WAKE: Creating a Trustless, Verified AIS Data Layer for Global Maritime Applications

Dr. Owen Taylor
WorldWide AIS Network

April 5, 2025

Abstract

The AIS (Automatic Identification System) serves as a vital tool for tracking maritime traffic, but the centralized nature of AIS data aggregation introduces significant risks of data manipulation and geographic gaps. WAKE (Worldwide AIS Network) provides an innovative solution through a decentralized, Web 3.0-based protocol that ensures data accuracy, transparency, and scalability. By utilizing blockchain and consensus algorithms, WAKE offers a tamper-proof and reliable maritime intelligence framework, rewarding contributors with WAKE tokens for submitting verified data. This transformative approach not only enhances the integrity of global AIS data but also drives efficiencies in maritime logistics, insurance, and regulatory compliance, empowering a new era of secure and transparent maritime operations.

Keywords: AIS, Blockchain, Maritime, Web3.0, Decentralization, Data Integrity

Contents

1 Executive Summary **3**

2 The Problem **3**

3 The WAKE Solution **3**

4 Network Architecture **4**

5 Token Utility and Economics **5**

6 Token Emissions & Supply Model **5**
- 6.1 Token Vesting & Unlock Mechanics . . . 6
- 6.2 Emission Curve & Node Incentives . . . 6
- 6.3 Token-Based Data Access Model . . . 6

7 Node Participation & Contributor Mechanics **7**
- 7.1 Hardware Requirements . . . 7

8 Reward Allocation and Emission Model **7**
- 8.1 Reward Modifier Function . . . 8
- 8.2 Governance of Reward Parameters . . . 8
- 8.3 Coverage Optimization in Practice . . . 8
- 8.4 Aligning Incentives with Integrity . . . 8

9 Fraud Detection and Data Integrity **9**
- 9.1 Layered Validation Framework: System Overview . . . 9
- 9.2 Geographic Validation Groups . . . 9

10 Peer Agreement and Adaptive Confidence Scoring **10**
- 10.1 Collusion Detection via Message Similarity . . . 11
- 10.2 Message Similarity Measurement . . . 11
- 10.3 Collusion Risk Score . . . 11
- 10.4 Collusion Response & Penalties . . . 12

11 Physics-Based Validation: Time-of-Arrival Checks **12**
- 11.1 Expected Propagation Delay . . . 12
- 11.2 Observed Time Difference . . . 13
- 11.3 Signal Timing Error . . . 13
- 11.4 Threshold and Flagging . . . 13

12 Node Geolocation and Motion Consistency **13**

13 Motion Modeling for All Nodes **14**
- 13.1 Correlation with AIS Beacon Data (When Available) . . . 14

14 Fraud Scoring and Economic Slashing **14**
- 14.1 Fraud Probability Score . . . 15
- 14.2 Severity-Weighted Slashing . . . 15

15 Early Deployment Logic and Retroactive Validation **16**

16 Less Trust, More Truth **16**

17 Integrity by Design: Attack Resistance and Limitations **16**

18 Data Storage & Integrity Architecture **17**

19 Data Access & Commercial Integration **17**

20 Conclusion **18**

1 Executive Summary

AIS (Automatic Identification System) data is the foundation of global maritime visibility — essential for everything from port logistics and insurance underwriting to real-time cargo tracking, environmental compliance, and financial trade analysis. As over 90% of global trade moves by sea, understanding ship movements is critical for economies, supply chains, and national security.

Today, AIS data is aggregated and sold by centralized platforms, many of which rely heavily on volunteers and hobbyists to provide terrestrial coverage. These contributors receive no financial reward, and are typically only offered access to premium tools in exchange for meeting uptime targets. Meanwhile, the platforms they support monetize the data — often without ensuring its authenticity. The system is vulnerable to spoofed signals, tampered data, and patchy coverage, particularly in underdeveloped regions.

WAKE (Worldwide AIS Network) is a decentralized data layer that fixes this broken model. It rewards contributors with WAKE tokens for submitting valid AIS data, using on-chain consensus to verify accuracy, freshness, and geographic coverage. The result is a trustless, tamper-proof, system for global maritime intelligence.

2 The Problem

AIS (Automatic Identification System) data underpins the modern maritime domain. It enables vessel tracking, supports search and rescue operations, ensures navigational safety, and is used across industries for logistics, compliance, security, and finance. However, the current system for collecting and distributing AIS data is deeply flawed.

The AIS data ecosystem is controlled by a small number of centralized platforms, which aggregate data from commercial satellites, shore-based receivers, and a global network of volunteers. A significant portion of terrestrial AIS data is collected by hobbyists who operate receiver stations at their own expense. These contributors receive no financial compensation, and in many cases are only rewarded with access to premium platform tools if they meet strict uptime requirements.

Despite monetizing this data, centralized aggregators offer no way to validate whether the data is authentic or tampered with. AIS signals can be easily spoofed or manipulated — leading to the emergence of "ghost ships," illegal fishing vessels, and military decoys that transmit false locations. This undermines trust in the data and can have serious implications for national security, regulatory enforcement, and commercial decision-making.

Furthermore, coverage gaps persist, particularly in:

- Remote regions far from major shipping lanes
- Developing nations with little infrastructure
- High-latitude zones where satellite congestion or blind spots reduce signal reliability

3 The WAKE Solution

WAKE (Worldwide AIS Network) is a decentralized protocol designed to reshape the global AIS data economy by rewarding contributors, ensuring data integrity, and expanding maritime visibility through trustless infrastructure.

