sending ship [9]. AIS is an international standard for communication between ships and ground stations developed to improve maritime safety, which will help ships avoid collisions by assisting Vessel Traffic Services (VTS) in controlling ships sailing near certain coasts and harbors [10]. Therefore, AIS has become an important system for assessing the risk of marine collisions [11][12] and conducting surveillance, including anti-piracy operations, or preventing illegal fishing [13].

To find out marine status or marine information, data visualization can be used, which can show the movement of ships to provide better insight into the movement of ships, and data is converted into information in a visual form so that it can be communicated effectively to the intended audience. Effectively conveying insights from data necessitates design proficiency, technological expertise, and experience. Data must be precisely represented through visually appealing graphics and compelling language to successfully convey the message to the target audience [14]. Effectively conveying information from data requires a creative integration of data visualization and storytelling techniques, connecting raw data with actionable insights and informed decisions [14]. Turning visualizations into data stories strengthens their communicative power, enabling audiences to better understand complex information, and know what to do and why [15].

Visualisations of AIS data offer insights into maritime traffic patterns, highlighting areas of high density. Dense locations are often monitored due to their vulnerability to threats, including collisions, congestion, and illegal activity. In addition, visual analysis can help identify anomalies, which are unusual patterns of ship movement, like ships that behave differently from other ships even though their status is the same, or a large ship carrying hazardous goods (cargo-hazardous type) speeding in a dense traffic area. Both of these are considered anomalies because these ships behave differently from other ships of the same status. Large ships should not be at high speed, especially in dense traffic areas with hazardous cargo types. Anomalies in AIS trajectories refer to behavior that deviates from the norm or, more precisely, is not expected during standard operations [16].

The identification of areas exhibiting high-density traffic can be achieved through the application of the grid method, which has previously been utilized in Tianjin Port to assess the density of ships within the port. The results show that the density of ships in Tianjin Port varies by location.

The harbor area has a high density, while the anchorage and lane areas show a lower density and a more uniform distribution [17].

Furthermore, the process of anomaly detection will be conducted by examining it through a visualization, wherein AIS data is transformed into a visual representation utilizing the Python programming language and the Matplotlib library. The objective of data visualization is to communicate data and visual information with enhanced clarity and efficiency by utilizing suitable graphical representations. Python represents a highly advantageous option, characterized by an abundance of third-party libraries, a robust open-source community, and continuously improved documentation specifically tailored for data visualization. Matplotlib, a library developed in Python, features a straightforward syntax, high rendering quality, and code that is easy to comprehend [18].

This study presents a novel approach by employing the grid-based method to ascertain ship density and subsequently identify anomalies in ship movement within the Tanjung Priok Port area. Through this research, the grid-based approach and anomaly detection are expected to contribute to community service, especially in supporting maritime monitoring by knowing the anomalies that have occurred and the density patterns in the Tanjung Priok Port area. The results of this research can serve as information for maritime agencies and fishing communities to improve shipping safety and reduce the risk of marine accidents.

2. RESEARCH METHOD

This research method aims to determine the area of ship density and detect ship movement anomalies based on AIS data as ship movement information. This information is useful to support maritime monitoring because marine inspectors can make a decision based on existing information. The approach includes several basic stages, starting from preprocessing, density analysis using Kernel Density Estimation (KDE) and grid-based techniques, to the discovery of ship movement anomalies. These stages are depicted in the research flow in Figure 1.

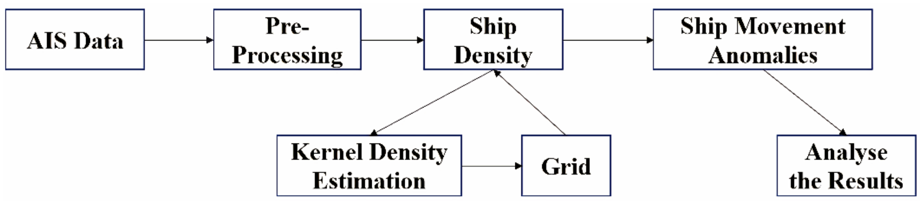

Figure 1. The process flow of the study

AIS Data

Automatic Identification System (AIS) data was obtained from Tanjung Priok harbor during the period 01-April-2021;00:00 to 01-April-2021;23:59. This data consists of information about the ship's position, speed, course, status, and many other columns. The raw data is not even in tabular form but already has column separators, so it is only necessary to create a table with these column separators. The AIS data obtained has many columns, and some data is null, so pre-processing is necessary.

AIS is a system that plays an important role in marine navigation and shipping, which allows ships to automatically communicate information using VHF radio waves [4]. The premise of AIS operation involves the use of transponders that receive and communicate information about a ship's identity, geographical position, speed, and direction of motion. Information provided through AIS data includes ship identity, navigation status, and additional information such as cargo and destination.

Standards and regulations related to the use of AIS are set by international groups such as the International Maritime Organization (IMO) to improve maritime safety and navigation efficiency, both regionally and worldwide.

