Privacy-aware Anomaly Detection and Notification Enhancement for VANET based on Collaborative Intrusion Detection System

Guhan Zheng, *Member, IEEE*, Qiang Ni, *Senior Member, IEEE*, and Yang Lu

***Abstract*—Collaborative Intrusion Detection System (CIDS) is an essential technology that enables vehicular ad hoc networks (VANET) to protect against malicious intrusions. CIDS, however, is unable to prevent accidents if an anomalous vehicle is detected. Detecting anomalies and notifying vehicles in the VANET rapidly is thus essential, considering technical challenges such as communication efficiency, vehicle velocity and privacy. In this paper, we propose a novel two-layer privacy-aware trust evaluation CIDS framework, termed 2PT-CIDS, tailored to VANET. In 2PT-CIDS, vehicles and roadside units (RSUs) cooperate communication-efficient to enhance anomalous vehicle detection and notification. Considering its potential privacy leakage, we then present two types of game-theoretic information incentive mechanisms. In the case of traffic congestion, the privacy-aware incentive mechanism is presented based on the Stackelberg game. A Barycentric Lagrange interpolation (BLI) based algorithm is then proposed to speedy achieve the Nash equilibrium (NE). In the case of traffic smooth, the varying high velocities of vehicles are involved and a noncooperative game-based mechanism is proposed. The optimal NE decision selection is reconstructed as a Markov decision process (MDP) and the NE point is obtained via the designed novel reward-shaping double duelling deep Q network (D3QN) learning algorithm. Simulation results highlight the superiority of 2PT-CIDS over existing CIDS and potential application algorithms for VANET, effectively enhancing anomaly detection and notification considering communication cost and vehicle privacy.**

***Index Terms*—privacy, trust evaluation, information incentive mechanism, Nash equilibrium.**

I. Introduction

A. Background

RAPID advances in communications techniques are progressively making autonomous driving a reality, which could provide a better driving experience. To improve the driving efficiency of autonomous vehicles, they need to communicate with other vehicles or roadside units (RSUs) to share and collect various travel-related information. In particular, communication with RSUs, not only allows vehicles to perceive road accidents, traffic conditions, navigation and other information but also enables vehicles to use strong computing and storage resources. This is because the RSUs are generally fitted with edge cloud servers that provide similar services to vehicles as cloud [1].

Manuscript received 11 January 2024; revised 11 June 2024 and 25 August 2024; accepted 07 October 2024. This work was supported in part by the Western O-RAN Deployment (ONE WORD) Project. (Corresponding author: Qiang Ni.)

G. Zheng, Q. Ni, and Y. Lu are with the School of Computing and Communications, Lancaster University, LA1 4WA, UK (Email:{g.zheng2, q.ni, y.lu44}@lancaster.ac.uk).

The increasing exchange of vehicle information simultaneously presents growing avenues for attackers to penetrate the vehicular ad hoc network (VANET). This poses an ever-increasing challenge for VANETs to detect malicious attacks and attackers (e.g., malicious vehicles). It is necessary to improve the capability of vehicles' intrusion detection systems (IDSs). The collaborative IDS (CIDS) is an effective technique for VANETs to tackle this security challenge [2]. It enables the IDS of an individual vehicle to work collaboratively by exchanging the required information with other vehicles. The intrusion detection of the individual vehicle is thus collaboratively strengthened and improves IDS accuracy and scalability.

B. Challenges

If a vehicle deploying CIDS detects an anomalous vehicle and fails promptly to notify other vehicles on the road, anomalous vehicles still seriously impact the driving experience of other vehicle users. For instance, in the case of a malicious vehicle and other vehicles training together on the same task within an RSU's edge server, it leads to misinformation that greatly hampers training efficiency [3]. Moreover, anomalous vehicles should give wrong vehicle control commands significantly compromise traffic security [4]. Thus, ensuring timely notification among vehicles in CIDS-based VANETs when an anomalous vehicle is detected has emerged as a critical challenge to be addressed.

Nevertheless, vehicles possess variable velocities as well as limited communication distances. This enables communication between vehicles is inefficient, and information cannot be delivered promptly through vehicle-to-vehicle communication [5]. It is thus necessary to design a communication-efficient and rapidly anomalous vehicle notification mechanism. Furthermore, the privacy of the vehicles needs to be considered. Vehicles are not always willing to share their information in case an anomalous vehicle is detected due to concern for their privacy. Even if the information is shared, persuading other vehicles of the detection result poses a challenge for the vehicle. Appropriate incentives for information sharing and detection confirmation are hence crucial.

Building upon the above considerations, the primary challenge associated with CIDS-based VANET is to devise a communication-efficient and privacy-considering scheme. This scheme requires swift detection and confirmation of anomalous vehicles, along with prompt alerting of vehicles with varying velocities in the VANET.

C. Related works

T. Nandy et. al [6] proposed a trust-based CIDS for VANET. They employed a k-nearest approach to improve the vehicle's local intrusion detection capability. G. Raja et. al [7] introduced an SP-CIDS for VANETs that used distributed learning to enhance the accuracy of CIDS. In [8], the graph node attention network was utilized to extract context-dependent features and thus improve the accuracy of vehicles' CIDS modules. In addition, R. Liu et. al [9] presented a privacy-preserving two-layered distributed machine learning framework. It can be used to train the machine learning model of CIDS. Nevertheless, these studies are all concerned with upgrading the intrusion detection accuracy via collaboration rather than rapidly alerting detected anomalous vehicles in the VANET.

To the best of our knowledge, there is no previous CIDS-based framework for VANET that effectively detects, confirms anomalous vehicles, and notifies all vehicles rapidly.

Several searching anomalous vehicle techniques for the VANET, blockchain and trust management, show promise for addressing these challenges. The decentralisation, anonymity and security properties of blockchain ensure that both parties are trustworthy when a transaction is generated [10]. A series of works have studied leveraging blockchain in VANETs to secure information exchange and uncover anomalous vehicles [11]–[13]. Nevertheless, blockchains become useless in case more than half of the vehicles in the environment are malicious. The high latency of blockchain also makes it difficult or even impractical for deployment in high-mobility VANETs.

An alternative promising technique is trust management. A decentralized trust management system was proposed in [15] for VANET. This approach is based on blockchain and hence encounters the same problem as blockchain in VANET. S. A. Siddiqui et. al [14] proposed a weight quantification scheme for VNAET to find the anomalous vehicle. Furthermore, A. Mahmood et. al [16] presented a trust management system for the identification and eviction of misbehaving vehicles. These two approaches, however, neglected the privacy of the vehicle.

To address the aforementioned privacy issue in trust computing, Xing et. al [17] proposed a two-layer intrusion detection framework with trust computing. It secures vehicles via intrusion detection systems while allowing a privacy-considering method for fast identification of anomalous vehicles and notification to the whole VANET. Nevertheless, the characterisations of vehicles and the VANET are disregarded, e.g., traffic conditions, vehicle velocity, and time variations. Moreover, these conventional trust management approaches require unique information defined by themselves to determine vehicle anomalies. Vehicles do not notify and confirm the attacker/anomalous vehicle at the same time when it is detected via the CIDS module. This requires an additional number of communications (i.e., communication cost), increases the detection time and the anomalous vehicle does not always respond accordingly.

D. Motivation and contributions

Given the above, we contend that there is an urgent need for a framework and algorithm facilitating anomalous vehicle detection and notification in CIDS-based VANET, with a specific focus on communication efficiency and privacy.

Consequently, to bridge the gap identified above, this paper proposes a novel two-layer privacy-aware trust evaluation framework for VANET, i.e., 2PT-CIDS, which integrates CIDS and trust computing. The detection of anomalous vehicles is simultaneous with the enhancement of intrusion detection capabilities. It allows vehicles to perform anomalous vehicle detection and to improve their real-time intrusion detection capability simultaneously via the CIDS module. Untrustworthy vehicle information is provided to the RSU for further identification and communication-efficient speedy notification. Considering privacy data-sharing issues in trust computing, we further propose new privacy-aware information incentive mechanisms for the 2PT-CIDS. They are based on game theories and take into account the vehicle velocity. We also propose different algorithms to find the Nash equilibrium (NE) points for these games.

