ANALYZING MOTION IRREGULARITIES IN AIS DATA FOR MARITIME TRAJECTORY BEHAVIOUR

Lavanya Vyas[1], Sapna Salian[2], Nishigandha Sarode[3], Sonal Kadam[4]

[1]Student, [2]Student, [3]Student, [4]Asst. Professor
Department of Computer Science and Technology
SNDT Women's University, Mumbai, India

Abstract: The Automatic Identification System (AIS) is widely used for maritime navigation and vessel tracking but remains vulnerable to data inconsistencies due to the absence of authentication mechanisms. This study presents an analysis of AIS data to examine motion characteristics associated with irregular vessel behaviour. A compact set of movement indicators, including displacement between positions, reporting intervals, and speed changes, is used to analyze inconsistencies in vessel trajectories. An Isolation Forest model is employed as an exploratory tool to identify observations exhibiting significant deviations in movement patterns. The analysis reveals that irregular behaviour is strongly associated with abrupt positional changes and inconsistent reporting intervals. These findings highlight the importance of continuity in vessel movement and provide insight into observable patterns that can support anomaly detection approaches in maritime systems.

***Index Terms* - AIS, Anomaly Detection, Isolation Forest, Maritime Data, Trajectory Analysis**

I. INTRODUCTION

The Automatic Identification System (AIS) is a radio-based communication protocol mandated under the Safety of Life at Sea (SOLAS) convention for large commercial vessels. AIS transponders broadcast navigational data that contains vessel identity, geographic coordinates, speed, and course at regular intervals, enabling maritime authorities and operators to monitor vessel activity across coastal and open-water environments [1]. Its widespread adoption has made AIS the primary data source for studying vessel movement, maritime traffic patterns, and operational behaviour at sea.

Despite its practical importance, AIS was not built with data integrity in mind. Messages are transmitted over open radio channels without encryption or authentication, so the records received by monitoring systems may carry inconsistencies from a range of sources: transmission errors, signal interference, GPS inaccuracies, and transponder configuration faults [2], [3]. Whatever their cause, such inconsistencies disrupt the continuity of vessel trajectories and can undermine the reliability of any analysis that depends on them.

Researchers have explored various ways to assess AIS data quality and to distinguish unusual observations from normal vessel behaviour [4], [5]. Among these, movement-based indicators, which are derived directly from consecutive position reports, have proven to be a practical and readable basis for this task. The distance between successive positions, the time elapsed between messages, and changes in reported speed all respond to breaks in movement continuity and require no external reference data or labeled examples to compute.

Unsupervised methods are well suited to AIS analysis because verified anomaly labels are rarely available. The Isolation Forest algorithm [7] is one such method: it identifies atypical observations by measuring how easily each point can be separated from the rest of the data through recursive random partitioning. It makes no assumptions about the shape of normal behaviour and scales well to large datasets, making it a natural fit for exploratory work on AIS records.

This study examines motion characteristics in AIS data to understand, at the feature level, how irregular observations differ from typical vessel behaviour. The aim is not to build a detection system but to characterize what irregularities look like in terms of observable movement indicators, and to consider what those patterns reveal about the data. The AIS records used here cover U.S. coastal waters during April 2024 and were obtained from the NOAA MarineCadastre repository [8].

The remainder of this paper is organized as follows. Section II reviews relevant prior work on AIS data analysis and movement-based characterization of vessel behaviour. Section III describes the dataset, preprocessing steps, and the movement indicators used in this study. Section IV presents the results and their interpretation. Sections V and VI discuss the findings and limitations, respectively. Section VII concludes the paper.

II. RELATED WORK

Research on AIS data analysis has grown along two main lines: work on the security and integrity of the AIS protocol itself, and work on modeling vessel movement to identify behavioural deviations. This study draws on both, with particular attention to approaches that characterize irregular behaviour through movement-derived indicators.

2.1 AIS Protocol Vulnerabilities and Data Integrity

The security limitations of AIS are well established. Balduzzi et al. showed through software-defined radio experiments that AIS messages can be intercepted and forged, allowing spoofing and identity falsification without any onboard alarm [2]. Kavallieratos and Katsikas examined these risks in the context of networked vessels, identifying the lack of authentication as a structural weakness that exposes maritime communication to a wide range of interference [3]. Together, these findings underscore the need for analytical tools that can assess data quality without depending on authenticated inputs.

