VANET Alert Endorsement Using Multi-Source Filters*

Tiffany Hyun-Jin Kim[§] Ahren Studer[§] Rituik Dubey[§] Xin Zhang[§] Adrian Perrig[§]
Fan Bai[†] Bhargav Bellur[†] Aravind Iyer[†]

[§] Carnegie Mellon University
{hyunjin, astuder, rituik, xzhang1, perrig}@cmu.edu

[†] General Motors Research
{fan.bai, bhargav.bellur, aravind.iyer}@gm.com

ABSTRACT

We propose a security model for Vehicular Ad-hoc Networks (VANETs) to distinguish spurious messages from legitimate messages. In this paper, we explore the information available in a VANET environment to enable vehicles to filter out malicious messages which are transmitted by a minority of misbehaving vehicles. More specifically, we introduce a message filtering model that leverages multiple complementary sources of information to construct a multi-source detection model such that drivers are only alerted after some fraction of sources agree. Our filtering model is based on two main components: a threshold curve and a Certainty of Event (CoE) curve. A threshold curve implies the importance of an event to a driver according to the relative position, and a CoE curve represents the confidence level of the received messages. An alert is triggered when the event certainty surpasses a threshold. We analyze our model and provide some initial simulation results to demonstrate the benefits.

Categories and Subject Descriptors: C.2.0 [Computer – Communication Networks]: General – *security and protection*; C.2.1 [Computer – Communication Networks]: Network Architecture and Design – *Wireless communication*
General Terms: Algorithms, Design, Security
Keywords: VANET, Misbehavior Detection

1. INTRODUCTION

Within the US, vehicular accidents result in over 34,000 deaths [12] and cost motorists 164.2 billion dollars a year [3]. A recent study found drivers stuck in traffic congestion in 2007 waited for 4.2 billion hours and wasted 2.8 billion gallons of fuel [19]. In the near future, vehicles will possess On-Board Units (OBUs) which wirelessly communicate with other OBUs or Road-Side Infrastructure (also called Road-Side Units (RSUs)). Applications in such Vehicular Ad-hoc Networks (VANETs) can improve roadway safety while reducing congestion through real-time traffic management [1]. However, security mechanisms are needed to prevent malevolent behavior. For example, the Electronic Emergency Brake Light (EEBL) application is meant to alert drivers of rapidly decelerating nearby vehicles which may not be visible (e.g., hidden from view by a large truck), reducing the chance of multiple vehicle collisions. However, a maliciously crafted EEBL message could cause a driver to suddenly decelerate or swerve, causing a dangerous situation or even an accident, that may not happen without VANET. The proposed IEEE 1609.2 standard uses a public-key infrastructure and digitally signed messages to secure VANET applications [11]. However, cryptography only enables origin authentication of a message that could ultimately identify a culprit only after an attack. Instead, a mechanism to detect spurious messages *during an event* is needed to ensure correct operation of applications, prevent VANET-induced accidents, and facilitate the successful adoption of VANETs.

Our goal is to propose a model that enables OBUs to accurately label VANET messages as legitimate or spurious. This is an important challenge since spurious messages can occur as a result of malevolence or a faulty sensor. Error-free and tamper-proof sensors and hardware could prevent the generation and distribution of spurious messages, but are prohibitively expensive.[1] Prior works have proposed misbehavior detection mechanisms for VANETs [6, 7, 8, 13, 14, 17], which individually leverage different sources of information to detect a variety of attacks. In this work, we describe a framework to combine these complementary sources as a means to utilize all of the information available to OBUs to corroborate the validity of a V2V message. Specifically, we examine the following 6 sources of information:

1. **Cryptographic Authentication:** Does the message include a valid digital signature and certificate?
2. **Source Location:** Is the sender in a valid and relevant location?
3. **Local Sensors:** Do my local sensors support the alert?
4. **Other Vehicles' Messages:** Do other vehicles' messages support or contradict a given message?
5. **Infrastructure Validation:** Do RSUs with sensors embedded in the road support the message?

*This research was supported in part by CyLab at Carnegie Mellon under grant DAAD19-02-1-0389 from the Army Research Office, grant CNS-0627357 from the National Science Foundation, and by support from General Motors through the GM-CMU Collaborative Research Laboratory. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of ARO, CMU, GM, NSF, or the U.S. Government or any of its agencies.

VANET'10, September 24, 2010, Chicago, Illinois, USA.

[1] For example, the IBM-4764 (a high-security cryptographic coprocessor) costs over $8,000 [10].

6. **Sender Reputation:** Did the sender previously broadcast spurious data?

In Section 3.2, we provide more detailed definitions and examples of the 6 sources. Our system combines the data from various sources to calculate a Certainty of Event (CoE) value. The OBU only notifies the driver if the CoE exceeds a threshold, which varies with distance to provide faster notification for nearby events while reducing the number of alerts associated with far away events that are irrelevant to the driver. Faster notification for nearby events is important as it provides drivers with more time to respond.

Contributions. This work provides a framework for VANET misbehavior detection that allows the combination of different sources of information in a systematic fashion. Rather than limiting ourselves to cryptography, local sensors, or reputation alone, we investigate a holistic synthesis that harmonizes different, possibly contradictory, sources of information. Evaluating the CoE versus our proposed threshold curve will dramatically reduce the number of spurious driver alerts, even in environments without attackers. We also present a theoretical analysis and a simulation-based evaluation to provide guidelines on how to configure the system and to demonstrate the benefits of the system. Rather than being a complete solution for VANET misbehavior detection, our system provides a general framework such that any future improvements to misbehavior detection that leverage a given source of information could be "plugged into" our system as a way to improve detection performance.

2. PROBLEM DEFINITION

In this section, we present a concise problem definition, attacker model, assumptions, and the trust model.

2.1 Problem Definition

Given a road with a small but non-negligible fraction of vehicles that transmit spurious messages which contain false information (either intentionally or accidentally due to malfunctioning units), we need to ensure that legitimate vehicles can filter out such spurious messages with high probability. In this paper, we introduce a message filtering model that verifies the validity of a received message (it is outside the scope of this work to distinguish whether a spurious message is transmitted by a malicious node or a malfunctioning node). We propose a filtering model that leverages multiple sources of information and alerts the driver only after some fraction of sources agree.

We will evaluate the fidelity of our message filtering model along two dimensions:

1. efficacy of the filtering model based on the collision percentage, the percentage of vehicles in a region which are involved in a vehicular collision, and
2. the delay between when a message is received and when the driver is alerted since delivering safety messages to drivers on time is critical.

2.2 Attacker Model

We consider active attackers who violate the integrity of messages (i.e., attackers create bogus alerts, or suppress legitimate messages). More specifically, attackers may inject malicious messages announcing invalid driving safety information and attempt to propagate bogus information to other vehicles on roads. Such active attacks may encourage legitimate drivers to change driving behavior; for example, legitimate drivers may slow down or decide to take alternative paths if the bogus message announces that some hazardous material is spilled on the road ahead. As a result, attackers succeed in disrupting normal driving behavior, clearing the road to lower congestion. Attackers may also suppress legitimate alerts of critical safety information from further propagation by simply dropping packets. This attack may prevent legitimate drivers from being warned about some critical safety information that they will encounter. Consequently, these legitimate drivers may not have enough time to properly react when they reach the relevant zone, and in the worst case, attackers succeed in further exacerbating the situation near the hazardous region.