These include the requirements for the use of AIS in the International Convention for the Safety of Life at Sea (SOLAS) as well as technical standards published by groups such as the International Electrotechnical Commission (IEC). The law also affects the use of AIS in ports, where the system is used to control ship traffic, guide ships to the right berth, and monitor safety on the water.

Data Pre-Processing

The pre-processing phase represents the initial stage in the analysis of AIS data prior to the next steps. At this stage, critical columns for ship movement information, namely ship ID / Maritime Mobile Service Identity (MMSI) and ship location (Longitude and Latitude), are examined for the presence of nan / null values in the AIS data. If such values are detected, those entries must be removed. Additionally, since the ship name column is redundant, as it is already represented by the MMSI column, this column is also eliminated to streamline data processing. Prior to the removal of the column, the relevant information within the dataset will be organized into a table or visualization. This will include data pertaining to ship type, status, and visual representations indicating the location of the port.

Before pre-processing, information retrieval from the data is first carried out such as creating tables and visualizations as a reference such as creating a table 'Ship Type', and 'status', an example table can be seen in Table 1 and a visual showing the location of the Port, where the location of the port of Tanjung Priok is at lat: 6.1 and lon: 106.9 can be seen in Figure 2. After that, pre-processing is carried out such as eliminating columns that are not needed, ensuring that the MMSI / Ship ID, longitude, and latitude columns are not null.

Table 1. ShipType and Status Code

Shiptype	Shiptype Text	Status	Status Text
80	Tanker - all ships of this type	0	Under way using engine
53.0	Port Tender	9	NA
70.0	Cargo - all ships of this type	1	At Anchor
60.0	Passenger - all ships of this type	12	Reserved
0.0	Not Available	2	Not under command
52.0	Tug	8	Under way sailing
72.0	Cargo - Hazardous Category B	5	Moored
30.0	Fishing	11	Reserved

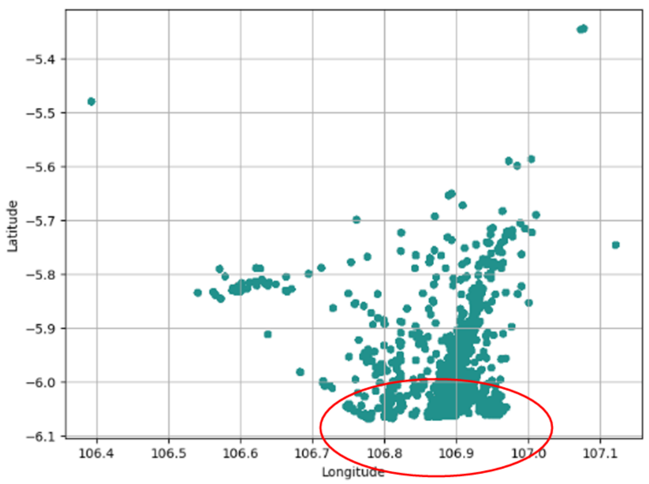

Figure 2. Visual representations indicating degree coordinates that specify the geographical location of the harbor (highlighted within the red circle)

During the Pre-Processing stage, we found the presence of empty, null, or NA rows within the ship type and status columns. Given that the important columns are MMSI, latitude, and longitude, we chose to keep the empty or null rows. It is not feasible to impute the null values, as doing so would result in an error due to the potential mismatch in data types for the values intended for entry.

Drop Columns/using only the required columns = Reduce storage space, this theory is proven in the grid-based AIS data storage method conducted by Yang et al., where the storage method only stores information needed to calculate ship density/stores only a few columns such as MMSI, Region, Enter Time, and Time. The grid-based AIS data storage method can save storage space since it keeps information that is necessary for determining ship density and ignores ship navigation details in the grid [17].

Ship Density

Ship density is built in two stages, first with KDE to create a density heatmap, and then a grid technique to visualize the density so that dense areas can be seen more clearly. Ship density and its distribution are fundamental notions that represent the actual state of marine traffic in a given water area, indicating the amount of activity and risk associated with maritime navigation. Ship density indicates the economic prosperity of the maritime industry [17]. Areas of high density are often monitored for their susceptibility to hazards, such as collisions, congestion, and illegal activities.

KDE is a statistical approach to estimate the probability density function of a data collection. KDE is a technique for probability density function estimation that is a must-have for users to assess the researched probability distribution better than when using typical histograms [19]. KDE is used to obtain the probability density function of ship density by selecting an appropriate kernel function and using an ideal bandwidth [17].

Table 2 shows the algorithm in which the coordinates were changed to Universal Transverse Mercator (UTM). UTM is a coordinate system that employs meter units, thereby improving grid resolution as it is characterized by whole numbers (e.g., 500 meters) rather than degree units. Thus, this approach has been termed a grid-based technique. The selected kernel function is Gaussiankde, as indicated by the findings of Huang et al. in their research. The Epanechnikov kernel exhibits the lowest degree of smoothness, whereas the Gaussian kernel demonstrates the highest degree of smoothness. Eight distinct kernels, namely Uniform, Triangular, Epanechnikov, Quartic, Triweight, Tricube, Gaussian, and Cosine, have been subjected to a comparative analysis in parallel [20].

