A Maritime Loitering Behavior Analysis with Multi-Stage Framework based on Extended Kalman Filter and HDBSCAN

Tuan Manh Tao, Ukhatov Aleksandr, Changha Lee, Thanh Lam Tran and Chan-Hyun Youn*
School of Electrical Engineering
Korea Advanced Institute of Science and Technology
Daejeon, Korea
{ttmanh, uhatov007, changha.lee, thanhla, chyoun}@kaist.ac.kr

***Abstract*—Loitering vessel behavior—marked by slow, irregular, or circling motion without navigational intent—is a key indicator of illicit activities such as illegal fishing or covert rendezvous. Detecting such behavior is challenging due to sparse AIS transmissions, noisy measurements, and the lack of labeled datasets. We propose an unsupervised framework that integrates trajectory reconstruction, adaptive segmentation, and anomaly detection to automatically label loitering segments without manual annotation. A Reconstructive Adaptive Extended Kalman Filter (R-AEKF) restores missing AIS points by modeling nonlinear vessel dynamics with adaptive noise estimation. Inter-adaptive sliding windows capture local kinematic and entropy features, which are clustered via HDBSCAN to isolate anomalous segments. Loitering behaviors are further verified using spectral and directional metrics. Experiments on real-world AIS data demonstrate that the proposed method consistently outperforms threshold-based and supervised baselines, offering a scalable solution for Maritime Domain Awareness.**

***Index Terms*—AIS trajectory reconstruction, anomaly detection, HDBSCAN, loitering detection, maritime surveillance, spectral analysis, unsupervised learning, vessel behavior modeling.**

I. Introduction

Maritime loitering behavior—where vessels engage in prolonged, irregular, and low-speed maneuvers within a localized area—often signals illicit or covert activities such as illegal fishing, rendezvous at sea, smuggling, or surveillance operations. Unlike straightforward transit or docking maneuvers, loitering lacks a clear navigational objective and is typically masked within dense maritime traffic or sparse oceanic regions. Its detection is therefore of critical importance for Maritime Domain Awareness (MDA), yet remains one of the most challenging behavioral patterns to identify reliably.

The Automatic Identification System (AIS), mandated by the International Maritime Organization (IMO) for large vessels, serves as the primary data source for monitoring maritime activity [1]. AIS broadcasts real-time position, speed, and course information that can reveal patterns of movement over time. However, AIS data is inherently noisy and frequently suffers from missing segments due to satellite blind spots, voluntary transponder shutdowns, spoofing, or deceptive broadcasting [2]. These characteristics severely limit the effectiveness of traditional rule-based or supervised methods in detecting subtle behavioral anomalies like loitering.

Loitering behavior is particularly difficult to quantify due to its contextual nature. For instance, a low-speed circling pattern may be routine near ports but highly suspicious in restricted or remote zones. Moreover, the absence of labeled datasets precludes the use of conventional supervised learning techniques. Thus, effective loitering detection systems must: (i) estimate vessel intent from partial observations, (ii) adapt to diverse motion patterns across vessel types and geographies, and (iii) operate without labeled behavioral categories.

In this work, we propose a unified unsupervised framework for loitering detection that bridges raw AIS data with interpretable behavioral insights. First, a Reconstructive Adaptive Extended Kalman Filter (R-AEKF) recovers missing or noisy AIS segments while preserving nonlinear vessel dynamics, ensuring reliable trajectory continuity. Next, an adaptive sliding window captures speed and heading shifts, and HDBSCAN clusters normal patterns while isolating loitering-like outliers. Finally, spectral and directional concentration analysis on Course Over Ground (COG) labels loitering segments and classifies vessel behaviors into six types, enabling fine-grained, interpretable anomaly detection without manual supervision.

Our contributions include: (i) a R-AEKF for reconstructing incomplete AIS trajectories; (ii) an adaptive sliding window segmentation and HDBSCAN-based pipeline for unsupervised loitering detection under AIS sparsity; and (iii) a COG-based spectral analysis for automatic labeling and classification into six distinct loitering types.

We validate our framework on large-scale AIS datasets collected in the Korean Exclusive Economic Zone (EEZ), demonstrating its ability to uncover subtle, spatially diverse loitering activities with high interpretability. By emphasizing behavior estimation under uncertainty and eliminating the need for labeled supervision, this work advances scalable and operationally relevant solutions for maritime surveillance.

