A Sybil Attack Traceability Detection Scheme Adapted to Vehicle Pseudonym Change Strategy

***Abstract*—In Vehicular Ad hoc Networks (VANETs), Sybil attacks can persistently occur while vehicles are moving at high speeds, posing a serious threat to cooperative driving. Existing detection schemes primarily focus on identifying Sybil nodes while neglecting the traceability of the attackers, which hinders timely prevention of attacks and leads to further damage. Additionally, vehicles employ periodically changing pseudonyms to protect sensitive private information, enhancing the stealthiness of Sybil attacks and complicating detection. To address these challenges, we propose a scheme that integrates vehicle behavior with spatiotemporal features, combining the core characteristics of attack behavior with three-dimensional data regarding time, space, and communication within the traffic flow into graph-structured data. This approach emphasizes the abnormal behavior of attackers and comprehensively reveals the dynamic interactions among vehicles, enabling near real-time detection of Sybil nodes and accurate traceability of attackers. Experimental results demonstrate that our proposed scheme can accurately track over 92.7% of attackers on average, with superior detection and tracking capabilities compared to existing schemes, effectively preventing ongoing Sybil attacks in VANETs.**

***Index Terms*—Sybil attacks, VANET, traceability, vehicle behaviors, spatiotemporal features.**

I. Introduction

With the advancement of the Internet of Vehicles (IoV), there is a global trend towards the construction of Intelligent Transport Systems (ITS) that enable cooperation between vehicles and infrastructure. Cellular Vehicle-to-Everything (C-V2X) is pivotal for establishing connected vehicles and ITS, facilitating wireless communication among various entities. However, while this technology enhances traffic efficiency, it also introduces significant safety risks. The Sybil attack, which utilizes fake identities to send erroneous messages, poses a serious threat to driving safety. In current Vehicular Ad hoc Network (VANET) scenarios, vehicles communicate using pseudonyms provided by authorities as temporary identities to protect private information. Sybil attackers illegally collect or capture these legitimate pseudonyms to masquerade as virtual nodes, increasing the stealth of the attack and complicating detection and tracking.

Attackers exploit multiple fake identities to create ”ghost vehicles” that do not actually exist, deceiving other vehicles with false information. These ”ghost vehicles” are referred to as Sybil nodes. By transmitting numerous erroneous messages under false identities, Sybil nodes compel other vehicles to alter their driving status, potentially leading to severe traffic accidents. Additionally, Roadside Units (RSUs) are misled by inaccurate messages, resulting in incorrect assessments of the traffic situation, which severely threatens the security of the VANET. Consequently, the rapid and accurate detection of Sybil nodes and tracing the source of the attack has become an urgent research priority.

Current research has proposed detection schemes for Sybil attacks in vehicular networks. Yang et al. [1] designed a bidirectional anonymous authentication and key agreement scheme based on elliptic curves, enhancing the efficiency of authentication through lightweight hashing and XOR operations. However, due to the need for vehicles in VANETs to frequently change temporary identities in short cycles, the key exchange process still incurs significant computational overhead, consuming resources from Roadside Units (RSUs). In realistic traffic scenarios, RSUs are responsible for basic traffic management functions, and the remaining computational resources are often insufficient to meet system demands. Devika et al. [2] proposed an anomaly detection framework based on Generative Adversarial Networks (GANs) called VADGAN, aimed at detecting abnormal behaviors in connected vehicles. Although deep learning frameworks demonstrate strong detection performance, they require substantial computational resources, making it difficult to achieve rapid and accurate detection, especially in scenarios with high traffic density.

Most existing detection schemes focus on detecting Sybil nodes and ignore attacker tracking, resulting in attacks that cannot be stopped on time. Therefore, the method proposed in this paper must address the following two **Key Questions** in a targeted manner, as shown in Fig. 1.

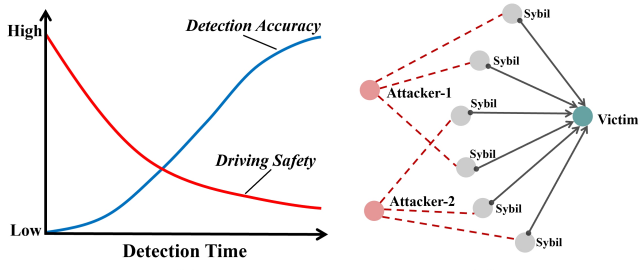

Fig. 1. Schematic of the Core Issues to be Addressed.

Q1. How can Sybil node detection schemes balance high accuracy and fast detection in scenarios where vehicles use pseudonymous communication?

Q2. How can an attacker be precisely traced in scenarios where vehicles communicate with each other using pseudonyms?

The **contributions** of this study, addressing the above questions, are presented below.

ISBN 978-3-903176-72-0 © 2025 IFIP

1) In this study, we proposed a Sybil attack detection and tracing scheme based on the fusion of vehicle behavior and spatiotemporal features. Using the fusion of the vehicle-driving state and spatiotemporal information related to traffic flow, the abnormal behavior of the attacker can be rapidly identified, and the detection and tracing of the attack can be achieved with the limit of detection data of only two messages.
2) This paper presents a method for constructing a spatiotemporal graph of traffic flow. It combines vehicle relationships in three dimensions (i.e., time, space, and communication) into graph-structured data and utilizes the subgraph structure between vehicle nodes to uncover their interactions. This approach enables the accurate tracing of a attacker, even when all vehicles communicate through pseudonymous identities.
3) In this study, experiments were conducted using the publicly available VeReMi Extension real-world simulated traffic dataset. The proposed scheme in this study achieves near real-time Sybil node detection capability with periodic changes in vehicle pseudonyms. The average recall rate of tracking attackers is as high as 92.7%, which is a significant improvement over existing schemes.

II. Related Work

Sybil attacks pose a serious threat to the security of cooperative driving, and many researchers have focused on Sybil attacks in VANETs [3]–[5]. The analysis of existing schemes has concentrated on two areas: Sybil attack detection and tracking of the attacker, to clarify the advantages and shortcomings of different techniques. In this study, we present a research methodology that integrates the features of attack behaviors.

A. Detection Schemes for Sybil Attacks

Distinguishing between Sybil and real vehicles is the key to accomplishing the detection task, as current detection methods focus on detecting fake Sybil nodes created by attackers, which do not exist in real scenarios. They can communicate with real vehicles and utilize error messages to impact the driving of other vehicles.

