Bayesian Trust Scheme: A Decentralized Safety Message Trust Method in Multi-Hop V2V Networks

Hanaa S. Basheer, Carole Bassil, and Bilal Chebaro

Abstract—Vehicular Ad hoc Networks (VANETs) have an important role in improving road safety especially when no infrastructure is available. Inter vehicle communications (IVCs) provide decentralized communications where vehicles cooperate together to disseminate road traffic data relying on broadcasting reports and warning messages. Trusting the data of every warning message must be accomplished during dissemination as its information is public. Many researches have concentrated on securing the system entities by adding authenticity to each vehicle or aggregating digital signature. These traditional security schemes at some point needed a central management. In this paper, we introduce an approach to trust the information of the warning message before message dissemination as its information is public. Many researches were published to improve system reliability and to disseminate warning messages properly between vehicles [1]. Since warning messages are created to be sent in real time and public information, that’s why maintaining data integrity is considered as a big challenge against authorized nodes that act maliciously. To avoid depending on center server, vehicles authorization and communications security have focused on trusting the propagation data. In 2008, Raya et al were the first to establish data trust instead of entities trust for ephemeral ad hoc networks to prevent authorized nodes from manipulating message information [2]. IVCs are going after less usage of infrastructure, so revocation of misbehaving nodes must be done by the node's private defense, based on evidences available from the cooperative of neighboring nodes at the same area. When we start thinking of building a dissemination protocol to propagate warning messages between vehicles without infrastructure, we do concentrate of how to cover a wide area with the warning message and assure other vehicles about its correct information. Trusting message's information was the solution to this issue. In this paper we are going to involve the Bayesian Network (BN) to help nodes through multi-hop V2V communication in making a decision, if to trust the incoming warning message or not. BN represents probability distributions in a computationally tractable way because it depends on few variables. A General Bayesian network “is a graphical structure that allows representing an uncertain domain, the strength of the relationship between variables is quantified by conditional probability distributions associated with each node” [3]. The rest of this paper is structured as follows. We start in Section II with describing to the main issues we are depending on to start our research. Section III presents our suggested two schemes to trust the warning message throughout a one hop and multi-hop communication. In Section IV, we discuss the technical presentation of Bayesian trust model. Section V presents the ways in defining the data distribution for the Bayesian trust scheme. Section VI shows the graphical representation of the scheme. Section VII will give an analyzing to our simulation results, before we sum up our paper with a conclusion in Section VIII.

I. INTRODUCTION

The increasing in roads congestion, leads to more roads accidents. The need for urgent solutions becomes substantial to help in saving human lives. Currently, intelligent vehicles are equipped with GPS, onboard units (OBU) and digital map to reach the goal of Inter Vehicle Communication (IVC) by connecting vehicles together as the Vehicle to Vehicle (V2V) communications. Vehicles also can connect to the internet throughout fixed Road Side Units (RSUs) to form a vehicle to infrastructure (V2I) communications. These units support the propagation of traffic information such as sending a real time safety report, which can be created by a vehicle, depending on its own sensors, to warn neighboring vehicles about an abnormal situation.

Several projects were conducted worldwide to address road safety and to benefit from vehicles wireless communication to build reliable networks with less time delay on the channel access protocols. Moreover, many
II. VANET ISSUES

In this paper we are interested in trusting warning message data before disseminating it to all neighboring nodes. Our goal is to transmit a trusted warning message through multi-hop V2V communication, without depending on a central server to secure the data entity. The following subsections will give a review for the issues that are related.

A. Dissemination Methods

The main advantage of the broadcast dissemination technique is that it does not need to know the IP address of the destination, which is thus effective in a very dynamic network such as VANET. Flowing data rapidly may cause a broadcast storm problem, while controlling data flow may not help in covering a wide area. Thus, adding conditions to the blind data flooding are used to propagate the warning message to all neighboring nodes in the same radio area. Passing the packet through router is done either as a single hop broadcast, or as a multi-hop broadcast. One of the suggested techniques in multi-hop broadcast is the use of a selected nodes set, to relay on to retransmit the message farther. This reduces messages Redundant Rate (RR) to a suitable value, since high RR may lead to a broadcast storm problem, while low RR may cause a loss for the message [4]. Another way is to choose the best recipients among many nodes to be the best candidates that can re-broadcast the message and ensure its proper propagation [4], [5]. To avoid different broadcast problems the researchers in [6], [7] suggested to classify the multi-hop broadcast methods to three categories basing on the way of messages disseminations; the time delay, the probability, and the network coding based.