The main contributions of this paper are summarized as follows:

- We design a 2PT-CIDS framework for VANET anomaly detection. Vehicles can enhance intrusion detection capabilities while detecting anomalous vehicles and notifying other vehicles considering communication efficiency, velocity and privacy.
- We present a new privacy-aware incentive mechanism based on the Stackelberg game for congested traffic scenarios. Moreover, a barycentric Lagrange interpolation (BLI) based algorithm is proposed to expedite the game to achieve the Nash equilibrium (NE) point.
- We propose a novel privacy-aware incentive mechanism based on the non-cooperative game for high-speed vehicle mobility scenarios. Furthermore, a reward-shaping double duelling deep Q network (D3QN) algorithm is presented to enable the vehicle to achieve the optimal decision quickly at varying times and velocities.

The rest of the paper is organised as follows. The traffic model and CIDS are introduced in Section II and the proposed 2PT-CIDS framework is presented in Section III. Section IV and Section V present the detailed design of our game-theoretic incentive mechanisms and the solutions for searching NE for traffic congestion and traffic smoothness, separately. Section V illustrates the simulation results to demonstrate the effectiveness of our algorithms. Finally, we conclude the paper in Section VI.

For easy reference, the main parameters and their description used throughout this paper are presented in Table I.

II. System Model

In this section, we first present the traffic model and then describe the CIDS.

A. Traffic model

A set of RSUs $\{1, 2, ..., p, ..., P\}$ are equipped with edge cloud servers beside the road and a set of vehicles

TABLE I: NOTATION DEFINITION

Symbol	Definition
P	Set of RSUs
Q	Set of vehicles
v	Vehicle velocity
ς_p	Number of vehicles arriving at the RSU p's range
n_p	Number of vehicles in RSU p's coverage
T	Trust value
U_p	Utility of RSU p
U_q	Utility of vehicle q
$l_{q,p}$	Size of report from q to p
$\mu_{q,p}$	RSU p's reward factor
R	Incentivise reward
Φ	Privacy loss cost
A	Report revenue
W	Bits of total report
U_q	Utility of vehicle q
Ω	Maximum long-term utility
V	Value-state function

$\{1, 2, ..., q, ..., Q\}$ are driving on the road. Each vehicle communicates with at least one nearby RSU. We assume the width of the road is uniform in each area. According to [19], the vehicles follow a Poisson distribution when they reach the coverage of each RSU. The number of vehicles arriving at the RSU p communication range can be expressed by

$$\varsigma_p = \overline{v_p}\frac{n_p}{L_p}, \tag{1}$$

where n_p is the number of vehicles in RSU p's coverage, L_p is the width of the road unit in meters where RSU p is located. Further, $\overline{v_p}$ is the average velocity (km/h) of vehicles in the coverage of RSU p, which is related to the extent of traffic congestion and can be denoted by

$$v_p = \max\{v_{p\ max}(1 - \frac{n_p}{n_{p\ max}}), v_{p\ min}\}, \tag{2}$$

where $v_{p\ max}$ is the velocity of the vehicle when the traffic is smoothest so that the vehicle is moving unimpeded on that road at maximum speed, and $n_{p\ max}$ is the number of vehicles in such circumstances. Similarly, $v_{p\ min}$ is the speed when the traffic is congested where the vehicles move slowly. Moreover, the same as [20], we suppose the speed of an arbitrary vehicle in the service range of an RSU is a normally distributed random variable.

B. CIDS

In CIDS, there are two types of information transmission required to implement intrusion detection enhancements, i.e., **challenges** and **requests**. They are defined as follows:

*1) **Challenges**:* are sets of alarms sent to test vehicles, which the sending vehicle knows in advance the severity of these. The **challenges** is similar to the question that knows answers in advance.

*2) **Requests**:* are similar to the **challenges**, but the sending vehicle does not know the severity in advance. The feedbacks of **requests** are used to "alarm aggregation" to improve the intrusion detection performance of the sending vehicle, which is the most unique characteristic of CIDS.

The **challenges** and **requests** are sent in a random manner so that tested vehicles are difficult to distinguish between **challenges** and normal **requests**. The vehicle first evaluates the trust value of the tested vehicle through judging satisfaction of the feedback of **challenges**. Upon determining that the vehicle is trusted, the "alarm aggregation" is executed. We assume vehicle j transmitting these two types of information vehicle i via CIDS modules. The trust value of the vehicle i can be denoted by

$$T_i^j = (\frac{\omega_s \sum_{k=0}^{n} F_{j,i}^k \lambda^{t_{j,i}^k}}{\sum_{k=0}^{n} \lambda^{t_{j,i}^k}} - T_s)(1-x)^d + T_s, \tag{3}$$

where $F_{j,i}^k \in [0, 1]$ is vehicle j's satisfaction of the received feedback k from vehicle i and n is the total number of feedbacks. Further, $t_{j,i}^k$ is the time period that vehicle j received the replies from vehicle i. Here, λ is a forgetting factor that assigns less weight to older feedback responses. Further, ω_s is a significant weight related to the total number of received feedbacks. If there are only a few feedbacks under minimum number m, then $\omega_s = \frac{\sum_{k=0}^{n} \lambda^{t_{j,i}^k}}{m}$, otherwise $\omega_s = 1$. Thus, $T \in (0, 1]$.

In addition, vehicle i is sometimes reluctant to send back unknown answers to guarantee its trust value. To encourage vehicle i to provide satisfactory feedback responses whenever possible, a "*don't know*" answer is set. It also can decrease their trust value. Hence, in Eq. (3), x is the percentage of "*don't know*" answer during a time period and d is a positive incentive parameter to control the punishment of "*don't know*" replies. Further, T_s is the default trust value of a stranger.

III. 2PT-CIDS

Our 2P-CIDS is based on a combination of direct and indirect trust. The overall process is shown in Fig. 1.

It consists of two layers: vehicles fitted with CIDS components form the first layer and RSUs form the second layer. As with pervasive CIDS communication, in the first layer, vehicles perform direct trust computing and enhance their intrusion detection capabilities through exchange **requests** and **challenges**

Nevertheless, in the VANET, it is challenging for a vehicle to notify an anomalous vehicle situation to other normal vehicles on a large scale in a short period of time. That is because of the high-speed mobility and limited communication distance ranges of vehicles. In addition, the judgment of a single vehicle is not always very credible, and it also needs to be judged again by others to confirm. In order to save communication resources and quickly notify all vehicles in the VANET, RSUs with a wider communication range and stronger detection and computing capabilities will be formed as a second layer. Vehicles are encouraged to transmit vehicle anomaly alerts to RSUs for information aggregation for trust computing verification.

In the second layer, RSUs are also fitted with CIDS modules and constantly interact with vehicles. We employ RSUs as judges and advocates. The reporting vehicle j needs to report **challenges** to RSU p that were sent to the suspicious vehicle i before. Further, vehicle j also needs to report its own corresponding answer and feedback from vehicle i. The RSU

p first uses Eq. (3) to verify the trustworthiness of vehicle j. We have the trust $T_{p,i}$ as:

$$T_{p,i} = \sum_{i=1} T_i^j. \tag{4}$$

Only if vehicle j is sufficiently trustworthy, the RSU p will proceed to verify the trust value of the reported anomalous vehicle i. After receiving the first anomaly report about vehicle i in a period of time T, RSU p performs a weighted summation of vehicle i's trust values obtained from all reporting vehicles and derive the RSU p's local trust value for vehicle i.

