Cyber Security And Artificial Intelligence For Cloud-Based Internet of Transportation Systems

S. Lakshmi Sravani[1], Dr.K.Sekhar[2], Dr. Jasmine Sabena[3]
[1, 2, 3]SV ENGINEERING COLLEGE

Abstract- *The Internet of Things (IoT) has major implications in the transportation industry. Autonomous Vehicles (AVs) aim at improving day-to-day activities such as delivering packages, improving traffic, and the transportations of goods. AVs are not limited to ground vehicles but also include aerial and sea vehicles with a wide range of applications. The IoT systems consisting of a collection of AVs have come to be known as the Internet of Transportation systems. While such IoT systems manage large quantities of sensor data, much of the data is also sent to a cloud for offline analysis. While there is great potential in AVs and the improvements it can make to the transportation industry, security and privacy concerns pose new challenges that need to be addressed as we move forward. In addition, Artificial Intelligence techniques are also becoming crucial for such IoT systems to be able to intelligently manage the AVs. This paper discusses AI and security for cloud-based Internet of Transportation Systems.*

Keywords- Cyber Security, Artificial Intelligence, AI, Cloud Internet of Transportation

I. INTRODUCTION

In recent years there has been an explosion of AVs. Companies are investing heavily in AVs. AVs evaluate their environment using a variety of sensors (e.g., camera, GPS, Inertial Measurement Unit [IMU], LiDAR, RADAR and ultrasonic sensors). While there is great potential in AVs and the improvements it can do to the transportation industry, security and privacy concerns pose new challenges that need to be addressed. The sensors are susceptible to malicious tampering (e.g., IMUs are susceptible to sound waves and GPS receptors are susceptible to spoofing signals). Vehicles should verify the veracity of sensor signals before acting upon them [1].

The IoT systems consisting of a collection of AVs have come to be known as the Internet of Transportation Systems. The Internet of Transportation Systems are subject to attacks (like any cyber physical system). Streaming data is being collected from such systems including autonomous and in the future driverless vehicles. As transportation systems go electric, they need energy conservation. Threats to the security of such systems could cause massive damage including accidents, loss of lives as well as being stranded on lonely highways due to attacks on energy management.

Data Science/ML techniques are being applied to analyze the data of AVs and a challenge is to apply the stream analytics/learning techniques for transportation data. For example, how can the ML techniques be applied to the massive amounts of sensor data emanating from the AVs? [2]. The Internet of Transportation Systems will also depend heavily on Data Science/AI/ML (Machine Learning) techniques for various applications including optimum directions, driving without a human in the loop and many more. The Adversary will learn the machine learning models that we use and try and thwart our models [3]. Finally, while massive amounts of data are collected by the Internet of Transportation Systems, the privacy of the individuals has to be protected. We envision that much of the data sharing and analytics will be carried out using the services running in the cloud integrated with the Internet of Transportation System [4].

This paper explores how Artificial Intelligence, Security and the Cloud can be integrated to develop Intelligent Internet of Transportation Systems. We first discuss the integration of cyber security and AI in Section II. Next, we discuss how a secure cloud may be utilized to carried out data analytics for the Internet of Transportation Systems in Section III. Section IV discusses security and privacy for the Internet of Transportation Systems. Section V discusses how the various components (e.g., AI, Security for Cloud) can be integrated to provide Intelligent and Secure Internet of Transportation Systems. Future directions are discussed in Section VI.

II. INTEGRATION OF CYBER SECURITY AND AI

There are three aspects to integrating cyber security and AI. One is to apply AI for cyber security, the second is to apply cyber security for AI and the third is to detect privacy attacks due to AI. Research began on applying AI for cyber security around the mid-1990s. The idea is to apply ML techniques for detecting unauthorized intrusions. This research was expanded in the 2000s to include malware analysis and

insider threat detection [5]. Massive amounts of attack data are being collected. This data has to be analyzed so that malicious attacks can be detected. Furthermore, we also need to predict how the malware could mutate so that the attacks can be prevented [6]. In addition, streaming data are being analyzed to detect malicious insiders.

The second area is securing the AI techniques. This area, now come to be known as adversarial machine learning, has become quite prominent over the past decade. We are increasingly depending on ML techniques for every aspect of our lives from healthcare to AVs. These ML techniques could be attacked and could result in catastrophic situations. Therefore, we need to examine the types of attacks and adapt the ML techniques. For example, in our work, we have examined support vector machines (SVM) and adapted the SVM techniques to detect some of the attacks. The adversary will learn about our models and adapt its behavior. Our adversarial support vector machine technique is able to learn what the adversary is doing and adapt itself so that it can detect the attacks. Over time it becomes game playing between the adversary and us.

The thirst aspect is the privacy violations that could occur to do the ML techniques. For example, it is now possible to integrate massive amounts of data and analyze the data and obtain various properties of individuals. This could result in the privacy of the individuals being compromised. Many privacy- aware machine learning (data mining) techniques have been developed [7]. The challenge is to enforce appropriate policies so that we can carry out policy aware data collection, storage, integration, analysis and sharing [8].

III. SECURE CLOUD-BASED IOT

As stated earlier, we envision that much of the data collected from the AVs will be sent to a cloud for further processing including carrying out analytics. That is, the massive amounts of data including attack data may be analyzed in the cloud using various ML techniques. Therefore, it is important that the cloud itself be secure especially if it has to carry out security critical operations.