AIS data is also affected by inconsistencies that arise independently of deliberate interference. Reporting gaps, duplicate records, coordinate errors, and irregular transmission intervals are common in real-world AIS streams and can distort trajectory reconstruction [4]. Identifying and accounting for these issues is a necessary step before any behavioural analysis can be carried out.

2.2 Trajectory Analysis and Vessel behaviour

A substantial body of work has examined how vessel trajectories can be modeled and how deviations from expected patterns can be identified. Pallotta et al. developed an unsupervised framework for extracting traffic patterns from AIS data, showing that typical vessel routes can be inferred from historical records and used to flag departures from normal behaviour [1]. Their work established that trajectory-level analysis can reveal behavioural structure without requiring labeled examples or expert annotation.

Mazzarella et al. examined intentional AIS on-off switching, where vessels suppress transmissions to hide their position [4]. Their analysis showed that this behaviour leaves characteristic gaps and discontinuities in reporting intervals which is observable through the timing of transmitted messages alone. This is directly relevant here, as reporting interval is one of the three motion indicators used in the present study.

Shi et al. studied AIS-based behavioural analysis using movement features derived from position reports, finding that speed variation and spatial displacement serve as practical indicators of unusual activity [5].Studies have also examined efficient handling and processing of AIS trajectory data. For instance, online methods for trajectory compression have been proposed to reduce data volume while preserving essential movement characteristics, enabling scalable analysis of large AIS datasets [6]. The present work uses a deliberately compact feature set in order to keep the analysis readable and grounded in observable behaviour.

2.3 Isolation Forest in AIS Analysis

Liu et al. introduced Isolation Forest as an ensemble method for identifying outliers by measuring how easily individual observations can be isolated through recursive random partitioning [7]. Observations that require fewer partitions to isolate tend to lie in sparse, peripheral regions of the feature space and are treated as anomalous. This behaviour suits movement data well, since irregular AIS observations typically occupy such low-density regions.

Zheng et al. applied Isolation Forest to AIS records and showed that spatial jump distance and temporal gap between messages are strong inputs for identifying kinematically inconsistent segments [9]. Their results confirm that movement-derived features provide a solid basis for unsupervised anomaly analysis in this domain.

Rong et al. approached maritime anomaly characterization through statistical analysis of trajectory attributes, including position change rates and temporal regularity [10]. Their work found that irregular behaviour in AIS data is consistently linked to disruptions in movement continuity. This finding directly motivates the feature-level analysis in this study.

Taken together, these studies show that a small set of movement indicators, examined through an unsupervised method, can surface meaningful structure in AIS data without labeled examples of anomalous behaviour. The present work follows this approach, focusing on interpretive analysis of three indicators: displacement between positions, reporting interval, and speed change.

III. METHODOLOGY

3.1 Data Source

The study uses AIS position records obtained from the NOAA MarineCadastre repository [8], covering U.S. coastal waters during April 2024. This dataset is distinct from AIS data used in prior related work and consists of vessel position reports including vessel identifier (MMSI), timestamp, geographic coordinates, speed over ground, and course over ground. The April 2024 dataset was selected to ensure that the analysis reflects a recent and independent observation period.

3.2 Data Preprocessing

The dataset is preprocessed to ensure trajectory reliability. Preprocessing steps include:

- Removing duplicate and incomplete records
- Validating geographic coordinates against permissible latitude and longitude ranges
- Sorting observations chronologically within each vessel trajectory

These steps establish a consistent basis for computing movement indicators and ensure that trajectory segments reflect actual vessel movement.

3.3 Movement Indicators

Three movement indicators are derived from consecutive AIS observations for each vessel:

- Displacement between positions: the great-circle distance (in nautical miles) between successive reported coordinates, referred to as jump distance.
- Reporting interval: the elapsed time in seconds between consecutive AIS messages from the same vessel.
- Speed change: the absolute difference in reported speed over ground between consecutive observations.

These indicators are computed directly from raw AIS fields without transformation or dimensionality reduction. They capture the continuity of vessel movement in space, time, and kinematics respectively, and are interpretable without domain modeling assumptions.

3.4 Exploratory Identification of Irregular Observations

An Isolation Forest model is used as an exploratory tool to identify observations that exhibit substantial deviations in the three-dimensional movement indicator space. The algorithm partitions the feature space through recursive random splits; observations requiring fewer splits to isolate receive higher anomaly scores and are treated as irregular.