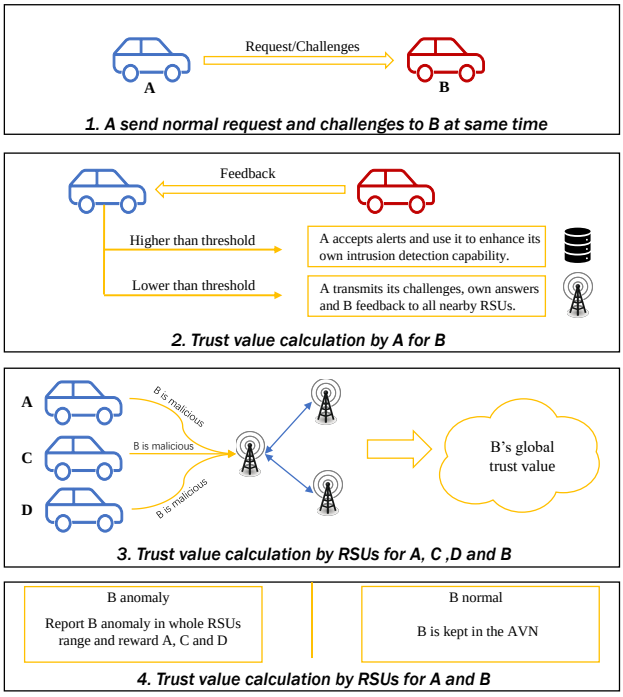

Fig. 1: Overall process of 2PT-CIDS.

The RSU p then communicates with other RSUs who may also receive the intrusion report about vehicle i and share their local trust value evaluation about vehicle i. RSU p therefore can get the trust evaluation of vehicle i from the RSUs network. We define this evaluation value as vehicle i's network trust value, which can be denoted by

$$T_{n,p,i} = \frac{\sum_{h=1,h\neq p}^{P} T_{p,i}\gamma_{p,h}^t}{\max_{h\in P,h\neq p}\gamma_{p,h}^t(I-1)}, \tag{5}$$

where I is the number of RSUs in the networks, $T_{p,i}$ is the local trust value of vehicle i calculated by RSU p before it shares with other RSUs. Further, $\gamma_{p,h}^t = |\Lambda_p^t \cup \Lambda_h^t|$ is the similarity of intrusion reports detected by RSU p and other RSU h by the time slot t. Further, Λ denotes the set of challenges answers detected differently by RSUs. Moreover, $|\cdot|$ is the number of challenging answers in a set. Therefore, the total global trust of vehicle i as assessed by RSU p can be expressed by

$$T_{g,p,i} = \alpha T_{n,p,i} + \beta T_{p.i}, \tag{6}$$

where weights parameters $\alpha + \beta = 1$, and they are utilized as weight parameters to trade off the preference of final trust value towards direct or indirect trust. Higher α than β means the device prefers its own trust assessing result. i.e., direct trust value, while higher β than α means the device prefers the trust value from other devices, i.e., indirect trust value. It can adaptively adjust based on evaluation time, vehicle velocities, devices' confidence in their intrusion detection capabilities, number of devices participating in indirect trust, etc. In case $T_{global,p,i}$ is below the predefined trustworthiness threshold, RSU p will inform all vehicles in its coverage about vehicle i anomaly and its trust value to protect the VANET.

IV. Incentive Mechanism: Traffic Congestion

A. Motivation of the game

From the previous section, in the first layer, if vehicle i refuses to answer requests or challenges, vehicle j will treat it as having the answer "*don't know*". It can decrease the trust value of vehicle i according to CIDS. Vehicles can also interact via V2V links to enhance intrusion detection capabilities. Thus they have a sufficient basis to implement trust computing. However, vehicle j hasn't got the reason/benefit to report anomaly situations to RSUs, because it can leak part of vehicle j's privacy. This makes it difficult for 2PT-CIDS to detect anomalous vehicles efficiently and accurately.

Therefore, we first propose a privacy-aware congestion scenario incentive mechanism based on Stackelberg game theory to encourage vehicles to give more ratio of the anomalous vehicle's information. We assume each report has the same degree of privacy and the BLI-based algorithm is presented to achieve the NE. RSUs will reward a vehicle by increasing its initial trust value when it reports and successfully detects anomalies. Because vehicles with a higher initial trust value can get faster responses about it **requests** and **challenges** when multiple vehicles are communicating with one vehicle according to CIDS [21].

B. Game at vehicles

The utility $U_{q,p}$ of the vehicle q in the service range of RSU p after reporting the trust computing information to RSU p at a fixed time slot can be expressed as received reward $R_{p,q}$ minus privacy leak $\Phi_{q,p}$. According to the conventional vehicle to RSU incentive scheme [5], we have

$$U_{q,p}(l_{q,p}) = R_{q,p}(l_{q,p}) - \Phi_{q,p}(l_{q,p}), \tag{7}$$

where $l_{q,p}$ is the size of reports from vehicle q, and $W_{q,p}$ is a total bits of reports. Further, $R_{q,p}(l_{q,p})$is the incentive reward obtained from RSU p, which can be denoted by

$$R_{q,p}(l_{q,p}) = \frac{\mu_p(1-T_{q,p})l_{q,p}}{\delta_p(1+W_{q,p})}, \tag{8}$$

where δ_p is the number of vehicles reporting anomaly alert information RSU p about the same anomalous vehicle, $T_{q,p}$ is the trust value of vehicle q for RSU p and μ_p is the reward factor. Further, $\Phi_{q,p}(l_{q,p})$ is the privacy loss cost of vehicle q, which we can be expressed as [22]:

$$\Phi_{q,p}(l_{q,p}) = \chi \log_2(1 + e^{1-\frac{W_{q,p}+1}{l_{q,p}}}), \tag{9}$$

where χ is the weighting parameter to convert the loss into a monetary consumption. The $\log$ function represents the privacy leakage versus the number of providing documents.

However, it is flawed to divide μ_p directly and equally in a proportional way to δ_p. vehicle q does not know how many other vehicles in the vicinity will transmit similar vehicle anomalies report in advance. It should be made in conjunction with RSUs who know this information. vehicles can only calculate by using the parameters they know. In addition, the conventional approach does not consider the reported revenue of RSUs, which makes it impossible for the VANET to formulate μ_p according to real-time. We therefore assume the utility as monetary gain. The ϕ is redefined as a monetary benefit parameter and Eq. (8) is rewritten as:

$$R_{q,p}(\mu_p, l_{q,p}) = z\mu_p(1 - T_{q,p})\frac{l_{q,p}}{W_{q,p}+1}, \tag{10}$$

where z represents the monetary benefit parameter. The calculation of μ_p is related to RSU p's utility, which will be mentioned later.

We can thus define the vehicle q's utility function as:

$$U_{q,p}(\mu_p, l_{q,p}) = z\mu_p(1-T_{q,p})\frac{l_{q,p}}{W_{q,p}+1} - \chi \log_2(1+e^{1-\frac{W_{q,p}+1}{l_{q,p}}}), \tag{11}$$

Problem 1:

$$\max_{l_{q,p}} \quad \mu_p(1 - T_{q,p})\frac{l_{q,p}}{W_{q,p}+1} - \chi \log_2(1 + e^{1-\frac{W_{q,p}+1}{l_{q,p}}}) \tag{12a}$$

$$s.t. \quad 0 < l_{q,p} < W_{q,p}. \tag{12b}$$

C. Game at the RSU

In this subsection, we define the RSU p's utility and present the game at the RSU p. Without loss of generality, we define the RSU p's utility function as revenue minus cost and denoted by

$$U_p(\mu_p, l_{q,p}) \triangleq A_p(\mu_p, l_{q,p}) - R_p(\mu_p, l_{q,p}), \tag{13}$$

where l_p is the total size of RSU received reports, $A_p(\mu_p, l_{q,p})$ is the reporting revenue (e.g., the accuracy of trust evaluation) gained from vehicles' reporting data and $R_p(\mu_p, l_{q,p})$ is the cost incurred due to reward to vehicles.

