Detecting false messages in vehicular ad hoc networks based on a traffic flow model

Jizhao Liu1, Weidong Yang2, Junbao Zhang1 and Changlin Yang1

Abstract
In vehicular ad hoc networks, inside attackers can launch a false information attack by injecting false emergency messages to report bogus events such as traffic accidents. In this article, a false message detection scheme is proposed and evaluated. First, traffic flow theory is employed to analyze vehicular behavior under a traffic accident scenario. It shows that a “bottleneck” phenomenon is triggered because the road capacity is reduced by blocked lanes at an accident site. The traffic parameters, such as vehicular density, exhibit a distinct statistical property compared to an accident-free site. Based on this, a false message detection algorithm is proposed in which the traveling vehicles are exploited as witnesses to collect traffic parameters, and their observation data are used as evidence to feed a traffic flow model. A Bayesian theorem–based method is used to calculate the likelihood for each traffic scenarios, and the actual traffic condition is estimated to determine whether the reported accident has actually occurred. Finally, the performance of the proposed scheme was verified through simulations in a realistic traffic scenario. It was shown that a higher detection accuracy could be obtained compared to previously proposed approach.

Keywords
VANETs, security, intrusion detection, traffic flow theory

Date received: 4 January 2019; accepted: 22 January 2020

Handling Editor: Antonella Molinaro

Introduction
In recent years, vehicular ad hoc networks (VANETs) have received much attention from academics and industry because a variety of VANETs applications have emerged for road safety, passenger comfort, and traffic efficiency. In VANETs, vehicles are equipped with wireless access vehicular environment (WAVE) devices, which enable them to communicate with each other (vehicle-to-vehicle, V2V), and with pre-deployed roadside units (RSUs; vehicle-to-infrastructure, V2I). Road safety applications are expected to decrease road accidents in VANETs, and this includes cooperative collision warning (CCW), road hazard notification (RHN), and post-crash notifications (PCN). In these applications, vehicles are exploited as a “moving sensor” for collecting traffic information, and they are allowed to broadcast two types of messages: (1) periodic beacon messages, which are used to show the present of a vehicle in the network, and (2) emergency messages, which are used to report the occurrence of damaging events (such as a traffic accident and congestion). These messages can help drivers beware of the traffic situation and hazard events that are beyond the horizon. However, these benefits can only be

1School of Computer Science, Zhongyuan University of Technology, Zhengzhou, P.R. China
2Henan University of Technology, Zhengzhou, P.R. China

Corresponding author:
Jizhao Liu, School of Computer Science, Zhongyuan University of Technology, No. 41, Zhongyuan Middle Road, Zhengzhou 450007, Henan Province, P.R. China.
Email: jizhao.liu@zut.edu.cn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
realized only if the messages are reliable, in other words, if they reflect real world honestly and correctly. Inside attackers can inject false messages into the network when they are motivated out of selfishness to make a malicious attack. For example, by reporting a bogus traffic congestion, a selfish vehicle may try to create the illusion of traffic congestion to mislead other vehicles to exit the current road in order to reach their destination faster. In addition, hardware faults can also result in incorrect data. The drivers may be misled by unreliable messages to take wrong reactions, such as hard brake or switching to an alternative route, which results in journey delay or traffic disturbance, and in extreme cases, an accident could occur. Hence, effectively detecting false messages in VANETs is very important.

Generally, a VANETs message can be considered to be reliable if the following two conditions are satisfied: (1) the sender is a valid node, and the message integrity is protected well against malicious injection, modification, and replay attacks, and (2) the information contained in the message honestly and correctly reflects the real world. Some existing VANETs security mechanisms focus on the first requirement by the use of digital signature and authentication technology. However, compared to launching attacks on a protocol stack, it is much easier attack for an inside attacker to inject false information with validly signed messages because an inside attacker commonly has the certificate and key distributed by the certificate authority.

In recent years, various schemes have been proposed to detect false messages injected by an inside attacker. In literature, the concept of reputation/trust is used to represent to what degree a vehicle can be determined to be trustworthy. Each vehicle observes the behavior of its neighbors and uses a numeric score to represent the node’s credibility. A central reputation server is responsible for storing and updating the reputation/trust value of vehicles in the network. However, VANETs have some unique features such as a large geographical range and network consisting of millions of nodes. High node mobility generally results in frequent topology changes that make it very difficult to maintain real-time querying and updating of the reputation/trust score. Therefore, applying a centralized reputation/trust mechanism in VANETs has long been debatable. To solve this problem, data-centric detection schemes have been proposed to reduce the network overhead and improve efficiency using fully distributed and localized detection algorithms, in which the vehicles located at the scene are exploited as “witness,” and they perform cooperatively and locally the detection algorithm to collect possible evidence for verifying the correctness of safety messages. Examples include the plausibility and consistency checking scheme, trajectory-based detection classifier framework, and heartbeat-based detection. However, these efforts focus on how the evidence are collected and used, and few works attempt to intensively analyze the inherent characteristics of the traffic data itself.

In real-world traffic scenario, the occurrence of abnormal traffic events can commonly be reflected by a change of traffic parameters. For example, in a car collision scenario, parts of the lanes may be blocked by crashed cars and the road capacity can be reduced by them. Since the arrived newly vehicles cannot pass through in a timely manner, they brake and change lanes to merge into the remaining unobstructed lanes. Queuing or traffic jams can be formed during rush hours, and a “bottleneck” phenomenon can be triggered at the accident site. As a result, traffic parameters, such as density and speed, exhibit abnormal fluctuations. These provide us with a feasible approach to infer whether an accident reported by an emergency message has actually happened. In this article, we extract two typical traffic patterns—lane-blocking and blocking-free—and employ a traffic flow model to estimate the probability density function (PDF) of traffic parameters. Based on this, a Bayesian inference–based false message detection algorithm is proposed. Our contributions in this article are summarized as follows:

1. Two typical traffic patterns are extracted, and a traffic flow model is built to analyze the characteristics of vehicular behavior and to estimate the PDF of traffic parameters under each traffic pattern.