In the wireless communication environment of VANETs, detection methods based on RSSI values do not require additional communication equipment and offer the advantages of lower communication costs and reduced latency. This capability enables the detection of Sybil nodes in vehicle-to-vehicle communication networks [6]. Li [7] proposed a Sybil detection technique that relies on RSSI sequences and vehicle traveling matrices (RSDM), assessing the differences between RSSI sequences and traveling matrices through dynamic distance comparisons. However, these detection methods are affected by environmental factors in real-world scenarios, making it difficult to adapt to high-density traffic flows and accurately locate Sybil nodes.

Another class of schemes [8] employs a reputation evaluation approach, generating a reputation score for each vehicle. Engoulou et al. [9] defined a set of locally aware behavioral reputation parameters to enable the distributed evaluation of vehicle reputation and exclude malicious vehicles. However, if the proportion of attackers is high, reliable evaluation metrics will not be available, undermining the reliability of the reputation system.

Data-centric systems based on communication messages can learn and adapt to new attack patterns. Hammi et al. [10] utilized Cooperative Awareness Messages (CAMs) sent by vehicles as a database, transforming these messages into feature vectors and classifying them using a machine learning model. This class of methods [11], [12] can effectively identify Sybil nodes; however, the detection effectiveness depends on the quantity of training data and is constrained by limited computational resources. The detection performance of the latest schemes [13] and [14] is limited by the duration of the accumulated data; the former experiences a decline in detection performance as the duration increases, while the latter requires more than 300 seconds of accumulated data to achieve acceptable detection precision. The experimental section of this paper compares these two schemes.

B. Traceability Schemes for Attackers

Current strategies for protecting vehicle identity privacy complicate the detection of attackers. Methods such as pseudonym exchange, identity information verification [15], and vehicle dynamic monitoring require significant computational resources while concealing the identities of attackers, leading to an inability to prevent attacks. Distributing substantial computation across multiple fog nodes presents a viable solution. Wazid [16] proposed an authenticated key management protocol for fog computing-based Internet of Vehicles (AKM-IoV), which generates system parameters using nonsingular elliptic curves to achieve two-way authentication between vehicle-to-fog (V2F), roadside-to-fog (R2F), and fog-to-cloud (F2C) network entities.

This class of methods [17]–[19] effectively reduces the computational resources required on the vehicle side; however, under the vehicle pseudonym strategy, it fails to address the issues of frequent pseudonym changes and the use of multiple false identities, thereby hindering the ability to detect the attackers.

III. Proposed Methods

In this section, we propose a detection and traceability scheme based on the fusion of vehicle behavior and spatiotemporal features, which consists of three parts: vehicle behavioral feature construction **(To solve Q1)**, traffic flow spatiotemporal graph construction **(To solve Q2)**, and feature fusion classification tracing, as shown in Fig. 2.

A. Vehicle Behavioral Feature Construction Method

This scheme leverages the constraints and delays in vehicle-following theory to aggregate driving behavior features, enabling rapid differentiation between Sybil nodes and real

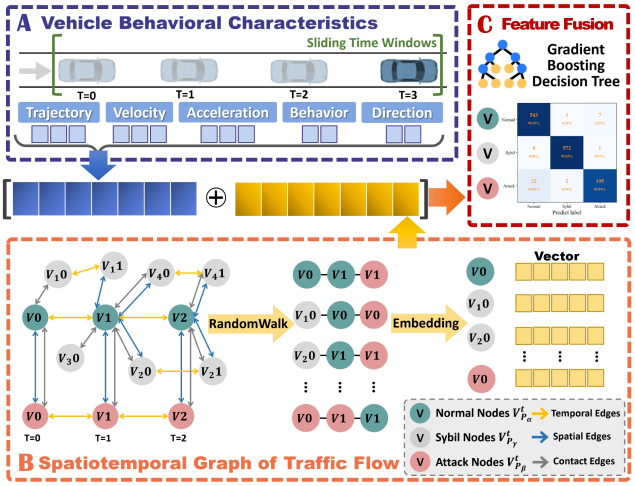

Fig. 2. Detection and Traceability Method Overview.

vehicles. The approach constructs behavioral features from the perspective of vehicle trajectory and speed changes, demonstrating high adaptability and reducing the influence of road types and traffic flow densities on detection performance. The specific construction method is detailed below.

We defined the set of Basic Safety Messages(BSM) received by the RSU in the target detection area and transmitted to each other by the vehicles as the basic data $D_{BSM} = \{M_1, M_2, \cdots, M_n\}$, where M is the BSM messages in the dataset. The driving state information contained in the messages is as follows:

$$M = \begin{pmatrix} TimeStamp, Pseudonym, Position, \\ Speed, Acceleration, Heading \end{pmatrix}$$

$TimeStamp$ refers to the time at which the message is sent, while $Pseudonym$ is a temporary identity that the vehicle can change periodically. $Position$, $Speed$, and $Acceleration$ represent the vehicle's driving information, and $Heading$ indicates the current direction of travel.

First, the vehicle's $Pseudonym$ and the corresponding set of messages need to be obtained from D_{BSM}. Then, based on the order of $TimeStamp$, a sliding time window method is applied to obtain a set of consecutive messages broadcast by the vehicle, with the time window length set to W seconds.

In this study, we use a vehicle with pseudonym $Pseudonym_\alpha$ (abbreviated as P_α) as an example to obtain its broadcast message set D_{P_α}. Subsequently, we employ the sliding time window method to extract the message data over the period $[t, t+W]$, generating a behavioral feature vector as the basic detection unit. In **Part-1**, five different classes of features will be described in terms of both construction methods and concrete implementations.

1) Vehicle path track features: By analyzing vehicle trajectory data, the following conclusions can be drawn: 1. The change in vehicle position is closely related to the driving path. 2. The moving distance of the vehicle within the time window is affected by different road conditions. 3. The spatial position of the vehicle can be described using the geometric center coordinates of its trajectory. Therefore, we find that the trajectory characteristics of Sybil nodes differ from those of real vehicles, particularly reflected in metrics such as the average moving distance.

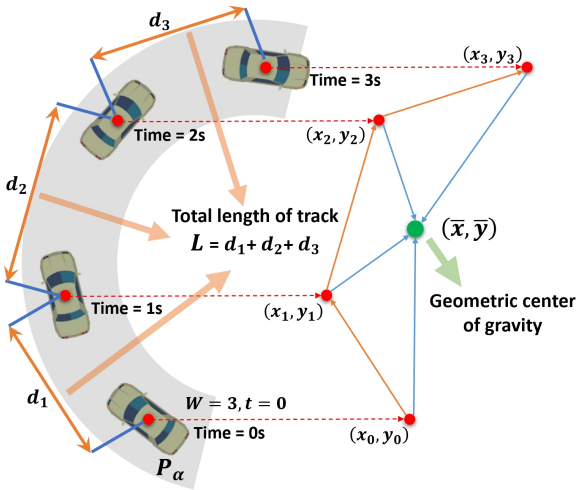

Fig. 3. Diagram of the Path and Center of Gravity.

