Intrusion Detection System for Platooning Connected Autonomous Vehicles

Dimitrios Kosmanos* dikosman@uth.gr
Apostolos Pappas* apopappas@uth.gr
Francisco J. Aparicio-Navarro† fnavarro@dmu.ac.uk
Leandros Maglaras† leandros.maglaras@dmu.ac.uk
Helge Janicke† heljanic@dmu.ac.uk
Eerke Boiten† eerke.boiten@dmu.ac.uk
Antonios Argyriou* anargyr@uth.gr

**Department of Electrical and Computer Engineering*
University of Thessaly
Volos, Greece
†*Faculty of Computing, Engineering and Media*
De Montfort University
Leicester LE1 9bH, UK

***Abstract*—The deployment of Connected Autonomous Vehicles (CAVs) in Vehicular Ad Hoc Networks (VANETs) requires secure wireless communication in order to ensure reliable connectivity and safety. However, this wireless communication is vulnerable to a variety of cyber atacks such as spoofing or jamming attacks. In this paper, we describe an Intrusion Detection System (IDS) based on Machine Learning (ML) techniques designed to detect both spoofing and jamming attacks in a CAV environment. The IDS would reduce the risk of traffic disruption and accident caused as a result of cyber-attacks. The detection engine of the presented IDS is based on the ML algorithms Random Forest (RF), k-Nearest Neighbour (k-NN) and One-Class Support Vector Machine (OCSVM), as well as data fusion techniques in a cross-layer approach. To the best of the authors' knowledge, the proposed IDS is the first in literature that uses a cross-layer approach to detect both spoofing and jamming attacks against the communication of connected vehicles platooning. The evaluation results of the implemented IDS present a high accuracy of over 90% using training datasets containing both known and unknown attacks.**

***Index Terms*—Intrusion Detection Systems, Connected Autonomous Vehicles, Vehicular Ad Hoc Networks**

I. Introduction

The deployment of Connected Autonomous Vehicles (CAVs) is considered the key factor to enhance road safety, increase the infrastructure efficiency, and reduce fuel consumption in Intelligent Transportation Systems (ITS) [1]. Adaptive Cruise Control (ACC) can automatically regulate parameters such as speed changes and gaps between vehicles by using on-board sensors.

Vehicle platooning is an application for semi-autonomous cooperative driving that comprises a leading vehicle and a group of following vehicles. The motion of the vehicles forming a platoon is determined by the Cooperative Adaptive Cruise Control (CACC) technology [2]. CACC is an enhancement to ACC that introduces Vehicle-to-Vehicle (V2V) communications, and allows vehicles to travel in more compact and stable platoons than ACC [1]. Most CACC systems require the following vehicle to communicate with its nearest preceding vehicle and/or the leading vehicle of the platoon [1].

Vehicle platooning is achieved by the exchange, in real-time, of information about the longitudinal (e.g. acceleration and braking) and lateral (e.g. steering) control system of the vehicles, as well as management protocols that supervise the formation of the platoon, driving maneuver and platoon disengagement [3]. This information is shared by the exchange of Cooperative Awareness Messages (CAMs) between the connected vehicles. These messages are transmitted several times per second using Dedicated Short Range Communication (DSRC) and Wireless Access in Vehicular Environments (WAVE) technology, based on the IEEE 802.11p standard.

The CAV infrastructure requires secure wireless communication channels in order to ensure reliable connectivity and safety. Connected vehicles are permanently interconnected by periodically broadcasting CAMs. However, these messages are vulnerable to a wide range of cyber threats, such as eavesdropping, spoofing and modification attacks. For example, a spoofing attack against the communication between CAVs could allow an attacker to change the distance between autonomous vehicles within the platoon, disrupting the flow of traffic and increasing the chances of accident. Moreover, the wireless communication channel is exposed to RF jamming (e.g. radio signals maliciously emitted to disrupt the legitimate communication) and Denial-of-Service (DoS) attacks [4]. Jamming against CAMs can be implemented easily, and can disrupt the performance of platooning [5]. For this reason, it is important to design innovative and robust cyber security solutions that can successfully protect the technology powering CAVs against cyber-attacks.

This paper describes the work conducted by the Cyber Technology Institute at De Montfort University, as part of the Innovate UK-funded research project *FLOURISH*.

In this paper, an Intrusion Detection System (IDS) based on Machine Learning (ML) developed to detect spoofing and jamming attacks in a CAV environment, with special focus on vehicles' platooning communication is presented. This is the first proposed ML IDS in the literature that can effectively detect both spoofing and jamming attacks. The detection engine of the proposed IDS is based on several ML algorithms like Random Forest (RF), k-Nearest Neighbour (k-NN) and One-Class Support Vector Machine (OCSVM), as well as the use of data fusion techniques in a cross-layer approach. both supervised learning techniques used are very popular, with the k-NN being robust against noisy training data like the ones obtained from a real-life urban environment and RF being one of the most accurate algorithms, due to the fact that it reduces the chance of over-fitting (by averaging several trees, there is a significantly lower chance of over-fitting). Moreover, except of supervised ML techniques, we also use a semi-supervised ML technique (OCSVM), in order for the proposed IDS to be applicable in cases where only one class data (normal data) exists. This is a common case, because vehicular communication datasets obtained by a real testbed comprising also traces of cyber-attacks are rarely available. Last, the proposed IDS has as a goal to achieve the fusion of the two used supervised ML algorithms using data fusion techniques with special purpose to enhance the overall performance.

The main contribution of this paper is the development of a novel cross-layer IDS for CAV platooning capable of detecting a spoofing attack and a reactive jamming attack too. The novelty of the proposed IDS is the cross-layer set of features that are utilized for both training and testing. The second contribution is that the implemented IDS can produce probabilistic results for both known and unknown attacks.

