Journal of Trust Management
a SpringerOpen Journal

RESEARCH Open Access

Detecting Sybil attacks in vehicular networks

Muhammad Al-Mutaz, Levi Malott* and Sriram Chellappan

*Correspondence: lmnn3@mst.edu
Department of Computer Science, Missouri University of Science and Technology, Rolla, MO (65401), USA

Abstract

A Sybil attack consists of an adversary assuming multiple identities to defeat the trust of an existing reputation system. When Sybil attacks are launched in vehicular networks, the mobility of vehicles increases the difficulty of identifying the malicious vehicle location. In this paper, a novel protocol for Sybil detection in vehicular networks is presented. Considering that vehicular networks are cyber-physical systems, the technique exploits well grounded results in the physical (i.e., transportation) domain to detect the Sybil attacks in the cyber domain. Compared to existing works that rely on additional cyber hardware support, or complex cryptographic primitives for Sybil detection, the protocol leverages the theory of *platoon dispersion* that models the physics of naturally occurring vehicle dispersion. Specifically, the proposed technique employs a certain number of roadside units that periodically collect reports from vehicles regarding their physical neighborhood. Leveraging from existing models of platoon dispersion, a protocol was designed to detect anomalously *close* neighborhoods that are reflective of Sybil attacks. To the best of the authors' knowledge, this paper is unique in integrating a well established theory in transportation engineering for detecting cyber space attacks in vehicular networks. The resulting protocol is simple, efficient, and robust in diverse attack environments.

Keywords: Vehicle-to-vehicle; Vehicle-to-infrastructure; Platoon dispersion; Sybil; Detection

Background and literature review

Introduction

Organizations in many countries today are investing in vehicular networks to leverage wireless networking support to improve state-of-the-art in road transportation. The US Federal Communications Commission (FCC) has allocated 75 MHz of spectrum in the 5.9 GHz band for Dedicated Short Range Communications, a set of protocols and standards for short to medium-range wireless communication for automotive use. Some recent vehicular networking efforts are the USDOT's Vehicle Infrastructure Integration (VII), which is a cooperative initiative between USDOT and automobile manufacturers, focusing on feasibility of deploying communications systems for safety and efficiency of road transportation systems. The ERTICO partnership is a multi-sector partnership pursuing development and deployment of Intelligent Transport Systems across Europe. Apart from these efforts, a variety of VANET test-beds have been set up in academia also for basic research and development of services.

This paper addresses a critical and emerging security problem in vehicular networks, namely detecting the presence of Sybil attacks. Sybil attacks are classified as an attack on

the trust of a peer-to-peer system by an attacker assuming many pseudonymous identities. Using these identities, the attacker can gain a disproportionately large influence on system functionality. In vehicular networks, the presence of a Sybil attack can have negative consequences. For instance, in an application like road safety, consider a single malicious vehicle, V_M, assuming a large number of fake identities incorrectly reporting road conditions. Other benign vehicles will tend to believe such a message, since it appears to be coming from multiple vehicles, and may adjust their routes. In such a case V_M can potentially obtain exclusive access to the road, which it otherwise could not. A number of other applications like content exchange, intelligent traffic signalling, and ramp metering can all be compromised in the presence of Sybil attacks. Unlike static networks like the Internet, vehicular mobilities make Sybil detection very difficult with the added spatio-temporal constraints.

Related work

The problem of detection Sybil attacks in VANETs has been previously studied. In [1] and [2], the proposed solution detects Sybil attacks when vehicles may only hold one valid pseudonym at a time. When a pseudonym need to be refreshed, a new pseudonym is obtained from a trusted Road-Side Unit (RSU). The consequence of this approach is a possibly complex pseudonym allocation mechanism implemented by the RSU network. Another technique leverages directional antennas to identify the location/direction of message arrival [3]. A vehicle launching a Sybil attack will likely be detected as many messages will arrive from the location/direction. However, in dense networks, localization errors can lead to frequent false positives. This scheme may be compromised as a smart attacker may use directional antennas to mislead its neighbors about its direction.

In [4], heavy-weight cryptographic techniques are leveraged for detecting Sybil attacks in VANETs. Specifically, each vehicle is given a list of pseudonyms to protect their privacy during communication. However, the pseudonyms of each vehicle are designed in such a manner wherein they are all hashed to a common value. By calculating the hashed values at Road Side Units, a central server can determine whether or not certain pseudonyms came from the same pool. Sybil attacks are detected if many pseudonyms from the same pool are detected in a short interval of time. Unfortunately, the computational complexity of cryptographic protocols in this technique is quite high.

In [5], GPS and RSSI signal measurements are used for detecting Sybil nodes. The proposed scheme uses Vehicle-to-Vehicle (V2V) communications to confirm reported positions of vehicles by referencing the RSSI measurements. To correct inaccuracies arising from RSSI measurement, caused by vehicle mobility, traffic patterns and support from roadside base stations are used. Specifically, statistical algorithms are implemented to verify the signal strength distribution of a suspect vehicle over time to significantly reduce the detection rate. In [6], analysis is performed to quantify performance of Sybil detection under assumptions like transmission range, antenna model, signal strength etc. Unfortunately, the un-reliability of RSSI measurements limits the practical reliability of these techniques [7]. In [8], inability of multiple vehicles to exhibit close temporal and spatial correlations at multiple locations is exploited for Sybil defense. The idea is to have RSUs sign location and timestamp information for vehicles as they move. Upon detecting groups of vehicles having many similar locations with similar timestamps, a Sybil attack is detected. The overhead in this scheme though is quite high, especially in the case of

urban networks. Significant cryptographic overhead is incurred as RSUs have to sign each received message.

We also would like to point out two other areas of work that are also closely related to our problem of Sybil attack detection in vehicular networks. The first area is secure localization in wireless networks like sensor and mobile ad hoc networks [9,10], wherein locations of nodes are determined in a secure manner. Our problem is similar, but orthogonal in the sense that we are attempting to verify integrity of relative location updates as vehicles move in the network. The other area we wish to highlight is the issue of detecting Sybil attacks and nodes in static networks like sensor, Internet scale, and social networks [11-14]. As can be observed, while the goal of these works are related to ours, the issue of vehicle mobility and unique mobility patterns of these nodes necessitates fundamentally new approaches for Sybil detection, which we attempt in this paper.

Contributions

Presented in this paper is an innovative protocol for Sybil detection in vehicular networks. Vehicular networks today are examples of cyber physical systems, where there is a clear integration of cyber and physical components. The premise of this paper starts with two simple questions: Can the natural physics of the underlying transportation domain be integrated with the Cyber domain in detecting Sybil attacks, and b) If so, can such an integration generate high quality solutions to detect Sybil attacks, while alleviating complexities (in the form of complex cryptography and additional hardware requirements) in the cyber domain. This papers yields a positive response to both questions.

The technique employs a certain number of road side units (RSUs) that periodically collect reports from communicating vehicles regarding this neighborhood. In the event of a vehicle performing Sybil attacks, the geographic proximity the Sybil identities will be long-term and repeating, while the geographical proximity of benign vehicles will short-term. To put it in terms of transportation engineering, Sybil identities will appear to "*platoon*" together, while identities of benign vehicles will eventually "*disperse*". The dispersion of vehicles in roads occurs due to a combination of road conditions, vehicle dynamics and human factors. This theory has been extensively studied by transportation engineers in the last five decades in the form of a theory called "*platoon dispersion*" [15-18]. Integrating platoon dispersion models provide an alternative method for Sybil attack detection. To detect attacks, RSUs compare models of naturally occurring dispersion among benign vehicles with anomalously occurring platoons among Sybil nodes. Using a combination of both theoretical analysis and simulations, the simplicity, efficiency, practicality and quality of the protocol for Sybil detection in vehicular networks is demonstrated. To the best of the authors' knowledge, this paper is unique in proposing an inter-disciplinary approach for addressing cyber space attacks in emerging vehicular networks.

