Dark Activity Detection in AIS-Based Maritime Networks

Bekir Nazmi Görkem, Burak Çağlayan, Erkam Karaca, Candar Karabulut and Ömer Korçak
Computer Engineering Department
Marmara University
Istanbul
Email: omer.korcak@marmara.edu.tr

***Abstract*—The Automated Identification System (AIS) is an indispensable tracking system employed in the maritime industry for vessel identification, location tracking, and collision avoidance. While AIS messages provide essential information for maritime traffic management, they also present challenges when vessels aim to conduct operations discreetly or evade observation. This phenomenon, referred to as "dark activity", involves intentional AIS deactivation by vessel operators seeking to conceal their actions, often related to illicit or illegal maritime activities such as smuggling, piracy or illegal fishing. The detection and monitoring of dark activities pose significant challenges for law enforcement and security agencies. This paper explores innovative approaches to address this issue by harnessing AIS data and incorporating rule-based techniques, as well as machine learning techniques to enhance maritime security efforts. We adopted a local approach where a dark activity of a vessel is detected by nearby ships depending on the previous signals. We implemented a detailed simulation environment based on real and realistic data to run the proposed algorithms. Simulation results show that while rule-based approach is successful in detecting dark activities, it tends to produce false alarms, and ML-based approach provides better overall accuracy.**

I. Introduction

In the vast expanse of open seas, vessels communicate critical information to ground stations through satellites or direct means. This communication takes the form of an AIS signal, which contains essential data like the vessel's current position, speed, course over ground, and heading [1]. These parameters are updated automatically and broadcasted throughout specific time intervals depending on vessel's movement and message type. However nothing can stop vessels from turning off their AIS transmitters and going "dark" intentionally. The authorities cannot tell for sure whether a ship switched off its AIS to hide its location for some illicit dark activities (e.g., smuggling, fishing in restricted areas, or unauthorized waste disposal) [2], or its AIS signal cannot be received due to natural reasons such as heavy whether condition or signal congestion. Particularly in the open seas, where satellite-based AIS (S-AIS) is used, the ships can always deny this dark activity because the signal can be lost the way to the satellite especially in congested waters. Thus, only a nearby ship can be aware of the dark activity. The ratio between lost signals unintentionally due to conditions beyond AIS and deactivating the transmitter on purpose is something between 1:10 and 1:20 depending on ship type and geographical area according to Ron Crean, vice-president for commercial at Windward Maritime Analytics [3]. However, we believe that vessels working cooperatively can overcome this issue; obviously, not if the AIS was turned off before sailing in the first place.

Recognizing the significance of AIS signals for maritime security, we embarked on a mission to develop a rule-based decision-making algorithm. This algorithm is designed to help vessels meeting specific range criteria determine if a neighboring vessel has entered a state of dark activity. To realize our vision, we crafted a simulation environment that vividly portrays vessel movements, including scenarios involving dark activity. We harnessed machine learning models to enhance the accuracy of predictions, utilizing data gathered from our simulation experiments.

This paper stands as a fundamental tool, contributing to the enhancement of maritime security by combatting smuggling and curbing illegal activities on the high seas. Moreover, it plays a pivotal role in safeguarding marine life and preserving the delicate ecological balance by detecting illegal fishing activities. Our primary objective revolves around the detection of illegal vessel activities within specific sea regions. Leveraging AIS signals, we pursued the following key steps:

- Creating a realistic simulation environment with a user-friendly interface, enabling the visualization of vessel movements and the adjustment of parameters related to dark activity.
- Designing and implementing a rule-based decision-making algorithm to identify vessels entering a state of dark activity.
- Gathering a dataset from the simulation environment to fuel machine learning algorithms.
- Simulating both algorithms within the environment and assessing their accuracy

Through these endeavors, we aim to strengthen maritime security, safeguard the environment, and ensure responsible and legal conduct on the high seas.

A. Related Work

In the literature, there are several studies on dark activity detection. Shahir et al. [4] addressed the critical issue of maritime domain awareness, emphasizing its significance in preventing smuggling and safeguarding vital sea-based structures. Their solution consisted of three phases: i) Engagement Detection: Vessels in close proximity were clustered together; ii) Detection of Candidates: Candidates for engagement were identified based on kinematic features, particularly slow speeds and converging courses or close proximity; iii) Scenario Detection: Leveraging the results of engagement detection, scenarios were represented by left-to-right Hidden Markov Models and classified using Support Vector Machines. Moreover, an additional phase was introduced for anomaly detection, extending the scope of scenario detection to rectify misclassified scenarios.

