international journal of critical infrastructure protection 51 (2025) 100812

Contents lists available at ScienceDirect

International Journal of Critical Infrastructure Protection

journal homepage: www.elsevier.com/locate/ijcip

Robust trust management in Intelligent Transportation System: A machine learning approach

Ahmed Danladi Abdullahi [a,*], Erfan Bahrami [b], Tooska Dargahi [a], Mohammed Al-Khalidi [a], Mohammad Hammoudeh [a]

[a] Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, United Kingdom
[b] Computer Engineering Department, Sharif University of Technology, Tehran, Iran

ARTICLE INFO

Keywords:
Intelligent Transportation System
Trust management
Machine learning
Feedforward neural network
Cyber–physical system
Vehicular networks

ABSTRACT

Intelligent Transportation Systems (ITS) are revolutionising modern mobility by leveraging advancements in 5G technology, smart sensors, and sophisticated data analytics. These advancements facilitate the exchange and decision making of information in real time, improving safety and efficiency. However, the heterogeneous and loosely connected nature of the ITS components presents significant challenges in evaluating and managing trust within the ecosystem. Traditional approaches, such as blockchain-based consensus mechanisms, peer-to-peer voting systems, and static rule-based trust models, struggle to evaluate trust uniformly across diverse components and data types in real time, leaving the system vulnerable to various threats. Recent studies explored Machine Learning (ML) techniques to address trust management in ITS. These advanced approaches offer promising solutions for processing large volumes of heterogeneous data, identifying complex patterns, and adapting to dynamic environments. However, most existing ML-based solutions focus on assessing trust for particular components, such as vehicles and roadside units (RSUs), rather than addressing the collective trust of the entire ITS ecosystem.

This paper proposes a novel ML-based dynamic trust management system termed MLT. It employs a feedforward neural network and the Levenberg–Marquardt Algorithm to dynamically assess the trustworthiness of ITS components. The system incorporates a dynamic time decay factor and continuously updates the trust scores, allowing effective identification and isolation of malicious actors. Through extensive simulations, MLT outperforms baseline models by up to 10% in precision and 9% in F-measure across various attack scenarios. These results highlight the superior performance of MLT in accuracy and robustness compared to existing trust management models.

1. Introduction

Intelligent Transportation Systems (ITS) are revolutionising the perception and experience of transportation, emerging as a cornerstone of modern and connected mobility networks. Rapid advancements and global adoption of cutting-edge mobile technologies, such as 5G, have paved the way for the integration of vehicles, infrastructure, and data, creating a dynamic ecosystem that promises to transform the future of transportation. The diverse sensors and sophisticated communication technologies embedded within modern vehicles are at the heart of this transformation. These advanced systems grant vehicles unprecedented access to information services, keeping drivers and other road users informed about critical factors, such as road conditions, weather patterns, and real-time safety messages. Simultaneously, these technologies enable vehicles to monitor and analyse their internal systems, from tyre pressure monitoring to collision avoidance and traffic signal recognition, empowering them to make split-second decisions that enhance safety, efficiency, and overall performance.

The convergence of these technologies forms the foundation of the connected vehicle ecosystem, a concept that goes beyond mere transportation and embraces the notion of intelligent, collaborative mobility. It is crucial to recognise that ITS is inherently heterogeneous, comprising a vast array of loosely connected devices and sensors. These include roadside units (RSUs), which serve as communication hubs along transportation corridors, as well as pedestrians equipped with smart devices, vehicles with advanced onboard systems, and a myriad of other road infrastructures and sensors that use channels such as vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-pedestrian (V2P) communications to exchange information and ensure

* Corresponding author.
E-mail address: ahmed.abdullahi2@stu.mmu.ac.uk (A.D. Abdullahi).

https://doi.org/10.1016/j.ijcip.2025.100812
Received 16 December 2024; Received in revised form 14 October 2025; Accepted 21 November 2025
Available online 25 November 2025

safety [1]. This heterogeneity poses unique challenges regarding interoperability, data integration, and trust [2]. With devices and sensors from various manufacturers, each with its protocols and security standards, ensuring seamless communication and data exchange becomes complex. Moreover, the loosely connected nature of these components raises concerns about the reliability and security of the information being transmitted, necessitating robust trust management mechanisms to maintain the integrity of the entire system.

In recent years, intelligent transportation systems have been increasingly targeted by sophisticated cyber attacks and inferences that exploit vulnerabilities in data integrity, communication protocols, and perception systems. Recent studies have identified a broad spectrum of attack vectors, including data poisoning and false information campaigns that manipulate vehicle sensor streams and message exchanges [3]. For example, Usha et al. [4] demonstrated the susceptibility of vehicular networks to distributed denial-of-service (DDoS) attacks, achieving over 94% detection accuracy using adaptive neuro-fuzzy models. In contrast, Wang et al. [3] provided an extensive survey of poisoning and inference attacks that undermine the integrity of ITS data flows. Other recent work explores game-theoretic deception models and sensor spoofing techniques that distort situational awareness and disrupt coordination among autonomous vehicles [5,6]. Furthermore, our previous work, LIFT the AV: Location InFerence aTtack on Autonomous Vehicle Camera Data [7], exposed how adversaries can infer sensitive location information from camera metadata, compromising vehicular privacy even without direct access to GPS signals. Collectively, these studies highlight the growing sophistication and diversity of threats targeting ITS infrastructures, reinforcing the need for adaptive, trust-based mechanisms capable of jointly assessing both the reliability of entities and the credibility of their shared information.

This paper specifically addresses the fundamental challenge of establishing and maintaining trust in ITS environments. The notion of trust has been conceptualised in multiple ways across the literature. Early studies defined trust as a subjective belief in the honesty, competence, and benevolence of an entity [8]. In cyber–physical and vehicular systems, trust has also been treated as a probabilistic measure of behavioural expectation [9], a reputation-based score aggregated from peer evaluations [10,11], or a belief-evidence combination using subjective logic [8,12]. These definitions differ mainly in whether they emphasise behavioural consistency or information credibility. For this work, we adopt a measurable and data-driven definition aligned with ITS requirements, where trust is quantified through two complementary dimensions: reliability, which captures the consistency and correctness of component behaviour, and credibility, which represents the authenticity and accuracy of the exchanged information. This dual perspective supports real-time machine learning-based assessment of trust across heterogeneous ITS components, where both the dependability of entities and the veracity of messages are critical for safety. In this context, trust is defined as a quantifiable measure of the reliability and credibility assigned to ITS components, representing the probability that an entity will behave as expected and provide accurate information [13]. Trust is particularly critical in ITS because unreliable or malicious communications can lead to safety hazards, traffic disruptions, and overall inefficiencies [14]. Given the heterogeneous and decentralised nature of ITS, developing robust trust mechanisms is essential to mitigate threats such as data tampering, message spoofing, and denial-of-service (DoS) attacks [15], all of which can compromise the safety and performance of transportation networks [15]. These trust mechanisms operate across multiple OSI layers: filtering malicious traffic at the network layer, prioritising legitimate connections at the transport layer, and allocating processing resources based on sender reputation at the application layer [16].

Building upon these challenges, machine learning and deep learning approaches have been increasingly applied to trust management in mobile ad-hoc networks and ITS. These advanced techniques offer promising solutions for processing large volumes of heterogeneous data, identifying complex patterns, and adapting to dynamic environments. For instance, Imana et al. [17] use Radial Basis Function Neural Networks to predict node trust in mobile ad-hoc networks, achieving high accuracy with dynamic adaptability. Trofimova et al. [18] adopt Multilayer perceptron and backpropagation algorithms for trust assessment based on Packet Delivery Ratio in ad hoc networks. Han et al. [19] propose a two-phase approach combining unsupervised k-means clustering and supervised Support Vector Machines for trust evaluation in underwater acoustic sensor networks. In a vehicular network scenario, El-Sayed et al. [20] develop an entity-centric trust model using decision trees and artificial neural networks. This demonstrates exemplary performance and robustness in trust evaluation for vehicle nodes. These approaches showcase the potential of machine learning to handle the complexity and dynamism of trust management in ITS environments.

Despite these efforts, several challenges remain in applying machine learning in trust management and evaluation in ITS. Most existing solutions focus on assessing trust for particular components of ITS rather than addressing the collective trust of loosely connected devices that comprise the entire system. This narrow focus fails to capture the complex interactions and interdependencies within ITS ecosystems [21,22].

To address these challenges, MLT employs a feedforward neural network that processes diverse inputs from heterogeneous ITS components, enabling real-time trust evaluations in the dynamic transportation environment [23]. This network's scalable architecture adapts to the evolving ITS landscape, accommodating new devices and data types as they emerge. Its pattern recognition capabilities allow for effective anomaly detection and threat identification within vast ITS data streams. At the same time, its non-linear problem-solving ability, facilitated by activation functions, models the complex interactions between various ITS components and trust factors [23,24].

The key contributions of this research are as follows;

- A robust trust management system, called MLT, which leverages machine learning to address the limitations of traditional trust evaluation methods in ITS. Traditional approaches often rely on fixed rules or static formulas, which struggle to adapt to the highly dynamic and heterogeneous nature of ITS. MLT, in contrast, uses a feedforward neural network to perform real-time trust assessments across diverse components such as vehicles, RSUs, and pedestrian-enabled IoT devices, ensuring a robust and adaptive trust evaluation process.
- A novel framework that incorporates both immediate trust assessments and the ability to adapt trust levels over time. The framework dynamically adjusts trust scores based on interactions and changing network conditions, enabling the system to effectively identify and isolate malicious actors. This adaptability is achieved using optimisation techniques that enhance the stability and reliability of trust evaluations, ensuring their applicability even in rapidly changing environments.
- Comprehensive validation of MLT through simulations incorporating three attack scenarios: simple attacks, recommendation attacks, and zigzag attacks. The experimental evaluation demonstrates MLT's superior performance, achieving approximately 90% precision, above 87% recall, and exceeding 90% F-measure even with 40% malicious nodes in the network. These results are supported by rigorous comparative evaluation against recent trust management models in the literature, demonstrating significant improvements in both accuracy and robustness.

The remainder of this paper is organised as follows. Section 2 provides an overview of related work on trust management in ITS. Section 3 presents the details of the system design. Section 4 describes the proposed ML trust management system model. Section 5 discusses the evaluation metrics and simulation results. Finally, Section 6 concludes the paper and outlines future research directions.