Paper organization

The rest of the paper is organized as follows. Section 'Platoon dispersion and its application to Sybil detection' presents a brief overview of platoon dispersion theory in transportation engineering, and its application for Sybil detection in vehicular networks. Section 'Research design and methodology' presents the formal attack model, problem statement, overall framework, and protocol for Sybil detection. Section 'Performance

evaluations' will demonstrate the performance of the protocol, and the paper concludes in Section 'Conclusions'.

Platoon dispersion and its application to Sybil detection

Provided first is a brief overview of how models of dispersion among vehicles that naturally occur in roads have been studied by transportation engineers. Afterwards, a simplified example of how to use platoon dispersion theory for Sybil detection is presented. The discussions will help guide the proposed Sybil detection protocol discussed in the next section.

Platoon dispersion theory in transportation engineering

A platoon is a group of vehicles traveling in close proximity for some amount of time as shown in Figure 1. Ideally, consistent vehicle platooning is preferable and improves critical transportation parameters like signal optimization, congestion avoidance, improved road safety, and capacity [19-23][a]. Under normal traffic, vehicle platooning is short-term. Clearly, if all vehicles in an existing platoon are traveling at a constant speed, a platoon will never disperse. However, due to physical factors like road friction, vehicle characteristics and signalling, human factors, lane changes, and fatigue [24] cause platoons to disperse over time. The longer the travel time between points the greater dispersion, due to the difficulty of maintaining constant speed over longer time scales. This phenomena is called *platoon dispersion*, a simple illustration of which is shown in Figure 2.

Platoon dispersion has been well studied in transportation engineering [15-17,25-31], via two mathematical models. One is the (more popular) Robertson's geometric distribution model [16] and the other is the Pacey's normal distribution model [15]. Both models assume that road segment travel times follow some probability distribution. The Robertson platoon dispersion model follows a shifted geometric series, and has been implemented in traffic-simulation software like SCOOT [32], SATURN [33] and TRAFLO [34]. The basic of Robertson recursive platoon dispersion model takes the following form:

$$q'_t = R \cdot q_{t-T_{min}} + (1 - R) \cdot q'_{t-\delta t}. \quad (1)$$

Figure 1 An eight-car platoon. Real-life example of eight cars in a platoon configuration.

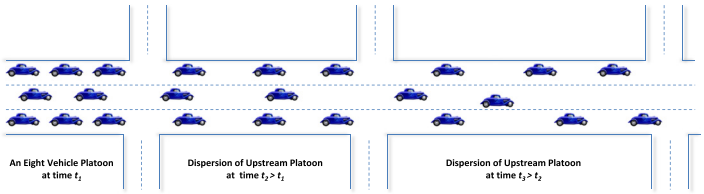

Figure 2 Formation and dispersion of platoons. Illustration describing the various states of vehicle platoons.

$$R = \frac{1}{1 + \alpha\beta T_{mean}}, \quad \text{where } 0 \leq R \leq 1. \tag{2}$$

A numerical procedure was developed for the Robertson model in [25] by rewriting Equation 1 as,

$$q'_t = \sum_{i=T_{min}}^{\infty} R \cdot (1-R)^{i-T_{min}} \cdot q_{t-i}. \tag{3}$$

where,

q'_t: arrival flow at the downstream location at time t-T (veh/hr);
q_t: departure flow at the upstream location at time t (veh/hr);
δt: time step duration;
T_{min}: minimum travel time on the roadway;
T_{mean}: mean roadway travel time, measured in units of time steps.
$\alpha = \frac{1-\beta}{\beta}$: dimensionless platoon dispersion factor depending on the level We are also investigating moreof friction along the roadway;
$\beta = \frac{2T_{mean}+1-\sqrt{1+4\sigma^2}}{2T_{mean}}$: dimensionless travel-time factor;
R: smoothing factor governing dispersion, where $0 \leq R \leq 1$;
σ: the standard deviation of link travel time assuming individual vehicle speeds follow normal distribution and are unchanged.

As can be seen from Equations 1, 2, 3 and definitions of parameters, all we need to know are the speed deviations σ among vehicles, and the mean travel time T_a between the upstream and downstream locations. If both can be determined (which is quite straightforward to obtain), one could compute platoon dispersion factors α and β. These parameters subsequently can be used to compute the smoothing factor R, from which the degree of how an upstream platoon will disperse at the downstream location can be computed.

Figure 3 shows an illustration of *upstream* platooning and its *downstream* dispersion, wherein the shaded portion represents *similar* vehicle speeds that tend to platoon together, while non-shaded portion represents *varying* speeds of vehicles that *disperse* from the original platoon. A numerical example of dispersion based on the Robertson model [16-18] is shown in Figure 4. Each observation (i.e., downstream) point is one mile apart, and the minimum travel time between each point is one minute. For small speed deviations, the dispersion in expected number of vehicles reaching the observation point

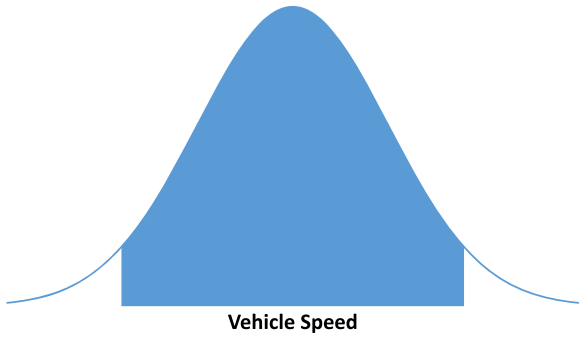

Figure 3 Illustration of platoon dispersion model.

is less than the case where the speed deviation increases. This leads to platoon sizes decreasing, progressively, as vehicles travel from one observation point to another.

An illustrative example of Sybil detection using platoon dispersion

Consider a case where there are 50 vehicles in an upstream platoon. Let each vehicle have a unique identity given by $\{V_1, V_2, \ldots, V_{50}\}$. Vehicle V_{50} is malicious and possesses 50 fake identities $\{\bar{V}_1, \bar{V}_2, \ldots, \bar{V}_{50}\}$. When all vehicle communicate with each other (including V_{50} with all of its identities to launch a Sybil attack), the up-stream platoon will appear to have 100 vehicles. With prior knowledge of road characteristics and (either currently sampled or prior estimates of) vehicle speeds, the dispersion parameters and the expected degree of dispersion at downstream can be computed. Say the smoothing factor is $R = 20\%$. If the Sybil, V_{50}, is part of the downstream platoon (recall shaded area in Figure 3), the number of identities actually seen in the downstream platoon is $n_d = 0.20 \times 50$ (benign vehicle identities) $+ 50$ Sybil identities $= 60$ identities. If the Sybil vehicle falls outside of the downstream platoon (recall the non-shaded area in Figure 3), then the number of identities actually seen in the downstream platoon is $n_d = 0.20 \times 50 = 10$ identities.

