Received 7 September 2022, accepted 8 October 2022, date of publication 14 October 2022, date of current version 26 October 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3214838

RESEARCH ARTICLE

An Improved Robust Misbehavior Detection Scheme for Vehicular Ad Hoc Network

MOHAMMED ALZAHRANI[1,2], **MOHD YAZID IDRIS**[1,3], **FUAD A. GHALEB**[1], **AND RAHMAT BUDIARTO**[4]

[1]Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru, Johor 81310, Malaysia
[2]Ministry of Communications and Information Technology, Riyadh 12211, Saudi Arabia
[3]Media and Game Centre of Excellence (MaGICX), Institute of Human Centered Engineering (iHumEn), Universiti Teknologi Malaysia, Johor Bahru, Johor 81310, Malaysia
[4]College of Computer Science and Information Technology, Al-Baha University, Al Bahah 65731, Saudi Arabia

Corresponding author: Mohammed Alzahrani (r1985@graduate.utm.my)

ABSTRACT Vehicular Ad Hoc Networks (VANETs) have emerged to improve road safety and traffic efficiency and provide passengers comfort. Most VANET applications rely on the cooperation among vehicles sharing their sensed information. However, misbehaving vehicles which send false information can disrupt the VANETs potential. Although many solutions have been proposed to defend against misbehaving vehicles in VANET, most of these solutions relays on honest majority assumptions and are thus vulnerable to collusion attacks. Colluding vehicles send fake information, and because detection depends on cooperation, such information misleads benign vehicles to make an accurate decision. This study proposes an improved Robust Misbehavior Detection Scheme (iRMDS) by replacing the statistics-based detection threshold, which assumes an honest majority, with a machine learning-based classifier. A Neuro-Kalman-Based Robust Misbehavior detection scheme is proposed in three phases. In the first phase, attackers-Independent features are extracted from signal properties such as the receive signal strength and signal direction and have been integrated with context information features. In the second phase, the Kalman filter algorithm has been designed to extract consistent patterns of context information for each vehicle. That is, the innovation errors of the Kalman filter have been utilized as the input features to train the misbehavior detection model. In the third phase, the artificial neural network algorithm is integrated with the outputs of the Kalman Filter algorithm to recognize the malicious pattern. Results show that the overall performance of the proposed iRMDS solution achieves a 3.44% improvement compared to the related work. Such enhancement is promising in realizing reliable VANET applications to improve road safety, traffic efficiency, and passenger comfort.

INDEX TERMS Misbehavior detection, neural network, Kalman filter, colluding attack, vehicular network, VANET.

I. INTRODUCTION

World Health Organization (WHO, 2020) reported that by 2030, road accidents would be the fifth leading cause of fatalities [1]. More than 1.35 million people die annually due to traffic; injuries are 50 times the fatalities [1], [2]. Every year, vast amounts of money (about 3% of global GDP) are expended on people recovering and compensating victims of traffic accidents who have damaged or lost their assets. The Vehicular Ad Hoc Network (VANET) has been established to enhance traffic safety and efficiency while providing passenger comfort. VANET is typically expected to be the main component of future Intelligent Transportation Systems [3]. Many VANET applications have been designed and analyzed, including cooperative active safety systems (CASS) [4], cooperative collision warning (CCWS) [5], and driver assistance systems (ADAS) [6]. The performance of VANET applications depends on the quality of the integrity of the cooperative awareness messages broadcasted by the nearby vehicles [2], [7], [8], [9]. However, misbehaving vehicles which broadcast inaccurate information can disrupt the fundamental operations of VANET applications [10], [11].

The associate editor coordinating the review of this manuscript and approving it for publication was Abderrahmane Lakas.

Thus, security is essential to release VANET applications' potential to improve road safety and traffic efficiency and provide passenger comfort [12], [13].

Many misbehaviors detection have been proposed for VANET in the literature [7], [9], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20]. Most of these solutions are based on data-centric that classify the vehicles based on the quality of the cooperative awareness messages shared among neighboring vehicles. It is a common belief that data-centric misbehavior detection is more effective than the entity-centric approach that classifies the vehicles based on their behavioral activities regarding the VANTE protocols (e.g., misusing communication protocol). Most recent data-centric solutions [8], [9], [15], [17] are based on honest majority assumptions in which a reference model is constructed online based on the majority. That is, vehicles that deviate much from the constructed reference are considered misbehaving ones. Although such an approach is suitable for vehicular context and outperforms the traditional predefined static approach, such an approach is vulnerable to collude attack [19], [21].

In collude attacks, misbehaving vehicles colludes to send false information to influence the honest majority-based solutions leading to false detection. Detecting such an attack is challenging due to the high similarity between benign and misbehavior data. Unfortunately, collude attacks have not yet been investigated in depth in VANET. A study in [19] proposed a misbehavior detection scheme by extracting independent attack features based on signal properties such as Angle of Arrival (AoA) and Signal Strength Indicators (RSSI). Such features are out of the attackers' control and can be used to distinguish misbehaving vehicles effectively. However, these solutions rely on predefined static detection thresholds, which make the classification of vehicles challenging, decrease the overall detection accuracy, and raise false alarms.

To this end, this study proposes a Neuro-Kalman Based misbehavior detection scheme called iRMDS. The proposed scheme extracts both the context information and signal properties and injects them into a Neuro-Kalman-based misbehavior detection constructed trained based on Kalman filter using an Artificial Neural Network for malicious pattern recognition. The proposed iRMDS consists of three phases: data acquisition, feature integration, and misbehavior detection. In the first phase, the signal properties and cooperative awareness messages are acquired from neighboring vehicles. In the second phase, a consistency algorithm based on the Kalman filter is used to evaluate the consistency of this information with its signal properties. In the third phase, the context reference, which is based on the honest majority, and the static detection threshold of the related RMDS have been replaced by an ANN-based classifier. This classifier was trained using a dataset collected from the innovation sequence of the Kalman filter of benign and misbehaving vehicles. Results show that the proposed iRMDS scheme outperforms the related work, a 3.44% improvement was attained in the overall performance.

1. An improved robust misbehavior detection scheme is presented to effectively detect a colluding attack on vehicular networks.
2. The signal properties-based features such as AoA and RSSI are derived and integrated with the message content-based features using the Kalman filter. The innovation errors of the Kalman filter have been used as the features vector that has been used to improve the detection accuracy of colluding attacks.
3. The decision rules that were based on the statistical detection thresholds have been replaced by a machine learning model trained using an artificial neural network algorithm.
4. Extensive experiments have been conducted to validate and evaluate the proposed model. The results show that the performance of the proposed scheme has been significantly improved compared with the performance of the state-of-the-art models.

The rest of this paper is organized as follows. A detailed description of the proposed solution is presented in section 2. The experimental design and performance evaluation are described in Section 3. The results and analysis are discussed in Section 4. The conclusion and future work are summarized in Section 5.

II. RELATED WORK

Many studies investigated the problem of misbehavior detection in VANET [7], [9], [12], [13], [14], [15], [16], [17], 18]. These studies can be categorized based on the actors used to be validated into two approaches data-centric and entity-centric. In the entity-centric approach, vehicles are classified based on their behaviors in terms of obeying the known rules or based on their types, such as police cars. Messages are evaluated based on the trust value of the entity [22], [23], [24]. These trust values are either constructed based on the vehicle past behavior or based on vehicle roles given by trust authority. For example, police vehicles and public buses have higher trust values compared to passenger vehicles. Thus, the trustworthiness of the information broadcasted by a vehicle is mapped directly to the trust value of the vehicle. This approach has many drawbacks. Firstly, trusted vehicles may misbehave suddenly by sending false information. Consider spoofed or stolen trusted vehicles. Accordingly, the trust values of vehicles need to be continuously updated by trusted authorities based on recent behavioral activities. Though, frequently updating the trust value of huge and decentralized networks is complex in VANET. Secondly, due to the distributed nature of VANET and the harsh vehicle environment, it is very difficult for a trusted authority to evaluate the vehicles based on their past behaviors due to the lack of the information and presence of false information. Thirdly, many entity-centric mechanisms include data-centric techniques to locally evaluate the vehicles based on the quality of their generated data which is a challenging research problem.

In the data-centric approach, a message is validated based on the quality of the generated information [25], [26]. The quality is always measured based on the consistency and plausibility of the message content. Unfortunately, the quality of the generated information is dynamic due to the dynamic noises in the vehicle's environment and communications losses due to traffic density which affect communication reliability and thus lead to a lack of information. Accordingly, many context-aware misbehavior models were suggested to validate the data shared among vehicles in the scene. In such an approach, a context reference is constructed online based on data collected locally from neighboring vehicles [2], [8], [9], [15], [17], [27], [28], [29], [30], [31], [32], [33]. A message that deviates much from the context is considered a malicious message and is accordingly used to decrement the trust value of the sender for the entity-centric approach. The following is a review of the existing context-aware misbehavior detection models.