Table 2. Step-by-step application algorithm of KDE

No	Algorithm 1
	Input: Kernel Density Estimation
	Result: Ship Density Heatmap
1	input: utm.coords = data[['lat', 'lon']]
2	kde = gaussian.kde(xy, bw.method=1.0)
3	density = kde(xy)
4	plt.figure
5	plt.show()

The bandwidth to be utilized is determined through the application of GridSearchCV. According to a study conducted in 2019 by Ranjan et al. [21], Grid SearchCV is a method that methodically constructs and evaluates models for every possible combination of algorithm parameters outlined in a grid. GridSearchCV is a function within Scikit-learn, a Python module that provides a comprehensive suite of sophisticated machine learning algorithms designed to address both supervised and unsupervised tasks at a medium scale [22]. Within the framework of KDE, GridSearchCV is employed due to its capability to methodically investigate various bandwidth values. Its comprehensive search across the bandwidth facilitates the identification of the most effective parameters [23]. The bandwidth values subjected to testing were 0.01, 1.0, 30, and 0.1. The values selected were random, as the research conducted by Yang et al. [17] indicates that the optimal choice of bandwidth remains an unresolved issue. Nonetheless, in numerous practical scenarios, the determination of a suitable bandwidth remains an unresolved issue. In practical applications, the suitable bandwidth may be determined based on empirical observations and specific contexts [17]. The optimal value determined by GridSearchCV is 1.0.

The grid organizes the data into square-shaped cells. Each grid represents the quantification of ships or the density level derived from KDE data. A grid measuring 500 meters was established, and to pinpoint areas of high density, a density threshold of 95% was implemented. Grid-based techniques are summarized in Table 3.

Ship Movement Anomalies

Anomalies in ship movements can signify illicit activity, security threats, or emergency conditions. Anomalies in this study relate to ship movements that are assumed to endanger surrounding ships and ships that behave differently from other ships, even though their status is the same. Anomalies in AIS trajectories refer to behavior that deviates from the norm or, more precisely, is not expected during standard operations[16]. Anomaly detection will be carried out by looking for it through a visualization where AIS data is converted into a visualized form with the Python language using the Matplotlib library.

Table 3. Step-by-step application algorithm of ship density

No	Algorithm 2
	Input: Grid-Based Technique
	Result: Ship Density Distribution with Grid Visualization
1	input: grid.size = 500 *(metre)
2	X, Y = np.meshgrid(x.grid, y.grid)
3	xy.grid = np.vstack([X.ravel(), Y.ravel()])
4	density.grid = kde(xy.grid).reshape(X.shape)
5	density.threshold = np.percentile(density.grid, 95)
6	dense.grid.mask = density.grid >= density.threshold
7	for x.center, y.center in zip(dense.grid.x, dense.grid.y):
8	rect = plt.Rectangle((x.center - grid.size / 2, y.center - grid.size / 2), grid.size, grid.size,
9	fill=False, edgecolor='red', linewidth=1.5)
10	plt.figure
11	plt.show()

The use of data visualization has become an attractive approach for deriving knowledge from massive ship itineraries. To enhance the monitoring of marine traffic, it is essential to thoroughly understand the vast ship traffic data from various analytical perspectives in practical applications. Advanced data visualization techniques facilitate the extraction of profound insights into marine traffic behavior derived from extensive ship traffic datasets [24].

A robust visualization system has the capability to disclose structures within vast datasets and to identify concealed relationships among extensive data succinctly, thereby enabling users to detect patterns and uncover anomalies [25]. Subsequently, the results were analyzed to validate the anomaly findings. The anomaly detection results are evaluated and validated using reference data/information captured prior to pre-processing to ensure the veracity and relevance of the results.

3. RESULTS AND DISCUSSION

3.1. Ship Density

Plt.figure() and plt.show(), frequently used in algorithms, serve as plotting functions from the Matplotlib library, which is widely recognized in Python for creating visualizations. The visualization of ship density can be seen in Figure 3.

Figure 3 illustrates that the area with the highest density is located in the Tanjung Priok port region, followed by the dock and ship channel areas. The results of this research align with those presented by Yang et al. [17] where the harbor area has a high density, while the dock and ship channel area shows a lower density and a more uniform distribution. When applied to the AIS data from other harbors, the results are likely to be similar, as ships tend to stop or anchor, creating an appearance of congestion, thereby consistently identifying the most congested regions. Despite establishing a ship density threshold of 95%, certain shipping lanes still overlap with congested regions.

As seen in Figure 3 a grid was successfully constructed using KDE to show the density of ship movements, where this visualization of ship density generates information that provides data on which areas have heavy traffic. Based on this information, maritime inspectors can subsequently determine which regions require additional patrol or surveillance, as these high-density zones are susceptible to risks such as collisions.

