GPS Spoofing Attack Detection on Intersection Movement Assist using One-Class Classification

Jun Ying
Purdue University
ying29@purdue.edu

Yiheng Feng
Purdue University
feng333@purdue.edu

Qi Alfred Chen
University of California at Irvine
alfchen@uci.edu

Z. Morley Mao
University of Michigan
zmao@umich.edu

***Abstract*—Intersection movement assist (IMA) is a connected vehicle (CV) application to improve vehicle safety. GPS spoofing attack is one major threat to the IMA application since inaccurate localization results may generate fake warnings that increase rear-end crashes, or cancel real warnings that may lead to angle or swipe crashes. In this work, we first develop a GPS spoofing attack model to trigger the IMA warning of entry vehicles at a roundabout driving scenario. The attack model can generate realistic trajectories while achieving the attack goal. To defend against such attacks, we further design a one-class classifier to distinguish the normal vehicle trajectories from the trajectories under attack. The proposed model is validated with a real-world data set collected from Ann Arbor, Michigan. Results show that although the attack model triggers the IMA warning in a short time (i.e., in a few seconds), the detection model can still identify the abnormal trajectories before the attack succeeds with low false positive and false negative rates.**

***Index Terms*—Anomaly detection, GPS spoofing attack, Intersection movement assist, Connected vehicles, One class classification**

I. Introduction

Connected vehicle (CV) technology has great potential to benefit the transportation system in terms of improving system efficiency, sustainability, and safety. Vehicle-to-vehicle (V2V) communication enables the CVs to send and receive real-time information from other nearby vehicles, for example, Basic Safety Messages (BSMs) to avoid collisions. BSMs play an important role in multiple CV applications, such as intersection movement assist (IMA), a widely implemented CV application to improve vehicle safety [12]. The IMA system can be applied when vehicles pass through unsignalized intersections. It receives other approaching vehicles' information such as location and speed to determine whether it is safe to enter the intersection. If a potential collision is detected, a warning message will be generated and sent to the driver (e.g., through an in-vehicle display or an audio warning). Among all information that is shared through V2V communication, vehicle position is critical in deciding whether to generate warnings to drivers. To obtain a vehicle's real-time position, a GPS receiver is commonly used for vehicle localization and navigation [1], [14]. Commercial-grade GPS receivers can get vehicle position within a meter accuracy [1] while AV-grade GPS receiver has centimeter-level positioning accuracy [13]. It is important to guarantee that the vehicle's localization module is accurate and reliable.

Existing studies show that GPS receivers are vulnerable to multiple cyber attacks. One major threat is the spoofing attack, which has been proved feasible both theoretically [15] and practically on various systems [3], including in autonomous vehicles [13]. To defend against GPS spoofing attacks, multiple detection methods have been proposed, including filter-based methods and observer-based methods [4] [17]. However, the proposed methods either rely on other onboard sensors or V2V information from surrounding vehicles, which may not be available in the CV environment with a low penetration rate. Our previous work [18] proposed a GPS spoofing attack detection method, which combines learning from demonstration and a decision tree classifier. The decision tree classifier needs to be trained using both ground truth trajectory and known attack trajectory. As a result, the proposed detection framework can be only applied to detect known attacks. However, in reality, new attacks are usually unknown to the detector, where the attack trajectories can not be obtained for training the classifier.

In this paper, we first propose a GPS spoofing attack model which aims to trigger the IMA warning of entry vehicles in a roundabout scenario. We further design a one-class classifier to distinguish the normal trajectories from the trajectories under attack, where only the normal trajectories are needed for training. Our work can be briefly summarized as follows. The attack model is formulated as an optimization problem, with specifically designed features to trigger the IMA warning while generating as normal driving behaviors as possible. The detection framework includes a feature extractor and a one-class neural network classifier. In the case study, a real-world data set, which is collected from a two-lane roundabout in Ann Arbor, Michigan [20] is applied to test both the CV threat model and the detection model. Results show that the proposed threat model can trigger the IMA warning in a short time (less than 1.7s) which poses great challenges to the detection. The online detection results denote that the proposed detection framework can differentiate the normal and abnormal trajectories before the attack succeeds time, with both low false

Symposium on Vehicles Security and Privacy (VehicleSec) 2023
27 February 2023, San Diego, CA, USA
ISBN 1-891562-88-6
https://dx.doi.org/10.14722/vehiclesec.2023.23038
www.ndss-symposium.org

positive rates and low false negative rates.

The main contributions of the paper are listed as follows:

1. We formulate the GPS spoofing attack as an optimization model, with the goal to trigger an IMA warning as well as generating smooth and realistic trajectories. The proposed threat model takes vehicle's initial state and road geometry into consideration and can be applied to different scenarios, not limited to the roundabout.

2. A generic detection framework is proposed. The detection framework combines a feature extractor with a one-class classifier. The feature extractor can be adjusted according to different driving scenarios. Besides, the proposed framework only requires normal trajectories for training, which enables it to detect unknown attacks.