2. Related work

Before examining ITS-specific approaches, we contextualise trust management within the broader Cyber–Physical Systems (CPS) domain, as ITS represents a specialised CPS application. Trust management in CPS has evolved through diverse paradigms, including reputation-based, probabilistic, fuzzy-logic, and blockchain-enabled frameworks. These approaches collectively aim to ensure reliability, resilience, and security in environments that tightly couple cyber components with physical processes. For example, blockchain-based CPS frameworks [25,26] ensure tamper-evident and transparent transactions. Reputation systems [27,28] provide distributed decision-making, and fuzzy-logic trust models [29–31] improve interpretability in uncertain environments. Although these methods improve system robustness in industrial and IoT-oriented CPS, their limited scalability and static configuration hinder effective adaptation to high mobility, data-intensive ITS contexts.

Recently, substantial efforts have been made to formulate a reliable trust management framework for ITS. These studies [32–36] primarily assess the trustworthiness of individual components of ITS, such as vehicles, sensors and RSUs. At the same time, some focus on the trustworthiness of the data generated and transmitted by these components, e.g., [37–41]. However, the highly dynamic, heterogeneous and mobile nature of ITS poses significant challenges for trust management. With vehicles constantly moving at high speeds, the network topology is subject to rapid changes, making it difficult for components to promptly evaluate the trustworthiness of all other components they interact with [16]. In addition, the vast amount of data generated by ITS, including traffic alerts, weather updates, and safety messages, further complicates the trust management process because traditional systems struggle to process and evaluate trust uniformly across diverse data types in real time [41]. The multitude of data points from varying sources with different reliability levels increases the risk of malicious inputs and requires constant trust reassessment [36]. This complexity overwhelms conventional methods, necessitating more sophisticated and dynamic approaches.

Chen et al. [33] propose a decentralised trust management system (DTMS) for intelligent transportation environments. The system leverages blockchain technology and trusted execution environments (TEEs) to ensure secure, transparent, and efficient trust evaluation and management, overcoming the limitations of traditional centralised trust management systems. DTMS consists of three key components: a trust evaluation model that employs a decentralised consensus-based approach, an incentive model that encourages participation and penalises malicious behaviour, and a consensus model that ensures the transparency and irreversibility of trust credits stored on the blockchain.

Ramesh et al. [32] propose a peer-to-peer trust management system for ITS based on Aumann's agreement theorem, aiming to establish the truth of consensus to identify malicious nodes without delay by considering the decision of each vehicle in finding trust values. The proposed dual-weighted trust architecture uses a voting-based intrusion detection system (IDS) to evaluate the trustworthiness of nodes, selecting vote participants using a hash function and grouping them based on their votes about a target node. The Aumann's agreement theorem is then applied to these groups to make the nodes converge towards a truth of consensus, and the decision parameter of the target node is used to update the trust values of voting participants, with malicious nodes having null trust values being identified and removed from the system. In [34], the authors propose an AI-enabled trust management system (AIT) for vehicular networks using blockchain technology to address the security challenges posed by erroneous or malicious traffic-related messages disseminated by compromised vehicles. The AIT system consists of five key steps: traffic data collection and message generation by vehicles, local trust level (LTL) calculation by vehicles and sharing with local RSUs, global trust level (GTL) calculation by local RSUs, trust validation and archiving using blockchain, and GTL voting and dissemination by all RSUs.

Cheng et al. [35] propose a trust-aware control framework for ITS, modelled as multi-agent systems (MAS) with both traffic participants and infrastructure components as interacting agents, to address the vulnerability of increased autonomy to compromised or malfunctioning agents that threaten the safety and efficiency of the entire system. The proposed framework embeds a trusted authority (TA) into the transportation infrastructure to systematically quantify the trustworthiness of agents using subjective logic based on direct and indirect observations of agent behaviours, and these quantified trustworthiness scores are then used to enable trust-aware coordination and control.

Müller et al. [36] propose a subjective logic-based trust management and misbehaviour detection mechanism for multi-agent systems (MAS), mainly focusing on ITS, to ensure the reliability of cooperative information shared among agents. The mechanism amends reliability information to the data shared in the MAS, fuses information from multiple agents reporting the same event, and detects and isolates misbehaving agents, including faulty and malicious ones. The authors specify an attacker model and consider trade-offs between security, generality, and resource consumption.

Qi et al. [37] propose a hybrid-trust-based emergency message dissemination (HTEMD) model for VANETs to guarantee the security and reliability of data transmission in the presence of malicious nodes that interfere with the proper judgment of vehicles through data tampering or spreading false messages. The HTEMD model integrates entity trust into data consistency evaluation, where the receiver calculates the probability of an event based on the sender's entity trust and event status and then uses information entropy to conduct consistent analysis of the results to help vehicles make correct decisions. The model also mitigates problems like channel collision and hidden nodes through a dynamic buffering and updating mechanism. In a similar work, the authors in [38] propose a multi-dimensional trust (MDT) model for misbehaviour detection in vehicular ad hoc networks (VANETs), which effectively integrates entity trust into data consistency evaluation to address the limitations of existing single trust models that are vulnerable to complex road environments, variable events, and intelligent attacks. The MDT model evaluates the trustworthiness of vehicles based on three aspects: behaviour trust, data trust, and recommendation trust, with data collection and trust calculation deployed in vehicles and the TA, respectively.

NOTRINO [40], a novel hybrid trust management scheme for Internet-of-Vehicles (IoV) that evaluates the trustworthiness of the received information in two steps. First, it classifies the message sender as trustworthy or malicious at the transport layer. Next, it evaluates the authenticity of the received messages at the application layer, with data trustworthiness being performed if the sender is classified as trustworthy at the lower layer. NOTRINO defines trust as a function of role-oriented trust, information quality, and effective distance, and it operates in a fully distributed manner to match the dispersed and distributed nature of IoV environments.

In [41], the authors propose A Trust Evaluation and Management Framework in Context-Enabled Vehicular Ad-Hoc Networks (TEAM). The TEAM framework uniquely combines asset-based threat modelling and ISO-based risk assessment to identify potential attacks and manage trust among vehicles in a decentralised manner.

The review of existing approaches reveals significant limitations in current trust management solutions for ITS. Blockchain-based approaches like DTMS [33] and AIT [32] focus narrowly on vehicle transactions and consensus mechanisms, failing to address trust relationships with pedestrians and other ITS components. While these solutions offer security through blockchain, they lack the flexibility needed for diverse ITS interactions. Similarly, peer-to-peer solutions, such as [32], while effective for vehicle-level trust, are restricted to direct participant interactions and cannot scale to ecosystem-wide trust management. Multi-agent approaches by Cheng et al. [35] and Müller

Table 1
Comparison of existing approaches across features. Annotations: "✓" indicates covered, and "X" indicates not covered.

Works	Decentralised architecture	Neural network-based	Temporal adaptation	Multi-component coverage	Attack pattern recognition	Real-time processing
[32]	✓	X	X	X	X	✓
[33]	✓	X	X	X	X	X
[34]	✓	✓	X	X	X	X
[35]	X	X	X	X	X	✓
[36]	✓	X	X	✓	X	✓
[37]	✓	X	X	X	✓	✓
[38]	✓	X	X	X	✓	✓
[40]	✓	X	X	X	✓	✓
[41]	✓	X	X	X	X	✓
Our work	✓	✓	✓	✓	✓	✓

et al. [36] rely on predetermined subjective logic rules and static trust quantification methods, making them insufficient for the dynamic nature of ITS environments. Their fixed trust calculation mechanisms cannot adapt to rapidly changing network conditions. Hybrid models like HTEMD [37] and MDT [38], despite considering multiple trust aspects, employ rigid trust calculations with predefined parameters, limiting their effectiveness in evolving ITS scenarios. More recent solutions like NOTRINO [40] and TEAM [41], while introducing layered and context-aware approaches, still suffer from inflexibility. NOTRINO's fixed two-step evaluation process and static role-oriented metrics cannot accommodate the fluid nature of ITS interactions. At the same time, TEAM's reliance on predetermined threat models and risk assessments restricts its adaptability. These limitations underscore the need for a more comprehensive and dynamic approach to trust management in ITS environments.

The existing trust frameworks for CPS exhibit several essential advantages. Blockchain-based models, such as DTMS [33] and AIT [34], provide decentralisation, immutability, and tamper-resistant record keeping, ensuring transparency and auditability of trust transactions. Reputation- and voting-based schemes, including peer-to-peer consensus architectures [32], offer lightweight operation and efficient trust aggregation without requiring central authorities. Multi-agent and subjective logic-based frameworks [35,36] contribute to interpretability by explicitly modelling uncertainty and combining direct and indirect observations, while hybrid models such as HTEMD [37] and MDT [38] combine behavioural and data trust to improve accuracy in detecting misbehaviour.

Despite these strengths, most CPS and ITS trust solutions still face notable drawbacks that limit their scalability and adaptability in dynamic environments. Blockchain frameworks, while secure, impose significant computational and communication overhead that can violate the real-time latency constraints of safety-critical ITS applications [42]. Reputation and voting-based mechanisms are inherently vulnerable to collusion, Sybil, and recommendation attacks because they depend primarily on peer feedback rather than content verification [30]. Subjective logic and fuzzy-based methods require manually tuned parameters and static belief assignments, reducing adaptability when the network topology or data distributions change rapidly [26]. Hybrid trust models, though more comprehensive, still rely on predefined feature weights and threshold values that do not capture evolving attack patterns [21,24]. These constraints highlight the need for an adaptive, learning-driven trust framework that can automatically learn complex interdependencies among features, adjust to environmental variations, and maintain accurate trust evaluations in real time.

The proposed MLT system addresses these limitations by leveraging a feedforward neural network to establish trust dynamically across loosely connected ITS components. A critical advantage of this approach is the introduction of an automated temporal decay mechanism, a capability notably absent in existing solutions. While current approaches like NOTRINO [40] and TEAM [41] attempt dynamic trust evaluation, they either maintain static trust values or require manual intervention to adjust trust scores over time. This lack of temporal adaptation is particularly problematic in ITS environments where trust relationships naturally evolve and degrade based on changing conditions and interactions. By incorporating both diverse input sources from ITS components and an automated time decay factor, MLT offers a more comprehensive, adaptive, and scalable solution to trust management in complex ITS ecosystems. This temporal awareness, combined with the system's ability to process inputs from various sources, including pedestrian-enabled IoT devices, enables MLT to maintain trust evaluations that accurately reflect the current state of the network while automatically adjusting for the passage of time. A comparative analysis of the key features of existing approaches, alongside our proposed methodology, is presented in Table 1 to highlight their individual components and capabilities.