* Corresponding author

II. Related Work

Maritime anomaly detection has been widely studied, focusing on identifying irregular vessel behaviors such as loitering, route deviation, and suspicious rendezvous. Early approaches relied on rule-based systems and handcrafted features, while recent works emphasize machine learning and unsupervised methods.

Nascimento et al. [3] proposed a hybrid surveillance framework combining vessel motion features with expert rules to detect illegal activities, highlighting the benefits of multi-source fusion. Lane et al. [4] presented one of the earliest threat assessment systems based on AIS patterns, laying the foundation for anomaly-driven maritime monitoring. More recent efforts leverage clustering: Kumar et al. [5] employed HDBSCAN+ for detecting abnormal behaviors, while Wijaya and Nakamura [6] quantified loitering via dynamic AIS message features.

Loitering behavior refers to the act of a vessel remaining in a limited geographic area for an extended period while performing repetitive, irregular, or non-navigational maneuvers. According to the loitering detection study, such behavior is often characterized by low-speed movement with high spatiotemporal redundancy—such as circling, zig-zagging, or random turning—without a clear route progression. These patterns are commonly observed in scenarios involving illegal fishing, patrol operations, unauthorized rendezvous, or surveillance activities near protected maritime zones.

Estimating loitering patterns is challenging due to the variability in ship behavior and the absence of ground truth labels. Typical loitering indicators include low mean speed ($\bar{U}$), high turn rate ($\bar{\omega}_\chi$), directional entropy ($H(\chi)$), and trajectory redundancy metrics such as:

$$\psi = \frac{D}{PM}, \tag{1}$$

where D is the trajectory length and PM is the perimeter of the minimum bounding rectangle. Low values of ψ indicate spatially confined and repetitively turning behaviors. Enhanced versions of this metric incorporate heading and course changes [6] to better capture dynamic motion.

However, in real-world monitoring scenarios, these indicators are often unreliable due to the fragmented nature of AIS data. When AIS signals are sparse or missing—whether due to low transmission frequency in open sea, signal occlusion in narrow straits, or deliberate AIS deactivation by vessels attempting to evade detection—the observed trajectories become disjointed or incomplete. This lack of temporal continuity complicates the estimation of behavior-based metrics such as speed, turn rate, or redundancy.

Discrete and irregular AIS signals often result in broken trajectories with missing segments, making it difficult to observe the full behavior of loitering maneuvers such as circling or drifting. These gaps hinder accurate computation of motion features and may cause partial or complete misclassification of loitering behavior. For instance, short disconnected segments may miss the turning or dwelling phases of a loitering pattern, resulting in under-detection or misclassification.

Moreover, loitering behavior can span across time windows larger than typical AIS reporting gaps. Loitering often spans beyond typical AIS gaps, with sporadic transmissions breaking the behavioral arc. Conventional trajectory measures,and even CNN classifiers,fail when missing segments distort patterns, e.g., circling paths reduced to incomplete arcs resembling normal transit.

Trajectory reconstruction has also been a key research direction. Perera and Soares [7] applied an Extended Kalman Filter (EKF) for vessel trajectory estimation, and Mieczyńska et al. [8] integrated DBSCAN for AIS data recovery. Building on this, Thi et al. [9] introduced an Adaptive EKF variant tailored to maritime dynamics, showing improved handling of missing transmissions. Similar filtering techniques have been extended to unmanned surface vehicles [10].

For loitering detection, Zhang et al. [11] used trajectory shape descriptors and CNNs, while Liang et al. [12] proposed AIS-based anomaly detection using knowledge-driven representations, and Li et al. [13] combined autoencoders with Temporal Convolutional Networks for spatiotemporal anomaly discovery.

Compared to these studies, our framework uniquely integrates (i) adaptive trajectory reconstruction, (ii) inter-adaptive sliding window segmentation, and (iii) intra-segments unsupervised loitering classification. This design enables automatic labeling of loitering behaviors without manual intervention, providing both scalability and interpretability for real-world maritime domain awareness.

III. Proposed Model

In this section, we propose a comprehensive framework for detecting and classifying loitering vessels using reconstructed AIS data.

As illustrated in Figure 1, the methodology consists of two major components: (i) AIS trajectory reconstruction using Reconstructive-Adaptive Extended Kalman Filter (R-AEKF) and inter-trajectory segmentation and clustering via Adaptive Momentum HDBSCAN, (ii) intra-trajectory loitering detection using spectral concentration in sliding windows and loitering pattern classification mainly based on course-over-ground (COG) variations.

