

Computational model of multi-source data fusion for submarine cable threat detection

Jinyao Liu

hd2008dev@sina.cn

Huadian(Fujian) Wind Power Co. Ltd

Boqing Li

Huadian(Fujian) Wind Power Co. Ltd

Tianling Chen

Fujian Zhongshisuo Electric Power Testing and Commissioning Co., Ltd

Fumin Bian

Fujian Zhongshisuo Electric Power Testing and Commissioning Co., Ltd

Research Article

Keywords: Submarine Cable Monitoring, Multi-source Data Fusion, Automatic Identification System Spoofing Attack, Intelligent Early Warning Model

Posted Date: September 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10692685/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Discover Artificial Intelligence on October 23rd, 2026. See the published version at <https://doi.org/10.1007/s44163-026-02461-z>.

Computational model of multi-source data fusion for submarine cable threat detection

Jinyao Liu*, Boqing Li

Huadian(Fujian) Wind Power Co. Ltd.,

Fuqing, 350318, China

Email: hd2008dev@sina.cn

Email: 978649244@qq.com

*Corresponding author

Tianling Chen and Fumin Bian

Fujian Zhongshisuo Electric Power Testing and Commissioning Co., Ltd.,

Fuzhou, 350026, China

Email: 1273818428@qq.com

Email: 13514091024@163.com

Abstract: Submarine cable monitoring relies on automatic identification system single data source, which is vulnerable to spoofing attack and leads to early warning failure. This paper proposes a multi-source data fusion computing model for submarine cable threat detection. The cable topology, automatic identification system trajectory and underwater acoustic data are mapped to a unified spatio-temporal graph structure, and cross-domain features are extracted by graph convolution and attention mechanism. Dempster-shafer evidence theory is introduced to construct conflict quantification and adaptive discount mechanism, and sonar evidence is preferentially accepted when digital and physical evidence conflict. Experiments based on real ocean data sets show that the accuracy of the model is 95.4%, which is 18% higher than that of the single source method. The F1-score maintains 0.910 under 30% automatic identification system spoofing attack. The model provides a scalable multi-source fusion computing framework for critical infrastructure security.

Keywords: Submarine Cable Monitoring; Multi-source Data Fusion; Automatic Identification System Spoofing Attack; Intelligent Early Warning Model

Biographical notes: Jinyao Liu is an Engineer at Huadian (Fujian) Wind Power Co., Ltd, China. He received a Bachelor's degree from Fujian normal University (2016). His research interests include new energy monitoring and power system automation.

Boqing Li is an Engineer at Huadian (Fujian) Wind Power Co., Ltd, China. He received a Bachelor's degree from Fujian Agriculture and Forestry University (2019). His research interests include power system automation.

Tianling Chen is an Engineer at Fujian Zhongshisuo Electric Power Testing and Commissioning Co., Ltd., China. She obtained a Master's degree from Xi'an Jiaotong University (2025). Her research interests include intelligent monitoring and operation and maintenance of offshore wind power.

Fumin Bian is an Engineer at Fujian Zhongshisuo Electric Power Testing and Commissioning Co., Ltd., China. He obtained a Bachelor's degree from Ningde Normal University (2017). Her research interests include intelligent monitoring and operation and maintenance of offshore wind power.

1 Introduction

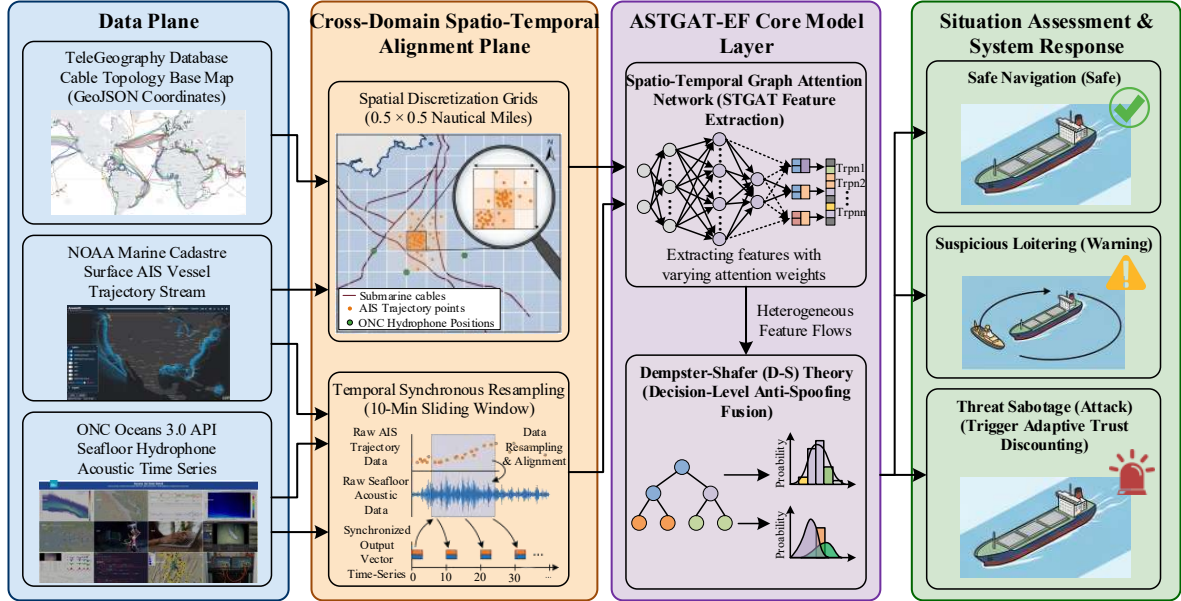
Undersea communication cables constitute the irreplaceable physical backbone of the global Internet ecosystem and transnational communication networks, carrying the vast majority of international data transmission in the world. With the rapid development of digital economy and cloud computing, the security of such Cyber-Physical Infrastructure is directly related to national security and the lifeblood of the global economy. Once the submarine cable is physically interrupted or sabotaging, it will not only lead to large-scale data flow blocking and communication island effect, but also cause serious secondary information security crisis, such as network route hijacking and sensitive data leakage (Bueger and Liebetrau, 2021). Therefore, the construction of a highly robust submarine cable channel security monitoring and early warning mechanism has become a key topic in the field of modern information security.

At present, most of the conventional protection methods for submarine cable channels are limited to single-dimensional physical space monitoring, and the most typical one is to rely on Automatic Identification System (AIS) for surface trajectory tracking to prevent physical damage such as bottom trawling and illegal anchor dropping. However, this single-source monitoring paradigm exposes fatal security vulnerabilities in complex information adversarial environments. Due to the plaintext broadcast of traditional AIS signals and the lack of underlying encryption and authentication mechanisms, the monitoring system is extremely vulnerable to AIS Spoofing attacks and trajectory data tampering. The inherent vulnerability of AIS data integrity has been systematically documented in maritime cybersecurity literature, where researchers have established comprehensive taxonomies for classifying and mitigating such threats (Goudossis and Katsikas, 2019). Malicious attackers or illegally operating vessels can create the illusion of "safe navigation" on the digital monitoring terminal by transmitting false latitude and longitude coordinates and navigation status, thus concealing their submarine cable sabotage behaviors in the physical world. Relying only on the highly contaminated single-dimensional AIS data can't provide accurate security warning, but will make the entire information security defense system useless. Empirical studies on vessel trajectory anomaly detection have demonstrated that purely data-driven approaches without cross-modal verification are fundamentally limited when facing adversarial data manipulation (Zhen et al., 2017).

In order to fundamentally break the security bottleneck of single-dimensional monitoring, the threat warning paradigm must evolve to the direction of cross-modal and multi-dimensional

data fusion. As shown in Figure 1, the cyber-physical submarine cable monitoring architecture constructed in this paper.