Total length of path: As shown in Fig. 3, we need to calculate the total distance traveled by the vehicle P_α within the time window $[t, t+W]$. First, we obtained the message set D_{P_α} and extracted the position coordinates corresponding to each moment $(x_0, y_0), (x_1, y_1), \cdots, (x_W, y_W)$. Then, the distance traveled by the vehicle under the adjacent moments is $d_k = \sqrt{(x_k - x_{k-1})^2 + (y_k - y_{k-1})^2}$, $k \in [t+1, t+W]$. Finally, the total length L of the traveled path in the current time window $[t, t+W]$ is calculated as follows.

$$L = \sum_{k=t+1}^{t+W} d_k \tag{1}$$

Average distance traveled: The value of this feature is the average distance $\bar{L} = \frac{L}{W}$ traveled by a vehicle with pseudonym P_α in the same time window $[t, t+W]$.

Coordinates of the geometric center of the trajectory: The track of the vehicle matches the shape of the real road. The difference between the coordinates of the track's geometric center point can be used to distinguish the specific road on which the vehicle is traveling, as shown in Fig. 3. We take a curve path as an example and use the continuous track coordinates $(x_0, y_0), (x_1, y_1), \cdots, (x_W, y_W)$ to compute the geometric centroid coordinates $(\bar{x}, \bar{y})$ as shown in Eq.(2).

$$(\bar{x}, \bar{y}) = \left(\frac{1}{W+1} \sum_{j=t}^{t+W} x_j \ , \ \frac{1}{W+1} \sum_{j=t}^{t+W} y_j \right) \tag{2}$$

2) Vehicle speed change features: Changes in vehicle speed can accurately describe the motion state. This method utilizes three speed sequence features to characterize Sybil attack behavior for identification purposes.

1. Average Speed: There are significant differences in average speeds of vehicles on different roadways, which can be used to distinguish vehicles on different roads and reduce the interference of road factors in the learning process of behavioral features.

2. Maximum Speed: Sybil nodes frequently change their position and speed according to their attack intent, resulting in

abnormal high speeds appearing in continuous BSM messages for short periods, as shown in Fig. 4.

3. Variance of Speed Difference: According to the theory of single-lane vehicle following, when vehicles are within 100-125 meters of each other, they typically influence one another, resulting in group braking or acceleration. Sybil nodes lack the awareness and control of a real driver, leading to significant differences in their speed change patterns compared to real vehicles, which can help distinguish the two.

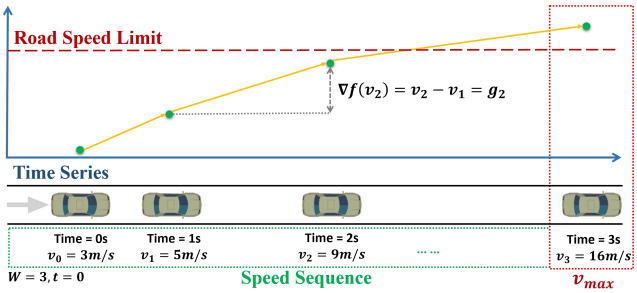

Fig. 4. Vehicle Speed Sequence and Maximum Instantaneous Speed Schematic.

The following are the specific computational steps.

Average speed: First, we capture the message information during the time window $[t, t+W]$ of D_{P_α}. We collect the instant velocities $v_0, v_1, v_2, \cdots, v_W$ at each moment and calculate the average velocity $\bar{v}$.

$$\bar{v} = \frac{1}{W+1} \sum_{j=t}^{t+W} v_j \tag{3}$$

Maximum speed: In the instantaneous velocity sequence at each of the above moments, the maximum instantaneous velocity of the vehicle in $[t, t+W]$ is extracted as the maximum velocity $v_{max} = max\left(v_0, v_1, v_2, \cdots, v_W\right)$.

Variance of the difference of the speed change series: Similarly, using the sequence of instantaneous velocities $v_0, v_1, v_2, \cdots, v_W$, the differentials of the instantaneous velocities are calculated.

$$g_k = v_k - v_{k-1}, k \in [t+1, t+w] \tag{4}$$

The difference series $g = [g_1, g_2, \cdots, g_W]$ is obtained, where $\bar{g}$ is the mean of the difference series, and the variance Var_v of the difference series is calculated to obtain the volatility index of the speed series change, as shown in Eq.(5).

$$Var_v = \frac{1}{W} \sum_{k=t+1}^{t+w} \left(g_k - \bar{g}\right)^2 \tag{5}$$

3) Vehicle acceleration change features: Acceleration is a measure of how rapidly a vehicle's speed changes over time. It can reflect changes in driving conditions and behavioral characteristics. We used the acceleration sequences $a_0, a_1, a_2, \cdots, a_W$ at each instant in $[t, t+W]$ to calculate the difference of its acceleration change $\Delta a = a_k - a_{k-1} = h_k$ and calculate the difference in acceleration over a period of time. Therefore, this method adopts the same calculation method as that used for calculating the characteristics of the velocity sequences and obtains the eigenvalues of the three acceleration sequences as follows:

Average acceleration:

$$\bar{a} = \frac{1}{W+1} \sum_{k=t}^{t+w} a_k \tag{6}$$

Maximum acceleration:

$$a_{max} = max\left(a_0, a_1, a_2, \cdots, a_W\right) \tag{7}$$

Acceleration changes differential variance:

$$Var_a = \frac{1}{W} \sum_{k=t+1}^{t+w} \left(h_k - \bar{h}\right)^2 \tag{8}$$

4) Vehicle driving behavior features: Analysis of the Sybil attack model reveals that an attacker generally maintains a significant distance from the fake Sybil node to hide their attack intent and reduce the risk of exposure. Consequently, roadside units located at the edge of the effective communication range are unable to consistently receive messages from Sybil nodes, while messages from normal vehicles remain uninterrupted. This method utilizes the occurrence of message interruptions to identify abnormal behavior of Sybil nodes.

The number and duration of interruptions were calculated as follows.

Number of interruptions: In a real communication scenario, a vehicle transmits information regarding its driving status to the outside world at a certain time interval Δt. In this study, we set $\Delta t = 1$s. The time window $[t, t+W]$ is sliced into units of 1s such that one message should be sent at each sliced moment. If no message is sent at a given moment, it is considered an interruption. We aggregated the moments in which no message is sent into a set S_τ and counted the number of time slots in S_τ (i.e., the number of vehicle trajectory interruptions $K = |S_\tau|$).

