Real-Time Detection and Mitigation of DDoS Attacks in Intelligent Transportation Systems

Ammar Haydari
Department of Electrical Engineering
University of South Florida
Tampa, Florida 33620
Email: ammarhaydari@mail.usf.edu

Yasin Yilmaz
Department of Electrical Engineering
University of South Florida
Tampa, Florida 33620
Email: yasiny@usf.edu

Abstract—Vehicular network (VANET), a special type of ad-hoc network, provides communication infrastructure for vehicles and related parties, such as road side units (RSU). Secure communication concerns are becoming more prevalent with the increasing technology usage in transportation systems. One of the major objectives in VANET is maintaining the availability of the system. Distributed Denial of Service (DDoS) attack is one of the most popular attack types aiming at the availability of system. We consider the timely detection and mitigation of DDoS attacks to RSU in Intelligent Transportation Systems (ITS). A novel framework for detecting and mitigating low-rate DDoS attacks in ITS based on nonparametric statistical anomaly detection is proposed. Dealing with low-rate DDoS attacks is challenging since they can bypass traditional data filtering techniques while threatening the RSU availability due to their highly distributed nature. Extensive simulation results are presented for a real road scenario with the help of the SUMO traffic simulation software. The results show that our proposed method significantly outperforms two parametric methods for timely detection based on the Cumulative Sum (CUSUM) test, as well as the traditional data filtering approach in terms of average detection delay and false alarm rate.

I. INTRODUCTION

Thanks to the recent improvements in vehicular technology, today’s vehicles tend to have more and more electronic components rather than being purely mechanic devices. This improvement leads to the birth of a new area called intelligent transportation systems (ITS) [1]. Vehicular Ad-hoc Network (VANET), which evolved from Mobile Ad-Hoc Network (MANET), is one of the major type of ITS applications. In VANET, communication among vehicles in traffic has significant impacts on public in terms of mobility and safety [2]. The applications of VANET can be divided into two categories, safety applications and service-oriented applications. Road Side Unit (RSU) based traffic management applications are good examples for safety applications, and internet based media sharing programs are good examples for service-oriented applications.

VANET collects and distributes many types of data packets such as information of emergency situation and vehicle conditions (e.g., position, average speed and behaviors on the road). The technological and societal aspects of VANET in real world makes it vulnerable to cyberattacks. Attacks on vehicular systems can be classified as inter-vehicle attacks and intra-vehicle attacks [3]. While inter-vehicle attacks aim to damage communication between vehicles and infrastructure, intra-vehicle attacks focus on inter-connection of devices within a vehicle. This study considers detecting inter-vehicle attacks, which can cause more severe damages to the entire network than intra-vehicle attacks.

There are various security approaches to ensure different objectives in ITS, such as availability, authenticity, integrity and non-repudiation [4]. Due to the highly dynamic characteristics, availability of network is one of the most important and challenging objectives in VANET, especially in safety related applications. There are various types and solution methods for Denial of Service (DoS) attacks, which target availability. DoS attack is basically performed by sending high volume of data packets (i.e., flooding the server) in order to interrupt network operations. Launching a high-rate (i.e., large amount of increase in data packets) DoS attack can cause significant damages to the system, but on the other hand, it is quite easy to detect such attacks, manifesting a trade-off for attackers. After detection, mitigation of attack would be also easy if attack originates from a single source. To prevent easy mitigation, attackers typically perform Distributed DoS (DDoS) attacks from a large number of compromised nodes in the network. In this research, we focus on low-rate DDoS attacks which is a way to perform stealth DoS attacks, e.g., [5], [6]. With a low-rate DDoS attack, attackers can make detection and mitigation quite challenging for the network operator by slightly increasing the data traffic from many nodes synchronously with respect to the nominal baselines, while achieving a sufficiently high data rate at the server that can interrupt the regular network operations at least in the
long-run. Such an attack is still a detrimental DDoS attack, but considering each node separately it may seem like no malicious activity takes place.

In this paper, we consider real time statistical detection and mitigation of flooding-based low-rate, as well as high-rate, DDoS attacks in ITS, specifically RSU-based VANETs. The proposed Intrusion Detection System (IDS) runs at RSU, which serves as the network center in a VANET, and monitors it for possible threats. Our contributions are listed below.

(1) To the best of the our knowledge, this work is the first one dealing with the timely detection and mitigation of low-rate DDoS attacks in a general-purpose VANET without specifying data-type and routing protocol. The proposed approach can be easily tailored for a specific-purpose VANET.

