Journal of Communications and Information Networks, Vol.2, No.2, Jun. 2017
DOI: 10.1007/s41650-017-0025-7

Research paper

Special Issue on Internet of Vehicle

On the security of warning message dissemination in vehicular Ad hoc networks

Jieqiong Chen*, Guoqiang Mao

School of Computing and Communications, Univiersity of Technology Sydney, Sydney NSW 2007, Australia

* Corresponding author, Email: jieqiong.chen@student.uts.edu.au

Abstract: Information security is an important issue in vehicular networks as the accuracy and integrity of information is a prerequisite to the satisfactory performance of virtually all vehicular network applications. We study the information security of a vehicular Ad hoc network whose message could be tampered by malicious vehicles. An analytical framework is developed to analyze the process of message dissemination in a vehicular network with malicious vehicles randomly distributed in the network. The probability that a destination vehicle at a fixed distance 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 provides insight on the design of security mechanisms to improve the security of message dissemination in vehicular networks.

Keywords: vehicular Ad hoc networks, Internet-of-Vehicles technology, message dissemination, malicious vehicles, security

Citation: J. Q. Chen, G. Q. Mao. On the security of warning message dissemination in vehicular Ad hoc networks [J]. Journal of communications and information networks, 2017, 2(2): 46-58.

1 Introduction

Interest is surging on vehicular networks and Internet-of-Vehicles technologies owing 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 infrastructures along roadsides (e.g., RSU (Road-Side Units)), equipping vehicles with on-board communication facilities (e.g., OBU (On-Board Units)), and with the assistance of DSRC (Dedicated Short-Range Communication)[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 requiring real-time information regarding 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].

Accompanying the convenience and advantage of wireless communications is the potential security threat that vehicular networks may present to transportation systems. Unlike 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 not been previously encountered. This may render traditional se-

Manuscript received Mar. 3, 2017; accepted Apr. 7, 2017

curity 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[4]. 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 including traffic congestion, traffic crashes, even loss of lives and must be thoroughly investigated before vehicular networks are deployed.

In this paper, we study information security in VANETs (Vehicular Ad hoc NETworks), 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 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, or congestion, sends a message informing other vehicles of the situation. The message is forwarded from the source vehicle in a broadcast and multi-hop manner to other vehicles. Instead of forwarding every received single piece of message, each vehicle would conduct a message fusion process to combine all its received messages before its forwarding, to conclude a comprehensive opinion on the accuracy of the received messages so that the accuracy of the message broadcast by them will be improved. In this manner, the accuracy and security of the forwarded message can be improved. We analyze the probability that a vehicle at a fixed distance, 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. Our results may provide insight into the design of security mechanisms, particularly secure routing algorithms and topology control algorithms, to improve information security in vehicular networks. 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 is analyzed.

2. Simulations are conducted to establish the accuracy of the analysis. Using the analysis, a relationship is revealed between the key parameters such as the probability of correct message reception 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.

The rest of this paper is organized as follows: Section 2 reviews related work. Section 3 introduces the system model and the problem formation. Theoretical analysis is presented in section 4. In section 5, we conduct simulations to validate the accuracy of our analysis and discuss its insight. Section 6 concludes this paper.

2 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 review the works on these three schemes separately.

Entity-centric misbehavior detection schemes focus on assessing the trustworthiness level of each vehicle to filter out malicious vehicles. The assessment process is commonly conducted at each vehicle by monitoring their instantaneous neighbors' behavior. In Ref. [5], 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 Ref. [6], rather than allowing all vehicles to assess trustworthiness, Khan et al. proposed a novel malicious node detection algorithm for VANETs that optimizes the selection of assessors to improve the overall network performance. In Ref. [7], Haddadou et al. proposed a distributed trust model for VANETs that 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 Ref. [8], to overcome the challenges of intermittent and Ad hoc monitoring and assessment processes 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 such that the monitoring and assessment processes can be more effectively conducted in a relatively stable environment.

Data-centric misbehavior detection schemes focus on the consistency check of the disseminated data to filter out false data. In Ref. [9], Dietzel et al. argued that redundant data forwarding paths are the most promising technique for effective data consistency verification in a multi-hop information dissemination environment. They proposed three graph-theoretic metrics to measure the redundancy of the dissemination protocols. In Ref. [10], 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 techniques. In Ref. [11], Huang et al. first demonstrated that information cascading and oversampling adversely affect the performance of trust management scheme in VANETs, and then proposed a novel voting scheme that considers the distance between the transmitter and receiver when assigning a weight to the trust level of the received data. In Ref. [12], Zaidi et al. proposed and evaluated a rogue node detection system for VANETs using statistical techniques to determine whether the received data are false. In Ref. [13], Radak et al. applied a cautious operator to deal with data received from different sources to detect dangerous events on the road. Their adopted cautious operator is an extension of the Demper-Shafer theory that is known to be superior in managing data originating from dependent sources.

