Deep Reinforcement Learning for Adaptive Cyber Defense in IoT-connected Autonomous Vehicle Networks

By Dr. Olga Volkova

Professor of Artificial Intelligence, National Research University – Information Technologies, Mechanics and Optics (ITMO)

1. Introduction

In particular, the novel cyber defence mechanism is based on a robust, Reinforcement Learning-based Intrusion Detection & Response (RLIDR) module, which receives encrypted Low-Level Telemetry & Control (LLT&C) data streams and has an interface with a Human-Machine Interface (HMI). The architecture is based on Long Short-Term Memory (LSTM) recurrent neural network and convolutional layers, useful to understand waveforms and other hidden and non-linear information. LSTM is useful in the detection mechanisms because it can process the incoming LLT&C data stream, monitoring latency, signal bitrate and deep messages sharing with respect to the classical sequence-to-sequence Terminator-Transformer model. Attack resistance properties are developed over time by having an adaptation phase to the network characteristics, and this means that a proper Reinforcement Learning (RL) algorithm was conceived and selected to adapt dynamic energy driven threat against the system, considering the different behavior of the threat found in time with respect to the vehicle dynamics. Finally, a simulation in a controlled IoT scenario is shown, based on real telecommunication data from the metropolitan area of the city of Milano, to give evidence that the proposed RLIDR is quite robust.

[1] [2] The latest trends in Intelligent Transportation Systems (ITS) have shown that to improve safety, efficiency, travel time, and reduce traffic congestion, the widespread utilization of autonomous vehicles is key. In this context, behind these applications there are advanced communication technologies able to integrate autonomous vehicles in an Internet of Things (IoT) scenario in order to provide an uninterrupted connectivity to the passengers and to the infrastructures through intelligent algorithms, which coordinates the vehicles movements and minimizes possible congestions. However, all these advantages can also

bring new threats: the systems that control the vehicles can be attacked with released energy that exceeds the kinetic energy that the vehicle is going to have, putting thus in danger the health or lives of the passengers. Last, the vehicles infotainment systems which are connected to the vehicles operational systems can be further used as a attackers premises into internal pathways to the pilot-controlled systems. A cyberattack resistance system for network based autonomous vehicles is then proposed in the present work, based on the Q-Learning evaluated at each time the dynamic nature of the energy released, and then the dynamic behaviour of the cyber threat injected in the networks is changing in time. The system, which upper layers are typified by deep learning models have been tested in a real scenario by involving telecommunications experimental data related to the metropolitan area of the city of Milano.

1.1. Motivation

The increased dynamism and an overdose of data in an unsupervised manner regarding networked traffic may readily be exploited by any adversaries. This makes the current detection algorithms vulnerable to security threats due to the inevitable negligence of certain learned features [2]. It has been a challenge to tune the detection algorithms for the consideration of hidden structures and extended temporal behaviors of the evil process. Reinforcement learning models are being widely utilized due to the attributes of adapting the hidden temporal behavior, carefully crafted detection rewarding features, and thinkable decision making forever re-learning and improving against the evolving threat paradise. Especially, the DRL models are chosen due to the attributes of interoperability of most of the intelligent platforms, computing resources, and instructions up to possible scalability and internet sandbox architecture, ready for quantized data involved adaptive activity-based learning and decision making. The enhanced models deployed as solutions in literature result of a host of successfully evaluated networked environment threats, ranging from the constant evolution of spam to the deliberate undetectable damaging adversarial crafted hardware exploited for hacking purpose [3].

I. INTRODUCTION Deep reinforcement learning (DRL), a variant of machine learning utilizing neural networks capable of empowering smart devices with the capability of learning and self-improvement provided they possess large amounts of data for training purposes, encapsulates this ongoing trend. In the emerging era of Internet-of-Things (IoT), especially in

vehicular scenario involving autonomous vehicles, DRL offers autonomous cyber security adaptability that can mitigate severe unfettered cyber threats. DRL, with the aid of adaptive cyber defense, helps in enabling automatic learning and decision making processes, which supplements the system especially in a modifying environment of self-scripted coordinated attacks, exhibiting a legitimate behavioral concept joining human and machine solutions for cyber defense and driving safety [4].