(2) Novel nonparametric, as well as traditional parametric, methods are presented for timely detecting DDoS (even low-rate) attacks in VANET while ensuring small false alarm probabilities (i.e., false positive rates).

(3) An effective statistical mitigation technique that successfully identifies attack locations is developed to overcome the effects of DDoS attack after detecting it.

The organization of the paper in remainder is as follows. Related works are discussed in Section 2. The traffic and attack models are given in Section 3. The proposed detection and mitigation methods are presented in Section 4. Numerical results are provided in Section 5. Finally, the paper is concluded in Section 6.

II. RELATED WORK

There is a number of works done for VANET safety. In [7], authors present a statistical detection based solution for DoS attacks in the IEEE 802.11 DCP protocol. This model, for each node, compares the received Clear-to-Send (CTS) packet rates to an adaptive threshold which is defined by a Markov chain. In [8], a DoS attack detector based on packet monitoring at a centralized node, similar to RSU, is proposed. By comparing SYN and ACK/SYN packets with predefined threshold values, DoS attack is detected. Another DoS detection mechanism for VANET is presented in [9], in which jamming attacks are detected through packet delivery ratio without needing centralized nodes. In [10], authors proposed a two-level method based on the Cumulative Sum (CUSUM) algorithm for statistical detection of DoS attacks in MANET. After calculating the first detection feature, they used CUSUM with the calculated value as the second detection feature.

There are several works on other attack types in VANET, such as false data injection attack, e.g., [11], [12], sybil attack, e.g., [13], and black hole attack, e.g., [14]. Machine learning based anomaly detection algorithms for VANET recently became popular. For instance, in [15], clustering together with Support Vector Machine (SVM) are used to detect malicious vehicles; and in [16], misbehavior classification through several features such as speed deviation and received signal strength (RSS) is studied.

III. SYSTEM MODEL

A. Traffic Model

In the considered traffic model (Fig. 1), system consists of vehicles with onboard units (OBU) for wireless communication and a road side unit (RSU). Vehicle to vehicle (V2V) and vehicle to RSU (V2I) communication are achieved by broadcasting. We focus on the V2I communication, in particular the communication from vehicles to RSU. Our proposed method does not specify any protocol, hence, it can be applied to all protocols. There is a variety of data packets which are transmitted in vehicular networks, such as position, average speed, condition of the road. For the sake of generality, we do not specify the type of data for our detection and mitigation model since vehicles may not have the same features and they may send different packets throughout the network.

For our detection model, RSUs collect packets within a range depending on the protocol that vehicles are using. Communication between vehicles and RSU is represented with lines in Fig. 1. The range is partitioned into \( D \) equal road segments and each data packet is labeled according to the received segment. If there are more than one car in the same segment of the road, regardless of the direction, their packets are binned together. We consider a periodic data communication, thus the packet rate depends on several factors such as the speed of the vehicle. For instance, if a vehicle is moving fast, the number of packets received on one segment of
the road will be less than that from a slow vehicle. Increasing number of cars will also increase the received packet rate at the RSU. If there is a traffic light on the road, the packet rate will also depend on the color of the traffic light and flow of traffic.

**B. Attack Model**

We consider DDoS attacks in VANET where attackers send high volume of data in order to make RSU unavailable at some point of time either through a highly-distributed low-rate DDoS attack or a high-rate DDoS attack (Fig. 2). As an example scenario, consider there is an accident within the monitored road segments. Normally, the accident information is received by RSU and conveyed to other RSUs in order to inform other vehicles who are far away from the accident. If attacker launches a DDoS attack from several vehicles to the closest RSU in the accident area, RSU cannot perform its regular operations, and thus cannot disseminate the accident information.

Data rate (packet/sec.) is a natural characteristic feature to consider in this type of DDoS attack in which some of the attackers increase their usual data rates. Since we do not specify the data type, discussion in this paper holds for any type of data including sum of all packet types. Attacker may target different types of packets, which will also increase the total number of packets.

In DDoS attacks, transmitted packets are legitimate, so no attack information can be derived from the packet contents. In a high-rate DDoS attack, the number of transmitted packets is highly anomalous (e.g., tens or even hundreds of times the nominal baseline), whereas in a low-rate DDoS attack, the number of packets transmitted from each vehicle may look nominal. The cumulative effect during the same time interval is what makes a low-rate DDoS attack coordinated among many vehicles detrimental to RSU. In addition to high-rate DDoS attacks, we specifically consider detecting and mitigating low-rate DDoS attacks from vehicles to RSU, that may stay undetected (i.e., stealth) to traditional IDS (e.g., firewalls) by slightly increasing (e.g., double the nominal baseline) the number of packets from a number of vehicles synchronously. On the other hand, such an attack can cripple RSU in the long-run or even earlier.