We also address attackers that have compromised other vehicles, and/or inject malicious messages from outside the area of relevance: attackers who are not physically located on the related roads, attackers who propagate malicious messages from the opposite direction, and attackers who are approaching the related area.

We consider Denial of Service (DoS) attacks, such as jamming, beyond the scope of this work.

2.3 Assumptions and Trust Model

In order to distinguish between spurious and legitimate messages, we assume that:

- Vehicles communicate using the Dedicated Short Range Communication (DSRC) technology.
- Vehicles adhere to the IEEE 1609.2 standard for VANET security [11]. More specifically, a Public Key Infrastructure is available and all OBUs can authenticate certificates to identify senders as valid VANET participants and signatures to validate the message contents.
- Vehicles are equipped with minimal local sensors, e.g., thermometer, GPS, accelerometer, etc.
- A limited percentage of vehicles are equipped with advanced sensing equipment. For example, radar and/or LIDAR equipment, which can independently detect stopped or slowly moving vehicles, may be available on some high-end vehicles.
- The majority of vehicles are honest, and malicious or malfunctioning vehicles represent a small fraction of the VANET population.

There may be roadside infrastructure, such as Road Side Units (RSUs) deployed by trusted authorities, and vehicles trust messages that RSUs generate. However, we do not make a bold assumption that RSUs exist on roads. Instead, we conjecture that RSUs will be deployed in the near future, and be available for VANET. In the beginning, RSUs may be deployed sparsely in general, probably with more RSUs in urban areas due to greater vehicle density.

3. MISBEHAVIOR DETECTION MODEL

To ensure that a vehicle's OBU delivers legitimate warnings that announce critical driving conditions and prevent drivers from being inconvenienced and disturbed by spurious warnings, we propose a general model to distinguish invalid messages from legitimate ones. Our model is based on a multi-source filtering model; given 6 complementary sources of information, we propose that an OBU tests the validity of a received message. More specifically, the OBU of a vehicle investigates a received message by aggregating the results from all applicable sources (among 6 sources). Only when the aggregated result indicates that the message is valid, the OBU confirms that the received message is announcing some real safety condition, and warns the driver. In Section

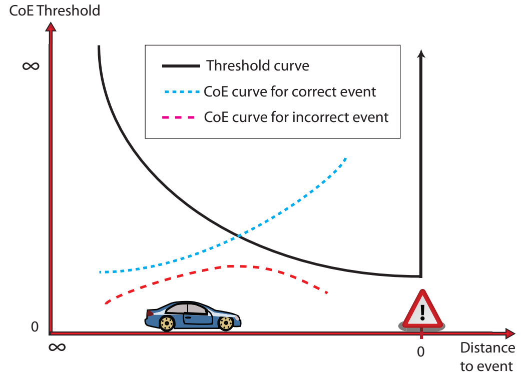


Figure 1: Threshold and CoE curves. We chose the direction of the x-axis to indicate a vehicle driving from left to right, approaching the event positioned at the 0 mark. A driver is alerted when the CoE curve crosses the threshold.

3.1, we provide an overview of our detection protocol, and in Section 3.2, we delineate 6 sources of filters in detail. In Section 3.3, we describe how the results from the 6 sources are combined to determine the validity of the received messages.

3.1 Overview

We propose an alert endorsement model that is based on two main components: a threshold curve, and a CoE curve. When vehicles exchange warning messages, it is critical that a receiving vehicle determines the validity of incoming messages and only alert the driver once the system has determined that the messages are legitimate. We suggest that an alert is triggered if the event certainty surpasses a threshold. We now present an overview of the two main components.

Threshold Curve. The importance of an event to a driver depends on the distance between the event and the driver; a nearby event is more important to the driver (There may be other factors that determine the importance of an event, such as the vehicle's current velocity, but in this work, we only consider the distance as an initial suggestion for deriving a threshold curve). Based on this observation, we suggest a threshold curve considering three aspects:

- Delayed event announcement: we suggest a threshold value that prevents an OBU from notifying the driver too early when the event is far away from the driver. Delayed announcements can also prevent invalid alerts from unnecessarily impacting driver behavior.
- Low false positives: a high threshold value is desirable to avoid accepting an invalid message. This is especially true when the corresponding event is far away.
- Low false negatives: a low threshold value is desirable in order to prevent any legitimate message from being filtered out. This is especially true as the driver approaches the corresponding event.

As a result, we suggest threshold values that diminish as the driver approaches the event. Figure 1 shows a sample threshold curve. Note that the threshold curve rises to infinity when the driver's position matches the event location since further alerts regarding the event at this location are meaningless as the driver passes the event.

CoE Curve. It is challenging to evaluate whether a received message reports a real critical event or not, unless the event is within the driver's vicinity. Hence, it is important that an OBU collects and corroborates messages which warn about a potential event. An interesting observation is that all vehicles that observe a real event tend to send messages to inform others that are approaching the event. Thus, an OBU of a vehicle approaching the event will receive multiple messages that either directly report the event or imply some abrupt change in traffic condition. On the other hand, if the reporting event is invalid, the OBU of a vehicle behind the event does not receive as many messages as it would for a real event since only malicious drivers will transmit such spurious messages. As a result, the OBU can gain confidence on the reported event's validity based on the number of valid messages received. We name the confidence of an event as Certainty of Event (CoE). The CoE value increases as the OBU approaches the real event since the number of messages reporting that event increases, and the OBU alerts the driver when the CoE curve crosses the threshold curve, as shown in Figure 1. On the other hand, the CoE value decreases as the OBU approaches a spurious event for the following reason: as the number of messages reporting the spurious alert does not increase, the alert is only useful for a limited time period. Hence, the importance of the alert decays over time. In this case, it is unlikely that the CoE curve crosses the threshold curve, as shown in Figure 1, and the OBU does not alert the driver.

Besides the number of other vehicles that report the same event, there are multiple criteria that an OBU may consider to determine the CoE value. We distinguish 6 complementary sources of inputs. An OBU considers each source to a received message to verify if the message satisfies the condition of the source, and adjusts CoE value for each source.

3.2 6 Sources of Inputs for CoE

In this section, we delineate 6 complementary sources for filtering in detail, and explain how each source is used to validate the legitimacy of received messages.

Source 1: Cryptographic Authentication. A vehicle's OBU authenticates received messages using public key cryptography in order to detect invalid safety messages. With IEEE 1609.2 [11], which provides a Public Key Infrastructure (PKI) for OBUs to authenticate messages, we can ensure that each vehicle has exactly one valid certificate at a given time (e.g., no Sybil attacks are possible), and receivers can use that certificate to authenticate any received messages. Unfortunately, cryptographic authentication is insufficient by itself because an adversary may have compromised the cryptographic keys of vehicles or may have altered the inputs to a vehicle's sensors.