1.2. Research Objectives

In this thesis, the main objectives are proposed to be experimental studies in order to enhance the security and performance of Autonomous Ground Vehicles (AGV) against attacks. Furthermore, we aim to propose security models related to the Internet of Vehicles (IoV) to increase the security of the autonomous transportation systems. The research objectives leading to an adaptable cyber security and performance enhance model can be broken down to the following tasks: (Task 1)AA study of secure autonomous vehicle communication strategies in a multi-agent environment—exploring the modular end-to-end multitask AI which includes: person 'detection' and 'identification' using CNN with LSTM. The proposed promotion-to defense strategy will be developed for adversarial detection with system performance evaluation using C-V2X with real vehicle data. (Task 2) Development of the control framework and adaptation for IoT-based autonomous system with secure and separate control and stability dynamic model for cyber security applications supporting Autonomous Car and associated IoT sensors, security metrics and cryptosystem.

[5], [6], [7]. As discussed in the above section, hacking and complex cyber-attacks can serve to obstruct the use of fully autonomous vehicles, including Drones, Cars, and Trucks. With an eye on enhancing security and performance of Autonomous Vehicles (AV) and developing advanced security solutions for the Internet of Autonomous Vehicles, this thesis aims to study a defense system for critical data in an congested Internet of Vehicles (IoV) system and model the person detection and identification process using convolutional neural networks (CNNs) and Long Short-Term Memory (LSTM). This would involve a framework for protecting the Vehicle-to-Everything (V2X) communications of AVs using reinforcement learning (RL) algorithm. Through interactions with a simulator, the malwares assessment and risk management will enforce specific cyber security policies for vehicle operations including

enabling/disabling sensors and processing input data, using a Hardware-in-the-Loop (HIL) real ADAS setup for verification.

2. Background

The main input factors for these ML-based systems would be the observable telemetry data such as speed, longitude, latitude, steering wheel data, engine motor, and others from various vehicular sensors. Conventional ML-based adaptive cybersecurity mechanisms such as Random Forest (RF), Neural Networks (NN), and Support Vector Machines (SVM) have been utilized to segregate between attackable and non-attackable data [8]. However, the adaptivity and accuracy of these systems could be enhanced by adopting deep end-learning mechanisms.

Most traditional cybersecurity mechanisms currently used in IoT-based vehicular networks could be easily handled, evaded, or attacked by the malicious entities [9]. Recent adapted research studies have emphasized the use of advanced artificial intelligence (AI) techniques to implement dynamic, smart, complex, and preventive cybersecurity defenses, such as machine learning (ML), Q-learning, reinforcement learning (RL) and deep learning (DL) [1]. With the recent advances made in these smart technologies, IoT-networked cars will have a high potential in terms of behaving securely not only against common and well-known cyber-physical attacks, but also against unknown and emerging threats by continuously assessing vulnerabilities.

2.1. Autonomous Vehicle Networks

Although we can have a learned two-level information about the environment with large quantities of data by allowing the network to influence the state of the environment, we are (temporarily) left with all the traditional RL problems of how to deal with the possibly high-variance, potentially infinite generalized state space, and the issue of state acquisition. Therefore, a natural question to ask is: in what circumstances, and at what cost, can we use the fact that the Bellman equation involves the network to convert these traditional RL problems to a supervised learning problem [10]. In this paper we give a formal treatment to this question and establish results of a rather general nature. A side remark is that it is not necessary to assume that vt belongs to the remarkable set, the only requirement for this result to hold is that the value function can be well approximated by the neural network. We thank Kamyar Tavakoli for a helpful comment.

