Security Engineering with Machine Learning for Adversarial Resiliency in Mobile Cyber Physical Systems

Felix O Olowononi [a], Danda B. Rawat*[a], Moses Garuba [a] and Charles Kamhoua [b]
[a]Data Science and Cybersecurity Center, Department of Electrical Engineering and Computer Science, Howard University, Washington DC 20059, USA. [b]U.S. Army Research Lab, Adelphi, MD, USA
felix.olowononi@bison.howard.edu, *danda.rawat@howard.edu, mgaruba@howard.edu, charles.a.kamhoua.civ@mail.mil

ABSTRACT

Recent technological advances provide the opportunities to bridge the physical world with cyber-space that leads to complex and multi-domain cyber physical systems (CPS) where physical systems are monitored and controlled using numerous smart sensors and cyber space to respond in real-time based on their operating environment. However, the rapid adoption of smart, adaptive and remotely accessible connected devices in CPS makes the cyberspace more complex and diverse as well as more vulnerable to multitude of cyber-attacks and adversaries. In this paper, we aim to design, develop and evaluate a distributed machine learning algorithm for adversarial resiliency where developed algorithm is expected to provide security in adversarial environment for critical mobile CPS.

Keywords: Vehicular Cyber Physical Systems, Machine learning, VANET, Security, False data injection, Resiliency, Adversarial, Bayesian model

1. INTRODUCTION

Cyber Physical Systems (CPS) operate by perceiving or sensing changes in their physical environment through sensors, analyze the input information and autonomously respond to the changes by making intelligent decisions through issuing commands to control physical objects or actuators. Furthermore, a CPS fuses the physical and cyber world in a real-time manner for making informed decision. CPS with the Internet-of-Things (IoT) is a core underlying concept of the fourth industrial revolution which is characterized by hyperconnectivity, hyperautomation and hyperintelligence[1]. CPS therefore is regarded a new technology that operates by combining already existing technologies like IoT, cloud computing, big data, machine-to-machine communications and multi-agent systems[2]. Since CPS applications include critical infrastructure such as smart grids (energy), implantable medical devices (healthcare), electronic warfare (military), industrial control systems (manufacturing) and intelligent transportation systems (transportation), they are expected to be free from vulnerabilities and immune to attacks to ensure their widespread deployment. However, this is not so as the complex cyber-physical interactions between the sensors, actuators and the computing nodes makes them susceptible to attacks. Unlike traditional information and communications technology services, the effects of delay and malfunctions of a CPS are very expensive both in terms of financial cost and the catastrophic effect they can have on the users. This paper particularly focuses on the security of mobile/vehicular cyber physical systems (mobile/vehicular CPS)[9].

A major reason for dominant research in Vehicular CPS, *also known as* intelligent transportation systems, stems from the need to reduce the rate of vehicular crashes, ensure a free flow of traffic as the number of cars continue to increase and also decrease the negative environmental impact of the transportation sector on the society[3]. Since most CPS are ubiquitous or remotely positioned, communication between the sensors and actuators is usually carried out via wireless communication links. According to Burg et al[4], some of the wireless communication standards used for CPSs and IoT include the Near-Field Communication (NFC), Zig Bee, Bluetooth, Cellular systems, Low-Power Wide Area Networks (LPWANs) traditional Wi-Fi (IEEE 802.11a/b/g/n), IEEE 802.11ad which is an extension of the set of Wi-Fi standards tailored towards millimeter-wave frequencies. Furthermore, IEEE 802.11p standard was developed to ensure the efficiency of vehicle-to-vehicle and vehicle-to-infrastructure communication. Almost all these technologies do not have security by design to provide resiliency in mobile/vehicular CPS.

Prior research works on the security of CPS have been done from the perspective of access restriction via physical protection of the sensors and actuators. However, as the attack surfaces have continued to increase due to the increasing

wireless connections and the autonomous nature of these CPS systems[5], the traditional approaches no longer work efficiently and effectively. Some of the security incidents for mobile CPS include eavesdropping, privilege escalation, man in the middle (MitM) attack and denial of service (DoS) with the Controller Area Network (CAN) bus and the engine control unit (ECU) being the attack surface. Furthermore, the speed sensor and the GPS receiver can also be targeted by ABS spoofing and GPS spoofing respectively[6].

