Voiceprint: A Novel Sybil Attack Detection Method Based on RSSI for VANETs

Yuan Yao*†, Bin Xiao†, Gaofei Wu*, Xue Liu‡, Zhiwen Yu*, Kailong Zhang* and Xingshe Zhou*
*School of Computer Science, Northwestern Polytechnical University, China
†Department of Computing, Hong Kong Polytechnic University, Hong Kong
‡School of Computer Science, McGill University, Canada
yuanyao8539@gmail.com;csbxiao@comp.polyu.edu.hk;gaofeiwu@nwpu.edu.cn;
xueliu@cs.mcgill.ca;{zhiwenyu, kl.zhang, xszhou}@nwpu.edu.cn

***Abstract*—Vehicular Ad Hoc Networks (VANETs) enable vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications that bring many benefits and conveniences to improve the road safety and drive comfort in future transportation systems. Sybil attack is considered one of the most risky threats in VANETs since a Sybil attacker can generate multiple fake identities with false messages to severely impair the normal functions of safety-related applications. In this paper, we propose a novel Sybil attack detection method based on Received Signal Strength Indicator (RSSI), Voiceprint, to conduct a widely applicable, lightweight and full-distributed detection for VANETs. To avoid the inaccurate position estimation according to predefined radio propagation models in previous RSSI-based detection methods, Voiceprint adopts the RSSI time series as the vehicular speech and compares the similarity among all received time series. Voiceprint does not rely on any predefined radio propagation model, and conducts independent detection without the support of the centralized infrastructure. It has more accurate detection rate in different dynamic environments. Extensive simulations and real-world experiments demonstrate that the proposed Voiceprint is an effective method considering the cost, complexity and performance.**

I. Introduction

Vehicular Ad Hoc Networks (VANETs) is a promising technology to address the challenging issues in the intelligent transportation system (ITS) such as accident avoidance, traffic monitoring and transport efficiency. VANETs enable a vehicle to directly communicate with neighboring vehicles (vehicle-to-vehicle, V2V) as well as roadside infrastructures (vehicle-to-infrastructure, V2I). According to a report published by National Highway Traffic Safety Administration, VANETs can provide a wide range of communication-based vehicle safety and non-safety applications in ITS such as intersection collision avoidance, cooperative collision warning, emergency electronic brake lights, traffic flow control and enhanced route guidance and navigation [1].

Dedicated Short Range Communications (DSRC) at 5.9 GHz is a set of protocols for VANETs issued by the Federal Communications Commission (FCC) in 1999. There are two kinds of communication devices defined in DSRC, namely the On Board Unit (OBU), which is installed in the vehicle, and the Road Side Unit (RSU), which is deployed on the roadside. Safety-related messages are broadcasted periodically on the Control Channel (CCH) by OBUs with the vehicles' identity, location, velocity, acceleration, direction and etc. Meanwhile, some useful information such as road condition, traffic density and accident zone are disseminated by RSUs to warn drivers within their vicinity.

The main purpose of VANETs is to improve the road safety as well as raise the traffic efficiency. Nevertheless, VANETs inherit all security vulnerabilities from the wireless networks, which becomes the major issue to apply this technology into practice. Many types of attacks can be launched in VANETs, but one of the most harmful is Sybil attack [2]. As aforementioned, many safety or non-safety applications in VANETs such as cooperative collision warning and enhanced route guidance and navigation need cooperation of other vehicles. This requires one vehicle gets enough credible information from legitimate vehicles. However, in Sybil attack, adversary (malicious node) generates multiple fake identities to create many untrusted virtual nodes (Sybil nodes) in VANETs. This violates the fundamental assumption in implementing those applications [3].

Due to the severe damage when Sybil attack happens, many detection methods are proposed by researchers. All these techniques can be classified into three categories: resource testing based, trusted certification based and position verification based mechanisms. The resource testing based methods may become invalid if the malicious node has more computation or communication resources, and they bring extra overhead to the system. Most of the trusted certification based methods run the detection algorithms in a centralized manner which are not suitable for the VANETs due to the fast changing dynamic topology. In addition, the deployment of public key infrastructure and the high complexity of cryptographic algorithms are also uncertain issues in this type of methods. Considering the low cost, wide availability and decentralized nature, the physical measurement based position verification methods are better for detecting Sybil attacks in the initial stage of VANETs.

In this paper, we propose a novel Sybil attack detection method based on RSSI, Voiceprint, to conduct a widely applicable, lightweight and full-distributed detection for VANETs. Unlike most of previous RSSI-based methods that compute the absolute position or relative distance according to the average RSSI values, or make statistic testing based on RSSI

distributions, Voiceprint uses the RSSI time series as the vehicular speech to compares the similarity among all these time series. This approach is based on the major observation in our real-world experiments that the RSSI time series of Sybil nodes have the very similar patterns. The main contribution of this paper is three-fold:

1) Voiceprint can be widely applied to real VANETs without any predefined radio propagation model. Extensive simulations and experiments show the applicability of the proposed method. It has high detection rate over 90% and low false positive rate under 10% in different dynamic environments. (model-free, widely applicable);
2) Voiceprint can make independent detection without any help of other vehicles, thus, it does not require to establish the credibility of neighboring nodes (trust relationship-free, lightweight);
3) Voiceprint is a fully distributed algorithm without any centralized control or support of RSU (infrastructure-free, fully distributed).

The rest of this paper is organized as follows. Section II reviews the related work of Sybil attack detection. Section III reveals several important observations from the real-world measurements that motivate our work. Section IV presents our proposed detection method in detail. Section V conducts simulations to evaluate our approach. Section VI carries out further experiments in a real DSRC testbed. Finally, Section VII draws the conclusion.

II. Related Work

Sybil attack is a very critical problem in distributed peer-to-peer systems. It was first introduced by Douceur [4] in the distributed storage system. Extensive works are done to detect the malicious node and Sybil nodes in these systems. The goal of these detection methods is to ensure each physical node is bound with a valid unique identity [5]. All these methods can be classified into three categories: (1) resource testing based [4][6], (2) trusted certification based [3], [7]–[12] and (3) physical measurement based mechanisms [13]–[19].

The resource testing based methods are in vain if the malicious node is equipped with sufficient resources and they usually bring extra overhead to the system when in testing. Trusted certification based methods are the most popular techniques to establish trust relationship among all nodes. This type of approaches usually uses the certificate authority, public key infrastructure, digital signatures and cryptographic algorithms to ensure the trustworthiness of each identity. They can find Sybil nodes at the beginning of the attack. However, this type of approaches usually requires a centralized trust party to issue digital signatures or certificates which cannot be easily applied in the initial stage of VANETs.

Due to the fast changing dynamic topology of VANETs and the high complexity of cryptographic algorithms, the lightweight and decentralized detection methods like position verification based methods are more suitable for the vehicular environment. These methods usually adopt some physical measurements such as Received Signal Strength Indicator (RSSI), Angle of Arrival (AoA) and Time Difference of Arrival (TDoA) to estimate the positions of the neighboring nodes. These measured values only depend on the hardware and physical environment that cannot be easily forged or modified by the malicious node.