The proposed detection model does not consider any further specification for attacker, such as details of data (i.e., attacker can send different types of information) and the duration of attack. This work also does not assume any specific traffic conditions so that the proposed approach is applicable to different conditions, such as one-way, two-way, high-velocity, low-velocity.

**IV. DETECTION AND MITIGATIION MODEL**

**A. Online Discrepancy Test**

Anomaly detection algorithms work by first learning baseline (no attack) behavior, and then detecting anomalies based on changes with respect to the baseline behavior. There are several challenges in anomaly detection for DDoS attacks in VANET. First, timely detection is highly important because secure traffic flow highly depends on healthy operation of RSU. Second, unknown attack patterns are a main challenge for detection algorithms. Specifically, unknown parameters such as the number of attacked nodes (i.e., road segments in the considered system model), the set of attacked nodes, and the magnitudes of attack vectors render the traditional signature-based detection approaches impractical. Third, for network-wide effective detection of low-rate DDoS attacks, joint monitoring of nodes is required, which brings about a curse-of-dimensionality challenge.

Parametric approaches to anomaly detection try to fit a suitable parametric probability distribution to the observed data. Due to the second and third challenges given above, as well as the difficulty of fitting a standard distribution to real data, parametric approaches are not favored for DDoS attack detection in VANET. Nonparametric methods are more preferred since they are typically free from assumptions such as probability distribution, number and identity of attacked nodes.

A recent successful nonparametric method is the Online Discrepancy Test (ODIT) [17], which is capable of quickly detecting even small anomalies in high-dimensional networks. ODIT is based on two algorithms, Cumulative Sum (CUSUM) test and Geometric Entropy Minimization (GEM). It combines the nonparametric nature of GEM with the timely detection capability of CUSUM.

CUSUM is a popular sequential change detection algorithm [18], which assumes probability distributions for both before-change and after-change observations. When the true probability distributions are exactly known with all parameters, CUSUM is minimax optimum in terms of minimizing expected detection delay subject to a false alarm constraint [19]. The practical version of CUSUM, called Generalized CUSUM (G-CUSUM), estimates the parameters from data. However, as we show in simulation results in Section V, it is not easy to design a probabilistic model even for the no-attack case in VANET, and especially for the attack case considering the high uncertainty in attack scenarios. On the other hand, GEM [20] is a nonparametric geometric method that decides whether each data sample is an outlier or not. GEM is optimal when the anomalous distribution is a mixture of the nominal and uniform distributions [20]. The lack of temporal aspect prevents GEM’s effective use in timely detection of DDoS attacks in VANET. As we will discuss next ODIT adapts GEM to sequential detection through a CUSUM-type testing procedure while maintaining the nonparametric and computationally efficient characteristics.

In the training phase, for computational efficiency, ODIT randomly splits the training dataset $\mathcal{X}^{N} = \{x_1, \ldots, x_N\}$ into two subsets $\mathcal{X}^{N_1}$ and $\mathcal{X}^{N_2}$, where $N_1 + N_2 = N$, similar to the Bipartite GEM algorithm [21]. The data vector $x_t = [x_1^t \ldots x_N^t]$ received by RSU at each time $t$ contains the total number of packets $x_i^t$ transmitted from each road segment.
to outliers and sensitivity to anomalies. Small \( k \) gives more robustness but less sensitivity. The \( M \) points are chosen according to minimizing the total edge length which is given for point \( m \) as follows

\[
L_m = \sum_{n=k-s+1}^{k} |e_{m(n)}|^{\gamma},
\]

where \( |e_{m(n)}| \) is the Euclidean distance between point \( m \) and its \( n \)th nearest neighbor in \( \mathcal{X}^{N_2} \), \( s \) is a fixed number between 1 and \( k \) defined for convenience, and \( \gamma > 0 \) is a weight typically chosen as 2. \( M \) is determined according to the outlier definition selected by the user. For instance, considering outliers at 0.05 significance level \( M \) is selected as the 95th percentile of data points in \( \mathcal{X}^{N_1} \), i.e., \( M = \text{round}(0.95N_1) \).

The total edge length \( L_{(M)} \) of the \( M \)th point will be used as a baseline statistic in the test phase.