3.2. An anomaly of a large ship speeding in a dense area

After identifying congested regions, the concentration of ships serves as data to detect subsequent anomalies. The density of ships has been added to the visual anomaly analysis to determine if the anomaly exists within a congested area. However, the visibility remains unclear due to the absence of a grid format, which would facilitate a direct comparison between the identified anomalies and the ship density data. This comparison is essential for validating the observations regarding large ship anomalies operating at increased speeds in congested regions.

For instance, to identify the specified anomalies, the dataset undergoes processing utilizing a conditional filter to isolate instances where the speed exceeds 20.3, which represents the maximum speed recorded within the dataset. Subsequently, the MMSI or ship ID is added through an iterative function applied to each ship that meets the speed criterion. The results are then presented visually through a graphical representation generated with the Matplotlib library, employing the plt.show() command to render the visualization. The vessel identified by MMSI: 255806085, classified under ship type code 72 (Cargo - Hazardous Category B), exhibits an anomaly characterized by excessive speed in a congested maritime zone.

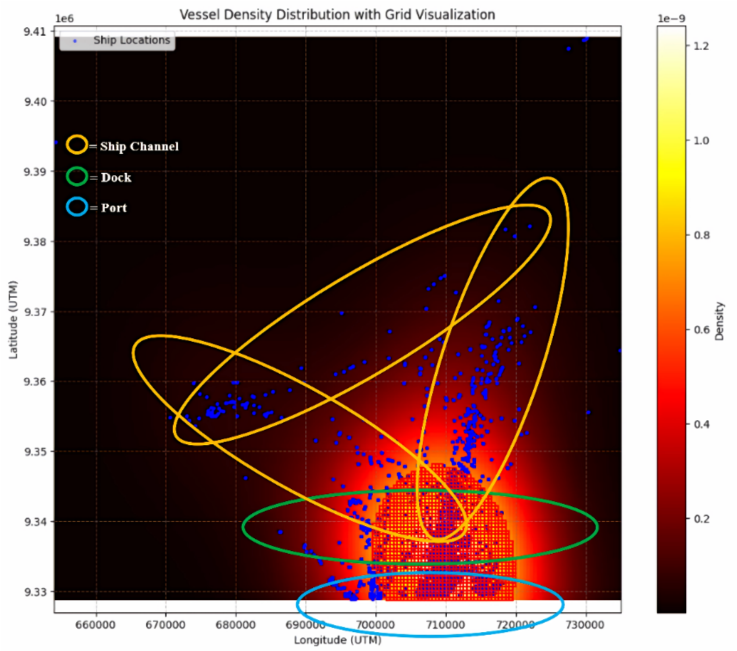

Figure 3. Visualization of ship density derived from the dataset. Blue dots represent the locations of ships. This site is located in Tanjung Priok Port.

Upon examination of the lookout table (reference), it was determined that this ship was operating at a high velocity (speed = 20.3, the maximum recorded in the dataset) and classified as a large ship type (cargo). Figure 4 illustrates the presence of a ship operating at a high velocity of 20.3, identified as a cargo ship. This particular cargo ship is transporting hazardous materials and navigating through a congested shipping lane. The implications of large ships carrying dangerous goods at elevated speeds in areas of heavy maritime traffic warrant further investigation, as there is a significant concern regarding the potential for incidents such as collisions.

The visual representation, specifically the graph depicted in Figure 4, indicates that this particular ship exhibits the highest recorded speed within the dataset at 20.3. This observation suggests that the ship is operating at an elevated velocity, as no other ships match this speed. In the reference data, speeds exceeding 18 are represented by only one ship at a speed of 19. Furthermore, this ship is classified as a Cargo-hazardous category B ship, which is substantial in size and transports dangerous goods. The implications of operating at high speeds in congested maritime traffic areas are significant, given the inherent risks associated with such conditions.

Based on the reference data that provided validation for the anomaly findings, the ship type identified is 72 (Cargo - hazardous category B) with a recorded speed of 20.3. In conjunction with the graph presented in Figure 4, this scenario can be classified as an anomaly involving a large ship transporting hazardous materials while exceeding speed limits in a congested region.

The anomaly was successfully found and validated, and this visualization generates data on anomalies that have occurred, which maritime supervisors can use to warn ships that speed in crowded areas because they are prone to collisions. This also raises awareness of potentially dangerous ships.