3. System design

This section presents the considered system model in this paper, including the assumptions made about the ITS environment and the communication protocols used. The attack model is also described, outlining potential attack scenarios and the adversary's capabilities. Additionally, the section introduces the key concepts and definitions that form the foundation of the proposed trust management system.

3.1. Assumptions

In this solution, every RSU is assumed to be equipped with MLT because it has sufficient computational resources for neural network processing, maintains stationary positions ideal for continuous monitoring and trust computation, and can effectively aggregate both local and global trust through its strategic placement and direct communication with multiple ITS components and other RSUs in the network. RSUs also contain a database that stores trust scores.

3.2. System model

As shown in Fig. 1, the considered system model contains three key components to highlight ITS heterogeneity, i.e. vehicles, RSUs, and pedestrians.

1. RSU: These are stationary units with considerable computing resources deployed along roads to facilitate communication between vehicles and other road infrastructures. RSUs are access points that compute the trust of a vehicle and determine its legitimacy, ensuring seamless data exchange.
2. Vehicles $(V_1, V_2, \ldots, V_n)$: These are the primary mobile units in the ITS, equipped with various sensors and communication capabilities to interact with the environment and other ITS components. Vehicles send data on traffic conditions, vehicle dynamics, and environmental factors to RSUs and pedestrians.
3. Pedestrians: Individuals equipped with smart devices that can communicate with vehicles and infrastructure, contributing to the data pool and enhancing situational awareness for all participants in the ITS.

The ITS ecosystem encompasses a wide array of interconnected components. Beyond the three primary components focused on (vehicles, RSUs, and pedestrians), numerous IoT devices contribute to the system's complexity. These include traffic signals with adaptive control systems, environmental sensors monitoring weather conditions, parking management systems, and various in-vehicle IoT devices such

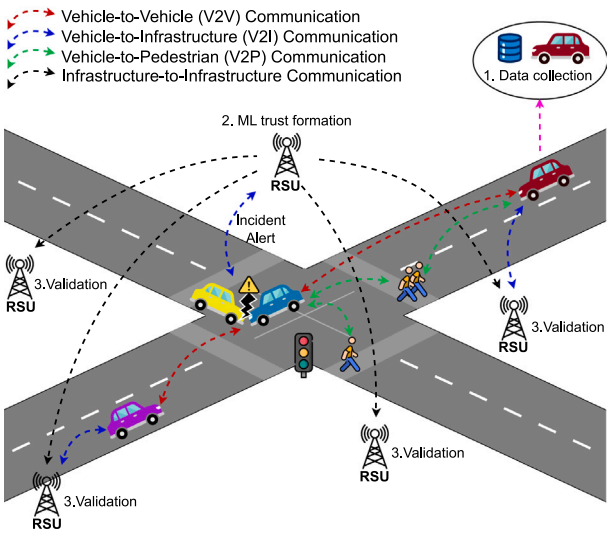

Fig. 1. An illustration of MLT's System model.

as tyre pressure monitors, engine management systems, and collision detection sensors. Each component type presents unique trust considerations: vehicle-integrated IoT devices must handle safety-critical data with minimal latency, infrastructure-based devices require long-term reliability and tamper resistance. In contrast, user-carried devices need to balance security with trust considerations. However, these diverse components can be effectively modelled through three primary categories: mobile nodes (vehicles), fixed infrastructure (RSUs), and human-carried devices (pedestrians). Vehicles represent high-mobility nodes with safety-critical requirements, RSUs serve as trusted anchors with consistent network presence, and pedestrians exemplify personal devices with varying trust requirements and security concerns. Through these components, the trust management model can be extended to handle the full spectrum of ITS devices while keeping the model implementable.

3.3. Attack model

In the proposed system, different ITS components are categorised based on their trust levels and behaviours. Trusted entities include RSUs, which are assumed to be secure and capable of performing trust evaluations without external compromise. These RSUs serve as central points for aggregating and disseminating trust scores across the ITS ecosystem. Honest-but-curious entities include vehicles and pedestrian devices that generally operate as legitimate nodes but may attempt to infer private information or manipulate trust evaluations for personal gain. Although these nodes do not engage in active attacks, they can introduce bias into the system by selectively sharing or withholding information.

Malicious nodes refer specifically to compromised vehicles and pedestrian devices that actively attempt to disrupt the ITS ecosystem. These malicious entities have the ability to falsify data, spoof identities, and execute coordinated attacks. Therefore, the following three attack scenarios are considered within the scope of this investigation. These three attacks were selected because they target critical data integrity and vulnerabilities inherent in ITS. These attacks are among the most common threats modelled by current state-of-the-art research, and have been considered in baseline approaches compared to MLT [40,43–48].

- Simple attack [49] (SA): Attackers use tactics such as black hole and DoS attacks to disrupt vehicle communications. Malicious nodes can drop the packets during transmission, leading to communication failures. They may also hinder normal nodes by repeatedly initiating communication requests, preventing them from connecting with the outside world.
- Recommendation attack [42] (RA): Malicious nodes often give false recommendations to disrupt the trust assessments of normal nodes. Recommendation attacks typically include bad-mouthing attacks, where negative evaluations lower a node's trust value, preventing its participation in normal network activities. Conversely, ballot attacks involve giving positive evaluations to colluding nodes, boosting their trust value to facilitate future malicious activities
- Zigzag attack [50] (ZA): An on–off attack, or intermittent attack, is a complex malicious tactic in which a node switches between normal and malicious behaviour to avoid detection. The malicious node builds trust at a low cost and then attacks the ITS network through false incident reports, dropping critical safety messages, or misrouting data packets. Afterwards, it quickly returns to normal behaviour, interacting regularly with legitimate nodes. This random timing, location, and target pattern make detection challenging for current trust management systems.

4. Proposed MLT

The neural network using the Levenberg–Marquardt backpropagation algorithm processes a mathematically rigorous feature space $\mathbf{X} \in \mathbb{R}^d$, where $d = 20$. These inputs encompass the location of the reporting vehicle $\mathbf{l} = (x, y, z)$, incident details $\mathbf{i}$, distance from the closest RSU r, traffic density ρ, visibility conditions υ, historical trust scores $\mathbf{h}$, and various other traffic-related metrics $\mathbf{m}$ generated through SUMO simulation similar to Zhang et al. [34].

The 20-dimensional input space was designed to comprehensively represent all critical aspects of trust evaluation in the ITS ecosystem. The input features cover spatial dimensions for vehicle location coordinates (x and y) that capture spatial context, along with event information, including incident type, severity, timing, and corroboration metrics. Infrastructure-related features cover RSU proximity, coverage, and signal quality, while the environmental metrics are represented through traffic density, visibility conditions, and weather state metrics. Historical data tracks previous interaction patterns, and reliability, complemented by network behaviour metrics for communication patterns, message consistency, and system security indicators. Table 2 describes the input features for MLT.

The network architecture implements a structured multi-layer design comprising an input layer processing the 20-dimensional feature vector, two hidden layers each containing 16 neurons (derived from the optimal capacity formula $\mathrm{h} = \left\lceil \sqrt{n \cdot m} \right\rceil$), and an output layer. The forward propagation through each layer l follows these formulae

$$\mathbf{Z}^{(l)} = \mathbf{W}^{(l)}\mathbf{H}^{(l-1)} + \mathbf{b}^{(l)} \tag{1}$$

$$\mathbf{H}^{(l)} = \phi(\mathbf{Z}^{(l)}) \tag{2}$$

where $\mathbf{W}^{(l)} \in \mathbb{R}^{n_l \times n_{l-1}}$ represents the weight matrices, $\mathbf{b}^{(l)} \in \mathbb{R}^{n_l}$ the bias vectors, and $\phi(\cdot)$ the ReLU activation function for hidden layers, chosen for its superior gradient propagation properties. The network uses weighted summation and sigmoid activation $\sigma(\cdot)$ in the output layer because its ability to output values between 0 and 1 aligns perfectly with the trust score range, providing natural normalisation.

The output layer produces two key results through the transformation

$$[TS, \lambda] = \sigma(\mathbf{W}^{(L)}\mathbf{H}^{(L-1)} + \mathbf{b}^{(L)}) \tag{3}$$

where TS represents the trust score for the message sender, bounded within $[0, 1]$, and λ determines the decay constant governing temporal trust evolution.

This study employs the Levenberg–Marquardt (LM) algorithm to address the challenges associated with traditional gradient descent

Table 2
Input features for MLT.

Category	Feature	Description
Spatial Context	X-coordinate	Vehicle position
	Y-coordinate	Vehicle position
Incident Information	Incident X-coordinate	Location of incident along East–West axis in metres
	Incident Y-coordinate	Location of incident along North–South axis in metres
	Distance to incident	Euclidean distance from vehicle to incident location in metres
Mobility	Current speed	Vehicle's instantaneous speed in m/s
	Acceleration	Rate of speed change in m/s^2
	Lane position	Vehicle's position along the current lane in metres
Network	Distance to nearest RSU	Euclidean distance to closest roadside unit in metres
	Vehicle density	Number of vehicles per kilometre in vicinity
	Edge occupancy	Percentage of road segment occupied by vehicles
Traffic Participants	Pedestrian density	Number of pedestrians in proximity to vehicle
	Pedestrian crossing activity	Number of active pedestrian crossings nearby
	Public transport presence	Number of buses/trams within communication range
Communication Metrics	Packet delivery ratio	Percentage of successfully delivered messages
	Signal strength	Network signal quality in dBm
	Communication delay	Round-trip time for message delivery in milliseconds
Historical Data	Previous interactions	Count of past communications with entity
	Reporting consistency	Correlation between reports and verified incidents
	Reporting frequency	Rate at which entity generates incident reports

methods in neural network training. These challenges include slow convergence, vulnerability to local minima, sensitivity to learning rate selection, inefficiency in plateau regions, zigzagging behaviour in narrow valleys of the error surface, and issues with vanishing or exploding gradients in deep networks [51]. The selection of LM is predicated on its proven efficiency in handling nonlinear data with superior convergence rates compared to standard techniques [52].