2. A false message detection algorithm is proposed, in which the vehicles located at the scene are exploited as witnesses to infer the actual traffic condition. Based on their observational data, a Bayesian approach is used to calculate the likelihood of each traffic pattern and to determine whether the accident reported by the emergency message has actually happened.

3. A simulation experiment was conducted for this article to validate the proposed algorithm.

The remainder of this article is organized as follows. We discuss some related works in section “Related works,” before describing the VANETs model and traffic flow model in section “System models.” In section “The proposed scheme,” we introduce the detailed design of the proposed algorithm. The results of performance evaluation are presented in section “Simulation.” Finally, we draw our conclusions in section “Conclusion.”

Related works

In recent years, there has been much researches aimed at detecting false messages in VANETs. In our opinion,
the methods can be classified into two categories: (1) node-centric and (2) data-centric schemes. Node-centric schemes try to analyze the senders’ behavior, such as the packet and message pattern, and compare them with the normal nodes to identify the attackers. Deviated nodes are deemed as malicious, for example, the nodes whose message sending rate is significantly higher than the average value. Furthermore, a numeric score, such as reputation value, is used to quantify the healthy degree of a node. Li et al. proposed a reputation-based announcement scheme. When an announcement arrives, the recipient determines whether to accept or reject it by using the sender’s reputation score, which reflects the extent to which the sender has announced reliable messages in the past. The reputation score of a sender is computed based on feedback from its neighboring vehicles. They tend to give positive feedback for reliable messages, or a negative feedback for unreliable messages, which increases or decreases the senders’ reputation score. A reputation server is responsible for collecting, updating, and certifying the reputation score. Even though reputation/trust-based schemes have been well studied in self-organized and peer-to-peer networks, they are impractical in our case because of the large network size and high node mobility of VANETs. During the detection process, reputation-based schemes require multi-round communication (vehicle-to-vehicle, vehicle-to-infrastructure, and vehicle-to-reputation server), which leads to high detection delay and heavy communication overheads.

Data-centric trust has been proposed to solve these problems. It attempts to evaluate the reliability of the messages rather than identify misbehaving nodes. In Rui et al., based on locally observed information, each vehicle independently decides whether a received message is false. To defend against sybil attacks, this scheme does not use any external inter-vehicle cooperation or majority voting mechanism. In LEAVE (local eviction of attacker by voting evaluators), each vehicle runs an intrusion detection system (IDS) to monitor the behavior of its neighbors. If an attacker sends false data, its neighbors will identify a significant deviation from the honest data reported by other neighbors. A voting procedure is triggered, and each vehicle launches an accusation against the sender. The sender is evicted from the network only if the number of accusations against it exceeds a given threshold. Threshold-based schemes are built on the assumption that each witness sends an alert message to report the observed traffic event. A vehicle accepts that a message is true only if the number of identical messages surpasses a predefined threshold. In practice, the main issue of threshold-based schemes is that the threshold is very difficult to choose. Too high of a value leads to valid messages being rejected; too low a threshold offers little defense against attacks launched by multiple colluding malicious vehicles. In addition, since there is typically some time period between the occurrence of a traffic event and the vehicles reporting it, the alert messages are not accepted until a sufficient number of vehicles have sent reports. This leads to an unnecessary delay. Sedjelmaci et al. proposed an efficient and lightweight intrusion detection mechanism for vehicular network (ELIDV) to detect internal malicious vehicles. ELIDV first evaluates the number of intrusion detection agents located within the wireless communication range and then uses a set of rules to evaluate the credibility of each node. Finally, it calculates and assigns a malicious level to detect malicious vehicles, and the vehicles are classified into one of the following classes: trustworthy, uncertain, or untrustworthy. Yao et al. considered the types of applications and the authority levels of nodes and proposed a dynamic entity-centric trust model by employing the experiences and utility theory. The scheme is simple enough for real-time trust evaluation. Analyses show that it can reflect the data trustworthiness objectively and help vehicles to detect the false or bogus data.

An RSU-aided data-centric trust establishment scheme was proposed in Grover et al., in which the trust relations between the reporting vehicles and the data-consuming vehicles are decoupled. The reported data are first collected by an RSU, rather than the data-consuming vehicles, and the RSU transforms the data into evidence, calculates a trust value, and provides the data and the trust value to the data-consuming vehicles. In heartbeat-based scheme, each vehicle continuously parses beacon messages of its neighbors and try to detect the possible inconsistency in disseminated information using consecutive beacons. Because a beacon message includes position, speed, and steering angle of the sender, the sender’s further position in a short period can be predicted by using speed and steering angle contained in the current beacon. The sender can be viewed with suspicion only if its reported position does not match the predicted position.

Zaidi et al. proposed a host-based intrusion detection scheme to detect false emergency messages in VANETs, in which the vehicles located in the scene are exploited to provide their observation data on the traffic condition, and a statistical approach is used to identify the malicious vehicles that broadcast a false message. The proposed scheme is based on a fact: if an accident occurs, the crashed vehicles block part of lanes and the traffic flow exhibits decreased trend. In the scheme, a hypothesis test method is exploited to identify the changes on traffic flow statistics. However, this decrease phenomenon has not been well modeled and analyzed. To what extent does traffic flow drop and the correlation of some parameters, such as the number
of lanes, number of blocked lanes, and vehicle density, have not been analyzed quantitatively. In real-world traffic scenario, traffic condition can be variable at different hours of a day. For example, traffic is light at midnight and heavy at rush hour. It should be considered fully and been analyzed quantitatively. In this article, the traffic flow theory is exploited to conduct a road capacity analysis, and several parameters, such inflow rate, number of lanes, and number of blocked lanes, are used as the input of the model. A more general model is proposed to estimate the actual traffic condition.