A combined use of the entity-centric and data-centric misbehavior detection scheme uses both the trust level of the vehicles and the consistency of the received data to detect misbehaving vehicles and filter out incorrect messages. Works adopting the combined scheme are limited. In Ref. [14], Dhurandher et al. 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 consistency check is conducted by comparing the received data with the sensed data by the vehicle's own sensors. In Ref. [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 address 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 validate their consistency, with a goal of detecting misbehaving nodes in the network. The proposed work is different from theirs in that we focus on theoretically characterizing the probability of correct message reception, and evaluate the impact of network topology and distribution of malicious vehicles on the probability.

3 System model and problem formation

3.1 Network model

We consider a vehicular Ad hoc network on a highway with bi-directional traffic flows. Vehicles in both directions are distributed randomly following Poisson point processes with spatial densities ρ_1 and ρ_2 respectively. As a ready consequence of the superposition property of Poisson processes[16], all vehicles on the highway are also Poissonly distributed with density $\rho = \rho_1 + \rho_2$. The Poisson distribution for vehicles has been supported by some empirical study that it can accurately characterize real traffic distribution in specific scenarios[17], and it is also commonly adopted by works on vehicular networks, e.g., Refs. [17-21]. Moreover, in actual road networks, there may be multiple lanes in each direction. Considering that 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[17,22].

3.2 Wireless communication model

We consider a general wireless connection model[19], where 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 its special cases. For instance, when $g(x)$ assumes the following form:

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

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,

$$g(x) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{10\alpha \lg (x/r)}{\sqrt{2\sigma^2}}\right)\right), \tag{2}$$

it becomes another widely known log-normal connection model[23-25], where α is the path loss exponent, σ is the standard deviation and r is the equivalent transmission range when $\sigma = 0$.

We consider a network with a sufficiently large vehicular density such that the generated vehicular network is a connected network[23]. Broadcast transmission is adopted such that each message can be received by multiple vehicles to increase the number of redundant data forwarding paths and reduce the message dissemination time. Furthermore, we assume that time is divided into time slots of equal length τ, and τ is sufficiently small such that we can regard vehicles as virtually stationary 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 selected to broadcast its received message. Such a broadcast protocol can be readily implemented in a distributed manner by having each vehicle waits a random amount of time identically and independently distributed following an exponential distribution before transmitting its received message. Each vehicle transmits its received message only once. Note that the radio propagation speed is considerably faster than the moving speed of the vehicles[20], which implies that during the message dissemination process considered in this paper, the distance moved by each vehicle is rather small. Therefore, we ignore the information propagation delay and assume that the topology of the vehicular network remains unchanged during the message dissemination process, i.e., the network topology is not affected by the mobility of vehicles during the message dissemination process.

3.3 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 modify the received message and alter its content. Further, we 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 modifies 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 a message from different vehicles before it broadcasts the message. Owing 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, however, 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 normal vehicles will broadcast the message in agreement with the greatest number of vehicles and discard the message conflicting with majority opinion; the malicious vehicles will broadcast the message conflicting with the majority opinion. When a tie occurs, all the vehicles will randomly choose one of the two messages (true or false message) 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.

3.4 Problem formation

Given the aforementioned background, we now present a formal definition of the problem considered in this paper.

Consider that a vehicle, termed the source vehicle V_S, detects an accident in front of itself and wishes to deliver a warning message to vehicles traveling in the same direction as V_S and behind V_S in that direction. 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 wish to investigate the probability that a vehicle, termed the destination vehicle V_D, located at distance L from V_S can receive the message from V_S correctly. We denote by $G(L, \rho, g)$ the sub-network we focus on, which is within the road segment $(0, L)$, with vehicular density ρ and a wireless connection model g. See Fig. 1 for an illustration.

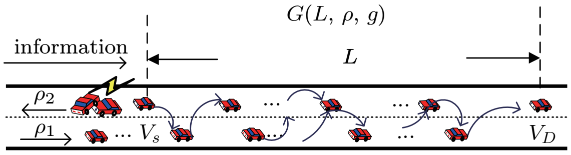

Figure 1 Illustration of the sub-network we focused on in this work, which starts from the location of vehicle V_S and ends at the location of destination vehicle V_D

Two kinds of messages are considered in this paper, $+1$ represents the true message (e.g., road is congested) and -1 represents the false message (e.g., road is not congested). In practice, the source vehicle may also be malicious, e.g., it may fabricate an incident message to deceive other vehicles. In this paper, we aim to investigate the probability that the message received by the destination vehicle is exactly the message broadcast by the source vehicle. It follows that our analysis will not be affected by the message type the source vehicle broadcasts. Therefore, without loss of generality, here we assume that the source vehicle V_S is a normal vehicle, i.e., the message broadcast by the source vehicle V_S is true. For malicious vehicles, as there is 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 (against the outcome of the majority voting rule), 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 equivalently after no further message is received after a period of time. Denote by M_D the concluded message after V_D has completed its data fusion. 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:

$$P_{\text{succ}} = \Pr(M_D = 1) \tag{3}$$

4 Theoretical analysis

In this section, we 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 in Eq. (3), P_{succ} can be expressed as follows as an easy consequence of the total probability theorem:

$$\begin{aligned} P_{\text{succ}} &= \Pr(M_D = 1) \\ &= \sum_{n=1}^{\infty} \Pr(M_D = 1|N = n) \Pr(N = n), \end{aligned} \tag{4}$$