The LM algorithm modifies the trust model's training process by introducing an adaptive damping factor μ that adjusts iteratively based on the trust ratio ρ which is defined as

$$\rho = \frac{E(\mathbf{w}) - E(\mathbf{w} + \delta\mathbf{w})}{L(0) - L(\delta\mathbf{w})} \tag{4}$$

where $E(\mathbf{w})$ represents the actual error and $L(\delta\mathbf{w})$ the predicted error reduction. This adaptive mechanism ensures stability and reduces the risk of divergence during optimisation. The trust ratio ρ serves as a crucial performance improvement factor in the LM algorithm within MLT. It represents the relationship between actual error reduction and predicted reduction based on linear approximation, effectively measuring how well the quadratic approximation models the error surface. When ρ approaches 1, it indicates the linear approximation accurately predicts error reduction, suggesting the algorithm operates in a well-behaved region. This allows for reducing the damping factor μ, enabling larger steps similar to Newton's method, accelerating convergence [53]. Conversely, when ρ is small or negative, it signals poor linear approximation, necessitating an increase in μ for smaller, gradient-descent-like steps that maintain stability. This adaptive mechanism improves MLT's performance by accelerating convergence in well-behaved regions, ensuring stability in complex trust landscapes, and automatically adjusting to different data characteristics without manual tuning [54]. The trust ratio enables MLT to efficiently process heterogeneous data streams typical in ITS environments, adapting its optimisation strategy to local error surface characteristics for more accurate trust evaluations and better generalisation compared to fixed-step methods.

Each input layer of the neural network is configured to receive specific traffic-related metrics, which are then weighted and processed through multiple hidden layers. The complete forward propagation process for an input pattern p follows

$$\begin{aligned} \mathbf{H}_p^{(0)} &= \mathbf{X}_p \\ \mathbf{Z}_p^{(l)} &= \mathbf{W}^{(l)}\mathbf{H}_p^{(l-1)} + \mathbf{b}^{(l)} \\ \mathbf{H}_p^{(l)} &= \phi(\mathbf{Z}_p^{(l)}) \\ [TS_p, \lambda_p] &= \sigma(\mathbf{W}^{(L)}\mathbf{H}_p^{(L-1)} + \mathbf{b}^{(L)}) \end{aligned} \tag{5}$$

The LM algorithm refines the network weights using the following update rule

$$\mathbf{W}_{t+1} = \mathbf{W}_t - [\mathbf{J}_W^T(t)\mathbf{J}_W(t) + \mu\mathbf{I}]^{-1}\mathbf{J}_W^T(t)\mathbf{e}_p(t) \tag{6}$$

where $\mathbf{W}_t$ and $\mathbf{W}_{t+1}$ are the weight matrices at consecutive training epochs, representing the state of the neural network at different points in the training process. $\mathbf{J}_W(t)$ is the Jacobian matrix of derivatives of the network errors with respect to the weights at time t, computed as

$$\mathbf{J}_W(t) = \begin{bmatrix} \frac{\partial e_1}{\partial w_1} & \frac{\partial e_1}{\partial w_2} & \cdots & \frac{\partial e_1}{\partial w_N} \\ \frac{\partial e_2}{\partial w_1} & \frac{\partial e_2}{\partial w_2} & \cdots & \frac{\partial e_2}{\partial w_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_M}{\partial w_1} & \frac{\partial e_M}{\partial w_2} & \cdots & \frac{\partial e_M}{\partial w_N} \end{bmatrix} \tag{7}$$

The damping factor μ adjusts according to

$$\mu_{t+1} = \begin{cases} \mu_t/\beta & \text{if } \rho > 0.75 \\ \mu_t\beta & \text{if } \rho < 0.25 \\ \mu_t & \text{otherwise} \end{cases} \tag{8}$$

where $\beta = 10$ (empirically determined). $\mathbf{I}$ is the identity matrix, which maintains dimensional consistency during weight adjustments. Finally, $\mathbf{e}_p(t)$ is the error vector at epoch t for pattern p, representing the difference between predicted and actual outputs

$$\mathbf{e}_p(t) = \mathbf{y}_p - \hat{\mathbf{y}}_p(t) \tag{9}$$

This comprehensive error-driven learning process ensures optimal weight adaptation while maintaining stability.

4.1. Dynamic trust score computation

Vehicles and pedestrians perform trust calculations when they receive updates from peer components. At the neural network's output layer, a trust score is computed for each component within the ITS. These scores are normalised between 0 and 1, representing the reliability of each component. The complete trust computation follows a multi-stage process incorporating both immediate evaluation and temporal dynamics.

The initial trust score for component i is computed through the neural network as follows

$$TS_i = \sigma\left(\sum_{j=1}^{n} w_j \cdot x_{ij} + b\right) \tag{10}$$

where TS_i represents the initial trust score of the ith component, establishing its baseline trustworthiness. The sigmoid activation function σ introduces essential non-linearity into the neural network and maps the output to a range between 0 and 1. The input features x_{ij} represent various metrics and data points fed into the network, weighted by w_j and bias b, which are continuously adjusted during training to minimise error.

As defined in Section 1, trust in MLT encompasses both reliability and credibility, which the neural network architecture explicitly captures through the 20-dimensional input feature space detailed in Table 2. Reliability assessment is achieved through features representing

the technical operational consistency of each component, including mobility metrics such as current speed, acceleration, and lane position, network performance indicators such as distance to nearest RSU, vehicle density, and edge occupancy, and communication quality parameters such as packet delivery ratio, signal strength, and communication delay. Conversely, credibility assessment relies on features that validate the authenticity of information, such as incident verification metrics like distance to incident and incident coordinates cross-referenced with vehicle position, reporting consistency, such as correlation between reports and verified incidents, and historical reliability, including previous interactions and reporting frequency. The unified trust score generated by the output layer, as defined in Eq. (10), is optimised using the Levenberg–Marquardt algorithm to balance the influence of both reliability and credibility-related features based on training data. This optimisation ensures that the resulting trust score reflects not only the technical dependability of component behaviour but also the informational integrity of its communications, fulfilling the security requirements of the ITS ecosystem.

To establish a time decay factor that captures the diminishing relevance of trust assessments over time, an exponential decay model is adopted, expressed by

$$D(t) = D_0 \times e^{-\lambda t} \tag{11}$$

where $D(t)$ represents the value at time t, D_0 is the initial value, and λ is the decay constant. This model effectively captures the diminishing influence of older ratings over time in the trust model, with λ bounded within empirically determined limits $[\lambda_{min}, \lambda_{max}] = [0.001, 0.1]$. The decay constant λ plays a crucial role in the temporal evolution of trust within the MLT system, determining how quickly trust scores diminish without reinforcing interactions. In dynamic ITS environments where conditions change rapidly, this temporal dimension maintains accurate trust assessments. Smaller λ values result in slower decay, appropriate for stable environments with consistent interactions, while larger values produce more rapid decay, suitable for highly dynamic scenarios where information quickly becomes obsolete. Rather than using a fixed value, MLT dynamically computes λ as part of its output, adapting the decay rate based on information criticality, environmental volatility, historical entity behaviour stability, and network interaction patterns. This adaptive approach enables MLT to maintain relevant trust assessments across diverse scenarios, from dense urban traffic to sparse rural environments with intermittent communications. The exponential decay function ensures that trust gradually decreases, mirroring real-world trust degradation patterns.

The exponential decay model is integrated into the feedforward neural network's output layer through an innovative approach where the network is designed to predict the decay constant λ as part of its output. This allows the network to learn and adjust λ dynamically based on the input data, thereby effectively predicting how quickly the trustworthiness of a message or a component should decay over time. The complete output formulation is

$$Output\ O = [TS, \lambda] = \sigma(\mathbf{W}^{(L)}\mathbf{H}^{(L-1)} + \mathbf{b}^{(L)}) \tag{12}$$

where TS represents the trust score and λ the decay constant. The application of the decay function yields the time-adjusted trust score

$$TS_{adjusted}(t) = TS \times e^{-\lambda t} \tag{13}$$

where $TS_{adjusted}(t)$ is the time-adjusted trust score at time t. Through this mechanism, the network learns to predict the appropriate λ value for each interaction, enabling the system to dynamically adjust the influence of trust scores based on their temporal relevance.

The trust score TS of a component is dynamically adjusted using this time decay factor, reflecting the decreasing reliability of data over time. When the initial TS exceeds a predefined threshold τ, determined through statistical analysis of historical trust distributions

$$\tau = \mu_{hist} + \alpha\sigma_{hist} \tag{14}$$

where μ_{hist} is the historical mean trust score, σ_{hist} is the standard deviation, and $\alpha = 1.96$ for 95% confidence, the component is considered trustworthy. However, to ensure the trust assessments remain current and accurate, the TS is subjected to a time-decay adjustment, as shown in Eq. (11).

After vehicles and pedestrians perform these trust calculations, they transmit a trust evaluation tuple to the nearest RSU

$$T_{eval} = (i, TS_{adjusted}, \lambda, t_{stamp}) \tag{15}$$

where i is the component ID, $TS_{adjusted}$ is the adjusted trust score, λ is the predicted decay constant, and t_{stamp} is the calculation timestamp. This allows the RSU to maintain a current trust profile P_i for each component i in its coverage area

$$P_i = \{T^1_{eval}, T^2_{eval}, \dots, T^k_{eval}\} \tag{16}$$

where k represents the number of recent evaluations maintained for each component.

The complete trust evaluation and management process is formalised in Algorithm 1, which integrates all components of the mathematical framework. The algorithm's implementation includes several critical supporting functions that handle specific aspects of trust computation

$$ConfidenceScore(TS_i) = (1 - \frac{\sigma^2_{hist}}{\mu_{hist}}) \cdot (1 - \frac{|\Delta TS|}{TS_{max}}) \tag{17}$$

where σ^2_{hist} represents the historical variance of trust scores, μ_{hist} their mean, and Δ TS the change from the previous score.

The RSU maintains a temporal trust profile matrix **P** for all components in its coverage area

$$\mathbf{P} = \begin{bmatrix} TS_{11}(t) & TS_{12}(t) & \cdots & TS_{1n}(t) \\ TS_{21}(t) & TS_{22}(t) & \cdots & TS_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ TS_{m1}(t) & TS_{m2}(t) & \cdots & TS_{mn}(t) \end{bmatrix} \tag{18}$$

where $TS_{ij}(t)$ represents the trust score of a component j as evaluated by the component i at time t.

The trust propagation through the network follows a weighted aggregation model

$$TS_{global}(i) = \frac{\sum_{j=1}^{m} w_j \cdot TS_{ji}(t)}{\sum_{j=1}^{m} w_j} \tag{19}$$

Algorithm 1 MLT algorithm.

