On the Security of Warning Message Dissemination in Vehicular Ad Hoc Networks

Jieqiong Chen and Guoqiang Mao

***Abstract*—Information security is an important issue in vehicular networks as the accuracy and integrity of information is a prerequisite to satisfactory performance of almost all vehicular network applications. In this paper, we study the information security of a vehicular ad hoc network whose message may be tampered by malicious vehicles. An analytical framework is developed to analyze the process of message dissemination in a vehicular network where the message may be tampered by malicious vehicles randomly distributed in the network. The probability that a destination vehicle at a fixed distance away can receive the message correctly from the source vehicle is obtained. Simulations are conducted to validate the accuracy of the theoretical analysis. Our results demonstrate the impact of network topology and the distribution of malicious vehicles on the correct delivery of a message in vehicular ad hoc networks, and may provide insight on the design of security mechanisms to improve the security of message destination in vehicular networks.**

***Index Terms*—Vehicular ad hoc networks, message dissemination, security.**

I. Introduction

Interest is surging on vehicular networks and Internet of vehicles technologies due to their increasingly important role in improving road traffic efficiency, enhancing road safety and providing real-time information to drivers and passengers [1]. By deploying wireless communication infrastructure along the roadside (e.g., road-side units (RSU)), equipping vehicles with on-board communication facilities (e.g., on-board units (OBU)), and with the assistance of dedicated short-range communication (DSRC) [2] and LTE technology, two wireless communication modes: vehicle-to-infrastructure and vehicle-to-vehicle communications, are supported in vehicular networks. Through wireless communications, messages can be disseminated for vehicular network applications, including safety applications like disseminating real-time information about traffic accidents, traffic congestion or obstacles in the road, and non-safety applications such as offering value-added services (e.g., digital maps with real-time traffic status) and in-car entertainment services [3].

Coming with the convenience and advantage of wireless communications is the potential security threat that vehicular networks may present to transportation system. Different from traditional security settings, in vehicular networks, information collection and dissemination are conducted by distributed vehicles. Quite often, information may be generated by or received from a vehicle that has never been encountered before. This renders traditional security mechanisms, largely based on cryptography and key management, or trust management, futile in vehicular networks. The situation is further exacerbated by the highly dynamic topology of vehicular networks where the connections may emerge opportunistically between vehicles and the associated network topology is constantly changing. All these features of vehicular networks pose unique challenges for vehicular network security and make vehicular networks prone to attacks by malicious and/or selfish attackers who may spread false messages, tamper or drop the received messages. These security threats are likely to result in severe consequences like traffic congestion, traffic crash, even loss of lives and must be thoroughly investigated before vehicular networks can be deployed.

In this paper, we study information security vehicular ad hoc networks (VANETs), where the message may be tampered by malicious vehicles randomly distributed in the network, by investigating the probability that a destination vehicle at a fixed distance away can receive the message correctly from the source vehicle. Specifically, consider that a vehicle (i.e., the source vehicle) detecting an abnormal situation, e.g., traffic accident, slippery road, congestion etc. sends a message informing other vehicles of the situation. The message is forwarded from the source vehicle in a multi-path manner using other vehicles. We analyze the probability that a vehicle at a fixed distance away, termed the destination vehicle, can receive the message correctly from the source vehicle in the presence of malicious vehicles in between which may modify the transmitted message. The novelty and major contributions of this paper are summarized as follows:

1) We develop for the first time an analytical framework to model the process of message dissemination in vehicular ad hoc networks in the presence of malicious vehicles randomly distributed in the network. The probability that a message is delivered correctly from the source vehicle to a destination vehicle at a fixed distance away is derived.
2) Simulations are conducted to establish the accuracy of the analysis. Using the analysis, relationship is revealed between key parameters such as the probability of correct message delivery and its major performance-impacting parameters. Discussions are presented on the impact of network topology and the distribution of malicious vehicles on secure message delivery in vehicular networks.
3) Our results may provide insight on the design of security mechanisms, particularly secure routing algorithms and topology control algorithms, to improve informations security in vehicular networks.

The rest of this paper is organized as follows: Section II reviews related work. Section III introduces the system model and the problem formation. Theoretical analysis is presented in Section IV. In Section V, we conduct simulations to validate the accuracy of our analysis and discuss its insight. Section VI concludes this paper.

II. Related Work

For secure message dissemination in vehicular networks, two major factors need to be considered: the trustworthiness of each vehicle, and the integrity of the transmitted message. Accordingly, three misbehavior detection schemes are commonly adopted for secure message dissemination: entity-centric misbehavior detection scheme, data-centric misbehavior detection scheme, and a combined use of both. In the following, we will review the works on these three schemes separately.