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

$$\Pr(N = n) = \frac{(\rho L)^n e^{-\rho L}}{n!}. \tag{5}$$

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 selected to broadcast. Denote by V_i the ith vehicle that broadcasts the message and denote its location by Y_i, where $Y_i \in (0, L)\,, i = 1, 2, \cdots n$ 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 as $y_0 = 0$. It follows that the destination vehicle 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

$$\begin{aligned} &\Pr(M_D = 1|N = n) \\ = &\int_0^L \cdots \int_0^L \int_0^L \Pr(M_D = 1|Y_1 = y_1, \\ &Y_2 = y_2, \cdots Y_n = y_n) \times f_{Y_1,Y_2,\cdots Y_n}(y_1, \\ &y_2, \cdots y_n) \mathrm{d}y_1 \mathrm{d}y_2 \cdots \mathrm{d}y_n, \end{aligned} \tag{6}$$

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

Combining Eqs. (4)-(6), it can be demonstrated 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 $y_i, i = 1, 2, \cdots n$, i.e., $\Pr(M_D = 1|Y_1 = y_1, Y_2 = y_2, \cdots Y_n = y_n)$, and the joint distribution of the locations of the 1st, 2nd, $\cdots$, and nth broadcast vehicles, i.e., $f_{Y_1,Y_2,\cdots Y_n}(y_1, y_2, \cdots y_n)$. In the following, we will calculate these two terms separately.

4.1 Calculation of $Pr(M_D = 1|Y_1 = y_1, Y_2 = y_2, \cdots Y_n = y_n)$

Denote by $h(y_i), i = 0, 1, \cdots n$ the indicator function that represents whether the destination vehicle V_D receives the message sent by the 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

$$h(y_i) = \begin{cases} 1, & g\,(L - y_i) \\ 0, & 1 - g\,(L - y_i) \end{cases}, i = 0, 1, \cdots n. \tag{7}$$

Denote by M_i the message broadcast by the ith broadcast vehicle V_i located at y_i, $i = 0, 1, \cdots n$. It follows that $M_0 = 1$ as we regard the source vehicle V_S is a normal vehicle that broadcasts the true message, and each $M_i, i = 1, \cdots n$ is a binary random variable taking its value from $\{+1,-1\}$. Assuming 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, $i=1,\cdots n$, can be calculated by:

$$\begin{aligned}
&\Pr(M_D=1|Y_1=y_1,Y_2=y_2,\cdots Y_n=y_n)\\
=&\Pr\left(\sum_{i=0}^{n}M_ih(y_i)>0\right)+\frac{1}{2}\Pr\left(\sum_{i=0}^{n}M_ih(y_i)=0\right)\\
=&\sum_{j=1}^{2^{n+1}}\left[\Pr\left(\sum_{i=0}^{n}M_ih^j(y_i)>0\right)\Pr\left(\boldsymbol{h}=\boldsymbol{h}^j\right)\right]\\
&+\frac{1}{2}\sum_{j=1}^{2^{n+1}}\left[\Pr\left(\sum_{i=0}^{n}M_ih^j(y_i)=0\right)\Pr\left(\boldsymbol{h}=\boldsymbol{h}^j\right)\right]\\
=&\sum_{j=1}^{2^{n+1}}\left\{\Pr\left(\sum_{i=0}^{n}M_ih^j(y_i)>0\right)\right.\\
&\times\left[\prod_{i=0}^{n}\left[g\left(L-y_i\right)h^j\left(y_i\right)\right.\right.\\
&\left.\left.\left.+\left(1-g\left(L-y_i\right)\right)\left(1-h^j\left(y_i\right)\right)\right]\right]\right\}\\
&+\frac{1}{2}\sum_{j=1}^{2^{n+1}}\left\{\Pr\left(\sum_{i=0}^{n}M_ih^j(y_i)=0\right)\times\right.\\
&\left[\prod_{i=0}^{n}\left[g\left(L-y_i\right)h^j\left(y_i\right)\right.\right.\\
&\left.\left.\left.+\left(1-g\left(L-y_i\right)\right)\left(1-h^j\left(y_i\right)\right)\right]\right]\right\}
\end{aligned}\tag{8}$$

where the vector $\boldsymbol{h}$ is defined by

$$\begin{aligned}
\boldsymbol{h}\ =&\ \{h(y_0),h(y_1),\cdots h(y_n)\\
&:h(y_i)\in\{1,0\},1\leqslant i\leqslant n\},
\end{aligned}\tag{9}$$

and 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 second step is obtained using the total probability theorem on $\boldsymbol{h}$. Note from Eq. (7) that each $h(y_i),i=0,1,\cdots n$ is a binary random variable. Therefore, the vector $\boldsymbol{h}$ can have 2^{n+1} possible values and we let $\boldsymbol{h}=\boldsymbol{h}^j$, $j=1,2,\cdots 2^{n+1}$ represents each possible value. The third step follows by plugging

$$\begin{aligned}
&\Pr\left(\boldsymbol{h}=\boldsymbol{h}^j\right)\\
=&\prod_{i=0}^{n}\left[g\left(L-y_i\right)h^j\left(y_i\right)\right.\\
&\left.+\left(1-g\left(L-y_i\right)\right)\left(1-h^j\left(y_i\right)\right)\right],
\end{aligned}$$

which readily results from the definition of each $h(y_i),i=0,1,\cdots n$ given as Eq. (7).