The network is built around three core innovations:

1. **Decentralized Data Collection**
   WAKE allows anyone to operate an AIS node — whether land-based, shipborne, buoy-mounted, or satellite-linked. Each node is cryptographically registered and tied to a unique on-chain identity. This open-access model dramatically increases coverage, especially in underserved regions where traditional AIS infrastructure is limited or absent.
2. **On-Chain Validation & Fraud Resistance**
   Every AIS transmission submitted to WAKE is validated through a decentralized consensus mechanism. Nodes cross-verify timestamps, geolocations, and signal patterns to detect anomalies or spoofed

data. In areas with dense coverage, a majority consensus approach is used. In low-density zones, node trust scores and historical accuracy factor into validation. To prevent fraud, all rewards are time-locked for a validation period. Nodes submitting inconsistent or manipulated data face penalties, including reduced reputation, reward slashing, or permanent exclusion from the network.

3. **Incentivized Network Growth**
   Once validated, contributors receive WAKE tokens as compensation. These tokens are dynamically allocated based on:

   - Geographic scarcity (higher rewards in low-coverage zones)
   - Data freshness
   - Node reliability

Third-party data customers access validated AIS datasets using WAKE tokens. These tokens serve as the required medium for interacting with the network and unlocking access to specific data products, including real-time streams, historical records, or filtered intelligence feeds.
WAKE tokens are acquired on the open market and used to engage directly with the decentralized data layer. This ensures that data access is transparently tied to protocol usage, creating a system where utility drives token demand.
By replacing centralized aggregation with cryptographic validation and decentralized contribution, WAKE establishes a global AIS data infrastructure that is verifiable, tamper-resistant, and accessible through open protocol interaction.

4 Network Architecture

The WAKE Protocol is underpinned by a globally distributed mesh of AIS receiver nodes, each capable of independently capturing vessel broadcasts and securely submitting them for validation. These nodes can be deployed across diverse environments — including coastal shorelines, onboard commercial vessels, autonomous buoys, aircraft, and satellite platforms — allowing for continuous, decentralized maritime coverage across both high-traffic and remote regions of the ocean.
Each participating node is uniquely identified through a cryptographically registered on-chain address, enabling secure interaction with the protocol while ensuring accountability. Upon receiving AIS signals from nearby vessels [2], the node timestamps and formats the data according to protocol standards. This raw AIS message is then transmitted to the WAKE network, where it enters a validation pipeline.
The WAKE validation process is built around a multi-layered, trustless consensus mechanism. In regions where AIS signals are received by several nodes concurrently, a majority consensus model is employed: independent receivers corroborate the location, time, and vessel identity information, ensuring that no single node can manipulate or fabricate data without detection. These submissions are cryptographically signed, and consensus is reached once a predetermined threshold of independently verified matches is achieved.
In contrast, in low-density or remote maritime regions where node coverage is sparse, the protocol adapts to a weighted validation approach. Here, node reputation plays a critical role. Historical accuracy, consistency of signal capture, and geographic uniqueness contribute to each node's trust score. Signals originating from isolated regions undergo time-based plausibility checks, and are cross-referenced against network baselines to ensure authenticity even in the absence of high node redundancy.
To protect the network against manipulation, each data submission is subject to a mandatory holding period before rewards are disbursed. During this time, additional network analysis occurs to detect inconsistencies, duplicated transmissions, or suspicious behavioral patterns. If anomalies are discovered, the submitting node may be penalized through partial slashing, withheld rewards, or in severe cases, permanent exclusion from the network.
At the core of the WAKE architecture is a hybrid storage model designed to balance scalability with verifiability. All data submitted to the network is divided into two components: on-chain metadata and off-chain raw AIS transmissions. The blockchain ledger maintains a tamper-proof record of validation hashes, timestamps, node identifiers, and verification outcomes. Meanwhile, the AIS messages themselves are stored off-chain in a scalable infrastructure optimized for maritime tracking applications. Each raw data file is cryptographically

linked to its corresponding on-chain metadata, ensuring that any future retrieval or audit can confirm the integrity of the original submission.
This architectural framework enables WAKE to function as a high-integrity, low-trust, and globally accessible AIS data network. By decentralizing both data collection and validation, and by implementing adaptive consensus models for both dense and sparse maritime zones, the WAKE Protocol offers a transformative upgrade over today's centralized and unverifiable AIS infrastructure. It creates a foundation not just for fair contributor compensation, but for a new paradigm of maritime situational awareness — one that is transparent, secure, and globally distributed by design.

5 Token Utility and Economics

The WAKE token serves as the native utility and incentive mechanism of the Worldwide AIS Network. It is designed not merely as a transactional medium, but as a structural pillar of the protocol's economic and validation model. The token coordinates behavior, enforces data integrity, and ensures that the individuals and organizations contributing to the network are fairly compensated for their efforts.
At its core, WAKE is used to reward contributors who submit valid AIS data. Each submission undergoes a multi-step validation process, and only after successful verification is a reward issued to the contributing node. These rewards are distributed in the form of WAKE tokens, which function as a utility within the network to compensate participants for their role in maintaining data integrity.
To ensure data quality and discourage manipulation, all rewards are subject to a mandatory holding period before being released. During this time, submissions are monitored for discrepancies, such as repeated signals, location inconsistencies, or time-based anomalies. If fraudulent behavior is detected, the submitting node may lose its reward, have its reputation score reduced, or be removed from the network altogether. This slashing mechanism is enforced on-chain, making it tamper-proof and transparent.
In addition to its reward function, the WAKE token also introduces an economic layer of accountability through its use as a collateral requirement. Nodes must maintain a small amount of WAKE in escrow in order to participate in the network. This creates a deterrent against low-quality data submission, since actors with skin in the game are incentivized to behave honestly. In the event of repeated violations, these collateral holdings may be partially or fully slashed.
Every AIS data submission also incurs a micro-fee — denominated in WAKE — to offset validation costs and reduce network spam. While negligible for legitimate contributors, these fees create a structural cost for malicious actors attempting to flood the network with fabricated data. This fee is deducted from the final reward payout, streamlining the user experience while enforcing protocol discipline.
Unlike inflationary systems that rely on continuous token issuance or speculative dynamics, WAKE is designed around fixed supply and real-world usage. Tokens are earned through protocol participation — specifically, by submitting valid AIS data that passes decentralized validation.
Validated data is made accessible to third parties through API-based services, which require WAKE tokens to interact with the network. As more users engage with the protocol to retrieve maritime intelligence, token demand increases proportionally to data access needs. This model ties contributor rewards directly to the usefulness of the network, ensuring that value flows from participation and utility — not speculation or centralized redistribution.