The rest of this paper is organised as follows. Section II provides an overview of related work in the domain of spoofing and jamming attack detection. Section III describes the topology and the types of the implemented attacks. Section IV describes the proposed probabilistic IDS. Section V presents the experimental evaluation setup, the impact of the implemented attack in V2V communication and finally, the experimental results. Section VI summarizes our findings and concludes our work.

II. Related Work

A. Spoofing Attacks

The literature in the area of cyber security for connected vehicles is divided in two distinctive areas of interest. Firstly, the techniques that use metrics from the Application layer (APP), e.g. speed-deviation, such as Acceptance Range Threshold (ART) [6]. Speed Deviation Verification at consecutive time intervals has been also used for the verification of each vehicle location. However, this metric is vulnerable against GPS spoofing attacks. Swaszek et al. [7] consider the use of range-only information to detect Global Navigation Satellite System (GNSS) spoofing of a platoon of vehicles equipped with inter-vehicle communications. However, this paper considers the use of short range only information communicated amongst a platoon of vehicles to detect GNSS spoofing. These methods are mainly based on upper layer metrics, the honesty of nearby vehicles and the traffic density of spoofing attackers.

Secondly, there is a specific area in which the publications also use metrics from the Physical (PHY) layer, such as the Received Signal Strength (RSS), and metrics from the Application (APP) layer, such as speed-deviation of nodes [8]. In various publications, the strength distribution analysis is used to detect Sybil or Spoofing attacks [9]. Last, the authors in [10] propose a solution to correct the wrong position given by the fake GPS. The correction is based on a validation process by comparing the given position to an RSU using the wireless Vehicle-to-Infastructure communication (V2I). However, the wireless communication between the transmitter and the Roadside Unit (RSU) can be impaired by fast fading characteristics that exist in VANETs.

All the above publications do not use ML approaches for detecting spoofing attacks. On the other hand, several articles such as [11], [12] introduce the Received Signal Strength Indicator (RSSI)-based schemes for detecting spoofing attacks in Wireless Sensor Networks (WSNs) using ML techniques without using a cross-layer architecture.

Additionally, extensive works present the applications of spatial processing methods for GPS spoofing detection and mitigation that use either Phase Delay Measurements [13] or the Angle of Arrival (AoA) estimation [14] from the PHY layer to verify the message originator. From an attacker perspective, an illegitimate node may intentionally falsify the information to achieve a certain goal that might be rational in some scenarios. A drawback of using metrics from the PHY layer is the incorrect GPS spoofing detection (e.g. false alarms) that may occur in situations where multiple correct satellite signals are received from similar directions and phase delay differences are below a predefined threshold.

B. RF Jamming Attacks

ML techniques for jamming attack detection in vehicular ad-hoc networks have been proposed in the past [15], [16]. The authors use the metrics Noise and Channel busy Ratio (CbR), Packet Delivery Ratio (PDR), Maximum Inactive Time (Max IT) and RSSI in order to detect attacks using ML techniques and examine both reactive and constant jammers.

Several recent works have proposed machine-learning based techniques for jamming attack detection in vehicular ad-hoc networks. Puñal et al. [15] used the above mentioned metrics in order to detect attacks using machine learning techniques and examine both reactive and constant jamming. Azogu et al. [16] proposed a new mechanism, called Hideaway Strategy according to which all nodes should remain silent while the network is under a jamming attack. The authors in [5] propose a data mining-based method for real-time detection of radio jamming DoS attacks in IEEE 802.11p V2V communications for a platoon of vehicles. State-of-the-art methods are compared with the proposed method under the realistic assumption of random jitter accompanying every

CAM transmission. However, authors did not use a cross-layer set of features for the training procedure because only features from the network layer are utilized.

In contrast to all the aforementioned works, we use an additional metric: the **relative speed (Δu)** between the sender and the receiver. The novelty of this metric from the APP layer is that it can be estimated by the wireless channel of the PHY layer using the effect of the Doppler phenomenon. This estimated metric can be combined with other metrics from the APP and the PHY layer leading to a cross-layer detection approach. The proposed cross-layer detection approach can be used without the need for a position verification model which is subject to statistical errors. Specifically, from the APP layer we can use the GPS coordinates indicating the location of the sender and from the PHY different metrics such as the RSSI, the Signal to Interference and Noise Ratio (SINR) and PDR for the effective detection of a spoofing attack or a jamming attack, differentiating also one from the other.

III. System Model

A. Topology

The simulated CAV enviroment comprises of a platoon of four connected vehicles, whose motion is determined by the CACC technology [2]. As described previously, most CACC systems require the following vehicle to communicate with its nearest preceding vehicle and/or the leading vehicle of the platoon. As represented in Fig. 1, the leading vehicle (Veh1) and second vehicle (Veh2) represent the Receiver (*Rx*) and Transmitter (*Tx*) of messages, respectively. The third vehicle (Veh3) is the attacker that conducts a reactive jamming attack, and the fourth vehicle (Veh4) in the platoon is the attacker that conducts the spoofing attack.

The experimental results analysis presented in this paper focus on assessing the effect of the different attacks in the communication between the *Tx* and *Rx* vehicles.

B. Spoofing Attack in V2V Communications

All connected vehicles in a platoon periodically broadcast CAMs, known as beacon messages, in order to inform neighbouring vehicles of their presence. Each CAM comprises of several fields such as Node Identifier (Vid), Time instance (Time), the MAC address and current vehicle GPS location. However, the Vid and the MAC address of sender in the WAVE Service Advertisement (WSA) frame [17] can be modified by a spoofer. .