The Internet of Things (IoT) introduces new security and privacy chall- enges in connected vehicle networks. IoT is increasingly connecting millions of computers, smart phones and smart end- devices to the Internet by offering various applications. Previously, Internet of Things applications had been particularly aimed at smart home, smart city, smart grid, health and smart transportation. Thus, IoT can be a viable model to determine an eventful future of the Internet. One of the core manifestations of IoT is called the Internet of Vehicles (IoV) [11]. The communication connections in IoT can be roughly divided into in-vehicle wireless communication (I2V),vehicle-to-vehicle (V2V), and vehicle-to-infrastructure (V2I).Thesubscriptionsilktimesgrows in multiple related fields, thereby providing a betting and evolving vehicle environment to improve our driving safety and the transportation ef- ficiency.

2.2. Internet of Things (IoT) in Autonomous Vehicles

None

2.3. Cybersecurity Challenges in IoT-connected Autonomous Vehicle Networks

To all estimated and consequent above-reported reasons, it is urgent an answer to a clear and actual automotive malware danger, seen in its updating tie to the next present. In particular, it is vital a successfully storage of the "purity" of the command issued by the Driver (bad thinking on the vehicle in trip, and over all of the clever techno evolution toward the car-to-car and to other physical-cybernetic surrounding, like big data concessions, potently repeating false and even dangerous processes and complications. In line of last observations, a re-establishment of the carrier of defense strategies formation is also, of course, suitable for other devices of the automotive train, such as structural parameters (like direction, resistances, etc.), impostations for systems of comfort (for internal climate, seat control), strategies for economy on energy, etc. [11].

General automakers, by translating all the commercial and global success obtained with the early threats of the first generations of automobiles, want to lose malwares but not the primary role in the automotive supply-chain. To do that, they can no longer use traditional automotive cybersecurity solutions, such as the static penetration testing approach, and have to explore new strategies. This directive in the business flow of the automotive supply-chain then opens a new and continually evolving field of research. The simplified vision of an endgame consists

of a fast and widespread competition to digitize the automotive control systems, framed inside a worldwide governance aiming at arm-wrestling the malware threats. Unlike Malware writers that already had united in being highly intelligent and totally unpredictable in incoming linings-up, the Malware challenge offers a clear side to grab in the joint advancement in the direction of Quality Assurance (QA) and of Quality Cipher with exclusive markings. Elicos is a team of established tests of the in-vehicle cybersecurity shield, that, mixed with a deepen oversight of main software delivery practices, highlights the key criteria for a malware-proof and First Class Automotive OEM Supply-Chain. The view of these investigations is some challenging article.

A Brisk Line of Research to deal with the Malware Attacks towards Automatic Vehicles Self-driving vehicles and smart cars have advanced, leading to the development of vehicular networks to manage traffic, parking, and accident avoidance [7]. In-vehicle networks (IVNs) have been designed to facilitate communication between various vehicle components and they employ protocols such as the controller area network (CAN), FlexRay, and Ethernet. However, the current CAN message frame, originally developed for simple electronic control units (ECUs), lacks authentication mechanisms, thus making any data inside an IVN susceptible to external attacks. As vehicle connectivity increases, purpose-made electronic components for different high-level and low-level functionalities have been devised, thereby increasing the overall complexity of the architecture. Consequently, there is a growing number of different electronic components that can be targets of any kind of safety, security and privacy (SSP) attacks carried out over an IVN, thus leading to the parallel increase in complexity of those eco-systems also. Overall, the evolving vehicle connectivity plays a fundamental role in the growth of the potential threats against targeted systems. Therefore, researchers have to face this increasing number of threats with a good aptitude both in spotting and neutralizing these cybersecurity evils for protecting the innocent drivers' happiness.