Interruption duration: The total interrupt time T_τ is obtained by summing all the time periods in the above set S_τ for which no message has been sent.

5) Vehicle direction of travel features: Analysis of traffic flow data indicates that under low traffic density, vehicles do not frequently change their driving direction and move smoothly. However, during a Sybil attack, a large number of Sybil nodes that obstruct traffic appear on the road, forcing normal vehicles to frequently change lanes or turn.

In this construction method, the message data in the time window $[t, t+W]$ are first intercepted from D_{P_α} and the sequence of vehicle travel direction vectors at each instant is extracted $\overrightarrow{u_0}, \overrightarrow{u_1}, \overrightarrow{u_2}, \cdots, \overrightarrow{u_W}$, $\overrightarrow{u_j} = (X_j, Y_j), j \in [t, t+W]$. Then, the vehicle direction vector $\overrightarrow{u_j}$ is transformed into a polar coordinate representation (ρ_j, θ_j), where $\rho_j = \sqrt{2}$ and θ_j is expressed in the equation $\theta_j = \text{atan2}\left(Y_j, X_j\right)$.

Finally, the driving operation of the vehicle is expressed through changes in angle: straight or turning. The calculation process is as follows.

Mean value of direction of travel: In the time window $[t, t+W]$, the vehicle driving direction is obtained to form a sequence of angle values $\theta_0, \theta_1, \cdots, \theta_W$, and then the average value of the sequence is calculated. Finally, the average traveling direction of the vehicle is obtained.

$$\bar{\theta} = \frac{1}{W+1} \sum_{j=t}^{t+W} \theta_j \tag{9}$$

Number of turns while driving: To calculate the number of times the vehicle steers in the time window $[t, t+W]$. First, define a set of S_η moments. Each of these moments satisfies $|\theta_{j+1} - \theta_j| > \theta_\eta$ and $\theta_\eta = 70$. This means that there is a large change in the direction of travel at this moment. Then, count how many moments $P_\eta = |S_\eta|$ meet the above conditions.

Based on the behavioral features of vehicles for detection and classification, we have distinguished between Sybil nodes and real vehicles, and these features can be constructed rapidly, enabling the model to have near real-time detection capability. However, relying solely on the behavioral information of individual vehicles makes it difficult to trace the identity of the attacker. Therefore, this method utilizes the data of messages sent and received among vehicles to construct a spatiotemporal graph of traffic flow, analyzing the associations among multiple vehicles from a holistic perspective to achieve accurate tracking of the attacker.

B. Traffic Flow Spatiotemporal Graph Construction Method

In **Part-2**, we focus on the difficulty of accurately tracking attackers. Attackers have behavioral characteristics similar to normal vehicles, and all vehicles communicate using pseudonyms, which increases the difficulty of tracking. This method combines the time, space, and communication network structure contained in the traffic flow, analyzing individual vehicles as part of the overall traffic flow. The proposed method for constructing a spatiotemporal graph of traffic flow can effectively distinguish between attackers and normal vehicles, and it is divided into three stages: preprocessing of traffic flow data, construction of the traffic flow spatiotemporal graph, and learning of node embedding representations.

1) Traffic flow data pre-processing: This method introduces the mapping relationship R of pseudonym identities for received and sent messages within the communication atmosphere between vehicles as the foundational data, which is used to construct the spatiotemporal traffic flow graph G_{TS} of the vehicle communication network structure.

$$R = \begin{pmatrix} SendTime, SenderPseudonym, \\ ReceiverPseudonym, Distance \end{pmatrix}$$

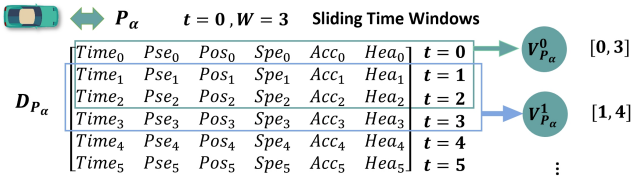

Fig. 5. Sliding Window Interception and Node Construction Schematic.

2) Construction of spatiotemporal maps of traffic flow: To achieve accurate tracing, this method constructs graph structure data related to time, space, and communication, and transforms it into a node classification task. The specific descriptions of node design and edge construction are as follows.

In the second part, this method uses the same sliding time window data of length W as in the first part. Using vehicles with pseudonyms $Pseudonym_\alpha$ and $Pseudonym_\beta$ as examples, abbreviated as P_α and P_β, respectively, the set of messages corresponding to pseudonym P_α is obtained from D_{BSM} as D_{P_α}. This study describes the graph construction method in a stepwise manner.

Node design

For vehicle P_α, message data intercepted within the time interval $[t, t+W]$ is recorded as the driving state node $V^t_{P_\alpha}$, and the node label is $L^t_{P_\alpha}$. By sliding the time window to segment the message data, a node sequence $V^0_{P_\alpha}, V^1_{P_\alpha}, V^2_{P_\alpha}, \cdots, V^m_{P_\alpha}$ is formed, referred to as the node set V_{nodes}, where m is the number of sliding windows. Figure 5 illustrates the node construction process.

Construction of temporal edges

Based on the unified timeline in the scene, the temporal edges $ET^t_{P_\alpha} = \left(V^t_{P_\alpha}, V^{t+1}_{P_\alpha}\right)$ are constructed for the node set of vehicle P_α, starting from $V^0_{P_\alpha}$. Here, T represents the temporal sequence. The collection of temporal edges for all vehicle nodes is denoted as ET, as shown in Fig. 6.

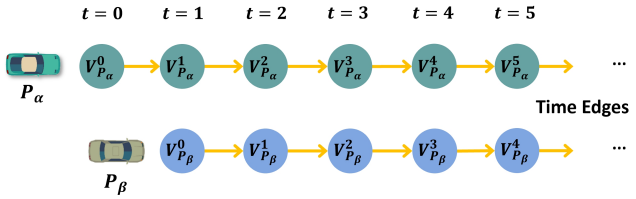

Fig. 6. Vehicle Temporal Edge Schematic.

Construction of spatial edges

The spatial edge reflects the spatial distance relationship between vehicles and the influence of neighboring vehicles during driving. In the attack region, the spatial edge helps detect the congregation of Sybil nodes and abnormal distribution patterns.

The spatial edge $ED^t_{P_\alpha}$ is constructed based on the spatial relationships of vehicle nodes in the current scene. The criterion for construction is the distance $d_{position}$ between two vehicle nodes within the time window $[t, t+W]$. If the distance is smaller than the threshold, a spatial edge is created. The set of spatial edges is denoted as ED, and its relationship with temporal edges is shown in Fig. 7.