B. Data Trusting Methods

Two kinds of messages need to be available for driver's benefit; the safety and non-safety messages. Our concern is to disseminate high priority warning messages quickly and correctly. Attacking messages during propagation by authorized participants in the networks considered as the weakest point in V2V communication. Meanwhile, a safeguard for connected vehicles must be against the insider misbehaving nodes as well as the outsider which can be achieved by adding a trust management model [8]. A qualitative risk analysis is presented in [9], that organized VANET assets rate from the highest important asset which is the transmitting of a safety message from the RSUs, followed by the distributing and protection mechanisms of the data, and then communication system (i.e. hardware, protocols, and media). To achieve a decentralized data trust scheme without losing nodes privacy, a study in 2011 manage to classify the trust models into three categories; the entity-oriented trust models, the data-oriented trust models, and the combination of trust models that uses the peer trust to evaluate the trustworthiness of data [10]. For data trustworthy many reputation systems of VANET were introduced such as the decentralized VARS (a vehicle ad hoc network reputation system), which based on collecting events for decision-making for messages confidential [11]. However VARS steps suffer from some unwanted effects when the set of opinion is empty. Another approach of self-organized trust management aims to prevent internal attackers from sending false messages based on trusting relations between direct connections as well as indirect connections through intermediate nodes [12]. In 2013 a decentralized scheme was presented to generate message’s trust value depending on Perron–Frobenius theorem. The scheme based on the assumptions that every vehicle has the ability to recognize the receiving message’s node situation (either as a source or an intermediate node), and also adds vehicle self-observation with the aggregating reputation of other vehicles about the messages’ nature and their priority [13]. The work seems to be promising; however it does not explain how vehicles depend on reputation and was not implemented for real world scenario yet. In 2015 another research depend on Perron-Frobenius theorem as well, called an Analysis Hierarchy Process (AHP). This approach trusts to exchange messages by coalesce various recommendation opinions from neighbor vehicles. To evaluate the trust values, the vehicles kinds vehicles (e.g. police, school van …etc.) and the message strength were taking into consideration [14].

From all the above related works we can conclude that almost all decentralized data trust schemes depend on the environment evidences to settle a trust relationship between vehicles.

III. DISSEMINATING TRUSTED MESSAGE MODEL

In this work we are focusing on building a strong decentralized trust method for a multi-hop V2V networks. Our model presents a two-stage data trust scheme to ensure the correctness of the warning message information during propagation. First stage trusted the message information when transmitted through a single hop and we called it as an endorsement trust scheme (ETS). The second stage of the trust scheme starts when trying to forward the message further to the next hop and we called it as a Bayesian trust scheme (BTS). At each stage a different data trust scheme is processed to ensure warning message information and to reduce the effect of the authorized nodes that might act maliciously.

The following two subsections will include details about the two stage data trust scheme.

A. Single Hop Trust Scheme: ETS

Using a proper message dissemination scheme with a data trust scheme at the same hop is addressed in a previous work of the authors [15]. The proposed work starts by assuming that the digital map of the highways is
divided into small fixed segments, each considered as a single hop as illustrated in Fig. 1. Each node $D_i$, which is placed at the front of others starts sending a packet to all neighboring nodes in the same hop telling them its status. The packet contains a peer of information $<B, M>$; the beacon ($B$) and the warning message ($M$), which will remain empty in the normal situation. When $D_i$ receives back information from the connected nodes about their statuses, it adds a weight value to each connecting link. Node $D_i$ calculated the weight value by summation the thresholds of three available values from the beacon’s fields, which are: the distance from the behind node, the behind node speed and the number of how many nodes it connected to ($C$). Any node $D_i$ can arrange an ordered list of all connected nodes starting from the node with the highest link weight value. This list is updated every interval of time. In case any sudden abnormal event shows, the nearest node will be the source node and starts creating a warning message to disseminate it immediately to all the nodes behind and wait for an acknowledgement (ACK) from the node with the highest weight value, which is the forwarder node [15]. When the forwarder ($F$) receives the warning message and sends back an ACK, it will keep checking of how many nodes did send the same message during the same interval time. Node $F$ chooses a random number $N$ bounded between the values $C/2$ and $C$. If node $F$ endorses this message by ($E$) nodes or more and $E > =N$, then the message information is promising and can be trusted to be rebroadcasted further to the next hop. Otherwise if $E$ is between $N$ and $\frac{N}{2}$ then the forwarder must make a decision based on a binomial distribution as in formula (1) to decide whether to trust the message or neglect it [16].

$$P(E) = \binom{N}{E} \times P^E \times (1-P)^{(N-E)}$$

(1)

where the value $E = 0, \ldots, N$. Upon testing our data sample we discovered that, with probability of 0.50 and permittivity of 0.025, node $F$ can make a decision of trusting the incoming message information to be rebroadcasted further.