The proposed outlier detection framework leverages Automatic Identification System (AIS) data to detect irregular vessel behaviors, with a particular focus on identifying loitering — defined as the act of vessels idling or maneuvering without clear navigational objectives. This behavior is often indicative of unregulated maritime activities such as illegal fishing, smuggling, or surveillance. The framework is structured into three core stages: trajectory reconstruction, outlier detection, and loitering behavior classification. Each component is designed to address the inherent challenges of AIS-based maritime surveillance, including incomplete data, lack of labeled training examples, and behavior variability across vessel types and conditions.

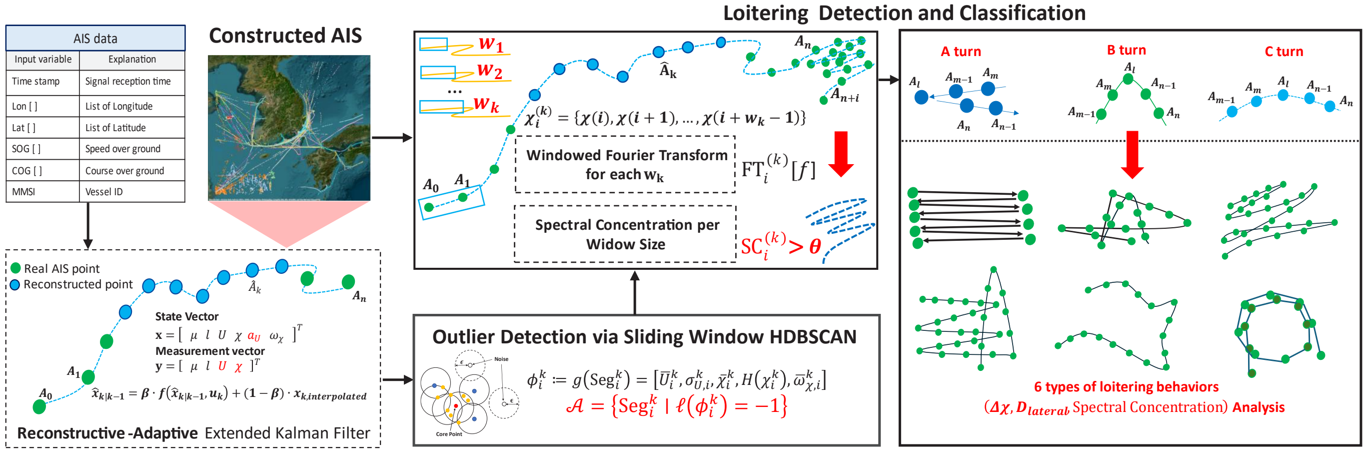

Fig. 1: Proposed framework for detecting and classifying loitering behaviors from reconstructed AIS data. The system integrates reconstruction **R-AEKF**, adaptive sliding window **HDBSCAN** segmentation, windowed spectral loitering detection, and classification based on angular variation and Fourier transform features.

Single AIS Trajectory Points: For a vessel uniquely identified by a Maritime Mobile Service Identity (MMSI), an AIS trajectory point is defined as:

$$A_i = \{t_i, \mu_i, l_i, U_i, \chi_i\} \tag{2}$$

where:

- $t_i \in \mathbb{R}$: Timestamp in Unix epoch seconds
- $\mu_i \in [-90°, 90°]$: Latitude
- $l_i \in [-180°, 180°]$: Longitude
- $U_i \in \mathbb{R}_{\geq 0}$: Speed Over Ground (SOG)
- $\chi_i \in [0°, 360°]$: Course Over Ground (COG)

A trajectory for vessel k is defined as $T^k = [A_1^k, A_2^k, \ldots, A_n^k]$.

Trajectory Reconstruction. Incomplete AIS data caused by signal loss or "going dark" results in discontinuities in T^k. We address this using a Reconstructive Adaptive Extended Kalman Filter (R-AEKF) to estimate the vessel's state vector $\hat{\mathbf{x}}_t = [\mu_t, l_t, U_t, \chi_t, a_{U,t}, \omega_{\chi,t}]^T$ from noisy or missing observations. For extended gaps, interpolation using cubic splines maintains kinematic continuity. The resulting trajectory $T^k_{\text{rec}} = [A_1^k, A_2^k, \ldots, A_m^k]$ satisfies $m \geq n$ and preserves realistic vessel dynamics.