Entity-Centric misbehavior detection scheme focuses on assessing on the trustworthiness level of each vehicle to filter out the malicious vehicle. The assessment process is commonly conducted at each vehicle by monitoring their instantaneous neighbors' behaviors. In [7], Gazdar et al. proposed a dynamic and distributed trust model to formalize a trust relationship between vehicles and filter out malicious and selfish vehicles. Their trust model is based on the use of a Markov chain to evaluate the evolution of the trust value. In [8], instead of allowing all vehicles to assess trustworthiness, Khan et al. proposed a novel malicious node detection algorithm for VANETs, which optimizes the selection of assessors to improve the overall network performance. In [9], Haddadou et al. proposed a distributed trust model for VANETs, which was motivated by the job market signaling model. Their trust model is able to gradually detect all malicious nodes as well as boosting the cooperation of selfish nodes. In [10], to overcome the challenge of intermittent and ad hoc monitoring and assessment process caused by the high mobility and rapid topology change in vehicular networks, Sedjelmaci et al. proposed a lightweight intrusion detection framework with the help of a clustering algorithm, where nodes are grouped into highly stable clusters so that the monitoring and assessment process can be better conducted in a relatively stable environment.

Data-centric misbehavior detection scheme focuses on the consistency check of the disseminated data to filter out the false data. In [6], Dietzel et al. argued that redundant data forwarding paths are the most promising technique to enable data consistency checks in a multi-hop information dissemination environment, and proposed three graph-theoretic metrics to measure the redundancy of dissemination protocols. In [11], Raya et al. proposed a framework for vehicular networks to establish data-centric trust, and evaluated the effectiveness of four data fusion rules: majority voting, weighted voting, Bayesian inference and belief propagation based technique. In [5], Huang et al. firstly demonstrated that information cascading and oversampling adversely affect the trust management scheme in VANETs, and then proposed a novel voting scheme that taking the distance between the transmitter and receiver into account when assigning weight to the trust level of the received data. In [12], Zaidi et al. proposed and evaluated a rogue node detection system for VANETs using statistical techniques to determine whether the data received are false. In [13], Radak applied a so-called cautious operator to deal with data received from different sources to detect dangerous events on the road. Their proposed cautious operator is an extension of the Demper-Shafer theory that is superior in handling data come from dependent sources.

A combined use of entity-centric and data-centric misbehavior detection scheme makes use of both the trust level of vehicles and the consistency of received data to detect misbehaving vehicles and filter out incorrect messages. Works adopting the combined scheme are limited. In [14], Dhurandher proposed a security algorithm using both node reputation and data plausibility checks to protect the network against attacks. The reputation value is obtained by both direct monitoring and indirect recommendation from neighbors; and the data plausibility check is conducted by comparing the received data with the sensed data by the vehicle's own sensors. In [15], Li et al. proposed an attack-resistant trust management scheme to evaluate the trustworthiness of both data and vehicles in VANETs, and to detect and cope with malicious attacks. They adopted the Dempster-Shafer theory to combine the data received from different sources, and then used this combined result to update the trust value of vehicles.

In summary, all the above works on security issues in vehicular networks focused on trust model establishment, trust model management, or methods to assess data from different sources to check their consistency, with the goal of detecting misbehaving nodes in the network. Our work is different from theirs in that we focus on theoretically analyzing the probability of correct message delivery, and evaluate the probability is affected by the network topology and the distribution of malicious vehicles in the network.

III. System Model and Problem Formation

A. Network Model

We consider a highway vehicular ad hoc network on a highway with bi-directional traffic flows. Vehicles in both directions are distributed randomly following Poisson point processes [16], [17] with spatial densities ρ_1 and ρ_2 respectively. As a ready consequence of the superposition property of Poisson processes [18], all vehicles on the highway are also Poissonly distributed with density $\rho = \rho_1 + \rho_2$. In actual road networks, there may be multiple lanes in each direction. Considering the width of a lane is typically small compared with the transmission range of vehicles, we ignore the road width and model multiple lanes in the same direction as one lane [16], [19].

B. Wireless Communication Model

We consider a general wireless connection model [20], where the probability that a receiver separated by a Euclidean distance x from a transmitter receives the message successfully with a probability $g(x)$, independent of transmissions by other transmitter-receiver pairs. There are two constraints on $g(x)$: 1) it is a monotonic non-increasing function of x and 2) $\lim_{x\to\infty} g(x) = 0$. This general wireless connection

model includes a number of widely-used wireless connection models as special cases. For instance, when $g(x)$ assumes the following form

$$g(x) = \begin{cases} 1, & 0 < x \leq r \\ 0, & x > r \end{cases},$$

it becomes the widely known unit disk model where a pair of wireless nodes are directly connected when their Euclidean distance is smaller than or equal to a threshold r, known as the transmission range. Alternatively, when $g(x)$ takes the following form,

$$g(x) = \frac{1}{2}\left(1 - \text{erf}\left(\frac{10\alpha \log_{10}\left(\frac{x}{r}\right)}{\sqrt{2\sigma^2}}\right)\right),$$

it becomes another widely known log-normal connection model , where α is the path loss exponent and σ is the standard deviation.

We consider a network with a sufficiently large vehicular density such that the generated vehicular network is a connected network. Besides, the broadcast transmission is adopted so that each message can be received by multiple vehicles to increase the number of redundant data forwarding paths. Furthermore, we assume that time is divided into time slots with equal length τ, and τ is sufficiently small such that we can regard vehicles are almost static during each time slot. After the message dissemination process begins, at each time slot, a vehicle among the set of vehicles that 1) have received at least one message and 2) are yet to transmit the message is randomly chosen to broadcast its received message. Such broadcast protocol can be readily implemented in a distributed manner by having each vehicle waits a random amount time identically and independently following an exponential distribution. Each vehicle only transmits its received message once. Note that the radio propagation speed is much faster than the moving speed of vehicles [22]. Therefore, we ignore the information propagation delay in this paper and assume that during the message dissemination process, the topology of the vehicular network remains unchanged.