It is easy to see that abnormalities in the physical domain will manifest in the form of abnormal platooning (and ensuing dispersion) under Sybil identities in cyber space. If all

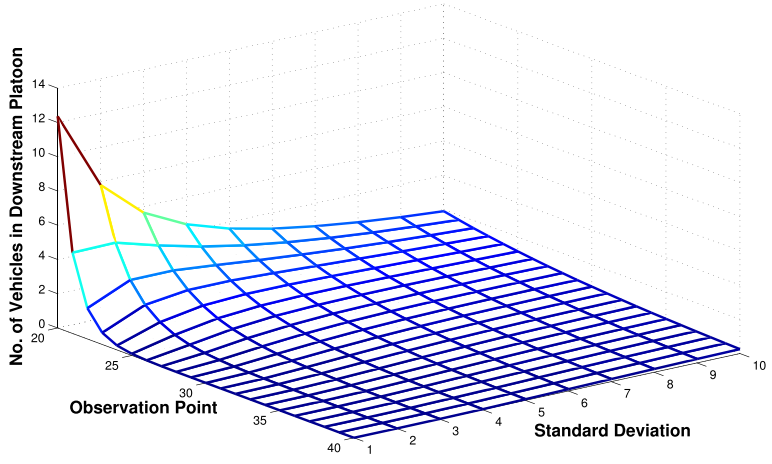

Figure 4 Numerical example of platoon dispersion. Graph describing the decrease in platoon size as platoons reach various road side units.

the identities upstream (i.e. 100 of them) are benign, the number of vehicle identities in the downstream platoon is expected to be $n_d = 0.20 \times 100 = 20$. Sufficient abnormalities in platoon dispersion that are straightforward to determine leading to a natural, elegant, and simple technique to detect Sybil attacks. To the best of the authors' knowledge, such a technique has not been attempted yet, and is formalized and elaborated in the next section.

Research design and methodology

In the previous section, a basic theory of platoon dispersion was illustrated, and how in principle can be leveraged for detecting Sybil attacks in vehicular networks. In this section, the protocol for detecting Sybil attacks via leveraging from platoon dispersion is presented. Several practical challenges in the actual design and implementation of the proposed protocol are addressed afterwards.

Vehicular network model

It is assumed that a certain number of Road Side Units (RSUs) are deployed in the vehicular network. The RSUs can communicate with each other and vehicles on the road. Communication is achieved through a 2-way radio, such as a DSRC (Dedicated Short Range Communications), to send and receive messages to other vehicles and RSUs. Each Vehicle V_x also has a unique identity that can identify it in the network. The identity for each V_x also acts as a unique public key P_x used for message encryption. Each vehicle maintains a secure private key for decryption of messages. Since energy is not a constraint, asymmetric encryption is feasible in vehicular networks [35]. Note that message confidentiality and privacy are not emphasized in this paper. It is assumed that some system exists for protecting confidentiality. If privacy is desired, techniques have been proposed for utilizing temporal pseudonyms that expire after a certain time [36], or using local coordination among vehicles and aggregating responses before forwarding messages. Such techniques do preserve vehicle privacy, and their usage will not affect the proposed protocol. Furthermore, a Central Coordination Authority (CCA) is involved in coordination among all RSUs, along with any key and pseudonym distribution among vehicles. The CCA and RSUs are assumed to be trusted.

For the proposed protocol, RSUs must have some prior information about vehicles and speed distributions along road segments. Such knowledge is reasonable and practical. In many countries across the world, including the US, efforts are being made to estimate vehicle densities and speed distributions for traffic management purposes. This is acheived using a variety of modern equipment like road side sensors, traffic light cameras and remote sensing imagery [37-40]. More discussion on this assumption and the accuracy of estimates are detailed in the next subsection during the description of the proposed protocol.

Simple attack model and problem definition

The attack model provides contextual information and elaborates the roles of the various agents. The attacker intends to subvert the integrity of peer vehicle communications by launching Sybil attacks. The attack has captured a certain number of legitimate identities (or keys) belonging to other vehicles. Such an attack model is practical, powerful, and has not been considered yet in related literature [1-6]. For example, a malicious parent could

easily steal identities of multiple members of their family that have identities already provisioned. Clearly with more sharing of vehicles today (e.g., rental cars and car pools), the feasibility of skimming attacks in vehicles (or in related hardware) to steal identities is even more practical. With multiple legitimate identities in hand, the potency of an attacker is much higher. The mobile nature of vehicular networks makes detection of such identity thefts quite difficult in practice. It is assumed that the attacker will utilize all its Sybil identities to attack the network in an attempt subvert the network integrity. Note that in this paper, we consider for simplicity one attacker in the network. The protocols we propose can be directly applied to multiple (but non-colluding) attackers. Analyzing and enhancing protocols for thwarting colluding attackers is part of future work[b].

Given the attack model, the formal problem statement is to rapidly detect the presence of Sybil attacks in a vehicular network.

Protocol 1 Protocol Executed by Each Vehicle V_x

1: **Each Vehicle V_x Executes the Following Steps when traveling between RSUs**
2: **For** every Vehicle V_y communicating with V_x
3: Store V_y
4: **End For**

5: **Each Vehicle V_x Executes the Following Steps when in range of RSU**
6: Forward Stored Vehicle Identities to RSU

Methods

Proposed technique for Sybil detection in vehicular networks

Overview

In Section Platoon dispersion and its application to Sybil detection, the basic overview of platoon dispersion and how it can be leveraged in principle for Sybil detection was presented. Unfortunately, there is one critical challenge to overcome before practical Sybil attack detection via leveraging platoon dispersion. Recall from Section 'Platoon dispersion theory in transportation engineering' that existing models for platoon dispersion assume that vehicle speeds from an upstream point to a downstream point are unchanged. These speeds are then used to derive σ (which is used to derive Smoothing Factor R in Equation 2), that subsequently determines dispersion phenomena. Anomalies in the dispersion phenomena naturally provide an ideal foundation for detecting Sybil attacks *in theory*. However, in practice, vehicle do not travel with constant speed and speed can vary widely between vehicles. This is solved by using measurement samples from peer vehicles and incorporating these data to determine platooning anomalies. A high level of confidence of detecting Sybil attacks is possible through this method.

Protocols description and analysis

Protocols 1 and 2 executed by vehicles and RSUs, respectively, illustrate the proposed technique for Sybil detection. As demonstrated in Protocol 1, each Vehicle V_x will store the identities of all vehicles with which it has communicated when traveling between RSUs. These identities are forwarded to an RSU when the vehicle is in range. Upon a vehicle transmitting its internally stored list of other communicating vehicles, the list may be

Protocol 2 Protocol Executed by Each Downstream RSU R_d

1: **Inputs:**
Definition of Platoon; Mean (μ_N^*) and Standard Deviation (σ_N^*) of Travel Times for N Vehicles in Upstream Platoon at R_u traveling to R_d based on Historical Information; CDF of Platoon Dispersion ($F_R(r) = \int_0^r \frac{1}{\sigma_N^* \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\sqrt{\frac{1-x}{x^2}} - \mu_N^*}{\sigma_N^*}\right)} d_x$; Confidence Interval of Sybil Detection (ϵ);

2: **Upon Receiving Vehicle Ids (Y_u) of a Platoon from Upstream RSU** R_u
3: Denote N as Number of Received Identities
4: Wait for Minimum Travel Time (T_{min}) between Location of R_u and Current Location
5: Receive Identities from every Forwarding Vehicle
6: Denote Y_d as Number of Vehicles Currently Platooning
7: Set m as Platoon Ratio ($m = \frac{Y_d}{Y_u}$)

8: Compute CDF of m vehicles in current platoon as $F_R(r = m) = \int_0^m \frac{1}{\sigma_N^* \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\sqrt{\frac{1-m}{m^2}} - \mu_N^*}{\sigma_N^*}\right)} d_x$

9: **If** $\frac{\epsilon}{2} < F_R(r = m) < 1 - \frac{\epsilon}{2}$

10: Flag No Sybil Attacks
11: **Else**

12: Flag the Detection of Sybil Attack
13: **End If**

14: Forward Current Platoon to Downstream RSU

deleted. Additionally, vehicles do not have to disclose their location for the protocol to work.