**System models**

**Network model**

As shown in Figure 1, we consider a road safety application scenario in which the vehicles broadcast an emergency message to report a traffic accident. The message is relayed to all the vehicles located within a predefined geographical range so that their drivers can make a timely reaction. Also, it is assumed that the vehicles are equipped with various kinds of on-board sensors (GPS, accelerometer, radar, etc.) that enable them to obtain the motion status of themselves and the vehicles in their vicinity. For example, a vehicle can count the number of neighboring vehicles and calculate the local traffic density within a perception radius $r$ using a wireless communication–based approach or heartbeat-based schemes. The local traffic density observed by a vehicle can be represented by $\hat{\rho} = \frac{\text{num}}{\lambda r}$, where $\lambda$ is the number of lanes and $\text{num}$ is the number of vehicles located in its perception radius.

**Adversaries model**

This article focuses on false message attacks launched by inside adversaries who are a valid VANET node and have the security parameters (key and certificate) distributed by the certificate authority. They launch a false message attack by broadcasting bogus emergency messages to claim a nonexistent accident. It also inserts manipulated low speed in its beacon messages to make an illusion of traffic congestion for enhanced deception. In addition, it is assumed that there may be colluded attackers, in which multiple adversaries launch a false message attack in a cooperative style such as reporting the bogus messages with the similar content to compromise the deployed IDS. In the most of the existing IDS, multiple reports with the same content can be regarded as a stronger signal of high credibility than a single report. The detection accuracy may be degraded, or even be compromised in some cases, which results in the greater damage to the reliability of VANETs.

Moreover, we assume the multiple pseudonyms–based privacy preservation scheme is used in VANETs, in which each vehicle stores multiple pseudonyms and uses pseudonym to signature the sent messages. It leads to a risk of sybil attacks, in which an attacker behaves as if it is a large number of nodes. The attacker can send multiple false alert messages to claim bogus event. Also, it can also provide false evidence to compromise the IDS by using multiple identities. The detecting of sybil attacks is out of scope of this article, so we assume a sybil attack detection protocol has been already deployed in VANETs.

**Trust model**

A trust model is built to represent the degree of trustworthiness of the data used in the detection process. The data can be classified into two categories: (1) completely trustworthy and (2) partially trustworthy. For a detecting vehicle, the data collected from its own on-board sensors, such as camera, lidar, and radar, are assumed as trustworthy completely. Besides, because we assume colluded attackers, there may be attackers in the detector’s neighbors. The evidence data provided by them are assumed as trustworthy partially.

**Traffic model**

For thoroughly analyzing “bottleneck” phenomenon, we introduce some concepts from traffic flow theory.
to model the vehicular behaviors and their macroscopic characteristics. As shown in Figure 2, a freeway with \( \lambda \) lanes is divided into \( N \) segments \( 1, 2, \ldots, i, \ldots, N \), and \( t = 1, 2, \ldots, k, \ldots \) are the discrete timeslots. Three parameters are used to characterize the vehicular macroscopic status in every single segment: speed \( v \) (in \( \text{km/h} \)), density \( \rho \) (in vehicles/km/lane), and flow \( q \) (in vehicles/hour/lane). Vehicles enter the freeway from segment \( N \) and they leave from 1 at constant flow rate \( q_{\text{in}} \) and \( q_{\text{out}} \), and \( q_i(k) \) denotes the flow rate at which the vehicles leave segment \( i \) and enter \( i + 1 \) in timeslot \( k \). When a traffic accident occurs, one or more lanes are blocked by the crashed vehicles. We assume that the blocking occurs at segment \( b \), and the number of blocked and remaining lanes are denoted by \( \lambda_b \) and \( \lambda - \lambda_b \), respectively. The blocked lanes, playing the role of a “bottleneck,” divide the freeway into two sections: \( 1, \ldots, b - 1 \) and \( b + 1, \ldots, N \).

The correlation between density and flow rate can be regulated by the following piecewise linear function

\[
q = \begin{cases} 
\rho \times \nu_f & \rho \leq \rho_c \\
Q_M - w(\rho - \rho_c) & \rho > \rho_c 
\end{cases}
\]  

(1)

where \( w \), \( \nu_f \), \( \rho_c \), and \( Q_M \) are constant parameters, and \( Q_M \) represents the maximum flow rate. When the density reaches the critical value \( \rho_c \), Equation (1) is also called the “fundamental diagram” (Figure 3), which divides the traffic status of a segment into two phases: free-flow and congestion. When the density is less than a critical density \( \rho_c \), vehicles can travel at their desired high speed, denoted by \( \nu_f \), due to large enough space between vehicles, and the flow rate can be calculated as \( q = \rho \times \nu_f \). This phase is called the free-flow state, and the desired speed \( v_f \) is called the free-flow speed. Vehicles are forced to reduce their speed and keep a safe distance to the vehicle in front of them when the traffic density exceeds critical value \( \rho_c \). This results in a decreased flow rate with increasing traffic density, and finally it drops to zero under maximum traffic density \( \rho_f \). This phase is called a congestion state, and the flow rate can be represented as \( q = Q_M - w(\rho - \rho_c) \), where \( w \) represents the rate of decline of flow rate. The maximum of flow rate of a road segment \( Q_M \), also called “capacity,” refers to the maximum flow rate that a segment can support. This is achieved when the density is equal to \( \rho_c \).