The remainder of the paper is arranged as follows. Section II reviews related literature on GPS spoofing attacks and related detection methods. Section III introduces the threat model, including the problem statement, objective function, and implementation framework. Section IV presents the detection methodology. Numerical experiments from the roundabout scenario are introduced in section V. Section VI concludes the work and lays out future research directions.

II. Literature review

In this section, we reviewed literature related to GPS spoofing attacks and related detection models.

A. GPS Spoofing Attack

GPS spoofing attack has been a long-existing problem. The attack broadcasts incorrect but valid GPS signals to mislead GPS receivers [11]. By providing falsified information, GPS spoofing attacks can deviate vehicles to random positions [3] or guide the vehicles to the wrong destinations. Zeng et al. [19] proposed a stealthy attack against the road navigation system. GPS locations were spoofed slightly to trigger the turn-by-turn navigation and guided the vehicle to the wrong destination without recognizing the attack. To prove the feasibility, the proposed GPS spoofing model was tested on real vehicles. Narain et al. [8] evaluated the INS-aided GPS system and developed algorithms to deviate the vehicle to alternative locations without being detected. The result showed that the proposed algorithm could deviate vehicles as far as 30km from the origin without raising alarms. Multi-Sensor Fusion (MSF) is usually considered one approach to defending GPS spoofing attacks. An MSF system combines inputs from multiple sensors for vehicle localization. It is highly unlikely that all sensors can be compromised at the same time. For example, Liu et al. [6] proposed an Extended Kalman filter (EKF) based algorithm to fuse the measurements from multiple sensors. The proposed algorithm performed well under GPS spoofing attacks where the GPS signal was deviated by a fixed bias. However, Shen at al. [13] proposed a GPS spoofing attack algorithm that penetrated the MSF based localization system with GPS, IMU and Lidar. The proposed algorithm only spoofed GPS to cause large deviations in the MSF output. It could deviate the vehicle from the original lane, or cross the road boundary, which may lead to collisions with other vehicles.

B. GPS Spoofing Attack Detection

To defend against GPS spoofing attacks, anomaly detection methods have been proposed. The detection methods can be divided into filter-based methods and observer-based methods [4]. Filter-based methods consider uncertainties and measurement noises and apply filters such as Kalman Filter to detect attacks on sensors. Van et al. [16] proposed an anomaly detection approach that combines convolutional neural network (CNN) and Kalman filtering with χ^2 detector. CNN was used for detecting anomalies in time-series sensor data and KF-based χ^2 detector is applied to detect the abnormal data which are undetected by CNN. Ju et al. [5] proposed a simple distributed Kalman filter based on neighboring vehicle measurement exchange. A Generalized likelihood ratio (GLR) detector was proposed to detect position sensor attacks based on the Kalman filter's result.

Compared with filter-based methods, observer-Based methods are usually applied to deterministic vehicle models. Wang et al. [17] proposed an observer-based method. An adaptive extended Kalman filter was applied to smooth vehicle sensor data based on a nonlinear car-following model. One Class Support Vector Machine (OCSVM) is applied to detect anomalies on sensors. He et al. [2] proposed an observer-based detection framework. A detector was developed according to the potentially compromised sensor measurements and the observer's estimation. Measurement data was discarded if it was larger than a threshold.

The existing methods either need input from multiple sensors [6], input from known attacks [18], or only detect the anomaly in the longitudinal vehicle dynamics [16]. Besides, surrounding vehicle states such as leading vehicles or information from other vehicles in the platoon are needed [2]. They may not be applicable to our case because 1) in the CV environment, there may not exist other onboard sensors to perform multi-sensor fusion or cross-validate the results from the GPS, especially if a vehicle is equipped with aftermearket safety devices (ASDs). 2) in the roundabout scenario, vehicles have lateral movement due to road curvature, which significantly increases the detection difficulty.

III. Threat model

In this section, the threat model towards the IMA application on the CV is presented. IMA is an important CV safety application [12] [7]. When approaching an intersection, the IMA system first receives information (i.e., BSMs) from other vehicles. According to the received data, the IMA system determines whether it is unsafe to enter an unsignalized intersection due to potential collision with other vehicles and sends warnings to drivers. Drivers receiving collision warning information from the IMA system should perform actions to avoid crashes at the intersection. In this work, it is assumed that the IMA warnings are triggered only based on received BSMs from other vehicles at the intersection. We propose

a threat model towards the IMA system at the roundabout scenario, which is a common type of unsignalized intersection.