Source 2: Geographic Location Validation. Vehicles are only concerned about safety warnings related to where the vehicle is headed (i.e., area of relevance) [17]. This implies that OBUs only need to verify messages that are generated by other OBUs that are located within the area of relevance. For example, a recipient can apply the Convoy Member Authentication and the Vehicle Sequence Authentication [17] to check whether senders are driving with and in front of the recipient. Furthermore, OBUs can use the senders' previous location claims to detect malicious attackers who announce messages from outside the area or relevance. More specifically, the receiving OBUs may utilize maps to investigate whether senders of the warning messages are traveling on the roads as indicated in the messages.

Source 3: Local Sensors. Vehicles are manufactured with various sensors, and OBUs may use data from those sensors for verification purposes. For example, an OBU can use a thermometer to invalidate a spurious message reporting icy road conditions. If the recipient vehicles are equipped with special devices, such as radars, the OBUs of such vehicles can also verify if the senders of the warning messages indeed exist in the location as indicated in the messages.

Source 4: Responses from Other Vehicles. All vehicles that are driving toward the security warning zone behave similarly when the message is valid. As a result, OBUs can confirm the validity of the received messages by checking how other vehicles respond as they approach the region. For example, an OBU can detect that a message about the existence of the debris is spurious if other vehicles near the debris do not slow down but rather drive through the debris.

Source 5: RSU Validation. If RSUs are deployed along the roads, RSUs can provide information on road conditions, traffic patterns, etc. that OBUs may use to verify messages. For example, an OBU can filter a malicious message that reports congestion in an area while RSUs do not indicate congestion in that area. An OBU can also deduce the validity of a message about traffic patterns, such as the sudden hard braking, if RSUs indicate the existence of a tight curve where vehicles tend to brake hard.

Source 6: Reputation.[2] OBUs may infer the validity of the received messages based on the reputation of the vehicles that report such data. In other words, if OBU A has received false information from some OBU B recently, then OBU A may not trust what OBU B reports thereafter since there is a high probability that OBU B may still lie or malfunction. As a result, the negative reputation of OBU B may aid OBU A to easily filter out its spurious messages.

3.3 Decision Making Procedure

In this section, we explain how an OBU determines the validity of a received message. Essentially, an OBU tests the received message based on the information from several sources, as described in Section 3.2, and combines their outputs. When the combined output implies that the message is relevant and legitimate, the OBU alerts the driver.

3.3.1 Prioritization of Sources

Several automotive applications that have been proposed to enhance safety and convenience in vehicular networks [1], and every application has different security and safety requirements. Given the application diversity, only a subset of our 6 sources may apply; consequently, OBUs may verify received messages based on selected applicable sources only. For example, a message for the Emergency Electronic Brake Light (EEBL) application, where a vehicle braking hard broadcasts a warning message, only relates to the message authentication (Source 1), location of the sender (Source 2), other vehicles' responses (Source 4), and the previous interaction with the sender (Source 6). Local sensors and RSU validation may not be applicable for EEBL events that are out of range of the local radar.

In order to minimize computational power, OBUs may prioritize the above 6 sources in specific orders based on the application. Even with efficient authentication mechanisms [9, 16], OBUs can further reduce computation if they can avoid the process of any cryptographic authentication. For example, a warning message of an icy road in the middle of the summer can be ignored even before cryptographically verifying the message. On the other hand, the same warning message from a nearby RSU, which is deployed by trusted authorities, can confirm that the road is indeed icy. Consequently, for RHCN messages (where a vehicle detecting a road hazard (e.g., fluid, ice, debris) notifies other vehicles within the potentially affected region), we suggest that OBUs apply all sources in the order of 5→2→1→3→6→4. Note that such ordering is a suggestion to harmonize different sources in a way to reduce the computational overhead, but the prioritization of the ordering can be changed depending on design criteria. The same ordering of sources may be applied for all event-driven applications whose purpose is to notify drivers about the actual road incidences (e.g., EEBL, SVA (Stopped/Slow Vehicle Advisor), PCN (V2V Post Crash Notification), and RFN (Road Feature Notification). For periodic-routine applications that warn drivers about potential collisions such as CCW (Cooperative Collision Warning) and CVW (Cooperative Violation Warning), such ordering does not apply.

Moreover, depending on the applications, outputs from certain sources may be more influential than those of other sources in determining the validity of received messages. For example, an OBU may consider RHCN messages from RSUs to be more meaningful than those from nearby vehicles because RSUs are governed by trusted authorities. This implies that the OBU may place more weight on Source 5 (RHCNs from RSUs) than on Source 1 (cryptographic authentication of RHCNs from nearby vehicles). Our framework accommodates fully or partially trusted RSUs. In this paper, we consider fully trusted RSUs.

3.3.2 Certainty of Received Messages

Determining the validity of a received message depends on two variables: relative location of the event and the time of the event. If an OBU receives a message about an event that is closer to the OBU, then it may consider this message to be more relevant for detection. For example, an OBU would categorize nearby EEBL messages as more relevant than more distant EEBL messages. Similarly, a recent message is more relevant for detection than an aged one, because the importance of an event decays over time.

We define CoE to represent the confidence level of an event at a specific point in time. The confidence level is derived by combining values of all outputs of applied sources, where these applied sources are determined by the specific application for the event. When the CoE value is greater than some threshold (which is application-specific) at a specific point of time, it alerts the driver.

We suggest two approaches for representing CoE: linear and cumulative. We now describe each in detail.

Linear Approach. In the linear approach, the CoE value for an event computed by vehicle v_n is determined by the alert messages input from other vehicles v_i at the specific time period regardless of the CoE values of other vehicles. An OBU considers a received message for CoE calculation *only when the message is verified by Sources 1 and 2, and by Sources 3 and/or 5 if they are available* (i.e., these sources act as pre-filters to remove obviously faulty data first). Then a vehicle v_n calculates the CoE for an event e as follows:

$$CoE_{(v_n,e)} = \sum_{i,j} l_{6(v_i)} \cdot (\alpha_j \cdot l_{4(v_i)}) \qquad (1)$$

where

- l_6 is the reputation value of the reporting vehicle v_i where $i \neq n$ (Source 6),
- l_4 indicates the reception of alert messages from vehicle v_i (Source 4) (e.g., for an EEBL application, $l_4 = 1$

[2] Various reputation systems are available to be incorporated into our filtering model. Their technical feasibility is another study which we do not address in this work.

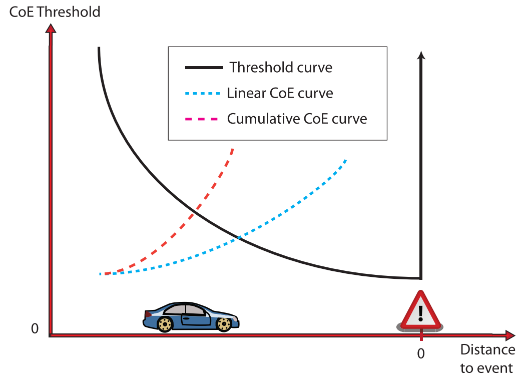


Figure 2: Comparison of linear and cumulative CoE.

for EEBL, RHCN, or SVA messages from v_i and 0 otherwise), and

- α_j is a weight based on the input message type j (e.g., $\alpha_{RHCN} = 1.2$, $\alpha_{EEBL} = 0.5$, etc.).