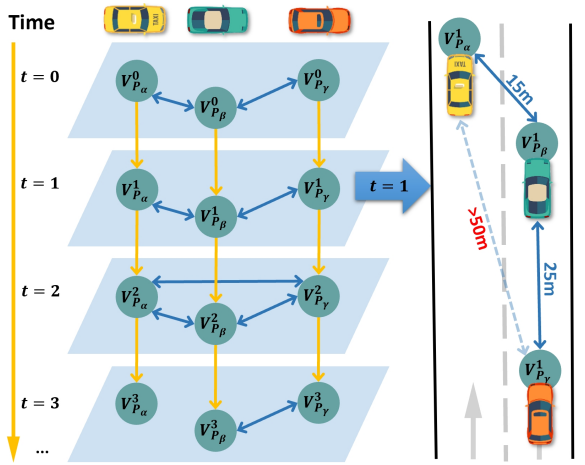

Fig. 7. Schematic Diagram of Vehicle Distance from Side Construction.

Construction of contact edges

The contact edges utilize the collected message data from vehicle nodes that send and receive communications to establish the communication relationships between vehicles. In the traceability analysis of Sybil attacks, contact edges can display unique communication patterns and behaviors between Sybil nodes and attacker nodes, highlighting the distinctions between attackers and other normal vehicles, and assisting in the precise identification of the attackers.

Both vehicles P_α and P_β contain multiple temporal nodes. The decision to construct the contact edge $EC^t_{P_\alpha P_\beta}$ is based on the communication messages sent or received between the two vehicles, directed from node $V^t_{P_\alpha}$ to the receiving node $V^t_{P_\beta}$.

Ultimately, the three different types of edges are combined to form the edge set $E_{TDC} = ET \cup ED \cup EC$, completing the construction of the spatiotemporal traffic flow graph $G_{TS} = (V_{nodes}, E_{TDC})$, as shown in Fig. 8.

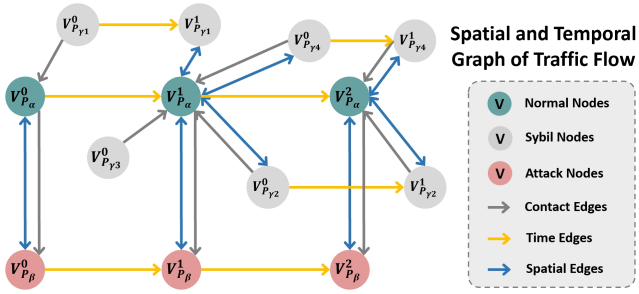

Fig. 8. Schematic Diagram of Traffic Flow Space and Time.

Node Embeddings and Representation Learning

The spatiotemporal graph has significant attributes and structures that enhance the connection between Sybil nodes and their attack source nodes, aiding in tracking the attackers.

Sybil nodes are fabricated and cannot provide an accurate list of messages; they only have temporal edges pointing to the current node from the previous moment, with an in-degree of 1. The attacker nodes do not receive messages from their fabricated Sybil nodes, and this unique subgraph structure simplifies the tracking of the identity and location of the attackers. In practical scenarios, accumulating data over a longer period to construct a larger graph helps in accurately tracing the identity of the attackers.

This approach employs the node2vec algorithm to obtain the embedded representation of each node. The algorithm utilizes a biased random-walking strategy controlled by two hyperparameters p and q, allowing adjustments to control the acquisition of the graph's homogeneity or global structure.

C. Summary of Feature Fusion Methods

The two methods proposed in this paper are summarized in this section. In **Part-1**, the construction of vehicle behavioral features for the rapid detection of Sybil attacks achieved nearly real-time attack detection performance **(Experiment 1)** with a detection window much smaller than the pseudonym change period **(To solve Q1)**. Meanwhile, according to the analysis of the following experimental results **(Experiment 2)**, in the scenarios with different traffic densities, the traceability of this method to the attacker is weak. It cannot accurately distinguish between attackers and normal driving vehicles, and the traceability of attackers needs to be further improved.

In **Part-2**, we propose the construction of a traffic flow spatiotemporal graph and an extraction method for node embedding representation features. By integrating the node features extracted from the vehicle behavior and traffic flow spatiotemporal relationship, the fusion feature vector of the target node was obtained, and the gradient boosting decision tree algorithm in the machine learning model was selected to complete the learning and classification. In the case where all vehicles use pseudonyms to communicate **(To solve Q2)**, with the excellent generalization ability of the gradient boosting decision tree algorithm, we obtain better traceability accuracy for the attacker **(Experiment 3)** and explore the impact on the scheme detection performance after employing different data sizes for training to obtain the appropriate data size to meet the detection requirements **(Experiment 4)**.

IV. Experiment Results

A. Experimental Preparation

1) Experimental data: This experiment is based on the GridSybil attack data from the open-source misbehavior dataset VeReMi extension dataset, generated by the misbehavior detection framework (F2MD) [20]. In this paper, all vehicles are set to communicate using pseudonyms, and the dataset is adjusted according to current regulations regarding the frequency of pseudonym changes [21] [22]. From the moment a vehicle enters the scenario, it automatically switches to a new pseudonym for communication every 60 seconds.

2) Evaluation metrics: In the experiments, the proposed scheme's performance in detecting Sybil nodes and tracing attackers was validated. The data was categorized into three classes: normal, Sybil, and attackers. Attacker vehicles are labeled as attackers only when they execute Sybil attacks; otherwise, they are labeled as normal. The experiments used $Accuracy$ to evaluate the overall performance of the model, while $Precision_{Sybil}$ and $Recall_{Attack}$ (shown in Fig. 9) were employed to measure the detection performance for Sybil nodes and the traceability of attackers.

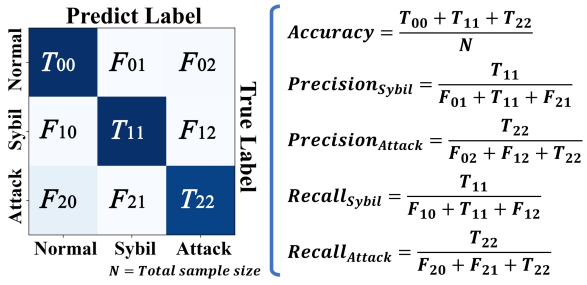

Fig. 9. Evaluation Indicator Formula.

3) Pre-experimentation: In the preliminary experiments, this study tested six commonly used machine learning algorithms. Based on the experimental results, the gradient boosting decision tree (GBDT) algorithm model was selected as the most suitable model for attack detection and traceability scenarios for the experiments.