We have designed and developed a layered architecture for a secure cloud [9]. At the lowest layer is the VNM (Virtual Network Monitor). Then we have the VMM layer (Virtual Machine Monitor) that carries out virtual machine introspection. Above that is the cloud storage layer based on technologies such as Hadoop/MapReduce. The data may be encrypted which means querying and analytics will have to be carried out on the encrypted data. Above this layer is the query layer for querying the cloud data. Finally, we have the application layer and in our example the applications are those that support the Internet of Transportation Systems.

IV. SECURITY AND PRIVACY OF THE INTERNET OF TRANSPORTATION SYSTEMS

One of the approaches to the security and privacy of the Internet of Transportation Systems is to build a reference monitor using a Physics-Based Anomaly Detection (PBAD) algorithm for ground and aerial AVs [1]. The algorithm will consist of three parts: (i) Building a model offline of the AV's physical invariants, (ii) Implementing an online tool to monitor expected and observed behavior to detect anomalies, and (iii) Raising an alarm if significant residual difference exists between executions. The techniques have been applied both for ground and Ariella AVs. Below we provide more details of the steps.

(i) Offline pre-processing: The AV's invariants are calculated using a well-known non-linear model for aerial and ground vehicles. Accelerometer, gyroscope and magnetometer sensor data on the x, y, and z axis is used for the aerial vehicle. Vehicle position and steering angle is used for the ground vehicle. (ii) Online stage: An Extended Kalman Filter (EKF) is used to predict AV's physical behavior by estimating unknown parameters from noisy sensor input. The algorithm is divided into two sections that predicts and corrects the estimation before it is compared against the sensor data. (iii) Anomaly detection: A CUSUM algorithm is then used to detect persistent attacks. An alarm is raised if the residual difference is larger than a predefined threshold.

Beyond the security of individual vehicles, the transportation sector could greatly benefit from a supporting infrastructure that allows communication between vehicles, motion sensors on lamp posts, and surveillance cameras (to name a few) to help identify traffic jams, re-route vehicles and increase vehicle safety. From the user's perspective, privacy concerns arise from all the information needed by such system that could lead to private information being exposed such as vehicle identification and driving patterns. Legislators, engineers and scientists should keep privacy concerns in mind as advances in IoT become more prominent in day-to-day activities. This will aid in improving the public perception, reduce hesitation from consumers and increase the adoption rate of new technologies [4].

V. INTEGRATING AI AND SECURITY FOR CLOUD-BASED INTERNET OF TRANSPORTATION SYSTEMS

Data Science/ML techniques are being applied to analyze the data and a challenge is to apply the stream analytics/learning techniques for transportation data. The main question is to understand the nature of the complex transportation data and adapt the stream analytics techniques and apply them on the massive amounts of heterogeneous sensor data being collected. Such data will often emanate as data streams. Therefore, many of the techniques for stream-based machine learning need to be examined [10]. In addition, deep-learning based techniques developed for IoT systems need to be examined [11].

The Internet of Transportation Systems will depend heavily on Data Science/AI/ML techniques for various applications including optimum directions, driving without a human in the loop and many more. The adversary will be learning the models used by the vehicles as well as learn about the data used in the training of the models. The adversary will attempt to thwart the vehicle's learning process. Therefore, the learning algorithms have to adapt to thwart the adversary's actions. Eventually it becomes game playing between the adversary and the vehicle's machine learning algorithms [3].

While massive amounts of data are collected by the Internet of Transportation systems, the privacy of the individuals have to be protected. As more and more sensor data are collected, the storage on the AVs will not be sufficient to store all of the data. We envision an encrypted cloud storage component where older data and/or less frequently accessed data are pushed to the cloud. Based on the access control policies, local applications running on the AVs will be given access to some of the collected data. When needed, these AVs will be allowed to access some of the encrypted data stored in the cloud via a simple query interface. We envision that much of the data sharing and analytics will be carried out using the services running in the cloud [8].

Another direction for enhancing security and at the same time ensure high performance computing is trustworthy analytics [12]. Computations over big data may require massive computational resources and, organizations (e.g., automobile companies) may use a third-party service to outsource some computations to be cost-effective. When a third-party server is used for computation, data inherently becomes available in untrusted environments, i.e., either observed by a man-in-the- middle during data transmission, or insider threat from adversaries at the third-party location where computation is performed. In these cases, data owners may need to protect their data and require cryptographic guaranties about data security and integrity of computational output from these third-party services. We are conducting research in Secure Encrypted Stream Data Processing and Trustworthy Analytics using advancements in embedded hardware technology (e.g., Intel SGX) to support trusted execution environment (TEE). We need to explore the applications of TEEs to Internet of Transportation and Infrastructures.

VI. SUMMARY AND DIRECTIONS

This paper has discussed the characteristics of the Internet of Transportation Systems with respect to AVs as well as the security and privacy concerns of such systems. Next, we discuss how AI and Security may be integrated. Cloud-based Internet of Transportation Systems were also discussed. Finally, we discussed how AI, Security and the Cloud may be integrated with the Internet of Transportation Systems.

We have only scratched the surface with respect to securing the Internet of Transportation Systems. We have to understand the various types of tracks and develop ML techniques to detect and prevent the attacks. We also have to examine how to handle the attacks on the ML techniques that are needed for the development of Intelligent Internet of Transportation Systems. Finally, we need to determine the types of data to send to the secure cloud for carrying out analytics.