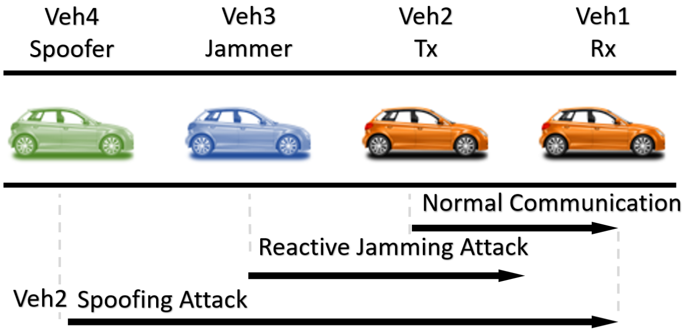

Figure 1: Schematic representation of the vehicles platooning topology: Receiver vehicle (Veh1), Transmitter vehicle (Veh2), Jammer vehicle (Veh3), and Spoofing vehicle (Veh4).

For the simulated attack scenario, initially, the *Tx* and *Rx* vehicles have a wireless connection established using the IEEE 802.11p MAC protocol, and drive in a platoon formation. The attacker Veh4 follows the *Tx* and *Rx* vehicles. When the distance between Veh4 and Veh2 is about 35m, the attacker intercepts the Vid and MAC address of Veh2 from the broadcasting CAMs and starts its spoofing attack. During the spoofing attack, Veh4 also broadcasts a CAM message every 0.1 seconds, using the Vid of Veh2, in order to inform the *Rx* about an incorrect GPS location and speed value. Since the attacker replicates the Vid and MAC address of Veh2, during the spoofing attack, there would be WSA frames showing discrepancies between the identity and the physical characteristic of the frames. The routing flow that is selected in transport layer is based on the the incorrect spoofed MAC address of the transmitter Veh2. This has as a consequence frame losses in PHY layer due to path losses and fast fading factors or due to the strict delay constraints of the backoff procedure in MAC layer. So many CAMs sent by the legal Tx are lost in MAC layer and are never acknowledged by the client, increasing the Packet Error Rate (PER) and decreasing the throughput too. So it is clear that the spoofing attack affects the communication channel. This attack can be also assumed as an another kind of a Denial of Service (DoS) attack. The communications problem that is provoked is discussed more thoroughly in Section IV-C. The designed IDS aims to detect these discrepancies in the communication channel.

As we previously said, the attacker exploits these fields to transmit false GPS location coordinates within the CAMs, which misdirects the platoon of connected vehicles to an incorrect location. This has as a consequence the observed RSSI values by the wireless communication between *Tx* and *Rx* to move to a different level, indicating the spoofing attack. Fig. 3 shows a comparison between the RSSI values and the distance between the *Tx* and *Rx* vehicles during the first two stages of the simulation (e.g. the initial period of normal traffic and the spoofing attack). When the position of the transmitter, which is the spoofer during the spoofing attack, is quite different from the legitimate's position the level of the RSSI values change significantly as can be seen in Fig. 3. This fact can indicate the spoofing attack, proving also that the RSSI maybe is a crucial metric for the detection of a spoofing attack using a cross-layer ML approach.

C. Jamming Attack in V2V communications

For the evaluation of the jamming attack scenario, a reactive radio frequency jammer has been implemented [18]. The RF jamming targets the IEEE 802.11p Orthogonal Frequency Division Multiplexing (OFDM) based PHY layer operating in the 5.85-5.925 GHz unlicensed national information infrastructure band, with 10 MHz bandwidth. In Fig. 4 shows the standard protocol WAVE IEEE 802.11p OFDM frame format, which consists of the OFDM PHY Layer Convergence Protocol (PLCP) preamble, PLCP header, PLCP Service Data Unit (PSDU), tail bits, and pad bits. In the PLCP preamble field, the preamble consists of ten identical short training symbols and two identical long training symbols. The OFDM signal has a fixed shape in the time domain and lasts $T_{const} = 64\mu s$. before the next OFDM signal can be transmitted, there is an idle time of $T_{prep} = 10\mu s$ required to set up the next transmission.

The jammer aims to block completely the communication between the pair of the *Tx* and *Rx* vehicles by transmitting in a reactive manner upon the detection of IEEE 802.11p frames in the communication channel, causing a DoS attack. Assunming the topology of Fig. 1, every time the *Tx* vehicle transmits a CAM message, Veh3 also transmits a CAM message to cause a collision. The reactive jammer starts the transmission of a CAM using OFDM signal with QPSK modulation upon sensing energy above a certain threshold. This threshold has been empirically set to -75 dbm for a certain time of $T_{detect} = 1.2\mu s$ as a good tradeoff between the jammer sensitivity and false transmission detection rate [4]. The reactive jammer starts transmitting when it is located at a distance of $35m$ from Veh1. Therefore, the *Rx* vehicle receives a combined signal from the jammer and the *Tx* vehicle with the form:

