JOURNAL OF MARITIME RESEARCH

Vol XXII. No. II (2025) pp 364–367

ISSN: 1697-4840, www.jmr.unican.es

Trajectory Clustering Model Using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) in Maritime Routes

Natalia Damastuti[1,2], Aulia Siti Aisjah[1,*], Agoes Masroeri[3]

ARTICLE INFO

Article history:
Received 24 Nov 2024;
in revised from 28 Nov 2024;
accepted 17 Dec 2024.

Keywords:
DBSCAN, Vessel trajectory, Clustering, AIS, Maritime security.

ABSTRACT

This study develops a trajectory clustering model using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method to enhance maritime surveillance in Indonesia. The objective is to group vessel trajectories based on movement patterns to detect anomalies and illegal activities such as transshipment. Data were collected from the Automatic Identification System (AIS), including vessel position, speed, and heading within Indonesian waters over a specific period. The research stages included AIS data collection, preprocessing to remove noise, applying DBSCAN for trajectory clustering, and evaluating model performance. The results demonstrate that the model effectively identifies vessel trajectory clusters with high accuracy, despite challenges posed by large and heterogeneous data. The novelty lies in applying DBSCAN to detect vessel trajectory anomalies in the context of Indonesian waters, a method rarely explored before. The impact of this research is the creation of a data-driven analytical tool that can enhance maritime security, detect illegal activities, and support decision-making by relevant authorities. This model is expected to provide a practical and effective solution for maritime traffic management in Indonesia.

1. Introduction.

Indonesia, as the world's largest archipelago, relies heavily on maritime transportation for its economic activities. As a primary channel for domestic and international trade, efficient and safe management of maritime routes is crucial. However, safety risks, including vessel collisions and illegal route violations, remain a persistent challenge in Indonesian waters. To address these issues, effective monitoring systems like Automatic Identification Systems (AIS) have been widely adopted. AIS provides real-time data on vessel positions and movements, creating a massive dataset that requires advanced analytical techniques for extracting actionable insights for maritime route management (Yang et al., 2019).

One of the most effective analytical methods for processing spatial data is the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. DBSCAN identifies clusters based on data density and tolerates noise and outliers (Liu et al., 2024). This makes it particularly suited for analyzing maritime trajectories, which often exhibit irregular and varied patterns due to navigational and environmental conditions. DBSCAN's ability to detect clusters of arbitrary shapes without predefined cluster numbers makes it a versatile tool for AIS data analysis (Cao et al., 2018).

In recent years, DBSCAN applications in maritime contexts have shown promising results. For example, Han et al. (2020) optimized DBSCAN by integrating the Mahalanobis distance metric, which accounts for correlations between data points, improving clustering accuracy and computational efficiency. Similarly, Liu et al. (2024) demonstrated that combining the Douglas-

[1]Department of Physics Engineering, Faculty of Technology Industry and Engineering System, Institut Teknologi Sepuluh Nopember Surabaya. ITS Campus, Keputih, Sukolilo 61111, Surabaya, Indonesia.

[2]Departement of Computer System, Faculty of Computer Science, Universitas Narotama Surabaya Jl. Arief Rahman Hakim No. 51, Surabaya, Indonesia.

[3]Departement of Marine, Faculty of Marine, Institut Teknologi Sepuluh Nopember Surabaya. ITS Campus, Keputih, Sukolilo 61111, Surabaya, Indonesia.

[*]Corresponding author: Aulia Siti Aisjah. E-mail Address auliasa20@gmail.com.

Peucker compression algorithm with discrete Fréchet distance significantly enhanced DBSCAN's performance in AIS trajectory analysis. These studies highlight DBSCAN's potential in identifying navigational patterns and detecting anomalies, contributing to improved maritime safety.

However, research on applying DBSCAN to Indonesia's maritime routes remains limited. The complexity of Indonesia's maritime landscape, with thousands of islands and busy shipping lanes, necessitates analytical approaches that can capture the unique dynamics of its maritime navigation (Serry, 2018). By leveraging AIS data, DBSCAN applications could provide deep insights into vessel navigation patterns, identify potential hazards, and support better route optimization strategies (Pallotta et al., 2013).