C. Malicious Vehicle distribution and Data Fusion Rule

We assume that vehicles along the highway can be classified into two categories: *normal vehicles,* which behave normally and will forward the received message without any alteration, and *malicious vehicles,* which may tamper the received message and alter its content. We further assume that the probability of each vehicle being a malicious vehicle is p_m, independent of the event that another distinct vehicle is a malicious vehicle. We further assume that the malicious vehicles act in a distributed manner and there is no central coordination among malicious vehicles. As a consequence of the assumption, each malicious vehicle simply modify the received message without evaluation of the true content of the message.

Following the broadcast dissemination scheme considered in the paper, each vehicle is likely to receive multiple copies of message from different vehicles before it broadcasts the message. Due to the existence of malicious vehicles, the received messages may not be the same. For example, one vehicle may detect a traffic incident and generate a message alerting other vehicles but this message may be modified by a malicious vehicle. In the situation of conflicting messages being received, a *majority voting* rule is employed by each vehicle to fuse their received messages. That is, the vehicle will broadcast the message in agreement with the most number of vehicles and discard the message conflicting with majority opinion. When a tie occurs, the vehicle will randomly choose one of the two messages with equal probability to broadcast. The simplicity of the majority voting rule allows us to focus on the topological impact of vehicular networks on the correct message delivery. It is part of our future work plan to investigate the optimum fusion rule for highly dynamic vehicular networks.

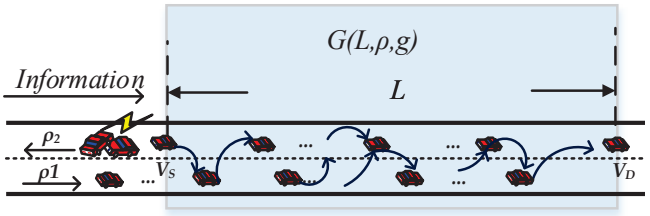

Fig. 1. An illustration of the sub-network we focused on in this work, which start from the location of vehicle V_S and ends at the location of destination vehicle V_D.

D. Problem Formation

Given the aforementioned background, we are now ready to give a formal definition of the problem considered in this paper.

Consider the a vehicle, termed the source vehicle V_S, detects an accident in front of it and wants to deliver a warning message to vehicles traveling in the same direction as V_S and behind V_S. Designate the location of V_S at the time instant when it broadcasts the message as the origin, and the direction of information propagation (in the opposite direction of the travel direction of V_S) as the positive direction. We want to investigate the probability that a vehicle, termed the destination vehicle V_D, located at distance L away from V_S can receive the message of V_S correctly. We denote by $G(L, \rho, g)$ the sub-network we focused on, which is with road segment $(0, L)$, vehicular density ρ and a wireless connection model g. See Fig. 1 for an illustration.

Two kinds of messages are considered in this paper, $+1$ represents the true message and -1 represents the false message. Here we assume that the source vehicle V_S is a normal vehicle, namely, the message broadcast by the source vehicle V_S is true. For malicious vehicles, as there are no central coordination among them, there is no way for a malicious vehicle to know the true content of the message. Therefore, it is assumed that a malicious vehicle simply modify the content of whatever message it receives, i.e., changing $+1$ to -1 and -1 to $+1$.

Finally, the destination vehicle V_D conducts its majority voting process after it has received all messages, or equivalent after no further message is received during a long time period. Denote by M_D the message after V_D has completed its data fusion its concluded message. In this paper, we are interested in investigating the probability that the destination vehicle V_D receives the correct message, denoted by P_{succ}, which can be expressed as follows:

$$P_{succ} = \Pr(M_D = 1) \tag{1}$$

IV. Theoretical Analysis

In this section, we will present our analysis on the probability that the destination vehicle receives the message correctly.

From the definition of the probability of correct message reception, which is given by (1), P_{succ} can be expressed as follows as an easy consequence of the total probability theorem:

$$\begin{aligned} P_{succ} &= \Pr(M_D = 1) \\ &= \sum_{n=1}^{\infty} \Pr(M_D = 1 | N = n) \Pr(N = n) \end{aligned} \tag{2}$$

where N denotes the random number of vehicles located in the sub-network $G(L, \rho, g)$. Due to the Poisson distribution of vehicles, we have

$$\Pr(N = n) = \frac{(\rho L)^n e^{-\rho L}}{n!}. \tag{3}$$

Recall that in our system, the source vehicle V_S located at the origin broadcasts its message first. After that, at each time slot, a vehicle among the set of vehicles having received at least one message *and* having not broadcast its message is randomly chosen to broadcast. Denote by V_i the ith vehicle that broadcast message and denote its location by Y_i, where $Y_i \in (0, L), i = 1, 2, ...n$ is a random variable representing the location of the ith vehicle broadcasting its message. We designate the source vehicle V_S as the 0th broadcast vehicle and its location is $y_0 = 0$. It follows that the destination vechile V_D then becomes the $n+1$th broadcast vehicle. Using the total probability theorem, the conditional probability that the destination vehicle V_D receives the correct message (after its fusion) given there are $N = n$ vehicles located in the sub-network $G(L, \rho, g)$, can be calculated by

$$\begin{aligned} &\Pr(M_D = 1 | N = n) \\ &= \int_0^L \cdots \int_0^L \int_0^L \Pr(M_D = 1 | Y_1 = y_1, Y_2 = y_2, ...Y_n = y_n) \\ &\quad \times f_{Y_1, Y_2, ...Y_n}(y_1, y_2, ...y_n) dy_1 dy_2 ... dy_n \end{aligned} \tag{4}$$

where $f_{Y_1, Y_2, ...Y_n}(y_1, y_2, ...y_n)$ is the joint distribution (probability density function) of the locations of the 1st, 2nd, ..., and nth broadcast vehicles.