3. Reinforcement Learning Fundamentals

The cyber-physical nature of the autonomous vehicle network depicts a distributed and interacting environment mainly emphasizing the interplay of various actors such as Intrusions and defenders corresponding to an ACD framework. The given scenario is formulated as a Markov decision process (MDP) and the optimal cyber defense policy is

learned using multi-agent deep Q-network (MADQN). Specifically, a goodness measure is introduced that captures the key insight of a malware attempting to persist in the network. The considered goodness measure is used as a basis to build a dickym MALWARE reputation model [12].

Transportation systems have recently taken advantage of autonomous vehicles (AVs) to enhance the safety and efficiency of transportation. At the same time, interconnected AVs can create a connectivity environment vulnerable to cyber-attacks [13]. Hence, security in AVs is crucial as critical systems can be put in jeopardy, potentially causing life-threatening scenarios. For this purpose, Adaptive Cyber Defense (ACD) is important to build an intrusion-resilient autonomous vehicle network. This paper aims to present a study of a multi-agent deep reinforcement learning (DRL) framework for building Adaptive Cyber Defense and detect Advanced Persistent Threats (APTs) in a connected and automated vehicle (CAV) network [14].

3.1. Basic Concepts

The secure operation of an intelligent autonomous vehicle and an efficient, adaptive, and practical solution can be found in the network including IoT connections by combining deep learning and reinforcement learning. The wireless multi-hop network management tool can help the self-organizing network to configure the reliable, low-cost, and low-delay communication routes needed for the connection by tracking traffic patterns and dealing with possible threats efficiently. The communication and monitoring systems of the vehicle network and the important system parameters that can be controlled are presented. Adaptive network defense tool is applied to wireless multi-hop communication networks including an intelligent vehicle and an IoT connection in the paper. The autonomous and decisive response and network hardening actions that need to be taken against malicious attacks or threats can be realized by applying deep reinforcement learning methods [15].

Deep Reinforcement Learning (DRL) is a powerful algorithm that combines reinforcement learning (RL) and deep learning, and has recently been applied to cyber security [16]. Deep Q networks (DQN) using DRL can process large-scale traffic logs to estimate the impact of MTD strategy on attackers and select suitable moving tools to increase the cost of attacks. Network hardening can help organizations to reduce the possibility of attacks by consuming more time or attacking resources. To achieve this, the attack surface is monitored by DRL, the effective

MTD strategy is developed against different attacks, but both of them do not explore the optimal interaction with potential threats [17].

3.2. Markov Decision Processes (MDPs)

A brief overview of the RL formalism is given here and readers are referred to a comprehensive overview [5]. An RL agent can interact with the environment in a sequence of discrete time steps t = 0, 1, . . . and the environment state s(t) evolves stochastically following a Markov chain (also called a MDP). The agent observes the environment state st and picks an action a(t) from some set of feasible actions A the environment state st which deterministically induces a reward rt = r(st, a(t)). The policy $\pi : S \times A \rightarrow [0, 1]$ is a conditional probability density, $\pi t(\cdot \mid st) = P(a(t) = \mid st)$, over the action space A. The policy is updated by the agent through learning from its interactions with the environment, following which the agent selects actions stochastically. The objective of the agent is to maximize the long-term expected return $\gamma > 0$ over a finite or infinite horizon, known as the discounted cumulative reward, $Rt = \infty\sum k=0 \gamma krt+k+1$. The factor γ weights immediate rewards r(t) heavily by repeating the action, the agent perceives the cumulative reward diminishing into the future.

Internet-connected systems with their dynamic nature and massive scale require responsive, adaptive, and scalable protection [18]. Both attackers and defenders are turning to AI techniques, particularly machine learning (ML), to gain an edge. Reinforcement Learning (RL), a branch of ML, is suitable for taking sequential actions optimally in the face of unknown and changing environments. RL models the problem as a Markov Decision Process (MDP) and learns to interact with the environment to maximize a long-term reward. DRL, which integrates deep learning, is particularly effective in securing large-scale Internet-connected systems like IoT networks, web services, etc.