$$\vec{y} = h_1\vec{x} + h_2\vec{s} + \vec{w} \qquad (1)$$

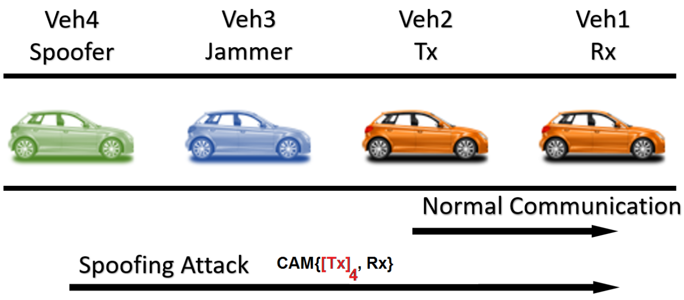

Figure 2: Spoofing Attack Scenario

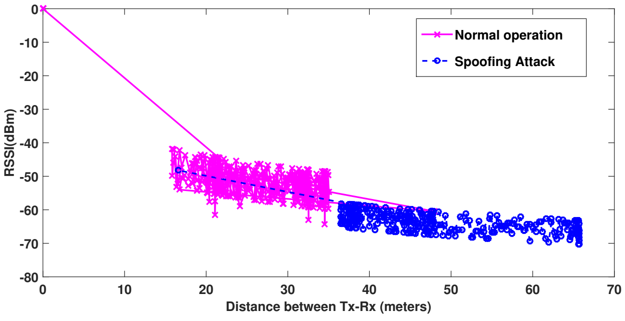

Figure 3: Different RSSI levels during the normal operation and spoofing attack

where $\vec{y}$ is the combined received baseband signal at the *Rx* vehicle, h_1 is the Rician fading channel between the *Tx* and *Rx* vehicles, h_2 is the Rician fading channel between the jammer Veh3 and the *Rx* vehicle, $\vec{x}$ is the valuable signal sent by the transmitter, $\vec{s}$ is the jamming signal sent by the jammer and $\vec{w}$ is white Gaussian noise. The total reaction $T_{reaction}$ is the sum of the detection time $T_{detect} = 1.2\mu s$ and the preparation time $T_{prep} = 10\mu s$. For the discussion of the results, we consider that the overall reaction time is, on average, $11.2\mu s$.

Although, the jamming misses the beginning of the IEEE 802.11p preamble, this noise signal overlaps with the PLCP, MAC and WSMP header of the IEEE 802.11p frame sent from *Tx* to *Rx*, as represented in Fig. 4. because of the reactive jamming, the *Rx* vehicle cannot process the CAM from the *Tx* vehicle due to insufficient SINR.

IV. Intrusion Detection Framework

An IDS is a fundamental element of security infastructure, aiming at identifying evidence of attacks or indications of suspicious activities in the system under protection. The use of ML techniques has gained wide interest in the area of network security and intrusion detection. An ML-IDS is based on models that allow the classification of the analysed information [19]. In the area of network security, the use of ML techniques has proven to improve the accuracy of an IDS [20]. We propose the use of an ML-IDS in the area of CAV communication security.

A. Machine Learning Techniques

The novel ML-IDS that we propose makes use of the supervised ML techniques k-NN and RF for the attack classification process. Additionally, experiments using the semi-supervised ML technique OCSVM have also been conducted.

The k-NN is a simple ML technique for pattern recognition, based on feature similarity [21]. When we say a technique is non-parametric, it means that it does not make any assumptions on the underlying data distribution. Therefore, k-NN could and probably should be one of the first choices for a classification study when there is little or no prior knowledge about the distribution data. k-NNN is also a lazy algorithm (as opposed to an eager algorithm). What this means is that it does not use the training data points to do any generalization. In other words, there is no explicit training phase or it is very minimal. This also means that the training phase is pretty fast.

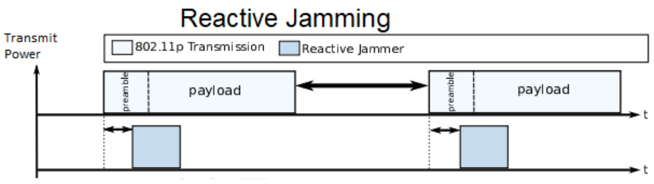

Figure 4: Reactive jamming against IEEE 802.11p frames.

So, the k-NN algorithm is also useful for non-linear data, which is the case for the data that we use for training.

The RF is a supervised learning algorithm, based on decision tree models that split a subset of features at training time and output the class that has the majority votes of the classes of the individual trees [22]. This supervised ML algorithm is preferred over others for the following reasons. Firstly, it can be used for both classification and regression tasks, providing high accuracy. Secondly, if there are more trees, it do not allow overfitting trees in the model. It has the power to handle a large data set with higher dimensionality. Last, RF classifier handles the missing values and maintains the accuracy of a large proportion of data.

The OCSVM is an effective semi-supervised classification technique that constructs the classification model of normal behaviour during the training process using only one type of samples (e.g. training datasets comprising of only non-malicious data). This feature makes OCSVM an ideal classification technique when only non-malicious training datasets are available Segmentation and clustering algorithms seem to be better choices because they do not need to know the signatures of the series. The shortages of such algorithms are that they always need parameters to specify a proper number of segmentation or clusters and the detection procedure has to shift from one state to another state. In order the above mentioned drawbacks to be minimized, the goal of an OSVM is to find the optimal separating hyperplane which maximises the margin of the training data and minimises the chance of accepting outliers [23].

B. Data Fusion Techniques

Each of the classification algorithms is able to generate accurate results when implemented independently. However, the combined use of these algorithms may help improve the overall performance of an IDS [24]. Different methodologies were evaluated to assess whether the classification results could be improved, for instance, by applying data fusion techniques.

Ensemble learning has been used to combine the outputs from different classification techniques. Ensemble learning is the process in which multiple classifiers are strategically selected and combined in order to solve a particular computational intelligence problem. Ensemble learning is primarily used to improve the classification performance of a model. One of commonly used ensemble learning algorithms is known as bagging [25]. In this algorithm, bootstrapped replicas of the training data for each classifier (RF, k-NN) are used. During the last step of bagging, the majority voting combination rule is used. Since the intended output of the IDS is a probabilistic IDS, the conditional probabilities are estimated for each classifier in the presented IDS using the bayesian rule as data fusion technique.