Combining (2) - (4), it can be shown that to obtain the correct message reception probability P_{succ}, it remains to calculate the conditional probability that the destination vehicle V_D receives the message correctly given that the ith broadcast vehicle is located at $y_i, i = 1, 2, ...n$, i.e., $Pr(M_D = 1 | Y_1 = y_1, Y_2 = y_2, ...Y_n = y_n)$, and the joint distribution of the locations of the 1st, 2nd, ..., and nth broadcast vehicles, i.e., $f_{Y_1, Y_2, ...Y_n}(y_1, y_2, ...y_n)$. In the following, we will calculate these two terms separately.

A. Calculation of $Pr(M_D = 1 | Y_1 = y_1, Y_2 = y_2, ...Y_n = y_n)$

Denote by $h(y_i), i = 0, 1, ...n$ the indicator function that represents whether the destination vehicle V_D receives the message sent by ith broadcast vehicle V_i located at $Y_i = y_i$. Following the general wireless connection model considered in the paper, it can be readily shown that

$$h(y_i) = \begin{cases} 1, & g(L - y_i) \\ 0, & 1 - g(L - y_i) \end{cases}, i = 0, 1, ...n. \tag{5}$$

Denote by M_i the message transmitted by ith broadcast vehicle V_i, $i = 0, 1, ...n$. It follows that $Pr(M_0 = 1) = 1$ and each $M_i, i = 1, ...n$ is a binary random variable taking value from $\{+1, -1\}$. Under the majority voting rule, the conditional probability that the destination vehicle V_D receives the message correctly given that the ith broadcast vehicle is located at y_i, $i = 0, 1, ...n$, can be calculated by:

$$\begin{aligned} &\Pr(M_D = 1 | Y_1 = y_1, Y_2 = y_2, ...Y_n = y_n) \\ &= \Pr\left(\sum_{i=0}^{n} M_i h(y_i) > 0\right) + \frac{1}{2} \Pr\left(\sum_{i=0}^{n} M_i h(y_i) = 0\right) \\ &= \sum_{j=1}^{2^n} \prod_{i=0}^{n} \left[g(L - y_i) h^j(y_i) + \right. \\ &\quad \left. (1 - g(L - y_i)) \left(1 - h^j(y_i)\right) \right] \\ &\quad \times \Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) > 0\right) + \frac{1}{2} \sum_{j=1}^{2^n} \prod_{i=0}^{n} \\ &\quad \left[g(L - y_i) h^j(y_i) + (1 - g(L - y_i)) \left(1 - h^j(y_i)\right) \right] \\ &\quad \times \Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) = 0\right) \end{aligned} \tag{6}$$

where the first step follows from the rule of majority voting, particularly noting that when a tie occurs, the destination vehicle will make a decision randomly with equal probability. The term $\sum_{i=0}^{n} M_i h(y_i) > 0$ in the first step implies that the number of vehicles reporting $+1$ is larger than the number of vehicles reporting -1, therefore, resulting $M_D = 1$. The term $\sum_{i=0}^{n} M_i h(y_i) = 0$ implies that the number of vehicles reporting $+1$ is equal to the number of vehicles reporting -1, therefore, the event $M_D = 1$ occurs with probability $\frac{1}{2}$. The second step is obtained by using the total probability theorem. Note from (5) that each $h(y_i), i = 0, 1, ...n$ is a binary random variable. Therefore, the vector $[h(y_0)\ h(y_1)\ ...h(y_n)]$ can have 2^n possible values. Each $[h^j(y_0), h^j(y_1), ...h^j(y_n)]$, $j = 1, 2, ...2^n$ represents one possibility of $[h(y_0)\ h(y_1)\ ...h(y_n)]$, and the term $\prod_{i=1}^{n} \left[g(|L - y_i|) h^j(y_i) + (1 - g(|L - y_i|)) \left(1 - h^j(y_i)\right) \right]$ is the probability of the occurrence of a particular $[h^j(y_0), h^j(y_1), ...h^j(y_n)]$.