Protocol 2 presents the steps executed by a downstream RSU R_d for Sybil detection. The inputs to the protocol are as follows. First, the Central Coordination Authority defines what constitutes a platoon. One simple way to define a platoon is to say all vehicles currently within the communication range of an RSU is a platoon. Alternate definitions can also exist depending on traffic models and does not change the protocol or its execution. Second, R_d will have prior models of vehicle densities and speed distributions from historical information. Such information can be obtained from a number of traffic management organizations. These organizations regularly collect information on vehicle volumes, densities, and speeds along road segments to improve congestion control, signalling, accident management, and other traffic attributes. A number of recent studies also propose innovative and accurate approaches to obtain such information including deployment of road side sensors and remote sensing imagery to obtain such information [37-40].

From such information, each RSU can derive the mean (μ_N^*) and standard deviation (σ_N^*) of vehicle travel time based on the number of vehicles N traveling from upstream. Note that even if such information may be historical and may not reflect existing trends, it is always possible to obtain samples in real-time from current traffic to build accurate more profiles. Recall from Equation 2, that the smoothing factor R governing platoon dispersion is given by $R = \frac{\sqrt{1+\sigma^2}-1}{2\sigma^2}$, where σ is the standard deviation of link travel time assuming individual vehicle speeds follow normal distribution and are unchanged. To obtain σ, Equation 2 can be rewritten as $\sigma = \sqrt{\frac{1-R}{R^2}}$. In practice vehicle speeds and link travel times on roads between RSUs fluctuate. Assuming the link travel times of each vehicle follows normal distribution [41,42] with parameters μ_N^*, σ_N^*, and N vehicles in an upstream platoon, the probability density function (pdf) of σ is

$$f_\sigma(x) = \frac{1}{\sigma_N^*\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_N^*}{\sigma_N^*}\right)} \tag{4}$$

Consequentially, the probability density function (pdf) for the smoothing factor R as $f_R(r) = \frac{1}{\sigma_N^*\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\sqrt{\frac{1-r}{r^2}}-\mu_N^*}{\sigma_N^*}\right)}$. The Cumulative Distribution Function (CDF) is then given by

$$F_R(r) = \int_0^r \frac{1}{\sigma_N^*\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\sqrt{\frac{1-r}{r^2}}-\mu_N^*}{\sigma_N^*}\right)} d_x. \tag{5}$$

The final input to the protocol is the confidence interval (ϵ) of Sybil Detection, that is predetermined by the Central Coordination Authority.

Once R_d obtains the number of vehicles n_d departing from R_u, it waits for T_{min}, the minimum travel time on the link. R_u starts receiving messages from vehicles regarding identities they had communicated with. Based on times of messages received, R_d determines the number of vehicles in the current platoon as Z. It then computes the platoon ratio as $m = \frac{Z}{Y}$ and $F_R(r = m)$ from the CDF of the smoothing factor. There are two cases of interest here. When the Sybil attacker is part of an upstream platoon and part of the downstream platoon, the number of vehicles Y_d in the downstream platoon will be abnormally large as the Sybil identities will not disperse. In the second case, the Sybil vehicle is not in the downstream platoon, causing the number of vehicles Y_d in the downstream platoon will be abnormally less. This is due to large number of Sybil identities being dispersed. In the first case, $F_R(r = m)$ will assume a large value as the platoon size will show little relative change, and in the latter case $F_R(r = m)$ will assume a much smaller value as the relative change in platoon size is high. This is captured in Step 9 of Protocol 2, where abnormal platoons in both cases are checked to indicate a Sybil attack.

Results and discussion

Only a limited number of numerically obtained parameters are used in the execution of the protocol. The critical parameters are μ_N^*, σ_N^* and ϵ. When μ_N^* is low, dispersion is low. This means that the likelihood of vehicles platooning together is higher, hence lowering the chances of detecting Sybil identities. On the other hand, σ_N^* denotes the standard deviation or error in the estimation of link arrival time. If σ_N^* is low, then the error in estimation of link arrival time is better, yielding a steeper CDF, which improves

detection accuracy. The parameter ϵ determines the degree of confidence in detection. Finally, all of these parameters are integral to the attacker strength in terms of number of Sybil nodes during detection. Next, the impacts of these parameters on the performance of the protocol are detailed.

Performance evaluations

In this section, the performance evaluation of our protocol to detect Sybil attacks in vehicular networks are reported. Section 'Preliminaries' presents preliminaries, while Section 'Analysis and simulation results' illustrates performance data.

Preliminaries

The simulations were performed on Simulation of Urban Mobility (SUMO), which is an open source traffic simulation package [43]. SUMO allows importing custom maps, user-defined vehicle trips, and detailed simulation output for every vehicle. A map consisting of 10 intersections by 10 intersections was constructed to simulate a typical inner-city topology. The distance between each intersection 400 ft or approximately 122 m. This gives a block of 400 ft $\times$ 400 ft. The map was then populated with varying number of vehicles generated with random source/destination pairs with the trip defined as the shortest route between the source and destination. The simulator provides a high degree of configurability for the protocol analysis while providing reports in easy-to-process formats. Additionally, the simulator provides fine-grain data as information for every vehicle for every second is saved. We wish to point out that the grid based topology we employ is widely used for simulation purposes and hence we adopted it in this paper. Extending our simulations to more complex and realistic topologies is part of future work.

A normal distribution with the current case parameters were used to randomly assign start and maximum speeds for vehicles. The simulator then creates a simulation report after completion. The output file includes every vehicle's speed, position, current road, position on road, and other data. Then these logs files were parsed to create individual files for every vehicle containing the timestep, vehicle speed, current position, and current road at that timestep.

The same data was used to create an output file containing the time a vehicle encountered a RSU and the speed of that vehicle. Note that RSUs were placed at each intersection. The simulator was utilized to learn μ_N^* and $\sigma_N^{*\,\mathrm{c}}$ for various values of number of vehicles N from 10 to 1000. Unless otherwise stated, the confidence internal was $\epsilon = 0.05$, and all vehicles within a 0.5 mile range were considered as part of a platoon. All results were averaged over 50 iterations, with a $\mu = 35$ mph. Due to space limitations, only simulation data are reported and discussed. However, it should be noted that the analysis data agrees well with simulations results.