3.3. Q-Learning

A basic Q-learning network is shown in Fig. 4 [13]. The network is characterized by an input layer, a hidden layer, and an output layer. The input layer consists of neurons, where each neuron corresponds to a feature. The hidden layer consists of neurons known as hidden units and it is responsible for transmitting information from the input layer to the output layer. The output layer consists of neurons that correspond to actions available. An action a is chosen by defining a target vector t according to Eq. (2) and then adjusting W and B using Eqs. (3) and

(4). After the training is executed properly, Q will be approximately equal to Q*. If the environment is stable and the network has converged properly, then the network is said to have learned a good policy.

DRQN is a deep Q network combined with a recurrent neural network (RNN) to facilitate temporal dependency in sequential tasks [19]. In the DRQN architecture, a Long Short-term Memory cell is used as the network's recurrent unit. The addition of DRQN to the DQN architecture allows it to memorize past events and hence facilitates a more informed action selection mechanism while dealing with sequential problems. In the Atari game-playing task, a DRQN has been shown to outperform a DQN.

4. Deep Reinforcement Learning

We propose a novel intrusion detection system (IDS) model adapted to IoT network environments, which we call DEvoid-Intrusion Detection System (DEvoid-IDS). This is to address limitations of other state-of-the-art intrusion detection systems for IoT network environments, such as heavy computation, very large memory requirements, and security risks such as attacks like adversarial attacks in the IoT network. Instead, we use state-of-the-art and very efficient deep reinforcement learning algorithm, which aspires to learn from the environment without human intervention and learn which policy to choose to accomplish a certain goal. REINFORCEMENT LEARNING is used to make systems autonomous and self-updating compared to traditional methods. The deep reinforcement learning finds a maximally optimal sequence of actions for a sequence of state within an environment in an IoT network where the attack records arrive in a specific time. The simulated results obtained shows high accuracy when the model is recognized with other models that use traditional recurrent neural networks and LSTM models. So, the DEvoid-IDS requires less storage, less computation time, and gives high accuracy for the intrusion detection especially in time-variant environment such as IoT networks.

Many IDS analyses propose the application of machine learning to IoT networks [15]. The well-known and widely used machine learning techniques are deep learning algorithms such as recurrent neural networks (RNN) with LSTM, bidirectional LSTMs with convolutional neural network (BiLSTM-CNN), and reinforcement learning, each with its advantages and disadvantages. Due to specific application requirements of each machine learning technique in the IoT network setting, in this work, we aim to propose a lightweight and efficient

intrusion detection model for the IoT network. We propose a lightweight intrusion detection model for IoT networks called DEvoid Intrusion Detection System (DEvoid-IDS) that uses state-of-the art deep reinforcement learning algorithm for capturing the time-based behaviour of IoT devices.

4.1. Deep Q-Networks (DQN)

When the current working mode of the system is selected as a threshold indicating the energy level, there does not exist an arbitrary adjacent threshold between them. The DQL-based scheme is effective in achieving a power-balanced state. It mainly considered the optimization function compliant with user networks and protected unlicensed fading and applied trend power economics.

The idea of the deep Q-network (DQN) algorithm is to address the dimensions that cannot be calculated by the antenna via supervised learning and introduce the concept of delay. The advantages of the algorithm are its strong reliability and it can be concurrently learning within the system to identify the system of observable parameters. Because of the intelligent control method of the proposed algorithm, it is capable of learning human control to stop the delay and the power system can learn how to discharge the electric vehicle and resolve the voltage frequency defect, enhancing the robustness of the power and reducing the drop in system voltage.VERTISEMENTS. The DQN stands out for its time relationship and can generate a fixed-effective solution that can deal with a high-value of no normalization and have experience reply throughout its inner state and there is no need for additional data for the SIRI created links. [20] The proposed autonomous deep reinforcement learning and routing method are validated via train system simulations. It trains the routing function of cyber attack traffic in large scale OT networks through model poison cyber attack traffic by using deep Q-Network inference competitions to test the robustness of the method. The simulation results show that a deep reinforcement learning algorithm driven by a Rayleigh Network has the best performance on the IOT-DDOS traffic under different attack energy parameters. Minimize the outage probability.

