Central Misbehavior Evaluation for VANETs based on Mobility Data Plausibility

Norbert Bißmeyer
Fraunhofer Institute for Secure Information Technology (SIT)
Mobile Networks
Darmstadt, Germany
norbert.bissmeyer@sit.fraunhofer.de

Joël Njeukam
Darmstadt University of Technology
Electrical Engineering and Information Technology
Darmstadt, Germany
jnjeukam@hrz.tu-darmstadt.de

Jonathan Petit
University of Twente
Distributed and Embedded Security Group
Twente, Netherlands
j.petit@utwente.nl

Kpatcha M. Bayarou
Fraunhofer Institute for Secure Information Technology (SIT)
Mobile Networks
Darmstadt, Germany
kpatcha.bayarou@sit.fraunhofer.de

ABSTRACT

Trustworthy communication in vehicular ad-hoc networks is essential to provide functional and reliable traffic safety and efficiency applications. A Sybil attacker that is simulating "ghost vehicles" on the road, by sending messages with faked position statements, must be detected and excluded permanently from the network. Based on misbehavior detection systems, running on vehicles and roadside units, a central evaluation scheme is proposed that aims to identify and exclude attackers from the network. The proposed algorithms of the central scheme are using trust and reputation information provided in misbehavior reports in order to guarantee long-term functionality of the network. A main aspect, the scalability, is given as misbehavior reports are created only if an incident is detected in the VANET. Therefore, the load of the proposed central system is not related to the total number of network nodes. A simulation study is conducted to show the effective and reliable detection of attacker nodes, assuming a majority of benign misbehavior reporters. Extensive simulations show that a few benign nodes (at least three witnesses) are enough to significantly decrease the fake node reputation and thus identify the cause of misbehavior. In case of colluding attackers, simulations show that if 37% of neighbor nodes cooperate, then an attack could be obfuscated.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General—*security and protection*

VANET'12, June 25, 2012, Low Wood Bay, Lake District, UK.
Copyright 2012 ACM 978-1-4503-1317-9/12/06 ...$10.00.

Keywords

C2X, C2C, VANET, IDS, misbehavior evaluation, intrusion detection, V2X, V2V, trust, confidence, reputation

1. INTRODUCTION

The communication between mobile nodes (e.g. vehicles) and infrastructure nodes (e.g. roadside units) in an Intelligent Transportation System (ITS) is primarily based on IEEE 802.11p wireless ad-hoc message transmission [11]. One of the main goals of ITS communication is the enhancement of traffic safety and efficiency. Vehicles and Roadside Units (RSU) are broadcasting messages with traffic related data. Due to the wireless property of the communication channel, the transmitted messages, used by ITS functions, have to be protected by cryptographic security solutions. As proposed by the IEEE 1609.2 draft standard [12], digital certificates are used to ensure the authentication and authorization of message senders. A Public Key Infrastructure (PKI) issues certificates, but only for authenticated nodes of a Vehicular Ad-hoc Network (VANET). Nevertheless, it is assumed that attacks on the VANET are possible as cryptographic keys could be misused or malicious software could be installed on some nodes. Also faulty nodes could disturb the functionality of ITS communication unintentionally but permanently. As a result, misbehavior detection and evaluation is necessary to keep overall functionality of ITS communications.

A local misbehavior detection system on the network nodes is able to detect inconsistencies in mobility data (i.e. absolute position, heading, speed) by applying plausibility checks. However, this detection is restricted to nodes that are in communication range of the attacker at current moment in time. Also, if local misbehavior detections are forwarded via multi-hop communication to neighbors, the long-term exclusion of attackers is not possible, as analyzed in more detail in Section 3.3. Furthermore, the local misbehavior evaluation suffers from pseudonym changes of attackers and therefore only a reduced set of information may be available. Hence, we propose the transmission of locally created misbehavior reports to a central evaluation entity. This permits to detect attacks, based on a larger set of information, and identifies attackers with higher probability. Our main concept is based on the computation of trust information regarding neighboring VANET nodes. Successful detections are used subsequently to exclude disturbing nodes from the VANET until they have proven their benignity. As great attention is given to scalability, flexibility and practicability, the proposed scheme aims to provide a basis for automated misbehavior detection in ITS.

Organisation: In Section 2, we present the related work for misbehavior detection in VANETs and trust management. Section 3 provides the system assumptions and Section 4 describes the proposed central misbehavior evaluation scheme. In order to show the feasibility of the model, Section 5 analyzes the performance and security issues and evaluates the scheme based on simulation. Section 6 concludes the paper and gives an outlook for future work.

2. RELATED WORK

Misbehavior detection is an important topic in VANETs and was studied in several publications. At first, related work regarding detection algorithms on network nodes is discussed and subsequently related proposals for trust management are analyzed.

2.1 Misbehavior Detection Approaches

Our central misbehavior evaluation scheme is based on detection algorithms applied on network agents. It is assumed that nodes are running mobility data plausibility checks as proposed by several authors [24, 13, 16, 7, 10]. In [21] the detection of attackers and their temporary local exclusion from communication is autonomously done on the network node. But, a permanently exclusion from the whole VANET can only be done by a central entity based on reported detections and a larger information base. Especially, if different spatially and temporarily distributed attacks are assumed, a reliable long-term exclusion seems not to be feasible by decentralized network nodes.

The authors in [28] and [6] propose misbehavior detection systems using roadside units that provide and forward special data and certificates in order to detect ghost vehicles simulated by an attacker. Assuming a VANET without a dense network of roadside units and no constant communication to back-end services, the protocols may not work reliable. Furthermore, communication of security-related data for misbehavior detection should not be transmitted between vehicles permanently, due to limited bandwidth. Though, a central misbehavior evaluation concept can concentrates on event based reports that contain information about observed inconsistencies in mobility data. In [4], a plausibility check detects overlaps between vehicles. This misbehavior detection schemes creates polygons for adjacent vehicle nodes based on their physical dimensions. As every vehicle occupies a certain space on the road, an overlap of two polygons triggers the generation of a misbehavior report. The design of central evaluation should be flexible so that different algorithms for misbehavior detection could be integrated (e.g. based on Received Signal Strength Indication [14] or directional antennas [22]).