To represent the status of the freeway at timeslot \( k \), we use the \( N \)-dimensional vector \( \rho_1(k), \ldots, \rho_N(k) \). The dynamic evolution of density in segment \( i \) with the time elapsing can be represented by the following difference equation

\[
\rho_i(k + 1) = \rho_i(k) + \frac{T_i}{I_i}(q_{i+1}(k) - q_i(k))
\]

(2)

where \( T_i \) and \( I_i \) are the length of the timeslot and the segment, respectively.

**The proposed scheme**

**Overview**

First, two typical traffic patterns, lane-blocking and blocking-free, are used to characterize a real-world traffic scenario with and without a traffic accident. Let \( y \in \{0, 1\} \) represent the credibility of the emergency message. \( y = 1 \) represents the traffic condition is under a lane-blocking pattern and when the reported accident actually happened, and \( y = 0 \) refers to a bogus accident and when the sender is an attacker. The detection region is defined as the upstream and downstream segments (see Figure 1) of the reported accident site. The vehicles in them are involved in the detection process by cooperatively evaluating the traffic density of the upstream and downstream segments, denoted by \( X = (\rho_{\text{up}}, \rho_{\text{down}}) \), as the evidence. Hence, the task of the detection algorithm is to calculate the conditional probability \( P(Y|X) \), which represents \( y \)'s probability under the given observation data \( X \).

As shown in Figure 4, the proposed algorithm is deployed at each vehicle node and it operates in a fully distributed style. When receiving an emergency message, the algorithm is triggered and an evidence

![Figure 2. Traffic model.](image)

![Figure 3. Fundamental diagram.](image)
collection process (Block I) is first performed, in which the vehicle collects its own sensor data and exchanges these data with the nearby vehicles to calculate \( \rho_{up} \) and \( \rho_{down} \). In Block II, a traffic model is responsible for estimating the probability distribution of \( \rho \) under each pattern. Finally, a Bayesian method (Block III) is employed to evaluate the likelihood of two patterns and to infer the actual traffic condition.

**Evidence collection**

First, each witness independently observes the traffic density of the segment it is located in and sends its observation data to all other witnesses. In addition, each witness receives observation data from others, and the traffic density is calculated as follows:

\[
\rho_{up} = \frac{1}{|v_{up}|} \sum_{v_{veh} \in v_{up}} \rho_{veh} \quad \text{and} \quad \rho_{down} = \frac{1}{|v_{down}|} \sum_{v_{veh} \in v_{down}} \rho_{veh},
\]

where \( v_{up} \) and \( v_{down} \) are the set of vehicles located in the upstream and downstream segments, respectively.

**Bayesian inference**

Observation data \( X \)'s distribution can be given by conditional probability \( p(X|y) \). The posterior probability, denoted by \( p(y|X) \), represents the likelihood that the current traffic condition is under a lane-blocking and blocking-free pattern with a given \( X \). One can turn the prior and conditional probability into the posterior by applying the Bayesian theorem, as follows

\[
P(y|X) = \frac{p(X|y)p(y)}{\sum_{y \in \{0,1\}} p(X|y)p(y)} \quad (3)
\]

\( P(y=0|X) \) and \( P(y=1|X) \) can be calculated by equation (3). For the upstream and downstream segments, two probabilities, \( P_{up}(y|X) \) and \( P_{down}(y|X) \), can be obtained. We use their mean values to represent the posterior probability, denoted by \( P(y|X) = 0.5 \times (P_{up}(y|X) + P_{down}(y|X)) \). The current traffic condition can be classified into the class whose posterior probability is highest with the following maximum likelihood principle

\[
y = \arg \max_{y \in \{0,1\}} \{p(X|y)p(y)\} \quad (4)
\]

The output \( y = 0 \) means that the emergency message is regarded as bogus. In this case, the message is dropped by the recipient, and its sender is evicted from the network. Otherwise, the message is accepted.

**Inferring the traffic condition**

For inferring the actual traffic condition, a remaining unsolved problem is how to obtain the conditional probability \( P(X|y) \). In this section, we analyze the dynamic evolution process of traffic density and use the concept of equilibrium to capture its stable status. The traffic density of a segment is modeled as \( \rho = \rho^e + \epsilon \), where \( \epsilon \sim N(0, \sigma^2) \) is normally distributed noise, which represents the random nature of vehicular maneuvers and possible measurement errors. \( \rho^e \) is the equilibrium value of traffic density under stable status. By using the proposed traffic model, we deduce the close-form solution of equilibrium under two traffic patterns.

**Blocking-free pattern**. In this pattern, there are no blocked lanes, and all the segments have the same capacity \( Q_M \). We use in-flow and out-flow functions, denoted by \( S_k(k) \) and \( R_k(k) \), to model the interaction of adjacent two segments

\[
q_k(k) = \min\{S_k(k), R_{k-1}(k)\} \quad (5)
\]

\[
S_k(k) = \min\{v_{y} \rho_k(k), Q_{M, i}\} \quad (6)
\]

\[
R_{k-1}(k) = \min\{Q_{M, i-1}, w(\rho_j - \rho_{k-1}(k))\} \quad (7)
\]

where \( S_k(k) \) represents the maximum outflow rate of segment \( i \). It is determined by two values: road capacity \( Q_{M, i} \) and the outflow rate of segment \( i \) under free-flow phase. Under congestion phase, \( S_k(k) \) always can reach its maximum outflow rate due to high traffic density. Hence, this phase is not considered in equation (6). \( R_{k-1}(k) \) is the maximum inflow rate of segment \( i-1 \) at time interval \([k-1, k]\). Because segment \( i-1 \) with free-flow phase is able to receive maximum traffic flow \( Q_{M, i-1} \), free-flow phase is not considered in equation (7).