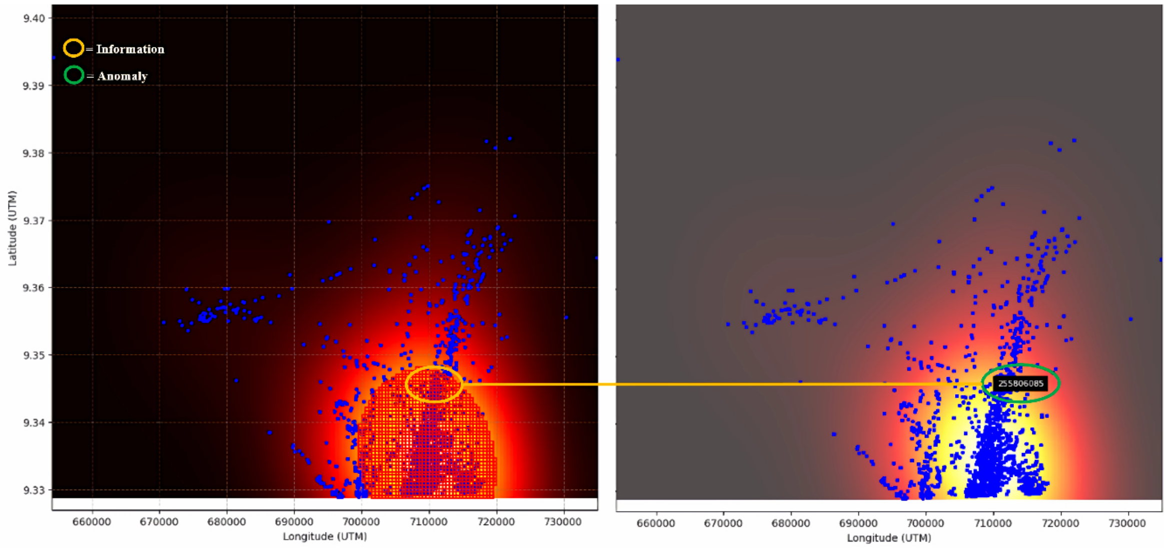

Figure 4. An unusual occurrence involving a large ship navigating at high velocity within a dense region

3.3. Anomaly Ship sails but does not update status

At the pre-processing stage, instances of null data were detected in the status variable where these data were retained (not dropped). This decision was taken based on the presence of non-null longitude, latitude, and MMSI data, indicating that the dataset retained its significance and should be retained. To identify the specified anomalies, the data is processed using the 'where' filter on the database command, specifically focusing on instances where the status is status.isnull(). Next, the MMSI or ship ID identification label is integrated via the 'for' function for all ships classified in status.isnull(). The resulting data will then be displayed in a visual format created using the Matplotlib library, using the plt.show() command to generate visualizations.

The MMSI of 525022166 is associated with a ship classified under ship type code 70, which indicates that the vessel is a cargo ship. The recorded speed was 6.9, while the status was recorded as NA, indicating a lack of data or information. The vessels exhibited abnormal sailing behavior without updating their status. Some ships did start sailing, but most of the ships with this null status were around the harbor. As seen in Figure 5, most of the ships that have not updated their status (nan/null ship status) are still inside the harbor and although some have started to move away from the harbour, the resulting visual is not very striking, in contrast to ships with MMSI=525022166 which look different in position / away from ships that have not updated their status.

According to the reference data used to validate the anomaly findings, the ship status is recorded as NA (nan/null (No Data/Information)). In addition, the graph presented in Figure 5 illustrates that this ship exhibits different behaviors independently. Therefore, this condition can be classified as an anomaly.

Anomalies are successfully identified and confirmed, with visualization of these anomalies generating information indicating the location of the anomaly. Upon detecting unusual ship behavior, operating differently from others despite having the same status, maritime inspectors can then choose to increase surveillance. These events signal unusual activity that requires further observation.

3.4. Anomaly of ship stopping at different places

After employing the highest speed data, the subsequent analysis will utilize the lowest speed or stationary ship data. Incorporating status data (status = 1, at anchor) is crucial for the identification of anomalies in ship movement. To identify the specified anomalies, the dataset is processed using a filter that sets speed = 0 (indicating ships that have stopped moving) and status = 1 (indicating ships that have come to a halt by using anchors). Following this, the MMSI or ship identification label is integrated via the application of a function for each vessel classified under a speed of zero. The resulting data is subsequently illustrated through a graphical representation generated using the Matplotlib library, utilizing the plt.show() command to display the visualization. In order to identify the ships in motion, supplementary tools are utilized, particularly Matlab from MathWorks, which facilitates interactive visualization. This enables the retrieval of data from the scatter plot, especially concerning points near the stationary ship.