Cumulative Approach. People tend to trust some news with higher confidence when they hear the same news from others at the same time. Similarly, an alert on a certain event is more trustworthy when it is reported by other vehicles. As a result, the cumulative approach places greater weight on an event that has been reported by preceding vehicles with high confidence about that event themselves, thus accumulating validity faster than the linear approach.

Unlike the linear approach, the cumulative approach requires each OBU to keep track of not only $l_{4(v_i)}$ generated by a preceding vehicle v_i for the same event, but also the alert messages that v_i received from others for the same event. Similar to the linear approach, an OBU considers a received message for CoE calculation *only when the message is verified by Sources 1, 2, 3, and 5*, and a vehicle v_n calculates the CoE for an event e as follows:

$$CoE_{(v_n,e)} = \sum_{i,j} l_{6(v_i)} \cdot (\alpha_j \cdot l^{cum}_{4(v_i)}) \tag{2}$$

where

- l_6 and α_j are the same as in the linear case,
- $l^{cum}_{4(v_i)}$ is the aggregated alert messages from vehicles v_k in front of v_i and from v_i itself (i.e., $k \leq i$) as follows:

$$l^{cum}_{4(v_i)} = \sum_{k,j} \alpha_j \cdot l_{4(v_k)} \tag{3}$$

Consider vehicles v_{i-1} and v_i driving in front of v_n and generating the same alert message (e.g., EEBL) (v_{i-1} is in front of v_i and generates the message first). When v_n's OBU first receives EEBL from vehicle v_{i-1}, the OBU increments $CoE_{(v_n,EEBL)}$ by $l^{cum}_{4(v_{i-1})}$, which includes the alert message from v_{i-1} itself ($l_{4(v_{i-1})}$). Within some short time τ, when v_n's OBU receives another EEBL from v_i containing $l^{cum}_{4(v_i)}$, the OBU increments $CoE_{(v_n,EEBL)}$ by $l^{cum}_{4(v_i)}$, which also includes v_{i-1}'s alert message $l_{4(v_{i-1})}$ (see Equation 3). In other words, v_n incorporates v_{i-1}'s alert message $l_{4(v_i)}$ twice for CoE calculation, increasing the weight of the messages.

Figure 2 shows linear and cumulative CoE graphs for a correct event. In this case, the cumulative approach endorses the alert faster than the linear approach.

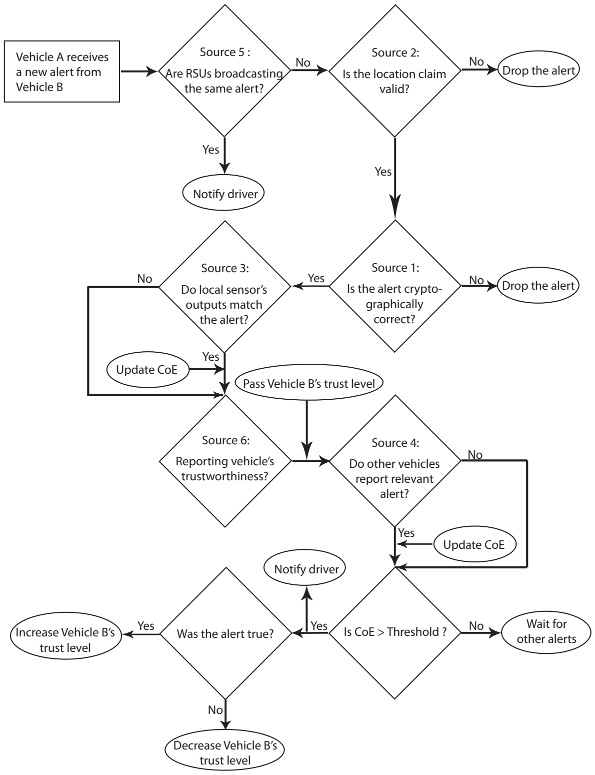

Vehicle A receives a new alert from Vehicle B
Source 5 : Are RSUs broadcasting the same alert?
No
Source 2: Is the location claim valid?
No
Drop the alert
Yes
Notify driver
Yes
Source 3: Do local sensor's outputs match the alert?
Yes
Source 1: Is the alert crypto-graphically correct?
No
Drop the alert
No
Yes
Update CoE
Pass Vehicle B's trust level
Source 6: Reporting vehicle's trustworthiness?
Source 4: Do other vehicles report relevant alert?
No
Yes
Update CoE
Notify driver
Increase Vehicle B's trust level
Yes
Was the alert true?
Yes
Is CoE > Threshold ?
No
Wait for other alerts
No
Decrease Vehicle B's trust level

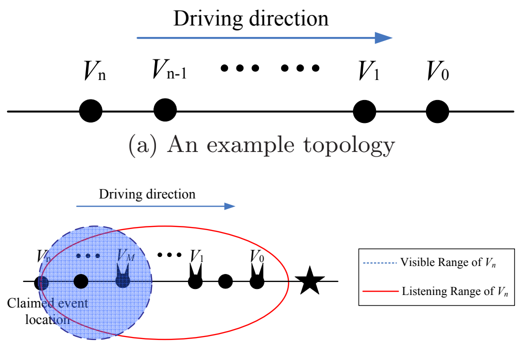
Figure 3: A complete data flow of endorsing an alert.

Threshold Function for CoE. Every CoE value is compared with some predetermined threshold that depends on the message deadline and the importance of events. The permissible latency for message authentication and the data origin properties vary widely between different safety applications. As a result, alerts may have different verification deadlines. For example, an EEBL notification is generated by a single vehicle and has a tight latency constraint, while an RHCN message requires the aggregate of a large number of vehicles and is not as time sensitive as the EEBL notification. Along with the permissible latency, the importance of events influences the threshold for determining the validity of messages; messages that report severe safety conditions must be distinguishable from those that do not report urgent conditions. As a result, the threshold to endorse an urgent safety alert may be tighter than the threshold to endorse a moderate safety alert. Note that thresholds for each application are predetermined by system designers/engineers.

Based on the CoE value and some predetermined threshold as explained above, an alert is evaluated as legitimate when the CoE exceeds the threshold, and is sent to the driver's attention. On the other hand, if the CoE does not exceed the threshold before the event lifetime, the alert is considered as fraudulent and is discarded. Figure 3 represents the complete data flow of endorsing an alert.

3.4 Discussion

In this section, we have proposed a general model to address the issue of how an OBU can endorse an alert given malicious vehicles. Our approach is based on a threshold curve and a CoE curve that is derived from 6 complementary sources. Our intention is that this general model assists other VANET research. In the following section, we theo-



(a) An example topology

(b) An attack scenario

Figure 4: Vehicle topology for EEBL application.

retically analyze our model, and in Section 5, we instantiate the model for the EEBL application. In Section 6, we show how prior work fits into our model.

4. THEORETICAL ANALYSIS

The design of the CoE and threshold curves directly affects the effectiveness and practicality of our model.In this section, we theoretically derive more general requirements and properties of the CoE and threshold curves.