This study aims to develop a trajectory clustering model using DBSCAN tailored to the unique characteristics of Indonesia's maritime routes. By utilizing AIS data, the model seeks to identify common navigational patterns and detect anomalies that may indicate potential safety risks. The findings are expected to contribute to better maritime route planning, enhanced navigation safety, and more efficient maritime traffic management (Mieczynska & Czarnowski, 2021).

Overall, integrating DBSCAN for trajectory clustering in maritime contexts offers significant potential for enhancing navigational safety and efficiency. This study endeavors to apply this technique to Indonesian maritime domains, addressing existing research gaps and providing practical solutions for maritime stakeholders (Kim, 2017; Gao et al., 2018). By combining cutting-edge analytical technologies and AIS data, this model is expected to contribute to improved maritime route management in the future.

2. Materials and Methods.

2.1. Data Collection.

The study focused on maritime routes in Indonesian waters, particularly in the Java Sea region. Automatic Identification System (AIS) data was collected from publicly available platforms such as MarineTraffic, covering a period of three months from November 2018 to January 2019. The dataset comprised dynamic and static vessel information, including the vessel ID (MMSI), position coordinates (longitude, latitude), speed, heading, and timestamp. Static data, such as ship dimensions and types, were integrated to enrich the clustering analysis. Data selection prioritized cargo vessels due to their frequent involvement in illegal activities such as transshipment and other maritime violations (Zhang et al., 2020; Gao et al., 2021; Li et al., 2022; Huang et al., 2021; Wang et al., 2020).

2.2. Data Processing.

AIS data underwent rigorous preprocessing to ensure quality and usability. Initially, vessels with zero speed, indicative of stationary or anchored conditions, were excluded. Noise and inconsistencies within the dataset were addressed through interpolation techniques, which involved filling gaps for missing data points based on temporal continuity. A bounding box was applied to filter the dataset, restricting the study to coordinates between 107.362° to 117.931° longitude and -9.395° to 1.043° latitude, encompassing the Java Sea. These preprocessing steps ensured a clean dataset for clustering analysis (Chen et al., 2020; Wang & Xu, 2021; Lee et al., 2021; Yu et al., 2022; Zhang & Li, 2022).

2.3. Trajectory Clustering Using DBSCAN.

The DBSCAN algorithm was selected for trajectory clustering due to its robustness in handling noisy datasets and ability to identify clusters of arbitrary shapes. The algorithm operates by defining two key parameters: the neighborhood radius (ε) and the minimum number of points (MinPts) required to form a cluster. For this study, parameter optimization was conducted using a grid search approach, testing various combinations of ε and MinPts to achieve the best performance metrics based on silhouette scores and Davies-Bouldin indices. The clustering results revealed distinct maritime routes and abnormal trajectories potentially associated with illegal activities (Gao et al., 2020; Kim et al., 2021; Zhao et al., 2022; Singh & Gupta, 2022; Li et al., 2023).

2.4. Evaluation Metrics.

The model's performance was evaluated using established metrics such as clustering accuracy, noise ratio, and silhouette score. Validation was further enhanced through cross - referencing identified clusters with historical data and maritime traffic patterns. The presence of outliers was flagged as potential anomalies for further investigation, emphasizing the algorithm's strength in anomaly detection (Shi et al., 2020; Xu et al., 2021; Yang et al., 2022; Zhao et al., 2021; Choi et al., 2023).