6 Token Emissions & Supply Model

The WAKE token is governed by a fixed-supply emission model designed to balance long-term sustainability with contributor incentives. The total supply is capped at 500 million tokens, with carefully structured allocations to ensure fairness, utility, and alignment across the ecosystem.

At genesis, the token supply will be allocated as shown in Table 1:
These token allocations are structured to support the operational needs of the network — including data validation, access infrastructure, and liquidity provisioning. All allocations outside the contributor reward pool are subject to predefined vesting schedules to discourage speculation and ensure sustained technical development. WAKE tokens do not represent equity, ownership, or any claim on network revenue. Their sole

Category Notes	%	Tokens
Mining / Node Rewards Distributed over 40+ years via 2.5% annual emissions	70%	350,000,000
Early Network Participation Tiered distribution with structured vesting	12%	60,000,000
DEX Liquidity Provisioning For initial liquidity pool creation	5%	25,000,000
CEX Liquidity Reserve Reserved for future exchange listings	3%	15,000,000
Founders & Advisors Vesting for contributors supporting protocol development and early infrastructure deployment.	10%	50,000,000

Table 1: Token Emissions and Supply Allocation

function is to enable on-chain interactions within the protocol, such as submitting data, accessing datasets, or participating in validation logic.

6.1 Token Vesting & Unlock Mechanics

To promote sustainable network growth and prevent early concentration of tokens, all non-contributor allocations are subject to structured vesting schedules. Tokens reserved for core protocol contributors will follow a 24-month vesting plan, with a 12-month cliff and subsequent linear monthly unlocks. Private token allocations will be distributed in phased tranches, each with defined lockup and vesting periods ranging from 6 to 18 months, based on participation timing and network readiness. Reserves allocated for decentralized and centralized exchange liquidity will remain locked until activated for deployment. At launch, only the contributor reward pool will be in active circulation.

6.2 Emission Curve & Node Incentives

WAKE uses a fixed annual emission rate of 2.5% of total supply, which equates to 12.5 million WAKE tokens per year. These emissions are distributed exclusively to node operators who contribute validated AIS data to the protocol.

Token emissions are expected to span a period of 40 years, gradually releasing the 350 million tokens allocated for mining rewards. Importantly, this process is non-inflationary beyond the initial supply cap — once the allocated emissions are fully distributed, no new tokens will be minted.

Reward distribution is dynamic and proportional: in the early stages, when fewer nodes are active, each participant receives a larger share of the fixed 2.5% annual emission pool. As the network scales and more contributors come online, individual rewards decrease in relative terms. This model is designed to support early adoption while maintaining long-term sustainability through protocol-defined emissions and ongoing token usage for data access.

6.3 Token-Based Data Access Model

WAKE tokens also function as the access mechanism for retrieving validated AIS data from the network. Third parties who wish to access this data — whether for commercial, regulatory, or analytical purposes — must use WAKE tokens to do so. Tokens are required to initiate queries, unlock datasets, or subscribe to filtered data streams through the network's access layer.

When a data purchase occurs, WAKE tokens are acquired and utilized as part of the access process. This creates a direct link between real-world demand for maritime intelligence and on-chain token usage.

Because token supply is capped and emissions are predictable, network sustainability is achieved without inflation. As more participants seek to interact with the network, token usage naturally scales in proportion to data demand — supporting long-term ecosystem health through practical application.

7 Node Participation & Contributor Mechanics

WAKE transforms AIS data collection from a volunteer-based effort into a decentralized, incentivized network of contributors — each operating independent receiver nodes around the world. But unlike traditional mining or staking systems, WAKE rewards are not based solely on signal volume or uptime. Instead, the protocol allocates token emissions through a mathematically defined incentive mechanism designed to optimize geographic coverage, reward data uniqueness, and preserve network integrity.
Contributors — referred to as nodes — participate by registering their AIS receiver to a blockchain wallet, establishing a verifiable on-chain identity. This identity enables reward tracking, accountability, and fraud prevention. Once active, a node captures AIS broadcasts from nearby vessels and transmits those signals (or cryptographic summaries thereof) to the protocol for validation.
However, WAKE's objective is not just to incentivize data — it is to incentivize the right data from the right locations. For that reason, the protocol includes a spatial and behavioral optimization layer that adjusts each node's reward eligibility based on how useful its contribution is to the global network.

7.1 Hardware Requirements

WAKE is designed to be inclusive and accessible — enabling participation from individuals, vessel operators, and communities around the world, including in developing regions. Any standard AIS reception setup that meets basic performance criteria can be used to operate a node.

To maintain data quality, each node must meet the following minimum requirements:

- A VHF AIS-capable receiver (e.g., SDR with support for AIS channels or dedicated AIS hardware [5])
- GPS-based timestamping, or an NTP-synchronized system clock
- Stable internet connectivity for data submission
- Consistent uptime and accurate geolocation reporting

Each hardware node is cryptographically linked to a single on-chain wallet. This ensures a one-to-one relationship between physical equipment and digital identity, enabling transparent reward attribution, behavior tracking, and fraud prevention. Multiple wallets cannot share the same hardware, and each wallet must operate independently validated equipment.
WAKE does not prescribe specific device models, but nodes must meet baseline requirements for signal fidelity, clock accuracy, and location consistency to qualify for rewards. This introduces a modest cost barrier — deterring spam or dishonest activity while remaining within reach for committed contributors.

8 Reward Allocation and Emission Model

WAKE emits a fixed 2.5% of total token supply annually, drawn from the 70% of tokens allocated for mining rewards. These emissions are distributed to active, validated nodes according to a proportional scoring system.