Behavior-Based Outlier Detection: Reconstructed trajectories are segmented using adaptive sliding windows. A segment $\text{Seg}_i^k \subset T^k$ is described by the behavioral feature vector:

$$\phi_m = [\bar{U}_m, \sigma_{U,m}, \bar{\chi}_m, H(\chi_m), \bar{\omega}_{\chi,m}] \tag{3}$$

where:

- $\bar{U}_m$: Mean SOG
- $\sigma_{U,m}$: SOG standard deviation
- $\bar{\chi}_m$: Mean COG
- $H(\chi_m)$: COG entropy
- $\bar{\omega}_{\chi,m}$: Mean turn rate

Let $\mathcal{S} = \{\text{Seg}_i^k\}$ be the set of trajectory segments and define a feature map $g : \mathcal{S} \to \mathbb{R}^d$ by

$$\phi_i^k := g(\text{Seg}_i^k) = [\bar{U}_i^k,\ \sigma_{U,i}^k,\ \bar{\chi}_i^k,\ H(\chi_i^k),\ \bar{\omega}_{\chi,i}^k] \tag{4}$$

We run HDBSCAN on the feature set $\Phi = \{\phi_i^k\}$, obtaining labels $\ell(\phi_i^k) \in \{-1, 1, \ldots, C\}$ (with -1 denoting noise). We then declare as anomalous those original segments whose features are labeled noise:

$$\mathcal{A} = \{\text{Seg}_i^k \mid \ell(\phi_i^k) = -1\} \tag{5}$$

Loitering Behavior Classification: Anomalous segments are further classified into 6 interpretable loitering categories: **Oscillatory Loitering, Transition Loitering, Hesitant Loitering, Steady Loitering, Gradual Drift Loitering, Circular Loitering.**

Each class is derived from ϕ_m and refined using domain-specific knowledge (e.g., vessel type, geography). This allows distinction between legitimate and suspicious loitering behavior.

Overall, the framework addresses AIS gaps via R-AEKF and interpolation, performs label-free anomaly detection with HDBSCAN, and categorizes loitering behaviors using interpretable pattern rules.

A. Trajectory Reconstruction and Behavior-Based Outlier Segmentation

We use a two-stage pipeline: (i) *trajectory reconstruction* with a Reconstructive Adaptive EKF (R-AEKF) and (ii) *behavior-based segmentation* with density clustering.

The R-AEKF operates in two main phases:

1) **Prediction Phase:** The state and its covariance are propagated using the vessel dynamics model and interpolation smoothing.
2) **Update Phase:** When AIS measurements are available, the predicted state is corrected using the observed measurement, weighted by the Kalman gain.

State and measurement vectors of **R-AEKF** are

$$\mathbf{x}_t = [\mu_t, l_t, U_t, \chi_t, a_{U,t}, \omega_{\chi,t}]^\top, \tag{6}$$

$$\mathbf{y}_t = [\mu_t, l_t, U_t, \chi_t]^\top. \tag{7}$$

Algorithm 1 Unified R-AEKF Reconstruction and Adaptive Sliding HDBSCAN Segmentation

Inputs: AIS per vessel; dynamics $f(\cdot)$; measurement $h(\cdot)$
State-Transition Jacobian $\mathbf{F}_t$, measurement Jacobian $\mathbf{H}_t$
$(L_{\min}, L_{\max}, k_U, k_\chi, \gamma)$; HDBSCAN params
Output: Reconstructed $\{T^k_{\mathrm{rec}}\}$; anomalous segments $\mathcal{A}$
Init $\hat{\mathbf{x}}_0, \mathbf{P}_0$; base covariances $\mathbf{Q}_0, \mathbf{R}_0$
for each vessel k **do**
  Init $\hat{\mathbf{x}}_{0|0}, \mathbf{P}_{0|0}$; $\mathbf{Q}, \mathbf{R} \leftarrow \mathbf{Q}_0, \mathbf{R}_0$; $T^k_{\mathrm{rec}} \leftarrow \emptyset$
  for each time t **do**
    Predict: $\hat{\mathbf{x}}_{t|t-1} = f(\hat{\mathbf{x}}_{t-1|t-1})$
    $\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}$
    if $\mathbf{y}_t$ available **then**
      $\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^\top (\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\top + \mathbf{R})^{-1}$
      Update: $\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{y}_t - h(\hat{\mathbf{x}}_{t|t-1}))$
      $\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1}$
      adapt $\mathbf{Q}, \mathbf{R}$ from recent residuals
    else
      Blend w/ interp (Eq. (8)):
      $\hat{\mathbf{x}}_{t|t} = \beta\, \hat{\mathbf{x}}_{t|t-1} + (1-\beta)\, \mathbf{x}_t^{\text{interp}}$
      $\mathbf{P}_{t|t} \leftarrow \mathbf{P}_{t|t-1}$
    end if
    Append $\hat{\mathbf{x}}_{t|t}$ to T^k_{rec}
  end for
  Adaptive windowing: compute L_m via Eq. (9)
  Smooth $L_m^{\text{final}} = \gamma L_{m-1}^{\text{final}} + (1-\gamma) L_m$
  Segmentation: split T^k_{rec} using L_m^{final}
  Segment-wise features form the vector ϕ_m (Eq. (3))
end for
Robust-scale $\{\phi_m\}$; run HDBSCAN; set $\mathcal{A} = \{\mathrm{Seg}^k_m \mid \mathrm{label}(\phi_m) = -1\}$
return $\{T^k_{\mathrm{rec}}\}$, $\mathcal{A}$