2.2 Trust Management Approaches

In [29], a survey is given that analyzes most important approaches for trust management in VANETs. This survey identifies the following main requirements: decentralization, sparsity, dynamics, scalability, confidence, security, privacy and robustness. Most trust management approaches discussed in this survey use the reputation as basis for trustworthy wireless communication via IEEE 802.11p. Another use case for applying trust information can be the identification of attackers assuming a majority of benign nodes that give bad ratings for possible attackers. In [2], the authors propose a routing protocol based on node's reputation. In this approach, the behavior of the nodes is assessed by direct neighbors. While Bella et al. do the node assessment based on detecting selfish routing behavior, previously discussed misbehavior detection schemes can also be used to rate neighbor nodes' trustworthiness in a VANET. In order to handle node reputation for misbehavior detection, a two-value pair approach is used that separate *trust* and *certainty*. This strategy is well known and has been described in more detail, for example in [23]. Related methods for handling and aggregating such two-pair values are proposed by Ebinger in [8]. Mármol et al. presents in [17] a Trust and Reputation Infrastructure-based Proposal for vehicular ad-hoc networks (TRIP), which computes a reputation score based on recommendations and a self-estimated reputation.

3. SYSTEM MODEL

A VANET consists of nodes (i.e. vehicles and roadside units) that broadcast messages frequently. The used message format is called Cooperative Awareness Message (CAM) [9] or Basic Safety Message (BSM) [25]. Every CAM contains information about the vehicle (e.g. current absolute position, speed, heading, station dimensions) and a pseudonym identifier that temporarily identifies the sender.

In order to protect the network against external malicious packet insertions or packet modifications, messages are digitally signed using a private key on sender's side and the receiver verifies the signature with the related public key. The public key can be extracted from the pseudonym-certificate as it is appended to every CAM. In general, only verified messages should be used by VANET nodes. Due to privacy protection requirements vehicles frequently and unexpectedly change their identifier as well as their signing certificate with related keys. The identifiers used in the ITS communication are derived from short-term pseudonym-certificates and can only be linked to the node's respective long-term ID by the PKI that has issued the pseudonym certificates. As a precondition for central misbehavior evaluation, the PKI can be used to get linking information of pseudonyms in order to ignore multiple reports from the same sender announcing the same misbehavior event.

Although nodes are probably equipped with several hundred pseudonym certificates valid for the same time, it is assumed that every node stores the related private keys in a tamper resistant storage. Therefore, the keys cannot be read by arbitrary users and system functions. Furthermore, it is assumed that the position information added to CAMs is very accurate and exceeds the quality of general Global Navigation Satellite System (GNSS, e.g. GPS) by applying optimization measures. Dead reckoning, differential GPS or relative positioning [1, 18] are some possible solutions to increase the quality of node's location knowledge. Additionally, sensors such as cameras, Radar or Lidar can be used to increase the position accuracy in CAMs. All these optimizations are necessary to reliably detect *vehicle overlaps*, according to [4].

For the central evaluation of locally generated misbehavior reports, it is necessary to have a connection between the nodes of the VANET and the central entity. But, the availability of such connection is assumed to not be permanent. Therefore, nodes would not be able to send misbehavior reports over a longer time. As we aim at excluding attackers in order to support long-term network functionality, a low latency between local detection and central evaluation is not a primary goal.

3.1 Adversary Model

In this paper, we focus on internal attacker that can forge messages to generate ghost vehicles. This malicious behavior is also known as Sybil attack.

An attacker, that places for example a non-existing broken down vehicle on a road segment, would be able to reroute other vehicles if their navigation systems process the faked messages, or worse, affect safety applications of neighboring vehicles. Based on this kind of malicious mobility data modification in broadcasted messages, a wide range of different attacks is imaginable. An exemplary situation is depicted in Figure 1. The claimed position of a faked vehicle $a \in V$ is overlapping a real benign vehicle $b \in V$. Other witness nodes $w_1, w_2, ..., w_5 \in V$ in the communication range are able to detect this inconsistency autonomously [4]. In the remaining sections, we also assume a and b as suspected nodes.

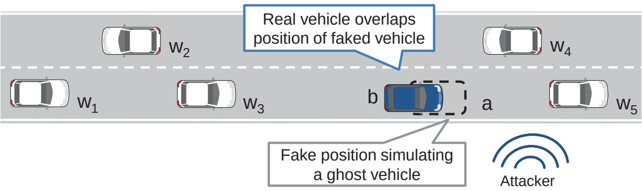
Real vehicle overlaps position of faked vehicle
W2
W4
W1
W3
b
a
W5
Fake position simulating a ghost vehicle
Attacker

Figure 1: Attacker simulates a faked ghost vehicle on the road that is overlapped by real vehicle positions

As this adversary model is relatively generic, we aim to detect a wide area of location based attacks by applying the proposed misbehavior detection and evaluation framework.

3.2 Local Misbehavior Detection

In order to detect the ghost vehicles, every node in the VANET is independently running a plausibility checker [24, 4, 13]. It is able to detect abnormal events such as overlaps, unexpected position jumps or suddenly appearing nodes. Our proposed central misbehavior evaluation system concentrates on the evaluation of virtual overlaps on the road as described in detail in [4] and depicted in Figure 1. Based on the traffic density, an attacker is not able to place ghost vehicles permanently onto a road without producing inconsistencies due to overlaps. A local misbehavior detection system is able to detect inconsistencies but has only restricted possibilities to identify the root cause. For example, in Figure 1 node a is overlapped by node b, which is detected by other nodes. But the identification of the originator or attacker node is problematic for the local system of neighboring nodes.

We can further show with the analysis in Section 3.3 that a central evaluation is able to work with a larger information basis. Hence, a central misbehavior evaluation is more promising for long-term exclusion of misbehaving nodes from the VANET than the use of local systems alone. However, local detection schemes are necessary to identify and prefilter obvious false positive detections.

3.3 Local vs. Central Misbehavior Evaluation

We propose to formally assess the number of inconsistencies that could be detected. First, we define the notation. A set of vehicles V is passing an area where a ghost vehicle $a \in V$ is simulated within the time frame $0, ..., n$. The attacker is able to change the pseudonym of the ghost vehicles arbitrarily. Therefore, node a may appear as $a', a'', a''', etc.$ The subset $COM_k \subset V$ denotes a set of vehicles that are within communication range of node a at time k where $k \in [0, n]$. A local misbehavior detection system, running on the observer node $obs \in V$, is able to detect all inconsistencies I_{obs} produced by the ghost vehicle a when another vehicle is overlapping a's occupied position. The maximum number of local detections is shown in Equation 1. Only if obs and a are both elements of COM_k and a uses the same pseudonym a' then detections can be linked to the same originator node. Different sizes of the subset COM_k are used in Equation 1 by introducing the variable i and m where $0 \leq i \leq m \leq n$. In general, every node that is an element of COM_k can produce an overlap with a's position but probably only a fraction is capable of doing so. Therefore, a variable $0 \leq \delta \leq 1$ is used.