Since our algorithm based on dividing the roads into segments, and each segment considered as a hop, then vehicles would response only to messages coming from other vehicles placed in the same hop. To avoid this restrict, we suggested that the nodes placed at the first few meters can have the ability to exchanging packets with the nodes from the front hop.

We called these few meters from each segment a safety zone to indicate the place where nodes have the ability to receive from the segment ahead. To understand our suggestion the node $V_i$ in Fig. 2 is placed in the safety zone and started exchanging packets with the nodes from the segment ahead. If the packets have a warning message, then $V_i$ immediately will start BTS to either trust the message and rebroadcast it further or not.

BTS helps in making a trust decision based on four variables, which are: global speed slowing dawn, the position of the nearest exit point from the highway, vehicles changing lane motion and then increasing in the road density. In the coming sections we are going to explain our suggested BTS in details.

IV. TECHNICAL PRESENTATION OF BTS

When an observation is made for new variables, an updating inference will be created upon the new information. This process is performed via a “flow of information” through network [3]. Our scenario begins when a node ($V_i$) reaches the safety zone as illustrated in Fig. 2, it will have the ability to exchange packets $<B, M>$ with the front hop. If $V_i$ receives $M$ that’s alert for an abnormal situation, then $V_i$ starts checking its validation (by checking the fields of $M$; the time stamp, the tag that indicates if the sending node is a source node and $M$ priority value). In addition $V_i$ must begin the process to trust $M$ depending on some variables.

Fig. 2. The forwarder node is rebroadcasting the message to the node placed at the safety zone from the next hop

First if $V_i$ records a global slowing speed, with a sporadic in cars flowing, then $V_i$ predicts this to be happened either because the existing of a near exist point from the highway or because of the existing of an abnormal situation. The node $V_i$ can specify these changing by checking the changing in the incoming packets’ speed and place fields and GPS information. In case that there is no highway exist point near, our model continues testing the next variable to make a decision to either trust the incoming warning message $M$ or not. Our
model based on the probability distribution of a huge data that where collected by Iraqi Police General Directorate. Thus secondly $V_i$ tested the following two variables simultaneously:

- Nodes rapidly changing their lane
- Increasing in road density

To demonstrate BTS procedure, we illustrate its steps in algorithm 1.

Algorithm 1: node $V_i$ Making decision

1. At the safety zone if message $M$ received from the front hop start
2. Check: if $V_i$ records speed slowing and a sporadic in cars flowing the go to 3, otherwise go to 6
3. Check: if a highway exist point is near then go to 6 otherwise go to 4
4. Making decision after checking all cases of both events; the changing in road density $\{T, F\}$ and rapidly lane changing $\{T, F\}$
5. Based on the probability distribution of all variables $V_i$ either trust M or not
6. End

V. DATA DISTRIBUTION FOR BTS

Every node $V_i$ depended on a conditional probability distribution tables that must establish from different resources. Here we based our simulator on the accidents data that were collected from the Iraqi Police General Directorate database for four years from 2011-2014, where we focused on the highway road that connects Baghdad with Ninava city. Moreover a previous study [17] mentioned that a car slows down 21% from its speed when deciding to change direction and to exit from a highway road to an urban city. The database information with the study result assists in building the conditional probability table of the two variables; speed slowing and existing accident $P(\text{ex}|\text{ss})$.

All conditional probability tables with the joint probability function will help $V_i$ in making the final decision of trusting the incoming warning message to either be forwarded further or not.

Previously many works used the Bayesian belief network techniques to predict cars crash as in [18], [19]. Where both references depended on a large dataset (e.g. dataset can be available from geographical information system GIS database) to build the posterior probability and evidences factor.

Another use for Bayesian network goes for trusting management model to reduce the effect of malicious nodes [20]. However, this approach depends on initial central values preset according to RSUs and also depends on vehicles’ different kinds, which are classified to be as an indicator of direct trust evidences. We will be the first to use the Bayesian network to build a decentralized data trust scheme.