Until this point, we explained the training procedure for the ODIT-based proposed detection method, in which nominal traffic conditions are observed in terms of the number of transmitted packets from road segments. In the test phase, considering the data vector \( \mathbf{x}_t \), which consists of the number of packets received from each road segment by RSU, its total edge length \( L_t \) with respect to the points in \( \mathcal{X}^{N_2} \) is computed, as shown in (1). Then, comparing \( L_t \) with the baseline distance statistic \( L_{(M)} \), we obtain an anomaly evidence at time \( t \) as

\[
D_t = L_t - L_{(M)}.
\]

Drawing upon the CUSUM test statistic, the ODIT statistic is recursively updated at each time \( t \)

\[
s_t = \max\{s_{t-1} + D_t, 0\},
\]

where \( s_t \) is selected to strike a balance between true positive rate and false positive rate (see the ROC curve in Fig. 9).

Fig. 3: Proposed detection procedure based on ODIT with \( N_1 = 6, N_2 = 9, M = 5, k = 2, s = 1, \gamma = 2 \). \( L_1 - L_{(M)} \) and \( L_2 - L_{(M)} \) are used to update the test statistic \( s_t \) and raise an alarm at time \( T \) as shown in (2)-(4). Training and test points are generated from a bivariate normal distribution with independent components, 0.5 mean and 0.1 standard deviation.

The total edge length \( L_t \) is computed, that segment \( t \) is attacked if \( D_t \geq \lambda \), where \( \lambda \) is defined for convenience, and \( L_t \) is the squared Euclidean norm of a \( d \)-dimensional distance vector whose \( d \)th entry \( t_{l}^{d} \) is the sum of distances of data from segment \( d \) at time \( t \) to its nearest neighbors in \( \mathcal{X}^{N_2} \). If \( t_{l}^{d} \) is large, that means segment \( d \) contributes to a large value of \( L_t \) towards the alarm. This provides an evidence that segment \( d \) is under attack. Specifically, we compute the following statistic for each segment \( d \) when an attack is detected:

\[
\bar{t}^{d} = (T - q)^{-1} \sum_{t=q}^{T} t_{l}^{d}
\]

where \( q \) is the time instance when the detection statistic \( s_t \) started to increase from zero before the alarm. Each road segment \( d \) is identified as attacked if \( \bar{t}^{d} \geq \lambda \). The threshold \( \lambda \) is selected to strike a balance between true positive rate and false positive rate (see the ROC curve in Fig. 9).
TABLE I: SIMULATION PARAMETERS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Area</td>
<td>9000 x 5000 m²</td>
</tr>
<tr>
<td>Simulation Time (Each Trial)</td>
<td>200s</td>
</tr>
<tr>
<td>Number of Trials</td>
<td>600</td>
</tr>
<tr>
<td>Average Number of Vehicle</td>
<td>250</td>
</tr>
<tr>
<td>Traffic Generation</td>
<td>Random</td>
</tr>
<tr>
<td>Route Generation</td>
<td>Random</td>
</tr>
<tr>
<td>Network Protocol</td>
<td>IEEE 802.11p</td>
</tr>
<tr>
<td>Beacon Rate</td>
<td>1s</td>
</tr>
<tr>
<td>Network Interface</td>
<td>OMNET++</td>
</tr>
<tr>
<td>Network Mobility Framework</td>
<td>Veins</td>
</tr>
<tr>
<td>Traffic Generator</td>
<td>SUMO</td>
</tr>
<tr>
<td>Map</td>
<td>Fowler Av. Tampa, FL</td>
</tr>
</tbody>
</table>

V. SIMULATION RESULTS

A. Simulation Setup

We tested our model using a compound of three softwares, OMNET++ [22], SUMO [23] and Veins [24]. OMNET++ is a network simulator providing interface to test network systems. Simulation of Urban Mobility (SUMO) is an open source traffic generator which creates mobility scenarios on real road maps based on the specified parameters. Veins is a special framework that connects SUMO with OMNET++. By the help of Veins, each vehicle is represented as a mobile node in the network. In this simulation, we consider the IEEE 802.11p vehicular communication protocol [25].

In order to have a realistic testbed, we simulate the traffic on a portion of the Fowler Ave. which lies on the southern edge of the University of South Florida (USF) campus in Tampa, as shown in Fig. 4. The considered portion is partitioned into 20 segments. Vehicle movements are not restricted. That is, vehicles may enter and exit the main road from all possible connecting side roads. Number of vehicles and route of vehicles are randomly generated. On average there are approximately 250 vehicles in each trial. Simulation parameters are summarized in Table I.