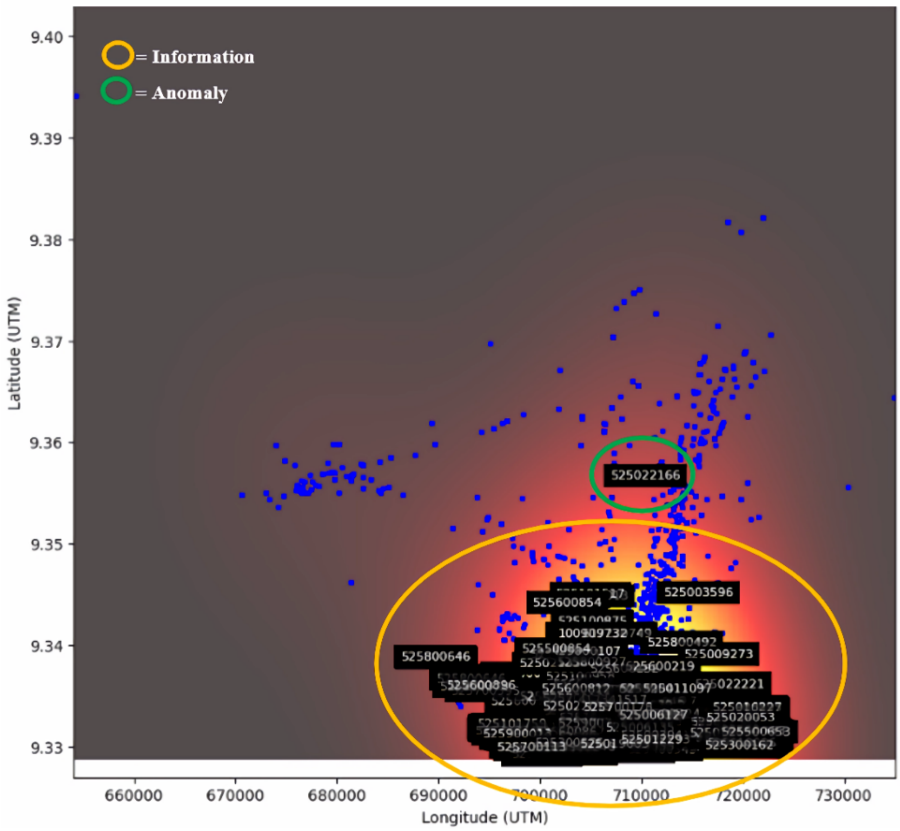

Figure 5. The anomaly ship is sailing; however, it fails to refresh its status information.

The ship identified by MMSI: 353999000, classified under shiptype code 70 (Cargo - encompassing all ships of this classification), is currently recorded at a speed of 0 and holds a status of 1 (At anchor). This anomaly ship exhibits a tendency to halt at various locations. Figure 6 illustrates the ship identified by MMSI: 353999000, which exhibits a pattern of stopping at various locations independently. This distinct behavior indicates an anomaly, particularly when analyzed in conjunction with the surrounding ships that maintain their movement, revealing atypical circumstances.

Utilizing the reference data that provided validation for the anomaly findings, the status of the ship is recorded as 1 (At anchor) with a speed of 0. The analysis of the resulting graph in Figure 6 reveals that this particular ship exhibits distinct behavior in comparison to other ships in motion (underway), indicating varying stopping positions. This scenario can be classified as an anomaly.

Anomalies are effectively identified and confirmed, with the visualization of these anomalies yielding insights into their locations. Upon detecting atypical ship behavior, characterized by a ship exhibiting distinct patterns despite having the same operational status as others, maritime supervisors can subsequently enhance surveillance measures. This response is warranted as such occurrences signify irregular activity within maritime operations. The findings from the analysis of ship movement anomalies can be utilized by maritime authorities and fishing communities to enhance their understanding of atypical ship activities.