From Eq. (8), to calculate $\Pr(M_D=1|Y_1=y_1,Y_2=y_2,\cdots Y_n=y_n)$, it remains to calculate the two terms $\Pr\left(\sum_{i=0}^{n}M_ih^j(y_i)>0\right)$ and $\Pr\left(\sum_{i=0}^{n}M_ih^j(y_i)=0\right)$ given each fixed $\boldsymbol{h}^j=\left\{h^j(y_0),h^j(y_1),\cdots h^j(y_n)\right\}$, $j=1,2,\cdots 2^{n+1}$. Using the joint distribution of M_1, M_2, $\cdots$ M_n, $\Pr\left(M_1=m_1,M_2=m_2,\cdots M_n=m_n\right)$, the above two terms can be obtained as follows:

$$\begin{aligned}
&\Pr\left(\sum_{i=0}^{n}M_ih^j(y_i)>0\right)\\
=&\sum_{\sum_{i=0}^{n}m_ih^j(y_i)>0}\Pr\left(M_0=m_0=1,\right.\\
&\left.M_1=m_1,\cdots M_n=m_n\right)\\
=&\sum_{h^j(0)+\sum_{i=1}^{n}m_ih^j(y_i)>0}\Pr\left(M_1{=}m_1,\cdots M_n{=}m_n\right),
\end{aligned}\tag{10}$$

and

$$\begin{aligned}
&\Pr\left(\sum_{i=0}^{n}M_ih^j(y_i)=0\right)\\
=&\sum_{\sum_{i=0}^{n}m_ih^j(y_i)=0}\Pr\left(M_0=m_0=1,\right.\\
&\left.M_1=m_1,\cdots M_n=m_n\right)\\
=&\sum_{h^j(0)+\sum_{i=1}^{n}m_ih^j(y_i)=0}\Pr\left(M_1{=}m_1,\cdots M_n{=}m_n\right).
\end{aligned}\tag{11}$$

According to the chain rule of probability, it can be readily obtained that the joint distribution of M_1,M_2 $\cdots$ M_n, is given by

$$\begin{aligned}
&\Pr\left(M_1=m_1,M_2=m_2,\cdots M_n=m_n\right)\\
=&\Pr\left(M_n=m_n|M_{n-1}=m_{n-1},\cdots M_2=m_2,\right.\\
&\left.M_1=m_1\right)\times\Pr\left(M_{n-1}=m_{n-1}|M_{n-2}=m_{n-2},\right.\\
&\left.\cdots M_2=m_2,M_1=m_1\right)\times\cdots\\
&\times\Pr\left(M_2=m_2|M_1=m_1\right)\Pr(M_1=m_1).
\end{aligned}\tag{12}$$

Note that the message fusion result of vehicle V_i is dependent on the messages M_0, M_1, $\cdots$ M_{i-1} broadcast by vehicles $V_S,V_1,\cdots V_{i-1}$. Therefore, the conditional distribution of each M_i, $i=1,2,\cdots n$ given

$M_1 = m_1, \cdots M_{i-1} = m_{i-1}$ can be obtained as follows:

$$\begin{aligned} &\Pr\left(M_i = 1 | M_0 = 1, M_1 = m_1, \cdots M_{i-1} = m_{i-1}\right) \\ &= \Pr\left(1 + \sum_{j=1}^{i-1} (m_j \cdot g(y_i - y_j)) > 0\right)(1 - p_m) \\ &+ \Pr\left(1 + \sum_{j=1}^{i-1} (m_j \cdot g(y_i - y_j)) < 0\right) p_m \\ &+ \frac{1}{2}\Pr\left(1 + \sum_{j=1}^{i-1} (m_j \cdot g(y_i - y_j)) = 0\right), \end{aligned} \tag{13}$$

and

$$\begin{aligned} &\Pr(M_i{=}-1 | M_0{=}1, M_1{=}m_1, \cdots M_{i-1}{=}m_{i-1}) \\ &= 1 - \Pr\left(M_i = 1 | M_0 = 1, M_1 = m_1,\right. \\ &\quad \left.\cdots M_{i-1} = m_{i-1}\right), \end{aligned} \tag{14}$$

where the three terms in Eq. (13) are the probabilities that vehicle V_i broadcasts message +1 under three different cases:

$$\begin{aligned} \text{case 1}: &\quad 1 + \sum_{j=1}^{i-1} (m_j \cdot g(y_i - y_j)) > 0, \\ \text{case 2}: &\quad 1 + \sum_{j=1}^{i-1} (m_j \cdot g(y_i - y_j)) < 0, \\ \text{case 3}: &\quad 1 + \sum_{j=1}^{i-1} (m_j \cdot g(y_i - y_j)) = 0 \end{aligned}$$

separately. Using the case 1 as an example to illustrate: when $1+\sum_{j=1}^{i-1} (m_j \cdot g(y_i - y_j)) > 0$, vehicle V_i would conclude from its majority voting process that the majority opinion of the message is +1. Considering each vehicle has probability p_m to modify the message (being a malicious vehicle), therefore, the probability for the vehicle to broadcast the correct concluded message (from the majority voting process) +1 would be $1 - p_m$, which leads to the term

$$\Pr\left(1 + \sum_{j=1}^{i-1} (m_j \cdot g(y_i - y_j)) > 0\right)(1 - p_m).$$

Specifically, from Eq. (13), when $i = 1$ we have

$$\Pr\left(M_1 = 1\right) = 1 - p_m, \tag{15}$$

and

$$\Pr\left(M_1 = -1\right) = p_m, \tag{16}$$

which can also be readily obtained as the 1st broadcast vehicle only receives the true message from the source vehicle.

Combining Eqs. (12)-(14), we can obtain the joint distribution of M_1, M_2, $\cdots$ M_n, $\Pr\left(M_1 = m_1, M_2 = m_2, \cdots M_n = m_n\right)$. Plugging this joint distribution in Eqs. (10) and (11), the two terms:

$$\Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) > 0\right), \Pr\left(\sum_{i=0}^{n} M_i h^j(y_i) = 0\right)$$

in Eq. (8) can be obtained, which in turn leads to the result of $\Pr(M_D = 1 | Y_1 = y_1, Y_2 = y_2, \cdots Y_n = y_n)$.

4.2 Calculation of $f_{Y_1,Y_2,\cdots Y_n}(y_1, y_2, \cdots y_n)$

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

$$\begin{aligned} &f_{Y_{m+1}|Y_1,Y_2,\cdots Y_m}(x | y_1, y_2, \cdots y_m) \\ &= \frac{1 - \prod_{i=0}^{m} \left(1 - g\left(|x - y_i|\right)\right)}{\int_0^L \left[1 - \prod_{i=0}^{m} \left(1 - g\left(|x - y_i|\right)\right)\right] dx}, \\ &\qquad m = 0, 1, \cdots n - 1, \end{aligned} \tag{17}$$

Eq. (17) is valid when $x \neq y_i, i = 1, 2, \cdots m$ as we assume each vehicle only broadcasts once. Particularly, when $m = 0$, we have the probability density function of the 1st broadcast vehicle's location

$$f_{Y_1}(x) = \frac{g(x)}{\int_0^L g(x) dx}. \tag{18}$$

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

$$
\begin{aligned}
& f_{Y_1,Y_2,\cdots Y_n}(y_1, y_2, \cdots y_n) \\
= \; & f_{Y_n|Y_{n-1},\cdots Y_2,Y_{n-1}}(y_n|y_{n-1}, \cdots, y_2, y_1) \\
& \times f_{Y_{n-1}|Y_{n-2},..Y_2,Y_1}(y_{n-1}|y_{n-2}, \cdots, y_2 y_1) \\
& \times f_{Y_{n-2}|Y_{n-3},\cdots Y_2,Y_1,}(y_{n-2}|y_{n-3}, \cdots y_2, y_1) \\
& \times \cdots \times f_{Y_2|Y_1}(y_2|y_1) \times f_{Y_1}(y_1), \qquad (19)
\end{aligned}
$$

where each conditional distribution in Eq. (19) is given by Eq. (17).

5 Simulation and discussion

In this section, numerical and simulation results are presented to discuss the relationship between the probability of correct message reception and its major performance-impacting parameters. Specifically, we adopt the unit disk model and the 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 range using DSRC[26]); for the log-normal connection model, we set the path loss exponent $\alpha = 2$, the standard deviation $\sigma = 4$[19] and the equivalent transmission range $r = 250$ m when $\sigma = 0$. Each simulation is repeated 5000 times and the average value is displayed in the plot.

Fig. 2 displays a comparison between the analytical result and the simulation result assuming the unit disk model, and indicates that the analytical result matches the simulation result well.

Fig. 3 and Fig. 4 illustrate 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, under different distance L between the source vehicle and the destination vehicle, and under different vehicular density ρ respectively. Specifically, we can see that $P_{\text{succ}} = 1$ when $p_m = 0$, which corresponds to 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 (0.5 in our system) when p_m is larger than a certain threshold p_{th}, e.g., $p_{th} = 0.2$ when $L = 3$ km and $\rho = 0.05$ veh/m. Beyond that threshold, a further increase in p_m has minimal 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 have 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. The minimum value of $P_{\text{succ}} = 0.5$ is because malicious vehicles in our network simply modify the received message without evaluation of the true content of the message. Therefore, when p_m is larger than its threshold, the message transmitted in the network will move between +1 and −1 alternatively, leading to the occurrence that P_{succ} converges to 0.5 instead of zero.