Authors in [8] proposed a context-aware data-centric misbehavior detection scheme for VANET. The uncertainty due to the environmental noises and communication losses have been considered as the main elements in constructing the context reference. Kalman filter was used to extract the features that context uncertainties and the Hampel filter were used to construct the context reference model. The temporal consistency of the messages generated by each vehicle has been estimated using the Kalman filter algorithm. The spatial consistency of a vehicle's data compared with the consistencies of neighboring vehicles is modeled using the Hampel filter. The messages that deviate much from Hampel filter boundaries (the context reference) are considered malicious messages. The context reference is reconstructed every 100 milliseconds, which is the period needed for exchanging new messages in VANET according to the IEEE standard. The main drawback of this approach is that the context-reference was built based on an honest majority assumption which doesn't hold in the presence of colluding vehicles that send fake but consistent information. Another limitation of such a model is the use of a parametric statistical model, which uses a predefined detection threshold to determine the outlier messages.

Authors in [18] trained a misbehavior detection model based on an artificial neural network (ANN). Kalman filter was used to extract the consistency features in terms of innovation error and communication characteristics. The main drawback of such an approach is it is too scenario-specific due to the assumption that the relationship between input and output features is stationary, which doesn't hold for a highly dynamic context.

Authors in [7] proposed a collaborative ensemble-based misbehavior detection scheme for VANET. Each vehicle independently trains a classifier based on its collected data using Random Forest (RF) algorithm and shares it with its neighboring vehicles. Each vehicle constructs the final ensemble misbehavior detection and uses it for the detection. The main drawback of this model is that the constructed detection model is composed of different contexts, which are not effective for VANET spatiotemporal context.

Authors in [15] proposed a context-aware-detection scheme called CA-EC-MDS, which considers the data consistency, plausibility, and behavioral features. Three feature sets were created based on data consistency, plausibility, and behavioral features. Kalman filter was utilized to extract the consistency features, while plausibility features were extracted using rules constructed based on overlapping positions. Meanwhile, the behavioral features were extracted from the broadcasting behavior of the vehicles. An ensemble of three classifiers was constructed, each of which was trained based on specific feature sets. The outputs of these classifiers were aggregated using the majority voting algorithm. The main drawback of this model is that it is vulnerable to colluding attacks. Another limitation is the use of parametric statistical techniques for highly dynamic networks.

Authors in [9] construct an ensemble of classifiers using the random forest (RF) algorithm utilizing the features extracted from the parameters of the statistical model proposed by [33]. However, the main drawback of this model is that the statistical parameters that were used to train the ensemble classifier can be manipulated by colluding vehicles due to the majority of honest assumptions used to construct the model.

Zhang and Chen [13] devised a trust-based misbehavior detection model using a support vector machine (SVM) and Dempster–Shafer theory (DST). Two models were built, one is data-centric, and the other one is entity-centric. SVM algorithm was used to train a classifier that can classify vehicles based on their broadcasting behavior. Each vehicle broadcasts the result of the classifications to neighboring vehicles. The trusted authority aggregates the reports collected for each vehicle using DST. The main limitation of such an approach is not considering the consistency and plausibility of the message content. Moreover, the context uncertainties have not been considered by the model. In addition, the model relies on reputation and long-term trust establishment, which is not suitable for early detection and new misbehaving nodes. Authors in [12] constructed a model consisting of two machine learning techniques, namely k-nearest neighbor (kNN) and random forest (RF) classifiers. However, the proposed model is scenario-specific and cannot be generalized. Moreover, the data consistency and plausibility features were not considered. Authors in [27] proposed a misbehavior detection scheme called FCA-MDS which is based on fuzzy logic. The detection rules were augmented by a fuzzy inference system to accurately adjust the detection thresholds. The thresholds are dynamically updated based on the context. However, the main limitation of this study is the honest majority assumption.

The author of this study presented an initial idea of a misbehavior detection scheme called RDMS in [19] that is robust against colluding attacks. The study argued that most of the existing misbehavior detection models depend on features that can be manipulated by vehicles. Thus, any reference constructed online based on these features is contaminated by the misbehaving vehicles data. The study aimed to enrich the consistent features that can be controlled by the sender vehicles, like the vehicle's mobility information, with features that cannot be manipulated by the sender vehicle; instead, they can be sensed directly by the receiver, such

TABLE 1. Limitations of the related work.

Reference	Limitations
[8]	• Context-reference was built based on an honest majority assumption • vulnerable to colluding attacks
[18]	• Scenario Specific which assumes that the relationship between input and output features is stationary and assume honest majority • vulnerable to colluding attacks
[7]	• The constructed detection model is composed of different contexts
[15]	• Use parametric statistical techniques for calculating predefined and static thresholds for highly dynamic networks. • vulnerable to colluding attacks
[9]	• Predefined statistical parameters and assume honest majority • vulnerable to colluding attacks
[13]	• The consistency and plausibility of the message content have not been considered. • Context uncertainties have not been considered by the model • relies on reputation and long-term trust establishment • vulnerable to colluding attacks
[12]	• scenario-specific and cannot be generalized. • vulnerable to colluding attacks
[27]	• honest majority assumption • vulnerable to colluding attacks
[19]	• uses predefined and static thresholds for detection • lack of deep investigation of the performance in different vehicular context

as signal strength and signal direction. The received signal strength indicator (RSSI) and the angle of the arrival (AoA), which are out of the attacker's control and modification, have been used to improve the consistent evaluation of the messages. Such mechanisms can thwart the colluding vehicles that misbehave to mislead the decision. However, the main limitation of this study is that it uses predefined and static thresholds for detection. Moreover, the study lack of deep investigation of the performance in different vehicular context.

To summarize, as shown in Table 1, most of the existing misbehavior detection systems proposed for VANET are vulnerable to colluding attacks in which vehicles can collide and send consistent but false information to cause decisions. Most of the solutions assume an honest majority either during the final decision using majority-based voting schemes or during the construction of the context reference in an online manner. Many models suggested machine learning techniques to train the detection model to avoid constructing a misleading reference due to the colluding attacks. However, such an approach is too scenario-specific and doesn't consider the dynamic uncertainty of the data in different vehicular contexts. Moreover, misbehaving vehicles can generate consistent data to overtake the detection. In this study, an improved Robust Misbehavior detection Scheme through the integration of the context information with signal properties. Kalman filter technique has been devised to extract the representative feature. The Artificial Neural Network (ANN) model is trained to detect malicious pattern recognition and accordingly predict the detection thresholds. The trained ANN model has been trained based on the extracted features from the Kalman filter. The trained ANN model replaces the predefined statistical thresholds that were presented in RDMS [19]. In the following section, a detailed description is presented.

III. PROPOSED DETECTION SCHEME

As shown in Figure 1, the proposed scheme comprises three phases: data collection, feature extraction, and misbehavior detection. In the first phase, vehicles receive cooperative awareness messages from neighboring vehicles and store them along with their signal properties in a local database (internal database in each vehicle). In the second phase, vehicles extract the features from the CAMs' content and form the signal properties using the Kalman filter algorithm. In the third phase, the outputs of the Kalman filter algorithm are used as input features to train an artificial neural network-based classifier to detect fake messages by recognizing malicious traffic patterns. A detailed description of these phases is presented in the following subsections.

A. DATA COLLECTION PHASE

In this phase, vehicles collect the Cooperative Awareness Messages (CAMs) from two sources: vehicles acquire their information from local sensors and collect information from their neighboring vehicles. The collected information is stored in a local database.

1) SELF CAMS ACQUISITION

Every 100ms (according to the IEEE standard in [34]), each vehicle in the network acquires the elements of CAMs messages from its local sensors such as positioning sensors, speedometer, and gyroscope to form a CAM message. CAMs messages consist of the mobility information of the sender vehicle, such as position (latitude and longitude), velocity, acceleration, and vehicle direction, as well as other information related to the status of the vehicle. This information is subject to environmental noises due to the vehicle's sensors in a harsh dynamic environment. For example, position information is subject to different kinds of noises due to the reflection and refraction of positioning signals on the obstacles such as constructions, buildings, and other moving vehicles. Position