From (6), to calculate $\Pr(M_D = 1 | Y_1 = y_1, Y_2 = y_2, ...Y_n = y_n)$, it remains to calculate the two terms $\Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) > 0\right)$ and $\Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) = 0\right)$

given each $[h^j(y_0), h^j(y_1), ...h^j(y_n)]$, $j = 1, 2, ...2^n$. Using the joint distribution of M_0, M_1, ... M_n, $\Pr(M_0 = m_0, M_1 = m_1, ...M_n = m_n)$, the above two terms can be obtained as follows:

$$\begin{aligned}&\Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) > 0\right)\\ =&\sum_{\sum_{i=0}^{n} m_i h^j(y_i) > 0} \Pr(M_0 = m_0, M_1 = m_1, ...M_n = m_n),\end{aligned} \tag{7}$$

and

$$\begin{aligned}&Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) = 0\right)\\ =&\sum_{\sum_{i=0}^{n} m_i h^j(y_i) = 0} Pr(M_0 = m_0, M_1 = m_1, ...M_n = m_n).\end{aligned} \tag{8}$$

According to the chain rule of probability, it can be readily obtained that the joint distribution of M_0, M_1, ... M_n, $Pr(M_0 = m_0, M_1 = m_1, ...M_n = m_n)$ is given by

$$\begin{aligned}&Pr(M_1 = m_1, ...M_n = m_n)\\ =&Pr(M_n = m_n | M_{n-1} = m_{n-1}, ...M_1 = m_1, M_0 = 1) \times\\ &Pr(M_{n-1} = m_{n-1} | M_{n-2} = m_{n-2}, ...M_1 = m_1, M_0 = m_0)\\ &\times ... \times Pr(M_1 = m_1 | M_0 = 1) \times Pr(M_0 = m_0).\end{aligned} \tag{9}$$

Note that the message fusion result of vehicle V_i is dependent on the message M_0, M_1, ... M_{i-1} broadcast by vehicles $V_S, V_1, ...V_{i-1}$. Therefore, the conditional distribution of each M_i, $i = 1, 2, ...n$ given $M_1 = m_1, ...M_{i-1} = m_{i-1}$ can be obtained as follows:

$$\begin{aligned}&Pr(M_i = 1 | M_0 = m_0, M_1 = m_1, ...M_{i-1} = m_{i-1})\\ =&Pr\left(\sum_{j=0}^{i-1} m_j > 0\right)(1 - p_m) + \frac{1}{2} Pr\left(\sum_{j=0}^{i-1} m_j = 0\right)(1 - p_m)\end{aligned} \tag{10}$$

and

$$\begin{aligned}&Pr(M_i = -1 | M_0 = m_0, M_1 = m_1, ...M_{i-1} = m_{i-1})\\ =&1 - Pr(M_i = 1 | M_0 = m_0, M_1 = m_1, ...M_{i-1} = m_{i-1}),\end{aligned} \tag{11}$$

where the two terms in the first step of (10) are the probabilities that vehicle V_i reporting message $+1$ under two different fusion result separately. When the fusion result is $\sum_{j=0}^{i-1} m_j < 0$, the malicious vehicle would not change the message, therefore, with probability 0 to report $M_i = 1$.

Combining (9) - (11), we can obtain the joint distribution of M_0, M_1, ... M_n, $Pr(M_0 = m_0, M_1 = m_1, ...M_n = m_n)$. Plugging this joint distribution in (7) and (8), the two terms $Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) > 0\right)$ and $Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) = 0\right)$ in (6) can be obtained, which in turn leads to the result of $Pr(M_D = 1 | Y_1 = y_1, Y_2 = y_2, ...Y_n = y_n)$.

B. Calculation of $f_{Y_1, Y_2, ...Y_n}(y_1, y_2, ...y_n)$

Let $K_m, m = 0, 1, ...n$ be the set of vehicles in the subnetwork $G(L, \rho, g)$ which have received at least one message after the mth broadcast vehicle V_m has broadcast its messages. Given the location of ith broadcast V_i as $Y_i = y_i, i = 0, 1, ...m$, a vehicle located at x, $x \neq y_i, i = 0, 1, ...m$ belongs to K_m implies that it connects to at least one vehicle that are located at $y_0, y_1, ...y_m$, which has the probability $1 - \prod_{i=0}^{m} (1 - g(|x - y_i|))$. Note that the $(m+1)$th broadcast vehicle V_{m+1} is randomly chosen from the vehicle set $K_m \setminus \{V_S, V_1, ...V_m\}$, therefore, given each $Y_i = y_i, i = 1, 2, ...m$, the location of $(m+1)$th broadcast vehicle Y_{m+1} has the conditional probability density function as follows:

$$\begin{aligned}&f_{Y_{m+1} | Y_1, Y_2, ...Y_m}(x | y_1, y_2, ...y_m)\\ =&\frac{1 - \prod_{i=0}^{m} (1 - g(|x - y_i|))}{\int_0^L \left[1 - \prod_{i=0}^{m} (1 - g(|x - y_i|))\right] dx}, \quad m = 0, 1, ...n\end{aligned} \tag{12}$$

Eq. (12) is valid when $x \neq y_i, i = 1, 2, ...m$ as we assume each vehicles can be chosen to broadcast only once. Particularly, when $m = 0$, we have

$$f_{Y_1}(x) = \frac{g(x)}{\int_0^L g(x) dx}.$$

As an easy consequence of the chain rule of probability, the joint distribution of $Y_1, Y_2, ...Y_n$ can be obtained as follows:

$$\begin{aligned}&f_{Y_1, Y_2, ...Y_n}(y_1, y_2, ...y_n)\\ =&f_{Y_n | Y_{n-1}, ...Y_2, Y_{n-1}}(y_n | y_{n-1}, ..., y_2, y_1)\\ &\times f_{Y_{n-1} | Y_{n-2}, ..Y_2, Y_1}(y_{n-1} | y_{n-2}, ..., y_2 y_1)\\ &\times f_{Y_{n-2} | Y_{n-3}, ...Y_2, Y_1,}(y_{n-2} | y_{n-3}, ...y_2, y_1) \times ...\\ &\times f_{Y_2 | Y_1}(y_2 | y_1) \times f_{Y_1}(y_1),\end{aligned} \tag{13}$$

where each conditional distribution in (13) is given by (12).