4.1 Requirements of Curves for Correctness

When an event is indeed valid, we need to guarantee that the CoE and the threshold curves *intersect* such that an alert is raised (**correctness**) when the vehicle is at least within the distance d^* away from the event. In this case, the driver has sufficient time to respond to the event (**timeliness**). The sufficient conditions for achieving the correctness are as follows: (1) The CoE value monotonically increases as the vehicle approaches the event, (2) the threshold monotonically decreases as the vehicle approaches the event, (3) within the distance d^* to the event, the CoE value should be greater than the threshold value, and (4) when the distance to the event is far away to be irrelevant, then the CoE value should be smaller than the threshold value.

We now prove that our model can achieve these conditions, thus guaranteeing correctness and timeliness.

4.2 Design of CoE and Threshold Curves

We consider the following scenario with an EEBL application. Assume that a sequence of vehicles $v_0, v_1, \ldots, v_{n-1}$ is in front of vehicle v_n and they are in v_n's radio range, as Figure 4(a) shows. Let $CoE_{(v_n,EEBL)}$ denote the CoE value computed by v_n. We introduce $l_4(v_i) \in \{0,1\}$ for vehicle v_i such that $l_4(v_i) = 1$ if v_n receives an EEBL from vehicle v_i, otherwise $l_4(v_i) = 0$. The goal is to compute the local $CoE_{(v_n,EEBL)}$ for v_n based on the received $l_4(v_i)$'s from the other n vehicles in v_n's radio range. To simplify the analysis, we assume that the reputations of all vehicles are the same (i.e., $l_{6(v_i)} = 1$). Using these assumptions, we next analyze the local CoE values using the linear and cumulative aggregation models.

Linear Approach. Equation 1 directly yields

$$CoE_{(v_n,EEBL)} = \sum_{i=0}^{n-1} l_4(v_i) \tag{4}$$

which is linear with respect to the number of received notifications from other vehicles. Therefore, given a valid event (in which case vehicles $v_0, v_1, \ldots, v_{n-1}$ send their messages), we have:

$$CoE_{(v_n,EEBL)} = O(n). \tag{5}$$

Cumulative Approach. In this approach, each vehicle v_i individually computes $l_4^{cum}(v_i)$ which is the number of messages from preceding vehicles *as well as its own alert* if v_i itself validates the event. If v_i confirms that the event is valid, it will further propagate its local $l_4^{cum}(v_i)$. Eventually, v_n will compute $CoE_{(v_n,EEBL)}$ based on all received $l_4^{cum}(v_i)$'s, instead of the primitive $l_4(v_i)$'s. In our example scenario, we have:

$$\begin{gathered} l_4^{cum}(v_i) = \sum_{j=0}^{i} l_4(v_j) \\ CoE_{(v_n,EEBL)} = \sum_{i=0}^{n-1} l_4^{cum}(v_i) \end{gathered} \tag{6}$$

For example, suppose there is a valid event, and vehicles v_0, v_2, …, v_{n-1} sequentially witness the event and believe it. Then we have: $l_4^{cum}(v_0) = 1, l_4^{cum}(v_1) = l_4(v_0) + l_4(v_1) = 2, l_4^{cum}(v_2) = l_4(v_0) + l_4(v_1) + l_4(v_2) = 3, \ldots, l_4^{cum}(v_{n-1}) = n$. And finally we have: $CoE_{(v_1,EEBL)} = 1, CoE_{(v_2,EEBL)} = 1 + 2, CoE_{(v_3,EEBL)} = 1 + 2 + 3, \ldots$. Therefore given a valid event in which case vehicles $v_0, v_2, \ldots, v_{n-1}$ send their messages, we have:

$$CoE_{(v_n,EEBL)} = 1 + 2 + \ldots n = O(n^2) \tag{7}$$

Correctness and Timeliness. As a preliminary analysis, we consider the given scenario where vehicles $v_0, v_1, \ldots, v_n$ are driving in the same lane in order. As v_n drives toward the event location (i.e., distance to the event location d is decreasing), more vehicles in its radio range have already driven past the event location and generated messages to report the event. Therefore, the value of $CoE_{(v_n,EEBL)}$ will increase as the vehicle is approaching the event in both linear and cumulative approaches. Correspondingly, we can also select a curve for the threshold to satisfy the correctness and timeliness requirements.

Resilience to Spurious and Suppressed Alerts. In the benign case, $CoE_{(v_n,EEBL)}$ increases faster in the cumulative approach than in the linear approach; hence, the cumulative approach generates an alert earlier (see Figure 2).

With the existence of the attackers sending bogus alerts, the cumulative approach is vulnerable to false alerts (false positives). Assume that there are M malicious vehicles within the radio range of v_n which collectively fabricate a bogus event with distance d_m to v_n (where d_m is outside the visible range of v_n), as shown in Figure 4(b). Suppose the threshold value with distance d_m is $TT(d_m)$. To cause a false alert on v_n, the number of malicious vehicles sending bogus alerts required in the linear approach is:

$$\begin{aligned} CoE_{(v_n,EEBL)} > TT(d_m) &\Rightarrow O(M) > TT(d_m) \\ &\Rightarrow M > O(TT(d_m)) \end{aligned} \tag{8}$$

In the cumulative approach:

$$O(M^2) > TT(d_m) \Rightarrow M > O(\sqrt{TT(d_m)}) \tag{9}$$

As the above equations show, it requires fewer attackers to cause a false alert in the cumulative approach compared to the linear approach. On the other hand, with the presence of attackers suppressing valid alerts, the linear approach is more subject to false negatives since it requires receiving more valid alerts to trigger a local alarm.

4.3 Relationship between CoE and Threshold

As explained in the previous section, a certain number of malicious vehicles can inject a false alert to a victim v_n. In some applications, once the alert is raised on v_n, it may automatically generate and further propagate its own alert. In this case, there exists a potential possibility of *cascading* the false alert propagation: once v_n is convinced and propagates its own message, a vehicle v_{n+1} following v_n can receive $m+1$ messages (m from the malicious attackers and one from v_n). Though v_{n+1} is farther away from the claimed event location (thus holds a higher threshold value), it also computes a higher CoE value since there are now $m+1$ messages. We denote the increase of the threshold on vehicle v_{n+1} as Δ_{TT} and the increase of the CoE value on v_{n+1} as Δ_{CoE}. Clearly, when $\Delta_{TT} < \Delta_{CoE}$, a false alert will be raised at v_{n+1} (and maybe so forth for $v_{n+2}, v_{n+3}, \ldots$), thus cascading the propagation of the false alert. To prevent this, we desire:

$$\Delta_{TT} > \Delta_{CoE} \tag{10}$$

This equation indicates the following guideline: The curve of the threshold should have a larger derivative in distance d than that of the CoE curve in the number of received messages n.

Summary. In this section, we have formalized the general requirements for the CoE and threshold curves and presented an in-depth analysis of the linear and cumulative approaches for computing CoE values with respect to correctness, timeliness, and resilience to spurious messages. Based on our analysis, we show that 1) both the linear and cumulative approaches can achieve correctness; 2) the cumulative approach has better timeliness since it can reach the threshold faster; and 3) the linear approach has better resilience to spurious messages. Finally, we also derive the relationship between the CoE curve and the threshold curve. After exploring these design trade-offs and requirements, we anticipate that these theoretical insights can help guide the practical design of the CoE and threshold curves.