In this paper we choose the Poisson Distribution (PD) to get the priori information, since this distribution is usually associated with rare events like cars accident. From Table II and the use of the probability mass function as in formula (2) we could predict the probability value of existing accident in both cases (true/false).

$$P(x) = \frac{m^x e^{-m}}{x!}$$

where (x) represents the average number of accident in each year and it can takes the values 0, 1, 2,…. The parameter (m) known as the event rate, and $m=(\text{total no. of accidents on the same highway area}) / (\text{total no. of accidents over period of time})$.

Depending on prior information illustrated in Table I, we calculated the event rate (m), and based on the predicated number of accident in each year we calculated the Poisson distributed probability as in Table II.

We now get that the probability distribution when no accident happened ($P(\text{ex})$) is equal to 0.30, thus $1-P(\text{ex}) = 0.70$ to indicate the probability distribution when accident is happened.

While the probability of global speed slowing dawn happening near of highway exit point is $P(\text{ss})$ and equal to 0.80 calculated based on the questioner of [17], so the probability distribution for no speed slowing is $1-P(\text{ss}) = 0.20$. The conditional probability table of the two discrete random variables ($P(\text{ss})$ and $P(\text{ex})$), $P(\text{ex}|\text{ss})$ for the variables in Fig. 3 are shown in Table III.

$m = \text{total no. of accidents on Ninava highway over four years}$

<table>
<thead>
<tr>
<th>Year</th>
<th>Total number of accidents</th>
<th>Total number of accidents on highway</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>426</td>
<td>124</td>
</tr>
<tr>
<td>2012</td>
<td>504</td>
<td>126</td>
</tr>
<tr>
<td>2013</td>
<td>529</td>
<td>209</td>
</tr>
<tr>
<td>2014</td>
<td>166</td>
<td>83</td>
</tr>
<tr>
<td>Total</td>
<td>1625</td>
<td>Total = 542</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Avg. Number of accidents (x)</th>
<th>Probability P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>1</td>
<td>0.359</td>
</tr>
<tr>
<td>2012</td>
<td>1</td>
<td>0.359</td>
</tr>
<tr>
<td>2013</td>
<td>2</td>
<td>0.219</td>
</tr>
<tr>
<td>2014</td>
<td>1</td>
<td>0.359</td>
</tr>
<tr>
<td>Total</td>
<td>1.5 no normalization</td>
<td></td>
</tr>
</tbody>
</table>
Different other datasets were used to build the remaining conditional probability tables.

For lane changing probability we depend on the number of alert messages that vehicles send to inform neighboring nodes about its new position. Meaning that if there is a rapidly changing in nodes y-axis positions, then the probability of lane changing P(lc) is increasing and vice versa. This dataset is also helped us in calculating the probability distribution of the changing in road density since vehicles can’t change lanes when the road traffic is dense. Road density is the second variable we depend on in the BTS. Fig. 4 illustrates the remaining conditional probability tables.

![Conditional probability tables for our model](image)

### VI. GRAPH PRESENTATION OF BTS

Our goal is to give the node at the safety zone the ability to process a trust scheme on the data of the incoming warning message from the front hop before rebroadcast it further. Depending on many variables we now can build a trust scheme based on the Bayesian network to help in making a decision (md) either to trust the incoming message or not. The variables have the probability of happening (true T) or not (false F), and they are with their probability symbols as follow:

- Road Global speed slowing P(ss)
- The position of the nearest exist point from the highway P(ep)
- The road density P(rd)
- Nodes action to change lane P(lc).

Depending on the general definition of Bayes formula (3) from [18], any prior probability P(w_i) can be converted to posterior probability P(w_i|z), P(z|w_i) which is the likelihood of w_i, and P(z) as the evidence factor.

\[
P(w_i|z) = \frac{P(z|w_i) \times P(w_i)}{P(z)} \quad (3)
\]

Based on the ability of graph presenting for BN, our BST can be represented as acyclic graph as in Fig. 5, which illustrates all the independent variables. The figure also shows also the symbol of each joint probability function. BST starts after receiving a warning message, and as we mentioned previously the node will begin checking the road global speed slowing. If there is a speed slowing dawn, the node must decide is it because of a nearby highway exit point P(ep|ss), or because of an abnormal situation occur P(ex|ss), which we concentrate on in our algorithm.

When the probability of accident existing is higher, the next step will start by testing the probability of road density P(rd|ex) and vehicles lane changing P(lc|ex) simultaneously.