Jin et. al used relative time measurements TDoA to determine the sender node's location and compared it with claimed coordinates of the sender node [13]. If they are different locations, the source node is judged as a Sybil node by receiver. TDoA-based method does not require time synchronization but it needs extra hardware (three receiving sensors mounted on different places of a vehicle). RSSI-based techniques, by contrast, are low-cost methods without any specialized hardware. They are on the basis of the idea that receiver can estimate distance from the sender according to RSSI values using theoretical radio propagation models. Demirbas et al. used RSSI-based localization method to detect Sybil nodes in a static Wireless Sensor Network (WSN) [14]. They adopted ratio of RSSIs from multiple receivers to overcome the time varying and unreliable nature of measured RSSI values. Wang et al. proposed a similar method by assuming a more realistic Rayleigh fading model [15]. Lv et al. proposed a Cooperative RSSI-based Sybil Detection (CRSD) scheme [16]. CRSD does not compute absolute positions, but relative distances among different nodes. Then, each node broadcasts the suspect identities with very closer distances. Finally, each node takes the largest intersection among all received groups as the detected Sybil nodes.

All above RSSI-based methods are decentralized techniques that each node runs the detection algorithm locally without the centralized infrastructure. However, these methods detect Sybil attack in a cooperative manner that each node needs the information from neighboring nodes, i.e., to get RSSI values observed by other nodes around to solve equations or compute the intersection of suspect groups. Therefore, the major problem in these methods is how to confirm the credibility and honesty of the neighboring nodes, since the Sybil nodes fabricated by the malicious node might send forged RSSI values to impede the detection. To avoid this problem, Bouassida et al. proposed an independent detection method based on RSSI [17]. In this scheme, the authors checked RSSI variations measured successively if they fall into a reasonable interval or not. The unreasonable nodes are labeled as Sybil nodes. But the authors only verify the proposed methods in a small scale testbed. Chen et al. proposed a centralized approach based on RSSI [18]. In this scheme, landmark as the trusted centralized party records all RSSI values transmitted by sensors and conducts a statistical testing for each two RSSI distributions. The nodes have similar RSSI distributions are considered as Sybil nodes. Xiao and Yu [20][19] proposed a cooperative detection method considering the trust relationship among all neighboring nodes. In this cooperative detection method, each vehicle first periodically broadcasts its identity and position as a claimer. After collecting enough information from witnesses (part of neighboring vehicles), one vehicle as a verifier estimates the position of all neighboring nodes accord-

ing to the received RSSI values and a predefined propagation model. To avoid some Sybil nodes provide forged location information, they assumed each vehicle can receive a position certification when passing through a RSU. And the witnesses only selected from the opposite traffic flow which has the issued position certification. According to this certification, this cooperative method can ensure that each node in the witness group is a trusted physical entity. However, it is not suitable for the sparse traffic and one-way roads.

The detailed comparisons of above RSSI-based methods as well as our proposed scheme are summarized in Table I.

TABLE I: Comparisons of RSSI-based detection methods

Methods	RPM	C/D	C/I	SoI	Mobility
Demirbas [14]	Free space	D	C	No	Static
Wang [15]	Rayleigh fading	D	C	No	Static
Lv [16]	Two-ray ground	D	C	No	Static
Bouassida [17]	Friis free space	D	I	No	Low Mobility
Chen [18]	Shadowing	C	-	Yes	Static
Xiao [20]	Shadowing	D	C	Yes	High Mobility
Yu [19]	Shadowing	D	C	Yes	High Mobility
Voiceprint	Model-free	D	I	No	High Mobility

Note: RPM: Radio Propagation Model; C/D: Centralized/Decentralized; C/I: Cooperative/Independent; SoI: Support of Infrastructure.

III. Measurements and Observations

As most of RSSI-based methods heavily rely on the assumed radio propagation models, we should first assess the effectiveness of such models in the real vehicular environment. In this Section, we conduct several real-world experiments using multiple vehicles equipped with DSRC radios in different scenarios.

A. Measurement Equipments

The experiment includes four vehicles that each one is equipped with an IEEE 802.11p compliant radio, namely the IWCU OBU4.2 produced by ITRI. The onboard equipments for each vehicle are composed of an IWCU OBU4.2 unit, a 5.9GHz antenna, a GPS module and a laptop which are shown in Figure 1.

IWCU OBU4.2 is a WAVE/DSRC communication device mounted in a vehicle. It is an embedded Linux machine (kernel 2.6.32) based on a 32 bits MIPS processor (Atheros AR7130) working at 300MHz. It has two Ethernet interfaces, a GPS connector and a DSRC radio based on the standard IEEE 802.11p-2010 [21]. IWCU OBU4.2 is connected to the 5.9GHz omni-directional antenna with a gain of 7dBi. The antenna is mounted on the roof of the vehicle. There is also a rooftop GPS receiver placed by the side of the antenna to log the vehicle's position. The IWCU OBU4.2 also connects to the laptop via an Ethernet interface, thus, the laptop can record the RSSI value of each successfully received packet. The details of measurement equipments are listed in Table II.

Fig. 1: Measurement equipments

TABLE II: Details of measurement equipments

Equipment	Details
Processor	Atheros AR7130 300MHz (MIPS 32bit)
DSRC radio	IEEE 802.11p, RX sensitivity: -95 dBm
Antenna	5.9GHz, 7dBi Omni
GPS module	50 channels, A-GPS support, sensitivity: -160 dBm, accuracy of time-pulse signal: 30ns (RMS), horizontal position accuracy: <2.5m (autonomous), <2.0m (SBAS)
Ethernet	10/100 Mbps (RJ45) port, full-duplex
TX Power	Max 32dBm (EIRP)
Channel width	10MHz/20MHz
Standards compliance	IEEE 802.11p-2010, IEEE 1609.2-v2-d9 3-2011-09, IEEE 1609.3-2010, IEEE 1609.4-2010

B. Measurement Scenarios

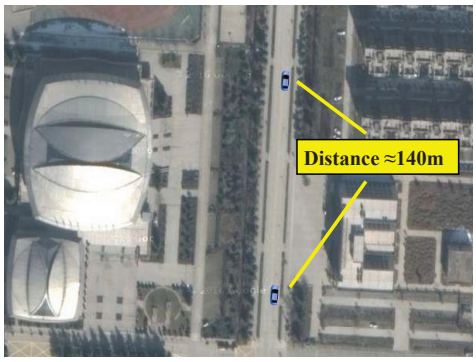
To assess the effectiveness of RSSI-based Sybil attack detection methods in VANETs, we conduct several experiments in different scenarios. Each vehicle adopts WAVE Short Message Protocol (WSMP) provided by IWCU OBU4.2 SDK software toolkit to send single-hop broadcast with its identity, GPS coordinates, direction and velocity. At receiver, the connected laptop records all received RSSI values via Ethernet. The basic communication parameter settings are shown in Table III.

TABLE III: Basic parameter settings

Parameter	Value
Transmission power	20dBm (EIRP)
Center carrier frequency	5.890MHz (CH 178 Control Channel)
Channel width	10MHz
Data rate	3Mbps
Packet size	500Bytes

(a) Campus (b) Rural area (c) Urban area

Fig. 3: Scenario 2 (Two vehicles communicate in different environments)



(a) Remain stationary

(b) Keep moving

Fig. 2: Scenario 1 (Two vehicles communicate in the campus)

***Scenario 1:* Two vehicles communicate in the campus.** This measurement is carried out in the campus. The scenario is shown in Figure 2a. Two vehicles keep stationary with each other at a distance about 140m. The sender broadcasts its information 10 packets per second, and the receiver records RSSI values from the sender. We conduct this experiment two times at different time periods, each one lasts 10mins. Another measurement is also carried out in the campus, but vehicles move around the schoolyard as shown in Figure 2b. The speed of vehicle approximately is 10-15 km/h.

***Scenario 2:* Two vehicles communicate in different environments.** In this case, we collect data from different areas including campus, rural area and urban area to illustrate the impact of the environment to the propagation models. Figure 3 gives snapshots of different environments.