A. Problem Statement

The proposed CV threat model generates falsified BSMs to trigger the IMA warning of CV at the entry of the roundabout. Figure 1 demonstrates the attack concept. The figure contains two vehicles, the vehicle under attack and the victim vehicle and both vehicles are CVs. The attack vehicle is the vehicle located in the inner lane of the roundabout. The victim vehicle is the vehicle located at the entry of the roundabout. The blue rectangles denote the real vehicle trajectory in the inner lane of the roundabout. The red rectangles denote the falsified BSM trajectory when the vehicle is under attack, changing lanes from the inner lane to the outer lane. The yellow rectangles denote the victim vehicle trajectory. A conflict point is defined as the intersection between the center line of the outer lane of the roundabout and the entry path. Note that lane changing is forbidden within the multi-lane roundabout. When the vehicle is not under attack, the BSM sent from the CV in the roundabout should be consistent with the real trajectory (the blue rectangles). There is no conflict between the real CV trajectory and the victim vehicle trajectory. The IMA warning will not be triggered for the entry vehicle. When the vehicle is under attack, the CV in the roundabout sends out falsified BSMs (the red rectangles) to trigger the IMA warning of the victim vehicle, without really controlling vehicle movement. The values within the rectangles denote the timestamps. The attack starts at time t_0 and the attack successfully triggers the IMA warning of the victim vehicle at time t_2.

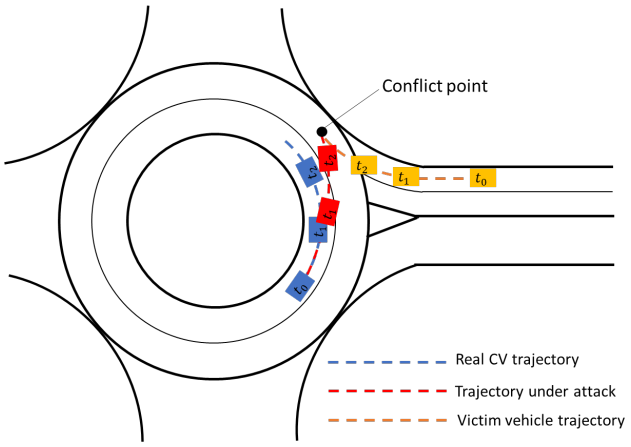

Fig. 1. Threat model on intersection movement assist system

To generate the falsified trajectory (the red rectangles), an optimization problem is formulated, as shown in Equation 1. The objective function is presented as $\theta^T f(\mathbf{s}, \mathbf{u})$. θ is the weight vector and $f(\mathbf{s}, \mathbf{u})$ is a function mapping a trajectory to feature vectors. $\mathbf{s}$ is the variable of the optimization model. $\mathbf{s} = (s_1, s_2, ..., s_N)$ denotes the set of trajectory points, where s_i is the trajectory point at time step i. Each trajectory point s_i consists of $(x_i, y_i, v_i, a_i, \psi_i)$, where x_i,y_i denotes vehicle's longitudinal and lateral coordinate at time step i. v_i and a_i represent vehicle speed and acceleration at time step i. ψ_i is the vehicle's heading angle. N is the planning horizon for the attack trajectory, which is determined by the estimated arrival time for the victim vehicle to reach the conflict point. $\mathbf{u}$ denotes the vehicle's initial state and road geometry. The vehicle's initial state includes its initial position and status (speed, heading, acceleration). Road geometry includes the radius of the inner and outer lanes of the roundabout and the coordinate of the conflict point. The feature selection for the objective function and the constraints are introduced in the following section.

$$\begin{aligned} \min_{\mathbf{s}} \quad & \theta^T f(\mathbf{s}, \mathbf{u}) \\ \text{s.t.} \quad & \text{vehicle dynamic constraints} \end{aligned} \tag{1}$$

B. Objective Function

1) Feature Vectors: The objective function contains two parts: 1) trigger the IMA warning of the victim vehicle. 2) generate a trajectory close to the normal driving behavior considering smoothness and comfort. A realistic attack trajectory will increase the difficulty in detection. To achieve the attack goal, five features are selected and elaborated as follows:

(1) Acceleration: $f_1 = \frac{1}{N}\sum_i a_i^2$. f_1 sums up the a_i^2 for the entire trajectory. Uncomfortable driving behavior such as large accelerations are penalized by minimizing f_1.

(2) Heading rate: $f_2 = \frac{1}{N}\sum_i (\dot{\psi}_i)^2$. ψ_i denotes the heading angle change rate at time step i. f_2 minimizes the difference of heading rate for two consecutive time steps.

(3) Curvature: $f_3 = \frac{1}{N}\sum_i (\sqrt{(x_i - x^c)^2 + (y_i - y^c)^2} - r^c)^2$. x^c and y^c denote the coordinate of the center of the roundabout. r^c denotes the radius of the roundabout. f_3 calculates the difference between the vehicle's distance to the center of the roundabout and the roundabout radius at time step i. f_3 guarantees the vehicle stays in the roundabout.

(4) Lateral terminal position: $f_4 = (x_N - x^{con})^2$ x_N denotes the vehicle's lateral coordinate at the end of the planning horizon. x^{con} represents the conflict point's lateral coordinate.

(5) Longitudinal terminal position: $f_5 = (y_N - y^{con})^2$ y_N is the vehicle's longitudinal coordinate at the end of the planning horizon. y^{con} is the conflict point's longitudinal coordinate. f_4 and f_5 push the vehicle to reach the conflict point at the end of the planning horizon and trigger the IMA warning of the victim vehicle.