$$I_{(obs,a')} = (\sum_{k=i}^{m} |COM_k|) \cdot \delta \tag{1}$$

The slower the observer obs is passing the communication range of a, the more inconsistencies can be detected in principle. If a is changing the pseudonym from a' to a'', then obs cannot link both detections. In contrast, a central misbehavior evaluation system is able to get the information that different pseudonyms (e.g. a', a'', a''') belong to the same node a. Furthermore, the central evaluation is able to consider inconsistencies from all time frames $0, ..., n$, whereby obs can only detect inconsistencies while being within communication range of a.

$$I_a = \sum_{obs}^{V} I_{(obs,a')} + \sum_{obs}^{V} I_{(obs,a'')} + \sum_{obs}^{V} I_{(obs,a''')}, obs \in V \tag{2}$$

The comparison of Equation 1 with Equation 2 shows that a central evaluation system is able to process a larger set of misbehavior detections than a local system: $I_a \geq I_{(obs,a')}$. Also if neighboring observers in communication range exchange misbehavior information, a central system is able to collect more reports due to possible pseudonym changes of the attacker: $I_a \geq \sum_{obs}^{COM_k} I_{(obs,a')}, obs \in COM_k$.

Another aspect is the false positive rate in the misbehavior evaluation. A local system, applied on a node, has to create decisions based on limited information (i.e. number of received messages), and within a restricted amount of time. Using Figure 1 as example, node w_1 and w_3 are approaching the attack scene and have to decide rapidly which of the nodes, a or b, can be trusted. A central misbehavior evaluation scheme in contrast is not forced to decide rapidly based on limited information in order to identity an attacker for long-term exclusion. Therefore, a central evaluation scheme can decrease the false positive rate by increasing the number of needed independent misbehavior reports regarding a suspected node.

4. CENTRAL EVALUATION SCHEME

We propose a Misbehavior Evaluation Authority (MEA) that is operated in the back-end infrastructure. It collects misbehavior reports from nodes that are directly involved in overlaps or have observed such an event as a witness. For the detection of ghost vehicles, the MEA uses trust information regarding nodes contained in misbehavior reports. The methods used in this paper, divide this trust information into two independent elements: *trust* and *confidence* (cf. also [23]). With this division, node assessment mechanisms can be used more efficiently as opinion and certainty about the trustworthiness of a node are separated.

Definition 1. Trust is modeled as the subjective probability that an entity behaves as expected. The *trust* that node $b \in V$ has regarding node $a \in V$ at time k is denoted as $t_{b,a}(k) \in \mathbb{R}$. Trust has values in the range $[-1, 1]$, where -1 denotes maximal distrust and 1 denotes maximal benignity. New nodes start with a balanced trust value of 0.

Definition 2. The *confidence* value is always related to an opinion (i.e. a trust value). According to [23], modeling the confidence of an opinion allows to provide information on how much evidence an opinion is based, or to state that there is no evidence available. Furthermore, it is possible to express, that one opinion might be supported by more evidence than another one. In our calculations, where opinions (i.e. trust values) are processed, the confidence value is used as weighting factor. A low confidence value means low consideration of related trust and high confidence means that the related trust can be used with high weight. The confidence value that node b assigns to the trust estimation of node a is denoted as $c_{b,a}(k) \in \mathbb{R}$ and has values in the range $[0, 1]$.

Definition 3. A *reputation* is used to express a combination of trust and confidence that node b assigns to node a. It is denoted as $r_{b,a}(k) \in \mathbb{R}$ with values in the range $[-1, 1]$.

4.1 Misbehavior Report

A Misbehavior Report (MR) is used to send information regarding possible misbehavior from network nodes to MEA. A report stores the type of detected misbehavior (e.g. node overlap, unexpected position jump), the pseudonymous ID of the reporter node, a list of overlapping nodes including their trust statements and a list of neighbor nodes surrounding the reporter. Figure 2 gives on overview of the misbehavior report structure and its content.

In every report an evidence of the observed event is added. In case of a position overlap, two signed CAMs are added that attest to the overlap of node polygons as detailed in [4]. The suspect nodes and relevant one-hop neighbor nodes are reported by providing their pseudonymous ID and a trust statement. The latter contains a trust value of the target that is calculated by the local misbehavior detection system of the reporter, as well as the contact duration and the traveled distance of this node. In order to attest to the values for distance and duration, an appropriate CAM is appended, and thus, can be verified by the MEA. Finally, the complete report is signed and encrypted before it is sent to the central MEA. In the case of dense traffic, the size of a misbehavior report could be limited by adding only relevant neighbors which has observed the misbehavior autonomously. Only selected one-hop neighbor nodes shall be added, prioritized on the distance between the node and the location of the detected inconsistency. The probability that nearby neighbors have also detected the inconsistency autonomously is higher than for distant neighbors.

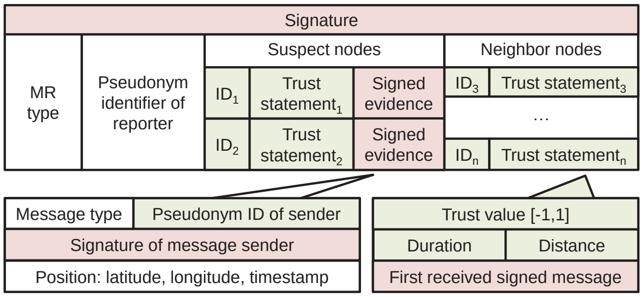

Figure 2: Structure of misbehavior report

Due to the signed evidence proving the overlap, an attacker cannot accuse arbitrary nodes in the network. Therefore, cooperative attacks with several malicious reporters are spatially and temporarily limited. Furthermore, the lightweight MR format allows the transmission to the infrastructure via roadside stations using IEEE 802.11p. Our implementation has shown that the size of a misbehavior report is approximately 1 KB and will be increased by 200 Bytes for every additional neighbor node. It has to be considered that the temporary MR storage on the network node should be persistent but has no requirements regarding security or tamper protection.

4.2 Certification of Misbehavior Reports

When receiving a new MR at the central MEA, the contained signatures are first verified by using the public keys of the related pseudonym certificates. As the MEA has a connection to the PKI that has issued the certificates, it is sufficient to store only short certificate IDs in the MR instead of the complete certificate structure. With the certificate ID, the MEA is able to request the appropriate certificate. In a second step, the evidence of overlaps is checked by verifying the signature of CAMs. The overlap scenario can be verified by comparing the position vectors of the appended messages as shown in Figure 2. Subsequently, all information of neighbor nodes that are appended to the MRs is verified by comparing the position vector of the signed CAM with given duration and distance values. If the MEA detects a noteworthy differences between claimed values for duration and distance and the CAM position, the report is discarded and will not be used in the further evaluation process. Furthermore, duplicated reports from the same node, using different pseudonymous IDs, are discarded. A possible infringement of privacy due to this required resolution of pseudonym ID to their related long-term ID is not further discussed in this paper. Nevertheless, the appropriate integration of misbehavior evaluation in a PKI solution is an important issue as identified in [29].