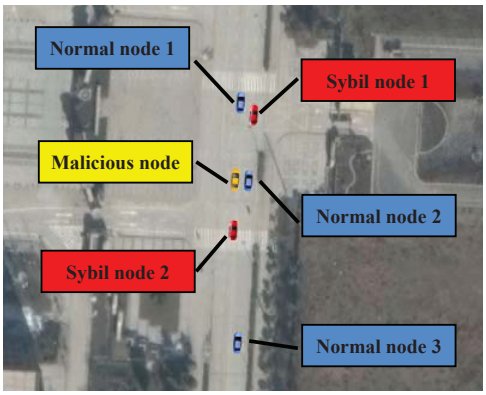


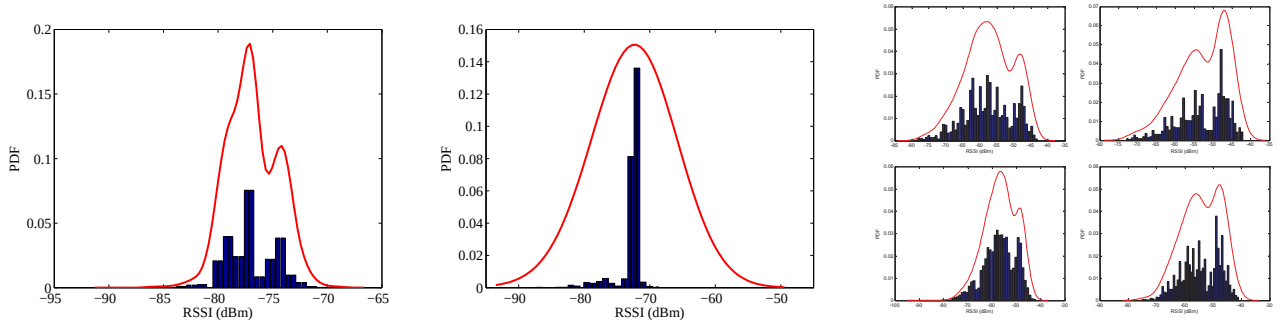
Fig. 4: Scenario 3 (Four vehicles simulate the Sybil attack)

***Scenario 3:* Four vehicles simulate the Sybil attack.** In this scenario, we simulate the Sybil attack with four vehicles as shown in Figure 4. There are three normal nodes (marked in blue) and one malicious node (marked in yellow) with motion at the same direction. The malicious node generates two fake identities i.e. Sybil nodes (marked in red) at false locations. During the experiment, the normal node 1 and 3 are ahead of and behind the malicious node respectively. The normal node 2 keeps moving with the malicious node side by side. The normal node 1 and 3 record all RSSI time series from the malicious node, the fabricated Sybil nodes 1 and 2 and the normal node 2.

C. Observations

We plot the RSSI distributions of ***Scenario 1*** in Figure 5. Figure 5a and 5b show the RSSI values recorded when two vehicles keep stationary in two different periods. Each distribution contains 6000 samples. The mean and standard deviation of two distributions are (-76.8600 dBm, 2.3266 dBm) and (-72.5390 dBm, 0.7654 dBm) respectively. According to Free Space Path Loss (FSPL) model and Two-Ray Ground Propagation (TRGP) model assumed in [14] and [16], the average distances between two vehicles are estimated to be 281.5m (FSPL in the first period) and 171.2m (FSPL in the second period), 263.9m (TRGP in the first period) and 205.8m (TRGP in the second period), respectively. Comparing to the real distance 140m, the estimated values are quite inaccurate.

Figure 5c gives four RSSI distributions of different segments randomly selected from ***Scenario 1*** that two vehicles move



(a) 2016/05/09 14:31:50-14:41:50 (stationary) (b) 2016/05/09 15:06:14-15:16:14 (stationary) (c) Four segments selected randomly (moving)

Fig. 5: RSSI distributions

around the campus. Each segment has 1 min long, thus, contains 600 RSSI samples. In some RSSI-based position verification methods [18][20][19], they assume the RSSI values follow the normal distribution according to the shadowing model. Actually, the RSSI values barely show the normal distribution in VANETs, especially when the vehicle keeps moving constantly.

From the results obtained by ***Scenario 1***, we get the first observation.

Observation 1: **Temporal variation of the channel in VANETs.** The channel quality changes over time in VANETs. Therefore, a predefined propagation model might lead to significant errors in position estimation or make false statistic testing based on the wrong assumption of RSSI distribution.

The empirical dual-slope piecewise linear model is widely used in VANETs [22] as shown in Equation 1.

$$P_r(d) = \begin{cases} P(d_o) - 10\gamma_1 log_{10}\left(d/d_o\right) + X_{\sigma 1}, d_0 \leq d \leq d_c \\ P(d_o) - 10\gamma_1 log_{10}\left(d_c/d_o\right) - 10\gamma_2 log_{10}\left(d/d_c\right) \\ +X_{\sigma 2}, d > d_c \end{cases} \tag{1}$$

where $P(d_o)$ is the known signal strength which is calculated using the free space path loss model at the reference distance d_o. γ_1 and γ_2 are the path loss exponents. d_c is the critical distance. $X_{\sigma 1}$ and $X_{\sigma 2}$ are zero-mean, normally distributed random variables with standard deviation σ_1 and σ_2 respectively.

Three data sets measured from ***Scenario 2*** in the campus, rural area and urban area are regression-fitted using least square method to obtain parameters of the model. We list fit parameters of the campus, the rural area and the urban area in Table IV.

Due to the sparsely distributed vehicles in campus and rural area, there is a dominant Line-Of-Sight (LOS) path between receiver and sender. Their breakpoint distances (d_c) are much longer than the value in the urban area since more densely distributed obstacles like vehicles and pedestrians on the road cause severe signal distortion at receivers in Non-Line-Of-Sight (NLOS) conditions. In addition, the signal attenuation in the campus environment seems much better than the rural area

TABLE IV: Fit parameters of the empirical model

Parameter	Value		
	Campus	Rural area	Urban area
d_o	1m	1m	1m
d_c	218m	182m	102m
γ_1	1.66	1.89	2.56
γ_2	5.53	5.86	6.34
$X_{\sigma 1}$	2.8dB	3.1dB	3.9dB
$X_{\sigma 2}$	3.2dB	3.6dB	5.2dB

because the effects of reflection and shadowing are probably more serious by those high and dense wayside trees (shown in Figure 3a and 3b).

Then, we have the second observation.

Observation 2: **Spatial variation of the channel in VANETs.** The channel conditions are not the same in different areas considering complex reflection, refraction, diffraction and multi-path effects caused by buildings, trees and other obstacles. For a predefined propagation model, it requires to set different parameters for different environments. However, it is very hard for a vehicle to sense the environment dynamically, and then to determine optimal parameters.

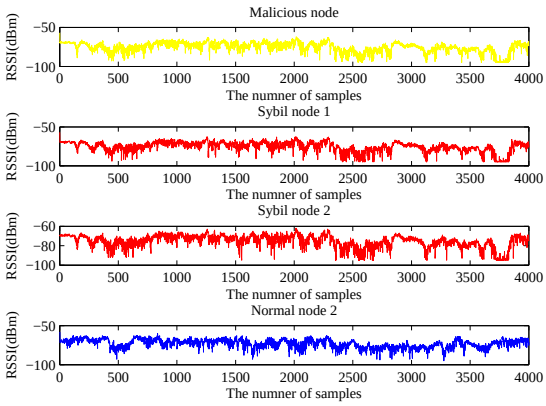


Fig. 6: RSSI time series recorded by the normal node 1

The Figure 6 and Figure 7 give the RSSI time series recorded by the normal node 1 and 3. Then, we have a

significant and interesting observation.