**Lane-blocking pattern**. When a car collision occurs, several lanes are blocked and the upstream traffic flow must be merged into the remaining unobstructed lanes. In the blocked segment, the capacity is reduced from its normal value of \( Q_M \) to the remaining value \( Q_M' \). We use in-flow and out-flow functions, denoted by \( S_k(k) \) and \( R_k(k) \), to model the interaction of adjacent two segments

\[
q_k(k) = \min\{S_k(k), R_{k-1}(k)\} \quad (5)
\]

\[
S_k(k) = \min\{v_{y} \rho_k(k), Q_{M, i}\} \quad (6)
\]

\[
R_{k-1}(k) = \min\{Q_{M, i-1}, w(\rho_j - \rho_{k-1}(k))\} \quad (7)
\]

where \( S_k(k) \) represents the maximum outflow rate of segment \( i \). It is determined by two values: road capacity \( Q_{M, i} \) and the outflow rate of segment \( i \) under free-flow phase. Under congestion phase, \( S_k(k) \) always can reach its maximum outflow rate due to high traffic density. Hence, this phase is not considered in equation (6). \( R_{k-1}(k) \) is the maximum inflow rate of segment \( i-1 \) at time interval \([k-1, k]\). Because segment \( i-1 \) with free-flow phase is able to receive maximum traffic flow \( Q_{M, i-1} \), free-flow phase is not considered in equation (7).
and jam density are also reduced to \( \rho' = \frac{\lambda - \lambda_b}{\lambda} \rho_c \) and \( \rho'_j = \frac{\lambda - \lambda_b}{\lambda} \rho_j \). Moreover, \( q(k) \) can be rewritten by equation (8)

\[
q_i(k) = \begin{cases} 
\min\{v_j \rho_i, Q_M - w[\rho_{i-1} - \rho_c]\} & 1 \leqslant i < b + 1 \\
\min\{v_j \rho_i, Q'_M - w[\rho_{i-1} - \rho_c]\} & i = b \\
\min\{v_j \rho_i, Q'_{M} - w[\rho_{i-1} - \rho'_c]\} & i = b + 1 
\end{cases}
\]

**Definition 1: Equilibrium.** An equilibrium is an \( N \)-dimensional state vector \( \rho^e = \{\rho^e_i \mid 1 \leqslant i \leqslant N \} \) that is a solution of equations (2), (5)–(8), and satisfies \( \rho^e_i(r) = \rho^e_i \) under the given time-invariant input \( q\).

Road segment \( i \) is said to be congested if \( \rho_i > \rho_c \), and uncongested if \( \rho_i \leqslant \rho_c \). An equilibrium is said to be congested or uncongested if and only if all segments are congested or uncongested. An equilibrium is said to be mixed if some segments are congested and others are uncongested.

**Lemma 1.** \( \rho^e \) is an equilibrium if and only if

\[
q_i = q_{i-1} \quad (\text{for } 1 \leqslant i \leqslant N)
\]

**Proof.** Define \( \Delta q_i(k) = q_{i+1}(k) - q_i(k) \), because \( \rho^e \) is an equilibrium, we have \( \rho_i(k) = \rho_i(k+1) \), \( (1 \leqslant i \leqslant N) \). According to equation (2), \( \rho^e \) being an equilibrium is equivalent to \( \Delta q_i(k) = 0 \) for any timestep \( k \). So we have \( q_{i+1} = q_i \) \( (1 \leqslant i \leqslant N) \) and \( q_0 = q_1 = \cdots = q_N \).

**Lemma 2.** Considering a blocking-free scenario in which \( \lambda_b = 0 \), there is a unique uncongested equilibrium \( \rho^c \), in which segments \( 1, 2, \ldots, N \) are all uncongested and \( \rho_1 = \rho_2 = \cdots = \rho_N = v_j^{-1} q_{in} \).

**Proof.** **Existence.** Define the traffic density under an uncongested state as \( \rho^c_i = v_j^{-1} q_i \) \( (1 \leqslant i \leqslant N) \) and \( q_i \leqslant Q_M \). We know that \( \rho^c_i(k) = \rho^c_i \) satisfies equations (2) and (5)–(7). Because we assume that segment 1 can expel traffic at the maximum flow rate \( Q_M \), equation (5) can be reduced to \( q_1 = \{v_j \rho_1, Q_M\} \) for segment 1. Thus, we have \( q_1 = v_j \rho_1 \leqslant v_j \rho_c = Q_M \) and segment 1 is uncongested. By induction, it can be proved that segments \( 2, 3, \ldots, N \) are also uncongested. It remains to prove that \( \rho^c \) is an equilibrium. Because \( \rho^c \) is uncongested, equation (5) can be reduced to

\[
q_i = \min\{v_j \rho_i, Q_M\}
\]

and the flow \( q_i \) is equal to \( v_j \rho_i \), which satisfies equations (2) and (5)–(7), so \( \rho^c \) is an equilibrium.

**Uniqueness.** Suppose \( \rho^c \) is an equilibrium. Because \( v_j \rho_i \leqslant v_j \rho_c = Q_M \), equation (5) reduces to

\[
q_i = \min\{v_j \rho_i, Q_M - w(\rho_i - \rho_c)\}
\]

and if \( q_i \neq v_j \rho_i \), it must hold that \( v_j \rho_i > Q_M - w(\rho_i - \rho_c) \) and \( \rho_i > \rho_c \). This contradicts the assumption that all segments are uncongested. Thus, \( q_i \) must be equal to \( v_j \rho_i \). Because segments \( 1, 2, \ldots, N \) are all uncongested, and we have \( q_1 = q_2 = \cdots = q_N \) from Lemma 1, then \( \rho_1 = \rho_2 = \cdots = \rho_c \).