Figure 1 Cyber-physical submarine cable monitoring and cross-domain data fusion architecture



This study proposes a novel multidimensional security early warning system. The system deeply integrates the physical topology benchmark data of submarine cables, real-time AIS surface trajectory features, and underwater acoustic sensor network data (such as heavy object falling and dragging noise detected by hydrophones). This cross-validation mechanism across the three domains of "surface, underwater and digital" can effectively detect single-dimensional data forgery. The integration of heterogeneous maritime sensors for anomaly detection has emerged as a critical research frontier, particularly when individual sensor modalities exhibit complementary strengths and weaknesses under adversarial conditions (Riveiro et al., 2018). However, multi-source heterogeneous sensors will inevitably produce a high degree of data conflict and uncertainty after spatio-temporal alignment—for example, when the received AIS digital coordinates show that a ship is moving away from the submarine cable corridor, but the underwater acoustic sensors accurately extract strong anchor physical features above the submarine cable. To solve this difficult problem, Dempster-Shafer (D-S) evidence theory is introduced in this paper, which provides a rigorous mathematical framework for conflict resolution and uncertainty reasoning in multi-sensor environments. By constructing an adaptive Basic Probability Assignment (BPA) model and conflict normalization rules (Denoex, 2019), the theoretical model can realize accurate traceability and high-confidence warning of potential safety threats in the harsh Marine environment with high noise and deliberate data pollution.

2 Related work

2.1 Security of marine critical infrastructure

As the nerve center of the global digital economy, the security protection of submarine communication cable has long been the focus of academic and industrial attention. Early studies mainly focused on the analysis of the physical layer vulnerability of submarine cables, and explored the mechanical damage caused by bottom trawling, illegal anchorage of ships, and seabed geological disasters. Some scholars have evaluated the probability distribution and

risk level of pure physical layer attacks on submarine cables in different sea areas by constructing physical tension models and historical failure databases. Multi-sensor communication fusion approaches have been proposed for submarine cable fault detection, combining distributed sensing modalities to improve detection reliability beyond single-sensor limitations (Chen et al., 2021). However, with the deep integration of information networks and physical systems, threats to Marine critical infrastructure are no longer limited to a single physical damage. The convergence of cyber and physical threats against critical maritime infrastructure demands resilience frameworks that explicitly account for cross-domain attack vectors and cascading failure modes (Francis and Bekera, 2014). In recent years, research has begun to emphasize the security perspective of "Cyber-Physical" Systems, pointing out that attackers may use vulnerabilities in information networks to cover up or assist their sabotage actions in the physical world. Such cross-domain security threats require defense mechanisms to evolve from traditional pure physical monitoring to a comprehensive security assessment system that includes data verification and information integrity checking.

2.2 Ship trajectory anomaly detection and data tampering threats

In the existing engineering practice of preventing physical damage of submarine cable, trajectory anomaly detection based on AIS is the most widely used technical means. Researchers have extensively adopted deep learning methodologies—including Long Short-Term Memory (LSTM) networks and Generative Adversarial Networks, to mine spatio-temporal patterns on massive AIS trajectories to identify high-risk ships that abnormally slow down or linger above the submarine cable corridor. Density-based clustering combined with recurrent neural networks has proven particularly effective for real-time maritime anomaly detection, offering robust pattern recognition capabilities for vessel behavior analysis (Zhao and Shi, 2019). Although these data-driven methods have made significant progress in the accuracy of trajectory prediction, their underlying assumptions are based on the assumption that the input data is completely real and reliable. The subsequent security research deeply reveals the fatal flaw of this assumption: AIS lacks rigorous encryption and identity authentication mechanism at the beginning of its design, which makes it vulnerable to data tampering and Spoofing attacks. When malicious vessels emit fake latitude and longitude coordinates to disguise their destructive operations in the submarine cable area, anomaly detection models relying only on single-dimensional AIS data will completely fail. Kinematic interpolation methods for AIS trajectory anomaly detection provide an alternative perspective by leveraging the inherent physical constraints of vessel motion to identify data inconsistencies without relying solely on supervised learning (Guo et al., 2021). The literature in this field deeply shows that stacking algorithms solely on AIS single-source data can no longer meet the requirements of high-level information security, and the introduction of multi-modal external data for cross-validation has become an inevitable trend.

2.3 Multi-source sensor data fusion mechanism and application of evidence theory

In order to overcome the limitations and vulnerability of single-source sensors, multi-source data fusion technology has been preliminarily applied in Marine environment perception and safety monitoring. Traditional data fusion methods mostly use Kalman filtering or probabilistic models based on Bayesian inference. While Kalman filtering remains prevalent in marine target tracking applications, its linearity assumptions often struggle with the nonlinear and non-Gaussian characteristics of underwater acoustic propagation environments. These methods perform well when dealing with low-dimensional and linear state estimation, but often struggle

to cope with complex spatio-temporal correlation when facing high-dimensional heterogeneous networks. Especially in complex Marine safety warning scenarios, there is often a high semantic gap and data conflict between digital trajectory signals from surface AIS and acoustically physical signals from hydrophones. In order to solve the problem of high conflict and uncertainty between heterogeneous sensors, D-S evidence theory is introduced into multi-sensor monitoring network. By establishing a BPA framework, the theory can effectively deal with ambiguous or contradictory sensor inputs in the absence of prior probabilities, and has shown excellent fault tolerance performance in scenarios such as coastal surveillance. The foundational mathematical framework of D-S theory has been extensively validated in sensor fusion applications for autonomous systems, where its ability to quantify and manage conflicting evidence provides critical robustness advantages over pure probabilistic approaches (Deng and Deng, 2013). However, the existing Marine applications based on D-S theory mostly stay at the top-level decision-level fusion, and lack the deep integration in the low-level spatio-temporal feature extraction stage.

2.4 Research gaps

Existing research exhibits three critical limitations. First, most works rely on single-dimensional AIS data or isolated physical monitoring, lacking systematic integration of surface trajectories, cable topology, and underwater acoustic sensing. Second, trajectory anomaly detectors are inherently vulnerable to AIS spoofing attacks due to the absence of cross-modal physical verification. Third, prior fusion approaches decouple feature extraction from conflict resolution, preventing the deep integration of spatio-temporal graph neural networks with D-S evidence theory. To bridge these gaps, this study proposes the Adaptive Spatio-temporal Graph Attention Network with Evidence Fusion (ASTGAT-EF), which aligns multi-source features at the representation level and resolves high-conflict evidence at the decision level.

3 Methodology

In order to achieve high robustness monitoring of submarine cable channels and security warning against data tampering, this study designed and constructed an ASTGAT-EF. This model breaks the limitation of traditional dependence on a single data source, and realizes a high degree of cross-modal information integration and conflict resolution from the underlying mathematical architecture. Table 1 systematically combs the key mathematical variables and their physical mapping implications throughout the theoretical derivation of this study.

Table 1 Definition table of core variables and parameters of ASTGAT-EF model

<i>Symbol</i>	<i>Meaning</i>	<i>Symbol</i>	<i>Meaning</i>
G	Topology of the discretization spatio-temporal map of the monitored sea area	α_{ij}	Normalized attention weight between node i and node j
V, E	The set of grid nodes and the set of spatial edges of spatio-temporal graphs	Θ	Identification Framework in D-S Evidence Theory (Security Level)
A	The spatial adjacency matrix between nodes	$m(\cdot)$	BPA for Feature Map Generation
$\mathbf{X}^{(0)}$	The initial multi-source heterogeneous feature input tensor	K	Data conflict coefficient of heterogeneous sensors such as AIS and hydrophones

$\mathbf{H}^{(l)}$	Hidden layer states of the l Graph Convolutional Network (GCN)	τ	The extreme conflict threshold that triggers the adaptive trust discount mechanism
e_{ij}	Original graph attention correlation coefficient between nodes	L_{CE}	Global cross-entropy loss function for early-warning model optimization

3.1 Threat modeling and data alignment

In the complex real ocean environment, the geographical coordinates of physical submarine cables, the navigation trajectories of surface ships, and the monitoring data of seabed acoustic sensors are significantly heterogeneous in spatial distribution and sampling frequency. If these raw data are directly concatenated, the neural network will not be able to capture their inherent physical associations. Therefore, the first step (and necessity) to construct a unified computing space is to discretize the target monitoring sea area into a regular two-dimensional spatial grid system and rigorously model it as a dynamic spatio-temporal graph topology. The mathematical expression of this topology is as follows:

$$G = (V, E, A) \quad (1)$$

where V represents the node set of discrete sea area grid after partition, and each node corresponds to a tiny sea area in a specific latitude and longitude range. E represents the set of spatial connectivity edges formed by the geographical adjacency between grid nodes; Then, A is the adjacency matrix that quantifies the degree of connectivity of these Spaces.