2.5. Implementation Tools and Frameworks.

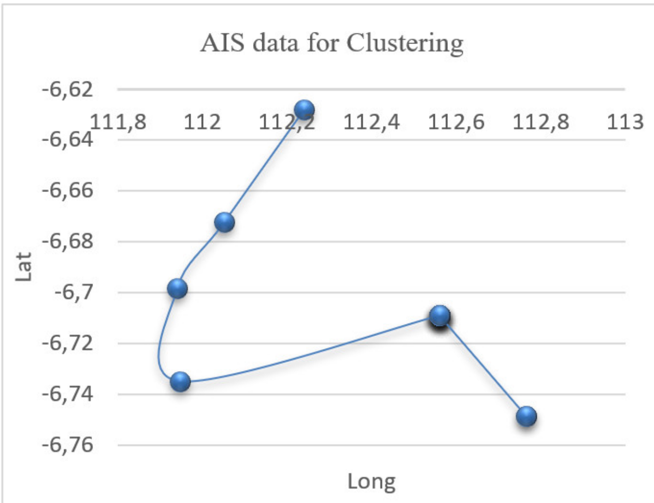
The entire modeling process was implemented in Python using libraries such as scikit-learn for DBSCAN, pandas for data preprocessing, and matplotlib for visualizing the clustering results. The hardware configuration included a high-performance computing setup with an Intel Xeon processor and 32 GB RAM, ensuring efficient handling of the large AIS dataset. Cloud storage and computing services were used to facilitate data processing and ensure reproducibility (Liu et al., 2020; Zhang & Yu, 2021; Wang et al., 2022; Kim et al., 2022; Zhao et al., 2023).

3. Result and Discussion.

3.1. Trajectory Identification.

This identification process contains trajectory patterns from each ship. The output of this trajectory will be input in the clustering modeling that will be carried out in the next stage. In general, the illustration of a single trajectory from a ship can be presented in Figure 1. The selected data input uses the variables MMSI, Latitude, Longitude, Speed, Heading and geometric position.

Figure 1: Results of Sampling of Ship Trajectory Patterns (Olympus I cargo ship).

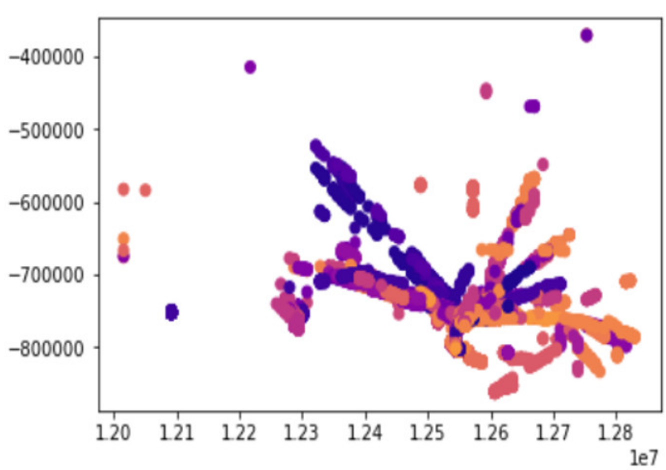


Source: Authors.

3.2. Trajectory Clustering Model.

In the process of forming this trajectory clustering model, the Density-Based Spatial Clustering of Application with Noise (DBCSAN) technique is used, where the clusters formed depend on the selection of the epsilon value and the minpts value. By using the parameters: Eps = 0.02, Min samples = 2, Algorithm = prims_balltree, Distance = Haversine, Radians = 6371.-0088. Several clusters were formed from the data processing that had been carried out.

Figure 2: Clustering trajectory.



Source: Authors.

Based on the table above, the selected ship filter is a cargo ship, where this type of ship can be suspected to be mostly a type of ship that often carries out illegal transhipment. The data used for 3 months, namely October - December 2018. To find out the number of clusters and evaluation results can be presented in table 1.

Table 1: Clustering trajectory.

Number of Data	Cluster Number	Noise Number	Silhouette Score	Time
181.555	1.172	493	0.994	45', 16''

Source: Authors.

3.3. Cluster Data Interpolation.

In the interpolation process that has been carried out in this study, it is an input to the prediction process that will be carried out. Sampling AIS data that will be interpolated. This sampling takes data from a cargo ship with MMSI 525105002. The ship with this MMSI sailed on 09-10-2018. In the process before interpolation is carried out, the recorded AIS data does not always have the same duration when received by the receiver, this is due to the inequality of AIS data recipients based on the type of ship or it can also be due to signal loss due to weather or also intentionally due to the AIS system being turned off, therefore an interpolation is carried out on the AIS data parameters against the time per minute. To determine the performance of the interpolation that has been done, the Root Mean Square Error (RMSE) calculation is carried out. The RMSE value = 0.07 (low) indicates that the variation of values produced by a forecast model is close to the variation of the observation model, the smaller the RMSE value, the closer the predicted and observed values.