The analysis is unsupervised: no ground-truth labels are available for AIS observations in the dataset, and the model is not trained to distinguish specific event types. The Isolation Forest output is used to partition observations into typical and irregular groups for the purpose of comparative feature-level analysis. The focus is on understanding how the two groups differ across movement indicators rather than on producing a classification system.

IV. RESULTS AND ANALYSIS

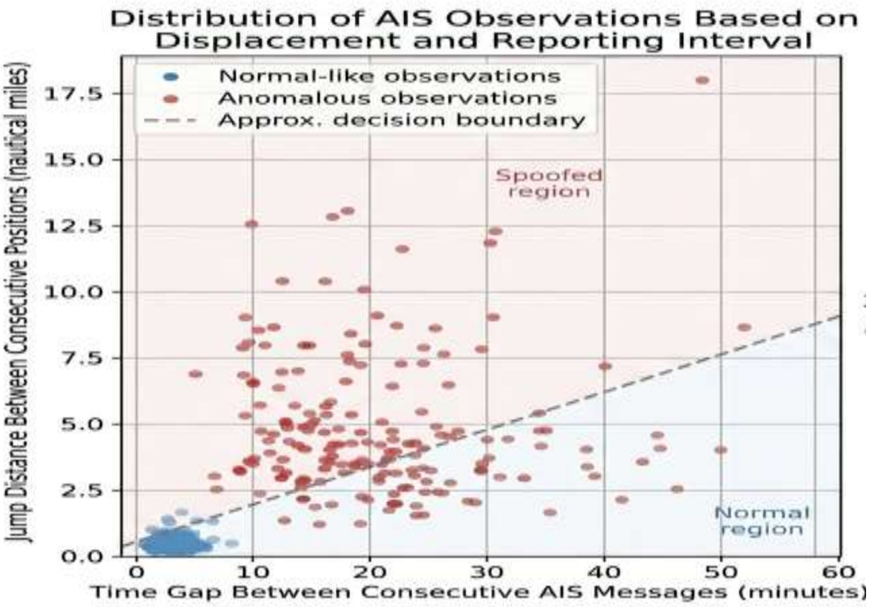

Figure 1: Distribution of AIS Observations Based on Jump Distance and Reporting Interval

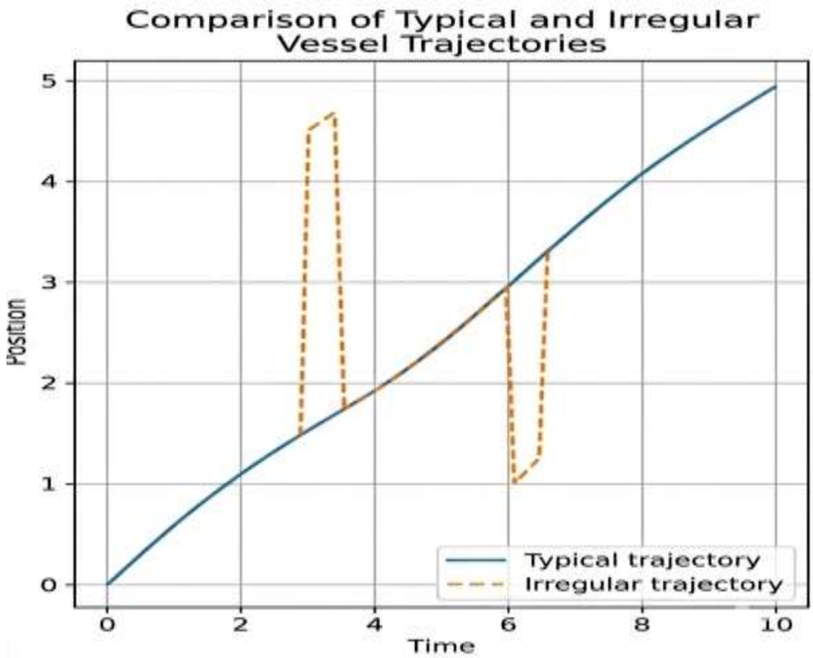

Figure 2: Comparison of Typical and Irregular Vessel Trajectories

4.1 Observed Movement Patterns

Vessel trajectories in the typical group exhibit smooth, continuous movement consistent with stable navigation behaviour. Positional updates are closely spaced, reporting intervals are regular, and speed changes between messages are small. In contrast, observations identified as irregular show pronounced discontinuities: abrupt positional shifts between consecutive messages, extended gaps in reporting, and sudden changes in speed. As shown in Figure 2, the trajectory of a typical vessel follows a coherent spatial path, while an irregular trajectory displays sudden deviations that are inconsistent with physically plausible vessel movement. These visual differences are consistent across the dataset and are not isolated to individual vessels.