C. CAV Communication Dataset

The presented IDS uses metrics from both, the PHY and APP layers. From the PHY layer we extract the RSSI, the SINR and the PDR. From the APP layer we extract the Relative Speed (Δu) and the GPS coordinates. All these metrics, listed in Table I, are used in a cross-layer approach to improve the detection accuracy of the detection system. Furthermore, for the training-testing procedure of the proposed IDS, the data have been divided into 70% for training and 30% for testing.

The Relative Speed (Δu), introduced in [26], indicates the relative speed between an attacker and the receiver Veh1:

$$\Delta u_A = |\vec{u}_A - \vec{u}_{Rx}| \tag{2}$$

where $\vec{u}_A$, $\vec{u}_{Rx}$ are the speed of the attacker and the receiver, respectively. The metric Δu can be effectively estimated by the RF signals' interchange in the PHY layer. The novelty of this metric is that it can be estimated by the physical properties of the wireless channel, using the effect of the Doppler phenomenon. For the jamming attack, we use again the Doppler effect in order to estimate the Δu between the jammer and the receiver from the combined value of the useful and the jamming signal at the receiver, as described in [27].

The scarcity of publicly available real vehicular communication datasets has been previously discussed in [28]. Currently, no vehicular communication dataset is available comprising traces of cyber-attacks. Computer simulation software becomes the only available alternative to conduct cyber security research in the area of connected vehicles.

We have developed an experimental simulation testbed using Veins [29] to evaluate the proposed IDS for a platoon of vehicles as depicted in Fig. 5 in a part of the Erlangen city. The simulation comprises a flow of four connected vehicles in a platoon formation. This simulation consists of 800 timesteps, of which 300 timesteps correspond to the normal operation of the system, 200 timesteps correspond to a spoofing attack and the remaining 300 timesteps correspond to a jamming attack.

In order to show the effect of the two attacks in the communication between the *Tx* and *Rx* vehicles, the Throughput has been plotted in Fig. 6. The Y-axis represents the throughput

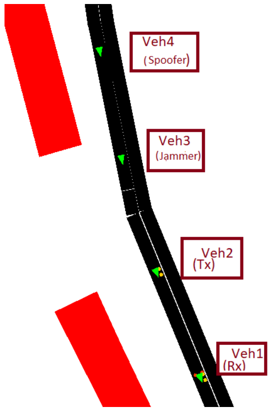

Figure 5: Close-up view of the Erlangen city map used for conducting the simulations. The four vehicles platooning are marked in green, moving south along the road.

in Mbps, whereas the X-axis represents the time in seconds. The normal (e.g. without attack) communication between the *Tx* and *Rx* vehicles is represented in $pink$, the spoofing attack is represented in $blue$, and the jamming attack is represented in $green$. The normal communication between the *Tx* and *Rx* vehicles occurs in two periods, between the time interval 15-25 second, and the time interval 75-95 seconds. The spoofing attack is launched during the time interval 25-45 seconds. Finally, the jamming attack is launched during the time interval 45-75 seconds.

As can be seen in Fig. 6, the average throughput for the normal communication is 18 Mbps, approximately. When the spoofing attack is launched, the average throughput drops to 10 Mbps. This change in the throughput clearly shows that the modification of the GPS coordinates and speed values within the CAM messages has a clear effect upon the communication between the connected vehicles. Even more noticeable is the effect of the jamming attack, where the average throughput reaches 0.5 Mbps. This makes the communication between the *Tx* and *Rx* vehicles almost impossible. In this simulation, the *Tx* vehicle broadcasts a CAM message every 0.1 seconds in order to inform the *Rx* about its current GPS location and speed.

The experiments have been conducted using the simulation parameters presented in above: The minimum distance between the jammer from Veh1 ($minDist_{Jx-Rx}$) is 25m, the minimum distance between Veh1 and Veh2 ($minDist_{T_x-R_x}$) is 15m, the minimum distance between the spoofer and Veh1 ($minDist_{Veh4-R_x}$) is 35m, Transmission Signal Strength is 100mW, Packet Length is 500 bits, and Data Rate is 18Mbps. both vehicles Veh1 and Veh2 move at a maximum speed of 10m/s, whereas both vehicles Veh3 and Veh4 move at a maximum speed of 25m/s. The jammer starts transmitting within a radius of 35m from Veh1. This simulation replicates a realistic example, in which the attackers show distinctive moving behaviours from the legitimate vehicles.

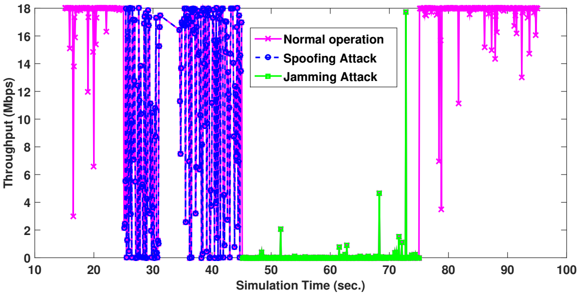

Figure 6: Throughput (Mbps) of the communication between the *Tx* and *Rx* vehicles during the experimental simulation. The normal communication without attack in $pink$, spoofing attack in $blue$, and jamming attack in $green$.

V. EXPERIMENTAL EVALUATION

We evaluated the performance of our attack detector by using a detection rate and receiver operating characteristic (ROC) curve. It is proved that the proposed IDS presents a high accuracy of over 90% using training datasets containing both known and unknown attacks.

A. Spoofing Attack Detection Results

The experiments to evaluate the efficiency of the developed IDS have been conducted using k-NN and RF. The ML techniques have been trained with both *Attack* and *No Attack* instances. The initial set of experiments has been conducted using single metric approach only, using the metrics described in Table I.