Wolf et al. studied that false data injection (FDI) attacks are an emerging class of attacks that are primarily aimed at CPS[7]. These attacks are carried out by compromising the data fed into CPS. Moreover, since CPS systems typically operate by implementing control loops from where the data inputs by the sensors lead to response and actions by the actuators, false injection of data into a CPS will lead to wrong decisions further leading to wrong actions by the system. These types of attacks therefore are a major way of making the CPS insecure and unsafe especially where it is operated in critical infrastructure in healthcare, transportation and industrial CPS. In the past, the threat of false data injection has been studied with a larger focus on power systems or smart girds. Some solutions have been proposed to make smart grids resilient to FDI attacks. For example, Yong et al[8] leveraged on the principles of robust optimization to propose an estimation method under adversarial attacks for resiliency. The IEEE 14-bus electric power system was used to test the effectiveness of the approach. Rawat et al.[13] proposed cosine similarity matching with Kalman filter based estimation and compared with the chi-square detector for detecting false data injection attacks in smart grids.

The concept of adversarial networks became widespread when it was noticed that effective or competent adversaries are beginning to use certain strategies to evade detection systems which are designed to operate based on machine learning algorithms. Their activities which are generally classified as adversarial attacks are intended to attack the integrity, availability and privacy of the targeted system. In vehicular CPS, activities of an adversarial attack might include the flagging of a number of activities that are normal as an attack thus making the detection system unnecessary busy and thus affecting the availability of the system. However, the injection of false data into the system thereby causing it to make wrong detections is one of the methodologies of adversarial attacks deployed in mobile CPS.

Machine learning and recently deep learning algorithms have been successfully used in cybersecurity. Three of the most prominent areas where this occurs include intrusion detection, malware analysis and spam detection. The success achieved in these areas led to the application of machine learning algorithms for security engineering in cyber physical systems. The application of machine learning algorithms however to secure or provide resiliency for mobile CPS systems is a budding area of research. Arman Sargolzaei et al.[10] proposed a neural network-based false data injection detection technique on the cruise control of connected vehicles. Fuad et al.[11] also used the feedforward and backpropagation algorithms to develop an effective classifier which was trained on historical data that contains a combination of attacker and normal traffic data. In this paper, we aim to design, develop and evaluate distributed machine learning algorithm for adversarial resiliency where developed algorithm provides resiliency for security in adversarial environment for critical mobile CPS[9,12].

The remainder of this paper is organized as follows. Section 2 presents a system model. Section 3 presents the analysis for the proposed approach for achieving resiliency in the vehicular/mobile CPS. Section 4 presents the performance evaluation and numerical results. Finally, Section 5 presents the conclusions.

2. SYSTEM MODEL

A typical system model to study adversarial resiliency in mobile CPS is shown in Figure 1. Vehicles are members of a networked system with the capability of communicating with one other through a Vehicular Ad Hoc Network (VANET). Furthermore, each vehicle is equipped with cameras (with which it can visualize its neighbor), radar technology, infrared sensors and GPS receivers. All vehicles are expected to report their speeds, locations, directions etc. about 8 times every second to the others so that they can move in a cooperative manner[9]. Based on the information received from vehicles travelling in the front and data observed by sensors/cameras, each vehicle must adjust its speed and other actions such as lane changing action using the observed information on the fly. As discussed earlier, such a communication system and data collection methods are vulnerable to various cyber-attacks[9,12]. Thus, there is an urgent need to build in a resilient

mechanism that can make an individual vehicle withstand the activities of an adversarial vehicle which tries to report or inject false data while reporting their data to other neighboring vehicles.