Traditional graph networks usually use a binary matrix composed of 0 and 1 to represent whether nodes are connected or not. However, in the submarine cable monitoring scenario, the physical wake generated by ship sailing and the sound wave generated by anchor show nonlinear continuous attenuation characteristics in the water body. Although the distant grids are not directly adjacent, there is still a weak physical influence. Therefore, a Gaussian kernel based on the actual geographical distance must be introduced to calculate the adjacency matrix:

$$A_{ij} = \exp\left(-\frac{dist(v_i, v_j)^2}{\sigma^2}\right) \quad (2)$$

where A_{ij} represents the spatial correlation weight between node i and node j . $dist(v_i, v_j)$ represents the actual Euclidean distance between two grid center points; σ is the standard deviation of the Gaussian kernel that controls the rate of signal decay. The necessity of this formulation lies in the fact that it truly restores the physical propagation law in the Marine environment, and the weight of the interaction increases exponentially with the closer grids.

After establishing the spatial topology, the problem of temporal dynamic alignment must be solved. The system aggregates multi-dimensional physical quantities (such as speed, heading, drain-depth) from AIS, as well as acoustic spectral features from seabed hydrophones, over the same historical sampling time window. From this, the initial multidimensional feature tensor for the network input is constructed as follows.

$$\mathbf{X}^{(0)} \in \mathbb{R}^{N \times T \times F} \quad (3)$$

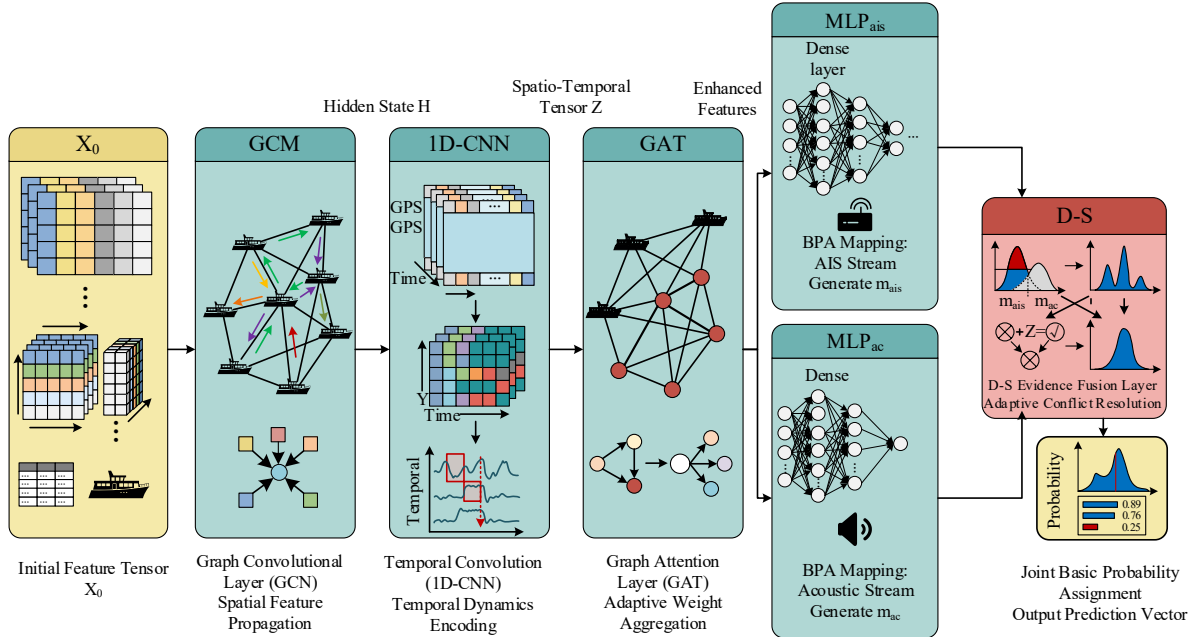
where $\mathbf{X}^{(0)}$ represents the initial state of the zeroth layer of the input neural network; N is the total number of partitioned grid nodes (spatial dimension); T is the time step of historical observation (time dimension); F is the total number of feature dimensions. The function of this tensor formulation is to provide a rigorous and three-dimensional aligned data base for heterogeneous data, so that subsequent deep learning algorithms can simultaneously perform joint calculations across the three dimensions of "time-space-feature".

3.2 Spatio-temporal feature extraction

In the massive trajectories, the vast majority of ships are sailing normally, and only a few trajectories (such as abnormal stagnation and low speed wandering over submarine cables) pose a real threat. In order to accurately identify high-risk behaviors that are potential threats to submarine cables from massive redundant trajectories, the core module of ASTGAT-EF model adopts a composite network architecture that deeply combines spatio-temporal convolution and graph attention mechanism.

As shown in Figure 2, the module first aggregates and reconstructs the features in the spatial dimension using GCN to capture the behavioral spillover effect between adjacent sea area grids.

Figure 2 The underlying architecture of ASTGAT-EF module with tensor flow diagram



First, the model needs to capture behavioral spillovers between adjacent ocean grids (e.g., the dynamics of A ship moving from grid A to grid B). This requires the use of GCN) for feature aggregation in the spatial dimension. The graph convolution propagation evolution rule for layer l is defined as follows:

$$\mathbf{H}^{(l)} = \sigma \left(\mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \mathbf{X}^{(l-1)} \mathbf{W}^{(l)} \right) \quad (4)$$

where $\mathbf{H}^{(l)}$ is the output feature matrix of the current layer; $\mathbf{A} = \mathbf{A} + \mathbf{I}$ indicates that the adjacency matrix introduces a self-loop. $\mathbf{W}^{(l)}$ is the weight matrix and $\sigma(\cdot)$ is the nonlinear activation function.

After spatial aggregation, the one-dimensional convolution operator is used to slide on the time axis to realize the joint encoding of spatio-temporal features:

$$\mathbf{Z}^{(l)} = \mathbf{H}^{(l)} * \mathbf{T} \quad (5)$$

where $\mathbf{Z}^{(l)}$ is the high-order spatio-temporal feature tensor that incorporates the time evolution dynamics, and $\mathbf{T}$ is the convolution kernel of the time dimension.

To prevent dilution of threat signals by uniform aggregation, a graph attention mechanism (Veličković et al., 2018) adaptively weights neighbor contributions:

$$e_{ij} = \text{LeakyReLU}(\mathbf{a}^T [\mathbf{W} \mathbf{h}_i \parallel \mathbf{W} \mathbf{h}_j]) \quad (6)$$

where e_{ij} quantifies the "threat impact degree" of grid j on grid i . The formula combines the hidden state feature vectors $\mathbf{h}_i$ and $\mathbf{h}_j$ through the concatenation operation $\parallel$, and uses the learnable attention weight vector $\mathbf{a}^T$ for inner product calculation. The necessity is that the model is able to learn autonomously from a large amount of historical data.

To make the attention scores comparable across different nodes, a global normalization must be performed using the Softmax function:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in N_i} \exp(e_{ik})} \quad (7)$$

where α_{ij} is the final normalized attention weight; The denominator computes the sum of the index scores of all direct neighbors (set N_i) of node i . Based on this weight, the final feature of the central grid will be updated by the attention weighted sum of all its neighbor features:

$$\mathbf{h}'_i = \sigma \left(\sum_{j \in N_i} \alpha_{ij} \mathbf{W} \mathbf{h}_j \right) \quad (8)$$

This threat-aware attention focuses computational resources on anomalous grids, improving sensitivity to weak threat signals.

3.3 D-S evidence fusion for attack detection

After extracting the high-dimensional features, the system enters the most important information security confrontation phase. When a malicious attacker tampers the AIS data (e.g., transmitting fake coordinates to indicate that the ship has left the submarine cable area), the neural network based on pure AIS data will output the false judgment of "safe", while the hydrophone can capture the real acoustic physical signal of the cable damage. Directly averaging these two predictions results in a failure of early warning. To achieve high confidence decision making (necessity) under information spoofing attacks, D-S evidence theory is introduced in this study to resolve the "extreme data conflict" between heterogeneous sensors (Denoeux, 2019).