2) Vehicle Dynamic Constraints: In this section, vehicle dynamic constraints are introduced. Constraint 2-5 denotes the evolution of vehicle state, including vehicle position, speed, acceleration, and heading angle. Equation 6-8 limits vehicle kinematic parameters within boundaries. Equation 6 bounds the vehicle's maximum acceleration and deceleration to be less than $8m/s^2$. Equation 7 limits vehicle's heading rate within range $(-\frac{\pi}{3}, \frac{\pi}{3})$. Equation 8 limits the vehicle's maximum speed.

$$x(i+1) = x(i) + v(i)cos(\psi(i))\tau \tag{2}$$

$$y(i+1) = y(i) + v(i)sin(\psi(i))\tau \tag{3}$$

$$\dot{\psi}(i) = \frac{(\psi(i+1) - \psi(i))}{\tau} \tag{4}$$

$$v(i+1) = v(i) + a(i)\tau \tag{5}$$

$$-8 \leq a(i) \leq 8 \tag{6}$$

$$-\frac{\pi}{3} \leq \dot{\psi}(i) \leq \frac{\pi}{3} \tag{7}$$

$$v(i) \leq 16.7 \tag{8}$$

C. Implementation Framework

The implementation framework of the attack model is described in this section. Details of the attack trajectory generation procedure is described in the following steps and shown in Figure 2.

Step 1: Collect vehicle state. The vehicle under attack (roundabout vehicle) collects its own state and the entry vehicle state, including vehicle speed and position.

Step 2: Calculate estimated arrival times to the conflict point for the roundabout vehicle and the entry vehicle, denote as t_r^{est} and t_e^{est}. D is vehicle's distance to conflict point. v denotes vehicle's current speed. The estimated arrival time is calculated using Equation 9, assuming the vehicle keeps current speed to reach the conflict point.

$$t^{est} = \frac{D}{v} \tag{9}$$

Step 3: Determine the attack start time. The attack starts when $|t_r^{est} - t_e^{est}|$ is less than 4s. If the criterion to start attack is satisfied, go to **Step 4**. Otherwise, go to **Step 1**. The threshold of launching attack is a hyper-parameter and needs to calibrated based on real-world data.

Step 4: Update entry vehicle state.

Step 5: Generate attack trajectory. The attack trajectory is generated according to Equation 1, assuming victim vehicle (entry vehicle) keeps a constant speed. The length of the attack trajectory (i.e., planning horizon) is set to be the same as the estimated arrival time of the roundabout vehicle to the conflict point, calculated in Equation 9.

Step 6: Determine whether the attack success criterion is satisfied. If the attack success criterion is satisfied, the attack ends. Otherwise, go to **Step 7**. Equation 10 and 11 denote the attack success criterion. D_{atk} is the attack vehicle distance to the conflict point. v_{atk} is the attack vehicle speed. r_{atk} is the distance between the attack vehicle trajectory point and the center of the roundabout. r_l denotes the lane boundary radius between the inner and outer lanes. The attack succeeds when the post encroachment time (PET) to the conflict point between the attack trajectory and the victim vehicle is less than T_g and the attack trajectory is deviated from the inner lane and crosses the road boundary. In this work, T_g is equal to $2s$. Both attack success criteria guarantee the two vehicles have a potential collision at the conflict point.

$$\frac{D_{atk}}{v_{atk}} \leq T_g \tag{10}$$

$$r_{atk} \geq r_l \tag{11}$$

Step 7: Determine when to stop generating attack trajectory. When the entry vehicle has reached the conflict point and the attack success criterion is not satisfied, continue attacking the vehicle will no longer trigger the IMA warning of the entry vehicle. Therefore, the attack should end. Otherwise, go to **Step 4**.

In order to minimize the prediction error, a rolling horizon framework is applied to update the entry vehicle information (speed and position) and calculate the planning horizon once the attack starts (**Step 4**-**Step 7**). The attack trajectory is generated for the whole planning horizon, but only the first 0.4s will be used.

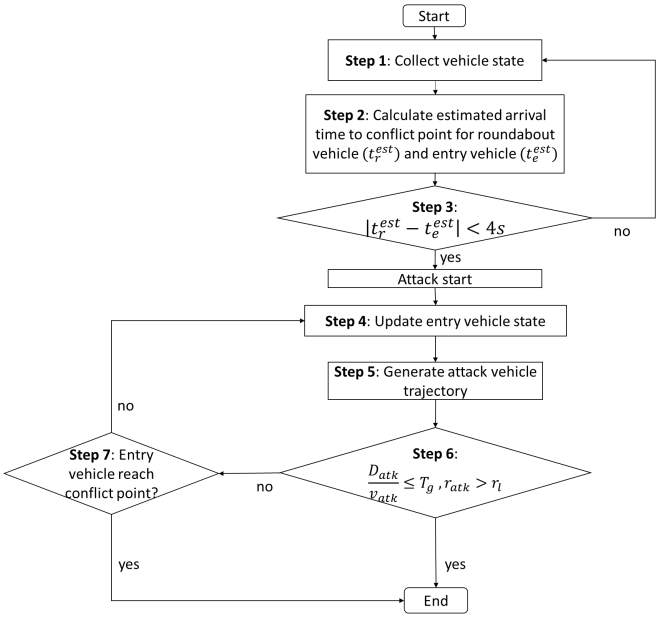

Fig. 2. Threat model implementation framework

IV. Detection Methodology

In this section, we introduce a one-class classifier that is designed to detection the GPS spoofing attack toward the IMA application.