Observation 3: **Similar patterns of RSSI time series.** The RSSI time series of the malicious node and the Sybil nodes have very similar patterns. The series of the malicious node and the normal node 2 are similar, but still have some differences even if they always keep very close distance (2.75-3.25m) during the motion.

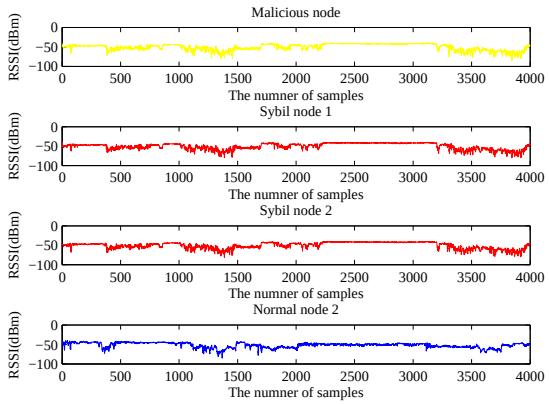


Fig. 7: RSSI time series recorded by the normal node 3

With this observation, we propose the Voiceprint method to transfer the Sybil attack detection into comparing the similarity between two time series which are like to recognize and differentiate the vehicular speech.

IV. The Proposed Scheme Voiceprint

In this section, we first describe the attack model and assumptions. Then, we introduce the similarity measures for time series. Finally, we give the detailed Sybil attack detection algorithm based on vehicular voiceprint.

A. Attack Model and Assumptions

In this paper, we focus on the simultaneous Sybil attack that each Sybil attacker concurrently creates multiple fake identities to disrupt normal functionalities of VANETs. Figure 8 shows a typical Sybil attack scenario in the highway environment. From this figure, the legitimate vehicle bounded with unique valid identity is referring to the normal node (marked in blue). The physical vehicle uses multiple forged identities is called malicious node (marked in yellow), and the claimed virtual identities are Sybil nodes (marked in red).

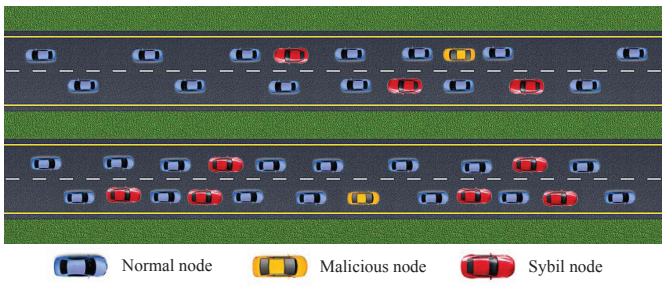


Fig. 8: An example of Sybil attack scenario in VANETs)

Assumption 1: We assume there may be several Sybil attackers in VANETs, but those malicious nodes do not collude to launch Sybil attacks. The attacker only creates new identities rather than stealing other vehicle's identity.

Assumption 2: In VANETs, every vehicle is equipped with only one DSRC radio and one 5.9GHz antenna that is able to broadcast its own information on CCH periodically. The OBUs may have heterogeneous configurations, but their broadcast frequency is constant (10Hz) according to the DSRC protocol. That means the malicious node will simulate and broadcast all Sybil nodes' information with the same frequency on CCH.

Assumption 3: Unlike most previous works considering the same transmission power for each node, we relax this assumption and allow different transmission power settings. The normal nodes may have different default TX Power settings or different antenna gains (heterogeneous OBUs). The malicious node may increase or decrease initial TX Power for each fabricated Sybil node. However, the TX Power remains constant during the transmission.

B. Similarity Measures for Time Series

A time series is a sequence of data points successively collected over time. With the ***Observation 3*** obtained from the real-world experiments, we find that the RSSI time series of Sybil nodes have very similar patterns. Therefore, we detect Sybil attack by measuring the similarity between two RSSI time series based on this important observation. Here, similarity is an absolute value computed by comparing or matching the resemblances between two series. Commonly, a distance function $D(X,Y)$ is defined to represent the similarity between time series X and Y.

Since time series similarity measures have been a major topic in data mining research for decades, many distance functions have been proposed in this domain. The classical form to compute the similarity is L_p norm as follows:

$$D_{Lp}(X,Y) = \left(\sum_{1}^{N} (x_i - y_j)^p \right)^{\frac{1}{p}} \tag{2}$$

where p is a positive integer, N is the length of two time series, x_i and y_i are the i^{th} element of time series of X and Y, respectively. When p equals to 2, it is the well-known Euclidean distance.

Another commonly used distance is called Dynamic Time Warping (DTW). DTW adopts dynamic programming technique to determine the best matching between two time series by warping the series in the temporal domain. Given two time series with different length N and M, $X_N(x_1, x_2, ...x_i, ...x_N)$ and $Y_M(y_1, y_2, ...y_j, ...y_M)$, DTW first establishes an N-by-M cost matrix C containing distance $c_{i,j}$ between each pair of points x_i and y_j. The cost $c_{i,j}$ usually uses Euclidean distance as:

$$c_{i,j} = (x_i - y_j)^2 \tag{3}$$

Then, DTW computes the minimum accumulated cost $D_{i,j}$ for each pairwise matching (i,j) between two series recursively by:

$$D_{i,j}=c_{i,j}+\min\{D_{i-1,j},D_{i,j-1},D_{i-1,j-1}\} \quad (4)$$

where $D_{0,0}$ is set to be 0 initially and other value in the accumulated cost matrix D are initialized to ∞.

After that, DTW constructs a optimal warp path $W = w_1, w_2, ..w_k, ...w_K$ ($w_k = (i,j)$ means the i^{th} element of X is matched to the j^{th} element of Y) with the minimum total accumulated cost. The optimal warp path W must start from $w_1 = (1,1)$ to $w_K = (N,M)$ to ensure all points of both series are matched. In addition, the warp path should also satisfy the monotonicity constraint which is defined as:

$$\begin{aligned} &IF\ w_k=(i,j)\,, w_{k+1}=(i',j')\,; \\ &THEN\ i\le i'\le i+1, j\le j'\le j+1 \end{aligned} \quad (5)$$

Finally, the DTW distance is measured as the total accumulated cost:

$$D_{DTW}(X,Y)=D_{N,M} \quad (6)$$

Here we give a simple example to illustrate how to compute the DTW distance as shown in Figure 9. The two time series are $X = \{1,1,4,1,1\}$ and $Y = \{2,2,2,4,2,2\}$ with total length of $N = 5$ and $M = 6$ respectively. The DTW distance is measured as 9 in this case.

Wang et al. make an extensive comparison for 13 different similarity measures using 38 data sets from various application domains [23]. The main conclusion drawn by the study is that the DTW distance is superior to the other newly proposed methods considering the accuracy in the vast majority of cases, and the well-established Euclidean distance is also a robust, simple, generic and efficient way to measure the similarity of time series. From above introduction of these two distances, we find that the Euclidean distance matches in the point-to-point way, which requires two time series having the same length. DTW distance overcomes this limitation that can tolerate the shifting, scaling and warping of series in the temporal domain, which is widely used in speech recognition to cope with different speeking speeds. Considering that packet loss often occurs in VANETs, we cannot always get two RSSI series with exactly the same length. Therefore, we use DTW distance to measure the similarity of RSSI time series like vehicular voiceprint recognition. However, despite the accuracy of DTW scheme, it has $O(N^2)$ time complexity in general since it should fill all cells in the NM cost matrix. So, we adopt FastDTW [24] in this paper. FastDTW speeds up DTW distance measure by adding constraints and data abstraction to limit the cost cell evaluation. Then, it achieves $O(N)$ time complexity while has only 1% loss of accuracy, which can meet requirements of the Sybil attack detection.