The verified MRs are stored in order to amass enough reports from nodes that are involved, or have observed an inconsistency, caused by node overlaps. Having enough reports for an evaluation, a *session* is created for every misbehavior scenario. This session contains a list of all involved nodes. The list of neighbors from the MRs is used to identify possible witnesses. Based on a policy, the number of needed witnesses can be defined before starting the further evaluation. It is therefore a requirement that every involved node should be able to detect an overlap autonomously, create a report and send it to the MEA.

4.3 Collection of Misbehavior Reports

Due to restrictions in communication range, shadowing effects or possibly missing communication links between nodes and the infrastructure, the central MEA may not be able to get all MRs from all nodes involved in a session. Additionally, attackers that try to blacklist benign network nodes by sending faked MRs without having an overlap scenario on the road, have to be considered. Therefore, the following considerations are checked before starting the evaluation of a session:

a) Either MRs from all overlapping nodes are received or the number of needed witness reports must be increased with respect to an upper bound. Assuming an overlap scenario where the MEA has received a report from a witness node w_1, stating that node a overlaps node b at time k, it is necessary that respective reports from node a and b are also received concerning the same overlap at time k, where $a, b, w_1 \in V$. This scheme should avoid blacklisting of benign nodes.

b) If an attacker is placing a ghost vehicle on the road, that is overlapped by real vehicles, then it cannot be assumed necessarily that a misbehavior report is sent by the attacker. In this case, the needed number of reports from witnesses has to be increased. As discussed in Section 3, colluding attackers have to spatially and temporarily synchronize. Therefore, the effort for colluding attacks increases with every cooperating malicious node. Finally, different pseudonyms used by a attacker in the communication can be identified and linked.

c) Misbehavior Reports have to be confirmed by witnesses. A received Misbehavior Report, stating an overlapping of node a and b at time k, has to be confirmed by witness nodes w_i with $i = 1, ..., N$. Determining the value of N is further discussed in Section 5 and has been addressed by related work [20].

4.4 Node Assessment Concept

The goal of the central misbehavior evaluation is the identification of an attacker inside a set of nodes that are actively or passively involved in an overlap scenario. The evaluation process is divided into five steps, as depicted in Figure 3.

4.4.1 Generation of Cooperative Trust-Confidence

As soon as enough reports are collected, the assessment process is started for the respective session. In order to use the reported information for detecting and identifying an attacker, the trustworthiness of all involved nodes is calculated. Figure 3 provides an overview of the process to calculate trust-confidence pairs that are reported by network nodes. In the first step shown in this figure, the confidence of reported trust is calculated using Equation 3. Node a provides the contact time with node b in form of $duration(a, b)$ and $distance(a, b)$. The variables γ and β determine the minimum value for common contact time and commonly driven distance to get a maximum confidence value. Example values for γ and β can be found in Section 5.4.

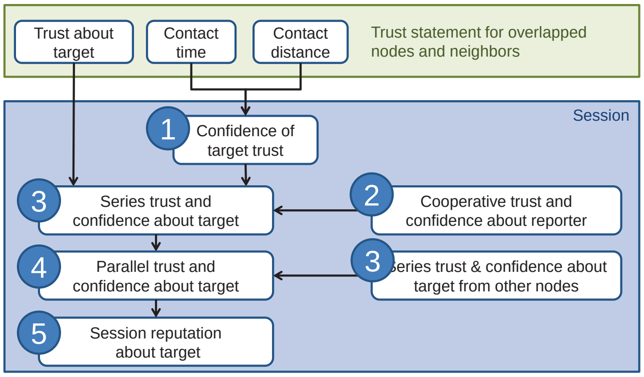
Trust about target
Contact time
Contact distance
Trust statement for overlapped nodes and neighbors
Session
1
Confidence of target trust
2
Cooperative trust and confidence about reporter
3
Series trust and confidence about target
3
Series trust & confidence about target from other nodes
4
Parallel trust and confidence about target
5
Session reputation about target

Figure 3: Overview of central processing of reputation information

$$c_{a,b} = min(1, \frac{\frac{1}{\gamma} \cdot duration(a,b) + \frac{1}{\beta} \cdot distance(a,b)}{2}) \quad (3)$$

The given trust and calculated confidence from Misbehavior Reports are used to express the trustworthiness regarding reported misbehaving nodes or witnesses. After extraction of the trust-confidence pairs for all session nodes, a cooperative trust and confidence value is calculated. This tuple determines the trustworthiness of every involved node as shown in step 2 of Figure 3. In the following process description, the set V contains all session nodes.

Cooperative trust $tc_e(k)$ and cooperative confidence $cc_e(k)$ for node $e \in V$ at time k are calculated where node e is evaluated by all other session nodes $a \in V$ and $a \neq e$.

$$tc_e(k) = \frac{\sum_{a,a\neq e}^{V} t_{a,e}(k) \cdot c_{a,e}(k)}{\sum_{a,a\neq e}^{V} c_{a,e}(k)} \quad (4)$$

In Equation 4, trust regarding all nodes in a session derived from all other nodes of the session, are combined with respective confidence as a weighting factor. In the formula, every trust value is multiplied by the associated confidence value and the sum of these values is divided by the sum of the confidence values of each session node.

Equation 5 shows the cooperative confidence of a node e, calculated from a combination of values from all other session nodes $a, b \in V$.

$$cc_e(k) = 1 - \frac{\sum_{a,b\in V,a\neq b}^{V} |t_{a,e}(k) - t_{b,e}(k)|}{2 \cdot |V| \cdot (|V| - 1)} \quad (5)$$

$$cc_e(k) = cc_e(k) \cdot \sum_{a,a\neq e}^{V} c_{a,e}(k) \quad (6)$$

This calculation ensures that the confidence value is high if the different nodes agree on similar trust levels (i.e. the gap between trust values is small) and the reverse, if the opinions differ a lot (i.e. trust value differentials are high). The fraction shown in Equation 5 expresses the mean value of the differences between all trust values provided by session nodes that have an opinion about node e. If the sum of confidence values regarding e is larger than 1, then the maximum value for $cc_e(k)$ is 1.