A. Detection Framework

Figure 3 demonstrates the detection framework. It consists of two parts, an offline training step and an online detection step. First, a training data set that includes historical normal trajectories is collected. A feature extractor is applied that maps trajectories to feature vectors, which represent different aspects of driving behaviors. A one-class neural network classifier is trained with extracted features from normal trajectories. The trained classifier is then applied to the online detection, as shown at the bottom of Figure 3. Given the observed trajectory, the same feature extractor is applied to extract driving related features. The extracted features are sent to the trained anomaly classifier, to determine if the vehicle is under attack or not.

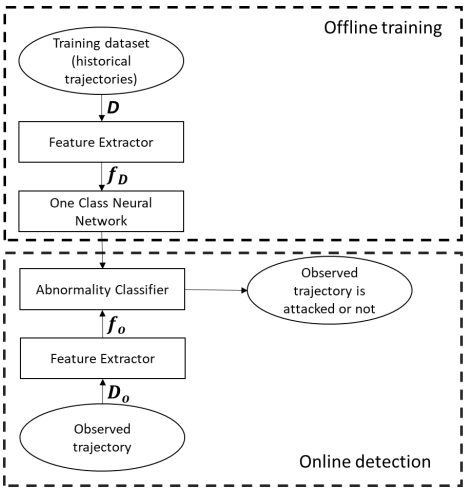

Fig. 3. Anomaly detection framework

B. One class classification

Figure 4 demonstrates the structure of the one-class classifier, which contains a feature extractor and a classifier network. The feature extractor is predefined and maps the input trajectory into a feature vector. Pseudo-positive data are generated from a Gaussian distribution $N(\overline{\mu}, \sigma^2)$. σ and $\overline{\mu}$ are parameters of the Gaussian distribution. Denote N is the number of input data and D is the dimension of the feature vector. The generated pseudo-positive data has the same dimension as the feature data set. The generated data are then combined with the extracted feature data set and fed into a classifier, following the method shown in [10]. The classification network is composed of three fully connected layers, followed by a sigmoid function. The output of the classifier is 0 or 1. 0 denotes that the data sample belongs to normal and 1 denotes that the data sample is abnormal. The binary cross entropy loss is applied as the loss function to train the classification network.

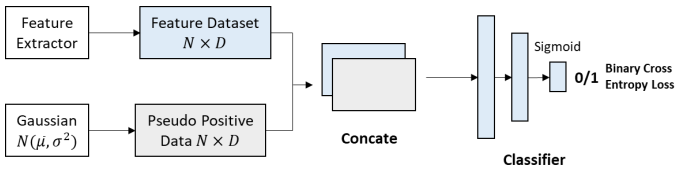

Fig. 4. One Class Classifier Framework

The classification network is optimized using the SGD optimizer, with the learning rate equals to 10^{-3} and the batch size equals to 64. μ equals 0 and σ equals 3 for the Gaussian distribution to generate pseudo-positive data.

C. Feature extractor

In the proposed anomaly detection model, ten features are designed to describe normal driving behavior, including both longitudinal and lateral behaviors. The designed features are elaborated as follows:

(1) Average lateral acceleration: $f_1 = \frac{1}{N}\sum_i^N |a_i \sin\psi_i|$. N is the trajectory length. ψ_i is the vehicle heading at time step i. f_1 calculates the average of lateral acceleration at each time step.

(2) Maximum lateral acceleration: $f_2 = \max_{i=1,...N}|a_i \sin\psi_i|$. f_2 is the max value of the lateral acceleration for the entire trajectory. f_1 and f_2 measure the smoothness of the lateral driving behavior.

(3) Average lateral speed: $f_3 = \frac{1}{N}\sum_i^N |v_i \sin\psi_i|$. f_3 calculates the average lateral speed.

(4) Maximum lateral speed: $f_4 = \max_{i=1,...N}|a_i \sin\psi_i|$. f_4 is the maximum lateral speed. f_3 and f_4 denotes the vehicle's lateral driving behavior.

(5) Average longitudinal acceleration: $f_5 = \frac{1}{N}\sum_i^N |a_i \cos\psi_i|$. f_5 calculates the average longitudinal acceleration.

(6) Maximum longitudinal acceleration: $f_6 = \max_{i=1,...N}|a_i \sin\psi_i|$. f_6 calculates the maximum longitudinal acceleration. f_5 and f_6 represent the smoothness of the longitudinal driving behavior.

(7) Average longitudinal speed: $f_7 = \frac{1}{N}\sum_i^N |v_i \cos\psi_i|$. f_7 is the average longitudinal speed.