**Lemma 3.** Consider a lane-blocking scenario in which \( \lambda_b > 0 \). There is a unique uncongested equilibrium if \( q_{in} \leqslant Q_M \), or there is a unique mixed equilibrium \( \rho^e \) if \( q_{in} > Q_M \), in which segments \( 1, 2, \ldots, b-1, b \) are uncongested and segments \( b+1, b+2, \ldots, N \) are congested.

**Proof.** If \( q_{in} \leqslant Q_M \), it is equivalent to a blocking-free pattern, because the in-flow rate does not surpass the reduced capacity \( Q_M \). From Lemma 2, we know that a unique uncongested equilibrium exists.

Then, we consider the case in which \( q_{in} > Q_M \). From Lemma 1, segments \( 1, 2, \ldots, b-1 \) are all uncongested. For segment \( b \), because segment \( b-1 \) is uncongested, we have \( q_b > q_{in} \). As \( \rho_b = (\lambda - \lambda_b)/\lambda \), we have that \( v_j \rho_b \leqslant Q_M \), and thus segment \( b \) is uncongested. For segment \( b+1 \), because the acceptable flow rate is \( Q_M < Q_M \), then if this segment is uncongested, it follows that \( q_{b+1} = \min\{v_j \rho_{b+1}, Q_M\} > v_j \rho_c = Q_M \), which contradicts \( Q_M < Q_M \). Thus, segment \( b+1 \) is congested. The results for \( b+2, b+3, \ldots, N \) hold by induction.

Under lane-blocking pattern, the traffic density in upstream and downstream segments exhibits distinct value. Due to reduced road capacity, upstream vehicles cannot pass timely the blocked site. They accumulate and queue at segment \( b+1 \), which turns the segment into congestion status. In addition, the reduced capacity also decreases the rate at which vehicles leave segment \( b \) and get into \( b-1 \). In segment \( b-1, b-2, \ldots, 1 \), the traffic flow becomes sparse and the condition turns to free-flow status.
**Proposition 1.** Suppose \( \rho_0^{\text{mix}} = \{ \rho_i^t | 1 \leq i \leq N \} \) is the mixed equilibrium given in Lemma 3. Then

\[
\rho_i^t = \frac{\lambda - \lambda_b}{\lambda} \rho_i (0 < i < b) \quad (11)
\]

\[
\rho_i^t = \rho_c + w^{-1} \left( Q_M - v_j \frac{\lambda - \lambda_b}{\lambda} \rho_c \right) (b < i \leq N) \quad (12)
\]

**Proof.** From Lemma 3, we know a fact that the traffic flow rate of whole freeway is determined by segment \( b \)'s flow rate \( q_b \). In its upstream segment, because the acceptable flow is restricted due to reduced capacity, the value of \( q_b + 1, \ldots, q_N \) under the equilibrium status will decrease gradually to \( q_b \) after the lane-blocking event occurs. However, the flow rate of downstream segment also reduces to \( q_b \) because segment \( b \) cannot supply them with enough flow. The results is that all values of \( q_1, q_2, \ldots, q_N \) will converge to \( q_b \), and the system condition turns from blocking-free equilibrium to lane-blocking equilibrium. We have \( q_1 = q_2 = \cdots = q_N = Q_M \). Then, \( \rho_i^t = v_j^{-1} Q_M = v_j^{-1} \frac{\lambda - \lambda_b}{\lambda} \rho_c (1 \leq i \leq b - 1) \).

For segment \( b \), because traffic flow is merged into the unblocked lanes, \( \rho_i^t = \frac{\lambda - \lambda_b}{\lambda} \lambda_b \rho_c = \rho_c \).

For \( b < i \leq N \), because segment \( i \) is congested, we have

\[
Q_M = v_j^{-1} \frac{\lambda - \lambda_b}{\lambda} \rho_c = Q_M - w(\rho'' - \rho_2) \quad (13)
\]

and so

\[
\rho_i^t = \rho_c + w^{-1} \left( Q_M - v_j^{-1} \frac{\lambda - \lambda_b}{\lambda} \rho_c \right) (b < i \leq N) \quad (14)
\]

Above analysis gives the equilibrium values of traffic density under blocking-free and lane-blocking pattern. By using them, the PDF of traffic density can be established easily (Figure 5), where the equilibriums are denoted by \( \rho_{0f} \) and \( \rho_{0b} \). In the detecting process, the PDFs can be used by vehicles to calculate the conditional probability \( p(X|y = 0) \) and \( p(X|y = 1) \) using their observed density \( X \) as inputs. Furthermore, the likelihood of each traffic patterns \( p(y|X) \) can be evaluated using equation (3). In summary, the detailed description of the proposed algorithm is given in Algorithm 1.

**Simulation**

**Simulation setup**

A simulation was conducted to evaluate the performance of the proposed algorithm. We used Simulation of Urban Mobility (SUMO) version 0.19.0 to generate the traffic scenario used in the simulation. SUMO is a traffic simulation software that has the ability to generate highly realistic vehicular behavior by specifying road type, speed limit, and traffic flow rate. In the simulation, a three-lane, one-way straight-line freeway with 1 km length was used, and a car following Krauss mobility model was used, which is the default vehicular mobility model in SUMO. To achieve heterogeneous vehicles, we set up three different types of vehicles: small, medium, and large size, with vehicle lengths of 5, 7, and 10 m, and maximum speeds were set as 19, 17, and 15 m/s, respectively. SUMO outputs a .xml file that contains the floating car data of all vehicles in the traffic scenario. The .xml file was converted to a NS2 mobility file with a Python script traceExporter.py provided by SUMO. The network simulation was performed by using network simulator NS2 2.35. In NS2, we implemented the proposed algorithm and turn it into simulations. The two-ray ground reflection model was used as wireless propagation model, and the wireless communication range was set as 250 m. Other parameter setting used in the simulations is given in Table 1. The simulation was run repeatedly for 50 times, and Figure 9 shows the average values of all the results.