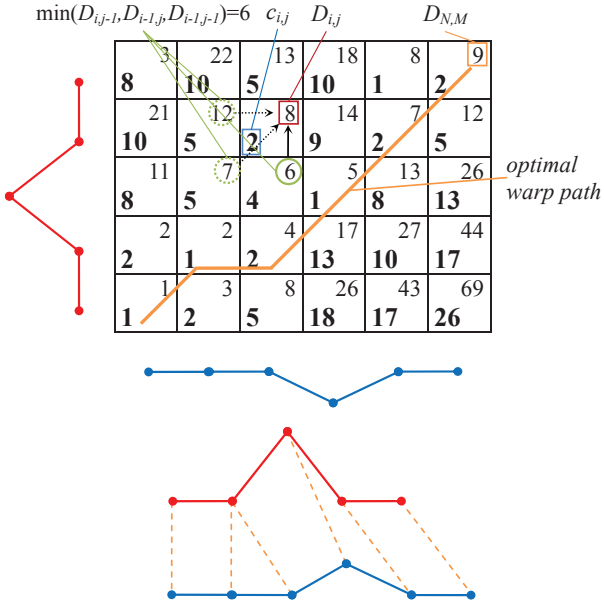


Fig. 9: A simple example of computing DTW distance

C. Proposed Detection Methodology

In this subsection, we present our Sybil attack detection method, Voiceprint, based on similarity measuring of RSSI time series. Voiceprint does not rely on any predefined radio propagation model, and it also does not require the support of centralized infrastructures (RSUs or base stations). Each vehicle conducts independent detection locally without establishing trust relationship among neighboring vehicles.

There are three phases in Voiceprint, collection, comparison and confirmation.

1) Collection: According to ***Assumption 2***, each vehicle mounts a DSRC compliant OBU to broadcast its basic information periodically on CCH. Generally, the basic information includes vehicle's identity, location, velocity, acceleration, direction and etc. Vehicles exchange the information from each other for safety-related applications. All neighboring nodes could receive these messages and measure the RSSI value for each successfully received packet. In the collection phase, one vehicle monitors the CCH and records all the latest messages within a constant interval (this interval is called observation time in this paper). Actually, for each packet, Voiceprint only needs to store a 2-tuple $< ID, RSSI >$, and then generates RSSI time series for each received IDs. Here, RSSI time series of vehicle i is denoted by $RSSI_i$.

2) Comparison: After a sufficient observation time for collection, each vehicle has enough data to make comparison between every pair wise RSSI time series. As aforementioned, we use DTW distance to measure the similarity of RSSI time series. However, based on ***Assumption 3***, if the malicious node deliberately increase or decrease the initial TX Power for different Sybil nodes, the similarity of RSSI time series among malicious node and Sybil nodes can be simply broken, because the relative distance between aligned points is enlarged. To solve this problem, we conduct a data preprocessing before the comparison which normalizes every RSSI time series by an enhanced Z-score normalization:

$$RSSI_i' = \frac{RSSI_i - \mu}{3\sigma} \quad (7)$$

where μ and σ are the mean and standard deviation of $RSSI_i$ respectively. This normalization makes 99.7% values fall into the range of (-1, 1). In this normalization, the whole shape and structure of RSSI time series cannot be changed, but the relative distances among Sybil nodes' RSSI series by spoofed transmission power are perfectly eliminated.

After data preprocessing, we compare every pairwise RSSI time series and measure the DTW distance. Then, we conduct a postprocessing for obtained DTW distances to normalize all values into the range of [0, 1] using min-max normalization:

$$D'_{DTW\,i,j} = \frac{D_{DTW\,i,j} - D_{DTW\,\min}}{D_{DTW\,\max} - D_{DTW\,\min}} \tag{8}$$

where $D_{DTW\,\min}$ and $D_{DTW\,\max}$ are the minimum and maximum values of all DTW distances respectively.

3) Confirmation: In the comparison process, each vehicle can get a group of DTW distances for all neighboring vehicles. Based on ***Observation 3***, DTW distances among all Sybil nodes should be very small that are closer to 0, while DTW distances between Sybil nodes and normal nodes or among all normal nodes should be much larger. However, from extensive simulations in Section V, we find that DTW distances are easily distinguishable in the low vehicle density, but have a small overlap when the density increases. There are two reasons for this phenomenon. First, when the traffic gets jammed, the average space between two vehicles is shorten, thus, the RSSI time series of malicious node and some normal nodes nearby are also very similar. Second, with the increasing traffic density, the number of nodes in VANETs is also increased. This leads to severe channel collisions that cause a lot of packet losses in the whole network. Thus, the similarity of RSSI time series among all Sybil nodes is decreased. The DTW distance overlap will reduce the detection rate and increase the false positive rate when the traffic density increases if we set a constant threshold. To deal with this problem, we just think of the threshold as a function of density. And the determination of the threshold can be transformed into a binary classification problem that finds the optimal decision boundary (actually a line in the two-dimensional condition) in the density-DTW distance plane. There are many methods such as perceptrons algorithm, linear classifier, logistic regression and support vector machines proposed to do classification in machine learning. In this paper, we use the Linear Discriminant Analysis (LDA) to determine the threshold. For an estimated density den and a measured DTW distance $D_{DTW\,i,j}$ between node i and node j, if $D_{DTW\,i,j} \leq k \cdot den + b$ is satisfied, the nodes i and j are detected as the Sybil nodes. Here k and b is the slope and intercept of the decision boundary respectively. These parameters can be obtained by training based on our simulation or experiment data. Each vehicle can estimate traffic density by:

$$den = \frac{N_{normal}}{2Dist_{\max}} \tag{9}$$

where N_{normal} is the number of normal nodes it can hear within the density estimation period (one vehicle can only use the total number of received nodes in the first estimation since it cannot recognize the legitimate ones at the beginning). $Dist_{max}$ is the maximum transmission range.

The procedure of Voiceprint is presented in Algorithm 1.

Algorithm 1 Voiceprint

Input:
$RSSI_n$: RSSI time series
ID_n: Corresponding IDs
den: Estimated traffic density
k: Slope of the decision boundary
b: Intercept of the decision boundary
Output:
$SybilIDs$: Suspect IDs of Sybil nodes

for $i = 1$ to n **do**
  $RSSI_i \leftarrow$ Z-score-normalization($RSSI_i$)
end for
for $i = 1$ to $n-1$ **do**
  for $i = 2$ to n **do**
    if $i<j$ **then**
      $D_{DTW\,i,j} \leftarrow$ FastDTW($RSSI_i$, $RSSI_j$)
    end if
  end for
end for
$D_{DTW} \leftarrow$ Min-max-normalization(D_{DTW})
for $i = 1$ to $n-1$ **do**
  for $i = 2$ to n **do**
    if $i<j$ **then**
      if $D_{DTW\,i,j} \leq k \cdot den + b$ **then**
        $SybilIDs$ = AddingIDs(i, j)
      end if
    end if
  end for
end for
return $SybilIDs$

V. Simulation Evaluation

In this section, we evaluate the performance of the proposed Voiceprint by NS2 simulations.

A. Simulation Setup

We conduct our simulation in the NS-2.34 simulator and use the empirical propagation model given in equation (1) [22]. To prove that Voiceprint does not depend on any predefined propagation model, we set a timer in NS2 and modify the parameters of the propagation model periodically. The simulation scenario is a 2km bi-directional highway with 2 lanes in each direction as shown in Figure 10 (Lane width is 3.6m). Vehicles re-enter the highway at the beginning of the other direction when they arrive at the end of one direction. For an individual simulation run, we randomly set 5% vehicles as malicious nodes, and each one generates 3-6 Sybil nodes. All nodes broadcast 10 packets per second on CCH, but the malicious node should send $10n$ packets if it fabricates n fake identities. The initial transmission power can be randomly selected from 17-23dBm for each node, but remains constant during the simulation.