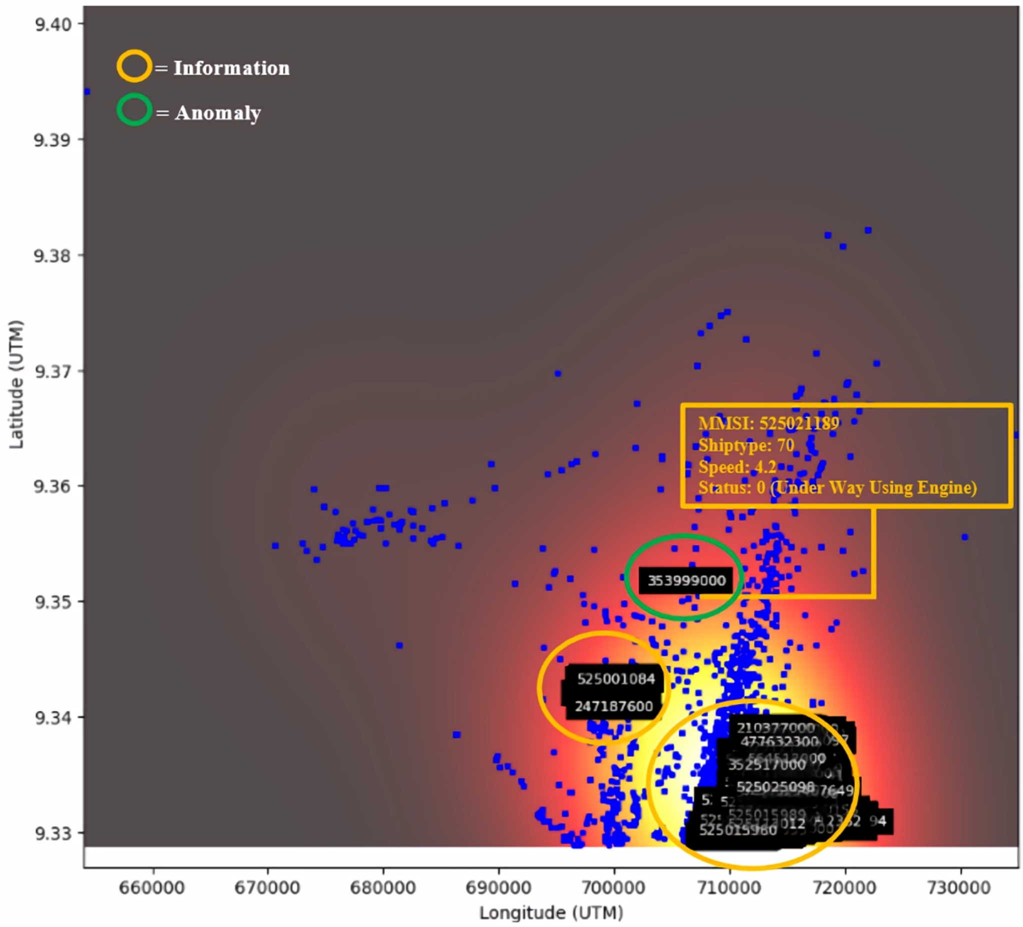

Figure 6. Anomaly behavior observed in ship halting at various locations

4. CONCLUSION

This study has demonstrated the application of a grid-based approach and anomaly detection techniques to analyze ship movement and congestion patterns within the Tanjung Priok harbor area, aimed at enhancing maritime monitoring capabilities. The study successfully pinpointed areas of high density by utilizing legacy AIS data, revealing that the greatest concentration occurs in proximity to the harbor, particularly surrounding berthing zones and shipping lanes. A grid was effectively constructed utilizing KDE to illustrate the density of ship movements. This visualization of ship density yields critical information regarding regions with elevated traffic levels.

Consequently, maritime inspectors can leverage this data to determine which areas require enhanced patrol or surveillance, as these high-density zones are susceptible to risks such as collisions. The visualization method provides significant insight into maritime traffic. Anomalies that have been identified are:

i. A cargo ship carrying hazardous materials traveled at high speed in a heavy traffic area, posing a significant safety risk, and
ii. A ship that behaves differently from other ships of the same status indicates unusual activity.

Anomalies have been successfully found and verified, with the visualization of these anomalies yielding information regarding their locations. The density data of ships provided by Yang et al. [17] can be utilized to find and verify anomalies associated with large ships navigating through congested regions. The occurrence of these anomalies constitutes important information, enabling maritime inspectors to issue warnings to ships navigating congested regions, which are susceptible to hazards like collisions. Upon recognizing anomalous ship behavior, which deviates from the norm despite its operational status, it is imperative to conduct additional surveillance to enhance security, as these occurrences signify unusual activity.

This finding emphasizes the necessity for increased maritime surveillance to improve port safety precautions. The anomalies identified are distinctly represented in a visual form. The condition of the AIS data significantly impacts the number of anomalies identified. However, time plays an even more critical role; a larger dataset may yield more anomalies, yet the process of exploring the data and searching for these anomalies is time-consuming. Additionally, the quality of the AIS data is paramount; if key information regarding ship movements, such as MMSI/ship ID, longitude, and latitude, is frequently missing or null, the amount of usable information will be severely limited.

Future research can look for other information from AIS data at Tanjung Priok harbor, such as extracting ship routes to find out marine information that can provide insight into maritime traffic patterns.

ACKNOWLEDGEMENTS

This research was funded and supported by the Research Organization for Earth Sciences and Maritime, the Indonesian National Research and Innovation Agency (BRIN), and Universitas Nasional.

REFERENCES

[1] Y. Rochwulaningsih, S. T. Sulistiyono, N. N. Masruroh, and N. N. Maulany, "Marine policy basis of Indonesia as a maritime state: The importance of integrated economy," *Mar Policy*, vol. 108, p. 103602, Oct. 2019, doi: 10.1016/j.marpol.2019.103602.

[2] I. Chapsos and J. A. Malcolm, "Maritime security in Indonesia: Towards a comprehensive agenda?," *Mar Policy*, vol. 76, pp. 178–184, Feb. 2017, doi: 10.1016/j.marpol.2016.11.033.

[3] X. Fu *et al.*, "AIS Data Analytics for Intelligent Maritime Surveillance Systems," in *Maritime Informatics*, Springer, 2021, pp. 393–411. doi: 10.1007/978-3-030-50892-0_23.

[4] D. Yang, L. Wu, S. Wang, H. Jia, and K. X. Li, "How big data enriches maritime research – a critical review of Automatic Identification System (AIS) data applications," *Transp Rev*, vol. 39, no. 6, pp. 755–773, Nov. 2019, doi: 10.1080/01441647.2019.1649315.