Here, we model the RSU p's benefit as:

$$A_p(\mu_p, l_{q,p}) = \lambda f_p(l_{q,p}), \tag{14}$$

where $f_p(l_{q,p})$ is the reporting revenue function (RRF), and λ is a monetary benefit parameter to convert RRF into a monetary benefit. We use a $\log$ function to model the RRF as:

$$f_p(l_{q,p}) \triangleq \log_2(1 + \sum_{q\in\delta_p} \frac{T_{q,p}l_{q,p}}{(W_{q,p}+1)}). \tag{15}$$

Though other functions can also be used to model $f_p(l_{q,p})$, the logarithmic function is shown in the literature to be more suitable to represent the relationship between the file number and computing value compliance [24], [25]. It is also obvious that when the sum of the reporting file is 0, the benefit is 0. Further, the accuracy of calculating trust values (benefit) increases with the number of reports and vehicles with higher trust values are likely to provide more valuable reports. These indicate that Eq. (15) can perfectly model the RRF.

The cost function $C_p(\mu_p, l_{q,p})$ consists of total reward paid to vehicles. We have

$$R_p(\mu_p, l_{q,p}) = \sum_{q\in\delta_p} R_{q,p}(\mu_p, l_{q,p}). \tag{16}$$

The utility function of RSU p thus can be expressed as:

$$U_p(\mu_p, l_{q,p}) = \lambda \log_2(1+\sum_{q\in\delta_p} \frac{T_{q,p}l_{q,p}}{(W_{q,p}+1)}) - \sum_{q\in\delta_p} R_{q,p}(\mu_p, l_{q,p}). \tag{17}$$

Problem 2:

$$\max_{\mu_p} \quad \lambda \log_2(1 + \sum_{q\in\delta_p} \frac{T_{q,p}l_{q,p}}{(W_{q,p}+1)}) - \sum_{q\in\delta_p} R_{q,p}(\mu_p, l_{q,p}), \tag{18a}$$

$$s.t. \quad \mu_p > 0. \tag{18b}$$

D. Equilibrium analysis and optimal solutions

NE existence: Problem 1 and Problem 2 together form a Stackelberg game. The key to this game is to find the NE point(s) between RSU p (leader) and vehicles (followers). It is observed from Problem 1 that the vehicles' strategy set is compact and convex and $l_{q,p}$ is continuous. Therefore, a pure strategy NE exists in this game according to the Debreu-Glicksberg-Fan theorem [26].

The aim of finding the NE is to get the best μ_p^* and $l_{q,p}^*$. Therefore, the NE of this model needs to satisfy the following conditions:

$$U_p(\mu_p^*, l_{q,p}^*) \geq U_p(\mu_p, l_{q,p}^*), \tag{19}$$

$$U_{q,p}(\mu_p^*, l_{q,p}^*) \geq U_{q,p}(\mu_p, l_{q,p}), \ \forall q. \tag{20}$$

When congestion occurs, vehicles have enough time to communicate with one RSU p. The RSU can calculate vehicles that wish to report the trust computing information and their trust value. For this situation, we use the backward induction method. That is, followers, obtain optimal strategies first and subsequently, the leader develops its own strategy based on the followers' optimal strategy.

E. Optimal strategies for traffic congestion

To find the NE, we first find the first derivative of the strategy space $l_{q,p}$ of vehicle q, that is

$$U'_{q,p} = \frac{z\mu_p(1-T_{q,p})}{W_{q,p}+1} - \frac{(W_{q,p}+1)\chi e^{1-\frac{W_{q,p}+1}{l_{q,p}}}}{l_{q,p}^2(e^{1-\frac{W_{q,p}+1}{l_{q,p}}}+1)\ln 2}. \tag{21}$$

In the case of Eq. (24) equals zero, $U_{q,p}$ can reach its maximum value. Hence, when $U_{q,p}^{'}=0$, we have the optimal $l_{q,p}$ as:

$$l_{q,p}=f_{q,p}(\mu_p), \tag{22}$$

where $f_{q,p}(\cdot)$ denote $l_{q,p}$ as the function of μ_p correlation, but due to the complexity of Eq. (24), we cannot specifically represent. We then substitute the optimal $l_{q,p}$ into Eq. (20) to find the optimal $\mu_{q,p}$. We have

$$\begin{aligned} &U_p(\mu_p) \\ &=\lambda\log_2(1+\sum_{q\in\delta_p}\frac{T_{q,p}f_{q,p}(\mu_p)}{(W_{q,p}+1)})-\sum_{q\in\delta_p}R_{q,p}(\mu_p,f_{q,p}(\mu_p)), \end{aligned} \tag{23}$$

where $f_{q,p}(\mu_p)$ denotes $l_{q,p}$ and we thus can achieve the NE by obtaining the maximum U_p. Nonetheless, similar to Eq. (22), the maximum U_p is also difficult to be obtained. The relationship between μ_p and $l_{q,p}$ thus is difficult to find and mathematically intractable. In general, the grid search method [27] (Algorithm 1) can solve this problem. Nevertheless, the numerical iterative method results in a huge amount of computation and requires long execution delays even at edge servers. We thus propose a BLI-based algorithm to estimate the optimal μ_p and $l_{q,p}$. The BLI [28] can construct a polynomial function similar to the original complex function by substituting some values of the independent variables and the corresponding dependent variables. This function enables the analytical solution of the optimal NE instead of the original numerical iteration solution. Hence, the approximate optimal value of the function can be obtained simply, quickly and efficiently. The new U_p, i.e., $U_{p,new}$ can be denoted by [28]

$$U_{p,new}(\mu_p)=\frac{\sum_{e=0}^{E}\frac{\omega_e}{\mu_p-\mu_e}U_{p,e}}{\sum_{e=0}^{E}\frac{\omega_e}{\mu_p-\mu_e}}, \tag{24}$$

where E is the number of substituting (interpolation) values, and $\mu_{p,e}$ and $U_{p,e}$ denote the value of the e-th substituting the value of μ_p and U_p. Furthermore, ω_e is the barycentric weight, which can be expressed by

$$\omega_e=\frac{1}{\prod_{e'=0,e'\neq e}^{E}(\mu_{p,e}-\mu_{p,e'})}, \tag{25}$$

where $\mu_{p,e'}$ is the substituting values except $\mu_{p,e}$.

In addition, in this paper, we choose Chebyshev nodes as the interpolation values. Because the distribution of values is denser at the two ends of the interval, sparse in the middle makes the polynomial interpolation vehicle good numerical stability [28]. Chebyshev nodes satisfy exactly this requirement. The selected μ_e can be thus denoted by

$$\mu_{p,e}=X-Y\cos\frac{e\pi}{n}, \tag{26}$$

where n is the polynomial degree, $[X,Y]$ represents the selected interval.

Since Eq. (24) is a polynomial function, it is straightforward to obtain the maximum value. We thus approximate the NE by obtaining the optimal μ_p, and the optimal $l_{q,p}$ after integer variable recovery. The process of the BLI-based algorithm is shown in Algorithm 2. It can be seen that Algorithm 1 undergoes two N rounds of iteration and the complexity thus is $O(N^2)$. Further, the complexity of our proposed approach (Algorithm 2) is $O(E^2)$. However, E is often much smaller than N, which we will prove in simulations.

V. Incentive Mechanism: Traffic Smooth

A. Noncooperative game design

Due to the high mobility, vehicles can switch RSU services frequently in a short period of time when the traffic is smooth. This not only makes it difficult for a single RSU to determine the number of vehicles in the game but also makes it prone to fluctuations in global trust value. Therefore, we change the Stackelberg game in traffic congestion to a noncooperative game for smooth traffic. vehicles estimate the number of likely reporting vehicles on the road based on their own average speed over a period of time and judge which mechanism to use based on such speed. Note that the time period we use to estimate the average velocity of vehicles is much less than the traffic condition change time in an RSU service range, and therefore the traffic condition change can be ignored.