We adopt a continuous-time stochastic mobility model to simulate vehicle motion. In this model, each vehicle's movement is divided into a sequence of random time intervals called mobility epochs. The epoch lengths are identically, independently distributed (i.i.d.) exponentially with mean $1/\lambda_e$.

TABLE V: Default parameter settings

Parameter	Value
Highway length	2km
Lanes	4
Lane width	3.6m
Density	10-100 vhls/km
Density estimate period	10s
Vehicle number	20-200
Model change period	30s
Frequency	5.9GHz
Bandwidth	10MHz
Transmission Power	17-23dBm
Date rate	3Mbps
Packet size	500Bytes
Packet generation rate	10Hz
Slot time	$13\mu s$
SIFS	$32\mu s$
Mobility epoch rate (λ_e)	$0.2s^{-1}$
Average speed (μ_v)	25m/s
Standard deviation of the speed (σ_v)	5m/s
Observation time	20s
Detection period	20s
Simulation time	100s

During each epoch, the vehicle moves at a constant speed which is an i.i.d. normal distributed random variable with mean μ_v and the standard deviation σ_v. The default parameters are given in Table V.

B. Metrics and Threshold

1) Metrics: We consider two main metrics to evaluate our scheme, i.e., detection rate (DR) and false positive rate (FPR). For a single normal node and one detection period, detection rate is the proportion of detected suspect nodes to the total number of illegitimate nodes within all its neighboring vehicles. False positive rate is the percent of normal nodes that are incorrectly detected as forged ones. For a single normal node i, it receives multiple packets from N_i different nodes during the observation time. Assume that in the k^{th} detection period, there are $N_{i,k}^n$ legitimate nodes, $N_{i,k}^m$ malicious nodes and N_j^s Sybil nodes generated by the j^{th} malicious node. If it correctly detects $N_{T,k}$ fabricated nodes and wrongly identifies $N_{F,k}$ normal nodes, then, the detection rate and false positive rate for node i in the k^{th} detection period are defined as follows:

$$DR_{i,k} = \frac{N_{T,k}}{N_{i,k}^m + \sum_{j=1}^{N_{i,k}^m} N_j^s} \tag{10}$$

$$FPR_{i,k} = \frac{N_{F,k}}{N_{i,k}^n} \tag{11}$$

Assume we have totally N_n normal nodes and each normal node detects K times during the simulation. Then, the average detection rate and average false positive rate can be calculated as follows:

$$\overline{DR} = \frac{1}{N_n K} \sum_{i=1}^{N_n} \sum_{k=1}^{K} DR_{i,k} \tag{12}$$

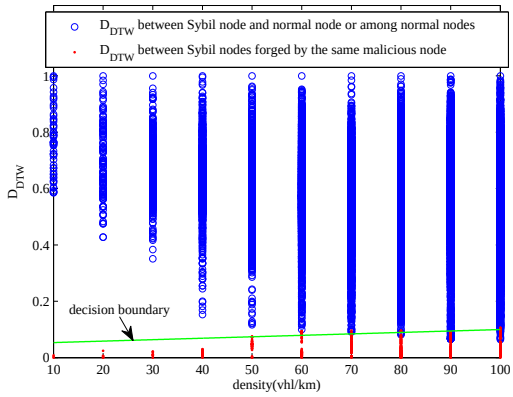


Fig. 10: The optimal decision boundary determined by LDA

$$\overline{FPR} = \frac{1}{N_n K} \sum_{i=1}^{N_n} \sum_{k=1}^{K} FPR_{i,k} \tag{13}$$

In the simulation, we use the average detection rate and average false positive rate to evaluate the performance of Voiceprint.

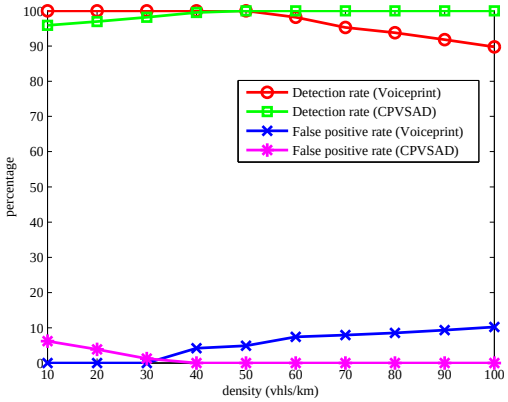
2) Threshold: In this paper, we leverage LDA to find the decision boundary. Each node can tune the threshold according to the estimated traffic density. We first conduct several simulations for different traffic densities (5 simulation runs at each density) and record all measured DTW distances. Then, we use these DTW distances as the training data to compute the optimal decision boundary, i.e. to determine the slope k and intercept b for the divider line. The results are shown in Figure 10. The blue cycle denotes the DTW distance between the Sybil node and the normal node or between two normal nodes. The red dot is the DTW distance between two Sybil nodes forged by the same malicious node. After training, the parameters of k and b are set to be 0.00054 and 0.0483 respectively.

C. Comparison and Results Analysis

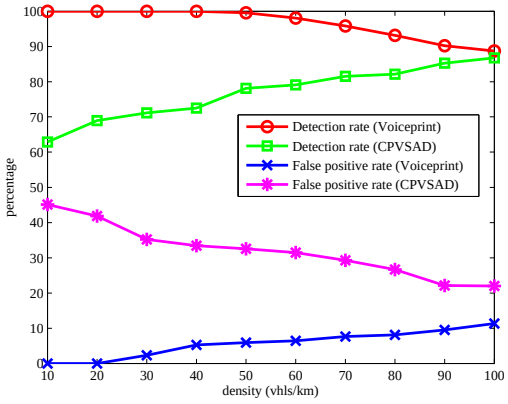
In our simulations, we compare the Voiceprint to the Cooperative Position Verification based Sybil Attack Detection (CPVSAD) scheme proposed in [19]. The observation time of CPVSAD is 10s, and the standard deviation of the predefined shadowing model is 3.9dB. The significant level is set to be 0.05.

Figure 11a shows the detection rate and false positive rate of two methods without propagation model change. The standard deviation σ_1 and σ_2 are both set to be 3.9dB during the simulation. From Figure 12a, we see that both Voiceprint and CPVSAD can achieve to 90% level detection rate and low false positive rate under 10%. The performance of CPVSAD improved with the increasing vehicle density, while Voiceprint has the opposite trend. This is because CPVSAD is the cooperative detection method. One vehicle conducts Sybil attack detection which not only uses the RSSI values observed by its own, but also adopts information received from neighboring

vehicles. With the increasing traffic density, each vehicle could collect more information from other vehicles nearby. However, to ensure all information to be correct, CPVSAD requires support of RSUs to establish trust relationship among neighboring nodes. Since Voiceprint is the independent detection scheme, one vehicle only uses RSSI time series observed locally. Therefore, with the increasing traffic density, the severe packet losses lead to less information obtained by each vehicle, thus, reduce the detection rate. Moreover, the dense traffic means the shorter average space among vehicles. Vehicles cannot easily distinguish malicious nodes from the normal nodes nearby that results in the increasing false positive rate.



(a) Without propagation model change



(b) With propagation model change

Fig. 11: Detection rate and false positive rate

Figure 11b gives the results with propagation model change. The model parameters are modified periodically during the simulation. We can observe that the performance of CPVSAD drops rapidly, while Voiceprint is almost immune to the change. This is because CPVSAD should conduct the statistical testing according to the predefined model parameters. It is hardly to get accurate results if the predefined parameter changes. However, Voiceprint does not rely on any propagation models. Thus, it is widely applicable for different environments and complex conditions.