First, the identification framework of the system (that is, the set of all possible and mutually exclusive security states of the system) is established:

$$\Theta = \{\text{Safe}, \text{Warning}, \text{Attack}\} \quad (9)$$

The ASTGAT-EF network maps AIS and underwater acoustic features to the probability support of each subset in Θ , namely the BPA, denoted as $m(A)$. The BPA generated by these sensors must strictly obey the probability normalization and empty set extrema constraints:

$$\sum_{A \subseteq \Theta} m(A) = 1 \quad (10)$$

$$m(\emptyset) = 0 \quad (11)$$

The above two formulas ensure that evidence theory has a rigorous foundation of probability theory. The system generates m_{ais} representing digital trace evidence on the surface and m_{ac} representing acoustic physical evidence on the ocean floor.

Next comes the core conflict quantification mechanism. The degree of contradiction between different sensor data is accurately evaluated by calculating the overall conflict coefficient K :

$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \quad (12)$$

The summation condition $B \cap C = \emptyset$ in the formula is extremely clever. It accumulates only those probabilistic judgments that are logically absolutely mutually exclusive. The necessity of this formulation is that it directly constitutes an AIS spoofing attack indicator. If the hacked AIS strongly reports the status as “Safe” and the hydrophone captures the anchor strike noise strongly pointing to “Attack”, the intersection is empty at this time and the K value will climb sharply and approach 1.

When the conflict coefficient K exceeds the preset threshold τ , the standard Dempster combination rule faces the risk of numerical instability. Numerous studies have addressed the conflict management problem in D-S theory, proposing improved combination rules that redistribute conflict mass based on evidence distance and dissimilarity measures to obtain more reasonable fusion results (Deng and Wang, 2020). In the classical D-S theory, when $K \rightarrow 1$, the normalization factor $1/(1-K)$ tends to infinity, which leads to the paradoxical conclusion of the combination result (Yager, 1987; Murphy, 2000). Therefore, this paper proposes an adaptive trust discount mechanism to dynamically reduce the weight of digital evidence in high conflict scenarios while maintaining the probabilistic framework of D-S theory. Contextual discounting strategies within the belief function framework enable refined modeling of sensor reliability, allowing dynamic adjustment of evidence weights based on observed performance characteristics and environmental conditions (Mercier et al., 2008).

Let the original BPA of sensor s ($s \in \{\text{AIS}, \text{Acoustic}\}$) be m_s and its credibility discount factor $\delta_s \in [0, 1]$ be determined by the conflict adaptation function as follows:

$$\delta_s = \begin{cases} 1 & \text{if } K < \tau \\ \exp\left(-\gamma \cdot \frac{K - \tau}{1 - \tau}\right) & \text{if } K \geq \tau \end{cases} \quad (13)$$

where $\gamma > 0$ is the attenuation coefficient ($\gamma = 2.0$ is taken in this paper), and $\tau = 0.85$ is the extremum conflict threshold determined through the validation of the training set. The discounted BPA is defined as follows.

$$m_s^{\text{discounted}}(A) = \delta_s \cdot m_s(A), \quad \forall A \subset \Theta \quad (14)$$

$$m_s^{\text{discounted}}(\Theta) = 1 - \delta_s \cdot \sum_{A \subset \Theta} m_s(A) \quad (15)$$

The discounted mass is reassigned to the frame of discernment as uncertainty, ensuring compatibility with standard D-S combination rules.

The selection of threshold $\tau = 0.85$ is based on 10-fold cross validation on the training set: when τ increases from 0.70 to 0.95, the F1-score of the model on the validation set containing 30% AIS spoofing samples reaches a peak of 0.951 at $\tau = 0.85$, and then decreases due to excessive discounting of normal AIS signals.

After quantifying conflicts and performing adaptive discounting, the heterogeneous evidence is fused using Dempster's orthogonal combination rule:

$$m_{1 \oplus 2}(A) = \frac{1}{1 - K} \sum_{B \cap C = A} m_1(B) m_2(C) \quad (16)$$

where $m_{1 \oplus 2}(A)$ represents the joint probability of the final output of the system for a specific threat state A after fusing multi-source data. The coefficient $\frac{1}{1 - K}$ in front of the formula assumes the critical task of conflict normalization. When the system detects that the forged AIS security coordinates and the anchor noise captured by the hydrophone produce a very high data conflict coefficient ($K \geq \tau$), the adaptive discount algorithm can quickly strip the decision weight of the tampered digital signal, and force the underwater acoustic sensing evidence with

physical non-tampering characteristics to dominate the calculation process of the molecules. Thereby neutralizing data spoofing attacks.

In order to clearly show the complete calculation logic of multi-source heterogeneous features from probability mapping, conflict detection to final security situation assessment, Algorithm 1 describes in detail the forward propagation execution process of the adaptive D-S evidence fusion module. The algorithm not only realizes the orthogonal combination of multi-source data, but also establishes a set of tamper-proof circuit breaker mechanism at the software level through the dynamic conflict threshold τ .

Algorithm 1 Adaptive D-S Evidence Fusion for Threat Assessment

Input: Enhanced spatio-temporal node features: $\mathbf{h}'_i$ (from STGAT); Frame of Discernment: $\Theta = \{\text{Safe, Warning, Attack}\}$; Extreme conflict threshold: τ ; Physical trust discount factor: $\gamma \in (0,1)$.

Output: Final threat prediction probability: P_{threat} ; System alert level: $Level$.

```

1: // Phase 1: Feature to BPA Mapping
2:  $\mathbf{m}_{ais} \leftarrow \text{Softmax}(\text{MLP}_{AIS}(\mathbf{h}'_i))$ 
3:  $\mathbf{m}_{ac} \leftarrow \text{Softmax}(\text{MLP}_{AC}(\mathbf{h}'_i))$ 
4:
5: // Phase 2: Global Conflict Coefficient Calculation
6:  $K \leftarrow \sum_{B \cap C = \emptyset} \mathbf{m}_{ais}(B) \mathbf{m}_{ac}(C), \quad \forall B, C \subseteq \Theta$ 
7:
8: // Phase 3: Adaptive Trust Discounting (Anti-Spoofing Mechanism)
9: if  $K \geq \tau$  then
10:     // Extreme conflict detected (e.g., AIS spoofing)
11:     // Degrade digital AIS evidence, enforce acoustic evidence dominance
12:      $\mathbf{m}_{ais} \leftarrow \gamma \cdot \mathbf{m}_{ais}$ 
13:      $\mathbf{m}_{ac} \leftarrow (1 - \gamma) \cdot \mathbf{m}_{ac}$ 
14: end if
15:
16: // Phase 4: Dempster's Orthogonal Combination
17: for each  $A \subseteq \Theta$  DO
18:      $\mathbf{m}_{fused}(A) \leftarrow \frac{1}{1 - K} \sum_{B \cap C = A} \mathbf{m}_{ais}(B) \mathbf{m}_{ac}(C)$ 
19: end for
20:
21: // Phase 5: Decision Output
22:  $P_{threat} \leftarrow \mathbf{m}_{fused}(\text{Attack})$ 
23:  $Level \leftarrow \arg \max_{A \in \Theta} \mathbf{m}_{fused}(A)$ 
24: return  $P_{threat}, Level$ 

```

In the backpropagation stage of model optimization, the target loss function needs to be constructed to guide the network to update parameters. To guide model optimization, the discrepancy between the fused output probability and the actual security label is measured via a global cross-entropy loss formulation:

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (17)$$

where y_i is the true security label (absolute 0 or 1), $\hat{y}_i$ is the predicted probability value of the model fusion output. When the predicted value deviates from the true value, the logarithmic term causes the penalty to increase exponentially, thus forcing the network weights to adjust in the direction of correctly identifying the threat. For the purpose of overfitting suppression, an L2 regularization component is integrated into the aggregated loss function.

4 Experimental setup

4.1 Dataset and preprocessing

In order to fully verify the effectiveness and robustness of the proposed ASTGAT-EF model in the complex Marine environment, the experimental data abandoned the idealized simulation generation, and instead constructed a multi-dimensional cross-domain data set completely based on the real physical world.