Let:

- $B(t)$: the total reward emission budget for time interval t
- N: the number of participating nodes during t
- R_n: the Reward Modifier Score of node n
- $W_n(t)$: the amount of WAKE tokens rewarded to node n during time interval t

Then, the reward distributed to node n is calculated as:

$$W_n(t) = B(t) \cdot \frac{R_n}{\sum_{i=1}^{N} R_i} \tag{1}$$

This ensures that the entire emission pool is distributed each epoch, but the share earned by each node depends on its relative contribution — not just its activity.

8.1 Reward Modifier Function

The Reward Modifier R_n is a composite score that captures how well a node enhances the network's overall coverage, trust, and utility. It is defined as:

$$R_n = \alpha \cdot U_n + \beta \cdot D_n + \gamma \cdot C_n + \delta \cdot N_n \tag{2}$$

Where each component is normalized to the range $[0, 1]$ across the network:

- U_n: Geographic Uniqueness Score
  Inverse function of local node density. Nodes in low-density, remote, or underserved areas receive higher scores.
- D_n: Distance Score
  Based on the mean distance to the next nearest node that received the same signal. Promotes spatial distribution.
- C_n: Signal Consistency Score
  Measures how regularly and reliably the node captures AIS signals, relative to expected regional coverage patterns.
- N_n: Signal Novelty Score
  Proportion of signals submitted by the node that were not detected by any other peer node during the same time interval.

The weights α, β, γ, δ are tunable constants set by the protocol, satisfying:

$$\alpha + \beta + \gamma + \delta = 1 \tag{3}$$

Different phases of the network's evolution may favor different weights — for example, emphasizing coverage expansion in early stages, or prioritizing signal integrity and consensus convergence in later stages.

8.2 Governance of Reward Parameters

To ensure long-term adaptability without compromising trust, the reward modifier parameters are configurable via runtime upgrades. In the early phase of network deployment, these parameters will be managed by a single protocol governor to allow rapid iteration and optimization. However, WAKE is designed for progressive decentralization. As the network scales and the token becomes more widely distributed, governance of key reward and emission settings will transition to a community DAO. This ensures that future protocol changes are transparent, auditable, and aligned with the collective interests of network participants.

8.3 Coverage Optimization in Practice

This design encourages strategic node placement and decentralized expansion. For example, a node deployed on a remote coastline with few neighbors, high uptime, and a reliable stream of unique signals will receive a disproportionately high share of the daily rewards — even if it submits fewer messages than nodes in congested areas.
Conversely, nodes that operate in already saturated regions, submit redundant data, or go offline intermittently will earn far less, despite similar hardware or effort. This architecture creates a self-balancing network topology that evolves toward global efficiency without central coordination.

8.4 Aligning Incentives with Integrity

The WAKE token model is built to compensate meaningful participation in the network — rewarding contributors based on the quality, uniqueness, and relevance of the AIS data they provide. Fixed emissions and dynamic reward weighting are used to maintain a sustainable incentive structure over time. However, meaningful participation also requires trust in the data itself. To ensure that contributions are valid and honestly earned, WAKE incorporates a multi-layered fraud detection system. This system is designed to verify each submission and convert raw AIS signals into trusted, verifiable maritime intelligence.

9 Fraud Detection and Data Integrity

9.1 Layered Validation Framework: System Overview

WAKE's approach to fraud detection is built around a core principle: data must be verifiable, even if the contributor is not trusted. The protocol assumes that any individual node may behave dishonestly — either through specific [3], collusion, or negligence — and therefore treats each AIS message as an object of scrutiny rather than trust.

To protect against these risks, WAKE implements a layered fraud detection framework, which evaluates each message using a blend of physical, geographic, statistical, and social verification techniques.

Each layer is designed to detect a specific type of fraud or anomaly as seen in Table 2:

Validation Layer	**Purpose**	**Example of Detected Attack**
Peer Agreement	Compare submissions against nearby nodes	Honest signal vs. synthetic injection
Collusion Detection	Identify replicated message sets from distant nodes	Virtual fleet mirroring
Time-of-Arrival (ToA)	Enforce physics-based timing constraints	Timestamp spoofing
Motion Modeling	Verify location consistency over time	Teleporting nodes, GPS jumps
AIS Correlation	Compare node location to vessel MMSI track	Remote relay pretending to be onboard
Fraud Scoring	Quantify behavioral risk over time	High % of flagged data = slashing

Table 2: Layered Validation Framework for Fraud Detection

These layers work together to create a self-healing, decentralized validation engine. By assessing each message through multiple independent filters, WAKE ensures that:

- Honest contributors are rewarded quickly and transparently.
- Isolated nodes are provisionally trusted but logged for retroactive analysis.
- Dishonest actors are economically disincentivized or removed.

This framework enables WAKE to maintain trustless data integrity without central oversight, even during the early phases of network deployment.

9.2 Geographic Validation Groups

WAKE must be able to trust the AIS data it receives — without trusting the nodes submitting it. To achieve this, the protocol uses a layered fraud detection system that verifies each message based on time, geography, physics, and peer consensus. This ensures that only authentic and useful data earns rewards, while dishonest behavior is penalized.

The system works by continuously comparing each node's submitted AIS data to those of nearby nodes and measuring how likely the data is to be valid. Below we explain the components of this system and how fraud is detected and punished in WAKE.

Every node in the network shares responsibility for verifying the data of nearby nodes. WAKE groups nodes into local validation zones based on their physical location.

Each node reports its current location using GPS coordinates. Based on these coordinates, the protocol forms a group of nearby nodes within a defined radius (typically 40 km, corresponding to AIS radio range). This group is known as the node's validation group.

If we let:

- $G_n(t)$ be the GPS location of node n at time t
- $d(Gn, Gj)$ be the geographic distance between node n and node j
- r be the radius of validation (e.g., 40 km)

Then node n's validation group is defined as:

$$V_n(t) = \{j | d(G_n(t), G_j(t)) \leq r\}$$

This means that all nodes within 40 km of node n at time t are included in n's local validation group. These peers are used to cross-check whether a submitted AIS message appears to be valid or not.