Analysis and simulation results

A road with a length of 10 miles was implemented for this study to allow for sufficient dispersion among vehicles. T_{min} and the average of σ^* (denoted as σ) are determined from simulation data, as shown in Figure 5. Then they were substituted in Equation 3 to determine the expected number of vehicles in the downstream platoon. As the number of vehicles increase, the trend remains the same. This sufficiently demonstrates the fidelity of the simulation environment in conducting further investigations. The trend

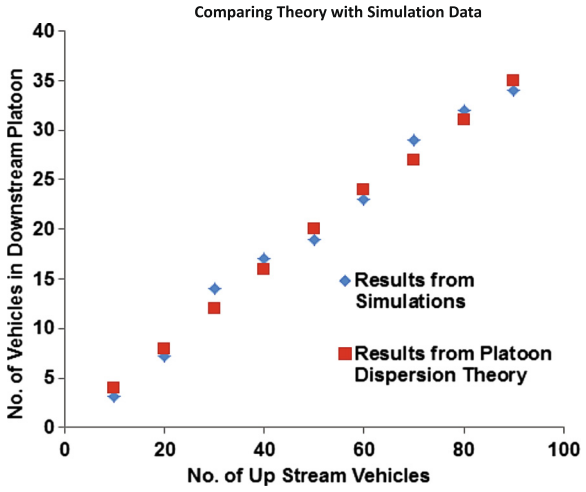

Figure 5 Comparing platoon dispersion theory with simulations. Graph of the expected results from the platoon dispersion model along with the simulation calculated results.

also remains for road segments of different lengths, which are not reported due to space limitations.

Figure 6 illustrates the basic features of our protocol, when the number of vehicles in upstream platoon $n_u = 50$. The Y-axis denotes the detection time, or number of RSUs traversed before a detection of Sybil detection is made. In this case, σ^* is constant and fixed as 1.0. When μ^* increases, Sybil attacks are detected faster. This high value of μ^* causes increased dispersion which increases the chances of detecting anomalous platoons. Also, varying the percentage of Sybil attackers varies the detection time. When the percentage of Sybil attackers increase, anomalous platoons are easier to detect, quickening the detection time.

Figure 7 illustrates the trend of false negatives in the protocol. The number of vehicles in upstream platoon is still set as $n_u = 50$, and $\sigma^* = 1.0$. As μ^* increases, the false negatives decreases. Again, the high value of μ^* increases dispersion and the chances of detecting anomalous platoons correctly which lowers false negatives. Interestingly, the false negatives rate decreases as the percentage of Sybil attackers increase. As more Sybil attackers

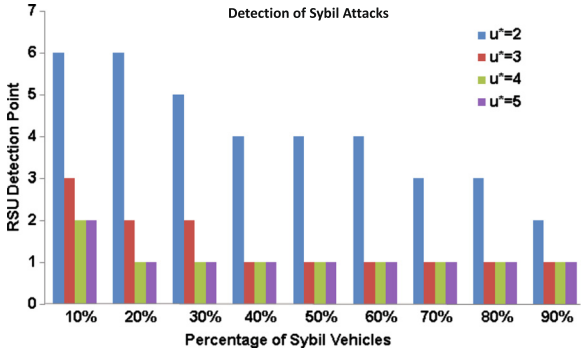

Figure 6 Sybil detection under varying percentage of Sybil vehicles. Chart characterizing at which road side unit Sybil nodes are detected against the percentage of vehicles classified as Sybil nodes.

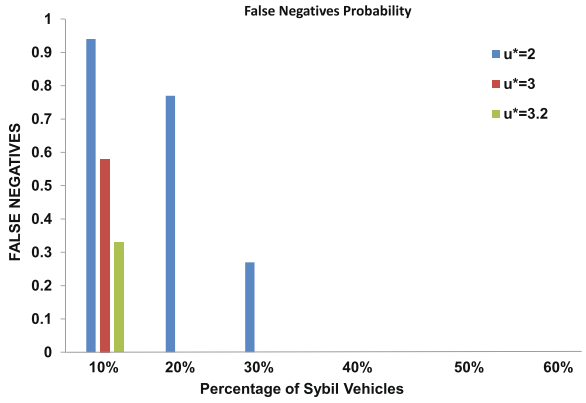

Figure 7 False negatives under varying percentage of Sybil vehicles. Chart detailing the trend of the number of false negatives as the percentage of Sybil vehicles increases.

leads to increased chances of detecting anomalous dispersion. For the case of Sybil attackers reaching 40% and above, the false negatives completely disappear, demonstrating that the protocol is robust against increasing degree of Sybil attacks.

In Figure 8, the trend of how increasing number of vehicles affects detection performance when the percentage of Sybil vehicles is 10% is shown. The observed trend is straightforward to interpret. When the number of vehicles decrease, σ^* increases. This causes increased platoon dispersion, thus accelerating the detection rate. When the number of vehicles increases, roads become congested and degree of dispersion decreases, causing an increased duration of Sybil detection.

In Figure 9, the trade-off between false negatives and false positives when $N = 50$, $\mu^* = 2$ and $\sigma^* = 0.4$ is shown. As we can see when the percentage of Sybil vehicles is low, the protocol yields excellent performance in terms of false positives and false negatives. Though, with increasing attack intensity the performance degrades. The false positive rate is decided by the parameter ϵ, which is user specified. When ϵ is low, the false positive rate is low and the false negative rate is higher, and vice versa. However, this is also sensitive to the number of Sybil vehicles. How to address this trade-off in practical settings is the topic of future research. One plan is to integrate long term predictions from RSUs

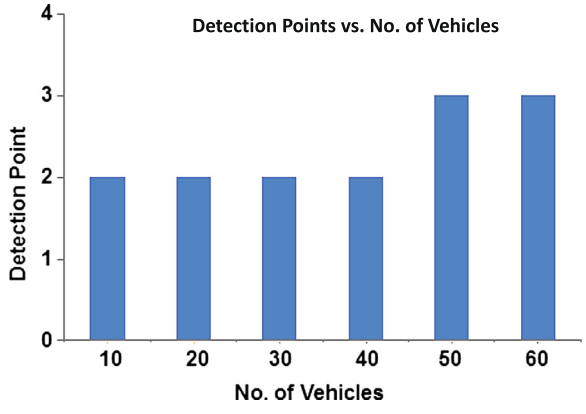

Figure 8 Sybil detection versus number of vehicles. Chart describing the road side unit Sybil nodes are detected at with various number of total vehicles.

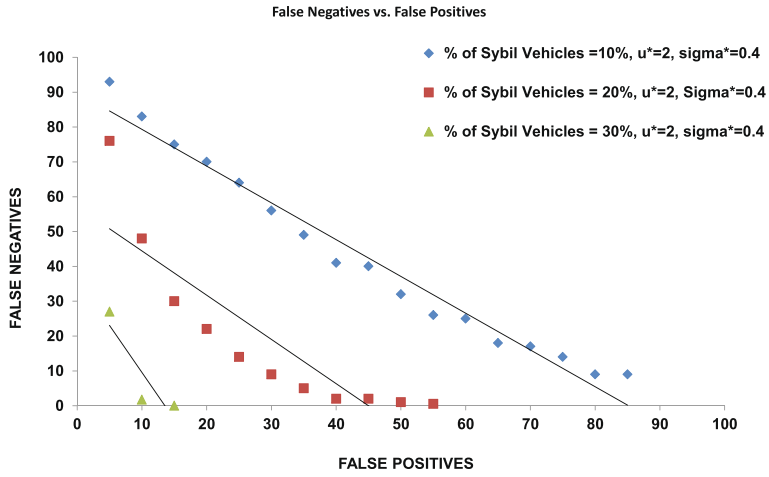

Figure 9 False negatives versus false positives. Chart comparing the number of false negatives versus false positives.

leveraging platoon dispersion with short term prediction models using distributed computation among local vehicles. Also, adapting the protocol to changing road conditions under traffic dynamics is also part of future research.