4.2 Distribution of Observations

Figure 1 illustrates how typical and irregular observations are distributed across jump distance and reporting interval. Typical observations occupy a compact region characterized by low jump distances and short, consistent reporting intervals. Irregular observations are dispersed across a substantially wider region of the feature space, with higher values on both axes. This spatial separation in the feature distribution is consistent with the Isolation Forest model's partitioning behaviour: irregular observations are those that occupy low-density regions and are therefore easier to isolate through random partitioning.

4.3 Relationship Between Movement Indicators

The distribution in Figure 1 also reveals a co-occurrence pattern between jump distance and reporting interval in irregular observations. As reporting intervals increase, the variability in jump distance tends to increase as well. This relationship is not present in the typical group, which remains tightly clustered regardless of minor variation in either indicator. The pattern suggests that irregularities in position reporting and irregularities in spatial displacement are not independent: when one indicator deviates from the norm, the other tends to deviate as well. This joint deviation strengthens the identifiability of irregular observations and indicates that movement continuity operates across multiple dimensions simultaneously.

4.4 Statistical Summary of Movement Indicators

Table 4.4: Statistical Summary of Movement Indicators for Typical and Irregular Observations

Indicator	Typical Observations (Mean ± Std)	Irregular Observations (Mean ± Std)	Interpretation
Displacement between positions (NM)	0.04 ± 0.07	1.16 ± 33.88	Irregular observations show significantly larger positional shifts
Reporting interval (seconds)	112.49 ± 61.93	281.67 ± 1527.42	Irregular observations exhibit inconsistent reporting intervals
Change in speed (knots)	0.06 ± 0.12	1.03 ± 2.50	Irregular observations demonstrate abrupt variations in speed

Table 4.4 presents the mean and standard deviation of each movement indicator for typical and irregular observations. The differences across all three indicators are substantial. The large variance reflects the presence of extreme deviations in a small subset of observations.

4.5 Interpretation of Irregular behaviour

The statistical results in Table 4.4 confirm that irregular observations are not characterized by modest deviations from the norm but by extreme and highly variable departures across all three indicators. The large standard deviations in the irregular group indicate that this category encompasses a wide range of deviation magnitudes, suggesting that irregularities in AIS data do not follow a single pattern but manifest in diverse ways. Despite this diversity, the common thread is disruption of movement continuity: sudden positional shifts occur without proportional increases in elapsed time, or extended intervals pass without meaningful positional change. Both scenarios are inconsistent with physically plausible vessel movement at typical operational speeds. The trajectory comparison in Figure 2 reinforces this interpretation. The irregular trajectory does not simply represent a vessel moving faster or along a different route; it represents a sequence of reported positions that does not form a coherent spatial path. This qualitative distinction, visible in the trajectory plot, is consistent with the quantitative separation observed in Table 4.4 and Figure 1.

4.6 Summary of Findings

The results demonstrate that displacement, reporting interval, and speed change collectively provide a clear and interpretable basis for distinguishing typical from irregular AIS observations. Irregularities are characterized by abrupt discontinuities rather than gradual deviations, and these discontinuities tend to co-occur across multiple indicators. The analysis shows that movement continuity is a fundamental and measurable property of normal vessel behaviour in AIS data, and that deviations from this continuity can be surfaced using a minimal set of features without requiring labeled training data.

V. DISCUSSION

The findings of this study highlight the value of feature-level interpretation in AIS data analysis. By examining how movement indicators differ between typical and irregular observations, the analysis provides insight into the nature of irregularities without relying on a complex model architecture or labeled ground truth. This interpretive orientation complements approaches that focus on detection performance, offering a basis for understanding what the data contains and why certain observations stand out.

A key observation is that irregular behaviour in AIS data is not uniform. The large variance in movement indicators among irregular observations indicates that the category encompasses multiple distinct deviation patterns. Some irregular observations are characterized primarily by large jump distances, while others show extended reporting intervals with little positional change. This diversity suggests that a single threshold-based rule would be insufficient to capture the full range of irregular behaviour, and that multi-indicator analysis is preferable.

The use of Isolation Forest as an exploratory tool is appropriate in this context because it makes no assumptions about the distribution of anomalies and does not require a predetermined decision boundary. Its outputs provide a principled basis for partitioning observations, which supports comparative analysis without overstating the certainty of individual classifications. The approach is transparent in that the features driving the separation are directly observable and interpretable, rather than embedded in a black-box representation.