5. SIMULATION

We simulate our system using ns-2 to demonstrate the use of our model when applied to EEBL. As a preliminary work, the main goal of the simulation is to select and tune the parameters of the filters and to demonstrate their impact. We quantify the impact of EEBL by measuring the collision percentage, defined as the fraction of the 100 vehicles that are involved in a collision [21]. We also measure the delay associated with the filters, the false positive rate, and the false negative rate.

5.1 Simulation Environment

We simulate a straight stretch of a road with a single lane and inject 100 vehicles. These vehicles travel along the road at 25 m/s periodically sending VANET messages and decelerate in response to hitting other objects or seeing other vehicles decelerating. The delay between when a driver sees the preceding vehicle decelerate and begins to decelerate herself depends whether the OBU alerted the driver.

Vehicle inter-arrival time is sampled from an exponential distribution [2] skewed to have a lower inter-arrival time cutoff of $0.7sec$ (otherwise spacing between the $5m$ long vehicles is unrealistically small).

Every $0.1sec$ each vehicle sends out a wireless message. The message contains information regarding the location and the velocity of the vehicle, and includes EEBL data only when the deceleration of the vehicle is greater than zero.

Vehicles only decelerate when they hit another vehicle or object or when the preceding vehicle is decelerating. When vehicles collide (i.e., the distance between the center of two vehicles is less that or equal to the length of a vehicle), the deceleration value depends on how rigid the collision is (i.e., how much distance a vehicle is allowed to move before coming to a complete stop, assuming that the vehicle is traveling at a speed of $25m/s$). We assume three possible cases:

1. A vehicle collides into a perfectly static target (e.g., a wall). We assume that the vehicle can move a quarter of its length ($1.25m$) before coming to rest. The deceleration is $250m/s^2$.
2. A vehicle collides into a partially yielding target (e.g., another stationary vehicle jammed against a wall). We assume the vehicle can move half of its length ($2.5m$) before coming to rest. The deceleration is $125m/s^2$.
3. A vehicle collides into a yielding target (e.g., another vehicle on brake). We assume that the vehicle can move the distance equal to its length ($5m$) before coming to rest. The deceleration is $60m/s^2$.

When a driver sees the preceding vehicle braking, the deceleration of a vehicle becomes a constant ($8m/s^2$) after the driver reaction time.

The reaction time is the delay from when a driver witnesses the preceding vehicle begins to decelerate to when the driver begins to apply the brakes. In our simulation scenarios, we assume that a driver requires a reaction time of $1sec$ when he has not seen an EEBL alert. We assume that after the driver is warned with an EEBL alert, the driver reaction time is reduced to $0.5sec$ [4].

5.2 Attack Scenarios

We simulate 3 scenarios:

- **Case 1:** No malicious vehicles are present,
- **Case 2:** Malicious vehicles are present that send out fake EEBL messages even when they are not decelerating,
- **Case 3:** Malicious vehicles are present that do not send out EEBL messages even when they are decelerating.

In real life, a combination of Cases 2 and 3 is possible, but for evaluation purposes we consider each of the cases separately as they tend to counter the effects of one another when they occur together. We assume 2 attackers in both Cases 2 and 3 and place them at the beginning of the vehicle train to assess their effects on the following vehicles. We choose 2 attackers to test collusion for Case 2; a single malicious vehicle that sends a spurious message can be easily filtered out, but with two colluding vehicles, filtering malicious messages becomes more challenging.

5.3 Measurement Metrics

We assume a linear threshold curve for evaluation purposes with two degrees of freedom: the slope and the intercept. The intercept (i.e., the value of the threshold curve at distance 0) is based on the assumption that a single EEBL suffices to exceed the threshold if two vehicles are apart by a minimum possible distance (i.e., the length of a single vehicle) without any collision. Consequently, the threshold curve is defined as $y = m \cdot x + c$, where m is a slope, c is an intercept, and x is the distance from the closest vehicle which

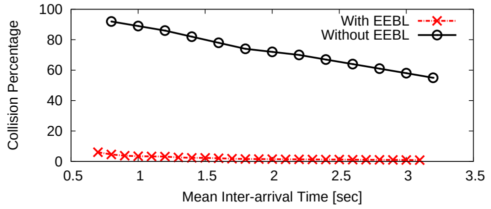


Figure 5: Comparison of collision percentage with and without EEBL.

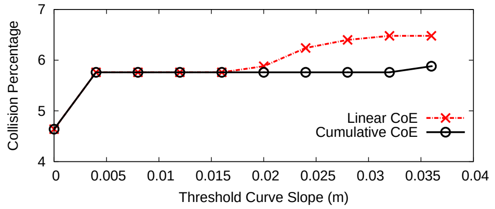


Figure 6: m for linear and cumulative CoE schemes.

reports an EEBL. We can select c for a particular value of m as follows:

$$c = 1 - m \cdot d_{min} \qquad (11)$$

where d_{min} is the length of a vehicle equal to 5 meters.

To measure the delay induced by the filters, we record three values: the instant at which a vehicle receives the very first EEBL, the time duration from this instant to trigger an alert, and the distance from that vehicle. A false positive occurs when a vehicle delivers an alert to the driver due to a fake EEBL message. A false negative occurs when the vehicle receives an EEBL message first, but the driver is never alerted and decelerates due to other external factors besides EEBL (e.g., an observable collision or a visual reaction to the braking of the immediately preceding vehicle). The rate for each is the fraction of the time such an error occurs.

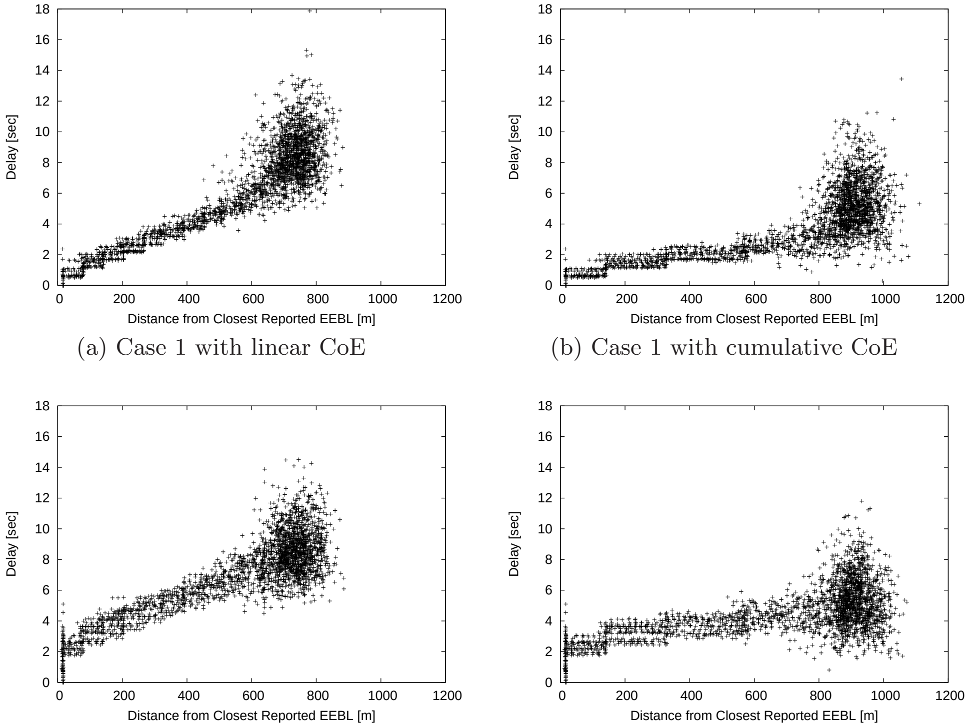
To measure the collision percentage, we first vary m; we choose a value for m which is as big as possible but which still gives a value of collision percentage comparable to the case without our filtering model. We call this value of m the operating point and measure the delay for the chosen operating point. We also measure the effect of the vehicle inter-arrival time on the collision percentage.