Table II presents the classification accuracy results when using single feature for each of the supervised ML techniques. As it can be seen by using metrics from the PHY layer, the IDS achieves the highest results, reaching over 91% accuracy with both techniques. The best detection is achieved by k-NN using the RSSI (e.g. 95.88% accuracy). Focusing on the metrics from the APP layer, the accuracy drops to 72.43% when using RF to analyse the metric Position Y (e.g. longitudinal control). Furthermore, it has been shown that the use of the k-NN algorithm outperforms the accuracy of the RF algorithm for the majority of metrics. Generally, in most cases k-NN performs better in dataset with low dimensional space. The results have been plotted using a ROC curve for each single feature when using k-NN and RF in Figs. 7 and 8, respectively.

Focusing on a multi-metric approach, using only features from the APP layer with k-NN, the detection reaches 98% accuracy. On the other hand, the detection results when using only features from the PHY layer reaches 99% accuracy. These results clearly show the improvement achieve by using a multi-metric detection approach.

Table II: Accuracy of the cross-layer classification

Metric	k-NN	Random Forest
RSSI	95.88%	95.47%
SINR	91.77%	92.18%
PDR	92.59%	94.24%
Estimated Δu	86.83%	78.6%
Position Y	84.77%	72.43%
Position X	94.24%	92.59%

Table I: Metrics that are jointly processed by the classification algorithms

ID	Model Feature	Short Description
1	Δu	Estimated relative speed between J_x-R_x (m/sec.)
2	GPS cords	GPS coordinated in x-axis, y-axis indicating the location
3	RSSI	Signal Strenght Indicator (dbm)
4	SINR	Signal Quantity Indicator (db)
5	PDR	Packet Delivery Ratio

Finally, the IDS that we have developed takes advantage on a cross-layer architecture, using metrics from both the PHY and the APP layers. The combination of all the considered features generate the best attack detection accuracy overall. by using cross-layer k-NN, the IDS generates 99.59% detection accuracy and 98% accuracy using the RF algorithm.

For the jamming attack detection, we have conducted two versions of the reactive Jamming attack. One generic version that transmits activity is sensed on the wireless channel. The other version is an intelligent reactive jamming attack that reduces the number of intended collisions in order to minimise its exposure to be detected by an IDS. For the former version of the jamming attack, the developed IDS achieves 100% accuracy using k-NN and 99% accuracy using RF. Whereas the IDS reaches 95% accuracy detecting the intelligent jamming version. The last detection result is achieved by both ML techniques with cross-layer approach.

B. Multi-Attack Detection Results

In order to evaluate the adaptability of the presented IDS, additional experiments have been conducted. These experiments combine the two implemented attacks, spoofing and jamming. by using the k-NN algorithm with a cross-layer approach, the IDS generates almost perfect detection (e.g. 99% accuracy), whereas the RF algorithm reaches 96% accuracy approximately. Fig. 9 represents the detection results using ROC curves when both attacks are included in the training and testing datasets. The ROC in pink represents the detection results for the experiments using only jamming attack and normal traffic in the training and testing dataset, the one in green represents the results using only spoofing attack and normal traffic, and the ROC in blue represents the results when only both attacks are used in the training and testing dataset. As can be seen, worst classification results are generated when only the two implemented attacks are considered. In this case, the proposed IDS is difficult to differentiate the two implemented attacks because a lot of metrics indicating the communication problem (such as SINR, PDR) between transmitter and receiver get quite low values for both attacks.

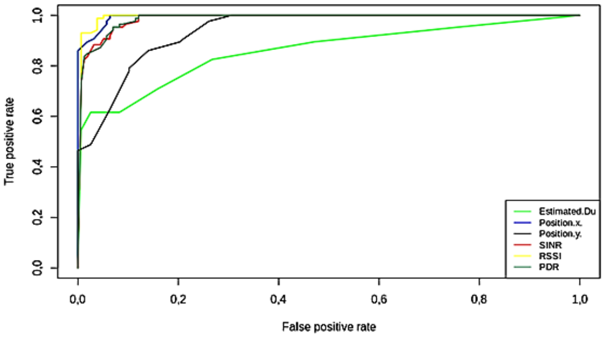

Figure 7: Spoofing Attack Detection: ROC curves for single feature spoofing attack classification using the k-NN algorithm.

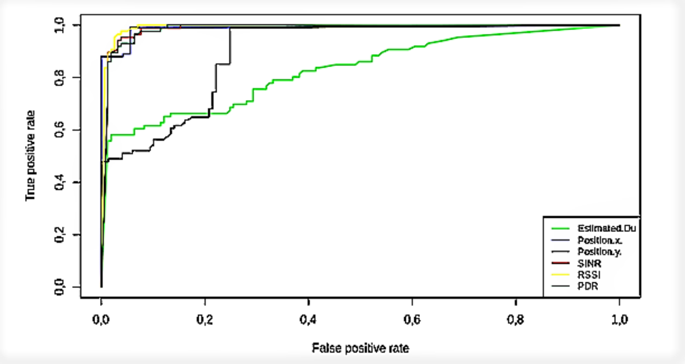

Figure 8: Spoofing Attack Detection: ROC curves for single feature spoofing attack classification using the RF algorithm.

C. Data Fusion Technique Results

In this subsection, we want to combine the two supervised ML classifiers (k-NN, RF) in order to achieve better performance than a single decision from one classifier.

In the proposed IDS, firstly, bootstrapped replicas of the training data for each classifier are used. The intended output of the IDS is a probabilistic IDS, which estimates the conditional probability of an observation to belong to class $Attack$ or $No\ Attack$, given the probability that each of the classifiers predicts the same class. For this reason, conditional probabilities are estimated for each classifier in the presented IDS using the bayesian rule as data fusion technique. Last, bagging is used in order to provide with a final prediction based on the aggregation of above predictions from the two supervised ML classifiers. The final prediction is estimated using the majority voting combination rule.