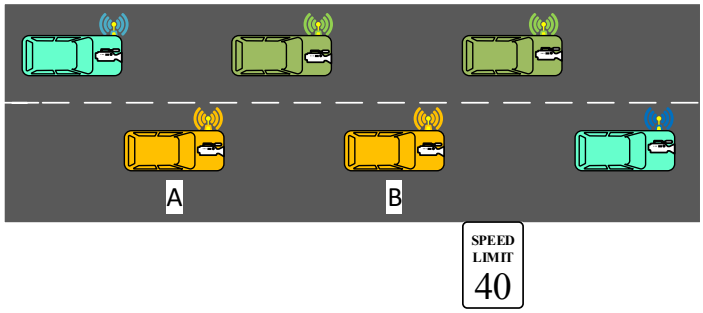

Figure 1: System model for mobile/vehicular CPS with posted speed limit where mobile CPS nodes (or vehicles are equipped with computing, communications, storage and controlling devices including cameras, sensors, radars, lidars, GPS devices, CAN, ECU).

In mobile CPS, beacon packets are broadcasted periodically over the network connecting vehicle-to-vehicle and vehicle-to-infrastructure. The packets broadcasted contain information such as the sender-id, time of transmission, the present position, direction, and the speed of a particular vehicle. This information is usually transmitted to other vehicles in the network many times in a second. For vehicles in mobile CPS, they are usually equipped with GPS receivers, cameras and sensors which help capture the speed and position of the vehicle at any point in time. Apart from being aware of its own information such as speed and location, a vehicle in the system using the camera, GPS location and speed sensor is able to measure or estimate the speed, location and distance of its neighbor vehicles.

From Figure 1, vehicle A can get its own speed information which can be obtained from the dash board of the vehicle. Vehicles own speed is denoted by a variable S. The speed of each vehicle along with the location info, direction etc. is broadcasted at certain intervals to other neighboring vehicles in the network. The speed of vehicle B when broadcasted to vehicle A, known as the reported speed is denoted with the variable α. Furthermore, by using its camera and/or radar technology, each vehicle, say vehicle A, is also able to estimate the speed of its immediate neighbor/vehicle ahead such as vehicle B for vehicle A. This estimated speed is known as the observed speed and is denoted with the variable β. Vehicle A can also calculated the speed and distance information based on location information received from periodic status messages. This is termed as the speed calculated based on coordinates and is denoted by G. In the case of an adversarial neighbor, there is expected to be a disparity between the reported speed and the observed speed. An adversarial neighbor will attempt to report a false speed or false location so that the honest neighbor can increase or decrease its own speed to keep in line and thus cause a crash. The resiliency mechanism in this system must therefore be able to estimate the trust of its neighbor and thus decide on the veracity of the reported speed. One guiding variable here is the speed limit of the road in which all the vehicles are travelling. This is termed with the variable P and can be obtained from the GPS module of the vehicle. Our goal in this set up is to find the approach that provides resiliency in mobile CPS against false data injection attacks.

All the five variables namely self-speed (S), speed limit of the road (P), speed estimated based on reported location coordinates (G), reported speed of neighbor (α) and observed speed of the neighbor (β) by using sensors and cameras are used by a vehicle to regulate its own speed thereby making it resilient to adversarial attacks caused by false data injection. This is true for all vehicles in the mobile CPS or platoon of networked vehicles.

3. ANALYSIS FOR BAYESIAN MODEL ENABLED MACHINE LEARNING

As an example scenario for analysis, we consider the combination of the five variables S, P, G, α and β where an individual vehicle determines whether to slow down, keep the same speed or increase its own speed. If the reported speed and observed speed are the same, the vehicle maintains a constant speed. However, in the case of a false reported speed, the vehicle puts different weight on the speed reported and the observed speed based on the trust values estimated over an observation period. From the given variables, the estimation of the trustworthiness of a given vehicle can be obtained from

the estimation of the suspicion level of the same vehicle at a given observation time. The estimated suspicion level of a vehicle at a given time can in turn be used to estimate the trustworthiness of the vehicle. Representing the type of vehicle as V_i; which could be either Adversarial (ε) or friendly (¥), and the suspicion measured over time t as M_t, the suspicion level of a given vehicle can be obtained using Bayes' theorem also known as Bayesian model enabled machine learning. The suspicion level of a vehicle *i* among *N* vehicles, denoted by τ_i at time *t* can be expressed as follows:

$$\tau_i(\mathrm{t}) = \frac{P\,(M_t|V_i = \varepsilon)\,P(V_i = \varepsilon)}{\sum_{m=1}^{N} P\,(M_t|V_m = \varepsilon)\,P(V_m = \varepsilon)} \tag{1}$$