As we mentioned from Eq. (2), i.e., $\overline{v_p}=v_{p\ max}(1-\frac{n_p}{n_{p\ max}})$ in case of traffic smoothly. We assume that vehicle q approximates the average vehicle's velocity and trust value in an RSU service range by its own velocity and trust value. Thus, the number of vehicle n in an RSU coverage can be denoted by

$$n=n_{p\ max}-\frac{\overline{v_q}n_{p\ max}}{v_{p\ max}}. \tag{27}$$

We then can derive the number of vehicles in this RSU communication range from Eq. (1) as:

$$\varsigma=\frac{\overline{v_q}v_{p\ max}n_{p\ max}-\overline{v_q}^2n_{p\ max}}{L_pv_{p\ max}}. \tag{28}$$

The reporting vehicle number δ therefore can be estimated in

$$\delta=\varphi\varsigma, \tag{29}$$

where $\varphi\in(0,1]$ is the weighting parameter, which denotes the possibility of detecting anomalous vehicles when vehicles communicate with each other.

We assume that vehicle q is unaware of the situations of other vehicles and estimate these via vehicle q 's velocity. The noncooperative utility function of RSU p thus can be estimated by vehicle q as:

$$U_{p,q}^t(\mu_p^t,l_q^t)=\lambda\log_2(1+\frac{l_q^t}{W_q+1}\delta_q)-z\mu_p^t(1-T_q^{t-1})\frac{l_q^t}{W_q+1}\delta_q, \tag{30}$$

where T_q^{t-1} is the trust value of the vehicle q in the previous sending, which will change by increasing the trust reward value. It can be updated by:

$$T_q^t=\begin{cases} T_q^{t-1}+R(\mu_p^t,l_q^t)/z, & T_q^{t-1}<1 \\ T_q, & t=0 \\ 0. & else \end{cases} \tag{31}$$

Algorithm 1 Grid search

Initialize: reward factor μ_p, number of iterations N, total report set W_q.

for n in N
  Substitute μ_p and get $l_{q,p}(q = 1, .., \delta_p)$ by Eq. (17)
  Count the value of U_p
end for
for each μ_p in N
  Find maximum $U_p(n)$
end for
  Output the optimal value of μ_p and $l_{q,p}$

Algorithm 2 BLI-based algorithm

Initialize: the set of substituting values $\mu_{p,e}(e = 1, \ldots, E)$, $U_{p,e}(e = 1, \ldots, E)$, set $\omega_e = 1$, $U_{p,new} = 0$,set $W_q(q = 1, .., \delta_p)$, R and D_i.

for each e in E
  for each $e^{'}$ in E
    if $e^{'} = e$
      continue
    else
      $\omega_e = \frac{\omega_e}{\mu_{p,e} - \mu_{p,e^{'}}}$
    end if
  end for
  $U_{p,new} = U_{p,new} + \frac{\frac{\omega_e}{\mu_{p,e} - \mu_{p,e}} U_{p,e}}{\frac{\omega_e}{\mu_{p,e} - \mu_{p,e}}}$
end for
Maximize $U_{p,new}$ and obtain corresponding μ_p
Each vehicles obtain optimal $l_{q,p}$ based on μ_p
if rounding $l_{q,p}$ satisfies constraint (12b) **then**
  Output the optimal value of μ_p and rounding $l_{q,p}$
else
  break
end if

B. Optimal strategies for traffic smooth

In case of traffic is congested, the number of participants in each game is different but determinable. Therefore, the vehicle aims to get the maximum utility on this determined occasion. However, as traffic is smooth, the environment is variable. The aim of vehicles can be represented as finding the optimal vehicle decisions in each time period game to obtain the maximum long-term utility (reward). It is defined as the weighted sum of the instantaneous utilities (rewards) over finite reporting time RT in the time slot T. We can denote it by

$$\Omega = \sum_{t=0}^{RT-1} \gamma^t U_{q,p}^t, \tag{32}$$

where $\gamma \in [0, 1)$ is the discount rate to discern the impact of future rewards and the smaller γ means the smaller the impact of subsequent rewards. We can only focus on the immediate reward when $\gamma = 0$. At any reporting time instant t, each vehicle's reward relies on its action and new state at the next reporting time. Here, a finite-state MDP is appropriate to be introduced to describe this game. We thus model this optimal decision issue as an MDP with states, actions, reward functions, and the state transition probability.

We define the state $s(t)$ to indicate the average velocity and trust value of the vehicle at the reporting time t. The action of each vehicle indicates how much information it needs to report, and the number is at most W which can be defined as $a(t)$. The corresponding μ of this action to maximise the $U_{q,p}^t$ can be derived by Eq. (21). Further, we use $r(t)$ to denote the reward function and this reward is the utility function $U_{p,q}^t$ of RSU q in reaching Nash equilibrium. Generally, we should use the utility $U_{q,p}^t$ of vehicle q to denote the reward function. However, due to the existence of the game, we set $r(t)$ to $U_{p,q}^t$ to ensure that the result obtained by Eq. (21) reaches NE at each time period. This is because there exists a unique $U_{q,p}^t$ reaching NE when $U_{p,q}^t$ reaches NE. This constructed reward aims to enable the vehicle to find the most suitable amount of data provided and improves the efficiency of anomaly detection and notification. In addition, the information on state transition probability is unknown for vehicles in each episode. It makes the integer programming method approach difficult to use to obtain the optimal policy π^*, i.e., the optimal number of report information at state $s(t)$.

Reinforcement learning is a promising method to obtain the optimal policy π^* to maximize the long-term reward [29]. It sets an agent to exchange actions and states with the environment to get the optimal policy. The optimal policy π^* in each state can be derived by the value-state function, we have

$$V(s, \pi) = \mathrm{E}[\sum_{t=0}^{RT-1} \gamma^t U_{p,q}^t | s(0) = s], \tag{33}$$

where $\mathrm{E}[\cdot]$ is the expectation operator. Hence, we have the following inequalities to get optimal decision:

$$V(s, \pi^*) \geq V(s, \pi), \tag{34}$$

where π^* is the optimal decision. At the optimal decision, the value-state function is greater than taking other decisions

According to the Bellman optimal equation, we obtain the optimal $V(s, \pi^*)$ as:

$$V^*(s, \pi) = V(s, \pi^*) = \log_a[U_{p,q}^{t,a} + \gamma \textstyle\sum_s P_{ss'} V(s^{'}, \pi^*)]. \tag{35}$$

To deal with this MDP problem, Q-learning, one of the reinforcement learning approaches is an effective method [30]. According to Q-learning, we can obtain optimal policy π^* from optimal Q-value function $Q^*(s, a)$ as:

$$V^*(s, \pi) = \log_a Q^*(s, a). \tag{36}$$

However, the training time for Q-learning rapidly grows as the status or action value grows. It is difficult for vehicles to handle such a large number of data. To deal with such

huge states and action spaces MDP problem, deep Q network (DQN) is one of the popular reinforcement learning methods [30]. It uses deep neural networks (DNN) to approximate action value $Q(s,a;\theta)$ that helps the agent to find the optimal approximate action. However, as the same values are used to select and evaluate an action in the DQN method, the Q-value function may be over-optimised for estimation. Thus, double DQN (DDQN) [31] is used to mitigate the above problem by decoupling the selection of the target Q-action and the calculation of the target Q-value to eliminate the problem of overestimation. The target value y of DDQN be defined as:

$$y = r(s,a) + \gamma Q(s^{'}, \mathrm{argmax}_{a^{'}} Q(s^{'}, a^{'}; \theta); \theta^{'}), \tag{37}$$

where $r(s,a)$ is the current reward, θ is the weight of the online network and $\theta^{'}$ represent the weight of target network. Specifically, both the online network and target network are used to compute the optimal $Q(s^{'}, a^{'}; \theta)$ by using the next state $s^{'}$. Then, the target value y is obtained. Finally, the error is calculated by subtracting the target value with the optimal action value function $Q(s,a;\theta)$ predicted by the online network. The weights are then updated by backpropagation.