In subsequent Fig. 10 the accuracy results achieved for the spoofing attack detection by the Data Fusion technique are compared with these that achieved using only k-NN or RF. The above reported accuracy results have been plotted using a ROC for each classifier and a different ROC for the Data Fusion technique. The Data Fusion approach combined the outcome of the two supervised ML classifiers. It is observed that it achieves the same accuracy with the k-NN algorithm, which in this case has an almost perfect result. It can also

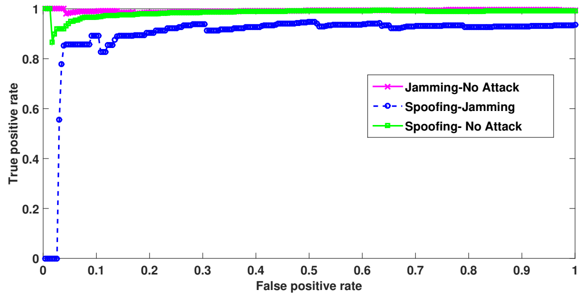

Figure 9: Multiclass results with class by class ROCs for the RF classifier

enhance the performance of the RF algorithm, which in our experiments shows the worst accuracy results.

D. Detection using Normal Data Training

The supervised ML algorithms require training datasets comprising *Attack* and *No Attack* instances. In cases where only normal (e.g. non-malicious) data is available, the use of semi-supervised ML algorithms is required. This section describes a final set of experiments focusing on the use of the semi-supervised ML algorithm OCSVM. The training dataset is 50% smaller than the testing dataset, and comprises normal data only. The training dataset comprises 20% of overall malicious data. The accuracy obtained by the OCSVM algorithm reaches 90% accuracy approximately. In this set of experiments using the OCSVM algorithm after training only with *No Attack* data, the IDS generates 40 False Negatives (FNs), 41 False Positives (FPs) alarms, 527 True Positives (TPs) indicating the $No\ Attack$ class and 200 True Negatives (TNs) indicating the $Attack$ class.

In most cases, no vehicular communication dataset is available comprising traces of cyber-attacks. From the above results, we can conclude that in cases where there is no malicious data the performance of the proposed IDS is not affected a lot.

VI. Conclusions

In this paper, we describe an IDS based on ML techniques designed to detect both spoofing and jamming attacks in a CAV environment. The IDS would reduce the risk of traffic disruption and accident caused as a result of cyber-attacks. The detection engine of the presented IDS is based on the ML algorithms RF, k-NN and OCSVM. To the best of the authors knowledge, the proposed IDS is the first in the literature that uses a cross-layer approach to detect both spoofing and jamming attacks against the communication of connected vehicles platooning. Various features from the APP and PHY layers have been extracted and analyzed.

In order to evaluate the efficiency of the developed IDS against different type of attacks, the vehicular network simulator Veins [29] has been considered. Although the experiments have been conducted using datasets from a simulated CAV environment, with vehicles platooning, the same IDS could be used to detect similar type of attacks launched from fixed location, for example, from a building on the side of the road.

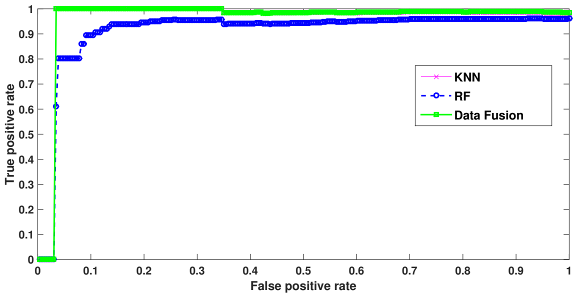

Figure 10: Accuracy evaluation for spoofing attack detection comparison among Data Fusion, k-NN, RF

The empirical analysis proves that both attacks impact on the communication between the transmitter and receiver vehicles. Overall, the features from the PHY layer weight out those from the APP layer in their contribution to the classification process, helping to detect correctly between *Attack* and *No Attack* using a cross-layer approach. In order to verify the adaptability of the proposed IDS, multiple sets of experiments have been conducted. The presented results shows that the proposed IDS can efficiently detect both attacks with high accuracy. The proposed IDS can also produce a measure of confidence or probabilistic classification result, instead of a binary classification (e.g. *Attack* or *No Attack*).

As future work, we will test the proposed IDS under more complicated scenarios with more wireless interference taking into consideration cases with more communicated nodes and attackers in the entire area.

References

[1] Z. Chen and B. B. Park, "Preceding vehicle identification for cooperative adaptive cruise control platoon forming," in *IEEE Transactions on Intelligent Transportation Systems (Early Access)*. IEEE, 2019, pp. 1 – 13.

[2] V. Milans, S. E. Shladover, J. Spring, C. Nowakowski, H. Kawazoe, and M. Nakamura, "Cooperative adaptive cruise control in real traffic situations," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, 2014, pp. 296–305.

[3] S. Santini, A. Salvi, A. S. Valente, A. Pescapè, M. Segata, and R. L. Cigno, "Platooning maneuvers in vehicular networks: A distributed and consensus-based approach," in *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 1, 2019, pp. 59–72.

[4] O. Punal, C. Pereira, A. Aguiar, and J. Gross, "Experimental characterization and modeling of RF jamming attacks on VANETs," in *IEEE Transactions on Vehicular Technology*, vol. 64, 2015, pp. 524–540.

[5] N. Lyamin, D. Kleyko, Q. Delooz, and A. Vinel, "Real-Time Jamming DoS Detection in Safety-Critical V2V C-ITS Using Data Mining," in *IEEE Communications Letters*, vol. 23, 2019, pp. 442–445.