- 1 Study area and data time range. The spatial intersection area of this dataset is selected as the sea area near Vancouver Island on the west coast of North America (geographic boundary: 48.0°N - 49.5°N , 123.0°W - 125.5°W). This area is not only the core landing hub of multiple trans-oceanic international submarine cables, but also the overlap area of high-density international shipping and bottom fishery activities. The data collection span from January 1, 2022 to December 31, 2022 (UTC time zone), covering the complete seasonal shipping cycle.
- 2 Physical topology data of submarine cables. Submarine cable physical topology datum extracted from open source TeleGeography authoritative database (<https://www.submarinecablemap.com/>), obtained four active undersea cables within the area of precise GeoJSON geographic coordinates and buried depth of information, including: Pacific Crossing 1 (PC-1), Tata TGN-Pacific, Canada-US, and Southern Cross Cable. The data access date is March 15, 2023.
- 3 AIS trajectory data. Extracted from the National Oceanic and Atmospheric Administration (NOAA) Marine Cadastre database (<https://marinecadastre.gov/ais/>), covering the area for up to 12 months in the full amount of history of ship AIS broadcast data. The dataset contains 12,847,293 high-frequency AIS records with a median sampling interval of 10 seconds, covering the vessel unique identification number, dynamic latitude and longitude, speed (SOG), course (COG), and draught state. Data acquisition followed the NOAA Open Data protocol without additional ethical approval, because AIS signals are publicly available VHF radio information that ships are forced to broadcast and do not contain personally identifiable information.
- 4 Underwater acoustic data were obtained through the Oceans 3.0 open API of Ocean Networks Canada (ONC, <https://data.oceannetworks.ca/>). The dataset comprises broadband audio recordings from an ICListen HF hydrophone (256 kHz sampling rate) and bottom current meter time series at the Barkley Canyon hydrophone node (BACEX.A1) of the NEPTUNE cabled seafloor observatory, deployed at the target sea area. Data were downloaded in NetCDF format via the Oceans 3.0 API between April 1 and May 30, 2023, under the Creative Commons Attribution 4.0 International License (CC BY 4.0).
- 5 Data alignment and preprocessing. To integrate the three heterogeneous data sources with vastly different sampling frequencies and spatial distributions into a unified computing framework, a strict grid-based spatio-temporal alignment algorithm was applied. The target sea area was discretized into a 0.5×0.5 nautical mile spatial grid (1,024 nodes), with time steps uniformly resampled into a 10-minute sliding window.

On this unified spatio-temporal basis, AIS trajectory points were mapped to corresponding grids with aggregated features (mean speed, course variance, and maximum draught). Concurrently, spectral energy density features (0.5–8 kHz, 64 frequency bands) extracted from the continuous broadband acoustic signals via Short-Time Fourier Transform (2,048-sample window, 50% overlap) were synchronously aligned to the same spatio-temporal slices.

Figures 3-5 visualize normal navigation density, high-risk anchoring hotspots, and ASTGAT-EF predicted risk scores, respectively. The extraction of spectral energy features from underwater acoustic recordings via short-time Fourier transform aligns with established practices in passive acoustic monitoring, where broadband hydrophone arrays provide continuous surveillance capabilities for seabed infrastructure protection (Mercier, et al., 2008). Figure 3 shows ship trajectories distributed along main channels and away from cable corridors. Figure 4 reveals high-risk anchoring hotspots precisely above the cable corridor, forming spatial separation from normal channels. Figure 5 demonstrates the model's accurate threat positioning and effective suppression of normal sea areas, validating the joint modeling approach.

Figure 3 Heat map of normal navigation density distribution

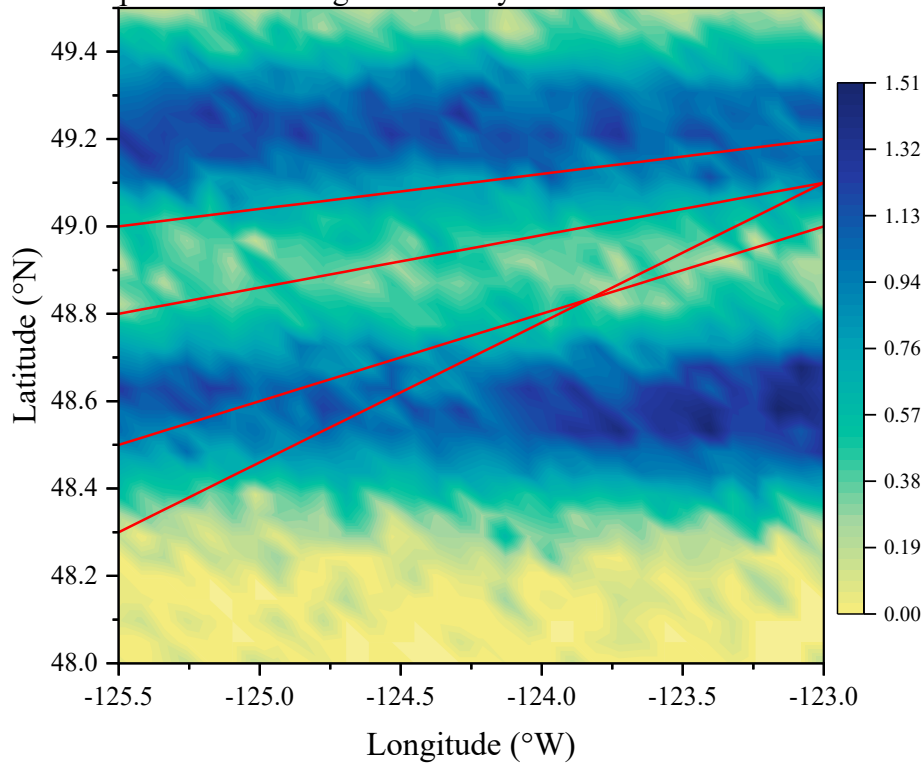


Figure 4 Spatial distribution of high-risk anchoring hotspots

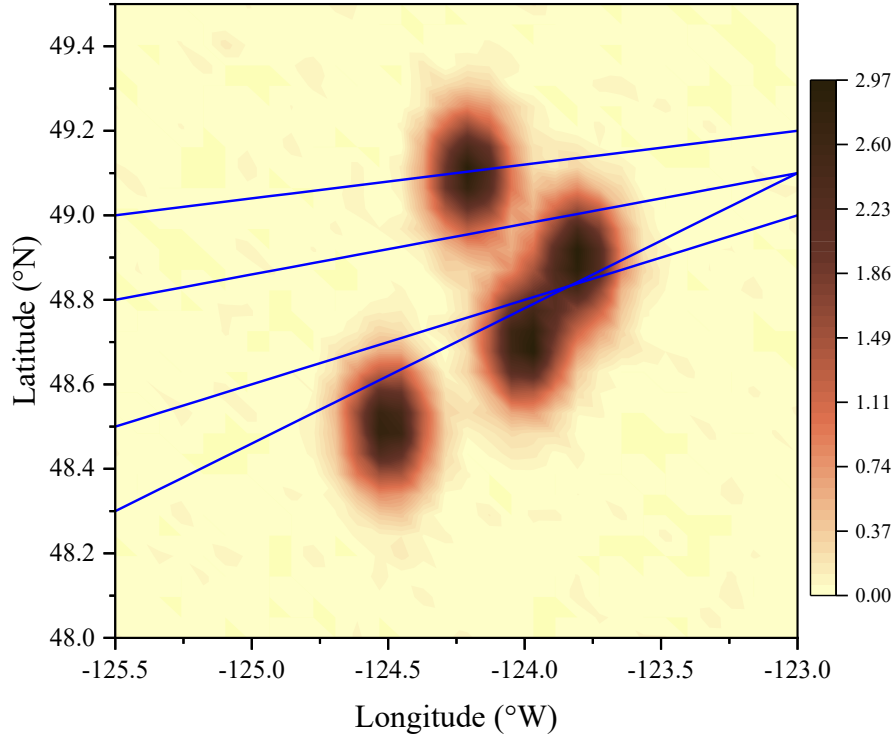
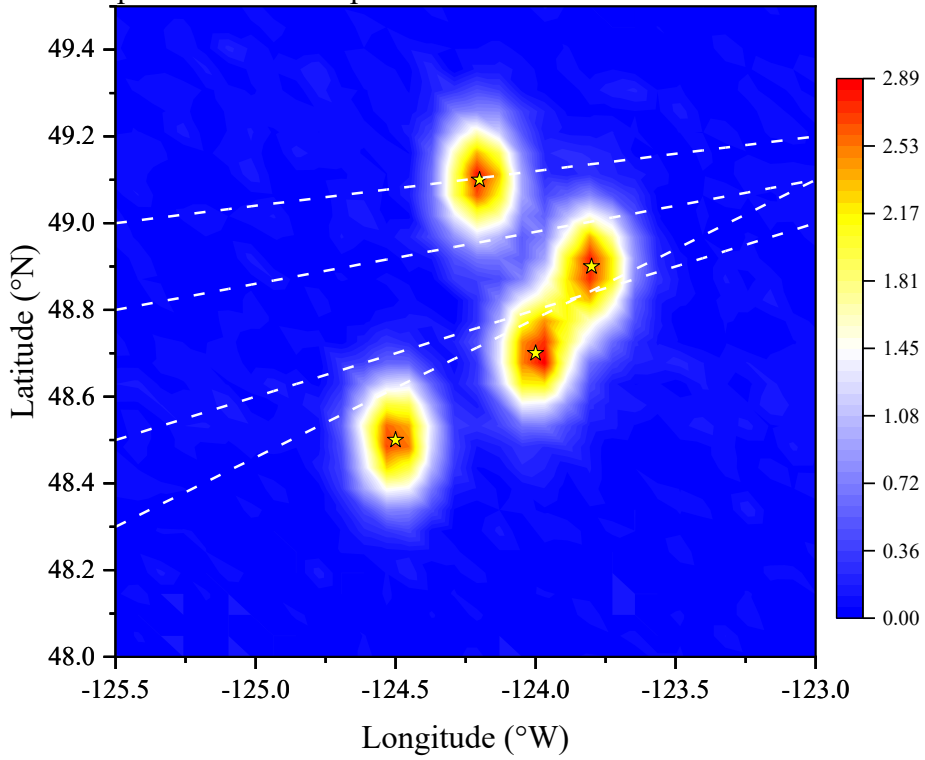