10 Peer Agreement and Adaptive Confidence Scoring

A foundational component of WAKE's validation mechanism is peer agreement — the process by which a node's AIS message submissions are cross-referenced with those submitted by nearby nodes. The presence of independent confirmations by geographically co-located peers significantly increases the likelihood that a message is valid. This provides a decentralized, trustless method for identifying legitimate AIS data, particularly in regions with multiple contributors.

Each node in the WAKE network belongs to a validation group, defined by a fixed radius (e.g., 40 kilometers) corresponding to the expected maximum reception range of AIS VHF transmissions. Messages submitted by a node are compared against those received by members of its validation group during the same time interval.

Let us define:

- $V_n(t)$ as the set of all nodes within radius r of node n at time t — i.e., its validation group
- $|V_n(t)|$ as the number of peers in that group
- $S_j(m)$ as an indicator function, where:

$$S_j(m) = \begin{cases} 1 & \text{if peer node } j \in V_n(t) \text{ submitted message } m \text{ during time } t \\ 0 & \text{otherwise} \end{cases}$$

Then the confirmation count C_m is calculated as:

$$C_m = \sum_{j \in V_n(t)} S_j(m)$$

This count represents the number of neighboring nodes that independently received the same AIS message.

To account for the number of available peers, WAKE defines a Trust Score T_m for each message, reflecting the network's confidence in its authenticity. This score is dynamically scaled based on the size of the validation group:

$$T_m = \begin{cases} 1, & \text{if } |V_n(t)| = 0 \\ \frac{C_m}{|V_n(t)|}, & \text{if } |V_n(t)| > 0 \end{cases}$$

Where:

- $T_m \in [0, 1]$ reflects the network's confidence in message m
- $T_m = 1$ when there are no nearby nodes, granting default trust in isolated regions
- $T_m = 1$ when all peers confirm the message
- Lower values of T_m indicate reduced peer consensus and increased suspicion

This function ensures that nodes operating in isolation — for example, early adopters or contributors in remote areas — are not unfairly penalized for lack of local coverage. Their submissions are provisionally trusted unless flagged by other mechanisms (e.g., signal timing or location anomalies).

However, in early deployment phases or remote areas, a node may operate in isolation or with only a small number of nearby peers. To accommodate this, WAKE applies a tiered trust model. A node with zero peers is not penalized. Instead, it receives provisional trust — allowing its data to be accepted initially, while being flagged for retroactive validation once additional peers come online in that region.

As the network matures and density increases, the required level of peer agreement scales accordingly. In high-density environments, a message that lacks validation from nearby peers will receive a low trust score and may be flagged for further inspection or penalization.

10.1 Collusion Detection via Message Similarity

While peer agreement within a validation group is a key indicator of data reliability, excessive similarity between nodes outside of their expected reception range may indicate collusion or dishonest behavior. In particular, nodes operated by the same entity or colluding actors may attempt to submit identical AIS data across distant locations to artificially boost their trust scores and mining rewards.
To counter this, WAKE continuously analyzes the similarity of message sets between all nodes — and adjusts their Collusion Risk Score based on both the degree of similarity and their physical separation.

10.2 Message Similarity Measurement

For any two nodes, i and j, WAKE defines their set of submitted messages during time window t as:

- $M_i(t)$: the set of unique AIS messages submitted by node i during time t
- $M_j(t)$: the same for node j

WAKE then computes the agreement ratio $A_{i,j}(t)$ as:

$$A_{i,j}(t) = \frac{|M_i(t) \cap M_j(t)|}{\min(|M_i(t)|, |M_j(t)|)}$$

Where:

- $|M_i(t) \cap M_j(t)|$ is the number of messages both nodes submitted in common
- $\min(|M_i(t)|, |M_j(t)|)$ normalizes for unbalanced message counts between the two nodes

This ratio gives a value between 0 and 1:

- $A_{i,j}(t) = 1$ means all messages submitted by the smaller node set are identical to those of the other node
- $A_{i,j}(t) = 0$ means no overlap, i.e., completely different message sets

In normal operating conditions, nodes that are geographically close may have high agreement due to overlapping AIS reception. However, if two distant nodes show a high degree of message overlap, the likelihood of coordination or collusion increases.

10.3 Collusion Risk Score

To quantify this, WAKE defines a Collusion Risk Score CRi,j(t) for each node pair:

$$CR_{i,j}(t) = A_{i,j}(t) \cdot \left(1 + \frac{d(G_i, G_j) - r}{r}\right)$$

Where:

- $A_{i,j}(t)$ is the message agreement ratio (as defined above)
- $d(G_i, G_j)$ is the geographic distance between the nodes' reported positions
- r is the expected AIS signal range (e.g., 40 km)

This equation increases the collusion risk score proportionally to both:

1. The agreement ratio, and
2. The excess distance beyond expected radio range

Interpretation:

- If two nodes are within signal range ($d(G_i, G_j) \leq r$), the multiplier is 1, and $CR_{i,j}(t) = A_{i,j}(t)$ — high agreement is expected and not penalized.
- If the nodes are far apart, the multiplier becomes greater than 1, and identical messages become increasingly suspicious.

For example, if two nodes are 80 km apart ($d = 2r$) and have an agreement ratio of 0.9, then:

$$CR_{i,j}(t) = 0.9 \cdot \left(1 + \frac{2r - r}{r}\right) = 0.9 \cdot 2 = 1.8$$

WAKE defines a global collusion threshold . If CRi,j(t) ¿ , the node pair is flagged for further review.