5.4 Simulation Results and Discussion

In this section, we describe the simulation results from a number of different experiments. We evaluate the performance of EEBL with and without our misbehavior detection model using EEBL messages and suggest the parameters for the threshold curve.

Experiment 1. In order to measure the effectiveness of EEBL messages before applying our filtering model, we set up an experiment to compare the collision percentages with and without EEBL. Basically, we set the foremost vehicle to collide into a static target at some point of time and measure the resulting collision percentage. We vary the inter-arrival time among vehicles and repeat the simulation 50 times for each inter-arrival time value to average out any noise in the observation. The average of all the runs is the resulting collision percentage as shown in Figure 5.

We observe that without EEBL, the collision percentage ranges from 92% to 54% as inter-arrival times increases



(a) Case 1 with linear CoE

(b) Case 1 with cumulative CoE

(c) Case 3 with linear CoE

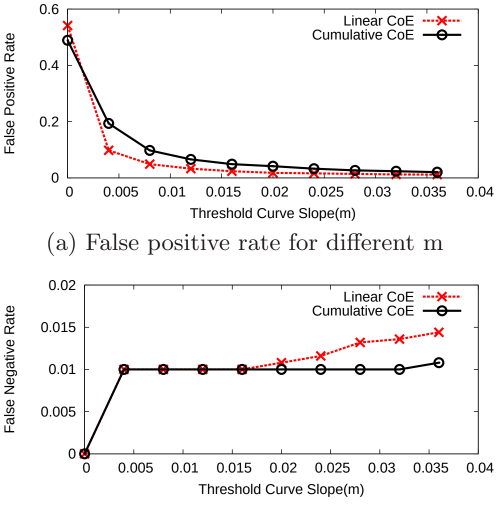
(d) Case 3 with cumulative CoE

Figure 7: Delay from Filtering Model.

from 0.7 to 3 seconds, respectively. When EEBL messages are broadcast, the collision percentage is very low (2∼6%) throughout the range of inter-arrival times. This low collision percentage implies that EEBL messages help mitigate human reaction delays by informing every vehicle in the radio range about the braking. We also observe that the collision percentage drops as the mean inter-arrival rate increases, which is intuitive since we expect fewer collisions at a given average speed as the vehicle density decreases. Thus, for the following experiments we keep the inter-arrival rate as 0.7*sec* to evaluate the worst-case behavior.

Experiment 2. In this experiment, we insert our filtering model into the EEBL framework. As mentioned in Section 5.3, we measure the collision percentage for both linear and cumulative CoE curves given a linear threshold curve and derive desirable m values for two CoE curves. The variation of collision percentage for different slope values of m is plotted in Figure 6. As shown in Figure 6, both linear and cumulative cases present 3 phases: an initial rise in collision percentage, followed by a flat portion, followed by a further rise. We would like to increase the slope of the curve (i.e., proportional to m) as much as possible without significantly increasing the collision percentage, choosing the highest slope point on the flat portion, to reduce alerts for far away events. Thus, we choose $m = 0.016$ for the linear and $m = 0.032$ for the cumulative approach. The delay induced by our filtering model for Case 1 (no malicious vehicles) under the linear CoE curve and the cumulative CoE curve using $m = 0.016$ are shown in Figures 7(a) and 7(b). Similarly, the delay results for Case 3 (malicious vehicles do not send out EEBL messages) under the linear and cumulative CoE curves using $m = 0.032$ are shown in Figures 7(c) and 7(d). Figures 8(a) and 8(b) show the false positive rate from Case 2 and false negative rate from Case 1 among 100 vehicles. The slopes for the linear and cumulative models deliver low false positive and false negative rates.

Based on the delay plots in Figure 7, we can infer that the delay due to our filtering model increases as the distance from the closest reported EEBL increases. This is a direct consequence of the threshold curve. The rate of rise of the delay with distance is greater in the linear CoE model than in the cumulative CoE model because in the latter case we have a quadratic accumulation of alerts for a linear rise in threshold, which leads to comparatively lower delay values for higher distances.



(a) False positive rate for different m

(b) False negative rate for different m

Figure 8: False Positive & False Negative Rate for Linear and Cumulative Approaches.

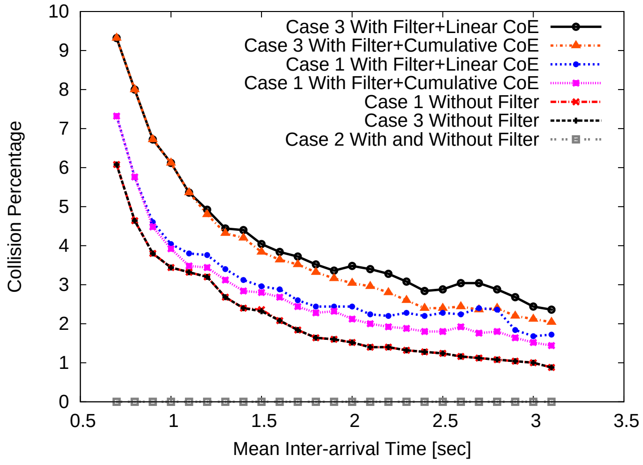


Figure 9: Collision percentages comparison.

Experiment 3. Based on the values of m from Experiment 2, we derive the values for c based on Equation 11 and Figure 9 for both linear and cumulative CoE approaches in Cases 1-3. We also plot the graphs for the case without our filtering models for Cases 1 and 3 in the same figure for comparison purpose.

There is a tension between minimizing false positives and false negatives. We conjecture that people would start ignoring VANET alerts when they sparsely reflect the real events. Therefore, reducing the number of false alerts is important such that drivers fully pay attention to the real alerts. In our simulation, we have engineered the filtering mechanism to mitigate attackers who craft spurious EEBL messages (Case 2), resulting in 0 collision given fake EEBLs. As a consequence, our filter discards legitimate EEBL messages and prevents fewer accidents in the other cases. However, Figure 9 shows that the presence of our filtering model does not lead to a significant increase in collision percentage (fewer than 3 collisions among 100 vehicles) for both linear and cumulative models throughout the inter-arrival range. The increase in collision due to our model is slightly higher in Case 3 because the threshold curve is more sensitive to attackers that fail to report they are braking compared to the scenario where an alert is raised on the first heard EEBL. However, this increase is still limited to fewer than 3 collisions per 100 vehicles. We leave it as a future work to find the optimal balance between security and functionality.