Mantecon et al. [5] address the challenge of maritime threats and illegal activities, and employ convolutional neural networks (CNN) to derive navigation patterns based on ship speed, direction, and maneuverability, using a dataset called DeepMarine, derived from historical AIS data. Then both AIS and Radar trajectories can be used to identify various vessel behaviors. The proposed method requires both the positional data and ship information to detect illegal activities such as fishing in non-allowable areas. The authors in [2] presented an anomaly detection methodology to discriminate between AIS messages that are not received by base stations due to communication channel-related effects and those that were not broadcasted at all to cover dark activities. The strength of the received signal RSSI is analyzed to detect On/Off switching of the ship's transponder. A training set of known good AIS data is used for comparison with received data, an alert is triggered when signal dropouts exceed a defined threshold. Eaton et al. [6] introduced a novel dark activity detection concept called Sensors and Platforms for Unmanned Detection of Dark Ships (SPUDDS), which combines hardware and software components. SPUDDS involves an autonomous buoy equipped with various sensors and software called CROWSNEST, designed for ship identification and classification. The system accurately categorizes detected ships, including sailboats, merchant ships, and fishing vessels, using a highly precise machine learning algorithm. CROWSNEST relies on a convolutional neural network and data-driven ODF for ship classification. The integration of a 360-degree camera on the buoy enhances its capabilities for safeguarding maritime security. Paolo et al [7] suggests using Synthetic Aperture Radar (SAR) images and automated machine learning in order to detect illegal fishing activities. They constructed and released xView3-SAR dataset for maritime object detection and characterization, and combine AIS and human annotations for labeling the data. Bereta et al. [8] employed satellite imaging techniques to achieve a 95% accuracy rate in detecting Dark Activities. They emphasized the limitations of relying solely on AIS signals, leading them to propose a hybrid approach using satellite data, specifically Copernicus Sentinel imaging, in conjunction with Marine Traffic AIS data to monitor ship density in areas of potential Dark Activity. The project involved acquiring data from the Alaska Satellite Facility and Copernicus Data Hub, followed by preprocessing to remove irrelevant image portions and filtering out cloud-obscured images, reducing data volume from terabytes to gigabytes. The data fusion step synthesized satellite and AIS data, aligning the satellite image time with AIS data from 30 minutes before and after the image capture. Utilizing the K-nearest neighbors (KNN) method, they matched satellite images with AIS data, identifying ships present in images but absent from AIS records as potential participants in Dark Activity.

This paper introduces a distinct approach from previous studies. Due to the fact that only nearby ships can notice a vessel turning off its AIS transponder, ships working co-operatively can be crucial in detecting dark activities. Thus, instead of depending on external sources like satellite imagery, buoys with 360-degree cameras, or X-band radar systems, we solely rely on local AIS signals. We put forward rule-based and machine-learning-based algorithms to identify dark activities of nearby vessels in real-time when their AIS transmitters are deactivated. Our approach offers a cost-effective solution compared to alternative methods and can be used alongside more expensive solutions to enhance the overall efficacy of illegal activity detection.

II. Methodological Background

We consider a system model where every vessel may verify the activities and status of nearby vessels through AIS signals. The horizontal range of vessel-to-vessel AIS signals is 20-30 nautical miles under most atmospheric conditions [9]. AIS is obligatory for ships that meet specific criteria. According to IMO (International Maritime Organization), passenger vessels irrespective of size, all ships engaged on international voyages with size of 300 gross tonnage, and cargo ships of 500 gross tonnage are required to have AIS transmitter[10]. AIS signals should be transmitted at intervals ranging from 2 to 180 seconds, with the specific interval determined by the velocity and the change in the course of the vessels [11].

When a vessel broadcasts an AIS signal, nearby vessels equipped with AIS receivers within the coverage area will receive it. Therefore, if vessel A receives the AIS signal from vessel B but then stops receiving it, there could be three possible reasons: i) Vessel A has moved out of the coverage area of vessel B due to mobility, ii) Signal collision and interference have occurred, iii) Vessel B has entered dark activity.