4.4.2 Assessment of Suspected Nodes

Having this aggregated cooperative trust-confidence pair of every node in a session, the evaluation of suspected nodes is started in the third step of Figure 3. All nodes that are reported to be overlapped by another node, are suspects. The trust regarding a suspect, provided by a witness node, is weighted with the cooperative trust-confidence pair of this witness node, by using the results from Equation 4 and 5. The computation of trust-confidence data in Equation 7 and 8 is denoted as *series* combination as the trustworthiness of a reporter node a is used to rate a target node e.

The function $ts_{ae}(tc_a(k), cc_a(k), t_{a,e}(k))$ with nodes $a, e \in V$ is used to calculate this series trust regarding node e using the cooperative trust and cooperative confidence about a at time k where $0 \leq ts_{ae}(k) \leq 1$. If $t_{a,b}(k) \geq 0$, then Equation 7 is used. Otherwise $ts_{ae}(k) = 0$.

$$ts_{ae}(k) = tc_a(k) \cdot cc_a(k) \cdot t_{a,e}(k) \tag{7}$$

The function $cs_{ae}(tc_a(k), cc_a(k), c_{a,e}(k))$ shown in Equation 8 with nodes $a, e \in V$ is used to calculate the series confidence of node e by using the cooperative values for trust and confidence of node a. The resulting confidence $cs_{ae}(k)$ is lower or equal than the confidence $c_{a,e}(k)$.

$$cs_{ae}(k) = cc_a(k) \cdot tc_a(k) \cdot c_{a,e}(k) \tag{8}$$

By using Equation 7 and 8, a cooperatively less trusted node has a lower impact in the assessment of suspects than a node with high cooperative trust. This assessment task is repeated for every suspected node, using the cooperative trust and confidence of every node involved in a session. In step 4 of Figure 3, all series trust-confidence pairs are combined in parallel in order to get the final values for trust and confidence of a suspect. This parallel combination uses results from Equation 7 and 8 to summarize the trustworthiness regarding a suspect from the different reporter view points.

The parallel trust $tp_e(k)$ and parallel confidence $cp_e(k)$, for node e at time k, are calculated using the results of the series combination of session nodes $a \in V$ where $a \neq e$.

$$tp_e(k) = \frac{\sum_{a,a\neq e}^{V} ts_{ae}(k) \cdot cs_{ae}(k)}{\sum_{a,a\neq e}^{V} cs_{ae}(k)} \tag{9}$$

In Equation 9, different series trust values considering the same suspect are combined with the respective confidence as a weighting factor. In the formula, every series-trust is multiplied by the associated series confidence value and the sum of these values is divided by the sum of the series confidence of each session node. This formula provides equally balanced results between session nodes, and therefore, no further weights are necessary.

Equation 10 shows the confidence of a node e, calculated from a combination of values from all session nodes. This calculation ensures that the confidence value is high if the different nodes agree on similar trust levels (i.e. the gap between trust values is small) and the opposite holds, if the opinions differ a lot (i.e. trust value differentials are high).

$$\begin{aligned} cp_e(k) &= 1 - \frac{\sum_{a,b\in V, a\neq b\neq e}^{V} |ts_{ae}(k) - ts_{be}(k)|}{2 \cdot |V| \cdot (|V|-1)} && (10) \\ cp_e(k) &= cp_e(k) \cdot \sum_{a,a\neq e}^{V} cs_{ae}(k) && (11) \end{aligned}$$

The fraction expresses the mean value of the differences between all series values and the cardinality $|V|$ defines the number of different session nodes a that have evaluated node e. If the sum of confidence values $\sum_a^V cs_{ae}(k)$ is larger than 1 then the maximum value $cp_e(k) = 1$ is given.

4.4.3 Identification of Attackers

In the last step, it is checked whether the suspected nodes are rated positively or negatively. Therefore, the results of the parallel combination of trust and confidence from Equation 9 and 10 are used in Equation 12. The higher the parallel confidence value $cp_e(k)$ the more the trust value $tp_e(k)$ is considered. If a reported trust of a suspicious node has low confidence or the reporter itself has a low confidence then a neutral reputation for the suspicious node $e \in V$ is given.

$$r_e(k) = tp_e(k) \cdot cp_e(k) \tag{12}$$

Based on the concept, presented in Figure 3, the reputation for suspicious nodes is combined in the fifth step so that a single reputation value $r_e(k)$ is generated. This reputation value contains aggregated information from all other involved reporters in a session. Based on a benign majority and local misbehavior detection mechanisms on the reporter nodes, a ghost vehicle is rated with a negative trust value and a real vehicle is rated with a positive trust value.

4.4.4 Example for Node Assessment

In this example, the adversary scenario presented in Section 3.1 is used to discuss the node assessment process. Assuming $P \subset V$ denotes the subset of nodes that contains all nodes that are participating in an overlap scenario (i.e. reporter, overlapped node or witness). According to Figure 1, $a, b, w_1, ..., w_5 \in P$. Every node $p \in P$ detects the overlap of a and b at time k autonomously and sends a report to the MEA as soon as a connection to the infrastructure is available. The report of every node contains the two overlapping nodes and the remaining nodes $w_1, ..., w_5$ as witnesses in the list of neighbors.

The MEA collects the reports, creates a session with all elements of P and checks with Equation 3 whether the reported confidence values in the MR are correct. At this stage, a trust $t_{a,b}(k)$ and confidence value $c_{a,b}(k)$ for every combination of $a, b \in P, a \neq b$ exists. Using Equation 4 and 5, a cooperative value for every session node $p \in P$ in generated. As a result, every session node is evaluated with a pair of cooperative trust and cooperative confidence. In order to identify whether node a or node b is the attacker, only these two nodes are evaluated in the following steps. Using the series combination of trust and confidence (see Equation 7 and 8) regarding the suspects a and b, two tuples $(ts_{pa}(k), cs_{pa}(k))$ and $(ts_{pb}(k), cs_{pb}(k))$ for every $p \in P, p \neq a \neq b$ are generated. Subsequently, the single opinions about the suspects are combined with Equation 9 and 10. The resulting two tuples $(tp_a(k), cp_a(k))$ and $(tp_b(k), cp_b(k))$ for the suspects $a, b \in P$ can be used in the last step to generate the two reputation values $r_a(k)$ and $r_b(k)$. Depending on policies and defined thresholds, the MEA can decide whether node a or b is a ghost vehicle. According to the discussion in Section 4.3, the MEA can also store the reputation results for a and b from this session and wait until further inconsistencies regarding these nodes are reported.

5. ANALYSIS AND EVALUATION

The following evaluation verifies the functionality and quality of the proposed misbehavior evaluation concept described in Section 4. After a performance and security analysis, we present a simulation using different scenarios based on the adversary model described in Section 3.1.