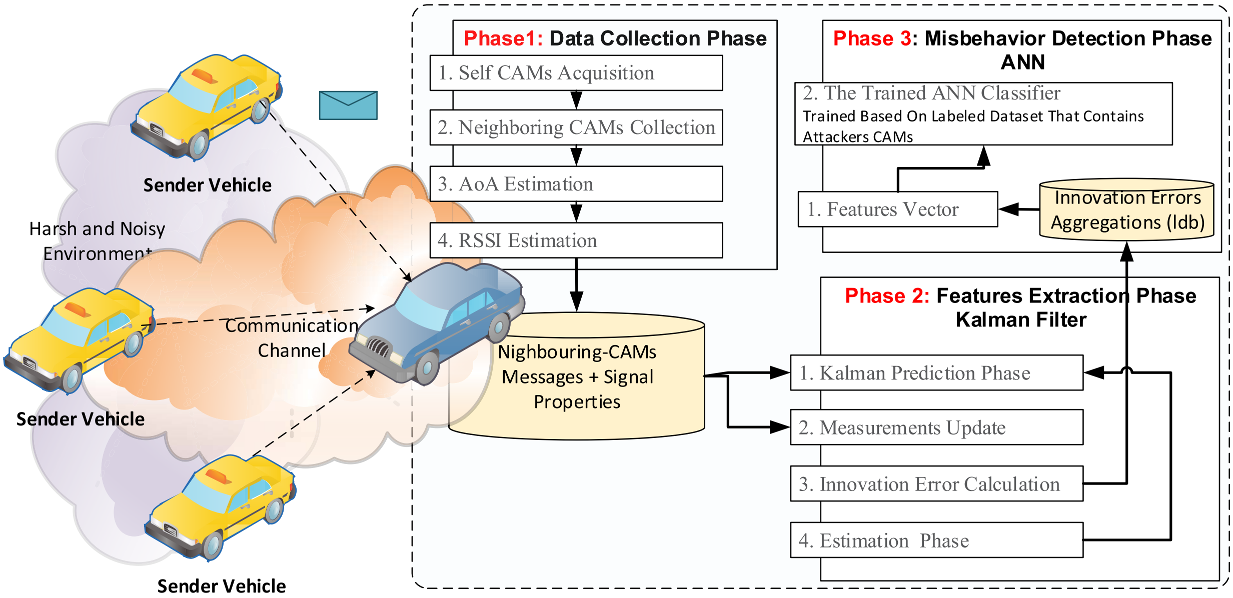
Phase1: Data Collection Phase
1. Self CAMs Acquisition
2. Neighboring CAMs Collection
3. AoA Estimation
4. RSSI Estimation
Phase 3: Misbehavior Detection Phase ANN
2. The Trained ANN Classifier
Trained Based On Labeled Dataset That Contains Attackers CAMs
1. Features Vector
Innovation Errors Aggregations (Idb)
Phase 2: Features Extraction Phase Kalman Filter
1. Kalman Prediction Phase
2. Measurements Update
3. Innovation Error Calculation
4. Estimation Phase
Nighbouring-CAMs Messages + Signal Properties
Sender Vehicle
Harsh and Noisy Environment
Communication Channel
Sender Vehicle
Sender Vehicle

FIGURE 1. The proposed improved Robust Misbehavior Detection Scheme (iRMDS).

signals are also subject to attenuations due to the water bodies such as trees, clouds, and rain. Such inaccurate information leads to inconsistencies and increases the false alarm rates of misbehavior detection. Therefore, an accurate acquisition algorithm should remove the noise and improve the CAMs content accuracy before it is broadcast to other vehicles. This study used the acquisition algorithm that was proposed in [35] to remove the noise and generate a more consistent mobility information stream. The algorithm in [35] continuously estimates the measurement's noise covariance based on an analysis of the correlations in the Kalman filter's innovation error. The presence of correlations in the innovation error of the Kalman filter increases the uncertainties about the accuracy of the positioning information. The correlation value is used to scale up or down the measurements' noise covariance and to adapt Kalman filter parameters to accurately estimate the vehicles' positions.

2) NEIGHBORING CAMS COLLECTION

The second source of information is collected from neighboring vehicles. According to the IEEE standard in [34], vehicles need to broadcast their CAMs messages every 100ms. Due to the congestion caused by vehicles' dynamic speeds and density, vehicles may not be able to broadcast 10 CAMs every second. A considerable amount of these messages will be lost depending on the context. CAMs messages are subject to interference and congestion, and attenuations due to high vehicle density and different vehicle speeds leading to high message loss and thus degrading the quality of the messages. Therefore, an effective and efficient broadcasting scheme should be used to accomplish this task.

Many broadcasted schemes were suggested for VANET with the concept of adapting the broadcasting rate of communicating vehicles based on the communication channel status [36], [37], [38], [39], [40], [41]. This led to an equal broadcasting rate for the neighboring vehicles. Some studies argue that the use of an equal broadcasting rate is inefficient and ineffective for VANET because, for safety reasons, emergency vehicles need a higher broadcasting rate compared to a vehicle with stable and predictable mobility movement [2], [3], [42], [43], [44]. In this study, the broadcasting scheme proposed by [44] was used due to the consideration of the driving situations of each vehicle in the broadcasting decision. A vehicle individually increases its broadcasting rate based on its deriving situation. For example, during the acceleration, deceleration, or lane change, a vehicle raises its broadcasting rate and reduces it once vehicle mobility information is predictable. As was argued by [2], [43], and [44] such a broadcasting scheme not only improves the message reliability but also preserves the quality of mobility patterns collected from the CAMs stream of each neighboring vehicle. However, due to the presence of misbehaving vehicles, CAMs contents are subject to be tampered with by attackers. Therefore, robust features that resist altering should be used along with mobility-based features.

In this study, because the CAMs messages are subject to falsification from misbehaving vehicles, features based on signal properties such as the received signal strength and the signal direction are collected using a directional antenna of the receiving vehicles. Such features are independent of the attackers. For example, attackers cannot misuse the signal direction because it is sensed by neighboring vehicles. It is out of the attacker's control, and thus, it is robust. Meanwhile, if the attacker modified the received signal strength indicators (RSSI), neighboring vehicles will detect directly because such signal strength will be higher than the standard value. Thus, new features based on the Angle of Arrival (AoA) and the Receiving Signal Strength Indicator (RSSI) were extracted.

3) AOA ESTIMATION

There are two more features extracted in this study: the actual and the expected. The actual AoA is estimated using an array antenna and signal arrival phase, while the expected AoA is calculated using the position claimed by the sender and receiver vehicle. The actual AoA can be estimated based on the time difference of signal arrival (TDOA), it can be measured based on the signal phase in the antenna array, or it can be estimated using the signal strength on a rotating antenna. To estimate the phase change from the source to destination vehicle antennas, let θ_t^v denotes the AoA of a vehicle as estimated using an array antenna d_1 is the distance between sender and receiver first antenna, d_2 is the distance between the sender and receiver second antenna, a is the distance between the antennas. The angle of arrival can be calculated for every two neighboring antennas in the antenna array as follows.

$$\theta_t^v = sin^{-1}\frac{d_2 - d_1}{a} \tag{1}$$

where the distance d can be obtained as follows.

$$d = \emptyset\frac{2\pi f}{c} \tag{2}$$

where $\emptyset$ is the phase of the received signal, f signal frequency, and c is light speed. Accordingly, the angle of arrival can be estimated as follows.

$$\theta_t^v = sin^{-1}(\frac{c}{2\pi f}\frac{(\emptyset_2 - \emptyset_1)}{a}) + \epsilon_s \tag{3}$$

Meanwhile, the expected AoA, denoted by $\hat{\theta}_t^v$ Can be derived using the position information extracted from the received CAMs, denotes the AoA as estimated from CAMs information.

$$\hat{\theta}_t^v = tan^{-1}\frac{x_t^s - x_t^r}{y_t^s - y_t^r} + \epsilon_p \tag{4}$$

where x_t^s, x_t^r denote to the latitude of the sender and receiver vehicles, respectively while y_t^s, y_t^r denote to the longitude of the sender and receiver vehicles, respectively and ϵ_p Is the uncertainty of the estimation.

4) RSSI ESTIMATION

The received signal strength indicator can be estimated in the physical layer of the wireless network interface. However, due to the non-isotropic nature of signal propagation, the RSSI measured by a single antenna is inaccurate. Therefore, the array antenna is perfect for this study. Two RSSI features will be used in this study. The actual RSSI is directly measured by the receiver. Meanwhile, the expected RSSI can be estimated using the positioning information in the CAMs messages and knowledge about the environment. The RSSI can be computed as follows.

$$RSSI = P_{tx} + G_{tx} + G_{rx} - path\ loss \tag{5}$$

where P^{TX} denotes the transmission power and A^g denotes the antenna gain. While the transmit power and antenna gain can be considered constant, the path loss, which is the reduction of the signal power due to obstacle in the medium where the signal transfer, is a function of the distance between the sender and the receiver, and the medium where the signal is transmitted. The free-space path loss (FSPL) can be derived as follows.

$$FSPL = \left(\frac{4\pi fd}{c}\right)^2 \tag{6}$$

where d denotes the distance from the transmitter to the receiver (meters), f is the signal frequency (Hz), and c is the light speed.

B. FEATURES EXTRACTION PHASE

In this phase, the data collected from the previous phase are integrated using the Kalman filter algorithm to extract the inconsistencies. Because CAMs and Signal properties information represent vehicles' state and such state should be consistent due to consistent movement of vehicles, state estimation algorithms such as the Kalman filter algorithm are utilized to fuse such data and extract the inconsistencies as in the pseudocode 1.