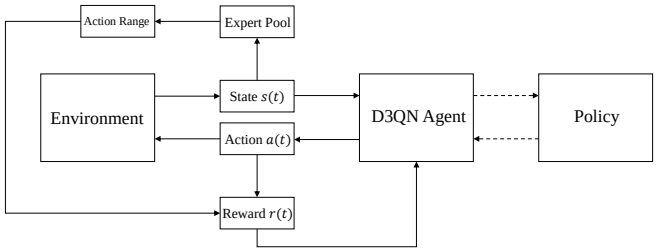

Fig. 2: Reward shaping D3QN process.

Moreover, noting that the Q-value function can describe how profitable an action a is taken at a state s. Another improvement called duelling DQN [32] is used to estimate value function $V(s,\pi)$ and advantage function $A(s,a) = Q(s,a;\theta) - V(s,\pi)$ respectively. The advantage function $A(s,a)$ is used to characterise the advantage of the action over other possible actions. Thus, the duelling architecture enables us to split the last layer of the DQN into two subnetworks to estimate value function $V(s,\pi)$ and advantage function $A(s,a) = Q(s,a;\theta) - V(s,\pi)$ respectively. The action value function $Q(s,a;\theta)$ can be estimated by combining $V(s,\pi)$ and $A(s,a)$.

In addition, learning efficiency can be improved if the optimal action ranges in certain situations are known in advance. It is named expert knowledge and used to guide the action selection process. Fortunately, we can get the results for different numbers of vehicles (i.e., different vehicle velocities) with $\gamma = 0$ by using Algorithm 2. It can be employed as an expert to assess the approximate range of results for γ at different values. However, expert knowledge cannot be used directly for algorithm training. Therefore, inspired by [31] and [32], in this paper, we propose a reward-shaping D3QN to obtain the optimal strategy. It combines DNN-based DDQN and duelling DQN as the training model. In particular, it utilises expert knowledge as an extra reward in reward value to improve training speed and accuracy (Fig 4). We employ the original reward function to synthesise expert knowledge rewards to obtain new reward values and feed them to the experience pool in the D3QN agent, completing the update of the weights for the network.

Algorithm 3 Reward-shaping D3QN for trust computing information reporting

Input: List of allowed actions to be taken by vehicle q.
Output: Optimal actions are required to achieve the maximum long-term utility, if feasible.
Initialize: parameters θ and $\theta^{'}$ in online network Q and target network $Q^{'}$ respectively and assign the Q network parameters to the $Q^{'}$ network, $\theta \longrightarrow \theta^{'}$. Target Q network parameter update frequency p. Replay memory $\mathcal{D}$.

Offline perform Algorithm 2 and obtain expert guidance pool.
for episode=1, M do
  Reset environment and state s_0
  for $t = 1, T$ **do**
   Selecting a range of action A^* from the expert guidance pool based on state s_t
    With probability ε select a random action a_t
    Otherwise select $a_t = \max_{a_t} Q^*(s_t, a; \theta)$
    Choose action a_t and observe immediate reward r_t
    if $a_t \in A^*$
      $r^* = r_t + \eta_1$
    else
      $r^* = r_t$
    end if
    Update s_{t+1} according to $\mathcal{P}$
    Store $(s_t, a_t, r_t^*, s_{t+1})$ in memory $\mathcal{D}$
    Sample random minibatch of $(s_t, a_t, r_t^*, s_{t+1})$ from $\mathcal{D}$ and calculate the y
    Calculate the loss by $(y - Q(s,a;\theta))^2$
    if $t\%p == 0 : \theta \longrightarrow \theta^{'}$
  end for
end for

The range of possible optimal value l_q^t can be estimated in some special cases of speed and W_q advance by Algorithm 2 and get the optimal report ratios τ^t, which can be denoted by

$$\tau^t = \frac{l_q^t}{W_q + 1}. \tag{38}$$

We can use the τ^t to estimate the result of l_q^t in other W cases. The expert bootstrap values previously calculated can provide an approximate range of optimal action A^*. When the action $a(l_q^t)$ falls within this range when training, the agent will receive an additional reward η_1. The detailed procedure of reward shaping D3QN is presented in Algorithm 3.

VI. Numerical Evaluation

A. Simulation settings

We simulate an autonomous driving scenario in which several vehicles driving on the road assisted with RSUs. We compare our proposed framework with several types of state-of-the-art anomalous vehicle detection frameworks to demonstrate the communication efficiency of our 2PT-CIDS. They are pure trust-based framework [14], blockchain-based framework [15] and two-layer based framework [5]. In terms of the privacy-aware information incentive mechanism for vehicles to RSUs, as the related work is limited, we compare our work with the conventional method, i.e., [5], [17]. Simulation settings are the same as [17], we assume different roads have different vehicle speed limits with the road's width $L = 20m$. The congestion velocity is 20 km/h and the maximum number of vehicles, $n_{p\ max}$, on all roads is 20. The possibility of detecting anomalous vehicles when vehicles communicate with each other ϕ is set as 1. In addition, we assume all vehicles are ready to report the anomaly to the nearest RSU p with the total available reports $W = 9999$. The trust values of vehicles on the road are different and normally distributed with the expectation of 0.8 to show the advantages of our proposed mechanism and algorithms in a more intuitive and simplified manner. The game parameters are selected as $\phi = 2$, $\lambda = 3$ and monetary benefit parameter $z = 3$.

B. Performance evaluation

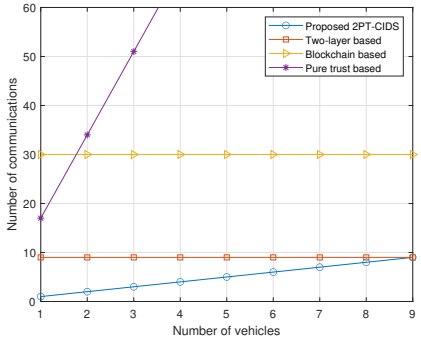

Fig. 3: Communication costs of various frameworks.

In Fig. 3, the minimum communication cost of different types of trust frameworks is shown. Since each framework transmits different content, we approximate the communication cost in terms of the number of communication hops. It can be seen that the communication cost of our proposed framework grows with the increase in the number of vehicles detecting abnormalities. This is due to the rise in the number of vehicles finding abnormalities and the rise in the number of vehicles reporting to the RSU. The number of increases depends on the number of vehicles reported to RSUs. In addition, 2PT-CIDS increases much less than the trust-based framework. Because it transfers the trust-based mechanisms' indirect trust evaluation requiring vehicles' multiple communications to RSUs performing. Our proposed two-layer-based framework also achieves lower communication costs relative to the conventional two-layer framework. Because of the deployment of CIDS and designed incentives, additional notification of other unreported vehicle participation in trust computing is not required.

Fig. 4 shows the relationship between the number of single-vehicle documents reported and the number of vehicles. We first assume all vehicles' trust values are fixed at 0.8 and 0.9 respectively. Moreover, μ in the conventional method [14] is the same as our proposed mechanism in case $\delta = 1$. In addition, to demonstrate the performance benefit of our mechanism, we assume that the vehicle in the conventional approach knows the number of vehicles sent to report even if it is not available to know it. It can be found that as the number of vehicles rises, the number of single vehicles reporting documents falls. Nevertheless, the proposed incentive mechanism declines more slowly at both trust values. Further, in contrast to the conventional approach, vehicles with a higher trust value transmit a higher number of reports. This is due to the presence of optimal dynamic incentives for variations in trust value and the number of vehicles in our scheme compared to the conventional method. This enables the RSU to get more valuable reports when the same number of vehicles are engaged in gaming. Our mechanism thus can improve anomaly detection accuracy.