10.4 Collusion Response & Penalties

When nodes exceed the collusion risk threshold, the protocol may:

- Reduce or suppress their trust scores
- Trigger a slashing penalty if message duplication is confirmed
- Mark the nodes as linked or untrustworthy, lowering their influence in validation
- Temporarily suspend rewards pending investigation

11 Physics-Based Validation: Time-of-Arrival Checks

AIS signals are transmitted over VHF radio frequencies — typically 161.975 MHz and 162.025 MHz [2] — and, like all electromagnetic waves, they propagate at the speed of light in air. WAKE leverages this physical constant to verify whether the timestamps reported by different nodes for the same message are physically plausible. By comparing when a message was received by two different nodes and how far apart those nodes are, the protocol can determine whether their timing is consistent with the known limits of signal propagation. This serves as a powerful anti-fraud mechanism — detecting spoofed, replayed, or misaligned signals that could otherwise go undetected by peer agreement alone. If two nodes are a certain distance apart, the signal cannot reach both of them instantly. It must take at least a minimum amount of time to travel that distance. If their reported timestamps are too close together — or too far apart — something is likely wrong.

Let us define:

- c_{air}: speed of light in air $\approx 3 \times 10^8$ m/s
- $G_i(t)$ and $G_j(t)$: GPS coordinates of nodes i and j at the time of reception
- $d(G_i, G_j)$: geographic distance between the two nodes in meters, calculated using great-circle or Euclidean distance
- $t_i(m)$: timestamp when node i received message m
- $t_j(m)$: timestamp when node j received the same message m

11.1 Expected Propagation Delay

This is the amount of time the signal *should* take to travel from one node to the other, if it moved at the speed of light:

$$\Delta t_{expected} = \frac{d(G_i, G_j)}{c_{air}}$$

This is a lower bound — real-world reception may involve delays due to processing, queuing, or minor atmospheric variations.

11.2 Observed Time Difference

This is the actual time difference between when node i and node j reported receiving the same message:

$$\Delta t_{observed} = |t_i(m) - t_j(m)|$$

If the two nodes are close together, this value should be small. If they are farther apart, a larger time difference is expected.

11.3 Signal Timing Error

WAKE then calculates the signal timing error, which is the difference between the expected and observed delays:

$$E_{i,j}(m) = |\Delta t_{observed} - \Delta t_{expected}|$$

This value tells us how well the real-world timestamps match the physics of signal propagation. If the error is very small, the data is likely genuine. If it is large, it suggests either timestamp spoofing, clock manipulation, or signal injection.

11.4 Threshold and Flagging

WAKE defines a small tolerance threshold ϵ (in seconds), representing the maximum allowed deviation due to hardware or software lag, clock sync drift, or atmospheric interference.

If:

$$E_{i,j}(m) > \epsilon$$

then the message m is flagged as physically implausible for the node pair (i, j). Repeated violations of this constraint across multiple messages or peer pairs will increase the node's fraud score (Section 7.6.6), and may ultimately trigger economic penalties or slashing4.

12 Node Geolocation and Motion Consistency

All WAKE nodes — whether fixed installations or mobile platforms — are required to submit accurate, periodic geolocation updates as part of their operation. This ensures the network can verify each node's physical position and assess the validity, uniqueness, and trustworthiness of its AIS data contributions.

Geolocation data is essential for:

- Forming validation groups (Section 7.6.1)
- Scaling peer trust scores (Section 7.6.2)
- Evaluating collusion likelihood (Section 7.6.3)
- Enforcing signal propagation constraints (Section 7.6.4)
- Detecting geographic spoofing and mobility fraud

In addition to broadcasting AIS messages, each node must report its position in the format:

$$G_n(t) = (lat_n(t), lon_n(t))$$

These coordinates must be updated at regular intervals, synchronized to the protocol's time epochs.

13 Motion Modeling for All Nodes

WAKE treats all nodes as position-aware actors, but applies different analytical models depending on whether the node is fixed or mobile.

A node is classified as mobile if its reported position changes significantly over time. In these cases, WAKE applies a motion consistency check to verify the plausibility of the node's movement.

Let:

- Δt be the time between successive geolocation submissions
- $d(G_n(t), G_n(t + \Delta t))$ be the measured distance moved
- $v_n(t) = \frac{d(G_n(t), G_n(t+\Delta t))}{\Delta t}$ be the implied velocity

WAKE validates that this velocity:

- Is within physical bounds (e.g., < 50 km/h for maritime platforms)
- Matches prior movement trends
- Reflects continuous, plausible trajectories

If abrupt location changes or discontinuous jumps are detected, the node is flagged for potential location spoofing or virtual node replication. If abrupt location changes or discontinuous jumps are detected, the node is flagged for potential location spoofing or virtual node replication.

13.1 Correlation with AIS Beacon Data (When Available)

In many cases, nodes will be hosted on vessels that also transmit AIS signals identifying their position and course. WAKE can correlate a node's declared geolocation $G_n(t)$ with the vessel's known AIS position at the same timestamp.

If a node claims to be aboard vessel MMSI X, but its geolocation deviates significantly from X's AIS beacon, the protocol may:

- Reduce the node's trust score
- Flag the discrepancy for manual review
- Apply slashing if fraudulent intent is confirmed

This mechanism binds AIS receivers to their physical environment — making it difficult to submit data remotely while claiming to be at sea.

14 Fraud Scoring and Economic Slashing

WAKE enforces data integrity by applying fraud scoring and conditional payment logic to every node. This ensures that contributors are economically accountable for the validity of the AIS data they provide — and that dishonest actors are penalized swiftly and transparently. At the core of this system is a 10-day reward lock: newly earned WAKE tokens are held in escrow by the protocol for 10 days after data submission. During this period, each node's submissions are evaluated for fraud risk based on validation logic described in previous sections.

14.1 Fraud Probability Score

Each node accumulates a rolling fraud probability score $P_n(t)$ that reflects the proportion of its AIS messages flagged as invalid. This score is calculated over a rolling 5-epoch window, allowing the protocol to detect sustained or repeated fraudulent behavior while tolerating occasional errors.