Require:
$\mathbf{X} \leftarrow$ Input features matrix from ITS components
$\mathbf{W}(t), \mathbf{b} \leftarrow$ Weight matrix and bias at time t
$\lambda \leftarrow$ Decay constant for trust score
$Threshold \leftarrow$ Minimum required trust score
Ensure:
$\mathbf{T} \leftarrow$ Trust status for each component
for each component c in ITS **do**
  Calculate initial trust scores:
  $TS = sigmoid(\mathbf{W}(t) \cdot \mathbf{X}_c + \mathbf{b})$
  if $TS \geq Threshold$ **then**
    Apply decay to trust score:
    $TS_{adjusted} = TS \times e^{-\lambda \cdot \Delta t}$
    if $TS_{adjusted} \geq Threshold$ **then**
      $\mathbf{T}(c) \leftarrow$ 'Trusted'
    else
      $\mathbf{T}(c) \leftarrow$ 'Not Trusted'
    end if
  else
    $\mathbf{T}(c) \leftarrow$ 'Not Trusted'
  end if
end for

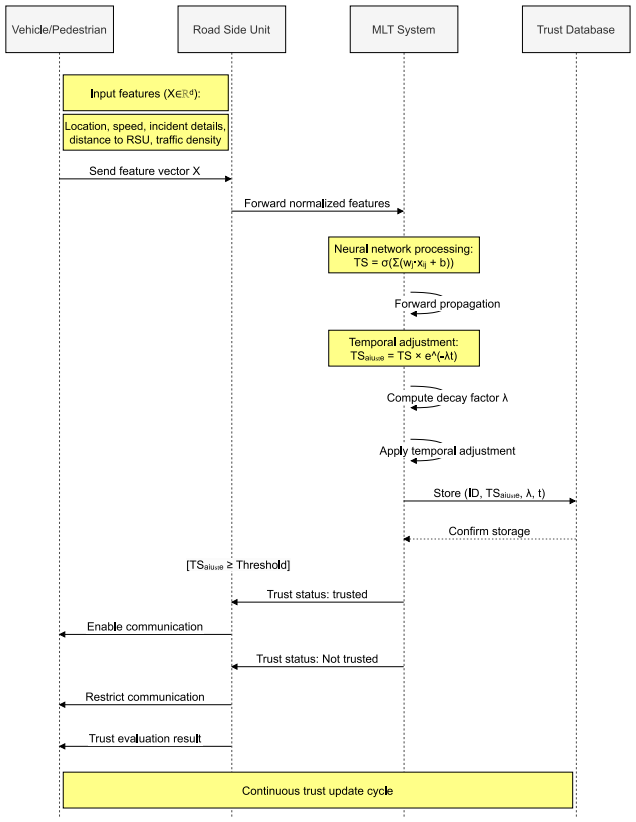

Fig. 2. Trust evaluation stages in MLT.

where w_j represents the weight assigned to each evaluating component based on its trustworthiness.

For anomaly detection, the system employs a statistical threshold based on the z-score of trust variations

$$z_i(t) = \frac{TS_i(t) - \mu_i}{\sigma_i} \tag{20}$$

A component is flagged for potentially malicious behaviour if

$$|z_i(t)| > \beta \tag{21}$$

where $\beta = 3$ represents a three-sigma confidence interval for anomaly detection.

The process sequence, as illustrated in Fig. 2, demonstrates how these mathematical components work together to provide a robust and adaptive trust management system. This framework ensures that trust evaluations remain accurate and relevant while adapting to the dynamic nature of ITS environments.

5. Experimental analysis

This section presents the simulation setup and performance analysis.

5.1. Simulation setting

Table 3 details the simulation parameters used to evaluate MLT. The proposed solution was validated on a Manchester City Centre, United Kingdom map generated from OpenStreetMap as shown in Fig. 3. Four RSUs were deployed at randomly selected locations on the map and experiments were conducted in a 2 km by 2 km area with 100 legitimate vehicles and 50 pedestrians, testing the system's performance against varying the percentage of malicious nodes. The simulation ran for 2000 s using IEEE 802.11p for wireless communication. The network environment was simulated using OMNET++ 6.0.2 [55] with VEINS 5.2

Table 3
A summary of the simulation configuration details.

Simulation parameter	Value
Simulation Area	2km by 2km
No. of legitimate vehicles	100
No of Pedestrian	50
% of malicious nodes	5, 10, 15, 20, 25, 30, 35, 40
Simulation time	2000 secs
Mac protocol	IEEE 802.11p
Network simulator	OMNET++ 6.0.2
Traffic simulation	SUMO 1.19.0
Vehicular network simulator	VEINS 5.2
AI Hyperparameter	70% training, 15% Validation, 15% testing
Neural network architecture	20 input layer, 8 input layer X2, 2 output layer

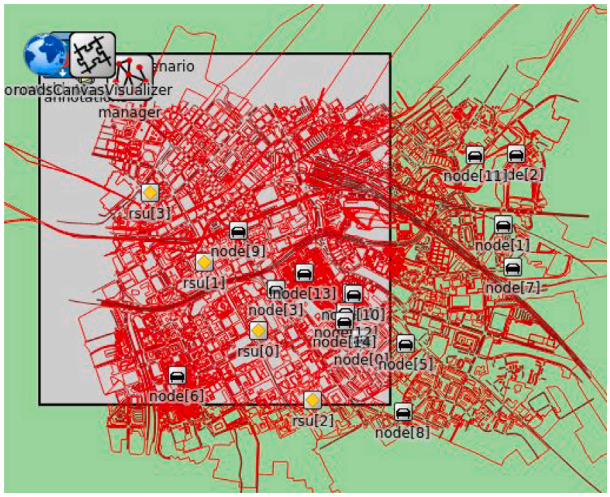

Fig. 3. Sumo simulation OpenStreetMap for Manchester City centre.

for vehicular network simulation and SUMO 1.19.0 [56] for the traffic modelling.

The neural network was implemented using Jupyter Notebook as the development environment and Scikit-learn, a widely used open-source library in Python, for the feedforward neural network with the LM algorithm and data preprocessing tasks. The data generated from the SUMO simulations was split into training and testing sets using a $70\% - 30\%$ ratio, where the training set was used to train the neural network, allowing it to learn the patterns and relationships between the input features.

5.2. Evaluation and results

The model is validated through simulations designed to emulate the three adversarial scenarios explained in Section 3.3. These scenarios test the models' ability to accurately identify genuine and adversarial inputs. A safety-related event, such as an accident or collision, is randomly generated within the network simulation through a V2V, V2P and V2I communication for the nodes to evaluate their genuineness, see Fig. 1. The key performance metrics used in this evaluation include Precision, Recall, and the F-measure, defined as follows:

- Precision (P): This metric quantifies the accuracy of positive predictions made by the trust model. It is defined as the ratio of true positive results to the total number of instances classified as positive

$$P = \frac{TP}{TP + FP} \tag{22}$$

where TP represents true positives and FP denotes false positives.

- Recall (R): Also known as sensitivity, this metric measures the model's ability to correctly identify actual positives from the data

$$R = \frac{TP}{TP + FN} \tag{23}$$

where FN represents false negatives.

- F1-Score: The F-measure is the harmonic mean of Precision and Recall, providing a balance between these metrics for cases where an equilibrium between precision and recall is required

$$F1 = 2 \cdot \frac{P \cdot R}{P + R} \tag{24}$$

In addition to the above metrics, the peculiarity of the proposed solution requires performance testing of the time decay factor. Therefore, the following metrics will analyse MLT's adaptation to the decay factor:

- Mean Response Time: This metric represents the average time taken by MLT to adjust trust levels in response to environmental changes, such as variations in traffic density or visibility conditions. A lower response time indicates that MLT can quickly incorporate new information and update trust scores, which is essential for real-time applications in ITS.
- Adaptation Accuracy: Adaptation accuracy measures the degree to which the MLT system's trust adjustments align with the theoretically optimal trust levels based on environmental changes. High adaptation accuracy ensures that MLT's trust scores accurately reflect real-world conditions, maintaining reliability even as those conditions fluctuate.
- Recovery Rate: Recovery rate measures how quickly MLT can return to optimal trust levels after a sudden change or disruption. This metric is important for assessing MLT's resilience in dynamic environments where disruptions might occur due to malicious behaviour.

The selection of evaluation metrics followed a structured process grounded in both theoretical relevance and empirical evidence from prior ITS and CPS trust management studies. A literature survey of recent trust frameworks [29,34,36,37] identified the indicators most commonly used to assess the reliability, robustness, and adaptability of the model in adversarial conditions. From this review, Precision, Recall, and F1-score were chosen as primary classification metrics because they directly quantify a model's ability to correctly detect malicious or trustworthy nodes, aligning with the goal of minimising false trust decisions in safety-critical environments. To evaluate the temporal responsiveness of the proposed MLT, complementary adaptation-oriented metrics such as Mean Response Time, Adaptation Accuracy, and Recovery Rate were selected. These quantify how effectively and quickly the model adjusts trust scores following environmental or behavioural changes, which is crucial for real-time vehicular operations. The inclusion of both classification- and adaptation-based metrics ensures a comprehensive evaluation that reflects not only the detection accuracy but also the dynamic performance of the trust model in continuously evolving ITS environments.

The performance of MLT was compared with three baseline approaches for trust management systems, which have been used in a number of research studies.

- ART [43]: In the ART scheme, vehicles gather and examine traffic data from networks to evaluate the reliability of information and nodes. The scheme employs evidence combination and collaborative filtering algorithms in trust calculations, which helps minimise the impact of misleading data and malicious attacks. However, the ART scheme's limitation lies in its independent assessment of trust attributes without considering their interrelationships.
- Weighted voting [44–48]: The weighted voting approach is a recognised method in trust management for recommender systems, as it mitigates the impact of false recommendations by considering inputs from neighbouring nodes. However, this method is vulnerable when there is a high proportion of malicious nodes among the neighbours or when malicious nodes engage in on–off attacks.
- NOTRINO [40]: introduces a novel hybrid trust model based on the IoT protocol stack, differing from previous models typically deployed at the application layer in VANET. It splits trust evaluation into two stages: assessing the sender's trustworthiness at the transport layer and verifying message authenticity at the application layer. This approach enhances flexibility and efficiency. However, its reliance on role-based trust evaluation without node cooperation makes it vulnerable to combined attacks [38].

It is important to note that the comparison in this work focuses on these established baseline approaches for several reasons. Firstly, these methods have been widely cited and influential in trust management for vehicular networks, providing a solid foundation for comparison. While more recent approaches exist, these selected baselines represent key concepts and methodologies that continue to inform current research.