VI. Field Test

In this section, we evaluate the performance of the proposed Voiceprint in the real-world field test.

A. Experiment Setup

In this field test, we use four vehicles equipped with DSRC radios and embedded with the Voiceprint application. We conduct a series of experiments under campus, rural area, urban area and highway environments shown in Figure 12. There are one malicious node (ID = 1) and three normal nodes (IDs = 2, 3 and 4), and the malicious node generates two Sybil nodes with two fake identities (IDs = 101 and 102). The setup is same to Figure 4 given in Section III-B ***Scenario 3***. Normal node 2 moves as close as possible to the malicious node during the test. The initial transmitted powers of all physical nodes (nodes 1-4) are 20dBm. The initial transmitted powers of Sybil node 101 and 102 are 23dBm and 17dBm respectively. The observation time is 20s and detection period is 1min. Since there are only four vehicles in the network, we just set the constant threshold to be $k = 0.05046$ at the traffic density of 4vhls/km.

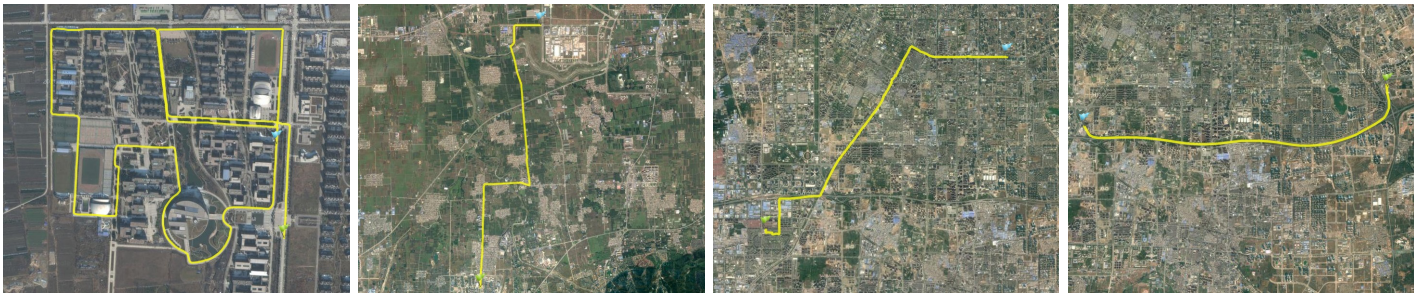
B. Results and Analysis

The durations of tests in different areas are 13min21s, 22min 40s, 34min46s and 11min12s respectively. Thus, the detections are conducted 14, 23, 35 and 11 times in campus, rural area, urban area and highway correspondingly. We store all measured DTW distances compared with the threshold. Figure 13 plots the results recorded by the normal node 3. Here, DTW(a, b) means the measured DTW distance of RSSI time series received from node a and b.

From Figure 13, we can find that the detection rate is 100% in all scenarios and the false positive rate is 0.95%, only one time that the normal node 2 is incorrectly detected as the Sybil node. In order to find out the cause of this false detection, we check the GPS trace and further analyze statuses, distances and speeds of all vehicles.

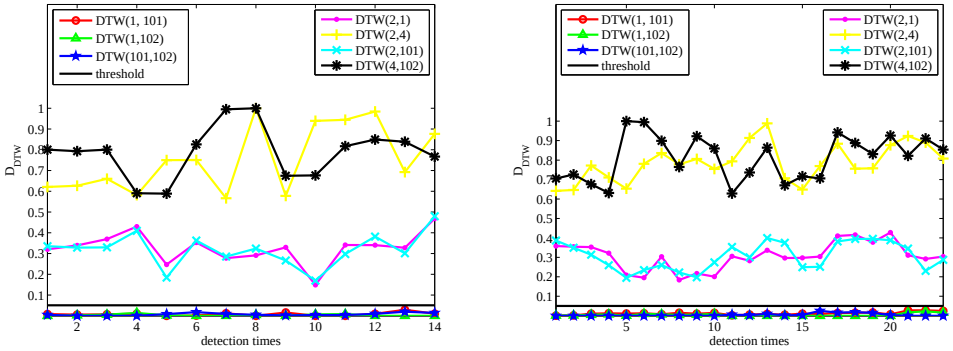
According to GPS traces of malicious node 1, normal node 2 and 3, we notice that all these nodes remain stationary without mobility at this detection period. Based on the locations of nodes on the map shown in Figure 14, we find that the false detection occurs at an intersection. The reasonable explanation is that all vehicles stop at the intersection waiting for a red light. The measured distances between each pair of nodes are 3.8m (node 1 and 2), 198.9m (node 1 and 3) and 195.2m (node 2 and 3) respectively. Therefore, the normal node 3 cannot distinguish two RSSI time series from malicious node 1 and normal node 2 since all nodes remain stationary in this detection period, which leads to very similar signal patterns between node 1 and 2 (Notice that node 3 is very far away from these two nodes, most of RSSI values are -95dBm which reaches the RX Sensitivity of our radio).

We also estimate the computational complexity of Voiceprint. The observation time is 20s and the transmission frequency is 10HZ. Hence, there are at most 200 RSSI values for each time series. The measured average time of comparing

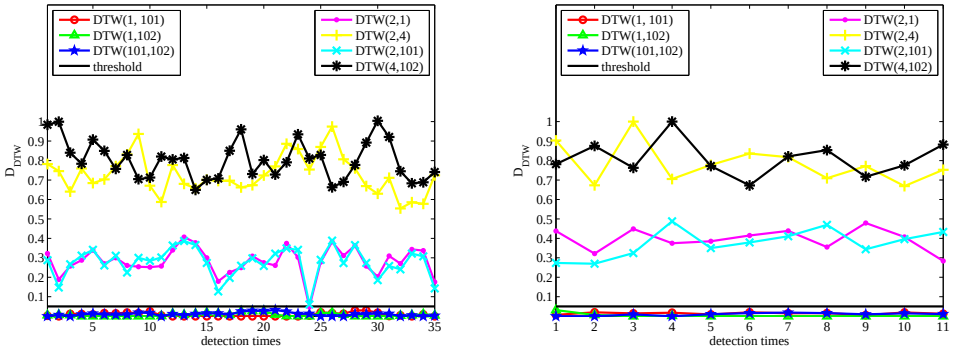


(a) Campus (3.5km) (b) Rural area (11.2km) (c) Urban area (15.4km) (d) Highway (12.7km)

Fig. 12: The routes in different areas



(a) Campus (b) Rural area



(c) Urban area (d) Highway

Fig. 13: The measured DTW distances in different areas

two RSSI time series is 0.1995ms. If there are 80 neighboring vehicles (suppose an extreme case in which traffic density is 200vhls/km and the transmission range is up to 400m), the total computing time is only about 630ms. This time complexity is affordable for our Sybil attack detection scheme.

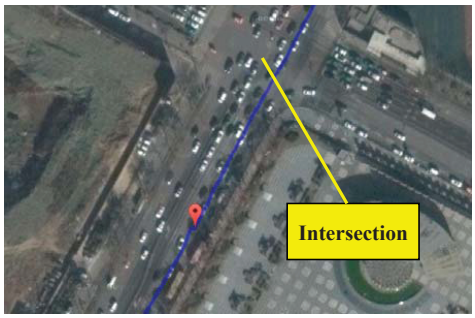
From the real-world experiments, we show that the proposed Voiceprint is suitable for different areas, especially in the rural and highway environments where vehicles can keep moving without long time stopping. Although, when vehicles stay stationary during the detection period, it may result in false alarms (some complex conditions in the urban area such as red light and traffic jam), Voiceprint is still an effective method considering the cost, complexity and performance. We suggest making a final determination of the Sybil node after several detection periods so as to reduce the false positive rate.