R-AEKF follows predict–update while *adaptively* tuning covariances from recent residuals. Long gaps are spline-filled; predictions are blended with interpolation for stability:

$$\hat{\mathbf{x}}_{k|k-1} = \beta\, f(\hat{\mathbf{x}}_{k-1|k-1}) + (1-\beta)\, \mathbf{x}_k^{\text{interp}}, \;\; \beta \in [0,1]. \tag{8}$$

Adaptive segmentation & clustering. Reconstructed trajectories are partitioned by an inter-adaptive window to determine window size for segment m, we define L_m as:

$$\max\left(L_{\min}, \min\left(L_{\max},\ k_U \cdot \frac{\bar{U}_{m-1}}{1+\sigma_{U,m-1}} + k_\chi \cdot \frac{1}{1+|\bar{\omega}_{\chi,m-1}|}\right)\right) \tag{9}$$

Each resulting segment is encoded by a behavioral feature vector and HDBSCAN groups dense routine behavior and flags low-density segments as anomalies.

B. Multi-Scale Loitering Detection and Classification

After inter-trajectory anomaly mining, we localize loitering within each vessel path by detecting periodic heading patterns in multi-scale windows. Let $\chi_i^{(k)} = \{\chi(i), \ldots, \chi(i + w_k - 1)\}$ be the COG series in a window of size w_k. We compute its discrete spectrum by applying the Fast Fourier Transform (FFT) to obtain the frequency-domain representation [14]:

$$\mathrm{FT}_i^{(k)}[f] = \mathrm{FFT}\left(\chi_i^{(k)}\right)[f], \tag{10}$$

power $\mathrm{PO}_i^{(k)}[f] = |\mathrm{FT}_i^{(k)}[f]|^2$, and the *spectral concentration* (periodicity score)

$$\mathrm{SC}_i^{(k)} = \frac{\max_{f\ge 1} \mathrm{PO}_i^{(k)}[f]}{\sum_{f=1}^{w_k-1} \mathrm{PO}_i^{(k)}[f]}. \tag{11}$$

A segment is flagged as loitering at scale k if $\mathrm{SC}_i^{(k)} > \theta$. Because loitering behaviors occur at multiple durations, we fuse detections across window sizes $w_k \in \{w_1, \ldots, w_K\}$:

$$LO = \bigcup_{k=1}^{K} \left\{\chi_i^{(k)} \mid \mathrm{SC}_i^{(k)} > \theta\right\}. \tag{12}$$

a) Geometric cues and labels.: To distinguish stationary loitering from slow drift, we pair spectral evidence with simple geometry. Heading variability is measured by

$$\Delta\chi = \max(\chi) - \min(\chi), \tag{13}$$

and lateral drift by

$$D_{\text{lat}} = \sqrt{(\mu_{A_i} - \mu_{A_0})^2 + (l_{A_i} - l_{A_0})^2}. \tag{14}$$

We coarsely bin turns into $A: [0, 30^\circ)$, $B: [30^\circ, 120^\circ)$, $C: [120^\circ, 180^\circ)$. Combining $\Delta\chi$ (via $A/B/C$), SC , and drift D_{lat} yields six interpretable types:

- **Oscillatory (AA):** sharp back-and-forth; high SC, low D_{lat}.
- **Transition (AB):** mix of sharp/moderate turns; mid SC, mid D_{lat}.
- **Hesitant (AC):** irregular sharp/gradual; low SC, higher D_{lat}.
- **Steady (BB):** consistent moderate turning; high SC, low D_{lat}.
- **Gradual Drift (BC):** moderate/gradual with drift; low SC, high D_{lat}.
- **Circular (CC):** near-constant turning; very high SC, $D_{\text{lat}} \approx 0$.

b) Refinement: For borderline cases we add a lightweight descriptor—the *peak ratio (PR)* between adjacent spectral peaks in the COG power spectrum and train a small Random Forest on $\{\Delta\chi,\ D_{\text{lat}},\ H(\chi), \sigma_U, \mathrm{PR}\}$ to refine rule-based labels, preserving interpretability while improving consistency on noisy AIS.

IV. Experimental Evaluation

This section presents the experimental results of our proposed self-supervised framework for unsupervised maritime anomaly detection using AIS data. The evaluation includes two main parts: a detailed description of the dataset, evaluation metrics, and results for both trajectory reconstruction and clustering-based anomaly detection as well as loitering classification.

A. Dataset Description & Evaluation Metrics

We evaluate our method on a large-scale AIS dataset collected within the Korean Exclusive Economic Zone (EEZ) between November 2023 and March 2024. This dataset comprises over 36 million AIS messages from more than 21,000 unique vessels, including fishing, cargo, and passenger ships. The data was collected via both terrestrial and satellite receivers, with each AIS point containing key fields: MMSI, timestamp, latitude, longitude, speed over ground (SOG), and course over ground (COG).

The dataset poses the following challenges:

- **Frequent Data Gaps:** Over 95% of trajectorie AIS gaps longer than 10 minutes, primarily intentional transponder shutdowns or signal occ
- **Behavioral Diversity:** Vessels of different ty functions show varying motion patterns, nece adaptive analysis.

Preprocessing involved deduplication of vessel II ing of stationary vessels (SOG $<$ 0.5 knots), and t segmentation based on a temporal threshold of 180 ı

We adopt quantitative metrics to evaluate two core

1) Trajectory Reconstruction Accuracy: Measur Root Mean Square Error (RMSE) based on the l distance between predicted and true vessel positions

2) Clustering Quality: We assess clustering qua two standard indices: the *Silhouette* score [15] and the *Bouldin* index [16]. Silhouette summarizes cohesion aration (values near 1 indicate well-separated cluster near 0/negative suggest overlap or misassignment). Bouldin measures the average worst-case ratio of cluster scatter to between-cluster separation (lower i We report mean Silhouette and DBI across all se windows and configurations.

B. Experimental Results

1) Trajectory Reconstruction via R-AEKF: We eva Reconstructive Adaptive Extended Kalman Filter (F against multiple baselines: linear and cubic spline lation, standard Kalman Filter (KF), Reconstructive KF), and traditional AEKF. Table I reports RMSE kilometers.

The R-AEKF (cubic spline) achieved the lowes of 0.176 km, showing robust performance in hand gaps and non-linear motion. Its dynamic noise adjustı extended state modeling notably improve reconstru noisy and sparse AIS scenarios.

Figure 2 illustrates the improvements of R-AEKF over traditional AEKF. In (a), the original AIS trajectory exhibits substantial gaps due to signal loss or AIS-off events. The basic AEKF (b) partially interpolates motion but suffers from deviation during complex maneuvers. The R-AEKF (c) significantly improves reconstruction by leveraging both model-based predictions and spline-smoothed interpolation, producing trajectories with more realistic dynamics and better alignment with surrounding points.

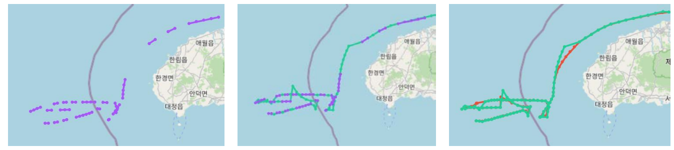

Fig. 2: Comparison of reconstructed vessel trajectories. (a) Initial AIS data with missing segments; (b) AEKF-based reconstruction vs. raw data; (c) Enhanced reconstruction by