In our simulation, while vehicles are moving on the roads in SUMO, they are identified as a node in OMNET++ by the help of Veins. Each node (vehicle and RSU) broadcasts packets to all nodes that are in their range. For training, 4 hours of traffic is observed which was sufficient to learn a baseline for the nominal traffic conditions. For the test part, we observed 33.3 hours of traffic and all the log files are saved. We have a single baseline in this case but we can generate different baselines for different time intervals such as in the early morning rush hour traffic, and in the afternoon free flow traffic. From collected log files, we computed data rates in MATLAB and obtained 600 trials with 200 seconds each. Attack data is generated in MATLAB for two different cases from uniform distribution. For the first case, lower and upper bounds are selected as 0.1 and 0.5 times the average number of packets in the nominal case and for the second case, these bounds are selected as 1 and 2 times the average number of packets in the nominal case. In each trial, attack data is added on top of the nominal data in 2 of the 20 road segments from 181st second to 200th second.

Fig. 4: Simulation map showing Fowler Ave.

Fig. 5: Histogram of number of packets for a road segment. First histogram represents the distribution of nominal data, whereas second and third represent attack cases with an average increase that is 0.3 and 1.5 times the baseline, respectively. Nominal and attack distributions are close to negative binomial distribution with extended tails under attacks.
Fig. 6: Histogram of number of packets for a road segment. First histogram represents the distribution of nominal data, whereas second and third represent attack cases an average increase that is 0.3 and 1.5 times the baseline, respectively. Nominal and attack distributions are close to normal distribution with extended tails under attacks.

B. Results

We compared our nonparametric model with Generalized CUSUM (G-CUSUM) assuming two different distributions since we insert the anomaly to two different road segments where both has different distributions. In one of the road segment, distribution of nominal data seems to be close to the negative binomial, which is indeed a Poisson distribution with conjugate prior (i.e., Gamma distribution) on the rate parameter, hence we firstly consider G-CUSUM with negative binomial assumption (Fig. 5). Along with negative binomial we also consider G-CUSUM with the normal (i.e., Gaussian) distribution, because the data in the other road segment is similar to normal distribution (Fig. 6). We also compared our statistical model with the classical data filtering approach which only considers the increase in the data rate without any statistical analysis.

Fig. 7: Average detection delay vs. False alarm probability for the proposed method, G-CUSUM and basic data filtering approach for an average increase that is 0.3 times the nominal mean data rate.

Different attack scenarios are considered for the test purposes. First, in order to test against low rate DDoS attacks, we increased the mean by only 30% on average using uniform distribution. Second, we considered 150% increase on average again from uniform distribution. For both scenarios, although the exact knowledge of the mean increase is assumed known by G-CUSUM, the proposed method significantly outperforms both G-CUSUM variants and basic data filtering model in terms of average detection delay vs. false alarm probability, as shown in Fig. 7 and Fig. 8. These results clearly demonstrate the advantage of the proposed nonparametric method over the parametric CUSUM-based methods, e.g., [10], and classical data filtering method, e.g., [8], for detecting low-rate DDoS attacks, which are typically much harder to detect than the high-rate attacks.

In the low rate DDoS attack scenario (Fig. 7), the identification performance of the proposed mitigation approach is shown by the ROC curve in Fig. 9. The attacked segments are successfully identified by the approach given in Section IV-B.

VI. CONCLUSION

Security of Intelligent Transportation Systems (ITS) is becoming more important as vehicles and smart infrastructure elements, such as Road Side Units (RSU) are getting more connected. We addressed the challenging low-rate DDoS attacks to RSU in VANET by presenting a novel detection and mitigation framework based on nonparametric anomaly detection. Our proposed method quickly detects low-rate DDoS attacks, successfully identifies the attack locations, and mitigates the attack by blocking the data traffic from attack locations. Extensive simulation results showed that standard parametric methods cannot model the data traffic in a real road scenario, thus they are significantly outperformed by the proposed nonparametric method which does not depend on probability distribution assumptions. Simulation data is generated using three softwares together, namely SUMO traffic simulator, OMNET network simulator, and Veins, which
connects SUMO and OMNET. The proposed method does not assume specific data type and protocol, hence it is applicable to a broad range of attack scenarios. Although we applied the proposed method to a single scenario, in practice it can be trained on different time intervals to learn different baselines for several traffic conditions, such as rush hour, weekend, accident, etc. In that case, depending on the time of the day the algorithm will test the incoming data against the corresponding baseline.

REFERENCES