The analysis in this work is based on a thorough review of the published algorithms and reported results of these approaches. These approaches were carefully modelled within the simulation environment, striving to accurately represent their core mechanisms and performance characteristics as described in their respective publications.

5.2.1. Robustness against attacks

To evaluate the robustness of the proposed model against the three baseline approaches, experiments were carried out with different adversarial models explained in Section 3.3. The results of the experiment are shown in Figs. 4–6. The simulation environment consists of 100 legitimate vehicles and 50 pedestrians as participants in the core network. The percentage of malicious nodes is calculated relative to the total legitimate node count of 150. Therefore, when 40% malicious nodes is indicated, this means that 60 malicious nodes of even distribution of vehicles and pedestrians are added to the network, creating a complete network of 210 nodes (100 legitimate vehicles, 50 pedestrians and 60 malicious nodes). These malicious nodes target both vehicles and pedestrians, attempting to disrupt critical safety communications between all network participants.

Fig. 4 compares MLT and the three considered baseline models against simple attacks in which the adversarial entities in the network disrupt both V2P and V2V communication through DoS, including black hole attacks involving 150. This attack is relatively straightforward to achieve. Figs. 4(a) and (b) show the precision and recall of MLT and other baseline approaches when the percentage of maliciousness in the network varies from 5% to 40%. The results show that as the percentage of malicious nodes increases, the precision and recall for all models generally decrease, reflecting the challenge posed by simple attacks. For precision, At 30% of malicious nodes, MLT performance shows significant resilience against NOTRINO and ART, showing a consistent trend in accurately identifying simple attacks. However, MLT ensures an improved accuracy of 89.03% at 40% above other baseline approaches owing to the application of a neural network and the ability to learn large complex data points in the vehicular network. MLT generally achieves high accuracy in handling and detecting simple attacks, as shown by F-measure in Fig. 4(c). MLT achieves an accuracy of 89.43% at 40% of malicious nodes.

Fig. 5 illustrates the accuracy of MLT against recommendation attacks whereby an adversary can interfere with other nodes' trustworthiness by sharing a false or defamed recommendation about a legitimate node. Figs. 5(a) and (b) show MLT demonstrates robust accuracy, maintaining 97.42% at 5% malicious nodes in terms of precision and dropping slightly to 90% at 40%, outperforming other

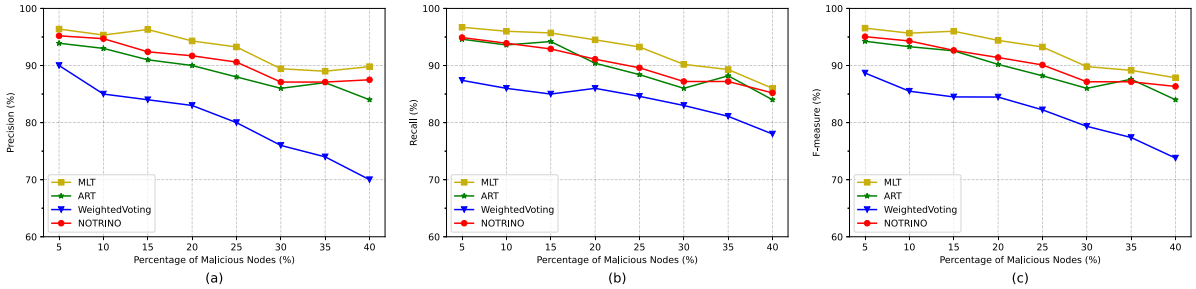

Fig. 4. Accuracy of the trust model under Simple attack. (a) Precision (b) Recall (c) F-measure.

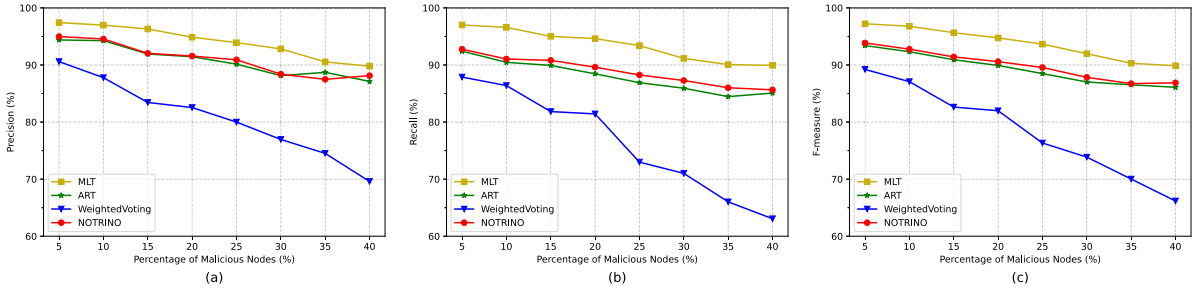

Fig. 5. Accuracy of the trust model under Recommendation attack. (a) Precision (b) Recall (c) F-measure.

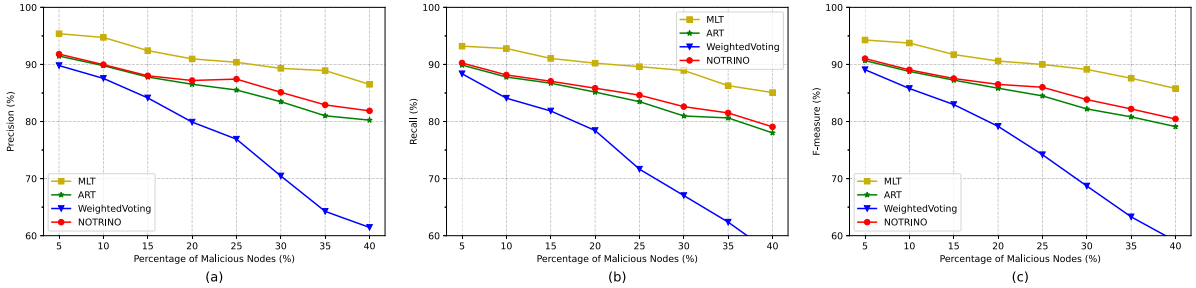

Fig. 6. Accuracy of the trust model under Zigzag attack. (a) Precision (b) Recall (c) F-measure.

approaches which decline more steeply. For recall, MLT effectively identifies relevant messages, starting at 95% and decreasing to 90% as threats increase, showcasing its reliability. The F-measure in Fig. 5(c) shows MLT excelling with values starting at 97.2% and holding at 89.8%, indicating a strong balance between precision and recall, even as the percentage of malicious nodes rises.

Similarly, Fig. 6 shows the performance of MLT against zigzag attacks. The attacks take the form of an on-and-off switch between a legitimate node and a malicious one, making it very difficult to detect. In MLT, the dynamically assigned trust decay factor and the robust ML allow the solution to adjust trustworthiness according to past behaviour and the time of the last trust value. When the trust value of the illegitimate node goes below the threshold, the node is isolated from participating in the vehicle network. In Figs. 6(a) and b, MLT maintains high precision and recall, starting at 94% with 5% malicious nodes and decreasing to 87% at 40%. The F-measure, shown in Fig. 6(c) combining precision and recall, similarly starts at 93.7% and holds around 85.7%, highlighting MLT's resilience and effectiveness against zigzag attacks.

5.2.2. *Environmental adaptability*

Here is a revised version for clarity: In Fig. 7(a), as the number of nodes increases, the precision of all models generally improves. MLT achieves approximately 91.6% precision for 50 nodes, which increases to 95% for 100 nodes, and further to 96.9% for 200 vehicles, indicating its scalability and robustness. Similarly, in Fig. 7(b), MLT demonstrates solid performance with around 93.1% for 50 vehicles, improving to about 94% for 100 vehicles and maintaining around 96% for 200 vehicles, showing consistency across varying component sizes. The F-measure results reflect a similar trend, with MLT starting at 92.1% for 50 vehicles, improving to 92.9% for 100 vehicles, and slightly increasing to 96% for 200 vehicles, confirming its effectiveness in balancing precision and recall in larger networks as shown in Fig. 7(c).

MLT consistently outperforms the other approaches across all speed categories and metrics. This superiority is particularly pronounced at higher vehicle speeds, indicating that MLT is more robust and adaptable to challenging, fast-moving scenarios in intelligent transportation systems. At low speeds (5 m/s), the performance gap between MLT and other methods is relatively narrow, suggesting that simpler approaches may be sufficient for slower-moving traffic. However, as speeds increase to medium (10 m/s) and high (20 m/s), MLT's advantage becomes more apparent. This trend is observed across all three performance metrics, highlighting the model's balanced improvement in both precision and recall as shown in Fig. 8(a) and (b). Interestingly, the performance of all models tends to decrease as speed increases, which is logical given the increased complexity and reduced reaction time in high-speed scenarios. However, MLT exhibits the least degradation in performance, maintaining a high level of accuracy even at 20 m/s. This resilience to speed variations is crucial in real-world applications where traffic conditions can change rapidly. The F-measure results in Fig. 8(c) confirm MLT's superiority across all speeds, showcasing its balanced effectiveness in precision and recall for diverse traffic scenarios.

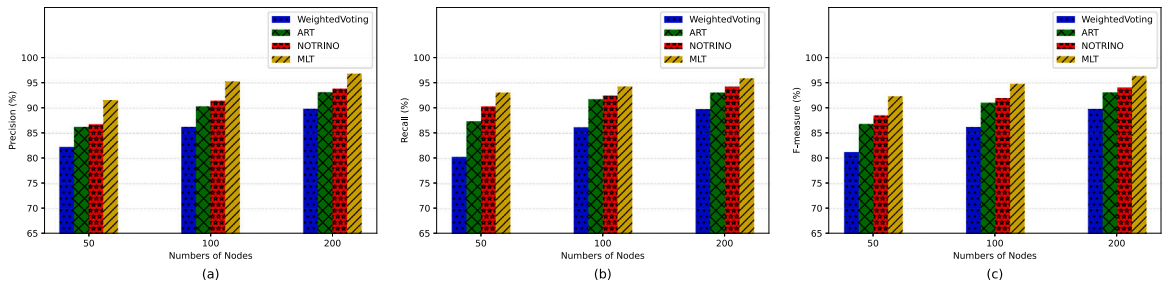

Fig. 7. Performance comparison of MLT versus baseline approaches with different numbers of vehicles. (a) Precision (b) Recall (c) F-measure.