1) KALMAN PREDICTION PHASE

$$\left[\ p_{x(k)}\ p_{y(k)}\ v_{x(k)}\ \ v_{y(k)}\ \theta_k\ RSSI_k\ \right] \rightarrow f_k \tag{7}$$

$$f_{k|k-1} = \begin{bmatrix} p_{x(k)} + \frac{d}{dt}p_{x(k)} + \frac{d^2}{dt^2}p_{x(k)} \\ p_{y(k)} + \frac{d}{dt}p_{y(k)} + \frac{d^2}{dt^2}p_{y(k)} \\ vx_k + \frac{d}{dt}v_{x(k)} \\ vy_k + \frac{d}{dt}v_{y(k)} \\ tan^{-1}\frac{p_{x(k)}^s - p_{x(k)}^r}{p_{y(k)}^s - p_{y(k)}^r} \\ P_{tx} + G_{tx} + G_{rx} - FSPL \end{bmatrix} \tag{8}$$

$$f_{k|k-1} = \begin{bmatrix} p_{x(k)} + Tvx_k + \frac{T^2ax_k}{2} \\ p_{y(k)} + Tvy_k + \frac{T^2ay_k}{2} \\ vx_k + Tax_k \\ vy_k + Tay_k \\ tan^{-1}\frac{p_{x(k)}^s - p_{x(k)}^r}{p_{y(k)}^s - p_{y(k)}^r} \\ P_{tx} + G_{tx} + G_{rx} - \left(\frac{4\pi fd}{c}\right)^2 \end{bmatrix} \tag{9}$$

2) MEASUREMENT VECTOR

The measurement vector contains the vehicle positions and speeds which are extracted from the received CAM message, the angle of arrival denoted by $\hat{\theta}_k$ Which is measured using the array antenna, and the received signal strength $\widehat{RSSI}_k$ denoted by $\widehat{RSSI}_k$ which is measured using the network interface card of the vehicles.

$$CAM_k = \left[\ \hat{p}_{x(k)}\ \hat{p}_{y(k)}\ \hat{v}_{x(k)}\ \ \hat{v}_{y(k)}\ \hat{\theta}_k\ \widehat{RSSI}_k\ \right] \rightarrow \hat{f}_k \tag{10}$$

where $\hat{p}_{x(k)}$ and $\hat{p}_{y(k)}$ the latitude and longitude as reported by the sender vehicle, $\hat{v}_{x(k)}$ and $\hat{v}_{x(k)}$ are vehicle velocities in different directions, $\hat{\theta}_k$ is the angle of arrival as measured by array antenna, and $\widehat{RSSI}_k$ Is the received signal strength as measured by the network interface.

3) INNOVATION ERROR

The innovation sequence of the Kalman filter, which is denoted by z_k represents the difference between information collected using measurements and the information predicted using the prediction model.

$$\hat{f}_k - f_{k|k-1} \rightarrow z_k \tag{11}$$

This information can be used to achieve two objectives. The first is improving the accuracy, and the second is ensuring security. To improve the information accuracy z_k can be used to correct the estimation of the vehicle position. Meanwhile, by analyzing the cointroduction between the information received by vehicles and the information measured by the same vehicle, the trustworthiness of the information can be estimated. For example, if the direction of the signal measured by the antenna deviates much from that predicted using the information obtained from the CAM content, this could be an indication of misbehavior. A detailed description of how to analyze to detect misbehavior is presented in the next section. The following sub-section demonstrates how the information can be accurately estimated even in the presence of misbehaving vehicles which send false information.

4) STATE ESTIMATION PHASE

In this phase, the Kalman filter gain K_k is calculated as follows.

$$K_k = P_{k|k-1} H^T \left(H P_{k|k-1} H^T + R\right)^{-1} \tag{12}$$

where $P_{k|k-1}$ denotes the prediction errors covariance (the prediction uncertainties), H is the unity matrix, and R is the measurements noise covariance. The estimated information based on the received CAM and

$$f_k = \hat{f}_k + K_k z_k \tag{13}$$

The estimation error covariance can be computed as follows.

$$P_{k|k} = (I - K_k H) P_{k|k-1} \tag{14}$$

As can be seen in pseudocode 1, in each time epoch $\boldsymbol{k}$, each vehicle uses the Kalman filter algorithm to predict or estimate (predict if it is missing and estimate if received) the currently expected messages using the previous actual messages $\boldsymbol{CAM_{i(k-1)}}$. Each vehicle $\boldsymbol{i}$ estimates the $\boldsymbol{RSSI_{i(k)}}$ ***and*** $\boldsymbol{\theta_{i(k)}}$ based on the $\boldsymbol{CAM_{i(k)}}$ messages and. Each received message is concatenated with its signal properties $(\boldsymbol{CAM_{i(k)}}||\boldsymbol{RSSI_{i(k)}}||\boldsymbol{\theta_{i(k)}} \rightarrow \hat{f}_{i(k)})$. Using the Kalman filter algorithm, a vehicle predicts the message $\boldsymbol{f_{i(k)}}$ based on the previous one $\boldsymbol{f_{i(k-1)}}$. T The innovation errors which is the difference between the predicted and the measured features $(\boldsymbol{f_{i(k)}} - \hat{f}_{i(k)})$ are calculated. For each vehicle, the innovation error is aggregated and stored in the local database (***ldb***). Similarly, in each time epoch, these features are aggregated and stored in the local database (***ldb***) to represent the context.

C. NEURO-KALMAN MISBEHAVIOR DETECTION PHASE

Pseudocode 1: Extracting Representative Features

```
∀ time epoch t do :
    ∀ i ∈
neigbouring vehicles in the CAMs database do:
    $CAM_{i(k)}||RSSI_{i(k)}||\theta_{i(k)} \rightarrow f_k$
    $KF_{predict(f_{k-1}:CAM_{i(t)},RSSI_{i(t)},\theta_{i(k)})} \rightarrow \hat{f}_{i(k)}$
    $f_{i(k)} - \hat{f}_{i(k)} \rightarrow KF_{innovation\_error} \rightarrow z_k$

  $Aggregate\left(KF_{innovation_{error}}\right) \xrightarrow{store\ to\ ldb} \mu_{i(t)}, std_{i(t)}$

    $Aggregate\left(KF_{innovation_{error}}\right) \xrightarrow{store} \mu_t, std_t$
Return ldb
```

In this phase, the features extracted from the previous phase are used as input to the ANN classifier to determine whether it contains malicious patterns. A malicious pattern indicates that the message is fake. To train the ANN classifier, a labeled dataset should be available. As long as VANET is not yet deployed, the dataset is generated using simulation. The Next Generation Simulation (NGSIM) dataset was used in this study to train the proposed model. NGSIM dataset has been replied in MATLAB simulation environment containing the surrounding noises affecting vehicular sensors' measurement accuracy. Different types of traffic scenarios were also simulated. Moreover, different types of misbehaviors were simulated and injected to misbehaving vehicle trajectories. The simulation details of creating the dataset are in the next section. The dataset has the message-id and the innovation sequence of the Kalman filter along with the CAM label fake or genuine. The dataset is split into two subsets 60% for training and 40% for testing. The proposed ANN classifier, namely the Multi-layer Perceptron classifier, consists of three layers: an input layer, a hidden layer, and an output layer. The Rectified Linear Unit (RelUe) is used as the activation function. The trained ANN model takes the aggregated outputs of the Kalman filter as input and the output is the decision of the misbehavior status of each vehicle. Vehicles that are predicted greater than zero are misbehaving, while vehicles which predicted smaller than zero are benign vehicles. The following equation shows the decision formula.

$$f\left(\mu_{i(t)}, std_{i(t)}\right) = \begin{cases} 1 & if \sum w_i x_i + b_i \geq 0 \\ 0 & if \sum w_i x_i + b_i < 0 \end{cases} \tag{15}$$

The following summarizes the role of integrating the Kalman Filter with the ANN model in the proposed scheme. As shown in Figure 2, the proposed scheme consists of three phases. In the first phase, vehicles collect the context information with their signal properties from neighboring vehicles. In the second phase, in each time epoch, a vehicle computes the statistical parameters of the Kalman filter innovation error of each feature. The results are time series data representing the behavioral activities of each vehicle. With the help of the innovation error resulting from CAMs, RSSI, and AoA feature, the statistical parameters of the misbehaving vehicles will be less consistent compared to benign vehicles. In the third phase, instead of human-based predefined thresholds as in the related works. The ANN model learns how to distinguish between benign and malicious patterns without the