V. Simulation and Discussion

In this section, numerical and simulation results are shown to discuss the relationship between the probability of correct message reception and its major performance-impacting parameters. Specifically, we adopt unit disk model and log-normal connection model as two special cases of the general wireless connection model respectively in the simulation. For the unit disk model, we set the transmission range $r = 250$m (typical radio ranges using DSRC [23]), and for the log-normal connection model, we set the the path loss exponent $\alpha = 2$ and the standard deviation $\sigma = 4$ [20]. Each simulation is repeated 5000 times and the average value is shown in the plot.

Fig. 2 and Fig. 3 show the relationship between the probability of correct message reception P_{succ} and the probability of each vehicle being malicious p_m assuming the unit disk model, with different distance L between the source vehicle and the destination vehicle, and with different vehicular density ρ respectively. Specifically, we can see that $P_{succ} = 1$ when $p_m = 0$, which is the case that all vehicles are normal vehicles; when p_m is small, P_{succ} decreases sharply with an increase of p_m and decreases to its minimum value 0 when p_m is larger than a certain threshold p_{th}. Beyond that threshold, a

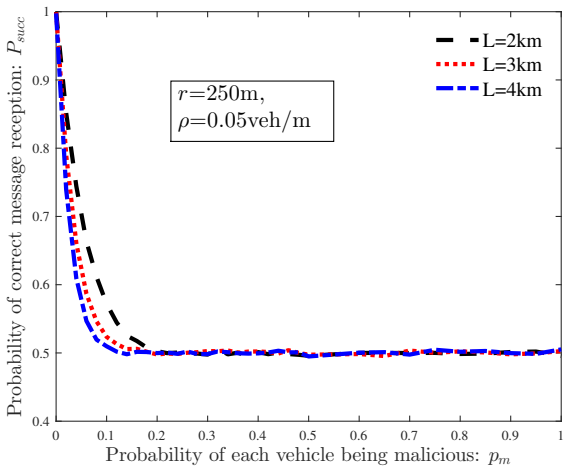

Fig. 2. The relationship between the probability of correct message delivery P_{succ} and the probability of each vehicle being malicious p_m assuming unit disk model, with different distance L between the source vehicle and the destination vehicle.

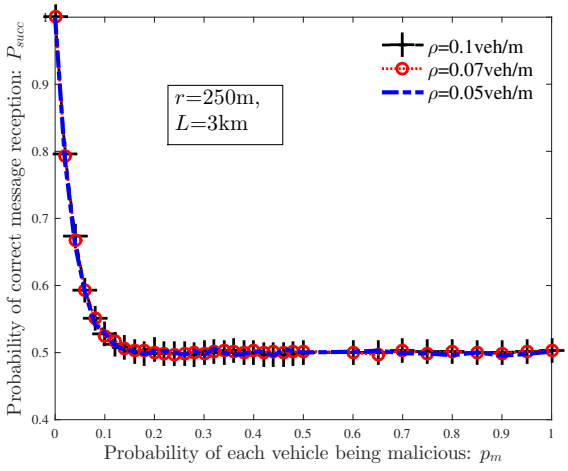

Fig. 3. The relationship between the probability of correct message delivery P_{succ} and the probability of each node being malicious p_m assuming unit disk model, with different vehicular density ρ.

further increase in p_m has little impact on P_{succ}. This can be explained by the fact that when $p_m < p_{th}$, the number of malicious vehicles in the network is small. Therefore, an increase in p_m will largely increase the number of malicious vehicles, which consequently, leads to a sharp decrease in the probability of correct message reception. When p_m is larger than its threshold, malicious vehicles play dominant roles in the majority voting scheme. In this case, for any vehicle in the network, the outcome of its message fusion result will be incorrect. This consequently, leads to that all the messages transmitted in the network are not the correct one, therefore P_{succ} converges to 0.

Fig. 2 shows that given a fixed vehicular density, when $p_m < p_{th}$, a larger distance L between the source vehicle and the destination vehicle will lead to a smaller P_{succ}. This is due to the fact that keeping other parameters constant, a larger L implies a larger number of malicious vehicles participating in tampering the message transmitted from the source vehicle to the destination vehicle. As a consequence, it leads to a smaller P_{succ}.