Figure 5 Heat map of ASTGAT-EF prediction risk



The label generation is performed in a semi-supervised manner: firstly, positive sample anchors are established based on historical submarine cable fault records, and then automatically labeled by domain expert rules. Finally, two experts with more than 5 years of maritime safety experience are used for manual review. The dataset contains $1,024 \text{ nodes} \times 52,560 \text{ time steps}$ (granularity of 10 minutes), and the proportion of threat samples is about 3.2%, showing a typical extremely unbalanced distribution of categories.

4.2 Baselines and metrics

After completing the alignment and labeling of heterogeneous datasets (semi-supervised labeling is performed according to historical cable fault records and abnormal behavior rules), in order to objectively evaluate the advancement of ASTGAT-EF model, four representative state-of-the-art Advanced Trajectory Anomaly detection and data fusion methods are carefully selected as the comparison baseline. The first is a trajectory anomaly detection model based on LSTM-Traj, which represents the traditional pure temporal deep learning method and only relies on single-dimensional AIS data for prediction. Trajectory prediction models based on bidirectional LSTM architectures have shown particular promise for maritime applications, though their reliance on clean AIS data remains a fundamental vulnerability in adversarial scenarios (Park et al., 2021). The second model is a single-modal Spatio-Temporal Transformer model (Transformer-S), which is used to verify the performance of pure self-attention mechanism in the absence of topological graph prior structure. The third is a multi-sensor Fusion scheme based on traditional Kalman filter (Kalman-Fusion), which is used to compare the ability of traditional mathematical statistics method and deep learning in dealing with nonlinear data conflict. The fourth one is the basic Spatio-Temporal Graph Convolutional Network (STGCN-Basic) stripped of the fusion layer of D-S evidence theory, and this baseline is used as the core reference for ablation experiments. Specifically used to quantify the performance gain brought by the D-S fusion layer in resisting information tampering attacks (Liang et al., 2024).

In order to comprehensively and multi-dimensionally measure the classification efficiency and ranking ability of the model on the extremely imbalanced class dataset (far more safe samples than corrupt samples), the experimental system integrates six stringent evaluation indicators: Area under the receiver operating characteristic curve (AUC), global Accuracy (ACC), Precision, Recall, comprehensive evaluation index F1-score and Normalized Discounted Cumulative Gain (NDCG@5). Among them, the introduction of NDCG@5 is particularly critical, which can accurately evaluate whether the most real and deadly physical threats can be ranked at the front of the system alert queue in the face of massive concurrent signals, which is of decisive significance for the response priority division in practical engineering deployment.

4.3 Robustness verification

To eliminate random fluctuations caused by initialization and data partitioning, all models are trained and tested 10 times with different random seeds on the same computing cluster. Paired t-tests are conducted to verify statistical significance ($p < 0.05$), and effect sizes are quantified via Cohen's d . To evaluate robustness against adversarial attacks, simulated AIS spoofing signals are injected into the test set at proportions ranging from 0% to 40%.

5 Results and analysis

5.1 Performance comparison

To comprehensively evaluate the practical effectiveness of the proposed ASTGAT-EF model in submarine cable channel monitoring, rigorous tests were conducted on a constructed real multi-source dataset of the West Coast of North America. Table 2 presents a detailed comparison of the core classification and ranking metrics of this model with four mainstream baseline methods on the test set.

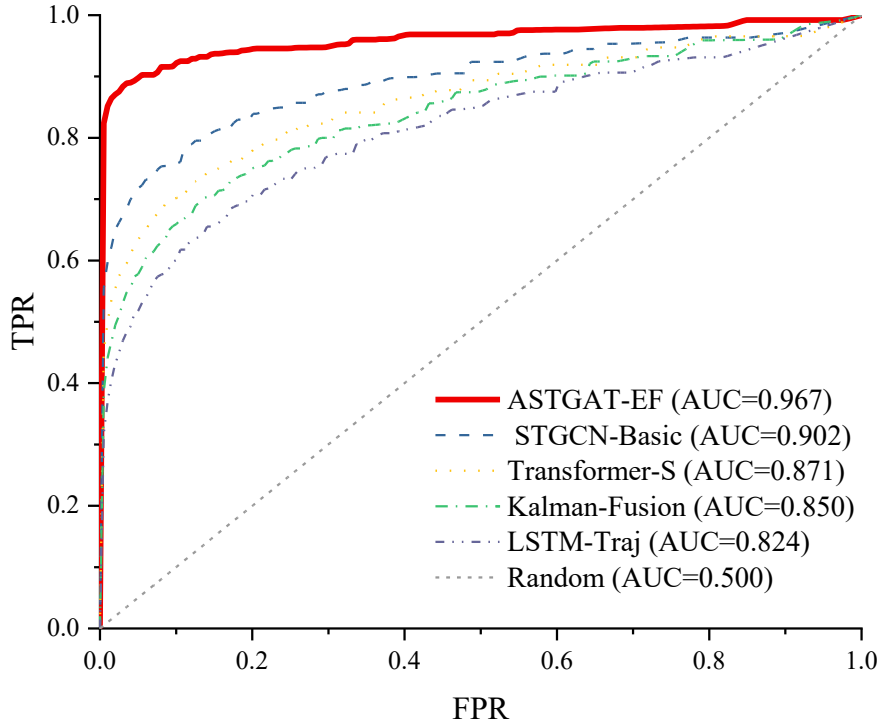
Table 2 Performance Comparison of ASTGAT-EF Model and Baseline Method on the Test Set

<i>Model method</i>	<i>ACC</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>AUC</i>	<i>NDCG@5</i>
LSTM-Traj	0.812	0.805	0.793	0.798	0.824	0.771
Transformer-S	0.865	0.852	0.841	0.846	0.871	0.820
Kalman-Fusion	0.843	0.838	0.829	0.833	0.850	0.805
STGCN-Basic	0.891	0.887	0.880	0.883	0.902	0.855
ASTGAT-EF	0.954	0.948	0.951	0.949	0.967	0.932

5.2 Resilience under AIS spoofing

In the security assessment of cyber-physical systems, the robustness of the model in the face of malicious data contamination is even more important than pure classification accuracy. In order to verify the core value of D-S evidence theory fusion layer in resisting network attacks, simulated AIS spoofing signals were injected at varying proportions. Figure 6 further presents the ROC curves of ASTGAT-EF and all baseline methods. The proposed model maintains superior discriminative capability across all classification thresholds, which is consistent with the high AUC value reported in Table 2.