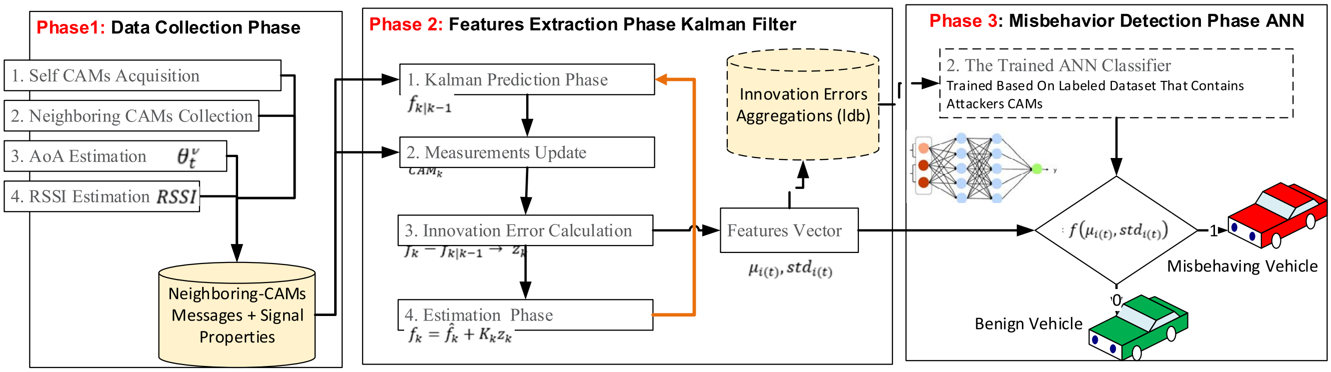
Phase1: Data Collection Phase
1. Self CAMs Acquisition
2. Neighboring CAMs Collection
3. AoA Estimation θ_t^v
4. RSSI Estimation $RSSI$
Neighboring-CAMs Messages + Signal Properties
Phase 2: Features Extraction Phase Kalman Filter
1. Kalman Prediction Phase
$f_{k|k-1}$
2. Measurements Update
CAM_k
3. Innovation Error Calculation
$f_k - f_{k|k-1} \rightarrow z_k$
4. Estimation Phase
$f_k = \hat{f}_k + K_k z_k$
Innovation Errors Aggregations (Idb)
Features Vector
$\mu_{i(t)}, std_{i(t)}$
Phase 3: Misbehavior Detection Phase ANN
2. The Trained ANN Classifier
Trained Based On Labeled Dataset That Contains Attackers CAMs
$f(\mu_{i(t)}, std_{i(t)})$
1
Misbehaving Vehicle
0
Benign Vehicle

FIGURE 2. The proposed improved Robust Misbehavior Detection Scheme (iRMDS) – the dotted lines denote one time process.

need for predefined thresholds. The ANN algorithm can accurately capture such differences without specifying predefined or statistical thresholds to distinguish between benign and misbehaving vehicles.

IV. EXPERIMENTAL PROCEDURE

In this section, the procedure that was conducted to validate and evaluate the proposed model is described. This procedure consists of three steps: dataset preparation, environmental noise injection, traffic scenarios, attack simulation, performance measures, and model evaluation. A detailed description of each procedure is in the following subsections.

A. DATASET PREPARATION

The Next Generation Simulation (NGSIM) dataset has been used to validate and evaluate the proposed model in this study. NGSIM dataset is commonly used by many related studies, such as in [2] and [27]. NGSIM dataset comprises ground truth trajectories of more than 5000 vehicles. The dataset includes different traffic scenarios that comprise different driving behaviors, speeds, acceleration patterns, and traffic flows. The datasets contain the mobility information of the vehicles collected every 100ms. This information includes vehicle positions in terms of longitude and latitude, vehicles' speed and acceleration, vehicles direction, and lane number. It is also included the velocity (longitude and latitude), acceleration, direction, vehicle type, dimensions, and lane number. The dataset is collected by cameras that are used to record real traffic for 45 minutes. NGSIM dataset contains some missed information and outliers due to the imperfection of extracting the information from the video stream. Therefore, it is a common research procedure to preprocess the dataset by filling in the missing data and replacing the outliers. A common procedure for prepossessing is the preprocessing described by Thiemann et al. [45], in which the missing and outlier data has been replaced using the exponentially weighted moving average method (sEWMA). The sEWMA method also is used to smooth the speed measurements. The acceleration was derived directly from the derivative of speed over time. In addition, the heading angle was derived from the derivative of position displacement in one axis over the displacement in the second axis.

B. NOISE INJECTION

NGSIM dataset contains ground truth mobility information which represents the ideal case of the VANET scenario. To make the experiment more practical, the position information should contain some environmental noises. Similarly, the uncertainties of the odometer and other sensors should also be considered. In this study, three types of noises have been injected into vehicle trajectories, more specifically to vehicle positions. These noises represent three common situations that can be found in vehicle positioning sensors such as GPS sensors and are widely used by VANET research in [2], [8], [9], [15], [27], and [35]. The first noise type can be found in the open sky environment, in which the error from positioning sensors can be represented as a random noise process that follows the normal distribution with zero mean. The second type of noise can be represented by random noises with dynamic variance. This type of error can be resulted from driving under trees and in cloudy and rainy situations. The third type of noise can be represented by white noise and the random walk process in which the variance grows with time. That is, the error accumulates with time. Such type of error occurs when a vehicle moves between long-story buildings and bridges or construction sites in which the positioning signal reflects on the objects creating random walk-type noise in the signal. In this study, a subset of the NGSIM dataset was replayed in a Matlab simulation environment to create the dataset. The vehicle trajectories were replayed to simulate vehicle movement, and these three types of dynamic environmental noises were injected into the vehicle's trajectories to represent a realistic environment. Vehicles use the algorithm in [35] to estimate the correct state by filtering the noises.

C. TRAFFIC SCENARIOS

In VANET, vehicle density and velocity have a significant impact on the reliability of the information shared among vehicles. Due to vehicle mobility, mobility information is outdated quickly. Thus, it is standardized that a vehicle sends its mobility information every 100ms. However, when vehicle density increases, vehicles compete on the communication channel, and thus some messages will be congested, leading to unreliable communication. Similarly, when vehicles have

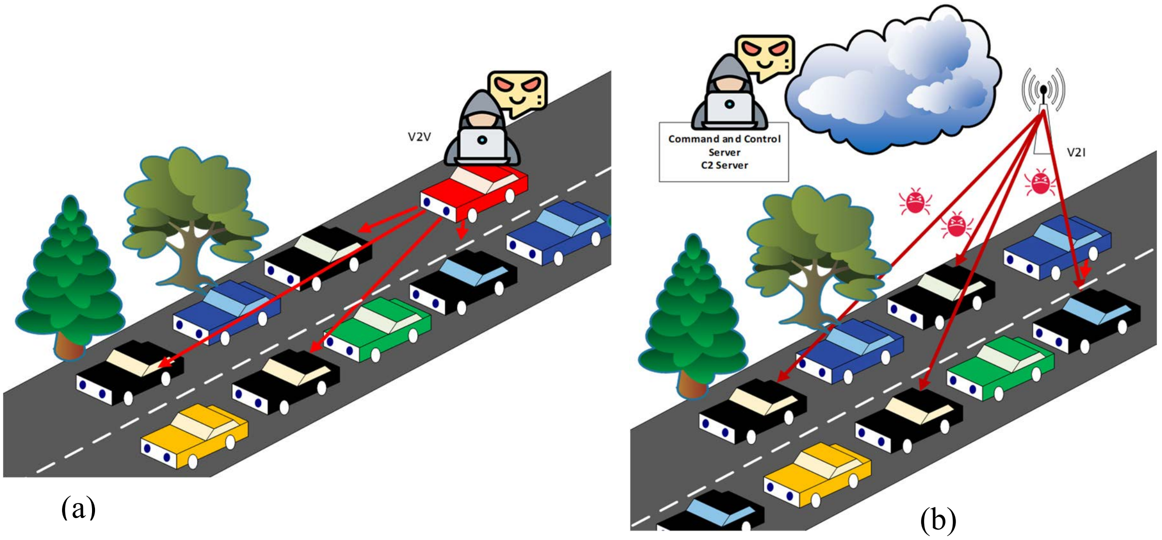

FIGURE 3. The attack scenarios (a) V2V-based C2 Server and (b) V2I-based C2 Server.

different speeds, vehicles go in and out of other communication ranges causing intermittent communication. Thus, applications will lack the granularity of mobile information leading to inaccurate information and wrong decision due to the uncertainty of the information. To address these problems, many broadcasting schemes have been designed for efficient broadcasting. Such schemes utilize the fact that vehicle movement is highly predictable except during maneuvering or road events. Thus, some broadcasting schemes encourage vehicles to broadcast their mobility information once it is necessary. This study uses the broadcasting algorithm in [44] for message broadcasting. The vehicles that run the iRMDS receive the CAMs messages and predict the lost or omitted CAMs by the broadcasting scheme to approximate the original movement patterns using the Kalman filter algorithm.

D. ATTACKERS MODEL

In this study, one thousand vehicles were simulated as attackers. Attackers collude to send false traffic patterns such as simulating braking, lane changing, accelerating, and stopping. Each colluded vehicle creates a fake message that is consistent with other colluded vehicles. For example, vehicles inject realistic noise data to their estimated locations to simulate a highly uncertain environment, thus causing vehicles wrongly construct the context reference. Similar to a botnet, we assume that an attacker (bot master in the command-and-control Server, i.e., C2 Server) infected a group of vehicles. The master creates a fake mobility pattern in the C2 Server and sends the fake information to the infected vehicle's sensors (zombie vehicles). The infected vehicles form the mobility messages and broadcast them. Because most of the recent context-aware misbehavior detection are constructed based on an honest majority, the context reference will be misled and wrongly constructed, causing high false alarms and low detection rate. Figure 3 shows the attacker scenario where the attacker can be on the road (V2V) or connected through infrastructure (V2I).