In our system model, vessels have the capability to verify the activities of nearby vessels and determine if they have entered dark activity. These vessels are referred to as detector vessels. To provide a clear explanation, we will concentrate on a scenario involving a single detector, represented as vessel 5 in Fig.1. As depicted in the figure, each vessel has a predefined broadcasting range for its AIS signal, which varies depending on its type. The green zones indicate areas where vessels' signals can be received by the selected detector vessel, while the red zones signify that their broadcasting range is insufficient to transmit their AIS signal to the detector vessel.

At time t, vessel 5 receives signals from vessels 1 and 4 but cannot receive signals from vessels 2 and 3. In the subsequent time step, vessel 3 comes within range, while vessel 4 goes out of range. Even though vessel 4's signal is no longer received, vessel 5 refrains from making a dark activity decision because

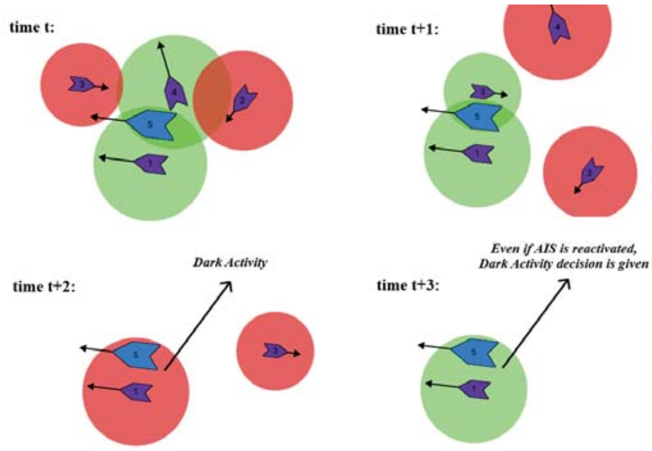

Fig. 1. Dark Activity Scenario: Vessel 1 enters dark activity, detected by vessel 5.

this situation was expected based on the vessel's location, course, and speed in the previous time step. However, at time t+2, the signal from vessel 1 is unexpectedly not received, prompting a dark activity decision. Even if vessel 1 reactivates its AIS signal after a period of dark activity, the decision of "possible dark activity" persists.

III. Dark Activity Detection

To accurately detect dark activities, a critical challenge lies in predicting a ship's future position based on the AIS data received at the current moment. A straightforward method involves utilizing vessel kinematics, considering factors like location, speed, and course. However, this approach may yield incorrect results if the vessel alters its course, speed, or other parameters. To enhance prediction accuracy, a machine learning approach becomes imperative. The subsequent subsections will elaborate on the proposed methods.

A. Rule-based Dark Activity Detection (R-DAD) Algorithm

A detector vessel continuously monitors AIS signals from nearby vessels, recording their transmitted parameters. It calculates the expected positions of these vessels in the next time step by applying fundamental physics principles to the received AIS data. These calculated positions, along with the corresponding AIS parameters, are stored. During each time step, the detector vessel checks whether the estimated positions of nearby vessels from the previous time step fall within its reception range. If a vessel's estimated position is within the detector vessel's reception range but no corresponding AIS signal is received, it triggers a potential alert for dark activity detection.

To monitor nearby vessels, the detector vessel maintains a database of received AIS signals, as depicted in the flowchart presented in Fig. 2. Subsequently, based on the AIS signal intervals, a designated time period is established to assess the potential occurrence of dark activities. Fig. 3 displays the procedural workflow of the R-DAD algorithm. When the AIS signal from a vessel was previously acquired but is

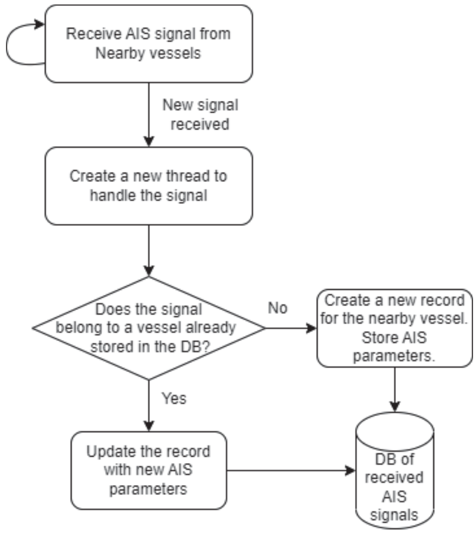

Fig. 2. Flowchart for the process of recording nearby AIS signals by the detector vessel

absent in the current temporal segment, a verification process is conducted concerning the estimated vessel location. This estimation is derived from previously acquired AIS parameters and utilizes kinematic principles for computation.