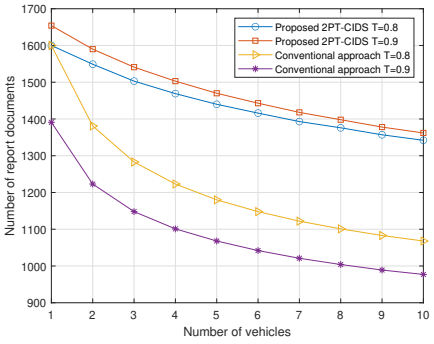

Fig. 4: Number of submitting reports from vehicle q in different numbers of vehicles.

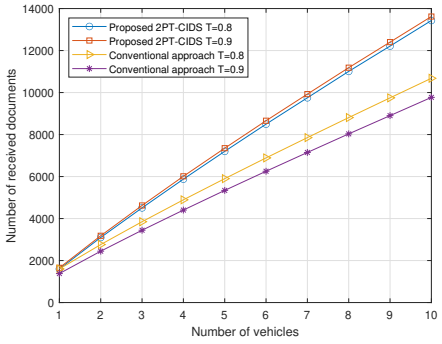

Fig. 5: Total number of documents received by RSU p in different numbers of vehicles (fined trust value).

Fig. 5 illustrates the relationship between the number of reports received by the RSU and the number of vehicles

TABLE II: Accuracy of BLI-based algorithm

polynomial degree	$l^{'}$, estimated value	l, search result	accuracy
3	1493	1600	93.31%
5	1619	1600	98.81%
10	1599	1600	99.94%

in the same situation as Fig. 3. It is observed that the number of documents received by the RSU increases with the number of vehicles. The proposed mechanism always obtains more amount and more valuable reported data at different trust values. This is also due to the existence of dynamic adaptations in our incentives and adapting to the environment to search for optimal decisions. This enables the RSU to assess vehicles more quickly and accurately. Furthermore, in Fig. 6, we randomly generate different trust values for the vehicles involved in the evaluation. The results are averaged over 20 experiments. As shown in Fig. 6, the same results as in Fig. 4 are still obtained in case the trust value is not fixed. Our mechanism always outperforms. These further demonstrate that our mechanism enables the VANET to improve anomaly detection accuracy.

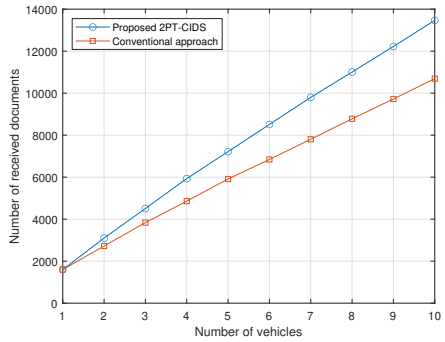

Fig. 6: Total number of documents received by RSU p in different numbers of vehicles (dynamic trust value).

Fig. 7 presents how monetary parameters influence the number of submitted reports. We set the trust value as 0.8. From Fig. 7, the number of vehicles submitting reports gradually decreases as χ increases. Furthermore, a larger λ provides an incentive for vehicles to submit more reports. Hence, the proposed mechanism's monetary parameters can be adjusted depending on the various deployment environments.

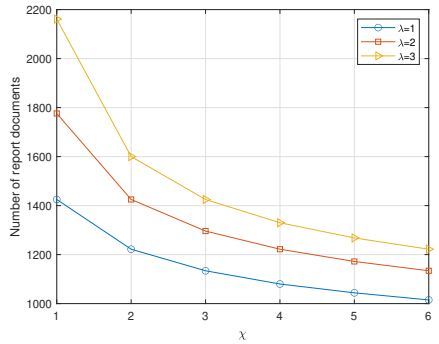

Fig. 7: Effect of different χ and λ on submitting reports.

In Table I, the effectiveness of the proposed BLI-based algorithm is demonstrated. We define the accuracy as the difference between the algorithm result $l^{'}$ and the iteration result l divided by l. Furthermore, the vehicle's trust value is set as 0.8, submitting reports the vehicle's number is set as 1 and $X = Y = 5$ and the result is rounded to the integer. It can be observed that as the polynomial degree (interpolation number) increases, the estimated value becomes increasingly close to the value derived from ten thousand searches. It is almost the same as the actual value in case the polynomial degree equals 10. Nevertheless, the computational number of our proposed BLI-based algorithm (complexity: $O(E^2)$) is much less than tens of thousands of searches (complexity: $O(N^2)$). Thus, our proposed solution is fast, accurate and efficient.

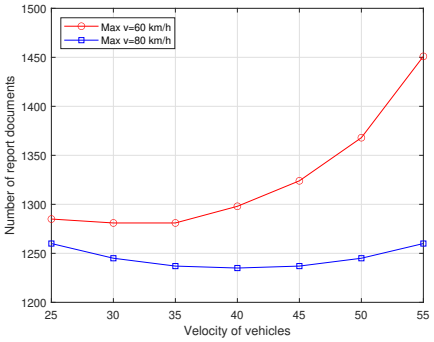

Fig. 8: Velocity versus reporting ratio.

In Fig. 8, we evaluate the relationship between the velocity of vehicles and the number of reports they submitted in case of vehicle's trust value is 0.8. Moreover, the congestion speed is set as 20 km/h. As shown in Fig. 8, although at different traffic speed limits, the number of documents reported tends to fall and then rise as the speed of the vehicle increases. This is due to the change in the number of vehicles on the road as their speed increases, resulting in changes in the chance of contact between vehicles. The number of documents reported thus changes as velocity and traffic conditions change.

Fig. 9 illustrates the advantages of our proposed reward-shaping D3QN with various baseline reinforcement learning methods (DQN, DDQN and D3QN) convergence speed. We set the $\gamma = 0.1$, $t = 3$ and the batch size is set as 128. For easy comparison, we count the total reward value every 100 episodes and calculate the average reward value for each step. Furthermore, to facilitate the comparison of convergence speeds between different algorithms, the average reward of the proposed algorithm has minus the additional reward value η_1 and normalised. It can be found that after a period of searching, all approaches eventually stabilise and converge. As a result of penalties given more than the expected range, the negative reward value occurs in our algorithm. It is worth noting that, our proposed reward-shaping D3QN algorithm reaches stability more quickly compared with typical reinforcement learning methods. Furthermore, after normalization and training, our proposed algorithm has the highest reward value, i.e., the highest utility of the vehicle. This is due to the presence of

reward-shaping, which encourages the vehicle to search within the optimal appropriate range. Therefore, our proposed method can obtain the optimal strategy more efficiently and rapidly.

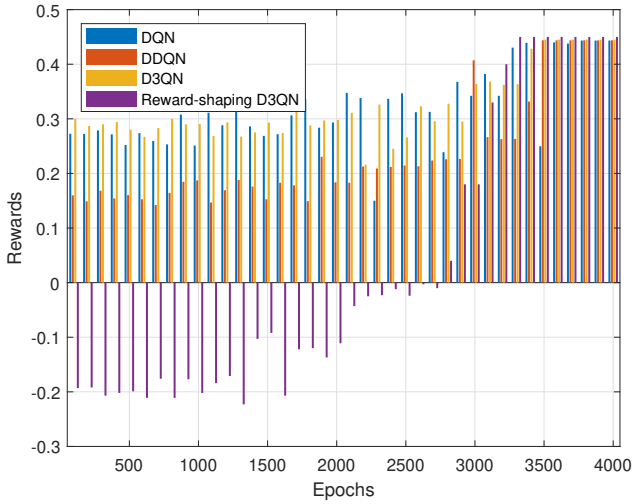

Fig. 9: Convergence speed and accuracy of different frameworks.

VII. Conclusions

In this paper, we proposed a novel privacy-aware 2PT-CIDS framework to improve anomaly detection. The 2PT-CIDS utilised the CIDS components to enhance vehicles' intrusion detection capabilities and compute vehicles' trust value by sending and verifying requests and challenges. It reduced the communication cost for trust computing. To enable the 2PT-CIDS to detect anomalous vehicles more accurately and efficiently, two types of game theories based incentive mechanism was proposed to encourage vehicles to report more messages to RSUs according to various traffic conditions, vehicle velocity, and time variations. A BLI-based algorithm and a reward-shaping D3QN algorithm are also presented to solve the different gaming problems. Numerical results demonstrated that our 2PT-CIDS and incentive mechanisms are highly effective. The proposed BLI and reward-shaping D3QN algorithms are also efficient and outperform the other baseline methods to achieve NE in various scenarios.