(8) Maximum longitudinal speed: $f_8 = \max_{i=1,...N}|a_i \cos\psi_i|$. f_7 and f_8 denote the driver's longitudinal driving efficiency at the roundabout.

(9) Maximum heading rate: $f_9 = \max_{i=1,...N}\left|\dot{\psi}_i\right|$. $\dot{\psi}_i$ denotes vehicle heading change rate at time step i.

(10) Average heading rate: $f_{10} = \frac{1}{N}\sum_i^N\left|\dot{\psi}_i\right|$. f_9 and f_{10} demonstrates vehicle's driving smoothness at the roundabout.

Note that for different driving scenarios, the features may be designed differently. A Greedy Algorithm is applied to select critical features from the designed feature list.

Algorithm 1 Greedy Algorithm

$S \leftarrow \emptyset$
$NS \leftarrow \{f_1, f_2, ...f_{10}\}$
$A^* = 0$
while $NS \neq \emptyset$ **do**
  for $f_i \in NS$ **do**
    Calculate the One class classifier accuracy using features $S \cup f_i$, the accuracy on testing set is denoted as $A(f_i)$
  end for
  $f^* = argmax_{f_i} A(f_i)$
  $a^* = A(f^*)$
  if $a^* > A^*$ **then**
    $S = S \cup f^*$
    $A^* = a^*$
    $NS = NS \setminus f^*$
  else
    Break
  end if
end while
return S

Algorithm 1 denotes the greedy algorithm that selects the critical features used for the one-class classification neural network. S denotes the selected feature set. S is initiated as an empty set at the beginning of the algorithm. NS denotes the feature set which is not selected. A^* denotes the highest accuracy, A^* is initiated with 0. While NS is not an empty set, the one class classifier is trained with $S \cup f_i$, where $f_i \in NS$. Testing accuracy is calculated and denoted as A_i. The algorithm loops through all features in NS and selects the feature that maximizes testing accuracy denoted as f^* and the relative testing accuracy is denoted as a^*. If $a^* > A^*$, then S will be updated by adding f^* into it, and f^* will be extracted from NS and A^* will be updated as a^*. If no improvement is made by adding feature $f_i \in NS$, the iteration stops and returns with S, which is the feature set that achieves the highest testing accuracy. f_1, f_2, f_3, f_5, f_9 are selected and used as feature extractor according to Algorithm 1. The selected features can describe both longitudinal and lateral driving behaviors.

V. Experiments

To validate the proposed anomaly detection framework, a roundabout data set collected at Ann Arbor, Michigan is applied. The data set is collected at a two-lane roundabout. The roundabout is equipped with infrastructure sensors such as radars and cameras are installed at the four corners of the roundabout. Vehicle trajectories approaching and within the roundabout are extracted from the video data with a time step equal to 0.4s [20].

This roundabout is of high-interest because of its high crash rates. 69 crashes happened at the intersection in year 2021. Among them, 66 crashes happened between the vehicle travelling inside the roundabout and the vehicle at the entry [9]. One possible solution to reduce the crash counts is to apply the IMA. However, if the IMA application is under cyber attack, generating fake warnings and/or canceling true warnings may even aggravate the crash risks.

In this section, we first show the results of the attack model and then evaluate the anomaly detection framework with the threat model.

A. Attack model results

In the experiment, qualified vehicle trajectory pairs in the roundabout data set are extracted and used to generate attack trajectories. Each trajectory pair consists of the vehicle traveling within the roundabout and an entry vehicle. A qualified vehicle pair must have a similar arrival time to the defined conflict point so that the driver of the entry vehicle would actually observe a real approaching vehicle in the roundabout. In this way, when the IMA warning is triggered, the driver of the entry vehicle may take real actions to avoid the (fake) conflict. Based on this criteria, a total number of 927 vehicle pairs are selected.

Using the attack model illustrated in section III, 744 attack trajectories are generated. Figure 5 shows an example of the attack trajectory. The red line represents the original vehicle trajectory. The green line represents the vehicle trajectory generated by the attack model. The black line denotes the victim vehicle trajectory. In this case, the average vehicle speed at the roundabout is around $7m/s$, which is consistent with the speed limit in the round about (15mph). At the end of the green vehicle trajectory, the IMA warning is triggered. The result shows that the proposed algorithm can generate falsified lane changing trajectory that follows roundabout's geometry to trigger the IMA warning of the entry vehicle.

Fig. 5. Attack Trajectory at the Roundabout

The overall attack success rate is 77.970% with the average attack success time of 2.096s. The attack success time is calculated as the difference between the attack start time and the time when the IMA warning of the victim vehicle is triggered. Given that the frequency of the trajectory data is 2.5HZ, the average attack succeeds at around the fifth time step.

One explanation for the attack failure is the vehicle speed variation before entering the roundabout. Figure 6 shows an example of the speed profile of an entry vehicle (victim vehicle). The proposed model fails to generate an attack trajectory to trigger the IMA warning in this case. The entry vehicle accelerates from $2m/s$ to $7m/s$ in two seconds and the estimated arrival time changes from 6.23s to 0.33s. Even though the estimated arrival time is updated every 0.4s, the large variation makes it impossible for the attack trajectory to reach the conflict point in time, without violating vehicle dynamic constraints.