5.1 Performance Analysis of Central MEA

In general, the scalability of a central entity has to be considered with special attention as we are talking about several hundred million vehicles in a VANET. Therefore, we stress that the number of processed misbehavior reports is not related to the number of nodes in the network. For example, the network consists of several million vehicles but only a handful of attackers are producing inconsistencies on the road that are detected by a handful of vehicles passing this area. At first, the network nodes can filter the inconsistencies where the local intrusion detection system assigns low confidence to. Only reliable detections are sent to the MEA. Secondly, the impact of a ghost vehicle attack is spatially restricted and therefore, only a very small subset of network nodes are able to send related misbehavior reports. Reports are created only if an overlap is autonomously detected as described in Section 3.2. In contrast to other related schemes [19, 5], the permanent report of node position and their system state is not needed. Indeed, by avoiding the generation of a "good behavior report", we reduce the dimensions of the infrastructure entities and make a constant communication link to the infrastructure unnecessary. The dimensions of the MEA is therefore related only to the number of mounted attacks plus false positive detections. As a result, we can state that the proposed central misbehavior evaluation scales well in the context of ITS communication.

5.2 Security Analysis of Central MEA

The goal of the MEA is to identify attackers in misbehavior scenarios as depicted by Figure 1 of Section 3.1. Nevertheless, additional vulnerabilities of the VANET should be avoided due to the introduction of a central misbehavior evaluation strategy. Therefore we analyze the most important security aspects regarding our proposed scheme.

The Denial of Service (DoS) is a well known attack to centralized services. As every MR is signed with the sender's pseudonym, the MEA checks in the first step the validity of the sender by verifying its pseudonym certificate and the message signature. Reports with an invalid signature are discarded after reception as described in Section 4.2. This strategy ensures that a DoS attacker must invest in cryptographic signing operations in order to flood the MEA with invalid reports. However, the verification of incoming reports should be equipped with enough processing power to be able to process a large number of incoming messages.

Another well known attack is the replay of messages or reports in case of central misbehavior evaluation. As shown in Figure 2 of Section 4.1, the observed misbehavior is described by signed messages containing position information and a timestamp. The combination of position and time allows the MEA to assign the report to a misbehavior session. Duplicates and replayed reports are detected and discarded. It has to be considered for both, the DoS attack and the replay attack, that the MEA is able to check whether different pseudonyms belong to the same node. Reports from the same node using different pseudonyms are also discarded.

Blackmailing is another attack that is considered in our scheme. The arbitrary generation of faked misbehavior reports is limited as previously discussed in Section 4.1. According to the type of observed misbehavior, the reporter has to attest the event by adding appropriate signed messages that cannot be faked by an attacker. Therefore, attackers are not able to blacklist nodes of the VANET arbitrarily.

5.3 Evaluation with Simulation

The MEA is implemented in Java, using a simulation tool that is able to create consistent misbehavior reports for different overlapping scenarios. The reference adversary scenario, depicted in Figure 1, can be used with variable number of benign and faked nodes. Further, the number of overlapping nodes can be assigned as well as the number of passive witnesses. In the following simulations, trust-confidence pairs in MRs are selected randomly and the evaluations are repeated 10 times. This is done in order to get a statistically sound value as well as the maximum and minimum of node assessments.

5.4 Configuration of Simulation

The trust statement in misbehavior reports is the primary information basis used for ghost vehicle detection. Therefore, given trust and confidence values have to be set appropriately by the simulator. The trust value is set by the local misbehavior detection system running on the nodes and the confidence value is set by processing distance and duration values. Full confidence should only be given if the reporter of a MR has traveled together with the evaluated target node a distance larger than the normal communication range. Using this approach, road side attackers may not be able to create fake MRs with high confidence values for involved benign nodes. The general concept to limit the power of static roadside attackers is proposed and discussed by Schmidt et.al in [26] and [24]. Furthermore, detecting mobile attackers is easier as faking a mobile ghost vehicle over time is more difficult (especially in dense traffic scenario).

5.4.1 Assign Appropriate Confidence

In the following simulation, a linear increase of confidence is calculated depending on distance and duration as shown in Equation 3 where $\gamma = 40$ seconds and $\beta = 1000$ meters. Maximum confidence is therefore given if nodes know each other for more than 40 seconds and traveled together more than 1000 meters. These exemplary values are selected for the simulation as we assume a maximal communication range of 1000 meters for IEEE 802.11p. Furthermore, an attacker is not able to fake the distance and duration as a signed message is used to attest the information as described in Section 4.2.

5.4.2 Assign Appropriate Trust

In order to estimate a useful value for trust in the MRs, a specific situation was simulated where a benign node overlaps a faked node. This overlap is detected by the benign node and the attacker where both entities send respective misbehavior reports. Furthermore, an increasing number of benign witness nodes are added to the scene that observe the overlap independently and report accordingly. Figure 4 shows the result of this simulation.

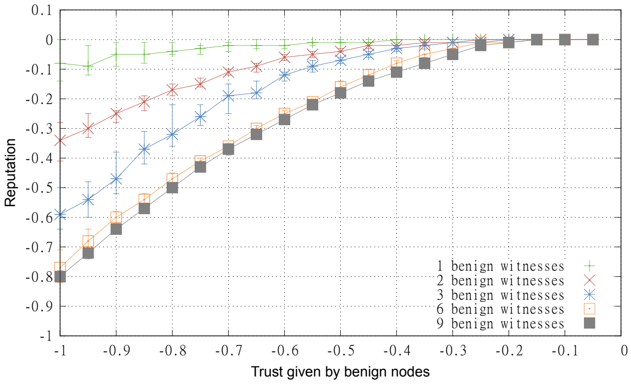

Figure 4: A ghost vehicle attack is detected by several benign nodes that provide increasing trust values

On the x-axis the provided trust value added to the MRs regarding the faked node is shown and the y-axis shows the calculated reputation of the ghost vehicle. In this reference test, all nodes are able to provide high confidence values in the upper bound (i.e. in the range [0.7, 1]) by setting appropriate duration and distance values in the MRs. Trust values are given according to the x-axis in the range [-1, 0]. As shown in Figure 4, decisions can be more reliable inside the MEA with an increasing number of benign witnesses and given trust values $t_{b,a} <= -0.5$.

The trust and confidence values between benign nodes and faked nodes are allocated in the following simulations, as shown in Table 1. Begnin nodes are providing positive trust values to other benign nodes and negative values for detected attackers. In contrast, attackers allocate every time a maximal positive trust value to other attackers and maximal negative values for benign nodes. As confidence depends on travel distance and duration of two nodes, the value cannot be faked arbitrarily by an attacker.