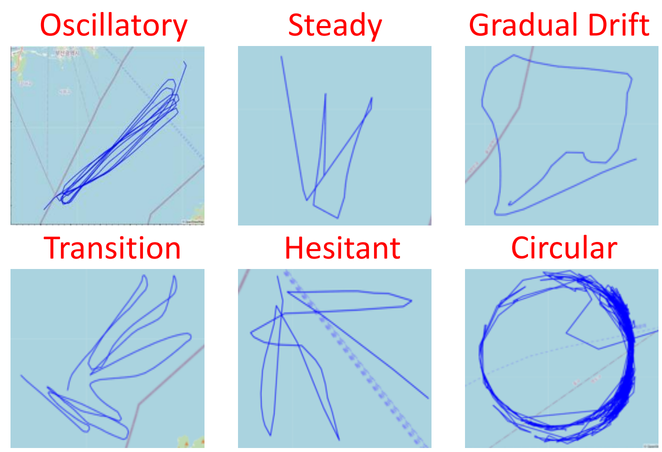

Fig. 3: Loitering Segments Trajectories Detection and six types Classifications

This refined estimation is particularly critical in anomaly detection, as precise path reconstruction affects downstream segmentation and loitering classification.

2) Outlier Detection with Inter-Adaptive Sliding Window HDBSCAN: To evaluate the effectiveness of our proposed inter-adaptive segmentation approach, we conducted a comparative analysis using three different trajectory segmentation strategies: (a) a fixed-length window with maximum size ($L_{\max} = 40$), b2) a fixed-length window with minimum size ($L_{\min} = 20$), and (c) our momentum-based adaptive windowing scheme described in equation (9).

Quantitative quality was assessed with Silhouette (cohesion/separation) and Davies–Bouldin (compactness vs. separation). With R-AEKF+HDBSCAN, the momentum-based adaptive windowing achieved the best scores (Silhouette = 0.459, DBI = 0.923), outperforming fixed-length baselines.

Adaptive windows resize to local dynamics (speed/heading changes), isolating brief or uneven behaviors (sharp turns, slow drift) more cleanly. Fixed windows misalign with transitions—either merging short anomalies or oversplitting steady motion—yielding weaker cluster structure.

3) Loitering Trajectory Classification: A total of 4,381 loitering segments were detected using the adaptive sliding window approach; examples of the six patterns (Oscillatory, Steady, Gradual Drift, Transition, Hesitant, Circular) are

Table I: Trajectory Reconstruction RMSE (km) Comparison

Method	RMSE (km)
Interpolation (Linear)	0.271
Interpolation (Cubic spline)	0.183
Kalman Filter (KF)	0.429
R-KF (Linear)	0.215
R-KF (Cubic spline)	0.198
Extended KF	0.354
R-AEKF (Linear)	0.179
R-AEKF (Cubic spline)	**0.176**

Table II: Clustering (R-AEKF+HDBSCAN) Results Analysis

Size	Clusters	Outliers	Silhouette	DB Index
L max (40)	2617	1881	0.432	0.928
L min (20)	2632	1967	0.440	0.945
L adaptive	2595	1837	**0.459**	**0.923**

shown in Fig. 3. We then compare a Random Forest (RF) with tuned hyperparameters against a Recurrent Neural Network (RNN) for loitering-type classification. Consistent with the qualitative distinctions in Fig. 3, the tuned RF achieves higher accuracy and F1 than the RNN while remaining lightweight for deployment. Hyperparameter tuning was performed on the RF model using GridSearchCV, optimizing for the best combination of parameters.

In Table III, the performance of different classification methods is presented. The RF (tuned) model achieve with the highest performance, achieving 89.57% accuracy. The RF model with tuning achieved the highest accuracy and F1-score, outperforming other methods like SVM, RNN, LSTM and GRU. Compared with deep learning baselines, the resulting pipeline has markedly lower computational and memory costs, enabling low-latency, real-time deployment without sacrificing performance.

V. Conclusion

We presented a streamlined, unsupervised AIS loitering-detection pipeline that stays robust with sparse, noisy, unlabeled data: R-AEKF reconstructs trajectories, adaptive windows segment them, HDBSCAN flags anomalies, and spectral/COG cues automatically label six loitering types—turning raw tracks into interpretable behaviors. By restoring continuity, adapting window sizes to local dynamics, and using density cues plus frequency signatures, the method reliably surfaces suspicious maneuvers such as illegal fishing, covert rendezvous, or spoofing.

The approach is modular and operationally ready: it can prioritize high-risk trajectories, reduce analyst workload, and ingest contextual layers (e.g., registries, restricted zones) to refine alerts. Its data efficiency and interpretability make it suitable for deployment in diverse maritime environments.