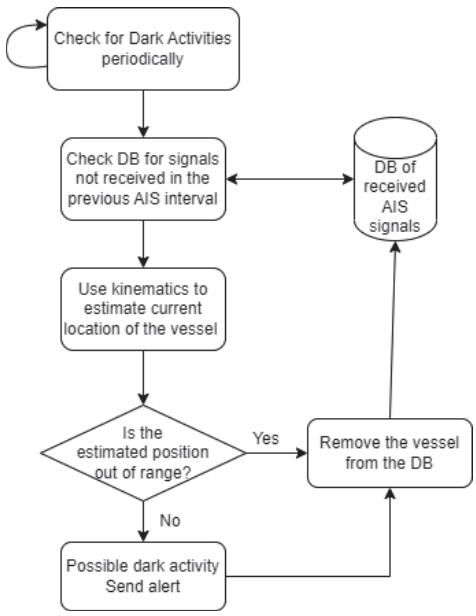

Fig. 3. Flowchart of the Rule-based Dark Activity Detection (R-DAD) algorithm

In this context, we employ the Haversine formula to compute both the expected vessel location and the distance between vessels. The Haversine formula serves as a precise method for calculating the distance d between two points on the surface of a sphere, based on their respective latitudes (ϕ_1, ϕ_2) and longitudes (λ_1, λ_2). It is defined as follows:

$$a = \sqrt{\sin^2(\frac{\phi_2 - \phi_1}{2}) + \cos\phi_1 \cdot \cos\phi_2 \cdot \sin^2(\frac{\lambda_2 - \lambda_1}{2})} \quad (1)$$

$$d = 2R \cdot \arcsin(a) \quad (2)$$

where longitudes and latitudes are defined in radians and R is the radius of the Earth.

To find the estimated location of a vessel, we first calculate the covered distance with its given speed, assuming that the speed is constant. Then given the initial latitude (ϕ_1) and longitude (λ_1), initial bearing (θ, clockwise from north), and the covered distance, the estimated location (ϕ_2,λ_2) is calculated with the following formula [12]:

$$\phi_2 = \arcsin(\sin\phi_1 \cdot \cos\delta + \cos\phi_1 \cdot \sin\delta \cdot \cos\theta) \quad (3)$$

$$\lambda_2 = \lambda_1 + \arctan 2(\sin\theta \cdot \sin\delta \cdot \cos\phi_1, \cos\delta - sin\phi_1 \cdot \sin\phi_2) \quad (4)$$

where δ is the angular distance d/R.

If the estimated location is out of range of the detector vessel, in other words if the distance calculated by (2) is more than the AIS range, this scenario is categorized as a typical operational condition. However, when the estimated location resides within the range, yet no AIS signal is detected, then this is interpreted as a potential dark activity, and an alert is triggered. It is noteworthy that one can establish a probabilistic assessment of the likelihood of dark activity based on the proximity of the anticipated position to the boundary of the AIS signal range. This approach necessitates further investigative exploration and research.

B. Machine Learning Based Dark Activity Detection (ML-DAD) Algorithm

In order to improve the precision of detecting dark activities, we employ a machine-learning technique. Initially, we create a dataset through a realistic simulation detailed in Section IV-A, which encompasses the attributes listed in Table I. Each data instance is also paired with a target value denoting whether it corresponds to a dark activity or not. The dataset comprises information about the detector vessel's course, heading and speed, as well as those of the selected nearby vessel. Additionally, the distance between two vessels and the AIS range of the nearby vessel is also used. Notably, a unique feature included in the dataset is the distance to the turn point, a parameter not typically conveyed in current AIS messages. It is worth mentioning that the AIS protocol accommodates a total of 64 message types, with 27 of them already allocated for specific purposes. The introduction of a novel message type featuring the distance to the turn point holds the potential to enhance maritime traffic control, and this study presupposes its utilization for improved accuracy.