Advanced attack model

The previous results and discussions have pertained to a simple attack model where the Sybil vehicles broadcast all stored pseudonyms during an attack. Once an RSU network attempts to identify a Sybil attack, the Sybil vehicles may change their message distribution algorithms to avoid detection. The following scenarios describes alternate attack methods a Sybil vehicle can implement to avoid detection. To determine the robustness of the proposed protocol, the results of the protocol against different Sybil attack schemes are presented.

Normal dispersion attack efficiency scenario

The main objective of the attacker is to maintain the efficiency of an attack. Attack efficiency is defined as ratio of Sybil pseudonyms to benign pseudonyms. In other words, the attack efficiency can be defined as

$$e = |V_s|/|V_b| \cdot 100\%, \text{ where}$$

$e =$ the attack efficiency

$V_s =$ set of Sybil pseudonyms

$V_b =$ set of benign pseudonyms

The attacker wants to maintain a high efficiency in order to masquerade Sybil pseudonyms as benign pseudonyms. For this to become possible, the attacker must have a priori road information. Say, the attack releases a subset of V_s, called $V_{s,1}$, at an upstream RSU. At the downstream RSU, the attack can release a subset of $V_{s,1}$ such that it appears the platoon dispersed under normal circumstances. The number of pseudonyms to "disperse" can be derived from the a priori road information and platoon dispersion theory.

Normal dispersion attack efficiency results

There are some practical limitations to this attack scenario, which the our protocol to detect Sybil attacks exploits. In order to have an effective attack, V_s (set of Sybil identities) must be large compared to V_b (set of benign identities) otherwise the Sybil attack will have little influence on the benign vehicles. Using the previous example of Sybil nodes incorrectly reporting road conditions, if exactly 50% of the pseudonyms were Sybil and reporting false information while the other 50% (benign pseudonyms) are reporting true information. No decision could be made on what information to believe. Therefore, an attacker needs more than 50% Sybil pseudonyms to launch an effective attack.

The attack efficiency, e, is held constant as both the Sybil and benign pseudonyms are dispersing at the exact same rate, as shown in Figure 10. This effect leads to an exhaustion point for the Sybil pseudonyms, where there are no more available pseudonyms to use. The attacker cannot reuse previous pseudonyms for some time interval as that would trigger an attack detection. Essentialy, reuse of previously dispersed pseudonyms leads anomalous behavior and is caught by the protocol. Figure 11 shows at which RSU the attacker has completely exhausted all available pseudonyms. At that point, the attacker must cease to launch an attack for some time or risk being detected.

Comparing the results of Figures 6, 7, 8, 9, 10 and 11, for the 10% and 20% Sybil pseudonyms cases, the attacker actually exhausts all available pseudonyms before the protocol could even detect an attack. The 30% to 60% cases show pseudonym exhaustion near the same point of detection, under the simple attack scheme. The highest percent Sybil pseudonym cases, 70% to 90% show a significant increase of attack duration before exhaustion. Under the simple attack method, these cases could be identified within 1 RSU. With the normal dispersion efficiency attack, the pseudonym exhaustion does not occur until the 6^{th} RSU. This may seem as a favorable scheme for the attacker, but even the best case of the simple attack method resulting in Sybil attack detection at the same RSU. The

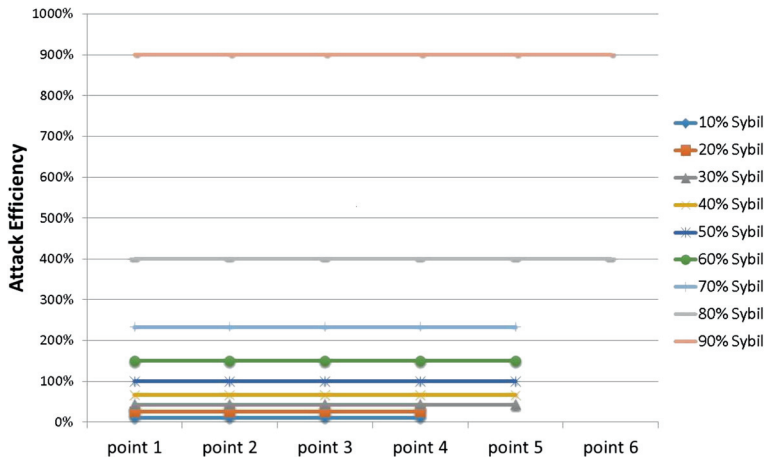

Figure 10 Effiency rate of Sybil attack under Normal Dispersion Efficiency Attack method. The efficiency rates of Sybil attacks under this method are held constant due to some pseudonyms being "dispersed" as the attacker moves. Benign vehicles disperse at the same rate leading to the constant efficiency.

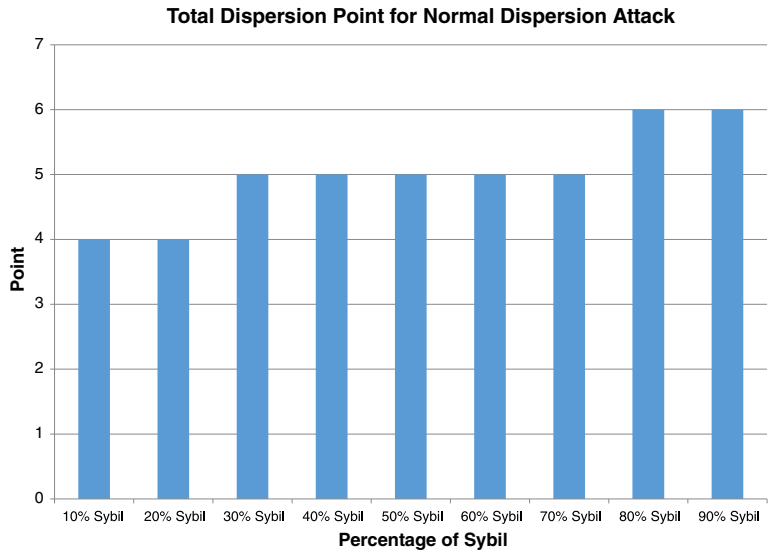

Figure 11 Id. of RSU where attacker exhausts all Sybil pseudonyms. The point of pseudonym exhaustion under the Normal Dispersion Efficiency Attack scheme. As the percentage of Sybil pseudonyms increase, the longer it takes for that set to exhaust. In the worst case, all pseudonyms are used by the 6^{th} RSU, showing comparable performance to the simple attack.

result of this simulation demonstrates the robustness of the protocol to multiple attack methods.

Minimum efficiency attack scenario

Similar to the Normal Dispersion Efficiency Attack, the Minimum Efficiency Attack (MEA) attempts to reduce the number of Sybil pseudonyms used while maintaining an influential attack on a vehicular network. The MEA reduces the number of pseudonyms used by computing the minimum number needed to gain an influence. The attack measures the number of vehicles neighboring platoon to obtain N_p. In order to influence the platoon, the attacker needs to use $N_p + 1$ pseudonyms at a minimum. This allows the attacker to launch a new attack once it knows the attack will be detected. For example under the 80% Sybil case and 100 vehicles, the attacker knows that $N_p = 20$. The attacker will then launch 21 of its 80 Sybil pseudonyms to start an advantageous attack. Knowing the detection scheme, the attacker also knows that this attack will be detected by the fifth RSU. Consequentially, the attacker retires the previous pseudonyms and launches a new attack with 21 of the remaining 59 pseudonyms available. This process continues until all of the attackers Sybil pseudonyms have been exhausted.