Also, since there is a probability of any vehicle on the road being adversarial or malicious, for any vehicle *i* and *m* $P\,(M_t|V_i = \varepsilon)P(V_m = \varepsilon)$. With this assumption, eq. (1) can be expressed as

$$\tau_i(\mathrm{t}) = \frac{P\,(M_t|V_i = \varepsilon)}{\sum_{m=1}^{N} P\,(M_t|V_i = \varepsilon)} \tag{2}$$

After the suspicion level of another vehicle is estimated by the given vehicle, the trust of a given vehicle can be estimated by using a suspicion level as

$$\psi_i\,(t) = 1 - \tau_i(\mathrm{t}) \tag{3}$$

This trust level learned is used by the vehicle to weight the message received from the given vehicle. Furthermore, using an algorithm given as *Algorithm 1*, the data (speed, location info) sent by an adversarial vehicle can be verified against its estimated speed using the local camera and sensors as well as speed estimated using reported coordinates. Weight are placed on the different input factors. When the suspicion level of a given vehicle is quite higher and above a certain threshold, the weight placed on this reported speed is very minimal while a larger weight dependence is on the locally estimated speed information which is expected to be very trustworthy. This is expressed below:

$$\text{Adjusted speed of a vehicle i} = w_1.f(S, \beta, G) + w_2.f(\alpha) \tag{4}$$

where the $w_1 + w_2 = 1$ and w_1 and w_2 are directly proportionately to trust level ψ_i and inversely proportional to suspicion level $\tau_i(\mathrm{t})$ for each vehicle. Each vehicle will have different weights while estimating its speed based on observed speed, estimated speed, and speed based on received periodic information of the neighbor.

Algorithm 1 – Algorithm for Adversarial Resiliency in mobile CPS run by each vehicle

- **Input:** Beacon messages sent from neighboring vehicles containing the reported speed, location, the estimated speed of the neighboring vehicle by the given vehicle, and the legal speed limit.
- **Output:** Trust level which is used to decide whether to increase, reduce or maintain the speed of the vehicle.
- For each vehicle *i*, to verify reported speed at time *t* **do**
- Estimate suspicion level as $\tau_i(\mathrm{t}) \leftarrow \frac{P\,(M_t|V_i = \varepsilon)\,P(V_i = \varepsilon)}{\sum_{m=1}^{N} P\,(M_t|V_i = \varepsilon)\,P(V_i = \varepsilon)}$
- Estimate trust level as $\psi_i\,(t) \leftarrow 1 - \tau_i(t)$
- **if** $\psi_i\,(t) \geq \gamma_{(T)}$ **then**
  - Weights w_1 and w_2 in (4) can be equal. Speed for the given time should be ok and will be based on self-estimated speed, estimated speed using reported coordinates and the speed reported by neighbor are highly correlated hence there is no need for corrections or adjustment
- **else if** $\psi_i\,(t) < \gamma_{(T)}$
  - Trust level of neighbor is too low hence the weight w_2 should be minimal compared to w_1 by maintaining $w_1 + w_2 = 1$
  - Adjusted speed ← $w_1.f(S, \beta, G) + w_2.f(\alpha)$
- **end if**

4. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed algorithm developed to enhance the resiliency in mobile CPS in the presence of adversaries who inject false data in the learning process. We used the variables such as self-speed (S), speed estimated based on reported location coordinates (G), reported speed information of neighbor vehicle (α) and observed speed of the neighbor (β) by using sensors and cameras. To simulate the different scenarios, we generated a random speed centered around 40 miles per hour by considering a Gaussian distribution[9] using MATLAB. We used Monte Carlo simulations.