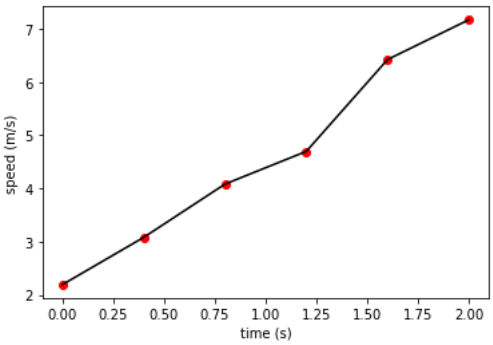

Fig. 6. Attack Failure example at Roundabout

B. Detection framework evaluation

To evaluate the detection framework, 2564 ground truth trajectories from the data set are extracted. 490 attack trajectories are generated with the proposed attack model. 40% of the ground truth are used to train the one-class classifier and the rest 60% are used for testing. All of the attack trajectories are used for testing. Both offline and online detection are conducted and the results are presented as follows.

In the offline mode, the detection is not performed until the full trajectory is observed. 99.59% (488/490) of the attack trajectory and 99.48% (1531/1539) of the ground truth trajectory are identified correctly. A false positive means that a ground truth trajectory is classified as an abnormal trajectory while a false negative means that an abnormal trajectory is identified as a normal trajectory. The false positive rate and false negative rate for the offline detection is 0.52% and 0.40% respectively. Figure 7 shows a false negative case in which the attack succeeds within only one time step. The short attack success time leads to little information can be used for detection. Therefore, the trajectory is not identified correctly.

Fig. 7. Misclassification example (FN case)

The online detection is more important in real-world implementations. The detection starts after sufficient number of trajectory points (e.g., 3 data points) and is conducted every time step until the trajectory is identified as abnormal or the attack succeeds. The trajectory will be identified as abnormal if it is classified as abnormal in two consecutive time steps. Therefore, trajectories used for online detection should be at least 5 time steps long. 314 attack trajectories and 1539 ground truth trajectories are used to test the online detection. The online detection performance is shown in Table I. The mean detection time is the elapsed time when the trajectory is identified as abnormal. The time to attack succeed is the difference between the attack success time and the detection time. Results show that the anomaly classifier can identify attack trajectories 0.49s before attack succeed in average, with a standard deviation equals to 0.22s.

Figure 8 shows a false positive case in the online detection. The attack trajectory speed profile shows that the vehicle's speed fluctuates between $1m/s$ to $7m/s$ in 1.2s. The average speed of the vehicle is $5.8m/s$. The large fluctuation under low travel speed are rare in the roundabout data set. As a result, the trajectory is classified as abnormal. A possible reason is due to the error in the trajectory processing. Even this trajectory is not under attack, its erroneous behavior indicates that it is not a normal trajectory and further attention is needed to identify the root cause.

TABLE I
PERFORMANCE OF ONLINE DETECTION

FP	FN	Mean attack success time (s)	Mean detection time (s)	Mean time to attack success (s)
14/1539 (0.91%)	0/314 (0%)	2.096	1.600	0.497

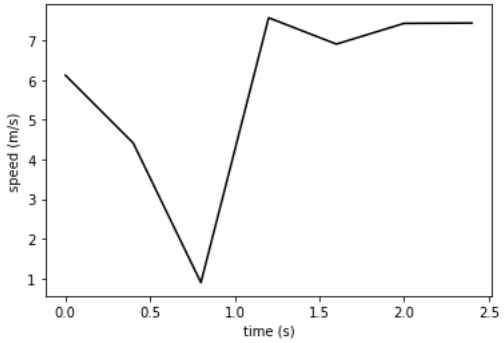

Fig. 8. Misclassification example (FP case)

VI. Conclusions and Discussions

In this paper, a GPS spoofing attack model towards the CV IMA application and an anomaly detection framework to detect such attacks using one class classification is introduced. An optimization model is formulated and served as the threat model which aims to trigger the IMA warning of the entry vehicle at the roundabout. Both models are evaluated with a real world data set and the results show that the threat model can generate falsified trajectory and trigger the IMA warning within a short time, which is very aggressive and raises challenges to the detection model. However, the detection result shows satisfactory performance that most of the abnormal trajectories can be identified correctly and in time.

Comparing with previous work on GPS spoofing attack, the proposed attack model in this paper is much more aggressive. For example, the average attack success time using algorithm proposed by [13] is 28.7s. In our work, the attack success time is only around 1.7s, which leaves little time for the detection model. In addition, the proposed detection model is more generic. Comparing with our previous work [18] which uses both ground truth trajectories and attack trajectories in the training process, the proposed detection framework based on one class classification only need ground truth trajectories for training, which makes it applicable to detect unknown attacks. Besides, the proposed anomaly detection model can be applied to multiple scenarios including IMA warning at the roundabout and unsignalized intersection, as well as Red

Light Violation Warning (RLVW) at signalized intersections, since the proposed model focus on learning the normal driving behaviors. As long as the normal driving behavior is affected, the proposed method can be applied to detect the anomaly.