[5] S. T. Sulistiyono and Y. Rochwulaningsih, "Contest for hegemony: The dynamics of inland and maritime cultures relations in the history of Java island, Indonesia," *Journal of Marine and Island Cultures*, vol. 2, no. 2, pp. 115–127, Dec. 2013, doi: 10.1016/j.imic.2013.10.002.

[6] S. Galić, Z. Lušić, and I. Skoko, "The role and importance of safety in maritime transportation," *Book of*, p. 186, 2014.

[7] P. Kaluza, A. Kölzsch, M. T. Gastner, and B. Blasius, "The complex network of global cargo ship movements," *J R Soc Interface*, vol. 7, no. 48, pp. 1093–1103, Jul. 2010, doi: 10.1098/rsif.2009.0495.

[8] M. Winther, J. H. Christensen, M. S. Plejdrup, E. S. Ravn, Ó. F. Eriksson, and H. O. Kristensen, "Emission inventories for ships in the arctic based on satellite sampled AIS data," *Atmos Environ*, vol. 91, pp. 1–14, Jul. 2014, doi: 10.1016/j.atmosenv.2014.03.006.

[9] K. Wolsing, L. Roepert, J. Bauer, and K. Wehrle, "Anomaly Detection in Maritime AIS Tracks: A Review of Recent Approaches," Jan. 01, 2022, *MDPI*. doi: 10.3390/jmse10010112.
[10] B. J. Tetreault, "Use of the Automatic Identification System (AIS) for Maritime Domain Awareness (MDA)," in *Proceedings of OCEANS 2005 MTS/IEEE*, IEEE, Sep. 2005. doi: 10.1109/OCEANS.2005.1639983.
[11] D. Liu, X. Wang, Y. Cai, Z. Liu, and Z. J. Liu, "A novel framework of real-time regional collision risk prediction based on the RNN approach," *J Mar Sci Eng*, vol. 8, no. 3, Mar. 2020, doi: 10.3390/jmse8030224.
[12] R. Zhen, M. Riveiro, and Y. Jin, "A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traffic surveillance," *Ocean Engineering*, vol. 145, pp. 492–501, Nov. 2017, doi: 10.1016/j.oceaneng.2017.09.015.
[13] F. Natale, M. Gibin, A. Alessandrini, M. Vespe, and A. Paulrud, "Mapping fishing effort through AIS data," *PLoS One*, vol. 10, no. 6, Jun. 2015, doi: 10.1371/journal.pone.0130746.
[14] A. Lo Duca and K. McDowell, "Using the S-DIKW framework to transform data visualization into data storytelling," *J Assoc Inf Sci Technol*, 2024, doi: 10.1002/asi.24973.
[15] B. Dykes, *Effective data storytelling: how to drive change with data, narrative and visuals*. John Wiley & Sons, 2019.
[16] R. L. Scania and R. Laxhammar, "Anomaly detection for sea surveillance," in *2008 11th International Conference on Information Fusion*, IEEE, Jun. 2008. [Online]. Available: https://www.researchgate.net/publication/4369995
[17] J. Yang, L. Ma, and J. Liu, "Modeling and application of ship density based on ship scale conversion and grid," *Ocean Engineering*, vol. 237, Oct. 2021, doi: 10.1016/j.oceaneng.2021.109557.
[18] S. Cao, Y. Zeng, S. Yang, and S. Cao, "Research on Python Data Visualization Technology," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Feb. 2021. doi: 10.1088/1742-6596/1757/1/012122.
[19] S. Węglarczyk, "Kernel density estimation and its application," *ITM Web of Conferences*, vol. 23, p. 00037, 2018, doi: 10.1051/itmconf/20182300037.
[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*, Apr. 2020, [Online]. Available: http://arxiv.org/abs/2004.13653
[21] G. S. K. Ranjan, A. Kumar Verma, and S. Radhika, "K-Nearest Neighbors and Grid Search CV Based Real Time Fault Monitoring System for Industries," in *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, IEEE, Mar. 2019, pp. 1–5. doi: 10.1109/I2CT45611.2019.9033691.
[22] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011, Accessed: Apr. 24, 2025. [Online]. Available: https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?source=post_page
[23] M. J. Eniola, A. C. Pingal, and C. Caño, "Enhancing Density-Based Spatial Clustering of Applications with Noise (DBSCAN) with Kernel Density Estimation," *The Mindanawan Journal of Mathematics*, vol. 6, no. 2, pp. 113–133, 2024, doi: 10.62071/tmjm.v6i2.530.
[24] K. Wang, M. Liang, Y. Li, J. Liu, and R. W. Liu, "Maritime Traffic Data Visualization: A Brief Review," in *2019 4th IEEE International Conference on Big Data Analytics, ICBDA 2019*, Institute of Electrical and Electronics Engineers Inc., May 2019, pp. 67–72. doi: 10.1109/ICBDA.2019.8713227.
[25] W. Chen, F. Guo, and F. Y. Wang, "A Survey of Traffic Data Visualization," Dec. 01, 2015, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/TITS.2015.2436897.