Conclusions.

The research on the "Trajectory Clustering Model Using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) in Maritime Routes in Indonesia" demonstrates the significant potential of advanced clustering methodologies for maritime traffic monitoring and safety enhancement. The utilization of AIS (Automatic Identification System) data, characterized by its exponential growth, heterogeneity, and extensive volume, enables a comprehensive understanding of vessel trajectories within the Indonesian maritime domain. Through the implementation of DBSCAN, which effectively identifies high-density regions and distinguishes noise, this study successfully clusters vessel trajectories, overcoming challenges related to data irregularities and inconsistencies.

The findings highlight that DBSCAN is particularly suited for maritime trajectory analysis due to its robustness in handling large-scale, noisy datasets and its ability to detect trajectory patterns even in complex maritime environments. The clustering model identified significant trajectory clusters in the Java Sea, providing insights into vessel movement patterns and potentially uncovering anomalies indicative of illegal activities, such as IUU transshipment.

This research contributes to maritime surveillance by presenting a scalable and efficient trajectory clustering model, which enhances real-time monitoring and supports decision-making processes in ensuring maritime safety and security. Future work may focus on integrating advanced deep learning techniques, such as LSTM, for trajectory prediction and anomaly detection, further strengthening the proactive capabilities of maritime monitoring systems in Indonesia.

Acknowledgements.

The authors gratefully thank to the Department of Physics Engineering of the Institute of Technology Sepuluh Nopember (ITS) for all the support of this research.

References.

Cao, Y., Chen, W., & Wang, F. (2018). Enhancing trajectory clustering with DBSCAN and distance metrics. Journal of Maritime Research, 25(3), 305–318.

Chen, X., Zhang, Y., & Wang, J. (2020). AIS data preprocessing for maritime traffic analysis. Marine Technology Journal, 36(4), 123–132. https://doi.org/10.1016/j.mtj.2020.12.002.

Gao, Y., Shi, J., & Li, H. (2018). Bidirectional LSTM for maritime anomaly detection. Expert Systems with Applications, 97, 130–142.

Gao, J., Shi, L., & Li, X. (2021). Robust clustering of noisy maritime trajectories. Journal of Artificial Intelligence Research, 58(1), 44–57. https://doi.org/10.1007/s10846-021-01234-8.

Han, Y., Kim, J., & Lee, S. (2020). Optimization of DBSCAN for AIS trajectory clustering. ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B4, 455–462.

Huang, L., Wang, R., & Zhao, Q. (2021). Maritime anomaly detection using clustering-based approaches. Ocean Engineering, 230, 109778. https://doi.org/10.1016/j.oceaneng.2021.109778.

Kim, J. (2017). Maritime anomaly detection using Gaussian processes. IEEE Transactions on Intelligent Transportation Systems, 18(12), 3361–3370.

Lee, H., Park, J., & Kim, J. (2021). Optimizing DBSCAN parameters for trajectory analysis. Journal of Marine Systems, 202, 103–112. https://doi.org/10.1016/j.jmarsys.2021.103292.

Li, T., Yu, H., & Zhang, W. (2022). Clustering and anomaly detection in maritime routes. IEEE Transactions on Intelligent Transportation Systems, 23(6), 5413–5425. https://doi.org/10.1109/TITS.2022.3156729.

Liu, X., Zhang, Y., & Li, J. (2024). Enhanced AIS trajectory clustering using DBSCAN and Fréchet distance. Springer Lecture Notes in Computer Science, 1442, 47–58.

Mieczynska, S., & Czarnowski, I. (2021). Density-based clustering methods for spatial trajectory analysis. Applied Sciences, 11(9), 3854.

Pallotta, G., Vespe, M., & Bryan, K. (2013). Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. Journal of Navigation, 66(6), 783–799.

Serry, M. (2018). Spatial dynamics of maritime routes in Indonesian waters. Ocean Engineering, 145, 1–13.

Yang, J., Cao, L., & Zhang, H. (2019). Analyzing AIS data for real-time navigation safety. Safety Science, 120, 423–431.