![Model Scenario after receiving a valid warning message from the front hop](image)

To start estimating the final decision of trusting scheme P(md) based on the joint probability function 4 [21]:

\[
P(md) = \sum_{ss, ex, rd, lc} P(ss, ex, rd, lc, md)
\]

\[
= \sum_{ss, ex, rd, lc} P(ss)P(ex|ss)P(rd|ex)P(lc|ex)P(md|rd, lc)
\]

### VII. ANALYZING THE SIMULATION RESULTS

After finishing the preparation of all the conditional probabilities for all the independent variables, we reach the final point of building our decision making model BTS. The model now is capable of answering the question: "Is the incoming warning message trusted or not?" when two or more confidence variables are available.

By using the Bayes formula (2) with different road situations and formula (3) we get the probability results of trusting the incoming warning message as illustrated in Table IV, noting that we take ss = true for all cases with the confidence of any other evidence.

![Trust probability values when 3 – 1 events available from 4 events](image)
We can conclude from the trust probability results in Table III that the more variables we depend on to trust the incoming message data, the more reliable decision we can get. In case 1 and case 2 three variables out of four are available, and thus a comfortable trust decision can be taken, while the other cases with only two variables available showed a disparity in probability values. Case 6 has the worst trust probability since only one variable is available, which leads to ignoring the incoming message as shown in Fig. 6.

Moreover, case 4 and case 5 with (ex = false) mean that the slowing down is due to some highway existing points. Here, the driver should be aware of the surrounding situation.

<table>
<thead>
<tr>
<th>Case</th>
<th>Variables</th>
<th>Probability</th>
<th>Prob. P(md=T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ex = true, lc=true, rd=false</td>
<td>( \sum_{x \in T} P(ss = T, ex = T, lc = T</td>
<td>md = T) )</td>
</tr>
<tr>
<td>2</td>
<td>ex = true, lc=false, rd=true</td>
<td>( \sum_{x \in T} P(ss = T, ex = T, rd = T</td>
<td>md = T) )</td>
</tr>
<tr>
<td>3</td>
<td>ex = true, lc=false, rd=false</td>
<td>( \sum_{x \in T} P(ss = T, ex = T</td>
<td>md = T) )</td>
</tr>
<tr>
<td>4</td>
<td>ex = false, lc=true, rd=false</td>
<td>( \sum_{x \in T} P(ss = T, lc = T</td>
<td>md = T) )</td>
</tr>
<tr>
<td>5</td>
<td>ex = false, lc=false, rd=true</td>
<td>( \sum_{x \in T} P(ss = T, rd = T</td>
<td>md = T) )</td>
</tr>
<tr>
<td>6</td>
<td>ex = false, lc=false, rd=false</td>
<td>( \sum_{x \in T} P(ss = T</td>
<td>md = T) )</td>
</tr>
</tbody>
</table>

VIII. CONCLUSIONS

One of the most important issues in VANET is to trust messages data before broadcasting them farther. In our model, we depend on two stages to trust the emergency messages; first stage begins when the chosen forwarder receives the emergency message from the source node that sensed the abnormal situation, and the second stage starts when a node placed at the safety zone of the next segment receives the message from the forwarder. Our model depends on roads traffic data that were collected by the Iraqi Police General Directorate during four years period from 2011-2014, but we focused on the highway road that connects Baghdad with Ninawa city only. Poisson distribution method is used for the three main environment events that our model depends on to help in building the conditional probability tables. This work is considered to be the first work that depends on Bayes network to help in taking a decision either to trust the incoming emergency message when the probability \( \geq (50% \text{ constant}) \%), or changing it to an attention message to aware the driver of a possible abnormal situation, this is when the probability is between (50-20)%). Otherwise message is ignored if probability is less.

REFERENCES


Hanaa S. Basheer was born in Iraq. She received her B. Sc. in Applied Science / Mathematics from Univ. of Technology/Iraq in 1991. From 1993 till 2003, she joined University of Baghdad and work as a computer laboratory supervisor besides teaching mathematics tutorial. During that period of time she received her Higher Diploma and MSc. in computer science / Data security from Institute for Post Graduate Studies in Information/ Iraq in 1998 and University of Technology/Iraq 2002 respectively. In 2003 she continued her career in Baghdad University as an instructor till now. In 2011 she started her PhD program and now she is a candidate researcher in VANET at Lebanese University.

Carole Bassil received her PhD in Computer science, and she is specialized in security and VANET fields. She is now an Associate professor at Lebanese University, Beirut, Faculty of Science II-Dept. of Computer and Statistics.

Prof. Bilal Chebaro, PhD. Applied Math, Lebanese University, Beirut, Faculty of Science I-Dept. of computer and Statistics.