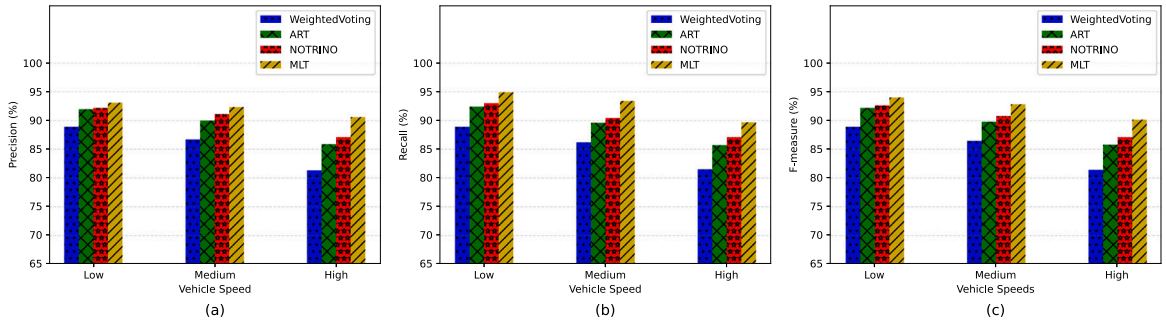

Fig. 8. Performance comparison of MLT versus baseline approaches with different travelling speeds of vehicles where low = 5 m/s, medium = 10 m/s, high = 20 m/s. (a) Precision (b) Recall (c) F-measure.

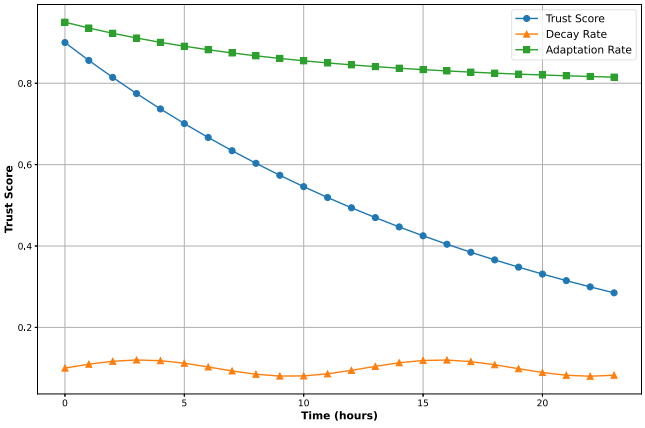

Fig. 9. Time decay performance.

Table 4
Adaptation response metrics.

Adaptation performance metric	Result
Mean Response Time	98.3ms ± 5.2ms
Adaptation Accuracy	96.7%
Recovery Rate	94.5%

Table 5
Temporal stability metrics.

Stability measure	Time window	Variance (σ^2)
Short-term Stability	1-hour	0.0025
Medium-term Stability	12-hour	0.0042

Table 6
Trust value retention over time.

Performance indicator	1-hour	6-hour	24-hour
Trust Value Retention	98.2%	92.7%	85.4%

5.3. Temporal performance analysis

The temporal evolution graph demonstrates how trust scores naturally degrade over time unless reinforced by positive interactions, with the decay rate dynamically adjusting based on environmental conditions and interaction patterns. As shown in Fig. 9, trust scores maintain stability during network perturbations while preserving the system's ability to adapt to changing conditions. This adaptive decay ensures that trust assessments remain relevant to current conditions while maintaining a historical perspective.

The temporal performance of MLT is particularly notable during trust level transitions, where the system achieves smooth adaptations without compromising security. This adaptability is facilitated by the *adaptation function* $A(t)$, defined in Eq. (12), which enables MLT to respond to environmental changes quickly. With a mean response time of 98.3 ms and an adaptation accuracy of 96.7%, MLT demonstrates the ability to adjust trust scores in real-time, ensuring that trust decisions remain both timely and secure, as shown in Table 4.

During sudden environmental changes, such as fluctuations in *traffic density* or *visibility conditions*, MLT maintains stable trust assessments by dynamically adjusting sensitivity. The **temporal stability metrics**, calculated across different time windows, further support this: variances of 0.0025 for short-term (1-hour) and 0.0042 for medium-term (12-hour) indicate the system retains a consistent trust evaluation while adapting to immediate shifts as enumerated in Table 5. These stability metrics show that MLT provides reliable and coherent trust scores even in dynamically changing conditions.

Additionally, the trust retention rates of 98.2% for a 1-hour window and 92.7% over 6 h, as presented in Table 6, illustrate MLT's capability to retain accurate trust levels over extended periods, mirroring realistic trust ageing in vehicular networks. This retention further demonstrates that MLT is well-suited to adapt and maintain accuracy over time in real-world ITS environments.

These results demonstrate MLT's capability to maintain reliable trust assessment over extended operational periods while remaining

responsive to dynamic changes in the ITS environment. The temporal performance of the system validates its suitability for real-world deployment, where maintaining accurate trust assessments over time is crucial for system security and reliability. The proposed MLT framework is designed to support real-time operation within intelligent transportation environments that generate heterogeneous and high-volume data streams. As shown in Table 2, the 20-dimensional input space employs feature aggregation, where raw data sources such as vehicle telemetry, incident reports, network metrics, and historical records are transformed into compact statistical representations. This approach substantially reduces the computational overhead associated with processing continuous ITS data flows. The feedforward neural network, composed of two hidden layers with sixteen neurons each, maintains a lightweight structure that enables efficient inference on resource-constrained RSU hardware. Furthermore, the distributed processing design allows each RSU to manage only the vehicles within its local coverage area, thereby avoiding centralised bottlenecks and promoting scalability. Trust score updates are performed on an event-driven basis rather than at fixed intervals, ensuring that computational resources are used adaptively and only when significant behavioural or contextual changes occur.

6. Conclusion

This paper presents the first trust management model for ITS that leverages the pattern recognition capabilities of neural networks to identify complex relationships between trust indicators across heterogeneous components with diverse requirements and features. The MLT framework introduces a temporal adaptation mechanism that enables trust relationships to evolve naturally over time. This mechanism addresses a critical gap in existing approaches that rely on static or manually updated trust calculations. By integrating vehicles, pedestrian devices, and RSUs into a unified trust ecosystem, MLT overcomes the fundamental challenge of reconciling components with vastly different mobility patterns, communication capabilities, and operational characteristics. The model mathematical framework provides both theoretical guarantees and practical performance advantages in dynamic ITS environments.

Our extensive simulation studies reveal that MLT consistently outperforms baseline models in various attack scenarios, including simple, recommendation, and zigzag attacks. This superior performance comes from several key architectural advantages. First, unlike traditional approaches that rely on fixed rules or predefined thresholds, MLT uses the pattern recognition capabilities of neural networks to identify complex relationships between trust indicators across different ITS components. This enables a more accurate distinction between legitimate and malicious behaviours, particularly for sophisticated attack patterns such as zigzag attacks that intentionally attempt to evade detection by alternating between benign and malicious behaviours.

Second, by leveraging a feedforward neural network alongside an exponential decay mechanism, the system ensures that trust evaluations remain current by emphasising recent interactions while gradually diminishing the influence of outdated data. This dynamic updating is crucial in an environment where vehicles are constantly in motion, RSUs serve as static but critical anchors, and pedestrian devices introduce additional mobility patterns and communication constraints. The temporal adaptation mechanism proves particularly valuable in scenarios where trust relationships must be established rapidly between previously uncounted nodes, a common occurrence in urban transportation environments.

Key limitations must be considered when interpreting these results. The computational requirements of neural network processing may present challenges for resource-constrained devices, though this is mitigated by implementing MLT primarily at the RSU level. Additionally, the current implementation assumes reliable communication between components, which may not always be guaranteed in real-world scenarios with signal interference or coverage gaps. Moreover, the performance evaluation of MLT was limited to three established baseline approaches. Although these models are commonly used and influential in trust management for vehicular networks, broader comparisons with more diverse techniques are necessary to further validate the effectiveness and generalisability of MLT. Future work will build on MLT's effectiveness to propose an enhanced authentication protocol for the ITS ecosystem that improves communication and data exchange across diverse ITS components.

CRediT authorship contribution statement

Ahmed Danladi Abdullahi: Writing – original draft, Investigation, Formal analysis, Conceptualization. **Erfan Bahrami:** Writing – review & editing, Validation, Formal analysis, Data curation. **Tooska Dargahi:** Writing – review & editing, Validation, Formal analysis, Conceptualization. **Mohammed Al-Khalidi:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Mohammad Hammoudeh:** Writing – original draft, Supervision, Investigation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

References

[1] D. Hahn, A. Munir, V. Behzadan, Security and privacy issues in intelligent transportation systems: Classification and challenges, IEEE Intell. Transp. Syst. Mag. 13 (1) (2021) 181–196.
[2] F. Nawshin, D. Unal, M. Hammoudeh, P.N. Suganthan, AI-powered malware detection with differential privacy for zero trust security in Internet of Things networks, Ad Hoc Netw. 161 (2024) 103523.
[3] F. Wang, Y. Li, X. Zhang, A survey of data poisoning attacks and defenses in intelligent transportation systems, 2024, arXiv preprint.
[4] S. Usha, R. Rajeswari, R. Prasanna, An adaptive neuro-fuzzy inference system for DDoS attack detection in intelligent transportation systems, Sci. Rep. 15 (1) (2025) 10215.
[5] Y. Zhou, H. Liu, Deceptive information attacks and resilience modeling in intelligent transportation systems, 2024, arXiv preprint.
[6] J. Li, M. Zhang, H. Sun, Adversarial manipulation of adaptive cruise control in connected autonomous vehicles, 2024, arXiv preprint.
[7] O. Adeboye, A.D. Abdullahi, T. Dargahi, M. Babaie, M. Saraee, LIFT the AV: Location InFerence aTtack on autonomous vehicle camera data, in: Proceedings of the 2023 IEEE 20th Consumer Communications & Networking Conference, CCNC, IEEE, Las Vegas, NV, USA, 2023, pp. 1–6.
[8] Y. Kim, S.-G. Sohn, K.T. Kim, H.S. Jeon, S.-M. Lee, Y. Lee, J. Kim, Exploring effective zero trust architecture for defense cybersecurity: A study, KSII Trans. Internet Inf. Syst. 18 (9) (2024).
[9] A. Pigola, F. de Souza Meirelles, Unraveling trust management in cybersecurity: Insights from a systematic literature review, Inf. Technol. Manag. (2024) 1–23.
[10] A. Mohammed, Building trust in driverless technology: Overcoming cybersecurity challenges, Aitoz Multidiscip. Rev. 2 (1) (2023) 26–34.
[11] V. Schlatt, T. Guggenberger, J. Schmid, N. Urbach, Attacking the trust machine: Developing an information systems research agenda for blockchain cybersecurity, Int. J. Inf. Manage. 68 (2023) 102470.
[12] R. Searle, K. Renaud, Trust and vulnerability in the cybersecurity context, in: Hawaii International Conference on System Science 2023, 2023.
[13] W. Najib, S. Sulistyo, Widyawan, Trust based security model in IoT ecosystem, in: Proceeding - 6th International Conference on Information Technology, Information Systems and Electrical Engineering: Applying Data Sciences and Artificial Intelligence Technologies for Environmental Sustainability, ICITISEE 2022, Institute of Electrical and Electronics Engineers Inc, 2022, pp. 195–199.
[14] K. Hasan, A. Overall, K. Ansari, G. Ramachandran, R. Jurdak, Security, privacy, and trust of emerging intelligent transportation: Cognitive Internet of Vehicles, in: Next-Generation Enterprise Security and Governance, 2022, pp. 193–226.