B. Experimental Results and Analyses

Experiment 1: Real-time detection capability based on vehicle behavioral features for Sybil node.

To test the detection capability of the Sybil attack detection method based on vehicle behavioral characteristics (**Part-1**) in a shorter time window, we used 11 sliding time windows of different sizes in our experiments to assess the detection speed. The results of the experiments are shown in Figs. 10-11.

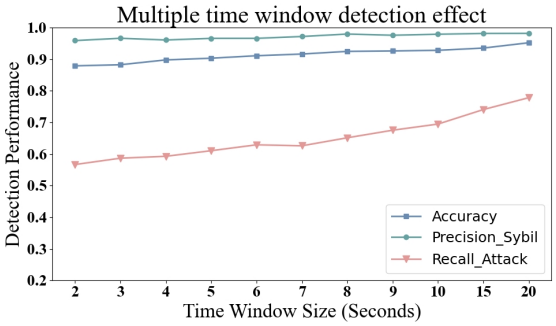

Fig. 10. Results for Different Time Windows.

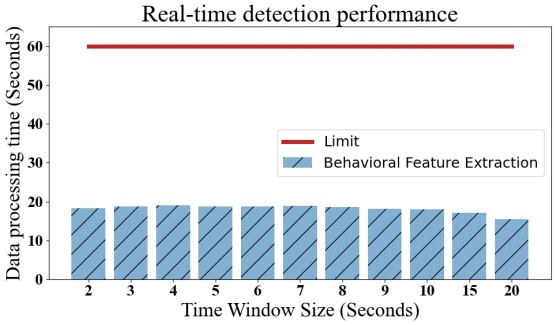

Fig. 11. Behavioral Feature Extraction Time.

In the results shown in Fig. 10, the proposed method achieved high detection precision and overall accuracy for Sybil nodes across different sizes of sliding time windows. Additionally, the detection performance improved with an increase in the time window size. However, shorter time windows are favorable for achieving rapid detection. When the detection time window was set to 2 seconds, the detection precision for Sybil nodes still exceeded **95%**, and the processing speed of the raw data was significantly higher than that of vehicle-generated data, as shown in Fig. 11. In the experimental data, the 2-second time window contained only **2 messages**, indicating that this method can almost satisfy all pseudonym change scenarios. This demonstrates that the method based on vehicle behavioral characteristics can achieve near-real-time detection speeds while maintaining high detection accuracy, addressing the challenge of balancing detection accuracy with real-time performance.

Experiment 2: Impact of different traffic flow densities on the performance of vehicle behavioral feature detection methods.

In this study, we used traffic flow data from different time periods to test **Part-1** of the vehicle behavior feature construction method in the proposed scheme. A sliding time window of 10s was applied, and the test results are shown in Fig. 12 indicate that the detection accuracy and precision of Sybil nodes were maintained at a high level. Fig. 12 contains the corresponding traffic flow densities for different time periods.

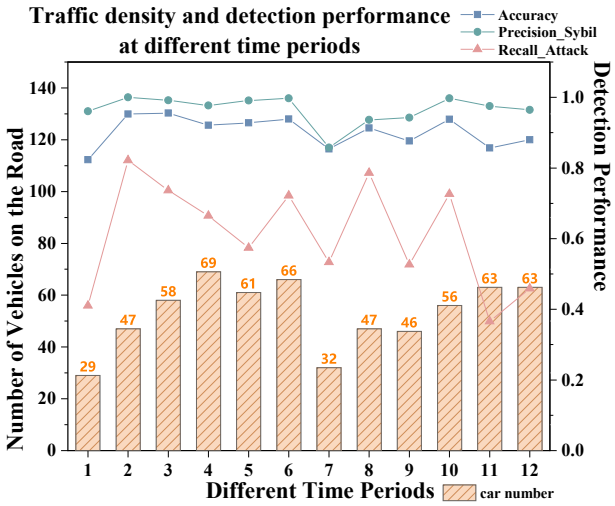

Fig. 12. Traffic Flow Density and Detection Performance under Different Time Periods.

According to the results, the proposed method achieved an average accuracy of 97.24% under different traffic flow conditions (12 time periods), demonstrating good generalization ability. However, as shown in Fig. 12, the method has poor traceability to attacker vehicles, with traceability recall significantly affected by variations in traffic flow, resulting in large fluctuations. At periods 1 and 7, the model encountered difficulties in distinguishing between attackers and normal vehicles due to the limited number of vehicles. Therefore, we considered the behavioral characteristics of attackers and proposed a method in **Part-2** to construct association features between attack behaviors and vehicle identities.

Experiment 3: Traceability based on fused features of traffic flow spatiotemporal graphs.

After obtaining the structural features of the spatiotemporal traffic flow graph, we combined them with vehicle behavioral features to create fusion features. Subsequently, we evaluated the detection and traceability capabilities of the proposed technique under different detection time window lengths. Fig. 13 to 15 display the experimental results, where N, S, and A represent normal vehicles, Sybil nodes, and attacker samples, respectively. The numbers 2-20 denote the time window sizes, Truth Label indicates the true labels of the samples, and Behavioural and Fusion represent the results of detection using behavioral features and fusion features, respectively.

Fig. 13 shows the differences in detection results between the vehicle behavioral features and fusion features in the proposed scheme. The detection results of the fusion features show significant improvement, with almost all Sybil nodes and attackers being accurately classified. The experimental results indicate that using behavioral features can effectively detect Sybil nodes, as shown in Fig. 14. However, there is no significant improvement in the model's detection capability from the "Behavioral" features to the "Fusion" features. In contrast, Fig. 15 illustrates a noticeable enhancement in the traceability capability of the fusion features, facilitating the swift tracing

of attackers. The average increase in the traceability recall rate exceeded 22.4%, demonstrating the exceptional traceability capability of this scheme.

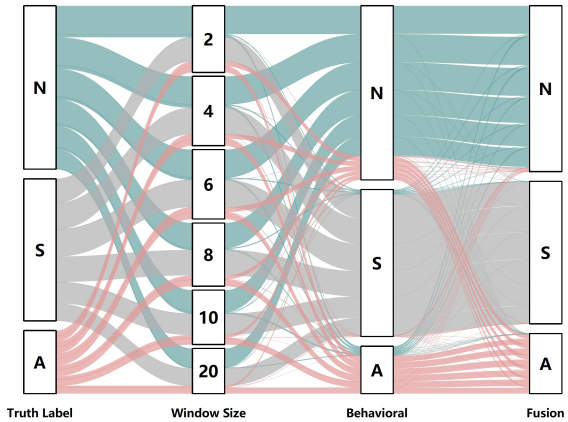

Fig. 13. Detection and Traceability of Results.