[6] R. W. van der Heijden, A. Al-Momani, F. Kargl, and O. M. F. Abu-Sharkh, "Enhanced position verification for VANETs using subjective logic," in *IEEE 84th Vehicular Technology Conference (VTC-Fall)*, 2016.

[7] P. F. Swaszek, R. J. Hartnett, and K. C. Seals, "Using range information to detect spoofing in platoons of vehicles," in *30th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS)*, 2017, pp. 2838–2853.

[8] J. Grover, N. K. Prajapati, V. Laxmi, and M. S. Gaur, "Machine learning approach for multiple misbehavior detection in vanet," in *International Conference on Advances in Computing and Communications*, vol. 192. SpringerLink, 2011, pp. 644–653.

[9] B. Yu, C.-Z. Xu, and B. Xiao, "Detecting sybil attacks in VANETs," in *Journal of Parallel and Distributed Computing*, vol. 73. Elsevier, 2013, pp. 746–756.

[10] B. Anouar, B. Mohammed, G. Abderrahim, and B. Mohammed, "Vehicular navigation spoofing detection based on V2I calibration," in *4th IEEE International Colloquium on Information Science and Technology (CiSt)*, 2016.

[11] E. M. de Lima Pinto, R. Lachowski, M. E. Pellenz, M. C. Penna, and R. D. Souza, "A machine learning approach for detecting spoofing attacks in wireless sensor networks," in *32nd IEEE International Conference on Advanced Information Networking and Applications (AINA)*, 2018.

[12] Y. Chen, J. Yang, W. Trappe, and R. P. Martin, "Detecting and localizing identity-based attacks in wireless and sensor networks," in *IEEE Transactions on Vehicular Technology*, vol. 59, 2010, pp. 2418–2434.

[13] J. Magiera and R. Katulski, "Detection and mitigation of GPS spoofing based on antenna array processing," in *Journal of Applied Research and Technology*, vol. 13. ScienceDirect, 2015, pp. 45–57.

[14] A. Abdelaziz, R. Burton, and C. E. Koksal, “Message authentication and secret key agreement in VANETs via angle of arrival,” in *2016 IEEE Vehicular Networking Conference (VNC)*, 2016.

[15] O. Puñal, I. Aktaş, C.-J. Schnelke, G. Abidin, K. Wehrle, and J. Gross, “Machine learning-based jamming detection for IEEE 802.11: Design and experimental evaluation,” in *15th IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2014, pp. 1–10.

[16] I. K. Azogu, M. T. Ferreira, J. A. Larcom, and H. Liu, “A new anti-jamming strategy for VANET metrics-directed security defense,” in *IEEE Globecom Workshops (GC Wkshps), 2013 IEEE*, pp. 1344–1349.

[17] C. Campolo, H. A. Cozzetti, A. Molinaro, and R. Scopigno, “Augmenting vehicle-to-roadside connectivity in multi-channel vehicular ad hoc networks,” in *Journal of Network and Computer Applications*, vol. 36, no. 5. Elsevier, 2013, pp. 1275–1286.

[18] D. Kosmanos, N. Prodromou, A. Argyriou, L. A. Maglaras, and H. Janicke, “MIMO techniques for jamming threat suppression in vehicular networks,” in *Mobile Information Systems*. Hindawi, 2016, pp. 1–9.

[19] P. García-Teodoro, J. Díaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, “Anomaly-based network intrusion detection: techniques, systems and challenges,” in *Computers & Security*, vol. 28, 2009, pp. 18–28.

[20] K. Ghanem, F. J. Aparicio-Navarro, K. G. Kyriakopoulos, and S. Lambotharan, “Support vector machine for network intrusion and cyber-attack detection,” in *Sensor Signal Processing for Defence Conference (SSPD)*, 2017, pp. 1–5.

[21] O. Sutton, “Introduction to k nearest neighbour classification and condensed nearest neighbour data reduction,” in *University lectures, University of Leicester*, 2012.

[22] A. Liaw, M. Wiener *et al.*, “Classification and regression by random forest,” in *R news*, vol. 2, no. 3, 2002, pp. 18–22.

[23] L. Maglaras and J. Jiang, “A real time OCSVM intrusion detection module with low overhead for scada systems,” in *International Journal of Advanced Research in Artificial Intelligence*, vol. 10, 2014, pp. 24–53.

[24] F. J. Aparicio-Navarro, K. G. Kyriakopoulos, and D. J. Parish, “A multi-layer data fusion system for wi-fi attack detection using automatic belief assignment,” in *World Congress on Internet Security (WorldCIS)*, 2012, pp. 45–50.

[25] D. P. Gaikwad and R. C. Thool, “Intrusion detection system using bagging ensemble method of machine learning,” in *IEEE International Conference on Computing Communication Control and Automation*, 2015.

[26] D. Karagiannis and A. Argyriou, “Jamming attack detection in a pair of RF communicating vehicles using unsupervised machine learning,” in *Vehicular Communications*, vol. 13. Elsevier, July 2018, pp. 56–63.

[27] D. Kosmanos, A. Argyriou, and L. Maglaras, “Estimating the Relative Speed of RF Jammers in VANETs,” in *Cryptography and Security of Computer Science*. Arxiv, 2018, pp. 1–13, unpublished.

[28] I. Mavromatis, A. Tassi, J. P. Robert, and N. Andrew, “A city-scale ITS-G5 network for next-generation intelligent transportation systems: Design insights and challenges,” in *17th International Conference on Ad Hoc Networks and Wireless, AdHoc-Now 2018*, vol. 11104. Springer International Publishing, pp. 53–63.

[29] C. Sommer, R. German, and F. Dressler, “Bidirectionally coupled network and road traffic simulation for improved IVC analysis,” in *IEEE Transactions on Mobile Computing*, vol. 10, 2011, pp. 3–15.