Figure 6 Comparison plot of ROC curves between ASTGAT-EF and baseline methods



However, when 30% AIS spoofing samples are injected into the dataset, the single-source model lacking cross-modal true-checking mechanism shows cliff-like performance decay. To systematically quantify this attenuation difference, Figure 7 presents the comprehensive performance evaluation results for different AIS spoofing ratios.

Figure 7 Comprehensive performance evaluation under different threats and noise levels:(a) F1-score decay curve under AIS spoofing attack; (b) Precision/Recall/F1 comparison under 30% AIS spoofing

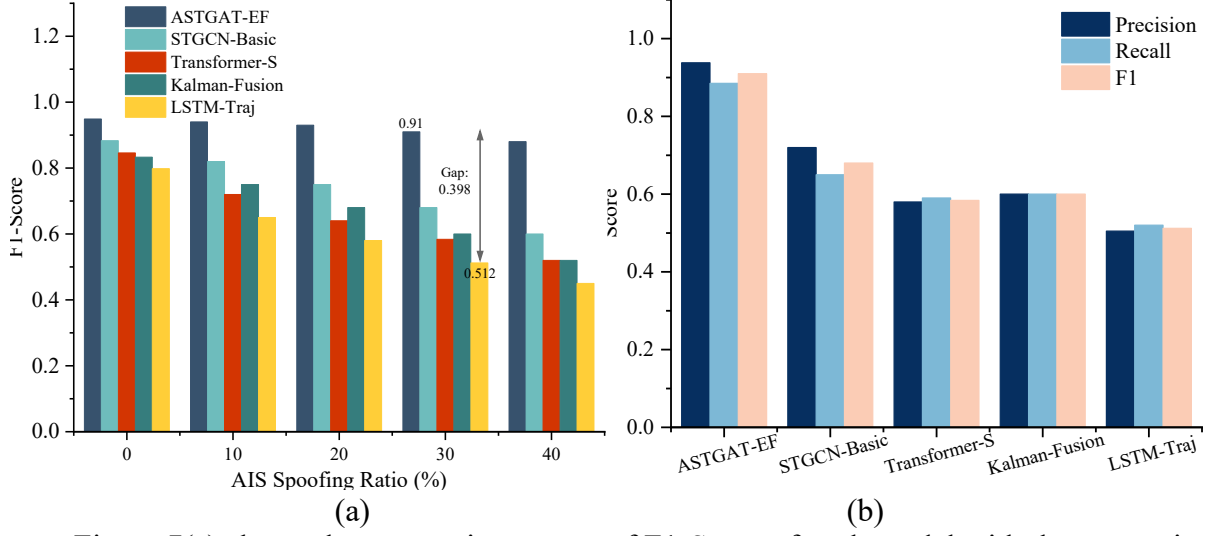


Figure 7(a) shows the attenuation curves of F1-Score of each model with the proportion of deception (0%-40%). At 30% spoofing, ASTGAT-EF maintains F1=0.910 ($\Delta < 4\%$), while single-source models collapse below 0.6, demonstrating the D-S fusion layer's anti-spoofing value. The robustness of spatiotemporal graph neural networks against data contamination has been systematically evaluated in industrial anomaly detection contexts, where attention-weighted architectures demonstrate consistent performance maintenance under significant noise injection (Tan et al., 2025).

Figure 7(b) further compares the performance of each model under the 30% deception condition from three dimensions of Precision, Recall and F1. Under 30% spoofing, ASTGAT-EF achieves Precision=0.938, Recall=0.885, and F1=0.910, significantly outperforming all baselines. This robustness stems from the D-S conflict normalization mechanism, which discounts tampered AIS evidence when it conflicts with hydrophone data ($k \rightarrow 1$).

5.3 Ablation study

Figure 8 shows the boxplots of the distribution of the four indicators ACC, Precision, Recall and F1-score between ASTGAT-EF and STGCN-Basic in 10 repeated experiments.

Figure 8 Boxplot of statistical significance of ablation experiments

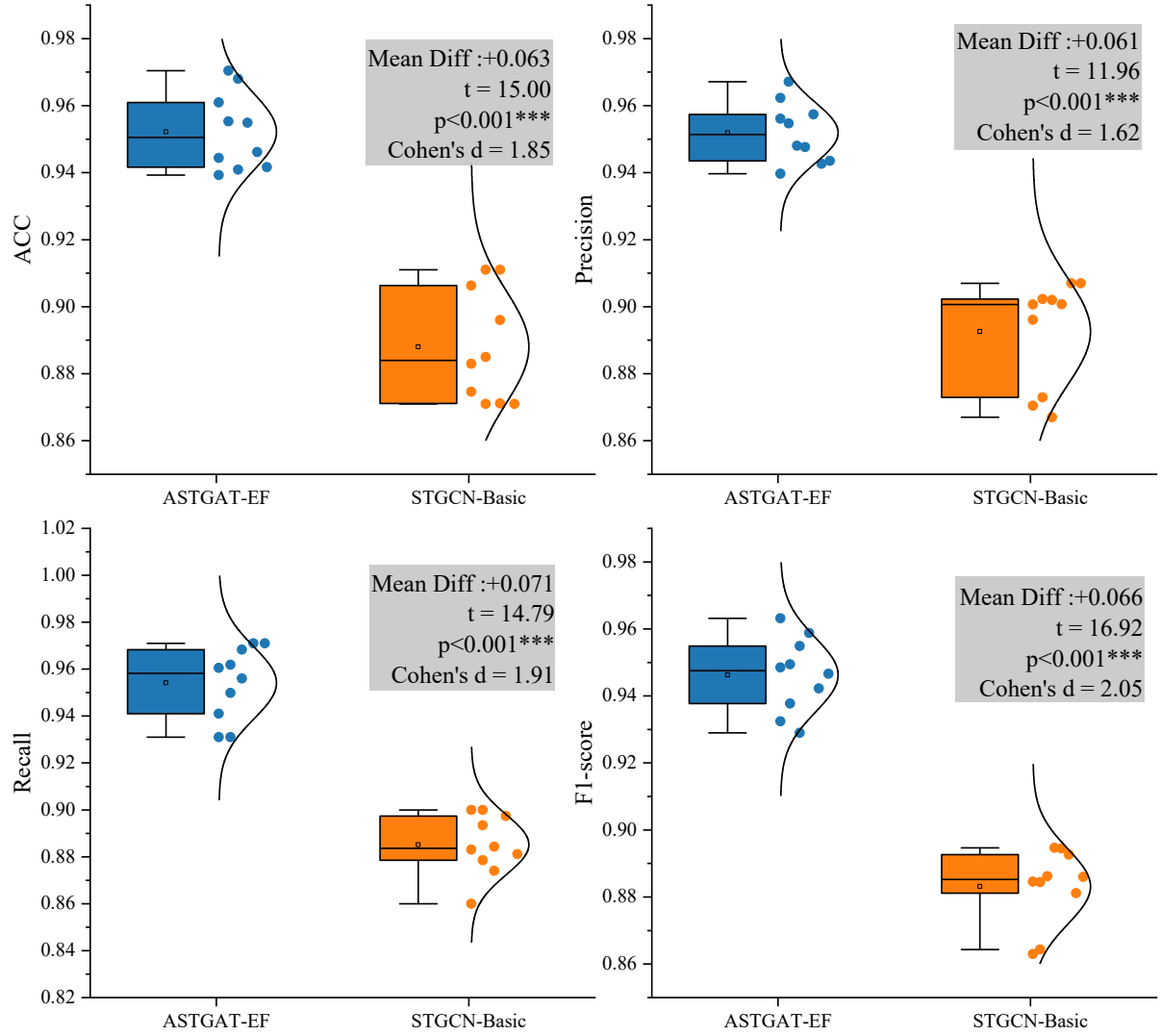


Figure 8 shows ASTGAT-EF achieves higher median and narrower interquartile range than STGCN-Basic across all metrics, confirming superior stability and statistical significance. tatistical validation frameworks incorporating effect size measures and repeated trials have become essential standards for evaluating maritime anomaly detection systems, ensuring that reported performance gains reflect genuine methodological improvements rather than experimental artifacts (Li et al., 2017).