Scenario 1: In this scenario, at every iteration, the values for the three variables related to speed given as S, P, α and β all fall within the same range because all the conditions are perfect. This implies that there is no adversarial attack or false data injection affecting the speed of the vehicle as obtained from the dashboard. Also, the speed reported to it by a neighbor has a close correlation with that reported by the speed obtained from the coordinates because the camera and sensor used for estimating the speed by the neighbor is working perfectly fine. As usual, a small difference is expected between the different variable despite the normal conditions because of high mobility and GPS accuracy issues[9]. This could either be one/two miles above or below the average speed to account for estimation errors due to the imperfection of the camera and GPS receiver with which we obtain the vehicle speed from the reported location coordinates. We plotted the variation of the speeds using different approaches vs the observation time, as shown in Figure 2. Three values are within the tolerance limit thus the given vehicle will maintain its current speed.

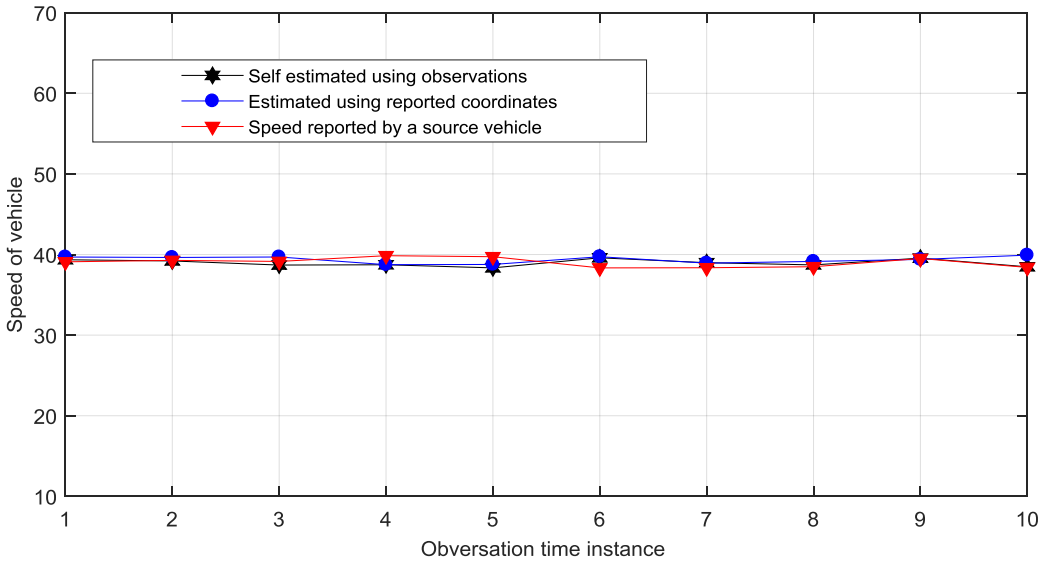

Figure 2: variation of speed vs the observation time where no adversarial attack is present in mobile CPS.

Scenario 2: This scenario differs from the *Scenario 1* in that the reported speed from the neighbor/vehicle ahead of a given vehicle is wrong and falsely reported to be higher than the self-speed with a hope that the following vehicle would speed up and crash into the front vehicle since we consider that there is adversarial vehicle who wants to mislead the machine learning model. The speed reported does not match the speed estimated by the given vehicles using cameras and sensors as well as speed estimated based on reported location coordinates. Normally, the vehicle behind is expected to increase its own speed to match up with the speed of the vehicle ahead so that it can keep up the distance. However, since the vehicle speed is actually slower than what is reported, an attempt to increase its speed will lead to a crash. Figure 3 shows the speed variation over observation time where reported speed of a front vehicle is way higher than its actual speed (speed estimated by using cameras, sensors and geolocation information by the following vehicles). The case represents an adversarial scenario where false data is injected into the mobile CPS to mislead the machine learning model. The CPS must therefore build resiliency by being able to remain safe and secure despite the injection of false data. Using the

algorithm presented as *Algorithm 1*, trust level is estimated to find the best weight and get the adjusted speed for a given vehicle. Because the trust level of the adversarial vehicle is low and does not correspond with what is obtained from the self-speed and GPS coordinates, weights are assigned to both the trusted and trusted values. As already highlighted, a higher weight is attached to the trusted data while a much lower weight is attached to the untrusted data.