Acknowledgment

This work was supported by the Korea Institute of Marine Science and Technology Promotion (KIMST) funded by the Korea Coast Guard, under the Integrated Satellite-Based Applications Development for Korea Coast Guard project, Grant RS-2023-00238652.

Table III: Loitering Trajectory Classification Performance Comparison

Method	Acc (%)	Pre (%)	Recall (%)	F1-Score (%)
SVM	75.22	67.34	70.74	68.99
RNN	64.57	64.82	61.70	63.20
LSTM	73.94	75.77	71.94	73.80
GRU	76.69	84.48	73.37	73.97
RF (tuned)	**89.57**	**86.09**	**87.78**	**86.94**

References

[1] Wolsing, K., Roepert, L., Bauer, J., & Wehrle, K. (2022). Anomaly detection in maritime AIS tracks: A review of recent approaches. Journal of Marine Science and Engineering, 10(1), 112.

[2] Nguyen, D., Vadaine, R., Hajduch, G., Garello, R., & Fablet, R. (2021). GeoTrackNet—A maritime anomaly detector using probabilistic neural network representation of AIS tracks and a contrario detection. IEEE Transactions on Intelligent Transportation Systems, 23(6), 5655-5667.

[3] do Nascimento, V. D., Alves, T. A., de Farias, C. M., & Dutra, D. L. C. (2024). A hybrid framework for maritime surveillance: Detecting illegal activities through vessel behaviors and expert rules fusion. Sensors, 24(17), 5623.

[4] Lane, R. O., Nevell, D. A., Hayward, S. D., & Beaney, T. W. (2010, July). Maritime anomaly detection and threat assessment. In 2010 13th International conference on information fusion (pp. 1-8). IEEE.

[5] Hari Kumar, R., C. P. Ramanarayanan, and K. S. R. R. P. Murthy. ”Detection of Abnormal Vessel Behaviours Based on AIS Data Features Using HDBSCAN+.” Defence science journal 73.4 (2023).

[6] Wijaya, W. M., and Y. Nakamura. ”Loitering behavior detection by spatiotemporal characteristics quantification based on the dynamic features of AIS messages.” PeerJ Computer Science 7 (2022): e1572.

[7] Perera, L. P., and C. Guedes Soares. ”Ocean vessel trajectory estimation and prediction based on extended Kalman filter.” 2010 2nd International Conference on Adaptive and Self-Adaptive Systems and Applications. 2010.

[8] Mieczyńska, Marta, and Ireneusz Czarnowski. ”DBSCAN algorithm for AIS data reconstruction.” Procedia Computer Science 192 (2021): 2512-2521.

[9] Thi, Q. A. M., Lee, C., Tao, T. M., & Youn, C. H. (2024, October). Tracking Vessel Activities with AIS Data using an Adaptive Extended Kalman Filter. In 2024 15th International Conference on Information and Communication Technology Convergence (ICTC) (pp. 349-354). IEEE.

[10] Fossen, S., and T. I. Fossen. ”Five-state extended Kalman filter for estimation of SOG, COG and course rate of unmanned surface vehicles: Experimental results.” Sensors 21.23 (2021): 7910.

[11] Zhang, Z., Huang, L., Peng, X., Wen, Y., & Song, L. (2022). Loitering behavior detection and classification of vessel movements based on trajectory shape and Convolutional Neural Networks. Ocean Engineering, 258, 111852.

[12] Liang, M., Weng, L., Gao, R., Li, Y., & Du, L. (2024). Unsupervised maritime anomaly detection for intelligent situational awareness using AIS data. Knowledge-Based Systems, 284, 111313.

[13] Li, Z., Sun, Y., Yang, L., Zhao, Z., & Chen, X. (2022). Unsupervised machine anomaly detection using autoencoder and temporal convolutional network. IEEE Transactions on Instrumentation and Measurement, 71, 1-13.

[14] Harris, Charles R., et al. ”Array programming with NumPy.” Nature 585.7825 (2020): 357–362.

[15] Rousseeuw, Peter J. ”Silhouettes: A graphical aid to the interpretation and validation of cluster analysis.” Journal of Computational and Applied Mathematics 20 (1987): 53–65.

[16] Davies, David L., and Donald W. Bouldin. ”A cluster separation measure.” IEEE transactions on pattern analysis and machine intelligence 2 (2009): 224-227.