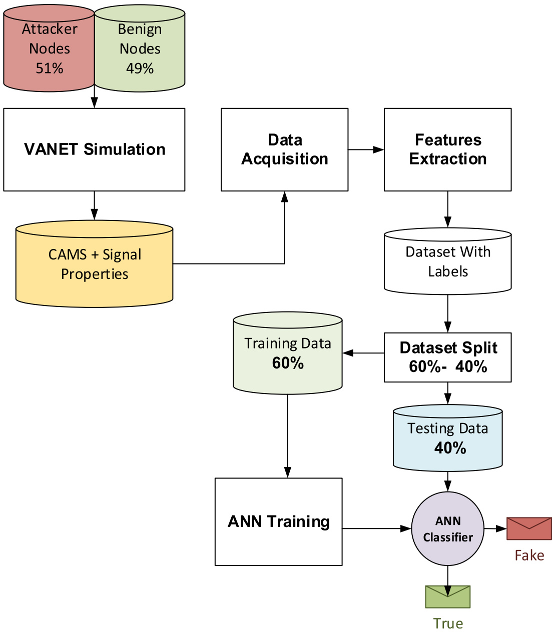

FIGURE 4. The experimental procedure.

E. TRAINING OF THE PROPOSED MDS MODEL

The mobility messages collected in each scenario and signal properties are integrated using the Kalman filter algorithm, and the innovation sequence is stored in the dataset with the labels, as shown in Figure 4. Then the dataset is splinted into two subsets 60% for training and 40% for testing. The ANN is trained and used for testing. As mentioned in Section III, part C, the proposed ANN classifier, namely the Multilayer Perceptron classifier, consists of three layers: an input layer with six input neurons, a hidden layer with 15 hidden

TABLE 2. Performance evaluation.

	Acc	Re	Pre	F-M	FPR	FNR
The Proposed iRMDS	**95.02**	**86.70**	**89.08**	**87.86**	**2.77**	**3.53**
RMDS	93.56	83.23	86.17	84.67	3.63	4.58
FCA-MDS	92.88	82.65	84.18	83.38	4.27	4.27
CA-EC-MDS	90.98	66.18	88.18	75.50	2.33	9.15

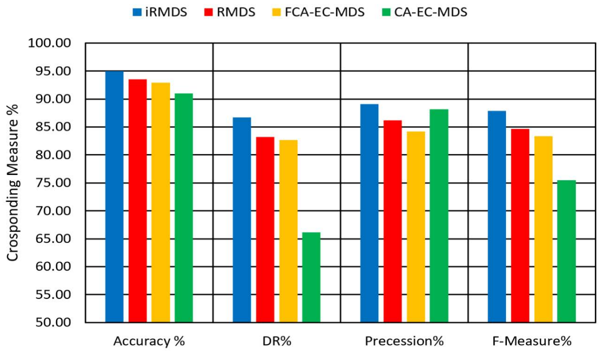

FIGURE 5. Results comparison in terms of accuracy, detection rate, precession, and F-measure.

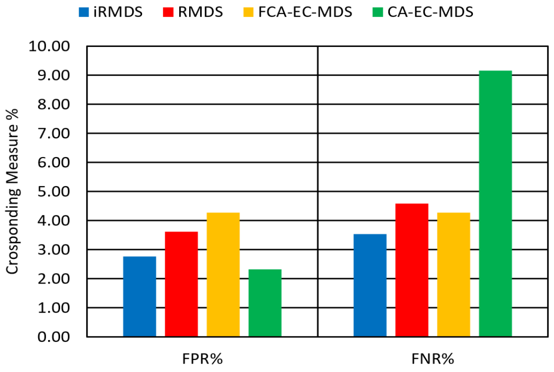

FIGURE 6. Results comparison in terms of false-positive and false-negative rates.

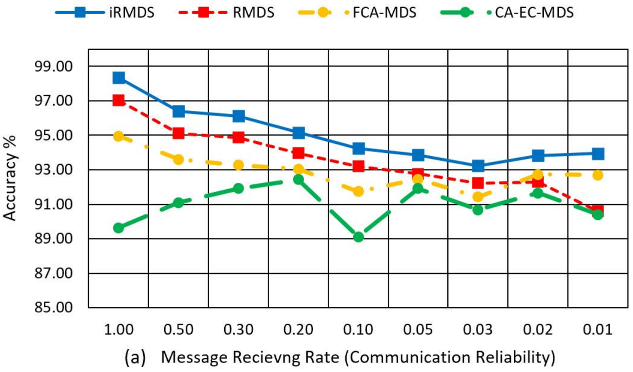

FIGURE 7. Impact of communication status on the detection.

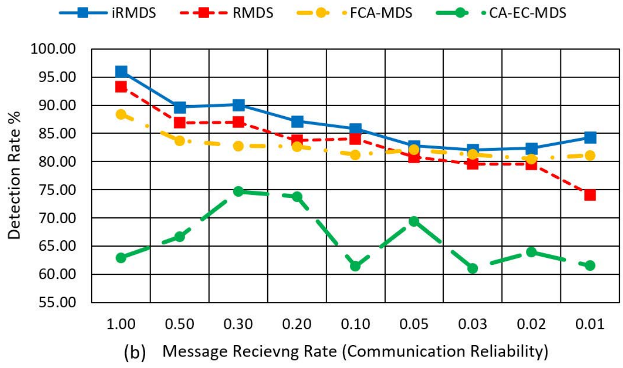

FIGURE 8. Impact of communication status on the detection rate.

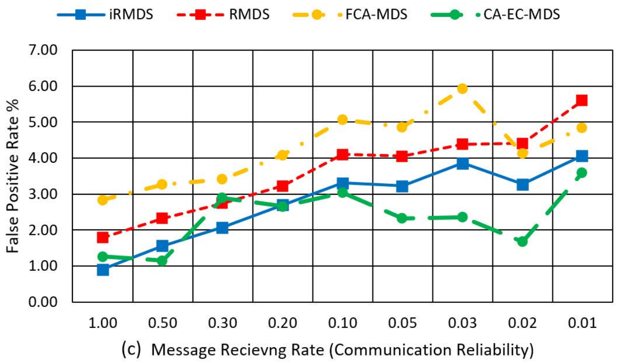

FIGURE 9. Impact of communication status on the false alarm rate.

neurons, and an output layer that contains one neuron for binary classification.

V. RESULTS AND DISCUSSION

The experimental results are presented in Table 2, Figures 5, 6, 7, 8, 9 and 10. Figure 5 presents a performance comparison between the proposed iRMDS and the related work RMDS in terms of accuracy, recall, precession, and F-Measure. Meanwhile, Table 3 lists the percentage of the improvement made by the proposed iRDMS as compared with the related work (RMDS). As mentioned earlier iRMDS has replaced the manually collaborated statistical threshold by ANN classifier. RMDS thresholds were tuned until the best accuracy was obtained (threshold = 3).

As can be seen in Table 2, the proposed iRMDS achieved the best performance measure compared to the existing models. It achieves 95.02%, 86.70%, 89.08, and 87.86% in terms of accuracy, recall, precession, and F-measure, respectively. Meanwhile, the false positive rate and the false-negative rate, which are achieved by the iRMDS model are reduced to 2.77% and 3.53%, respectively. Figures 4 and 5 depict the Comparison between the proposed model and the related studied models. In Figure 5, accuracy, recall, precession, and F-measure are presented, while Figure 6 presents the Comparison in terms of false-positive and false-negative rates.

Figures 7, 8, 9, and 10 demonstrate the impact of context change in VANET on the performance of the proposed iRMD model. As can be seen in Figure 6, the accuracy of the model has been tested based on nine context scenarios. These context scenarios represent the different road conditions that impact data accuracy and communication reliability. The first

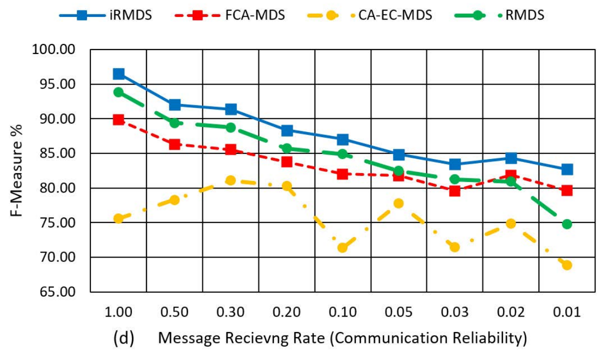

FIGURE 10. Impact of communication status on the f-measures.

TABLE 3. Improvement percentage.