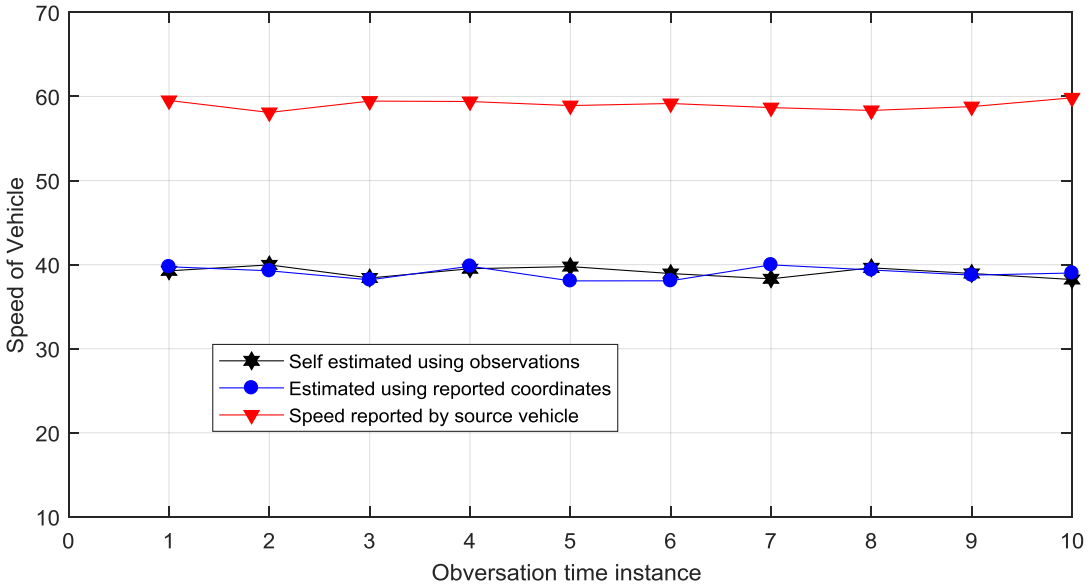

Figure 3: variation of speed vs the observation time where an adversarial attack is present in mobile CPS.

Figure 4 shows the result of applying the resilient algorithm while adjusting the speed of the given vehicle. We can observe that despite the injection of false data (as shown in Figure 3), the speed reported by neighbor (source vehicle) has been corrected and now lies within the tolerable average speed of 40 miles per hour in this case. With this, the vehicle decides to keep its own speed and does not increase to be in line with the false reported speed given by the adversarial vehicle.

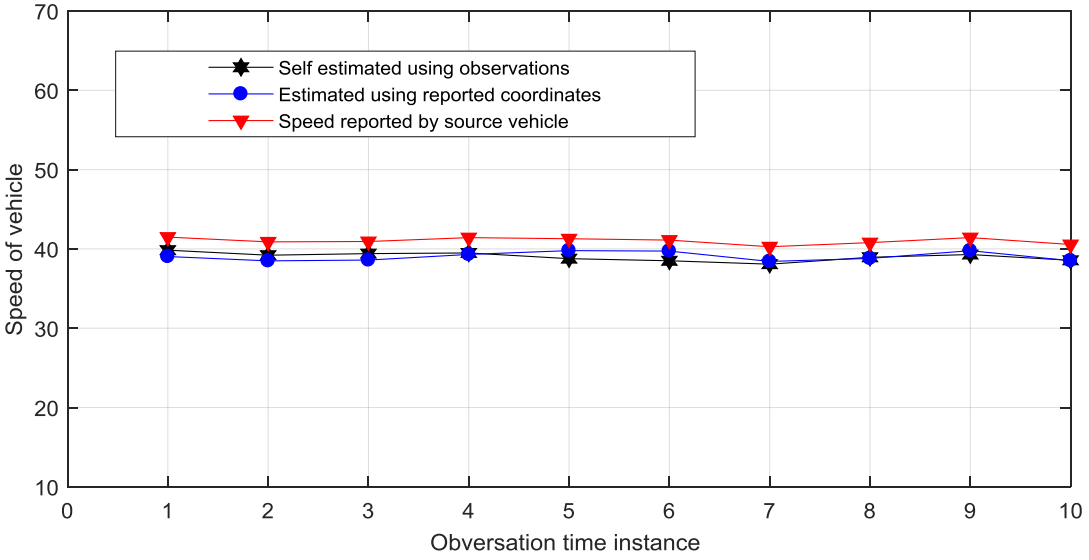

Figure 4: Corrected estimated speed by a given vehicle by using Algorithm 1 when an adversarial attack is present in the mobile CPS.