[15] D. Osorio, Towards 6G-enabled Internet of Vehicles: Security and privacy, IEEE Open J. Commun. Soc. 3 (2022) 82–105.
[16] M. Adam, M. Hammoudeh, R. Alrawashdeh, B. Alsulaimy, A survey on security, privacy, trust, and architectural challenges in IoT systems, IEEE Access (2024).
[17] E.Y. Imana, F.M. Ham, W. Allen, R. Ford, Proactive reputation-based defense for MANETs using radial basis function neural networks, in: The 2010 International Joint Conference on Neural Networks, IJCNN, 2010, pp. 1–6.
[18] Y. Trofimova, A.M. Moucha, P. Tvrdik, Application of neural networks for decision making and evaluation of trust in ad-hoc networks, in: 2017 13th International Wireless Communications and Mobile Computing Conference, IWCMC, 2017, pp. 371–377.
[19] G. Han, Y. He, J. Jiang, N. Wang, M. Guizani, J.A. Ansere, A synergetic trust model based on SVM in underwater acoustic sensor networks, IEEE Trans. Veh. Technol. 68 (11) (2019) 11239–11247.
[20] H. El-Sayed, H.A. Ignatious, P. Kulkarni, S. Bouktif, Machine learning based trust management framework for vehicular networks, Veh. Commun. 25 (2020) 100256.
[21] J. Wang, X. Jing, Z. Yan, Y. Fu, W. Pedrycz, L.T. Yang, A survey on trust evaluation based on machine learning, ACM Comput. Surv. 53 (2020).
[22] S.A. Siddiqui, A. Mahmood, Q.Z. Sheng, H. Suzuki, W. Ni, A survey of trust management in the Internet of Vehicles, Electronics 10 (18) (2021).
[23] R. Eldan, O. Shamir, The power of depth for feedforward neural networks, in: V. Feldman, A. Rakhlin, O. Shamir (Eds.), 29th Annual Conference on Learning Theory, in: Proceedings of Machine Learning Research, vol. 49, Columbia University, New York, New York, USA, 2016, pp. 907–940, PMLR, 23–26.
[24] K. Wong, R. Dornberger, T. Hanne, An analysis of weight initialization methods in connection with different activation functions for feedforward neural networks, Evol. Intell. 17 (2024) 2081–2089.
[25] K. Selvi, G. Dilip, Enhancing cyber–physical systems security: A review of deep learning and blockchain integration, in: 2024 5th International Conference on Image Processing and Capsule Networks, ICIPCN, IEEE, 2024, pp. 725–734.
[26] D. Das, S. Banerjee, P. Chatterjee, U. Ghosh, U. Biswas, W. Mansoor, Security, trust, and privacy management framework in cyber–physical systems using blockchain, in: 2023 IEEE 20th Consumer Communications & Networking Conference, CCNC, 2023, pp. 1–6.
[27] M.E. Seno, A. Zaidi, B. Gupta, R. Avacharmal, K.S. Yogi, M. Tiwari, F.A. Reegu, N. Shavkatov, M. Soni, A hybrid trust management strategy for reliable cyber–physical system in intelligent transportation, IEEE Trans. Intell. Transp. Syst. 26 (9) (2025) 14383–14392.
[28] H. Xia, F. Xiao, S.-S. Zhang, X.-G. Cheng, Z.-K. Pan, A reputation-based model for trust evaluation in social cyber–physical systems, IEEE Trans. Netw. Sci. Eng. 7 (2) (2020) 792–804.
[29] A. Noor, N. Tariq, M. Asim, F.A. Khan, J.A. Khan, A. Mylonas, A fuzzy logic-based trust framework against sybil and rank attacks in cyber–physical systems, Int. J. Inf. Secur. 24 (2025) 196.
[30] J. Queiroz, P. Leitão, E. Oliveira, A fuzzy logic recommendation system to support the design of cloud–edge data analysis in cyber–physical systems, IEEE Open J. Ind. Electron. Soc. 3 (2022) 174–187.
[31] S. Zou, M. Sun, G. Zhong, X. He, Fuzzy adaptive learning secure control for nonstrict-pure-feedback cyber–physical systems subject to malicious attacks, IEEE Trans. Ind. Cyber-Phys. Syst. 2 (2024) 626–638.
[32] T.R. Ramesh, M. Vijayaragavan, M. Poongodi, M. Hamdi, H. Wang, S. Bourouis, Peer-to-peer trust management in intelligent transportation system: An Aumann's agreement theorem based approach, ICT Express 8 (3) (2022) 340–346.
[33] X. Chen, J. Ding, Z. Lu, A decentralized trust management system for intelligent transportation environments, IEEE Trans. Intell. Transp. Syst. 23 (1) (2022) 558–571.
[34] C. Zhang, W. Li, Y. Luo, Y. Hu, AIT: An AI-enabled trust management system for vehicular networks using blockchain technology, IEEE Internet Things J. 8 (5) (2021) 3157–3169.
[35] M. Cheng, J. Zhang, S. Nazarian, J. Deshmukh, P. Bogdan, Trust-aware control for intelligent transportation systems, in: 2021 IEEE Intelligent Vehicles Symposium, IV, 2021, pp. 377–384.
[36] J. Müller, T. Meuser, R. Steinmetz, M. Buchholz, A trust management and misbehaviour detection mechanism for multi-agent systems and its application to intelligent transportation systems, in: 2019 IEEE 15th International Conference on Control and Automation, ICCA, 2019, pp. 325–331.
[37] J. Qi, N. Zheng, M. Xu, P. Chen, W. Li, A hybrid-trust-based emergency message dissemination model for vehicular ad hoc networks, J. Inf. Secur. Appl. 81 (2024) 103699.
[38] J. Qi, N. Zheng, M. Xu, X. Wang, Y. Chen, A multi-dimensional trust model for misbehavior detection in vehicular ad hoc networks, J. Inf. Secur. Appl. 76 (2023) 103528.
[39] X. Liu, O. Ma, W. Chen, Y. Xia, Y. Zhou, HDRS: A hybrid reputation system with dynamic update interval for detecting malicious vehicles in VANETs, IEEE Trans. Intell. Transp. Syst. 23 (8) (2022) 12766–12777.
[40] F. Ahmad, F. Kurugollu, C.A. Kerrache, S. Sezer, L. Liu, NOTRINO: A NOvel Hybrid TRust Management Scheme for INternet-of-Vehicles, IEEE Trans. Veh. Technol. 70 (9) (2021) 9244–9257.
[41] F. Ahmad, V.N.L. Franqueira, A. Adnane, TEAM: A trust evaluation and management framework in context-enabled vehicular ad-hoc networks, IEEE Access 6 (2018) 28643–28660.
[42] A. Hbaieb, S. Ayed, L. Chaari, A survey of trust management in the Internet of Vehicles, Comput. Netw. 203 (2022) 108558.
[43] W. Li, H. Song, ART: An attack-resistant trust management scheme for securing vehicular ad hoc networks, IEEE Trans. Intell. Transp. Syst. 17 (4) (2016) 960–969.
[44] C.V.L. Mendoza, J.H. Kleinschmidt, A distributed trust management mechanism for the Internet of Things using a multi-service approach, Wirel. Pers. Commun. 103 (2018) 2501–2513.
[45] R. Feng, X. Xu, X. Zhou, J. Wan, A trust evaluation algorithm for wireless sensor networks based on node behaviors and D–S evidence theory, Sensors 11 (2) (2011) 1345–1360.
[46] S. Nie, A novel trust model of dynamic optimization based on entropy method in wireless sensor networks, Clust. Comput. 22 (Suppl 5) (2019) 11153–11162.
[47] K.N. Qureshi, A. Iftikhar, S.N. Bhatti, F. Piccialli, F. Giampaolo, G. Jeon, Trust management and evaluation for edge intelligence in the Internet of Things, Eng. Appl. Artif. Intell. 94 (2020) 103756.
[48] F. Zawaideh, M. Salamah, An efficient weighted trust-based malicious node detection scheme for wireless sensor networks, Int. J. Commun. Syst. 32 (3) (2019) e3878.
[49] A. Talpur, M. Gurusamy, Machine learning for security in vehicular networks: A comprehensive survey, IEEE Commun. Surv. Tutor. 24 (1) (2022) 346–379.
[50] M. Najafi, L. Khoukhi, M. Lemercier, Decentralized prediction and reputation approach in vehicular networks, Trans. Emerg. Telecommun. Technol. 33 (7) (2022) Cited by: 6.
[51] J. Bilski, J. Smoląg, B. Kowalczyk, K. Grzanek, I. Izonin, Fast computational approach to the Levenberg-Marquardt algorithm for training feedforward neural networks, J. Artif. Intell. Soft Comput. Res. 13 (2) (2023) 45–61.
[52] J.d.J. Rubio, Stability analysis of the modified Levenberg–Marquardt algorithm for the artificial neural network training, IEEE Trans. Neural Netw. Learn. Syst. 32 (8) (2021) 3510–3524.
[53] A. Fischer, A.F. Izmailov, M.V. Solodov, The Levenberg–Marquardt method: An overview of modern convergence theories and more, Comput. Optim. Appl. 89 (1) (2024) 33–67.
[54] W.-Y. Shao, J.-Y. Fan, Global convergence of a stochastic Levenberg–Marquardt algorithm based on trust region, J. Oper. Res. Soc. China (2024) 1–23.
[55] O.D.E. Simulator, OMNeT++ discrete event simulator, 2024, (Accessed 23 February 2024).
[56] S. Eclipse, Simulation for urban mobility, 2022, (Accessed 19 May 2022).