Minimum efficiency attack results

As the platoon travels, benign vehicles will disperse but the attack will maintain the original pseudonym number of $N_p + 1$. Under this scheme, the attack efficiency of every case becomes over 100% efficient, as shown in Figure 12. This is ideal for an attacker that cannot obtain many pseudonyms from the network. Even the lowest Sybil percentage case (10%) increases to over 100% attack efficiency. This type is practical and ideal for attackers with very few Sybil pseudonyms. Some cases do not reach the high level of efficiency as in Normal Dispersion Efficiency attacks, but those cases have an increased duration before pseudonym exhaustion.

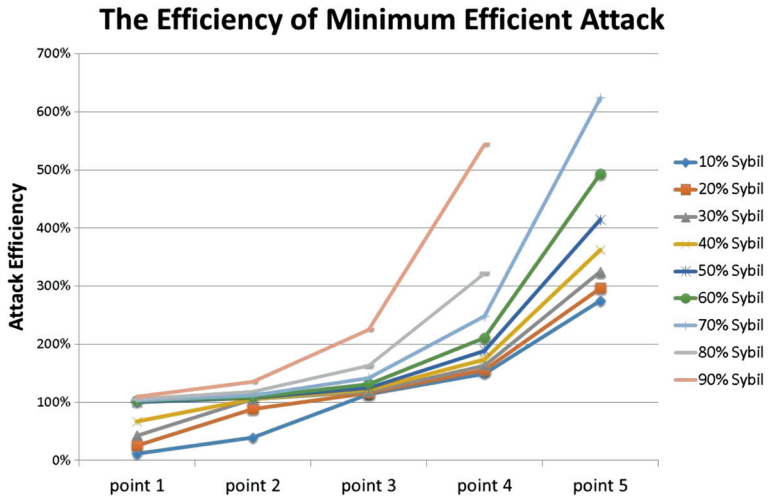

Figure 12 Minimum Efficient Attack efficiency as platoons reach RSUs further away. As an attack is launched, the attacker determines the minimum number of Sybil pseudonyms necessary to gain an advantage in the platoon. As the platoon disperses, the attacker increases its efficiency while minimizing the number of Sybil pseudonyms needed.

For the cases of less than or equal to 50% Sybil pseudonyms, the attacker cannot launch an effective attack. Each of these cases can only launch one attack and are detected at the sixth RSU, as shown in Figure 13. So even though those cases reach an attack efficiency of over 100%, it is short-term as another attack cannot be launch. The efficiency immediately drops to below 100% after the fifth RSU when the attacker attempts to launch another attack. The cause of this drop is the number of Sybil pseudonyms remaining is less than the number of vehicles in the platoon.

When the Sybil pseudonym percentage reaches 60%, detection time increases dramatically. The attacker knows when an attack will be detected, can retire the current Sybil

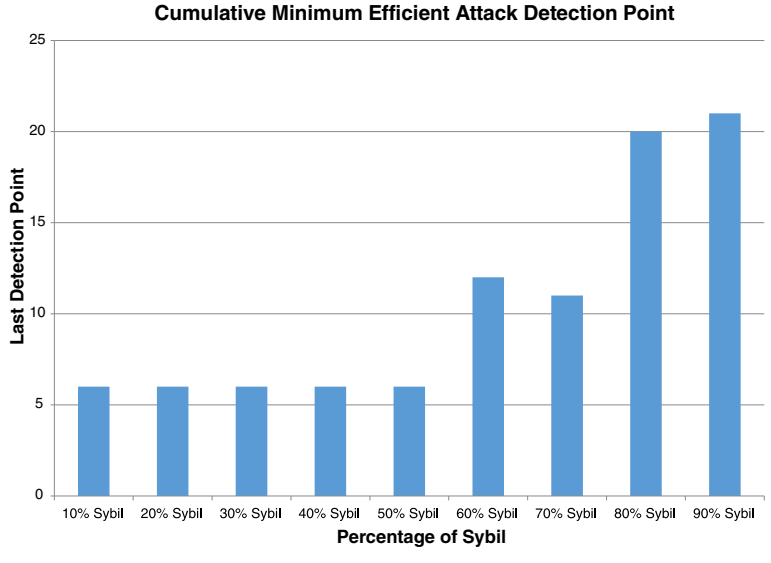

Figure 13 Minimum Efficient Attack detection points as the percentage of Sybil pseudonyms increase. As the number of Sybil pseudonyms increase, the detection point of an attack increases. In each case, multiple attacks may be launched so the last detection points are shown.

pseudonyms, and use fresh pseudonyms from V_s. This action is equivalent to launching another attack even though the motive may be the same. For example, in the 60% case the attacker can launch two undetectable attacks until the number of Sybil pseudonyms is less than the platoon size. At this point, the attacker uses the remaining pseudonyms to launch a final attack. This attack will be detected in a few RSUs. The summation of total RSUs traversed before the final detection is shown in Figure 13. The high-percent-Sybil cases show an apparent defeat in the proposed protocol. This attack scheme could be countered with the addition of mechanism to detect additions of many vehicle pseudonyms. It should be noted, though, obtaining such a high percent of Sybil pseudonyms could be a difficult task in real life.

Conclusions

In this paper, a novel protocol for defending against Sybil attacks in vehicular networks is presented. The novelty comes from the fusion of physical phenomena and the cyber domain to detect Sybil attacks. The combination of physical and cyber environments makes the protocol effective, practical, efficient, and simple. Additionally, this paper presents advanced attack methods where the attacker knows the detection scheme and has a priori road information. The protocol shows similar performance for Normal Dispersion Efficiency Attack model, while the Minimum Efficiency Attack model may remain undetected at high Sybil percentages. Future work involves integration of machine learning algorithms with platoon dispersion and wireless communication support to not only detect the presence of Sybil attacks, but identifying which are the Sybil nodes. Advanced collusion attacks to validate our protocol performance are also under investigation.

Endnotes

[a] In a recent 2011 study, platooning was exploited for content management in vehicular networks [44].

[b] The issue of detecting attacks where identities are selectively used (in which case attack potency is lowered) is part of future work.

[c] The subscript N in μ_N^* and σ_N^* is dropped in this section for ease of readability.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
MAM adapted the platoon dispersion Model to vehicular networks. He also conducted the theoretical analysis and also led the design of the proposed protocol. LM led all aspects of simulation studies, and results. He also conducted related work on Sybil attacks. SC defined the problem, and participated in the theoretical analysis. He also led the discussions on interpretations of the results and paper writing. All authors read and approved the final manuscript.