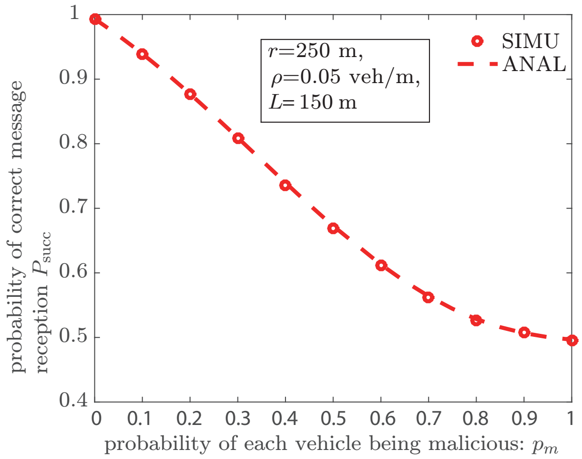

Figure 2 Comparison between analytical result and simulation result assuming the unit disk model

Fig. 3 indicates that given a fixed vehicular density, when $p_m < p_{th}$, a larger distance L between the source vehicle and the destination vehicle leads to a reduced P_{succ}. This is due to the fact that other things being equal, a larger L implies a larger number of malicious vehicles participating in modifying the message transmitted from the source vehicle to the destination vehicle. As a consequence, it leads to a reduced P_{succ}.

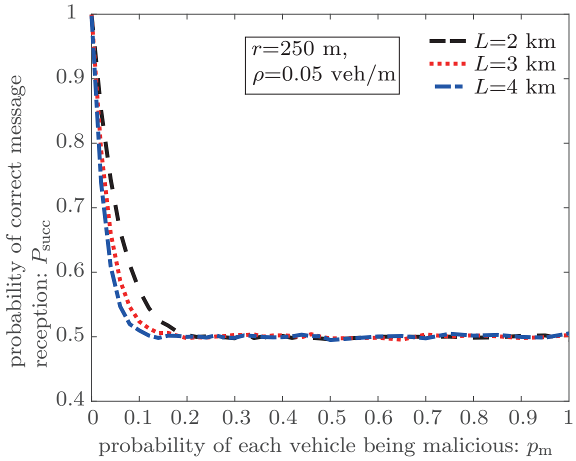

Figure 3 Relationship between probability of correct message reception P_{succ} and probability of each vehicle being malicious p_m assuming the unit disk model, with different distance L between the source vehicle and the destination vehicle

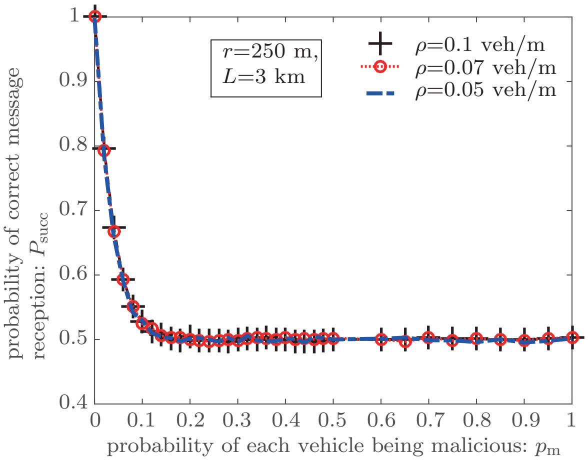

Figure 4 Relationship between probability of correct message reception P_{succ} and probability of each node being malicious p_m assuming the unit disk model, with different vehicular density ρ

Fig. 4 illustrates that in our system, a larger vehicular density ρ has minimal impact on P_{succ}. Intuitively, a larger ρ will lead to a greater P_{succ} because a larger ρ implies 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, other things being equal, the probability of correct message reception P_{succ} will increase. However, when a vehicle is randomly selected 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 to conduct a robust data fusion. 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 minimal impact on the P_{succ}.

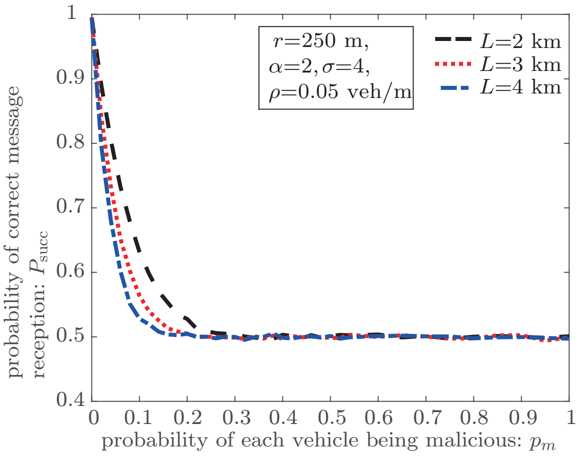

Figure 5 Relationship between the probability of correct message reception P_{succ} and p_m assuming log-normal connection model, with different distance L between the source vehicle and the destination vehicle.