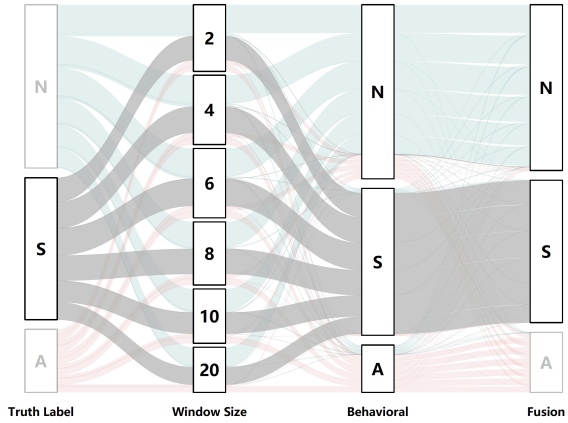

Fig. 14. Sybil Node Detection Results.

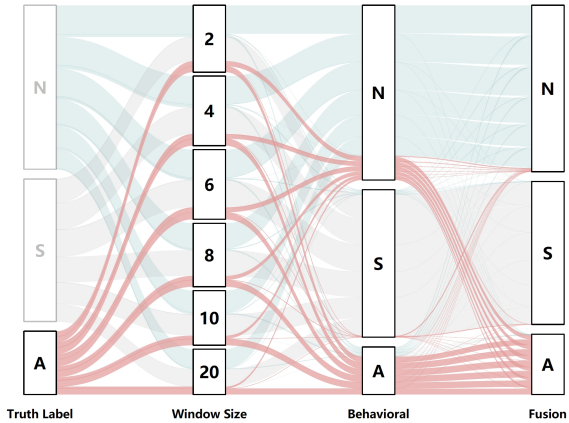

Fig. 15. Attacker Traceability Results.

Experiment 4: Influence of different accumulated data lengths on the fusion feature detection effect. The longer the accumulation time of the traffic flow spatiotemporal graph data, the more nodes and edges the graph contains. Therefore, this experiment set different data accumulation times to construct graph structure data of varying scales to test **Part-2** of the proposed scheme, with results shown in Fig. 16.

The experimental results indicate that as the accumulation time increases, the overall detection accuracy tends to stabilize,

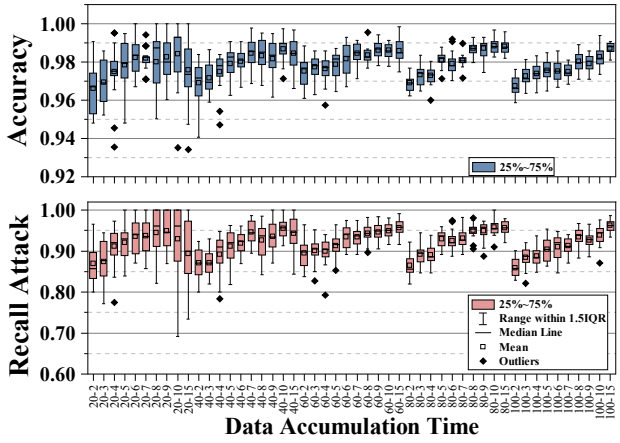

Fig. 16. The Effect of Different Graph Sizes on the Traceability of Detection. In This Experiment, We Chose Five Different Cumulative Time Lengths for the Data (i.e., 20, 40, 60, 80, and 100), and Sliced Them Using Different Sliding Time Window Sizes From 2s to 15s (e.g., the Label 20-2 for the Horizontal Coordinate Means That the Data With a Total Length of 20s Is Sliced Using a Sliding Time Window of 2s).

and the recall rate for tracing attackers remains at a high level. The experiment demonstrates that when the data accumulation time exceeds 60 seconds, the proposed scheme can achieve stable detection performance and accurately trace the identity information of attackers.

Experiment 5: Comparison with other existing detection schemes.

In the comparison experiment, we selected two existing attack detection schemes to compare with the scheme proposed in this paper [13] [14]. In [13], a marine predator algorithm based on teaching-learning was proposed, which constructs a multilayer fully connected neural network for training and is capable of identifying hidden attack characteristics against known and unknown threats. The other scheme [14] utilizes the unique sending source characteristic of BSM packets to detect and trace Sybil attacks without the need for training a machine learning model. Based on the reproduction of the above schemes, this paper conducted training and testing on a public dataset, with comparison results shown in Table I.

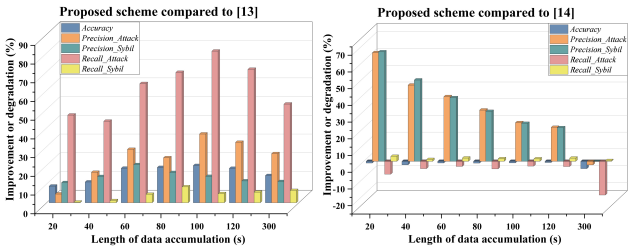

Fig. 17. Performance Gap Compared to Schemes [13] and [14].

This experiment utilizes the performance gap approach to clearly present the comparative results with schemes [13] and [14]. Fig. 17 shows that in comparison with [13], the proposed scheme in this paper comprehensively outperforms that scheme. The advantage lies in the further construction and extraction of features from BSM message data, which can more accurately reveal the anomalous behavior of attackers and is less affected by pseudonym strategies. In comparison

TABLE I
EXPERIMENTAL RESULTS OF COMPARISON WITH EXISTING SCHEMES

DataTime	Schemes	Acc(%)	Pre_A	Pre_S	Rec_A	Rec_S
20s	[13]	89.69	95.35	89.26	43.16	99.69
	[14]	**99.59**	35.05	34.51	**97.70**	97.03
	proposed	98.57	**100.00**	**100.00**	90.00	**100.00**
40s	[13]	86.39	80.09	85.73	50.00	97.72
	[14]	**99.58**	50.75	51.06	**97.95**	97.77
	proposed	97.49	**96.35**	**99.71**	93.56	**98.83**
60s	[13]	80.13	70.63	78.90	31.35	95.26
	[14]	**99.56**	60.45	61.21	**98.07**	97.81
	proposed	98.57	**99.14**	**99.25**	95.03	**99.76**
80s	[13]	79.63	74.85	83.28	24.05	91.24
	[14]	**99.55**	68.19	69.56	**98.24**	98.23
	proposed	98.48	**98.85**	**99.41**	93.72	**99.73**
100s	[13]	78.70	61.35	85.28	14.17	95.03
	[14]	**99.53**	74.99	76.66	**98.06**	98.42
	proposed	98.59	**98.07**	**99.34**	95.20	**99.87**
120s	[13]	80.43	65.84	88.11	23.71	94.31
	[14]	**99.54**	77.81	79.70	**98.14**	98.07
	proposed	98.80	**98.12**	**99.78**	95.01	**99.97**
300s	[13]	80.61	68.33	88.21	25.61	93.28
	[14]	**99.50**	**96.69**	**99.63**	**98.49**	99.41
	proposed	95.12	94.55	99.52	78.34	**99.83**

with scheme [14], the detection accuracy of the proposed scheme is roughly on par, with a slight decrease in recall but a significant improvement in precision. The results indicate that the proposed scheme has overall performance superior to existing schemes, particularly in maintaining high traceability capabilities under short data accumulation lengths, providing a solid foundation for timely blocking of attacks.