Model	Acc	Re	Pre	F-M	FPR	FNR
RMDS	1.46	3.47	2.91	3.19	-0.86	-1.05
FCA-MDS	2.14	4.05	4.9	4.48	-1.5	-0.74
CA-EC-MDS	4.04	20.52	0.9	12.36	0.44	-5.62

context scenario is the ideal context scenario where there is no communication loss. The absence of communication loss leads to an accurate prediction by the Kalman filter and thus clear distinguishing between benign and fake messages. As can be noticed in all tested models, generally, the accuracy degrades when the communication loss increases. The accuracy of the CA-EC-MDS fluctuant due to the use of statistical thresholds that are constructed based on the honest majority. The attackers can influence the thresholds and lead to a decrease the accuracy. Similarly, the accuracy achieved by the FCA-MDS slightly fluctuant when the communication loss increases. Both RMDS and iRMDS are more stable due to the use of robust features that the attacker cannot manipulate. The degradation of accuracy is due to the increase in the discrepancy of innovation error. That is, in highly reliable communication scenarios, the innovation error of the Kalman filter resulted from benign messages becoming more consistent and thus more distinguishable from fake messages due to high inconsistencies, while in low communication reliable scenarios, the discrepancy resulting from benign messages increases, which led to lower accuracy. In other words, in low communication reliability scenario, the discrimination between benign and fake messages is difficult. This interprets the low detection rate in the CA-EC-MDS and the high false alarms produced by the FCA-MDS, as shown in Figures 8 and 9.

As can be noticed from Figure 10, the overall performance in terms of F-Measure indicates that the proposed model achieves the highest performance compared to other studied models. The reason is that the proposed model strikes a better balance between the precession and recall while the other models improve one measure but favor the other performance measures, as seen in Figures 8 and 9.

To sum up, the proposed iRMDS outperforms the other studied models, RMDS, FCA-MDS, and CA-EC-MDS, with all tested scenarios. The iRMDS attains 98.71%, 97.57%, 98.01%, and 97.79% for accuracy, precession, recall, and F-measure respectively (see Table 2). The iRMDS achieves 2%, 2.97%, 3.90%, and 3.44% for accuracy, precession, recall, and F-measure, respectively (see Table 3). The false alarm rates and false-negative rates are reduced by 2.93% and 1.61%, respectively. These results prove that the proper design of the MDS with independent attack features such as AoA and RSSI improves the overall permeance and decreases the false alarms (FPR) and the false-negative rate.

VI. CONCLUSION

This study presents an improved robust misbehavior detection scheme. The static detection threshold, which is based on an honest majority, has been replaced by the artificial neural network-based classifier. This classifier has been trained based on attackers-independent features that created the signal properties of the cooperative awareness messages. The proposed scheme consists of three phases: a data acquisition phase features extraction phase and a misbehavior detection phase. In the data acquisition phase, Kalman filter-based acquisition algorithms were used to filter the environmental noises before being broadcasted to neighboring vehicles. A driving-situation-aware adaptive broadcasting scheme is used to broadcast critical cooperative awareness messages to neighboring vehicles. In the features extraction phase, signal properties such as the angle of arrival (AoA) and the received signal strength indicator (RSSI) are integrated using the Kalman Filter algorithm with the CAMs information for feature extraction. The inconsistencies in terms of Kalman innovation errors were used as features for the next stage. In the detection phase, an artificial neural network-based classifier was trained using the innovation error of the Kalman filter and used for detection. Experimental results using a realistic dataset, namely the NGSIM dataset, shows that the proposed model outperforms the related work. The overall accuracy performance improved by 3.44%.

Although these results are promising, the main drawback of this study is the use of a supervised learning approach in a highly dynamic environment. Such results may change according to the scenario, and in-depth analysis is required. In addition, this study didn't consider the attackers who tamper with the transmission range. In our feature research, more investigation will be carried on when such an attack is considered with different scenarios. In addition, this study is limited to the traffic flow in the NGSIM dataset in which vehicles' velocities are less than 100 Kilometers per hour. Further investigations should be carried out to evaluate the proposed model in high-mobility scenarios.

REFERENCES

[1] WHO. (2020). *10 Facts on Global Road Safety*. [Online]. Available: http://www.who.int/features/factfiles/roadsafety/en/

[2] F. A. Ghaleb, B. A. S. Al-Rimy, A. Almalawi, A. M. Ali, A. Zainal, M. A. Rassam, S. Z. M. Shaid, and M. A. Maarof, "Deep Kalman neuro fuzzy-based adaptive broadcasting scheme for vehicular ad hoc network: A context-aware approach," *IEEE Access*, vol. 8, pp. 217744–217761, 2020.

[3] Z. Lin, Y. Sun, Y. Tang, and Z. Liu, "An efficient message broadcasting MAC protocol for VANETs," *Wireless Netw.*, vol. 26, pp. 6043–6057, Jul. 2020.

[4] M. Sepulcre, J. Gozalvez, and J. Hernandez, "Cooperative vehicle-to-vehicle active safety testing under challenging conditions," *Transp. Res. C, Emerg. Technol.*, vol. 26, pp. 233–255, Jan. 2013.

[5] C.-M. Huang and S.-Y. Lin, "Cooperative vehicle collision warning system using the vector-based approach with dedicated short range communication data transmission," *IET Intell. Transp. Syst.*, vol. 8, no. 2, pp. 124–134, Mar. 2014.

[6] K. Bylykbashi, E. Qafzezi, M. Ikeda, K. Matsuo, and L. Barolli, "Fuzzy-based driver monitoring system (FDMS): Implementation of two intelligent FDMSs and a testbed for safe driving in VANETs," *Future Gener. Comput. Syst.*, vol. 105, pp. 665–674, Apr. 2020.

[7] F. A. Ghaleb, F. Saeed, M. Al-Sarem, B. A. S. Al-Rimy, W. Boulila, A. E. M. Eljialy, K. Aloufi, and M. Alazab, "Misbehavior-aware on-demand collaborative intrusion detection system using distributed ensemble learning for VANET," *Electronics*, vol. 9, no. 9, p. 1411, Sep. 2020.

[8] F. A. Ghaleb, M. A. Maarof, A. Zainal, M. A. Rassam, F. Saeed, and M. Alsaedi, "Context-aware data-centric misbehaviour detection scheme for vehicular ad hoc networks using sequential analysis of the temporal and spatial correlation of the consistency between the cooperative awareness messages," *Veh. Commun.*, vol. 20, Dec. 2019, Art. no. 100186.

[9] F. A. Ghaleb, M. A. Maarof, A. Zainal, B. A. S. Al-Rimy, A. Alsaeedi, and W. Boulila, "Ensemble-based hybrid context-aware misbehavior detection model for vehicular ad hoc network," *Remote Sens.*, vol. 11, no. 23, p. 2852, Dec. 2019.

[10] X. Zhang, C. Lyu, Z. Shi, D. Li, N. N. Xiong, and C.-H. Chi, "Reliable multiservice delivery in fog-enabled VANETs: Integrated misbehavior detection and tolerance," *IEEE Access*, vol. 7, pp. 95762–95778, 2019.

[11] J. Kamel, M. Wolf, R. W. van der Hei, A. Kaiser, P. Urien, and F. Kargl, "VeReMi extension: A dataset for comparable evaluation of misbehavior detection in VANETs," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[12] S. Ercan, M. Ayaida, and N. Messai, "Misbehavior detection for position falsification attacks in VANETs using machine learning," *IEEE Access*, vol. 10, pp. 1893–1904, 2022.

[13] C. Zhang, K. Chen, X. Zeng, and X. Xue, "Misbehavior detection based on support vector machine and dempster-Shafer theory of evidence in VANETs," *IEEE Access*, vol. 6, pp. 59860–59870, 2018.

[14] A. Saidi, K. Benahmed, and N. Seddiki, "Secure cluster head election algorithm and misbehavior detection approach based on trust management technique for clustered wireless sensor networks," *Ad Hoc Netw.*, vol. 106, Sep. 2020, Art. no. 102215.

[15] F. A. Ghaleb, M. A. Maarof, A. Zainal, B. A. S. Al-Rimy, F. Saeed, and T. Al-Hadhrami, "Hybrid and multifaceted context-aware misbehavior detection model for vehicular ad hoc network," *IEEE Access*, vol. 7, pp. 159119–159140, 2019.

[16] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, "Misbehavior detection and efficient revocation within VANET," *J. Inf. Secur. Appl.*, vol. 46, pp. 193–209, Jun. 2019.

[17] F. A. Ghaleb, A. Zainal, M. A. Maroof, M. A. Rassam, and F. Saeed, "Detecting bogus information attack in vehicular ad hoc network: A context-aware approach," *Proc. Comput. Sci.*, vol. 163, pp. 180–189, Jan. 2019.

[18] F. A. Ghaleb, A. Zainal, M. A. Rassam, and F. Mohammed, "An effective misbehavior detection model using artificial neural network for vehicular ad hoc network applications," in *Proc. IEEE Conf. Appl., Inf. Netw. Secur. (AINS)*, Nov. 2017, pp. 13–18.