Fig. 5 and Fig. 6 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, under different distance L between the source vehicle and the destination vehicle, and under 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 the unit disk model.

Therefore, we omit the duplicate discussion here.

Fig. 7 presents a comparison of the correct message reception probability P_{succ} achieved assuming the unit disk model (labeled as UDM) and the log-normal connection model (labeled as LSM). It is demonstrated that when $p_m < p_{th}$, the system assuming the log-normal connection model has a marginally 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 to other vehicles separated by greater distances. Therefore, other things being equal, each broadcast vehicle assuming the log-normal connection model can receive more copies of a message from other vehicles than that assuming the unit disk model, which leads to an improved message fusion result for each vehicle and consequently, results in a higher correct message reception probability.

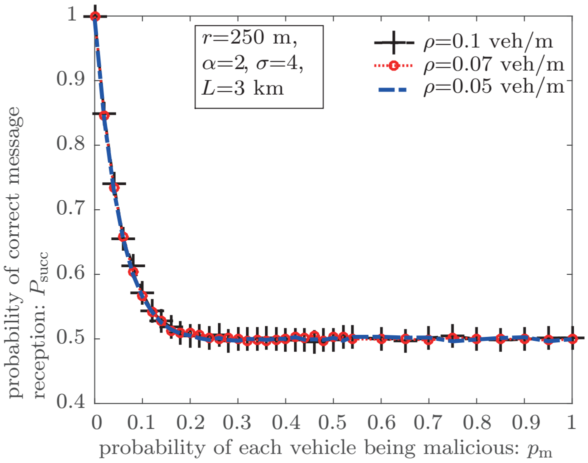

Figure 6 Relationship between the probability of correct message reception P_{succ} and p_m assuming the log-normal connection model, with different vehicular density ρ.

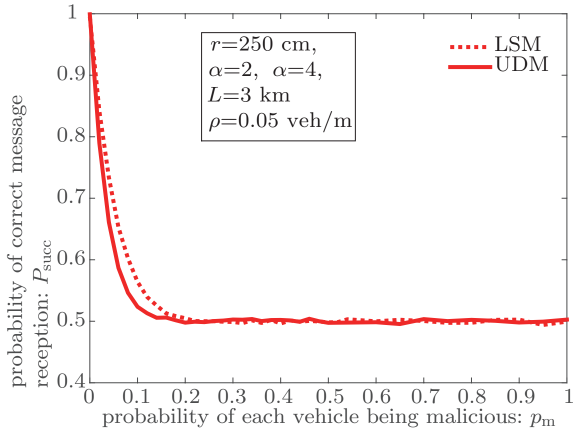

Figure 7 Comparison between probability of correct message reception P_{succ} achieved assuming the unit disk model and log-normal connection model.

6 Conclusions

This paper studied a vehicular Ad hoc network where a fraction of vehicles were malicious vehicles and these malicious vehicles were distributed randomly in the network. Furthermore, there was no central coordination among these malicious vehicles and consequently a malicious vehicle could simply modify its received message irrespective of its true value. An analytical framework was developed to model the process of secure message dissemination in the network, and the probability that a vehicle, located at a fixed distance from the source vehicle, could receive the message correctly was obtained. Simulations were conducted to establish the accuracy of the analytical results and 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. Further, a smaller distance between the destination vehicle and the source vehicle leads to a greater 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, Q. Zheng, P. Chatzimisios, et al. Heterogeneous vehicular networking: a survey on architecture,

challenges, and solutions [J]. IEEE communication survey & tutorials, 2015, 17(4): 2377-2396.

[2] J. Kenny. Dedicated Short-Range Communications (DSRC) standards in the United States [J]. Proceedings of the IEEE, 2011, 99(7): 1162-1182.

[3] S. Ilarri, T. Delot, R. Trillo-Lado. A data management perspective on vehicular networks [J]. IEEE communication survey & tutorials, 2015, 17(4): 2420-2460.

[4] G. Mao, B. D. O. Anderson. Graph theoretic models and tools for the analysis of dynamic wireless multihop networks [C]//IEEE Wireless Communications & Networking Conference, 2009: 1-6.

[5] T. Gazdar, A. Rachedi, A. Benslimane, et al. A distributed advanced analytical trust model for VANETs [C]//IEEE Global Communications Conference (GLOBECOM) 2012: 201-206.

[6] U. Khan, S. Agrawal, S. Silakari. Detection of Malicious Nodes (DMN) in vehicular Ad hoc networks [J]. Procedia computer science, 2015, 46: 965-972.

[7] N. Haddadou, A. Rachedi, Y. Ghamri-Doudane. A job market signaling scheme for incentive and trust management in vehicular Ad hoc networks [J]. IEEE transactions on vehicular technology, 2015, 64(8): 3657-3674.

[8] H. Sedjelmaci, S. M. Senouci. An accurate and efficient collaborative intrusion detection framework to secure vehicular networks [J]. Computers and electrical engineering, 2015, 43: 33-47.