Utilizing the dataset we constructed, we train a machine-learning model through a supervised learning algorithm. Subsequently, for each nearby vessel whose AIS signal was received in the previous time step but is currently not being received, we employ the model to make predictions regarding the presence of dark activity. Fig. 4 illustrates the ML-DAD algorithm. Instead of location estimation, dark activity decision is done according to the trained model.

TABLE I. FEATURES USED IN MACHINE LEARNING

Features	Target
ais_range	dark activity
distance_from_nearby_vessel	
nearby_vessel_heading	
nearby_vessel_course	
nearby_vessel_speed	
nearby_vessel_distance_to_turn_point	
heading	
course	
speed	
distance to turn point	

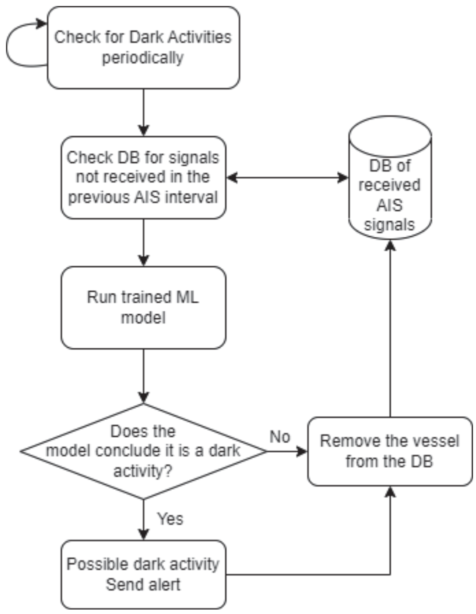

Fig. 4. Flowchart of the Machine-Learning-based Dark Activity Detection (ML-DAD) algorithm

IV. SIMULATION

A. Simulation Setup

To overcome the cost associated with obtaining real-time S-AIS data, we created a realistic dataset by leveraging actual port locations, established routes, and real-time vessel location data from MarineTraffic [10]. This dataset is stored in JSON format and comprises three static data types: ports, routes, and vessel data.

The **ports** data represents widely used ports and includes three fields: name, latitude (lat), and longitude (long). We initially compiled this static data by observing real maritime networks.

The **routes** data, also static, contains five fields: ”from,” ”to,” ”density,” ”noise,” and ”coordinates.” ”From” and ”to” denote the starting and final ports, respectively. ”Coordinates” is an array containing the latitude and longitude of intermediate turning points along the route. ”Density” and ”noise” are

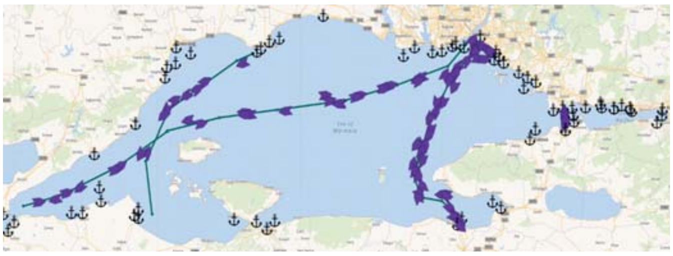

Fig. 5. A snapshot from the simulation map

arrays in which density[i] and noise[i] specify vessel density and noise levels between coordinates[i] and coordinates[i+1].

Vessel data is randomly generated along the predefined routes. The number of vessels between coordinates is determined by the corresponding density value from the route data. These vessels are randomly distributed between two coordinates, with their distance from the route randomly chosen within the range of -noise to +noise. Vessel data encompasses attributes such as Maritime Mobile Service Identity (mmsi), type, AIS range, AIS Interval, course, heading, speed, longitude, and latitude. The course of the vessel is calculated according to the following equation:

$$\theta = \arctan\left(\frac{\sin\Delta\lambda \cdot \cos\phi_2}{\cos\phi_1 \cdot \sin\phi_2 - \sin\phi_1 \cdot \cos\phi_2 \cdot \cos\Delta\lambda}\right) \quad (5)$$

where ϕ_1, λ_1 and ϕ_2, λ_2 are latitudes and longitudes (in radians) of starting point and ending point respectively, and $\Delta\lambda$ is the difference in the longitudes. Additionally, a binary variable named "dark_activity" is defined to simulate vessel dark activity.

We focused our study on the Marmara Sea area and simulated a scenario involving 87 vessels following predefined routes. Circular routes were designated for fishing vessels, and Fig. 5 provides a snapshot of a specific time instance within this scenario. Our simulation environment allowed for the selection of any vessel as the detector vessel, while any vessel other than the detector could be chosen to enter dark activity.