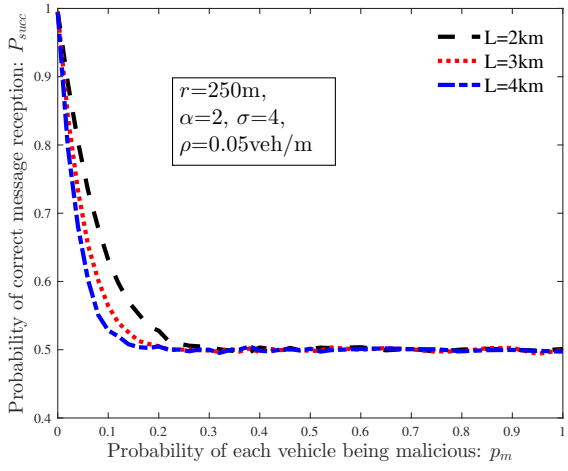

Fig. 4. The relationship between the probability of correct message reception P_{succ} and p_m assuming the log-normal connection model, with different distance L between the source vehicle and the destination vehicle.

Fig. 3 shows that in our system, a larger vehicular density ρ has little impact on P_{succ}. Intuitively, a larger ρ will leads to a larger P_{succ} due to the fact that a larger ρimplies a larger number of messages received by each vehicle, which is beneficial for vehicles to conduct data consistency checks. Therefore, when the traffic density increases, the message fusion result of each vehicle will be more accurate. Consequently, keeping other parameters constant, the probability of correct message reception P_{succ} will increase. However, when a vehicle is randomly chosen among the set of vehicles that have received at least one message to broadcast, it may not have received a sufficient number messages from other vehicles. This follows that even with an increase in traffic density ρ, the message fusion result of each broadcast vehicle does not improve. Therefore, a larger vehicular density ρ has little impact on the P_{succ}.

Fig. 4 and Fig. 5 show the relationship between the probability of correct message reception P_{succ} and the probability of each vehicle being malicious p_m assuming the log-normal connection model, with different distance L between the source vehicle and the destination vehicle, and with different vehicular density ρ respectively. We can see that with the increase of p_m from 0 to 1, the trend of P_{succ} is the same as that assuming unit disk model. Therefore, we omit the duplicate discussion here.

Fig. 6 gives a comparison of the correct message reception probability P_{succ} achieved assuming the unit disk model (labeled as UDM) and that achieved assuming the log-normal connection model (labeled as LSM). It is shown that keeping other parameters constant, when $p_m < p_{th}$, the system assuming the log-normal connection model has a slightly higher correct message reception probability P_{succ} than that assuming the unit disk model. The reason behind this phenomenon is that the log-normal connection model introduces a Gaussian variation of the transmission range around the mean value, which implies a higher chance for the vehicles to be connected

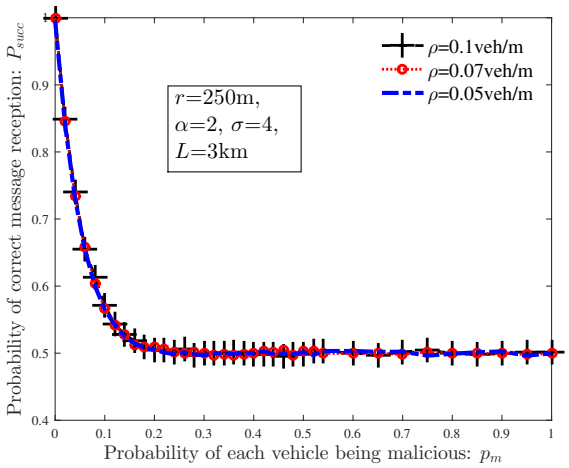

Fig. 5. The relationship between the probability of correct message reception P_{succ} and p_m assuming the log-normal connection model, with different vehicular density ρ.

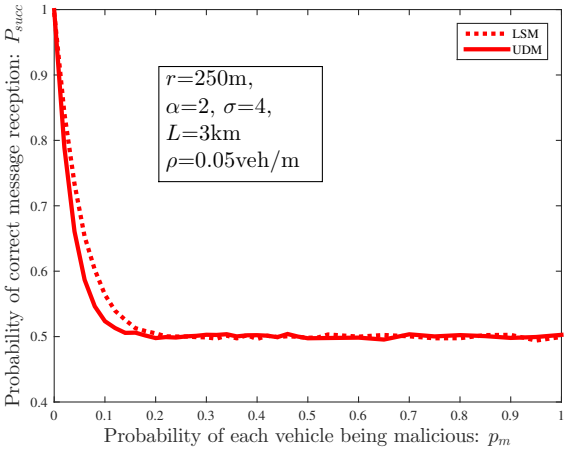

Fig. 6. A comparison between the probability of correct message delivery P_{succ} achieved assuming the unit disk model and that assuming the log-normal connection model.

to other vehicles separated further away. Therefore, assuming log-normal connection model will lead to a better message fusion result, resulting in a higher correct message delivery probability P_{succ}.

VI. Conclusions

This paper studied a vehicular ad hoc network where a certain fraction of the vehicles are malicious vehicles and these malicious vehicles are distributed randomly in the network. Furthermore, there is no central coordination among these malicious vehicles and consequently a malicious vehicle simply modify its received message irrespective of its true value. An analytical framework is developed to model the process of secure message dissemination in the network, and the probability that a vehicle, located at a fixed distance away from the source vehicle, can receive the message correctly is obtained. Simulations were conducted to establish the accuracy of the analytical results. Both simulation and numerical results demonstrate that the probability of correct message delivery reduces to its minimum after the proportion of malicious vehicles in the network increases beyond a threshold. Besides, a larger vehicular density and a smaller distance between the destination vehicle and the source vehicle will lead to a larger probability of correct message reception. Our results may provide insight on the design of security mechanisms, particularly secure routing algorithms and topology control algorithms, to enhance secure message dissemination in highly dynamic vehicular networks.