[9] S. Dietzel, J. Petit, G. Heijenk, et al. Graph-based metrics for insider attack detection in VANET multihop data dissemination protocols [J]. IEEE transactions on vehicular technology, 2013, 62(4): 1505-1518.

[10] M. Raya, P. Papadimitratos, V. D. Gligor, et al. On data-centric trust establishment in ephemeral Ad hoc networks [C]//IEEE Infocom the Conference on Computer Communications, 2008: 1238-1246.

[11] Z. Huang, S. Ruj, M. A. Cavenaghi, et al. A social network approach to trust management in VANETs [J]. Peer-to-peer networking and applications, 2014, 7(3): 229-242.

[12] K. Zaidi, M. B. Milojevic, V. Rakocevic, et al. Host-based intrusion detection for VANETs: a statistical approach to rogue node detection [J]. IEEE transactions on vehicular technology, 2014, 65(8): 6703-6714.

[13] J. Radak, B. Ducourthial, V. Cherfaoui, et al. Detecting road events using distributed data fusion: experimental evaluation for the icy roads case [J]. IEEE transactions on intelligent transportation systems, 2016, 17(1): 184-194.

[14] S. K. Dhurandher, M. S. Obaidat, A. Jaiswal, et al. Vehicular security through reputation and plausibility checks [J]. IEEE system journal,2014, 8(2): 384-394.

[15] W. Li, H. Song. ART: an attack-resistant trust management scheme for securing vehicular Ad hoc networks [J]. IEEE transactions on intelligent transportation systems, 2016, 17(4): 960-969.

[16] R. Nelson. Probability, stochastic processes, and queueing theory: the mathematics of computer performance modeling [M]. New York: Springer-Verlag, 1995.

[17] N. Wisitpongphan, B. Fan, P. Mudalige, et al. Routing in sparse vehicular Ad hoc wireless networks [J]. IEEE journal on selected areas in communications, 2007, 25(8): 1538-1556.

[18] A. B. Reis, S. Sargento, F. Neves, et al. Deploying roadside units in sparse vehicular networks: what really works and what does not [J]. IEEE transactions on vehicular technology, 2014, 63(6): 2794-2806.

[19] Z. Zhang, G. Mao, T. Han et al. Cooperative information forwarding in vehicular networks subject to channel randomness [C]//IEEE international conference on communications, 2014: 324-329.

[20] Z. Zhang, G. Mao, B. D. O. Anderson. Stochastic characterization of information propagation process in vehicular ad hoc networks [J]. IEEE transactions on intelligent transportation systems, 2014, 15(1): 122-135.

[21] M. Azimifar, T. Todd, G. Karakostas, et al. Vehicle-to-vehicle forwarding in green roadside infrastructure [J]. IEEE transactions on vehicular technology, 2016, 65(2): 780-795.

[22] K. Abboud, W. Zhuang. Stochastic analysis of a single-hop communication link in vehicular Ad hoc networks [J]. IEEE transactions on intelligent transportation systems, 2014, 15(5): 2297-2307.

[23] G. Mao. Connectivity of communication networks [M]. New York: Springer, 2017.

[24] G. Mao, B. D. O. Anderson. Towards a better understanding of large scale network models [J]. IEEE/ACM transactions on networking, 2012, 20(2): 408-421.

[25] G. Mao, B. D. O. Anderson. Connectivity of large wireless networks under a general connection Model [J]. IEEE transactions on information theory, 2013, 59(3): 1761-1772.

[26] Z. Haibo, L. Bo, T. H. Luan, et al. ChainCluster: engineering a cooperative content distribution framework for highway vehicular communications[J]. IEEE transactions on intelligent transportation systems, 2014, 15(6): 2644-2657.

About the authors

Jieqiong Chen [corresponding author] received the B.E. degree in engineering from Zhejiang University, Zhejiang, 310027, China, in 2012. She is now pursuing the Ph.D. degree in engineering with the University of Technology Sydney, Sydney, Australia. Her research interests include the area of wireless communications and intelligent transportation systems. (Email: jieqiong.chen@student.uts.edu.au)

Guoqiang Mao joined the University of Technology Sydney, in 2014 as a professor of Wireless Networking and the Director of Center for Real-time Information Networks. Prior to that, he was with the School of Electrical and Information Engineering, The University of Sydney. The Center is among the largest university research centers in Australia in the field of wireless communications and networking. He has published about 200 papers in international conferences and journals, which have been cited over 4500 times. His research interest includes intelligent transport systems, applied graph theory and applications in telecommunications, Internet of Things, wireless sensor networks, wireless localization techniques, and network performance analysis. He is an editor of the IEEE transactions on wireless communications (since 2014), the IEEE transactions on vehicular technology (since 2010) and received the Top Editor Award for outstanding contributions to the IEEE transactions on vehicular technology in 2011, 2014, and 2015. He is a Co-Chair of the IEEE Intelligent Transport Systems Society Technical Committee on Communication Networks. He has served as a Chair, Co-Chair, and TPC Member in a large number of international conferences. (Email: g.mao@ieee.org)