References

[1] G. Zheng, Q. Ni, K. Navaie, H. Pervaiz and C. Zarakovitis, "A Distributed Learning Architecture for Semantic Communication in Autonomous Driving Networks for Task Offloading," *IEEE Communications Magazine*, vol. 61, no. 11, pp. 64-68, November 2023.

[2] G. Raja, S. Anbalagan, G. Vijayaraghavan, S. Theerthagiri, S. V. Suryanarayan and X. -W. Wu, "SP-CIDS: Secure and Private Collaborative IDS for VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4385-4393, July 2021.

[3] Z. A. E. Houda, A. S. Hafid and L. Khoukhi, "MiTFed: A Privacy Preserving Collaborative Network Attack Mitigation Framework Based on Federated Learning Using SDN and Blockchain," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 4, pp. 1985-2001, 1 July-Aug. 2023.

[4] J. Cui and G. Sabaliauskaite, " US 2: An unified safety and security analysis method for autonomous vehicles ", *Proc. Future Inf. Commun. Conf.*, pp. 600-611, Apr. 2018.

[5] R. Xing, Z. Su, N. Zhang, Y. Peng, H. Pu and J. Luo, "Trust-Evaluation-Based Intrusion Detection and Reinforcement Learning in Autonomous Driving," *IEEE Network*, vol. 33, no. 5, pp. 54-60, Sept.-Oct. 2019.

[6] T. Nandy, R. M. Noor, M. Yamani Idna Bin Idris and S. Bhattacharyya, "T-BCIDS: Trust-Based Collaborative Intrusion Detection System for VANET," *2020 National Conference on Emerging Trends on Sustainable Technology and Engineering Applications (NCETSTEA)*, Durgapur, India, 2020.

[7] G. Raja, S. Anbalagan, G. Vijayaraghavan, S. Theerthagiri, S. V. Suryanarayan and X. -W. Wu, "SP-CIDS: Secure and Private Collaborative IDS for VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4385-4393, July 2021.

[8] J. Xiao, H. Chen and F. Zhong, "A Novel Feature Extraction Framework Using Graph Node Attention Network for In-Vehicle Networks Intrusion Detection," *IEEE Systems Journal*, vol. 18, no. 1, pp. 150-161, March 2024.

[9] R. Liu and J. Pan, "CRS: A Privacy-Preserving Two-Layered Distributed Machine Learning Framework for IoV," *IEEE Internet of Things Journal*, vol. 11, no. 1, pp. 1080-1095, 1 Jan.1, 2024.

[10] Y. Dai, D. Xu, S. Maharjan, Z. Chen, Q. He and Y. Zhang, "Blockchain and Deep Reinforcement Learning Empowered Intelligent 5G Beyond," *IEEE Network*, vol. 33, no. 3, pp. 10-17, May/June 2019.

[11] X. Zhang and X. Chen, "Data Security Sharing and Storage Based on a Consortium Blockchain in a Vehicular Ad-hoc Network," *IEEE Access*, vol. 7, pp. 58241-58254, 2019.

[12] J. Kang et al., "Blockchain for Secure and Efficient Data Sharing in Vehicular Edge Computing and Networks," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4660-4670, June 2019.

[13] C. Oham, R. A. Michelin, R. Jurdak, S. S. Kanhere, S. Jha, "B-FERL: Blockchain based framework for securing smart vehicles" *Information Processing & Management*, Volume 58, Issue 1, 2021.

[14] S. A. Siddiqui, A. Mahmood, Q. Z. Sheng, H. Suzuki and W. Ni, "Trust in Vehicles: Toward Context-Aware Trust and Attack Resistance for the Internet of Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 9546-9560, Sept. 2023.

[15] X. Chen, J. Ding and Z. Lu, "A Decentralized Trust Management System for Intelligent Transportation Environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 558-571, Jan. 2022.

[16] A. Mahmood, Q. Z. Sheng, Wei Emma Zhang, Y. Wang, and S. Sagar, "Towards a Distributed Trust Management System for Misbehavior Detection in the Internet of Vehicles,"*ACM Transactions on Cyber-Physical Systems*, vol. 7, no. 16, pp. 1-25, May 2023.

[17] R. Xing, Z. Su and Y. Wang, "Intrusion Detection in Autonomous Vehicular Networks: A Trust Assessment and Q-learning Approach," *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Paris, France, 2019, pp. 79-83.

[18] T. Zhang and Q. Zhu, "Distributed Privacy-Preserving Collaborative Intrusion Detection Systems for VANETs," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 1, pp. 148-161, March 2018.

[19] Y. Kuo, K. Lai, F. Y. Lin, Y. Wen, E. H. Wu and G. Chen, "Multirate Throughput Optimization With Fairness Constraints in Wireless Local Area Networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 5, pp. 2417-2425, Jun 2009.

[20] M. J. Khabbaz, W. F. Fawaz and C. M. Assi, "A Simple Free-Flow Traffic Model for Vehicular Intermittently Connected Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1312-1326, Sept. 2012.

[21] C. J. Fung, O. Baysal, J. Zhang, I. Aib and R. Boutaba, "Trust management for host-based collaborative intrusion detection", *Proc. Int. Workshop Distrib. Syst. Oper. Manage.*, pp. 109-122, 2008.

[22] G. Zheng, Q. Ni, K. Navaie and H. Pervaiz, "Semantic Communication in Satellite-borne Edge Cloud Network for Computation Offloading," *IEEE Journal on Selected Areas in Communications*.

[23] W. Li, Y. Wang, J. Li, M. H. Au, "Toward a blockchain-based framework for challenge-based collaborative intrusion detection" *the 1st International Workshop on Application Intelligence and Blockchain Security (AIBlock 2019) in conjunction with ACNS 2019*, pp. 122–139, 2019

[24] R. Mitchell and I. Chen, "Adaptive Intrusion Detection of Malicious Unmanned Air Vehicles Using Behavior Rule Specifications," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 5, pp. 593-604, May 2014.

[25] G. Zheng et al., "Mobility-Aware Split-Federated With Transfer Learning for Vehicular Semantic Communication Networks," *IEEE Internet of Things Journal*, vol. 11, no. 10, pp. 17237-17248, 15 May15, 2024.

[26] D. Fudenberg and J. Tirole, Game Theory. MIT Press, 1993.

[27] X. Kang and S. Sun, "Incentive mechanism design for mobile data offloading in heterogeneous networks," *2015 IEEE International Conference on Communications (ICC)*, London, UK, 2015, pp. 7731-7736.
[28] J.- P. Berrut and L. N. Trefethen, "Barycentric lagrange interpolation," *SIAM Rev.*, vol. 46, no. 3, pp. 501–517, 2004.
[29] Ö. Tutsoy and M. Brown, "Reinforcement learning analysis for a minimum time balance problem," *Transactions of the Institute of Measurement and Control*, vol. 38, no. 10, pp. 1186–1200, Jul. 2016.
[30] Ö. Tutsoy, "Graph theory based Large-Scale machine learning with Multi-Dimensional constrained optimization approaches for exact epidemiological modeling of pandemic diseases," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 8, pp. 9836–9845, Aug. 2023.
[31] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," 2015, arXiv:1509.06461. [Online]. Available: https://arxiv.org/abs/1509.06461
[32] Z. Wang, T. Schaul, M. Hessel, H. van Hasselt, M. Lanctot, and N. de Freitas, "Dueling network architectures for deep reinforcement learning," 2015, arXiv:1511.06581. [Online]. Available: https://arxiv.org/abs/1511.06581