A message is only marked invalid if it fails two or more independent validation checks, such as:

- Low peer agreement (signal not seen by nearby nodes)
- Inconsistent geolocation or motion
- Beacon mismatches (conflicting MMSI or GPS)
- Message duplication across distant receivers (potential collusion)
- Time-of-arrival anomalies (radio signal delay inconsistency)

The rolling fraud score is defined as:

$$P_n(t) = \frac{f_n(t_{-4} \to t)}{T_n(t_{-4} \to t)}$$

where:

- f_n is the number of failed messages from node n in the last 5 epochs
- T_n is the total number of messages submitted by node n during the same period

This value lies in the range $[0, 1]$ and indicates how much of a node's recent output is considered potentially invalid.

Reward Lock and Validation

Each node's earned rewards are held in a locked state for 10 epochs. During this time:

- The network assesses fraud risk using the rolling score $P_n(t)$
- Rewards are released only if the fraud score remains within acceptable bounds

Nodes may view their pending rewards, but cannot transfer or use tokens until validation is complete and the lock period has elapsed.

Slashing and Banning Logic

- Slashing: If a node's $P_n(t) \geq 0.30$ for two consecutive epochs, all pending rewards are slashed.
- Permanent Ban After three total slashing events, the node is permanently removed from the network.

Slashing is applied automatically and without appeal. Wallets are tied to hardware identity, and banned nodes cannot rejoin or resubmit data.

14.2 Severity-Weighted Slashing

If a node's data is found to be invalid, the protocol applies a slashing penalty against the locked reward — reducing or eliminating it before release.

Let:

- $P_n(t)$ = fraud probability score
- V_n = severity multiplier, based on violation type (e.g., 1 for redundancy, 3 for spoofing)
- $R_n(t)$ = total WAKE reward earned by node n during epoch t (and currently locked)

Then the slashing amount is:

$$S_n(t) = P_n(t) \cdot V_n \cdot R_n(t)$$

The final amount released to the node after 10 days is:

$$R_n^{released}(t + 10) = \max(R_n(t) - S_n(t), 0)$$

If no fraud is detected, the full reward is released. If slashing is triggered, only the remainder is paid out.

15 Early Deployment Logic and Retroactive Validation

In low-density phases of the network, or when operating in remote regions, a node may be the only one active in its area. To encourage early deployment without compromising protocol integrity, WAKE adjusts its validation logic accordingly.

If a node has fewer than 3 peers in its validation group, the following applies:

- Peer-based slashing is disabled
- Physics-based checks still apply
- Data is logged and stored for retroactive validation

As new nodes come online in the same region, WAKE re-evaluates older data using updated peer consensus. If the data is found to be fraudulent, penalties may be applied retroactively.

16 Less Trust, More Truth

WAKE is a trustless system. Nodes are not assumed to be honest — their data must prove itself.

In dense regions, peer consensus exposes fraud through cross-verification. In remote zones, physics and motion constraints ensure that even isolated submissions remain accountable. All data is stored with metadata for retroactive validation as new peers come online.

WAKE combines geospatial correlation, signal timing, behavioral scoring, and statistical consensus to filter truth from noise — without relying on identity, reputation, or central oversight.

This is not just a decentralized AIS feed. WAKE is the first global maritime data layer where trust is computed, not claimed.

17 Integrity by Design: Attack Resistance and Limitations

WAKE is built for a world in which no node can be trusted — only the data they provide. This adversarial mindset is fundamental to the protocol's architecture. From economic slashing to physics-based validation, WAKE is engineered to resist manipulation at every layer.

The protocol's design assumes that some participants will attempt to game the system for profit. It accounts for actors who may spoof messages, replicate data, fabricate GPS positions, or attempt to flood the network with low-value submissions. Each of these threats is met with a multi-layered mitigation strategy.

Peer agreement mechanisms prevent isolated nodes from fabricating consensus. Collusion scoring detects duplicated message sets across long distances. Signal timing validation uses the speed of light as a hard constraint on spoofed timestamps. Node motion modeling checks for improbable geographic jumps. And all economic rewards are time-locked — providing a window for these systems to catch abuse before payment is released.

WAKE is also economically hardened. By requiring all contributors to lock tokens for ten days before earning rewards — and by slashing that collateral if fraud is detected — the protocol makes dishonest behavior unprofitable at scale. The cost of deploying and maintaining hundreds of spoofing nodes outweighs

the potential rewards, especially in a network designed to downrank duplicate data and favor geographic diversity

While no system can eliminate all edge cases, WAKE is designed to detect inconsistency at scale. By requiring geographic, temporal, and behavioral consistency across multiple independent nodes, the protocol makes manipulation increasingly difficult to sustain. Even in edge scenarios — such as isolated or mobile deployments — WAKE's architecture ensures that all data is subject to retrospective validation and economic accountability.

18 Data Storage & Integrity Architecture

WAKE employs a hybrid storage model to ensure scalability without compromising validity. Given the high volume and frequency of AIS broadcasts, storing raw data entirely on-chain would be inefficient and cost-prohibitive. Instead, WAKE separates each data submission into two layers:

- **Off-Chain Payloads:** Raw AIS messages — including MMSI, timestamp, latitude/longitude, SOG, and COG — are stored off-chain in a decentralized but scalable data infrastructure. This allows the network to handle millions of daily submissions while keeping storage costs and bandwidth requirements manageable.
- **On-Chain Metadata:** For every submission, a cryptographic hash of the AIS message, the submitting node's ID, timestamp, validation results, and reward amount are committed on-chain. This creates a tamper-proof audit trail that links each AIS message to its origin and validation history, without exposing sensitive or bulky data directly on-chain.

Each off-chain file is cryptographically bound to its on-chain reference, ensuring the two cannot diverge. If a message is altered or deleted off-chain, it will fail integrity checks against its on-chain hash — making tampering immediately detectable.