In order to evaluate the traffic model and the proposed algorithm, we used different scenarios. The first
task is to verify whether the traffic model can predict accurately the vehicular behavior. According to the above two traffic patterns, an accident-free scenario and a traffic accident scenario are set in SUMO, the outputted .xml file was parsed by a Python script, and the traffic densities were calculated and plotted. Furthermore, we inserted a percentage of attackers into the accident-free scenario to demonstrate how the attacker behavior occurs and how well the proposed algorithm works.

Two metrics were used to evaluate the performance of the proposed scheme: detection rate (DR) and false positive rate (FPR). DR refers to the ratio between the attacks that were successfully detected and all attacks.

Table 1. Parameter setting.

| Parameter       | Value                     |
|-----------------|---------------------------|
| $v_f$           | 50.0 km/h                 |
| $\rho_c$        | 38.2 vehicles/km/lane     |
| $\rho_f$        | 125.0 vehicles/km/lane    |
| $\omega$        | 22.0 km/h                 |
| $\lambda_b$     | 2                         |
| $Q_M$           | 1910.0 vehicles/hour      |
| $q_{in}$        | 840.0 vehicles/hour       |
| Simulation duration | 600 s                   |
| Protocol        | 802.11p                   |
| Bandwidth       | 10 MHz                    |
| Frequency       | 5.9 GHz                   |
| Data rate       | 6 Mbps                    |
| Wireless communication range | 250 m                  |

Figure 6. Observation density under various densities: (a) in-flow rate = 640 vehicles/hour, (b) in-flow rate = 840 vehicles/hour, and (c) in-flow rate = 1040 vehicles/hour.
The FPR is the ratio between wrong attack alerts that honest data detected as bogus data and all alerts.

Simulation results

Verifying traffic model. First, we set up an accident scenario and collected vehicular data from it to examine whether the traffic flow model could accurately predict vehicular behavior. The duration of the simulation was 600 s. At the beginning, vehicles enter the freeway from segment N. At 130 s, an accident occurs and two crashed vehicles stop at 500 m, which block two of three lanes. We collected the observed density $^r\rho$ from several sample vehicles and plotted them in Figure 6. As shown, the curve can be divided into three phrases. A sample vehicle travels in a free-flow pattern when it enters into the freeway, and a blocking-free equilibrium can be observed where the theoretical value of the traffic density is 16.8, which is consistent with the observed data. Then, the curve sharply rises and stays around 38.2, which is consistent with the lane-blocking equilibrium. This means the sample vehicle arrived at the “bottleneck” site, in which the vehicles accumulate and result in increased density. This phenomenon can be proved by the third evidence: the observed density rapidly drops to a low value of approximately 12.7 after that, which is significantly smaller than the free-flow density. This validates Lemma 3 that the vehicles become sparse because the road capacity is reduced by blocked lanes, and the flow rate at which vehicles enter the downstream region is also decreased. In addition, it can also be seen that the shape of curves have no significant difference when the in-flow rate changes, as shown in Figure 6(a)–(c), which proves that the “bottleneck” phenomenon is stable and can be clearly observed by passing vehicles.

Furthermore, we define the observation area as 350–650 m, which consists of upstream area (350–500 m) and downstream area (500–650 m). The distribution of $\hat{\rho}$ in them is shown in Figure 7. As expected, the data are close to a normal distribution, and the mean value is also close to the theoretical equilibrium values of 16.8 (blocking-free pattern), 12.7 (lane-blocking pattern, downstream site), and 38.2 (lane-blocking pattern, upstream site). Meanwhile, it can be clearly seen that the value of $\hat{\rho}$ is obviously different under two patterns, which proves that it is feasible for recognizing the traffic pattern only depending on traffic density observation.

Evaluating detection algorithm. A collusion attack scenario is simulated by inserting a percentage of attacker into the accident-free scenario, who send false emergency messages to claim a traffic accident that occurs at 500 m. The attackers take part in the detection procedure and inject bogus observed density for misleading the detection algorithm to get the wrong results. In order to eliminate these false evidence, an honest detector first uses its sensor data to calculate the observation density $^r\rho$ and estimates the actual traffic pattern by using equation (3). If the result rejects the emergency message, it means that there may be a potential attack and the reported accident could be nonexistent. The detector establishes the PDF of blocking-free pattern using Lemma 2 and filters out the received density whose difference from the equilibrium value is greater than 2 standard deviations. The remaining densities are used to calculate the average density.

We run the simulations both with and without the proposed detection algorithm, and the results are shown in Figure 8. First, the average density $\rho_{up}$ of the upstream area with 10% and 20% attackers are plotted in Figure 8(a) and (b). The attackers start to inject false evidence at 240 s, and the false value is chosen randomly in the interval $[\frac{30}{40}, C_{138}]$, which is the typical traffic congestion density in the real accident scenario. It can be can seen that there is a significant deviation after 240 s. In comparison, Figure 8(c) gives the average density with the proposed detection algorithm, and it can be seen that the density value remains stable after the start of the attack, which proves that our algorithm works well and the false evidence can be effectively
filtered out. Finally, the results of every single detector were aggregated using a vote approach proposed in Raya et al.\textsuperscript{15}

We calculate the DR and FPR of the proposed algorithm and compare it with the previously proposed scheme in Rajesh and Soumya.\textsuperscript{13} Figure 9(a) gives the DR of two algorithms under varying attacker proportion. It can be observed that the proposed scheme worked well, and all the bogus messages were detected successfully under a small proportion attackers. The DR started to drop when the attacker proportion surpassed 0.2. This is mainly due to the attackers taking advantage in number, and they provide bogus evidence to mislead honest vehicles into obtaining wrong results. However, the heartbeat-based scheme worked well under small attacker proportion, but its performance was poor when the proportion became larger. The explanation for this is that the proposed algorithm can exploit the observation data provided by both upstream and downstream vehicles. Hence, our algorithm have a better performance in resisting collusion attack.