Table 1: Configuration of MEA for simulations

Direction of rating (provider → target)	Trust as range	Confidence as range
benign node → benign node	[0.5, 1]	[0, 1]
benign node → faked node	[-1, 0]	[0, 1]
faked node → benign node	-1.0	[0.1, 0.7]
faked node → faked node	1.0	[0.1, 0.7]

5.5 Quality of Malicious Node Detection

Based on the previously discussed simulation configuration, different quality measurements of the central MEA concept are presented in the following. In order to work with realistic simulation data, the evaluations presented in this Section are working with incomplete sets of Misbehavior Reports. It is assumed that 30% of all nodes are not able to transmit a recorded Misbehavior Report to the infrastructure. Therefore, this fraction of neighbors is not listed in Misbehavior Reports used in the simulation. In Figure 5, a situation is evaluated with one faked node that is overlapped by one benign node. This attack is also illustrated and described in Section 3.1. In order to get the information on “how many witness nodes are needed for reliable detection of an attacker”, the number of benign witnesses is increasing on the x-axis. The three graphs, shown in Figure 5, describe the reputation of the benign and the faked node as well as the mean reputation of all witness nodes. As shown in this graph, the decrease of faked nodes’ reputation attenuates with 7 witness nodes. In Figure 6, a similar situation is simulated. A fixed number of 10 benign nodes are generating misbehavior reports where 5 nodes are overlapping actively the one faked ghost vehicle. The other 5 benign nodes are acting as witnesses. In this simulation the number of malicious reporters is increased in order to measure the impact of several cooperating attackers. Using realistic configurations as described in Section 5.4, it is sufficient if 37% of the participants are cooperating attackers in order to hide a real attack on the road as shown in Figure 6. However, the effort for an attacker is very high to mount an attack where several manipulated vehicles are at the same location at a given time. Furthermore, the MEA is able to link different pseudonyms that are used by the same node.

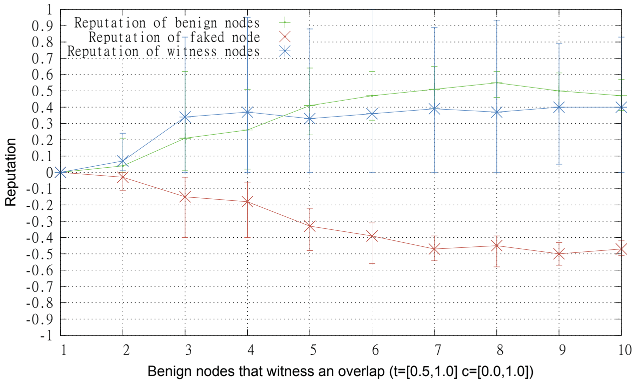

Figure 5: Ghost vehicle attack with increasing number of benign witnesses

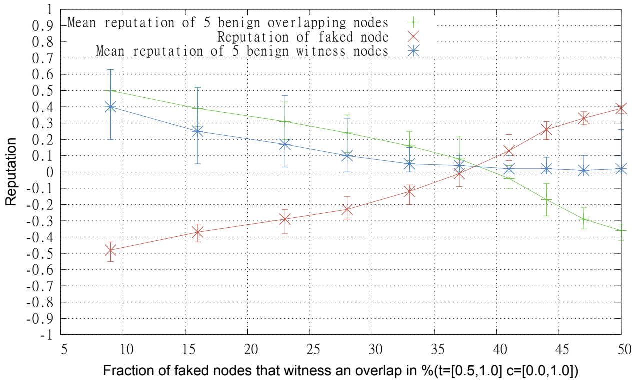

Figure 6: Ghost vehicle attack with increasing number of faked witnesses

5.6 Consolidated Findings

Based on the results, shown in Figure 4 and Figure 5, the number of needed MR reporters is dependent on their provided trust and confidence regarding a ghost node. This is also consistent to the assumptions made in Section 4.3 regarding the processing of incoming MRs at the MEA. If both involved nodes of an overlap scenario has transmitted misbehavior reports, then the number of needed witnesses may be lower. Otherwise, a higher number of benign witnesses is needed which would automatically increases the effort for a cooperative attack (see Figure 6). The proposed concept for central Misbehavior Report evaluation is used to decide which node is probably a faked ghost vehicle created by an attacker and which is a benign node. It has to be considered, that false positive detections where two benign nodes overlap each other virtually on the road, are not detected by this entity. Therefore, it is proposed to collect the final reputation of suspicious nodes in a global reputation table. Only if the same node has several overlaps with different other nodes, in a specific time frame, a reaction may be reasonable to protect the network against malicious or defective nodes.

6. CONCLUSION AND OUTLOOK

The positioning of ghost vehicles on the road by broadcasting messages with faked position information is one of the most critical attacks on VANETs from today's perspective [15]. With misbehavior detection mechanisms on all network nodes, the existence of ghost vehicles can be detected by finding position overlaps between network nodes. Nevertheless, a decision about which node is an attacker, can only be reliably decided at a central place. Based on Misbehavior Reports sent by network nodes, a central Misbehavior Evaluation Authority is able to reliably detect an attacker assuming a majority of benign reporters. Furthermore, the proposed design for Misbehavior Reports avoid cooperative attacks where fake reports are sent in order to blame benign nodes arbitrarily. Additionally, cooperative attacks entail high effort as they need a spatial and temporal reference. The efficiency and scalability of the proposed mechanism is supported by small internal report structure. Only if a misbehavior event is detected autonomously by a network node, a report is created and sent to the infrastructure as soon as an access point is available. Additional message transmission between network nodes via ad-hoc communication link is not necessary. Also a permanent transmission of position reports to the central evaluation entity is not required. Both aspects increases the system reliability and efficiency.

In future work, the central Misbehavior Evaluation Authority should be integrated into the proposed PKI concepts with consideration for privacy. As the resolution of communication pseudonyms is a basic requirement for misbehavior detection, a well-considered integration is necessary in order to preserve driver's privacy. Furthermore, additional misbehavior detection algorithms, as discussed in [24] and [13], could be integrated in order to decrease the spectrum of possible mobility data based attacks in VANETs. The adversary model presented in Section 3.1 is obviously not able to detect specific ghost vehicles, e.g. faked broken down vehicles on the shoulder lane. Finally, the security system architecture should be extended using results of different field operational tests (i.e. PRESERVE [3] and sim$^{\mathrm{TD}}$ [27]). With recorded measurements from FOTs, the quality of the presented central misbehavior evaluation approach could be shown related to realistic application scenarios.

7. ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union's Seventh Framework Programme project PRESERVE under grant agreement n°269994.

8. REFERENCES

[1] C. Basnayake, G. Lachapelle, and J. Bancroft. Relative positioning for vehicle-to-vehicle communication-enabled vehicle safety applications. In *ITS World Congress*. ITS America, October 2011.

[2] G. Bella, G. Costantino, and S. Riccobene. Managing reputation over manets. In *Information Assurance and Security (ISIAS)*, September 2008.

[3] N. Bißmeyer, J. Petit, D. Estor, M. Sall, J. P. Stotz, M. Feiri, R. Moalla, and S. Dietzel. PRESERVE d1.2 v2x security architecture. Deliverable, PRESERVE consortium, November 2011.

[4] N. Bißmeyer, C. Stresing, and K. Bayarou. Intrusion detection in vanets through verification of vehicle movement data. In *IEEE Vehicular Networking Conference (VNC)*, December 2010.

[5] Z. Cao, J. Kong, U. Lee, M. Gerla, and Z. Chen. Proof-of-relevance: Filtering false data via authentic consensus in vehicle ad-hoc networks. In *INFOCOM Workshops 2008*. IEEE, April 2008.

[6] C. Chen, X. Wang, W. Han, and B. Zang. A robust detection of the sybil attack in urban vanets. In *Distributed Computing Systems Workshops, 2009. ICDCS Workshops '09*, pages 270 –276, June 2009.

[7] G. D. Crescenzo, Y. Ling, S. Pietrowicz, and T. Zhang. Non-interactive malicious behavior detection in vehicular networks. In *IEEE Vehicular Networking Conference (VNC)*, December 2010.

[8] P. Ebinger and N. Bißmeyer. Terec: Trust evaluation and reputation exchange for cooperative intrusion detection in manets. In *Proceedings of the 2009 Seventh Annual Communication Networks and Services Research Conference*, CNSR '09, pages 378–385, May 2009.

[9] ETSI - European Telecommunications Standards Institute. Intelligent transport systems (ITS); vehicular communications; basic set of applications; part 2: Specification of cooperative awareness basic service. Technical Standard TS 102 637-2, ETSI, April 2010.

[10] M. Ghosh, A. Varghese, A. Kherani, and A. Gupta. Distributed misbehavior detection in vanets. In *Wireless Communications and Networking Conference (WCNC)*. IEEE, April 2009.

[11] IEEE Computer Society. IEEE standard for information technology – telecommunications and information exchange between systems – local and metropolitan area networks – specific requirements – Part II: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. Technical report, IEEE Std 802.11p, 2010.

[12] IEEE Computer Society. Draft standard for wireless access in vehicular environments - security services for applications and management messages. Technical Report 1609.2 - 2011 (D9), Institute of Electrical and Electronics Engineers, May 2011.

[13] A. Jaeger, N. Bißmeyer, H. Stübing, and S. A. Huss. A novel framework for efficient mobility data verification in vehicular ad-hoc networks. *International Journal of ITS Research, ITS Japan*, 9(3), September 2011.

[14] C. Laurendeau and M. Barbeau. Probabilistic localization and tracking of malicious insiders using hyperbolic position bounding in vehicular networks. *EURASIP J. Wirel. Commun. Netw.*, 2009:2:1–2:13, February 2009.

[15] T. Leinmüller, R. K. Schmidt, E. Schoch, A. Held, and G. Schäfer. Modeling roadside attacker behavior in vanets. In *GLOBECOM Workshops, 2008 IEEE*, pages 1 –10, December 2008.

[16] T. Leinmüller, E. Schoch, and F. Kargl. Position verficiation approaches for vehicular ad hoc networks. *Wireless Communications, IEEE*, 13(5):16 –21, October 2006.

[17] F. G. Mármol and G. M. Pérez. Trip, a trust and reputation infrastructure-based proposal for vehicular ad hoc networks. *Journal of Network and Computer Applications*, 35(3):934 – 941, May 2012. Special Issue on Trusted Computing and Communications.

[18] M. Obst, R. Schubert, and N. Mattern. Gnss-based relative localization for urban transport applications within the covel project. In *ITS World Congress*, Orlando, USA, October 2011. ITS America.

[19] B. Ostermaier, F. Dötzer, and M. Strassberger. Enhancing the security of local dangerwarnings in vanets - a simulative analysis of voting schemes. In *Availability, Reliability and Security (ARES)*, pages 422 –431, April 2007.

[20] J. Petit, M. Feiri, and F. Kargl. Spoofed data detection in vanets using dynamic thresholds. In *IEEE Vehicular Networking Conference (VNC)*, November 2011.

[21] M. Raya, P. Papadimitratos, I. Aad, D. Jungels, and J.-P. Hubaux. Eviction of misbehaving and faulty nodes in vehicular networks. *Selected Areas in Communications, IEEE Journal*, 25(8):1557 –1568, October 2007.

[22] Z. Ren, W. Li, Q. Yang, S. Wu, and L. Chen. Location security in geographic ad hoc routing for vanets. In *Ultra Modern Telecommunications Workshops (ICUMT)*, pages 1 –6, October 2009.

[23] S. Ries. Certain trust: a trust model for users and agents. In *Proceedings of the 2007 ACM symposium on Applied computing*, SAC '07, pages 1599–1604, March 2007.

[24] R. K. Robert K. Schmidt, T. Leinmüller, E. Schoch, A. Held, and G. Schäfer. Vehicle behavior analysis to enhance security in vanets. In *Proceedings of the 4th IEEE Vehicle-to-Vehicle Communications Workshop (V2VCOM)*, June 2008.

[25] SAE International TM. Surface vehicle standard - dedicated short range communications (DSRC) message set dictionary. Technical report, SAE J2735, November 2009.

[26] R. Schmidt, T. Leinmüller, and A. Held. Defending against roadside attackers. In *16th World Congress on Intelligent Transport Systems*, Stockholm, Sweden, September 2009.

[27] C. Weiß. Safe and intelligent mobility test field germany; deliverable d21.5; specification of IT security solution. Technical report, Fraunhofer-Institut SIT, Darmstadt, Germany, October 2009.

[28] B. Xiao, B. Yu, and C. Gao. Detection and localization of sybil nodes in vanets. In *Proceedings of the 2006 workshop on Dependability issues in wireless ad hoc networks and sensor networks*, DIWANS '06, pages 1–8, New York, NY, USA, September 2006. ACM.

[29] J. Zhang. A survey on trust management for vanets. In *Advanced Information Networking and Applications (AINA)*, pages 105 –112, March 2011.