6. RELATED WORK

A number of researchers have examined misbehavior detection and prevention in VANETs, but often focus on limited sources of information for the message validation.This work represents an attempt to combine the different sources available to an OBU and to systematize the decision process. In this section, we discuss how prior works relate to each of our 6 sources or the synthesis of different information.

In addition to the IEEE standard [11], researchers have proposed more efficient authentication mechanisms [9, 16] or privacy preserving key management [13, 18]. Such work can provide the cryptographic verification for Source 1.

Localization of an OBU within VANETs has received limited attention. Golle et al. [8] introduce a framework to detect and correct incorrect location claims or claims of fake vehicles in VANETs. The approach relies on individual nodes using their local sensor data coupled with a model of nominal VANET operation, and sharing this information with nearby vehicles. This approach allows a vehicle to verify an OBU's location (Source 2) using local sensors (Source 3) and reports from other vehicles (Source 4). Studer et al. also explore the authentication of the physical location and movement of vehicles [17]. Their approach leverages time of flight of VANET messages and continued presence to determine the position and direction of travel of other OBUs.

Reports from other vehicles (Source 4) help provide "data centric trust" [5, 14, 20] in VANETs. These works examine the data in messages from multiple vehicles and calculate how likely an event is based on reports from all of the vehicles in a region. Raya et al. [14] propose that vehicles in this framework use a decision logic system like Dempster-Shafer, Bayesian inference, or voting to integrate the related reports much like our CoE. However, their work assumes vehicles leave radio range rapidly, removing any use of reputations (Source 6), and ignore information from local sensors (Source 3) or verification of a location claim (Source 2).

Ghosh et al. [6, 7] and Schmidt et al. [15] construct reputation models for other vehicles (Source 6) based on the claims from sending vehicles, a model of normal behavior, and data from local sensors (Source 3), which provide a form of ground truth. In such work, an OBU determines whether or not a safety message was spurious by analyzing how the driver behaves in response to the event. For example, the OBU will consider a PCN (Post-Crash Notification) invalid if the vehicle subsequently drives through the claimed crash site. Such techniques help an OBU determine the truth about a notification and assign reputations to other vehicles (Source 6), but fail to provide any kind of misbehavior filtering while the driver is approaching an event.

7. CONCLUSION

Researchers have accomplished much progress over the past decade to secure VANETs. Misbehavior detection appears to be the final major problem that needs to be solved. Unfortunately, misbehavior detection appears fundamentally impossible to address, because ultimately we cannot determine the true root cause for an alert message: the alert may be legitimate even though highly unlikely (e.g., a truck could have dropped ice on the road in summer), the sensor of the vehicle may be malfunctioning, or an attacker may have created the malicious alert. Fortunately, we can leverage the assumptions that the majority of vehicles are not malicious, and that multiple vehicles observe the same event. Consequently, a vehicle can combine several information sources to corroborate the validity of the alert. In this paper, we present a basic framework, which we found to be useful when we applied it to misbehavior detection for EEBL messages. We anticipate that other researchers can leverage our model to address misbehavior detection in other applications.

8. REFERENCES

[1] F. Bai, T. Elbatt, G. Hollan, H. Krishnan, and V. Sadekar. Towards Characterizing and Classifying Communication-based Automotive Applications from a Wireless Networking Perspective. In *IEEE AutoNet*, 2006.

[2] F. Bai and B. Krishnamachari. Spatio-temporal Variations of Vehicle Traffic in VANETs: Facts and Implications. In *ACM VANET*, 2009.

[3] Cambridge Systematics, Inc. Crashes vs. Congestion: What's the cost to society? http://www.aaanewsroom.net/Assets/Files/20083591910.CrashesVsCongestionFullReport2.28.08.pdf.

[4] K. Chang and K. Chon. A Car-Following Model Applied Reaction Times Distribution and Perceptual Threshold. *Journal of the Eastern Asia Society for Transportation Studies*, 6:1888–1903, 2005.

[5] L. Eschenauer, V. D. Gligor, and J. Baras. On Trust Establishment in Mobile Ad-Hoc Networks. In *Proceedings of the Security Protocols Workshop*, 2002.

[6] M. Ghosh, A. Varghese, A. Gupta, and A. A. Kherani. Distributed misbehavior detection in VANETs. In *WCNC*, 2009.

[7] M. Ghosh, A. Varghese, A. Gupta, A. A. Kherani, and S. N. Muthaiah. Detecting Misbehaviors in VANET With Integrated Root-Cause Analysis. *Ad Hoc Networks*, 2010.

[8] P. Golle, D. Greene, and J. Staddon. Detecting and Correcting Malicious Data in VANETs. In *ACM VANET*, 2004.

[9] Y.-C. Hu and K. P. Laberteaux. Strong VANET Security on a Budget. In *ESCAR*, 2006.

[10] IBM. IBM 4764 PCI-X cryptographic coprocessor. http://www-03.ibm.com/security/cryptocards/pcixcc/order4764.shtml.

[11] IEEE. 1609.2: Trial-use standard for wireless access in vehicular environments-security services for applications and management messages. IEEE Standards, 2006.

[12] National Highway Traffic Safety Administration (NHTSA). Fatality Analysis Reporting System. http://www-fars.nhtsa.dot.gov/.

[13] M. Raya and J.-P. Hubaux. The security of vehicular ad hoc networks. In *ACM SASN*, 2005.

[14] M. Raya, P. P. Papadimitratos, V. Gligor, and J.-P. Hubaux. On Data-Centric Trust Establishment in Ephemeral Ad Hoc Networks. In *IEEE Infocom*, 2008.

[15] R. K. Schmidt, T. Leinmuller, E. Schoch, A. Held, and G. Schafer. Vehicle Behavior Analysis to Enhance Security in VANETs. In *V2VCOM*, 2008.

[16] A. Studer, F. Bai, B. Bellur, and A. Perrig. Flexible, Extensible, and Efficient VANET Authentication. In *ESCAR*, 2008.

[17] A. Studer, M. Luk, and A. Perrig. Efficient Mechanisms to Provide Convoy Member and Vehicle Sequence Authentication in VANETs. In *SecureComm*, 2007.

[18] A. Studer, E. Shi, F. Bai, and A. Perrig. TACKing Together Efficient Authentication, Revocation, and Privacy in VANETs. In *IEEE SECON*, 2009.

[19] Texas Transportation Institute. Urban Mobility Report. http://mobility.tamu.edu/ums/report/.

[20] G. Theodorakopoulos and J. S. Baras. On Trust Models and Trust Evaluation Metrics for Ad-Hoc Networks. *IEEE Journal on Selected Areas in Communications*, 24(2):318–328, 2006.

[21] Y. Zang, L. Stibor, H.-J. Reumerman, and H. Chen. Wireless Local Danger Warning Using Inter-vehicle Communications in Highway Scenarios. In *14th European Wireless Conference*, 2008.