A similar trend can be seen in Figure 9(b), which gives the FPR of the two schemes. The number of attackers has a significant impact on the performance of the detection algorithm. The FPR starts to rise when the attacker proportion reached 0.25. However, it is only a worse situation scenario. In practice, it is very hard to place plentiful attackers into a pre-selected scenario and arrange them in strategic positions.

**Conclusion**

In this work, a traffic flow model–based false message detection scheme was proposed and tested. The simulation results showed that the proposed algorithm exhibits better performance under collusion attacks compared to the previously proposed heartbeat-based scheme. The proposed scheme demonstrates the effectiveness of the traffic flow model on determining if the emergency message data are bogus based on the observation data collected from traveling vehicles. Using a traffic flow model, vehicular behavior and the value of traffic parameters under free-flow or accident scenarios can be accurately estimated, and the actual traffic pattern can be accurately inferred. It proves the feasibility in applying the traffic flow model on the false message detection problem in VANETs, without the need for any pre-deployed infrastructure.
Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (Grant No. 61772173), Key Research Project of Higher Education of Henan Province (Grant No. 18A520052), and Innovation Talent Support Program Project of University Science and Technology of Henan Province (Grant No. 19HASTIT027).

ORCID iD
Jizhao Liu https://orcid.org/0000-0001-6264-2939

References
1. Zeadally S, Hunt R, Chen YS, et al. Vehicular ad hoc networks (VANETS): status, results, and challenges. Telecommun Syst 2012; 50(4): 217–241.
2. Khekare GS and Sakhare AV. Intelligent traffic system for VANET: a survey. Int J Adv Comp Res 2012; 2(4): 99–102.
3. Arshad M, Ullah Z, Ahmad N, et al. A survey of local/cooperative-based malicious information detection techniques in VANETs. EURASIP J Wirel Comm 2018; 2018(1): 62.
4. IEEE. IEEE standard for wireless access in vehicular environments—security services for applications and management messages - redline. In: IEEE std 1609.2-2016 - redline, March 2016, pp. 1–884.
5. Van Der Heijden RW, Dietzel S, Leinmüller T, et al. Survey on misbehavior detection in cooperative intelligent transportation systems. IEEE Commun Surv Tut 2016; 21: 779–811.
6. Li Q, Malip A, Martin K, et al. A reputation-based announcement scheme for VANETs. IEEE T Veh Technol 2012; 61: 4095–4108.
7. Yang Y, Gao Z, Xuesong Q, et al. A hierarchical reputation evidence decision system in VANETs. Int J Distrib Sens N 2015; 2015: 1–11.
8. Ruj S, Cavenaghi MA, Zhen H, et al. Data-centric misbehavior detection in VANETs. In: IEEE Vehicular technology conference, San Francisco, CA, 5–8 September 2011.
9. Sedjelmaci H, Senouci SM and Abu-Rgheff MA. An efficient and lightweight intrusion detection mechanism for service-oriented vehicular networks. IEEE Internet Things 2017; 1(6): 570–577.
10. Ghosh M, Varghese A, Gupta A, et al. Detecting misbehaviors in VANET with integrated root-cause analysis. Ad Hoc Netw 2010; 8(7): 778–790.
11. Tao Y, Wei X, Yu L, et al. MisDis: an efficient misbehavior discovering method based on accountability and state Machine in VANET, 2013, https://link.springer.com/chapter/10.1007%2F978-3-642-37401-2_57.
12. Zaidi K, Milojevic M, Rakocевич V, et al. Host based intrusion detection for VANETs: a statistical approach to rogue node detection. IEEE T Veh Technol 2016; 65: 6703–6714.
13. Rajesh B and Soumya G. Heartbeat message based misbehavior detection scheme for vehicular ad-hoc networks. In: 2012 IEEE international conference on connected vehicles and expo (ICCVE 2012), Beijing, 12 December 2012, pp. 29–34.
14. Jin J and Ran B. Automatic freeway incident detection based on fundamental diagrams of traffic flow. Transp Res Record 2009; 2099: 65–75.
15. Raya M, Papadimitratos P, Aad I, et al. Eviction of misbehaving and faulty nodes in vehicular networks. IEEE J Sel Area Comm 2007; 25: 1557–1568.
16. Chen L, Ng SL and Wang G. Threshold anonymous announcement in VANETs. IEEE J Sel Area Comm 2011; 29: 605–615.
17. Shao J, Lin X, Lu R, et al. A threshold anonymous authentication protocol for VANETs. IEEE T Veh Technol 2016; 65: 1711–1720.
18. Yao X, Zhang X, Ning H, et al. Using trust model to ensure reliable data acquisition in VANETs. Ad Hoc Netw 2016; 55: 107–118.
19. Grover J, Gaur MS, Saxmi V, et al. Detection of incorrect position information using speed and time span verification in VANET, 2012, https://dl.acm.org/doi/10.1145/2388576.2388583.
20. Panichpapiboon S and Pattara-Atikom W. Exploiting wireless communication in vehicle density estimation. IEEE T Veh Technol 2011; 60: 2742–2751.
21. Song H and Lee HS. A survey on vehicle density estimation in vehicular safety communications and its challenging issues. In: International IEEE conference on intelligent transportation systems, The Hague, 6–9 October 2013.
22. Feng W. A review of traffic flow theory. J Syst Eng 1998; 3: 71–82.