Received: 12 February 2014 Accepted: 20 March 2014
Published: 11 May 2014

References

1. Parno B, Perrig A (2005) Challenges in securing vehicular networks. In: Workshop on Hot Topics in Networks (HotNets-IV). ACM, College Park, Maryland, pp 1–6
2. Studer A, Shi E, Bai F, Perrig A (2009) Tacking together efficient authentication, revocation, and privacy in vanets. In: IEEE Sensor, Mesh and Ad Hoc Communications and Networks (SECON). IEEE, Rome, Italy, pp 1–9
3. Golle P, Greene D, Staddon J (2004) Detecting and correcting malicious data in vanets. In: Proceedings of the 1st ACM international workshop on vehicular ad hoc networks. ACM, Philadelphia, PA, USA, pp 29–37
4. Zhou T, Choudhury RR, Ning P, Chakrabarty K (2011) P2dap sybil attacks detection in vehicular ad hoc networks. IEEE J Sel Area Comm 29(3): 582–594
5. Xiao B, Yu B, Gao C (2006) 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. ACM, Los Angeles, CA, USA, pp 1–8

6. Guette G, Ducourthial B (2007) On the sybil attack detection in vanet. In: IEEE inernational conference on Mobile Adhoc and Sensor Systems (MASS). IEEE, Pisa, Italy, pp 1–6
7. Parameswaran AT, Husain MI, Upadhyaya S (2009) Is rssi a reliable parameter in sensor localization algorithms: an experimental study. In: Field Failure Data Analysis Workshop (F2DA09). IEEE, Niagara Falls, NY, USA
8. Park S, Aslam B, Turgut D, Zou CC (2009) Defense against sybil attack in vehicular ad hoc network based on roadside unit support. In: IEEE Military Communications Conference (MILCOM). IEEE, Boston, Massachusetts, pp 1–7
9. Capkun S, Hubaux J-P (2005) Secure positioning of wireless devices with application to sensor networks. In: IEEE INFOCOM. IEEE, Miami, FL, USA, pp 1917–1928
10. Jadliwala M, Zhong S, Upadhyaya S, Qiao C, Hubaux J-P (2010) Secure distance-based localization in the presence of cheating beacon nodes. IEEE Trans Mobile Comput 9(6): 810–823
11. Newsome J, Shi E, Song D, Perrig A (2004) The sybil attack in sensor networks: analysis & defenses. In: Proceedings of the 3rd international symposium on information processing in sensor networks. ACM, Berkeley, CA, USA, pp 259–268
12. Douceur J (2002) The sybil attack. Peer-to-peer Systems. Lecture Notes in Computer Science 2429: 251–260
13. Kurve A, Kesidis G (2011) Sybil detection via distributed sparse cut monitoring. In: 2011 IEEE International Conference on Communications (ICC). IEEE, Kyoto, Japan, pp 1–6
14. Yu H, Kaminsky M, Gibbons PB, Flaxman A (2006) Sybilguard: defending against sybil attacks via social networks. ACM SIGCOMM Comput Commun Rev 36(4): 267–278
15. Pacey G (1956) The progress of a bunch of vehicles released from a traffic signal. RN/2665/GMP, Transport and Road Research Laboratory, Growthorne, UK. Research note No. Rn/2665/GMP. Road Research Laboratory, London
16. Robertson D (1969) Transyt-a traffic network study tool. rrl report lr 253. London: TRRL
17. Yu L (2000) Calibration of platoon dispersion parameters on the basis of link travel time statistics. Trans Res Rec: J Trans Res Board 1727(-1): 89–94
18. Rakha H, Farzaneh M (2006) Issues and Solutions to Macroscopic Traffic Dispersion Modeling, pp 555-564
19. Robinson T, Chan E, Coelingh E (2010) Operating platoons on public motorways: an introduction to The SARTRE platooning programme: 1–11. World Congress on Intelligent Transport Systems
20. SARTRE-Consortium The SARTRE Project. http://www.sartre-project.eu/en/Sidor/default.aspx accessed 10 October 2012
21. Varaiya P (1993) Smart cars on smart roads: problems of control. IEEE Trans Automatic Control 38(2): 195–207
22. Frankel J, Alvarez L, Horowitz R, Li P (1994) Robust platoon maneuvers for avhs 10. Manuscript, Berkeley, November
23. Tongue BH, Yang Y, White MT (1991) Platoon collision dynamics and emergency maneuvering i: Reduced order modeling of a platoon for dynamical analysis
24. Van Winsum W (1999) The human element in car following models. Trans Res Part F: Traffic Psychol Behav 2(4): 207–211
25. Seddon P (1972) Another look at platoon dispersion: 3. the recurrence relationship. Traffic Eng Contr 13(10): 442–444
26. Salter RJ, Hounsell NB (1996) Highway traffic analysis and design. Palgrave Macmillan
27. Denney RWJr (1989) Traffic platoon dispersion modeling. J Transport Eng 115: 193
28. El-Reedy T, Ashworth R (1978) Platoon dispersion along a major road in sheffield. Traffic Eng Contr 19: 186–189
29. Abbas M, Bullock G, Rhodes A (2001) Comparative study of theoretical, simulation, and field platoon data. Traffic Eng Contr 42(7): 232–236
30. Liu HX, Bhimireddy S (2009) Evaluation of integrated platoon-priority and advance warning flasher system at high-speed intersections. Trans Res Rec: J Trans Res Board 2128(-1): 121–131
31. Li J, Wang H, Chen QY, Ni D (2010) Traffic viscosity due to speed variation: modeling and implications. Math Comput Model
32. Hunt P, Robertson D, Bretherton R, Winton R (1981) Scoot-a traffic responsive method of coordinating signals. Technical report
33. Hall M, Van Vliet D, Willumsen L (1980) Saturn-a simulation-assignment model for the evaluation of traffic management schemes. Traffic Eng Contr 21(4): 168–176
34. Lieberman EB, Andrews B (1980) Traflo: a new tool to evaluate transportation system management strategies. Transport Res Rec 772
35. Wasef A, Lu R, Lin X, Shen X (2010) Complementing public key infrastructure to secure vehicular ad hoc networks [security and privacy in emerging wireless networks]. IEEE Wireless Comm 17(5): 22–28
36. Gerlach M, Guttler F (2007) Privacy in vanets using changing pseudonyms-ideal and real. In: IEEE Vehicular Technology Conference (VTC), pp 2521–2525
37. Tan Q, Wei Q, Yang S, Wang J (2009) Evaluation of urban road vehicle detection from high resolution remote sensing imagery using object-oriented method. In: Urban Remote Sensing Event, pp 1–6
38. Grant C, Gillis B, Guensler R (2000) Collection of vehicle activity data by video detection for use in transportation planning. J Intell Transport Syst 5(4): 343–361
39. Oh S, Ritchie SG, Oh C (2002) Real-time traffic measurement from single loop inductive signatures. Trans Res Record: J Trans Res Board 1804(-1): 98–106
40. Haoui A, Kavaler R, Varaiya P (2008) Wireless magnetic sensors for traffic surveillance. Transport Res C Emerg Tech 16(3): 294–306
41. Roess RP, Prassas ES, McShane WR (2004) Traffic engineering. Pearson/Prentice Hall
42. Yousefi S, Altman E, El-Azouzi R, Fathy M (2008) Analytical model for connectivity in vehicular ad hoc networks. IEEE Trans Veh Tech 57(6): 3341–3356
43. Krajzewicz D, Hertkorn G, Rossel C, Wagner P (2002) Sumo (simulation of urban mobility) In: Proc. of the 4th middle east symposium on simulation and modelling, pp 183–187
44. Zhang Y, Cao G (2011) V-pada: Vehicle-platoon aware data access in vanets. IEEE Trans Veh Tech 99: 1–1

doi:10.1186/2196-064X-1-4
Cite this article as: Al-Mutaz *et al.*: **Detecting Sybil attacks in vehicular networks.** *Journal of Trust Management* 2014 **1**:4.