VII. Conclusion

In this paper, we proposed a RSSI-based detection method, Voiceprint, against Sybil attacks in VANETs. The motivation of implementing the Voiceprint is based on our observation that the RSSI time series have very similar patterns among Sybil nodes and malicious attacker node. Voiceprint does not depend on any radio propagation model that makes it widely suitable for various environments (model-free, widely applicable). In addition, it conducts independent detection that does not require establishing the trust relationship of neighbor-

Time	Latitude	Longitude	Elevation	Heading	Speed			
2016/05/09-04:54:15.220			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:15.321			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:15.423			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:15.524			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:15.626			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:15.727			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:15.829			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:15.930			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.032			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.133			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.235			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.336			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.438			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.540			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.642			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.743			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.845			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:16.946			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:17.048			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:17.150			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:17.251			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:17.353			402.100000	27.400000	0.000000	2191	0	0.000000
2016/05/09-04:54:17.454			402.100000	27.400000	0.000000	2191	0	0.000000

(a) The GPS trace of the normal node 2



(b) The location of the false detection

Fig. 14: Analysis of the false detection

ing nodes (trust relationship-free, lightweight). Furthermore, Voiceprint does not need the support of the centralized nodes such as base stations or RSUs (infrastructure-free, full-distributed). The simulation and field test results illustrate the effectiveness of Voiceprint.

We will continue our work on several directions to extend Voiceprint. First, comparing to some cooperative detection methods, Voiceprint needs longer observation time to collect more RSSI values since it only uses the local information. As the maximum safety message rate defined in DSRC on CCH is 10Hz, each vehicle can only receive at most 10 packets, i.e., 10 RSSI values per second from one neighboring vehicle. In future work, we will take the Service Channel (SCH) into account. Since there is no strict restriction of beacon rate for SCH, we can increase the beacon rate and broadcast the samples from SCH much quicker. Second, as same as all RSSI-based methods, Voiceprint cannot identify the malicious node if it adopts power control. We will conduct more real-world experiments to extract other features or other measurable parameters to prevent smart attacks with power control.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China (61502394 and 61572403), the Fundamental Research Funds for the Central Universities (3102015JSJ0002).

REFERENCES

[1] "Vehicle Safety Communications Project: Task 3 Final Report: Identify Intelligent Vehicle Safety Applications Enabled by DSRC," *Vehicle Safety Communications Consortium*, 2005.

[2] D. Kushwaha, P. K. Shukla, and R. Baraskar, "A Survey on Sybil Attack in Vehicular Ad-hoc Network," *International Journal of Computer Applications*, vol. 98, no. 15, pp. 31–36, 2014.

[3] S. Park, B. Aslam, D. Turgut, and C. C. Zou, "Defense against Sybil Attack in Vehicular Ad Hoc Network based on Roadside Unit Support," in *Proc. MILCOM*, 2009, pp. 1–7.

[4] J. R. Douceur, "The Sybil Attack," in *Proc. the 1st International Workshop on Peer-to-Peer Systems*, 2002, pp. 251–260.

[5] C. Kumar Karn and C. Prakash Gupta, "A Survey on VANETs Security Attacks and Sybil Attack Detection," *International Journal of Sensors, Wireless Communications and Control*, vol. 6, no. 1, pp. 45–62, 2016.

[6] J. Newsome, E. Shi, D. Song, and A. Perrig, "The Sybil Attack in Sensor Networks: Analysis Defenses," in *Proc. the 3rd International Symposium on Information Processing in Sensor Networks*, 2004, pp. 259–268.

[7] M. Raya, P. Papadimitratos, and J. p. Hubaux, "Securing Vehicular Communications," *IEEE Wireless Communications*, vol. 13, no. 5, pp. 8–15, 2006.

[8] C. Chen, X. Wang, W. Han, and B. Zang, "A Robust Detection of the Sybil Attack in Urban VANETs," in *Proc. IEEE ICDCS Workshops 2009*, 2009, pp. 270–276.

[9] S. Chang, Y. Qi, H. Zhu, J. Zhao, and X. Shen, "Footprint: Detecting Sybil Attacks in Urban Vehicular Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 6, pp. 1103–1114, 2012.

[10] T. Zhou, R. R. Choudhury, P. Ning, and K. Chakrabarty, "P2DAP-Sybil Attacks Detection in Vehicular Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 582–594, 2011.

[11] K. Mekliche and S. Moussaoui, "L-P2DSA: Location-based Privacy-Preserving Detection of Sybil Attacks," in *Proc. the 11th International Symposium on Programming and Systems*, 2013, pp. 187–192.

[12] H. Rasheed and O. Heekuck, "On Secure and Privacy-Aware Sybil Attack Detection in Vehicular Communications," *Wireless Personal Communications*, vol. 77, no. 4, pp. 2649–2673, 2014.

[13] D. Jin and J. Song, "A traffic flow theory aided physical measurement-based sybil nodes detection mechanism in vehicular ad-hoc networks," in *Proc. the 13th IEEE/ACIS ICIS*, 2014, pp. 281–286.

[14] M. Demirbas and Y. Song, "An RSSI-based Scheme for Sybil Attack Detection in Wireless Sensor Networks," in *Proc. IEEE WOWMOM*, 2006, pp. 566–570.

[15] J. Wang, G. Yang, Y. Sun, and S. Chen, "Sybil Attack Detection Based on RSSI for Wireless Sensor Network," in *Proc. WiCOM*, 2007, pp. 2684–2687.

[16] S. Lv, X. Wang, X. Zhao, and X. Zhou, "Detecting the Sybil Attack Cooperatively in Wireless Sensor Networks," in *Proc. CIS*, vol. 1, 2008, pp. 442–446.

[17] M. Bouassida, G. Guette, M. Shawky, and B. Ducourthial, "Sybil Nodes Detection Based on Received Signal Strength Variations within VANET," *International Journal of Network Security*, vol. 9, no. 1, pp. 22–32, 2009.

[18] Y. Chen, J. Yang, W. Trappe, and R. P. Martin, "Detecting and Localizing Identity-Based Attacks in Wireless and Sensor Networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2418–2434, 2010.

[19] B. Yu, C. Xu, and B. Xiao, "Detecting Sybil Attacks in VANETs," *Journal of Parallel and Distributed Computing*, vol. 73, no. 6, pp. 746–756, 2013.

[20] B. Xiao, B. Yu, and C. Gao, "Detection and Localization of Sybil Nodes in VANETs," in *Proc. ACM Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks*, 2006, pp. 1–8.

[21] "IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, March, 2012," *IEEE Std 802.11p-2010*, pp. 1–51, 2010.

[22] L. Cheng, B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige, "Mobile Vehicle-to-Vehicle Narrow-Band Channel Measurement and Characterization of the 5.9 GHz Dedicated Short Range Communication (DSRC) Frequency Band," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 8, pp. 1501–1516, 2007.

[23] X. Wang, A. Mueen, H. Ding, G. Trajcevski, P. Scheuermann, and E. Keogh, "Experimental Comparison of Representation Methods and Distance Measures for Time Series Data," *Data Mining and Knowledge Discovery*, vol. 26, no. 2, pp. 275–309, 2013.

[24] S. Salvador and P. Chan, "Toward Accurate Dynamic Time Warping in Linear Time and Space," *Intelligent Data Analysis*, vol. 11, no. 5, pp. 561–580, 2007.