More broadly, the results suggest that continuity of vessel movement across position, time, and speed is a reliable and operationally meaningful criterion for assessing the regularity of AIS observations. This principle can inform the design of data quality filters, preprocessing pipelines, and more sophisticated behavioural models in future work.

VI. LIMITATIONS

The analysis is conducted without verified anomaly labels. The Isolation Forest model identifies observations that are statistically atypical relative to the dataset, but this does not imply that they represent confirmed events of a specific type. Some irregular observations may reflect data quality issues such as GPS errors or transmission noise rather than deliberate manipulation or operationally significant events. The analysis does not attempt to distinguish between these possibilities.

The three motion indicators used in this study capture important aspects of movement behaviour but do not represent the full range of features that could be relevant. Heading changes, acceleration, and contextual factors such as proximity to port or weather conditions are not included, and their omission may limit the completeness of the analysis. The approach is intended as an interpretive study rather than a comprehensive detection system and should be understood accordingly.

Finally, the results are specific to the April 2024 dataset and may not generalize to other geographic regions, vessel types, or time periods without further validation.

VII. CONCLUSION

This paper presents a feature-level analysis of movement indicators derived from AIS position data to characterize typical and irregular vessel behaviour. Using three compact indicators, namely displacement between positions, reporting interval, and speed change, along with an Isolation Forest model as an exploratory tool, the study demonstrates that irregular observations are consistently associated with disruptions in movement continuity across spatial, temporal, and kinematic dimensions.

The analysis reveals that irregular behaviour in AIS data is characterized by abrupt and often co-occurring deviations across multiple indicators, rather than isolated or gradual departures. Typical vessels maintain coherent and continuous trajectories with stable reporting behaviour, whereas irregular observations break from this continuity in identifiable and interpretable ways. These findings provide a transparent basis for understanding what distinguishes unusual AIS observations from the norm, without requiring labeled training data or complex model architectures.

The study contributes to ongoing efforts to improve the interpretability of AIS data analysis and establishes a foundation for future work on movement-based characterization of vessel behaviour. Potential directions include extending the feature set to incorporate additional kinematic attributes, applying the approach to datasets from different maritime regions and time periods, and integrating contextual information to support more fine-grained behavioural analysis.

VIII. ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Department of Computer Science and Technology at SNDT Women's University, Mumbai, for providing the academic environment and support necessary for this study. The authors also acknowledge the NOAA MarineCadastre platform for providing access to publicly available AIS data used in this analysis.

REFERENCES

[1] Pallotta, G., Vespe, M. and Bryan, K. 2013. Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy*, 15(6): 2218–2245.

[2] Balduzzi, M., Pasta, A. and Wilhoit, K. 2014. A security evaluation of AIS automated identification system. *Proceedings of the 30th Annual Computer Security Applications Conference*: 368–377.

[3] Kavallieratos, G. and Katsikas, S. 2020. Managing cyber security risks of the cyber-enabled ship. *Journal of Marine Science and Engineering*, 8(10): 768.

[4] Mazzarella, F., Vespe, M. and Vollero, A. 2017. A novel anomaly detection approach to identify intentional AIS on-off switching. *Expert Systems with Applications*, 78: 1–13.

[5] Shi, Y., Long, C., Yang, X. and Deng, M. 2022. Abnormal ship behavior detection based on AIS data. *Applied Sciences*, 12(9): 4635.

[6] Liu, Z., Yuan, W., Liang, M., Zhang, M., Liu, C., Liu, R. W. and Liu, J. 2024. An online method for ship trajectory compression using AIS data. *Journal of Navigation*, 77(1): 1–22.

[7] Liu, F. T., Ting, K. M. and Zhou, Z. H. 2008. Isolation Forest. *Proceedings of the IEEE International Conference on Data Mining*: 413–422.

[8] NOAA. 2026. AIS Data Repository. Available: https://hub.marinecadastre.gov/

[9] Zheng, H., Hu, Q., Yang, C., Mei, Q., Wang, P. and Li, K. 2023. Identification of spoofing ships from automatic identification system data via trajectory segmentation and Isolation Forest. *Journal of Marine Science and Engineering*, 11(8): 1516.

[10] Rong, Y., Zhao, Z. and Zhang, X. 2017. Maritime anomaly detection using trajectory clustering and statistical analysis. *Journal of Navigation*, 70(3): 37–58.