Fig. 6 shows another snapshot from simulation screen, where the detector vessel is selected and shown by red color, and the nearby vessels whose signals are detected are shown by yellow color. AIS signals of the vessels that are shown in dark blue color do not reach to the detector vessel, either due to long distance, or due to switching of the AIS transmitter and entering to dark activity.

Fig. 7 illustrates the basic system architecture of the simulation environment.

The simulation was executed through a total of 10 iterations, each spanning 1000 time steps. Every 10 time steps, a detector vessel was randomly chosen, and during each time step, a vessel was selected at random to engage in dark activity. We conducted experiments employing both R-DAD and ML-DAD algorithms. Within the ML-DAD algorithm, we utilized a variety of supervised machine learning techniques, including K-Nearest Neighbors, Decision Trees, Artificial Neural Networks

Fig. 6. A snapshot from the simulation screen depicting the AIS signals detected by a detector vessel

(with bagging), Support Vector Machines, Logistic Regression, and AdaBoost Random Forest. 30% of the data is allocated for the training set, while the remaining 70% is designated for the test set.

B. Numerical Results

Following the execution of the simulation involving both R-DAD and ML-DAD algorithms, with the latter employing various machine learning models, we assess the outcomes based on the accuracy of dark activity detections. We categorize these assessments into the following cases:

- True Positive (TP): A vessel is not involved in dark activity, and the prediction is accurate.
- False Negative (FN): A vessel is not involved in dark activity, yet the prediction is incorrect.
- False Positive (FP): A vessel is involved in dark activity, but the prediction is incorrect.
- True Negative (TP): A vessel is involved in dark activity, and the prediction is correct.

Then, the accuracy is calculated according to the following equation:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

Accuracy ratio for R-DAD algorithm is found to be 0.892, while the accuracy of the ML-DAD algorithm can be seen in Table II. AdaBoost random forest emerges as the best-performing supervised machine-learning algorithm with an

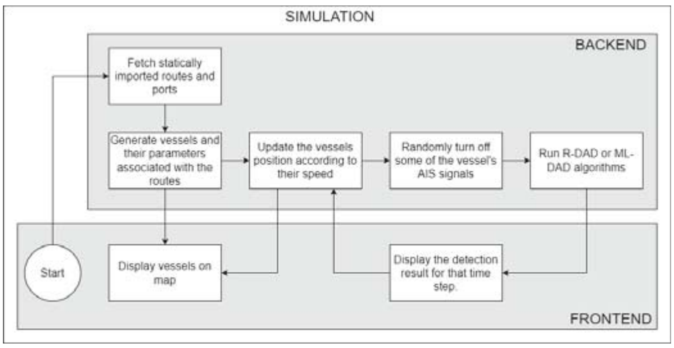

Fig. 7. System architecture of the simulation environment

TABLE II. ACCURACY RESULTS OF ML-DAD FOR VARIOUS MACHINE LEARNING ALGORITHMS

ML Algorithm	Accuracy
K-Nearest Neighbors	0.942
Decision Tree	0.952
Artificial Neural Networks	0.898
Suppor Vector Machines	0.942
Logistic Regression	0.864
AdaBoost Random Forest	**0.961**

accuracy ratio of 0.961. Table III displays the confusion matrix for the R-DAD algorithm, while Table IV presents the confusion matrix for the ML-DAD algorithm utilizing the AdaBoost random forest model. It is worth noting that, in the machine-learning-based approach, although there is a slight increase in false positive cases, there is a significant decrease in false negatives, resulting in a substantial improvement in the overall accuracy ratio.

TABLE III. CONFUSION MATRIX FOR R-DAD ALGORITHM

		Predicted condition: Positive	Predicted condition: Negative
Actual Condition	**Positive**	504	74
	Negative	5	151

TABLE IV. CONFUSION MATRIX FOR ML-DAD ALGORITHM

		Predicted condition: Positive	Predicted condition: Negative
Actual Condition	**Positive**	574	4
	Negative	20	136

V. CONCLUSION

This paper has addressed the crucial issue of detecting dark activities within AIS-based maritime networks, with a primary focus on enhancing security and safety at sea, particularly in identifying illicit engagements. Our examination primarily centers on a localized detection scenario, where vessels that deactivate their AIS transmitters are identified by nearby vessels based on historical signals received from them.