[19] M. Alzahrani, M. Y. Idris, F. A. Ghaleb, and R. Budiarto, "Robust misbehavior detection scheme for vehicular network," in *Proc. Int. Conf. Data Sci. Appl. (ICoDSA)*, Oct. 2021, pp. 54–60.

[20] R. Sultana, J. Grover, and M. Tripathi, "A novel framework for misbehavior detection in SDN-based VANET," in *Proc. IEEE Int. Conf. Adv. Netw. Telecommun. Syst. (ANTS)*, Dec. 2020, pp. 1–6.

[21] N.-W. Lo and H.-C. Tsai, "Illusion attack on VANET applications—A message plausibility problem," in *Proc. IEEE Globecom Workshops*, Nov. 2007, pp. 1–8.

[22] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, "Trust model for secure group leader-based communications in VANET," *Wireless Netw.*, vol. 25, no. 8, pp. 4639–4661, Nov. 2019.

[23] X. Li, J. Liu, X. Li, and H. Li, "A reputation-based secure scheme in vehicular ad hoc networks," *Int. J. Grid Utility Comput.*, vol. 6, no. 2, pp. 83–90, 2015.

[24] S. A. Soleymani, A. H. Abdullah, W. H. Hassan, M. H. Anisi, S. Goudarzi, M. A. R. Baee, and S. Mandala, "Trust management in vehicular ad hoc network: A systematic review," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, pp. 1–22, Dec. 2015.

[25] M. Fogue, F. J. Martinez, P. Garrido, M. Fiore, C.-F. Chiasserini, C. Casetti, J.-C. Cano, C. T. Calafate, and P. Manzoni, "Securing warning message dissemination in VANETs using cooperative neighbor position verification," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2538–2550, Jun. 2015.

[26] S. A. Soleymani, A. H. Abdullah, M. Zareei, M. H. Anisi, C. Vargas-Rosales, M. K. Khan, and S. Goudarzi, "A secure trust model based on fuzzy logic in vehicular ad hoc networks with fog computing," *IEEE Access*, vol. 5, pp. 15619–15629, 2017.

[27] F. A. Ghaleb, F. Saeed, E. H. Alkhammash, N. S. Alghamdi, and B. A. S. Al-Rimy, "A fuzzy-based context-aware misbehavior detecting scheme for detecting rogue nodes in vehicular ad hoc network," *Sensors*, vol. 22, no. 7, p. 2810, Apr. 2022.

[28] X. Y. Tian, Y. H. Liu, J. Wang, W. W. Deng, and H. Oh, "Computational security for context-awareness in vehicular ad-hoc networks," *IEEE Access*, vol. 4, pp. 5268–5279, 2016.

[29] H. Vahdat-Nejad, A. Ramazani, T. Mohammadi, and W. Mansoor, "A survey on context-aware vehicular network applications," *Veh. Commun.*, vol. 3, pp. 43–57, Jan. 2016.

[30] K. Emara, W. Woerndl, and J. Schlichter, "CAPS: Context-aware privacy scheme for VANET safety applications," in *Proc. 8th ACM Conf. Secur. Privacy Wireless Mobile Netw. (WiSec)*, Jun. 2015, pp. 1–12.

[31] X. Lin and R. Lu, "Context-aware cooperative authentication," in *Vehicular Ad Hoc Network Security and Privacy*. Hoboken, NJ, USA: Wiley, 2015, p. 216.

[32] J. Wan, D. Zhang, S. Zhao, L. T. Yang, and J. Lloret, "Context-aware vehicular cyber-physical systems with cloud support: Architecture, challenges, and solutions," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 106–113, Aug. 2014.

[33] S. Al-Sultan, A. H. Al-Bayatti, and H. Zedan, "Context-aware driver behavior detection system in intelligent transportation systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4264–4275, Nov. 2013.

[34] *IEEE Standard for Wireless Access in Vehicular Environments Security Services for Applications and Management Messages*, IEEE Standard 1609.2-2013 (Revision of IEEE Standard 1609.2-2006), 2013, pp. 1–289.

[35] F. A. Ghaleb, A. Zainal, M. A. Rassam, and A. Abraham, "Improved vehicle positioning algorithm using enhanced innovation-based adaptive Kalman filter," *Pervasive Mobile Comput.*, vol. 40, pp. 139–155, Sep. 2017.

[36] S. Najafzadeh, N. Ithnin, S. A. Razak, and R. Karimi, "BSM: Broadcasting of safety messages in vehicular ad hoc networks," *Arabian J. Sci. Eng.*, vol. 39, no. 2, pp. 777–782, Feb. 2014.

[37] K. Z. Ghafoor, K. A. Bakar, M. van Eenennaam, R. H. Khokhar, and A. J. Gonzalez, "A fuzzy logic approach to beaconing for vehicular ad hoc networks," *Telecommun. Syst.*, vol. 52, no. 1, pp. 139–149, 2013.

[38] F. Lyu, N. Cheng, H. Zhu, H. Zhou, W. Xu, M. Li, and X. Shen, "Towards rear-end collision avoidance: Adaptive beaconing for connected vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 1248–1263, Feb. 2021.

[39] Y. Allouche and M. Segal, "Cluster-based beaconing process for VANET," *Veh. Commun.*, vol. 2, no. 2, pp. 80–94, Apr. 2015.

[40] S. Zemouri, S. Djahel, and J. Murphy, "A short-term vehicular density prediction scheme for enhanced beaconing control," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–7.

[41] K. Z. Ghafoor, J. Lloret, K. A. Bakar, A. S. Sadiq, and S. A. B. Mussa, "Beaconing approaches in vehicular ad hoc networks: A survey," *Wireless Pers. Commun.*, vol. 73, no. 3, pp. 885–912, May 2013.

[42] A. Mchergui, T. Moulahi, M. T. B. Othman, and S. Nasri, "Enhancing VANETs broadcasting performance with mobility prediction for smart road," *Wireless Pers. Commun.*, vol. 112, pp. 1629–1641, Jan. 2020.

[43] A. Srivastava, A. Prakash, and R. Tripathi, "Fuzzy-based beaconless probabilistic broadcasting for information dissemination in urban VANET," *Ad Hoc Netw.*, vol. 108, Nov. 2020, Art. no. 102285.

[44] F. A. Ghaleb, A. Zainal, M. A. Rassam, and F. Saeed, "Driving-situation-aware adaptive broadcasting rate scheme for vehicular ad hoc network," *J. Intell. Fuzzy Syst.*, vol. 35, no. 1, pp. 423–438, Jul. 2018.

[45] C. Thiemann, M. Treiber, and A. Kesting, "Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2088, no. 1, pp. 90–101, Jan. 2008.

MOHAMMED ALZAHRANI received the bachelor's degree in computer engineering from the Albaha College of Science, Saudi Arabia, in 2008, and the M.Eng. degree in internetworking from Dalhousie University, Halifax, NS, Canada, in 2018. He is currently pursuing the Ph.D. degree with Universiti Teknologi Malaysia. He is currently a Systems Engineer with the Ministry of Communications and Information Technology, Saudi Arabia. His research interests include communication networks, particularly security, and trust in vehicle ad hoc networks.

MOHD YAZID IDRIS received the M.Sc. degree in software engineering, in 1998, and the Ph.D. degree in information technology (IT) security, in 2008. In software engineering, he focuses on the research of designing and development of mobile and telecommunication software. He is currently an Associate Professor with the Faculty of Engineering and the Faculty of Computing, Universiti Teknologi Malaysia. His main research interest includes IT security in the area of intrusion prevention and detection (IPD).

FUAD A. GHALEB received the B.Sc. degree in computer engineering from Sana'a University, Sana'a, Yemen, in 2003, and the M.Sc. and Ph.D. degrees in computer science information security from the Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia, in 2014 and 2018, respectively.

From 2004 to 2012, he was a Lecturer of network and computer engineering at Sana'a Community College, Sana'a. He is currently a Senior Lecturer of cybersecurity with the Faculty of Computing, Universiti Teknologi Malaysia. He is the author of 68 articles related to information and network security. His research interests include vehicular network security, cyber threat intelligence, intrusion detection, data science, data mining, and knowledge discovery. He was a recipient of many awards and recognitions, such as the Postdoctoral Fellowship Award, the Best Postgraduate Student Award, and the Best Presenter Award from the Faculty of Engineering, School of Computing, UTM; and the Best Papers Awards from IICIST, Kuala Lumpur, Malaysia, and Effat University, Jeddah, Saudi Arabia.

RAHMAT BUDIARTO received the B.Sc. degree in mathematics from the Bandung Institute of Technology, Indonesia, in 1986, and the M.Eng. and D.Eng. degrees in computer science from the Nagoya Institute of Technology, Japan, in 1995 and 1998, respectively. He is currently a Full Professor with the College of Computer Science and Information Technology, Al-Baha University, Saudi Arabia. His research interests include intelligent systems, brain modeling, IPv6, network security, wireless sensor networks, and MANETs.

• • •