ACKNOWLEDGMENT

This research is supported in part by the U.S. National Science Foundation through Grant SaTC #1930041, CNS-1929771, CNS-2145493 and USDOT UTC Grant 69A3552047138. The views presented in this paper are those of the authors alone.

REFERENCES

[1] S. Campbell, N. O'Mahony, L. Krpalcova, D. Riordan, J. Walsh, A. Murphy, and C. Ryan, "Sensor technology in autonomous vehicles : A review," in *2018 29th Irish Signals and Systems Conference (ISSC)*, 2018, pp. 1–4.

[2] X. He, E. Hashemi, and K. H. Johansson, "Distributed control under compromised measurements: Resilient estimation, attack detection, and vehicle platooning," *Automatica*, vol. 134, p. 109953, 2021.

[3] T. E. Humphreys, B. M. Ledvina, M. L. Psiaki, B. W. O'Hanlon, P. M. Kintner *et al.*, "Assessing the spoofing threat: Development of a portable gps civilian spoofer," in *Proceedings of the 21st International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2008)*, 2008, pp. 2314–2325.

[4] Z. Ju, H. Zhang, X. Li, X. Chen, J. Han, and M. Yang, "A survey on attack detection and resilience for connected and automated vehicles: From vehicle dynamics and control perspective," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 4, pp. 815–837, 2022.

[5] Z. Ju, H. Zhang, and Y. Tan, "Distributed deception attack detection in platoon-based connected vehicle systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4609–4620, 2020.

[6] Q. Liu, Y. Mo, X. Mo, C. Lv, E. Mihankhah, and D. Wang, "Secure pose estimation for autonomous vehicles under cyber attacks," in *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2019, pp. 1583–1588.

[7] M. Maile, Q. Chen, G. Brown, and L. Delgrossi, "Intersection collision avoidance: From driver alerts to vehicle control," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, 2015, pp. 1–5.

[8] S. Narain, A. Ranganathan, and G. Noubir, "Security of gps/ins based on-road location tracking systems," in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2019, pp. 587–601.

[9] M. O. of Highway Safety Planning. Michigan traffic crash facts. [Online]. Available: https://www.michigantrafficcrashfacts.org/querytool#q1;0;2021

[10] P. Oza and V. M. Patel, "One-class convolutional neural network," *IEEE Signal Processing Letters*, vol. 26, no. 2, pp. 277–281, 2018.

[11] S. Parkinson, P. Ward, K. Wilson, and J. Miller, "Cyber threats facing autonomous and connected vehicles: Future challenges," *IEEE transactions on intelligent transportation systems*, vol. 18, no. 11, pp. 2898–2915, 2017.

[12] H.-S. Seo, D.-G. Noh, C.-J. Lee, and S.-S. Lee, "Design and implementation of intersection movement assistant applications using v2v communications," in *2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2013, pp. 49–50.

[13] J. Shen, J. Y. Won, Z. Chen, and Q. A. Chen, "Drift with devil: Security of {Multi-Sensor} fusion based localization in {High-Level} autonomous driving under {GPS} spoofing," in *29th USENIX Security Symposium (USENIX Security 20)*, 2020, pp. 931–948.

[14] Z. Tian, Y. Cai, S. Huang, F. Hu, Y. Li, and M. Cen, "Vehicle tracking system for intelligent and connected vehicle based on radar and v2v fusion," in *2018 Chinese Control And Decision Conference (CCDC)*, 2018, pp. 6598–6603.

[15] N. O. Tippenhauer, C. Pöpper, K. B. Rasmussen, and S. Capkun, "On the requirements for successful gps spoofing attacks," in *Proceedings of the 18th ACM conference on Computer and communications security*, 2011, pp. 75–86.

[16] F. van Wyk, Y. Wang, A. Khojandi, and N. Masoud, "Real-time sensor anomaly detection and identification in automated vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 1264–1276, 2020.

[17] Y. Wang, N. Masoud, and A. Khojandi, "Real-time sensor anomaly detection and recovery in connected automated vehicle sensors," *IEEE transactions on intelligent transportation systems*, vol. 22, no. 3, pp. 1411–1421, 2020.

[18] Z. Yang, "An infrastructure-based cooperative driving framework for connected and automated vehicles," Ph.D. dissertation, 2022.

[19] K. C. Zeng, S. Liu, Y. Shu, D. Wang, H. Li, Y. Dou, G. Wang, and Y. Yang, "All your {GPS} are belong to us: Towards stealthy manipulation of road navigation systems," in *27th USENIX security symposium (USENIX security 18)*, 2018, pp. 1527–1544.

[20] R. Zhang, Z. Zou, S. Shen, and H. X. Liu, "Design, implementation, and evaluation of a roadside cooperative perception system," *Transportation Research Record*, p. 03611981221092402, 2022.