We introduced two distinct approaches to address this challenge. Firstly, we proposed a rule-based method grounded in kinematic estimations. Subsequently, we presented a machine learning approach leveraging a dataset created from a realistic simulation environment, which was designed with a graphical user interface using real data. Both approaches were evaluated within a scenario wherein a random subset of vessels engaged in dark activity.

The rule-based dark activity detection algorithm demonstrated strong performance in identifying dark activities but exhibited limitations, particularly in instances where AIS signals were not received due to changes in vessel movement parameters, leading to false alarms. On the other hand, the machine-learning-based dark activity detection approach yielded improved prediction accuracy and a reduced occurrence of false predictions in lawful cases, provided an appropriate supervised model was selected.

This research contributes valuable insights into the detection of dark activities in maritime networks, shedding light on the effectiveness and limitations of rule-based and machine-learning-based approaches. Further exploration of these methods and their integration into real-world maritime security systems holds promise for enhancing safety and security at sea. A potential avenue for future research involves the development of a collaborative system wherein multiple detector vessels engage in intercommunication to collectively identify proximate illicit activities. Another promising area of investigation entails the utilization of machine learning techniques to ascertain regions with elevated susceptibility to covert activities. Subsequently, this spatial information can be incorporated into dark activity detection algorithms to enhance the overall accuracy of the obtained results.

ACKNOWLEDGMENT

This work is supported by Marmara University, Faculty of Engineering.

REFERENCES

[1] D. Yang, L. Wu, S. Wang, H. Jia, and K. Li, “How big data enriches maritime research – a critical review of automatic identification system (ais) data applications,” *Transport Reviews*, vol. 39, pp. 1–19, 07 2019.

[2] F. Mazzarella, M. Vespe, A. Alessandrini, D. Tarchi, G. Aulicino, and A. Vollero, “A novel anomaly detection approach to identify intentional ais on-off switching,” *Expert Systems with Applications*, vol. 78, 02 2017.

[3] J. B. at Global Trade Review. (2020) Dark activity: As sanctions pressure rises, maritime trade turns to tech. Retrieved 12 September 2023. [Online]. Available: https://www.gtreview.com/magazine/volume-18-issue-3/dark-activity-sanctions-pressure-rises-maritime-trade-turns-tech/

[4] H. Y. Shahir, U. Glasser, A. Y. Shahir, and H. Wehn, “Maritime situation analysis framework: Vessel interaction classification and anomaly detection,” in *2015 IEEE International Conference on Big Data (Big Data)*. Ieee, 2015, pp. 1279–1289.

[5] T. Mantecón, D. Casals, J. J. Navarro-Corcuera, C. R. del Blanco, and F. Jaureguizar, “Deep learning to enhance maritime situation awareness,” in *2019 20th International Radar Symposium (IRS)*. IEEE, 2019, pp. 1–8.

[6] R. S. Eaton, S. German, and A. Balasuriya, “Maritime border security using sensors, processing, and platforms to detect dark vessels,” in *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*. IEEE, 2018, pp. 1–5.

[7] F. Paolo, T.-t. T. Lin, R. Gupta, B. Goodman, N. Patel, D. Kuster, D. Kroodsma, and J. Dunnmon, “xview3-sar: Detecting dark fishing activity using synthetic aperture radar imagery,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 37 604–37 616, 2022.

[8] K. Bereta, I. Karantaidis, and D. Zissis, “Vessel traffic density maps based on vessel detection in satellite imagery,” in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2022, pp. 2845–2847.

[9] D. Kocak and P. Browning, “Real-time ais tracking from space expands opportunities for global ocean observing and maritime domain awareness,” 10 2015, pp. 1–7.

[10] marinetraffic.com. Marine traffic. Retrieved 15 September 2023. [Online]. Available: https://marinetraffic.com

[11] M. Redoutey, E. Scotti, C. Jensen, C. Ray, and C. Claramunt, “Efficient vessel tracking with accuracy guarantees,” 12 2008, pp. 140–151.

[12] M. T. Ltd. Calculate distance, bearing and more between latitude/longitude points. Retrieved 12 September 2023. [Online]. Available: https://www.movable-type.co.uk/scripts/latlong.html