V. CONCLUSIONS

This study describes the behavioral features and modes of Sybil attacks and proposes a traceability detection scheme based on the fusion of vehicle behavioral features and spatiotemporal associations. Experimental results indicate that the proposed scheme achieves near-real-time detection speed while maintaining a high accuracy rate for detecting Sybil nodes. In response to the challenge of attackers enhancing their concealment through pseudonym strategies, this scheme reveals the complex relationships between vehicle nodes, demonstrating strong attack traceability performance. In comparisons, the proposed model outperforms existing optimal schemes, proving its effectiveness in practical applications. In the future, we will explore the structural relationships between vehicles in traffic flow and utilize dynamic graph methods to construct vehicle characteristic data to enhance real-time attack tracking capabilities.

REFERENCES

[1] Q. Yang, X. Zhu, X. Wang, J. Fu, J. Zheng, and Y. Liu, "A novel authentication and key agreement scheme for internet of vehicles," *Future Generation Computer Systems*, vol. 145, pp. 415–428, 2023.

[2] S. Devika, R. R. Shrivastava, P. Narang, T. Alladi, and F. R. Yu, "Vadgan: An unsupervised gan framework for enhanced anomaly detection in connected and autonomous vehicles," *IEEE Transactions on Vehicular Technology*, 2024.

[3] B. Hammi, Y. M. Idir, S. Zeadally, R. Khatoun, and J. Nebhen, "Is it really easy to detect sybil attacks in c-its environments: a position paper," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 18273–18287, 2022.

[4] M. M. Hamdi, M. Dhafer, A. S. Mustafa, S. A. Rashid, A. J. Ahmed, and A. M. Shantaf, "Effect sybil attack on security authentication service in vanet," in *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, pp. 1–6, IEEE, 2022.

[5] S. Kumar, A. Vasudeva, and M. Sood, "Sybil attack countermeasures in vehicular ad hoc networks," in *2022 International Conference on Communications, Information, Electronic and Energy Systems (CIEES)*, pp. 1–6, 2022.

[6] Y. Yao, B. Xiao, G. Yang, Y. Hu, L. Wang, and X. Zhou, "Power control identification: A novel sybil attack detection scheme in vanets using rssi," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2588–2602, 2019.

[7] W. Li and D. Zhang, "Rssi sequence and vehicle driving matrix based sybil nodes detection in vanet," in *2019 IEEE 11th International Conference on Communication Software and Networks (ICCSN)*, pp. 763–767, IEEE, 2019.

[8] S. Gyawali, Y. Qian, and R. Q. Hu, "Machine learning and reputation based misbehavior detection in vehicular communication networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8871–8885, 2020.

[9] R. G. Engoulou, M. Bellaiche, T. Halabi, and S. Pierre, "A decentralized reputation management system for securing the internet of vehicles," in *2019 International Conference on Computing, Networking and Communications (ICNC)*, pp. 900–904, IEEE, 2019.

[10] B. Hammi, M. Y. Idir, and R. Khatoun, "A machine learning based approach for the detection of sybil attacks in c-its," in *2022 23rd Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 1–4, IEEE, 2022.

[11] D. E. Laouiti, M. Ayaida, N. Messai, S. Najeh, L. Najjar, and F. Chaabane, "Sybil attack detection in vanets using an adaboost classifier," in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 217–222, IEEE, 2022.

[12] C. H. Quevedo, A. M. Quevedo, G. A. Campos, R. L. Gomes, J. Celestino, and A. Serhrouchni, "An intelligent mechanism for sybil attacks detection in vanets," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2020.

[13] S. Saudagar and R. Ranawat, "An amalgamated novel ids model for misbehaviour detection using vereminet," *Computer Standards Interfaces*, vol. 88, p. 103783, 2024.

[14] Z. Zhang, Y. Lai, Y. Chen, J. Wei, and Y. Wang, "Detection method to eliminate sybil attacks in vehicular ad-hoc networks," *Ad Hoc Networks*, vol. 141, p. 103092, 2023.

[15] T. Sapala, R. C. S. R. Penumallu, R. S. Kiran, M. Rajesh, and B. K. Devi, "A survey on vanet attacks and its security mechanisms," in *2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pp. 435–440, IEEE, 2022.

[16] M. Wazid, P. Bagga, A. K. Das, S. Shetty, J. J. Rodrigues, and Y. Park, "Akm-iov: Authenticated key management protocol in fog computing-based internet of vehicles deployment," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8804–8817, 2019.

[17] S. M. Faisal, B. K. Gupta, and T. Zaidi, "A hybrid framework to prevent vanet from sybil attack," in *2022 5th international conference on multimedia, signal processing and communication technologies (IMPACT)*, pp. 1–6, IEEE, 2022.

[18] E. F. Cahyadi, A. F. Isnawati, K. T. Putra, and H. Wijayanto, "A light reconstruction on cpp-bat in vanets," in *2022 IEEE International Conference on Consumer Electronics-Taiwan*, pp. 01–02, IEEE, 2022.

[19] U. Tariq, "Security-aware malicious event detection using multivariate deep regression setup for vehicular ad hoc network aimed at autonomous transportation system," in *2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, pp. 354–358, IEEE, 2022.

[20] J. Kamel, M. R. Ansari, J. Petit, A. Kaiser, I. B. Jemaa, and P. Urien, "Simulation framework for misbehavior detection in vehicular networks," *IEEE transactions on vehicular technology*, vol. 69, no. 6, pp. 6631–6643, 2020.

[21] T. ETSI, "Etsi ts 102 867 v1. 1.1-intelligent transport systems (its); security; stage 3 mapping for ieee 1609.2," *Standard, TC ITS*, 2012.

[22] D. SAE, "J2735 dedicated short range communications (dsrc) message set dictionary," *Society of Automotive Engineers, DSRC Committee*, 2009.