5. CONCLUSION

Mobile/vehicular CPS just like most other CPS applications transmit information wirelessly to other CPS nodes and individual CPS nodes adjust their operating parameters on the fly thereby increasing their attack surfaces. Due to the critical nature of these systems and the adverse effect of failures, there is a need to enhance security and safety by making the CPS resilient to the injection of false data by adversaries. In this paper, we have developed an algorithm that can be used to enhance the resiliency of mobile CPS. The simulations results have corroborated our claims through security engineering with machine learning for adversarial resiliency in mobile CPS.

ACKNOWLEDGMENT

This work is partly supported by the U.S. National Science Foundation (NSF) under grants CNS 1650831 and HRD 1828811. However, any opinion, finding, and conclusions or recommendations expressed in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the NSF.

REFERENCES

[1]. German National Academy of Science and Engineering (ACATECH), “Cyber-physical systems: Driving force for innovation in mobility, health, energy and production,” Tech. Rep., Dec. 2011.

[2]. P. Leitao, S. Karnouskos, L. Ribeiro, J. Lee, T. Strasser and A. W. Colombo, "Smart Agents in Industrial Cyber–Physical Systems," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1086-1101, May 2016.

[3]. K. Evers, R. Oram, S. El-Tawab, M. H. Heydari and B. B. Park, "Security measurement on a cloud-based cyber-physical system used for Intelligent Transportation," *2017 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, Vienna, 2017, pp. 97-102.

[4]. A. Burg, A. Chattopadhyay and K. Lam, "Wireless Communication and Security Issues for Cyber–Physical Systems and the Internet-of-Things," in *Proceedings of the IEEE*, vol. 106, no. 1, pp. 38-60, Jan.2018.

[5]. S. Checkoway, D. McCoy, D. Anderson, B. Kantor, S. Savage, K. Koscher, A. Czeskis, F. Roesner, and T. Kohno, “Comprehensive Experimental Analysis of Automototive Attack Surfaces,” Proceedings of the USENIX Security Symposium, San Francisco, CA, August 2011.

[6]. A. Humayed, J. Lin, F. Li and B. Luo, "Cyber-Physical Systems Security—A Survey," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1802-1831, Dec. 2017.

[7]. M. Wolf and D. Serpanos, "Safety and Security in Cyber-Physical Systems and Internet-of-Things Systems," *Proceedings of the IEEE*, vol. 106, no. 1, pp. 9-20, Jan. 2018.

[8]. S.Z. Yong, M.Q. Foo, E. Frazzoli, "Robust and resilient estimation for cyber-physical systems under adversarial attacks", *American Control Conference*, pp. 308-315, July 2016.

[9]. D. B. Rawat and C. Bajracharya, “Vehicular cyber physical systems: Adaptive connectivity and security,” Springer, 2016.

[10]. A. Sargolzaei, C. D. Crane, A. Abbaspour and S. Noei, "A Machine Learning Approach for Fault Detection in Vehicular Cyber-Physical Systems," *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Anaheim, CA, 2016, pp. 636-640.

[11]. F. A. Ghaleb, A. Zainal, M. A. Rassam and F. Mohammed, "An effective misbehavior detection model using artificial neural network for vehicular ad hoc network applications," *2017 IEEE Conference on Application, Information and Network Security (AINS)*, Miri, 2017, pp. 13-18.

[12]. D. B. Rawat and K. Z. Ghafoor, "Smart Cities Cybersecurity and Privacy", Elsevier Press, ISBN: 9780128150320, November 2018.

[13]. D. B. Rawat and C. Bajracharya, "Detection of False Data Injection Attacks in Smart Grid Communication Systems," in *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1652-1656, Oct. 2015.