This model ensures that:

- Nodes cannot manipulate stored data after submission
- Buyers and auditors can verify provenance without trusting the submitter
- Storage infrastructure remains scalable across global deployments

19 Data Access & Commercial Integration

WAKE is not just a decentralized network — it's a commercial-grade data layer built for real-world use. Once AIS data is validated through the protocol's multi-layered consensus engine, it is made available through secure and structured delivery channels:

- **API Access:** Real-time and historical AIS data can be accessed by enterprise clients through tiered subscription APIs, supporting applications across insurance, logistics, environmental monitoring, and port operations.
- **Bulk Licensing:** Government agencies, academic institutions, and analytics providers can license complete datasets, backed by cryptographic proofs of origin and validation.
- **Custom Intelligence Streams:** Users with advanced needs — such as compliance, risk scoring, or illicit behavior detection — can subscribe to filtered data feeds tailored by geography, vessel type, or behavioral profile.

All data products are priced in fiat to ensure seamless enterprise adoption. Revenues are managed by the WAKE Foundation (or a delegated entity) and used to strengthen the token economy via market buybacks.

20 Conclusion

WAKE redefines the foundations of global maritime intelligence. By decentralizing AIS data collection and anchoring validation in cryptographic truth rather than institutional trust, it introduces a radically more resilient, transparent, and equitable model for how maritime data is captured, verified, and monetized.

This protocol is more than a technical upgrade — it is a structural correction to a broken system. For too long, AIS data has been harvested from unpaid contributors, processed by black-box platforms, and sold without any accountability to those who made it possible. WAKE fixes this at every layer: contributors are rewarded, data is validated through peer consensus and physics-based constraints, and the economic loop is closed through real-world demand and token scarcity.

The implications extend far beyond shipping lanes.

- **Supply Chain Optimization:** Logistics companies can access verified vessel positions in real-time, improving ETAs, port scheduling, and multimodal coordination.
- **Insurance & Risk Analytics:** Underwriters and actuaries gain access to validated movement patterns, incident detection, and compliance data to improve modeling and reduce exposure.
- **Commodities & Trade Finance:** Traders, banks, and analysts can monitor cargo flows, detect anomalies in global movement, and verify delivery timing — enabling more accurate pricing, counterparty verification, and risk scoring.
- **Environmental Monitoring:** NGOs and regulators can track emissions, identify illegal fishing, and monitor vessel behavior in ecologically sensitive zones with confidence.
- **Security & National Defense:** Governments and intelligence agencies can spot spoofed signals, detect "dark fleet" behavior, and corroborate ship positions with independent, tamper-proof signals.

WAKE offers a universal, permissionless data infrastructure designed to meet the needs of diverse maritime stakeholders — all while ensuring data integrity through cryptographic validation. It introduces not only a new model for contributing AIS data, but a new standard for verifying its authenticity and provenance.

As adoption increases, WAKE has the potential to serve as foundational infrastructure for a global maritime data ecosystem — one where trust is built on computation rather than assumption, where contributors are recognized for their participation, and where access to information is earned through transparent, protocol-defined interaction.

References

[1] Raymarine. AIS - Automatic Identification System [Internet]. Raymarine; 2024 [cited 2025 Apr 5]. Available from: `https://www.raymarine.com/en-gb/view/blog/what-is-ais`

[2] International Telecommunication Union. Technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile band. ITU-R M.1371-5. Geneva: ITU; 2014.

[3] Polestar. AIS Spoofing: Detecting Maritime Deception [Internet]. Polestar; 2023 [cited 2025 Apr 5]. Available from: `https://www.polestarglobal.com/news/ais-spoofing-detection`

[4] He B, Li Y. Slashing mechanisms in blockchain consensus: a survey and taxonomy. ACM Computing Surveys. 2023;56(2):34.

[5] Miltech Marine. What Hardware Do I Need to Receive AIS? [Internet]. Miltech Marine; 2023 [cited 2025 Apr 5]. Available from: `https://www.miltechmarine.com/pages/faqs`

[6] Shi W, Zhang J, Gong H. Carbon footprint of blockchain consensus mechanisms: A review. Renewable and Sustainable Energy Reviews. 2022;153:111777.

[7] Lys A, Daniels C, Mølgaard P. The Nothing-at-Stake Problem in Proof-of-Stake Blockchains. Distributed Computing. 2021;34(4):789–804.

Appendix: Glossary of Terms

AIS (Automatic Identification System): A maritime radio-based tracking system that broadcasts a vessel's identity, position, speed, and course to enhance safety and transparency at sea.

API (Application Programming Interface): A structured interface that allows external applications to programmatically access WAKE's validated AIS data streams.

COG (Course Over Ground): The actual direction in which a vessel is moving over the Earth's surface, expressed in degrees relative to true north.

Cryptographic Hash: A unique digital fingerprint produced from input data, used to ensure the authenticity of data and detect any tampering.

Decentralized Protocol: A system managed by a distributed network of participants rather than a central authority, which ensures openness and resilience.

Emission Curve: The predetermined schedule dictating the rate at which new WAKE tokens are issued to contributors over time.

GPS (Global Positioning System): A satellite-based navigation system that provides precise location and time data to receivers on Earth.

MMSI (Maritime Mobile Service Identity): A unique nine-digit identifier assigned to a vessel's AIS transponder for identification in AIS transmissions.

Node: A participant in the WAKE network that operates AIS reception hardware and submits validated data in exchange for token rewards.

Off-chain / On-chain: Data stored off-chain resides outside the blockchain, while on-chain data is recorded directly on the blockchain ledger to provide a tamper-proof audit trail.

Slashing: A penalty mechanism by which a portion or all of a node's reward is forfeited due to detected protocol violations, such as submitting falsified or inconsistent data.

SOG (Speed Over Ground): The actual speed at which a vessel moves over the Earth's surface, as determined by GPS.

Trust Score: A dynamic metric that reflects a node's historical data quality, reliability, and overall performance in the network.

Validation Group: A collection of geographically proximate nodes that cross-check each other's submitted AIS data to ensure authenticity.

VHF (Very High Frequency): A radio frequency range (typically 30–300 MHz) used by AIS systems, with typical AIS transmissions occurring in the 161.975–162.025 MHz band.