References

[1] K. Zheng, et al., "Heterogeneous Vehicular Networking: A Survey on Architecture, Challenges, and Solutions," *IEEE Communication Survey & Tutorials*, vol. 17, no. 4, pp. 2377-2396, Fourth Quarter, 2015.

[2] J. B. Kenney, "Dedicated Short-Range Communications (DSRC) Standards in the United States," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1162-1182, 2011.

[3] S. Ilarri, T. Delot, R. Trillo-Lado, "A Data Management Perspective on Vehicular Networks," *IEEE Communication Survey & Tutorials*, vol. 17, no. 4, pp. 2420-2460, Fourth Quarter, 2015.

[4] IEEE Trial-Use Standard for Wireless Access in Vehicular Environments- Security Services for Applications and Management Messages, IEEE Std. 1609.2-2006.

[5] Z. Huang, S. Ruj, M. A. Cavenaghi, M. Stojmenovic, and A. Nayak, "A social network approach to trust management in VANETs," *Peer-to-Peer Networking and Applications*, vol. 7, no. 3, pp. 229-242, 2014.

[6] S. Dietzel, J. Petit, G. Heijenk, and F. Kargl, "Graph-Based Metrics for Insider Attack Detection in VANET Multihop Data Dissemination Protocols," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 4, pp. 1505-1518, May. 2013.

[7] T. Gazdar, A. Rachedi, A. Benslimane, and A. Belghith, "A distributed advanced analytical trust model for VANETs," in *IEEE Global Communications Conference (GLOBECOM)* 2012.

[8] U. Khan, S. Agrawal, and S. Silakari, "Detection of Malicious Nodes (DMN) in Vehicular Ad-Hoc Networks," *Procedia Computer Science*, vol. 46, pp. 965-972, 2015.

[9] N. Haddadou, A. Rachedi, and Y. Ghamri-Doudane, "A Job Market Signaling Scheme for Incentive and Trust Management in Vehicular Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 8, pp. 3657-3674, Aug. 2015.

[10] H. Sedjelmaci, and S. M. Senouci, "An accurate and efficient collaborative intrusion detection framework to secure vehicular networks", *Computers and Electrical Engineering*, vol. 43, pp. 33-47, 2015.

[11] M. Raya, P. Papadimitratos, V. D. Gligor, and J. P. Hubaux, "On Data-Centric Trust Establishment in Ephemeral Ad Hoc Networks," in *IEEE INFOCOM* 2008.

[12] K. Zaidi, M. B. Milojevic, V. Rakocevic, A. Nallanathan, and M. Rajarajan, "Host-Based Intrusion Detection for VANETs: A Statistical Approach to Rogue Node Detection," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6703-6714, Aug. 2014.

[13] J. Radak, B. Ducourthial, V. Cherfaoui, and S. Bonnet, "Detecting Road Events Using Distributed Data Fusion: Experimental Evaluation for the Icy Roads Case," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 184-194, Jan. 2016.

[14] S. K. Dhurandher, M. S. Obaidat, A. Jaiswal, A. Tiwari and A. Tyagi, " Vehicular Security Through Reputation and Plausibility Checks," *IEEE system journal*, vol. 8, no. 2, Jun. 2014.

[15] W. Li, and H. Song, "ART: An Attack-Resistant Trust Management Scheme for Securing Vehicular Ad Hoc Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 960-969, Apr. 2016.

[16] N. Wisitpongphan, et al., "Routing in Sparse Vehicular Ad Hoc Wireless Networks," *IEEE Journal on Selected Areas in Communications,* vol. 25, no. 8, pp. 1538-1556, Oct. 2007.

[17] A. B. Reis, et al., "Deploying Roadside Units in Sparse Vehicular Networks: What Really Works and What Does Not," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 6, pp. 2794-2806, Jul. 2014.

[18] R. Nelson, Probability, Stochastic Processes, and Queueing Theory: The Mathematics of Computer Performance Modeling. *New York: Springer-Verlag*, 1995.

[19] K. Abboud and W. Zhuang, "Stochastic Analysis of a Single-Hop Communication Link in Vehicular Ad Hoc Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2297-2307, Oct. 2014.

[20] Z. Zhang, G. Mao, T. Han and B. D. O. Anderson, "Cooperative Information Forwarding in Vehicular Networks Subject to Channel Randomness," in *Proceedings of IEEE ICC*, 2014.

[21] Z. Zhang, G. Mao, and B. D. O. Anderson, "On the hop count statistics in wireless multi-hop networks subject to fading," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 7, pp. 1275–1287, 2012.

[22] Z. Zhang, G. Mao, and B. D. O. Anderson, "Stochastic Characterization of Information Propagation Process in Vehicular Ad hoc Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 122-135, Feb. 2014.

[23] Z. Haibo, L. Bo, T. H. Luan, H. Fen, G. Lin, L. Ying, Q. Yu, and X. Shen, "ChainCluster: Engineering a Cooperative Content Distribution Framework for Highway Vehicular Communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no.6, pp. 2644-2657, Dec. 2014.