Extended tests against LSTM-Traj, Transformer-S, and Kalman-Fusion confirm consistent significant improvements. To further quantify the statistical significance and practical effect sizes of these improvements, Table 3 presents the detailed paired t-test results between ASTGAT-EF and each baseline model.

Table 3 Statistical analysis results of paired t-test between ASTGAT-EF and each baseline model

Comparison model	Evaluation metrics	Difference of means	Standard Error	t value	Cohen's d
STGCN-Basic	ACC	0.063	0.0042	15	1.85
	Precision	0.061	0.0051	11.96	1.62
	Recall	0.071	0.0048	14.79	1.91
	F1-score	0.066	0.0039	16.92	2.05
LSTM-Traj	ACC	0.142	0.0063	22.54	3.12
	Precision	0.143	0.0072	19.86	2.89

Transformer-S	Recall	0.158	0.0059	26.78	3.45
	F1-score	0.151	0.0061	24.75	3.28
	ACC	0.089	0.0054	16.48	2.01
	Precision	0.096	0.0063	15.24	1.95
	Recall	0.11	0.0051	21.57	2.67
Kalman-Fusion	F1-score	0.103	0.0048	21.46	2.64
	ACC	0.111	0.0058	19.14	2.34
	Precision	0.11	0.0067	16.42	2.08
	Recall	0.122	0.0054	22.59	2.78
	F1-score	0.116	0.0052	22.31	2.75

6 Discussion and conclusion

In this paper, the D-S evidence theory is deeply embedded into the underlying spatio-temporal feature space, and the immutability of the physical world (such as the acoustic physical characteristics towed by heavy objects on the sea floor) is used as the truth check anchor in the digital world (such as the AIS broadcast data that can be easily tampered by hackers), which completely changes the passive situation of traditional monitoring nodes in the face of trajectory spoofing attacks.

Although this study shows excellent performance in resisting single-dimensional information deception, there is still room for exploration in the face of advanced security threats in the future. If an adversary uses adversarial example attack to pollute AIS data while simultaneously using the Unmanned Underwater Vehicle to release fake environmental floor noise to cover up the real damage voiprint, the current evidence trust model may face the risk of misleading the underlying features at the same time. Future research should focus on improving the defense resilience of graph neural networks under multimodal adversarial sample attacks, exploring the introduction of robust adversarial training mechanism into the early stage of spatio-temporal feature extraction, and constructing a nonlinear dynamic decay function based on historical trust evaluation in the D-S evidence theory framework.

In summary, this study proposes a security warning model based on ASTGAT-EF to solve the single-dimensional monitoring vulnerability and data tampering threat of submarine communication cable in complex cyber-physical environment. The model successfully embedded D-S evidence theory deeply into the underlying spatio-temporal feature space of graph neural network, and constructed a multi-source conflict resolution mechanism spanning physical topology reference, surface digital trajectory and underwater acoustic sensing from the mathematical mechanism. Rigorous verification and statistical analysis based on real multi-dimensional ocean data sets show that the fusion model has achieved breakthrough results in the performance of early warning classification. The more critical research finding is that, in the extreme information security confrontation scenario where up to 30% trajectory spoofing signals are injected, the system relies on the forced cross-checking truth of acoustic physical evidence to successfully control the attenuation rate of its warning effectiveness within 4%, which completely overcomes the fatal defect of the traditional single-source trajectory monitoring system that is easy to collapse when faced with digital signal pollution. This defense architecture combines dynamic high-order spatio-temporal map feature extraction and multi-modal evidence conflict normalization, providing a robust defense paradigm for marine critical infrastructure protection.

Funding Declaration

This work is supported by the Science and Technology Project of China Huadian Corporation LTD. “Research and Application of Wind Farm Regional Monitoring and Early Warning Technology Based on Multi-source Information Fusion” (No. CHDKJ23-02-248).

Declarations

Consent to Participate declaration: not applicable.

Consent to Publish declaration: not applicable.

Ethics declaration: not applicable.

Clinical Trial Registration: not applicable.

All authors declare that they have no conflicts of interest.

References

- Bueger, C. and Liebetrau, T. (2021) ‘Protecting hidden infrastructure: The security politics of the global submarine data cable network’, *Contemporary Security Policy*, Vol. 42, No. 3, pp. 391-413.
- Chen, Y., Li, X., Cai, C., Wu, C., Zhang, W., Huang, X., Guo, Q. and Jiang, D. (2021) ‘Submarine cable detection method based on multisensor communication’, *Journal of Sensors*, Vol. 2021, No. 1, p. 1176347.
- Deng, X. and Deng, Y. (2013) ‘Multisensor information fusion based on dempster-shafer theory and power average operator’, *Journal of Computational Information Systems*, Vol. 9, No. 16, pp. 6417-6424.
- Deng, Z. and Wang, J. (2020) ‘A novel evidence conflict measurement for multi-sensor data fusion based on the evidence distance and evidence angle’, *Sensors*, Vol. 20, No. 2, p. 381.
- Denœux, T. (2019) ‘Decision-making with belief functions: A review’, *International Journal of Approximate Reasoning*, Vol. 109, pp. 87-110.
- Francis, R. and Bekera, B. (2014) ‘A metric and frameworks for resilience analysis of engineered and infrastructure systems’, *Reliability Engineering & System Safety*, Vol. 121, pp. 90-103.
- Goudossis, A. and Katsikas, S.K. (2019) ‘Towards a secure automatic identification system (AIS)’, *Journal of Marine Science and Technology*, Vol. 24, No. 2, pp. 410-423.
- Guo, S., Mou, J., Chen, L. and Chen, P. (2021) ‘An anomaly detection method for AIS trajectory based on kinematic interpolation’, *Journal of Marine Science and Engineering*, Vol. 9, No. 6, p. 609.
- Li, H., Liu, J., Liu, R.W., Xiong, N., Wu, K. and Kim, T.-h. (2017) ‘A dimensionality reduction-based multi-step clustering method for robust vessel trajectory analysis’, *Sensors*, Vol. 17, No. 8, p. 1792.
- Liang, M., Weng, L., Gao, R., Li, Y. and Du, L. (2024) ‘Unsupervised maritime anomaly detection for intelligent situational awareness using AIS data’, *Knowledge-Based Systems*, Vol. 284, p. 111313.
- Mercier, D., Quost, B. and Denœux, T. (2008) ‘Refined modeling of sensor reliability in the belief function framework using contextual discounting’, *Information Fusion*, Vol. 9, No. 2, pp. 246-258.
- Murphy, C.K. (2000) ‘Combining belief functions when evidence conflicts’, *Decision Support Systems*, Vol. 29, No. 1, pp. 1-9.
- Park, J., Jeong, J. and Park, Y. (2021) ‘Ship trajectory prediction based on bi-LSTM using spectral-clustered AIS data’, *Journal of Marine Science and Engineering*, Vol. 9, No. 9, p. 1037.

- Riveiro, M., Pallotta, G. and Vespe, M. (2018) 'Maritime anomaly detection: A review', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 8, No. 5, p. e1266.
- Tan, C., Hu, J., Li, J., Miao, M., Hu, W. and Wang, S. (2025) 'Dts-mixnet: Dynamic spatiotemporal graph mixed network for anomaly detection in multivariate time series', *Big Data and Cognitive Computing*, Vol. 9, No. 10, p. 245.
- Yager, R.R. (1987) 'On the Dempster-Shafer framework and new combination rules', *Information Sciences*, Vol. 41, No. 2, pp. 93-137.
- Zhao, L. and Shi, G. (2019) 'Maritime anomaly detection using density-based clustering and recurrent neural network', *The Journal of Navigation*, Vol. 72, No. 4, pp. 894-916.
- Zhen, R., Jin, Y., Hu, Q., Shao, Z. and Nikitakos, N. (2017) 'Maritime anomaly detection within coastal waters based on vessel trajectory clustering and Naïve Bayes Classifier', *The Journal of Navigation*, Vol. 70, No. 3, pp. 648-670.