[5] The DQN algorithm was proposed by DeepMind in 2013, and it utilized CNN to predict the Q-value function, where the method was to get the maximum expected value of the Q-value function with the current state as input. CNNs can address the relationship between the reward function and the state-action space, and generate complex mapping interactions,

which can reduce the time-domain correlation relationship between data, and charge the Q-value label more effectively. The neural network can learn the Q-value mapping under noise input, and the correlation of the temporal difference error extends the target function. In the early stage, the deep reinforcement learning method required a large number of samples to explore the environment, but DQN resolved it through some targeted technology, such as experience replay, which can retain data related to the sequence of experience, and fix the instability of the network, enhancing network learning and prediction accuracy.

4.2. Policy Gradient Methods

Under the framework of Policy Gradient, various algorithms are designed to update the policy directly according to the observed returns. These algorithms are referred to as Monte Carlo Policy Gradient, whose parameter update rule is so that C = R − V, where V ∈ Bel and R = ∑t j=t d γ t−j r. For identity R, a variant of Monte Carlo called REINFORCE is derived. These algorithms are variance by nature and it usually takes many iterations to obtain a good return estimate. During learning, as the policy changes, the states distribution changes. Therefore, it is of importance to derive an update such that the states and the actions from the updated policy do not heavily depend on the policy that in the future may be obsolete, leading to collapsed learning.

[21] [22]Policy Gradient methods, also known as asynchronous advantage actor-critic (A3C) or Proximal Policy Optimization (PPO), are a class of reinforcement learning algorithms that optimize an objective function directly with respect to the policy it seeks to train [19]. They are gradient-based methods that update the policy towards states that have a high advantage or value based on the observations that are experienced. A key insight in this work is that the policy is updated with large gradient steps along those states with a low action value and with small gradient steps along those with a high action value. This differentiation between the high-valued experience and low-valued expérience is referred to as bootstrapping, which is a technique recommended by Sutton and Barto in their seminal book, which contains comprehensive coverage of reinforcement learning algorithms.

4.3. Advancements in Deep Reinforcement Learning

The results of the experiments illustrate that IoT-reinforcement with autonomous vehicles is superior to the IoT-reinforcement learning for the smart grid without autonomous vehicles,

in terms of enhanced cumulative reward [23]. Moreover, protection against higher number of attacking agents is offered by the IoT-reinforcement with a single generation of autonomous vehicles in the vicinity of the CNS, against the IoT-reinforcement with no autonomous vehicles, i.e., the simplest case of using IoT-reinforcement learning. In light of these observations, it is observed that IoT-reinforcement learning outperforms the on-vehicle ID-IT scheme, as the reward offered by the ID-IT based vehicles is significantly lower than that of IoTreinforcement-based vehicles. The emphasis in this work is on developing security mechanisms that can withstand several attacks, and the approaches' compatibility with security blacklists, intrusion detection systems and AI-based detection techniques designed for thwarting different attacks are considered in the future as an extension to this work.IoT-reinforcement with deep learning thereby improves the data collection and execution time with virtual reality as compared to in reality simulations.

Reinforcement learning is gaining popularity for securing IoT systems as it can learn the environment with minimal information and adapt parameters on-the-fly. It is proposed as an emerging solution for IoT security, with the potential to combat various cyber-attacks [24]. Reinforcement learning with autonomy can be designed for adaptive defense in the IoT-connected smart grid with autonomous vehicle networks, composing Cyber-Physical Systems (CPSs). However, there is a scarcity of studies, which introduces deep reinforcement learning to secure IoT-connected autonomous vehicles in the smart grid [25]. In this paper, the scope includes attacks on the system components viz. authentication gateways, navigation control stations, Data Initially Deployed (DID) energy meters, IoT-connected plug-in electric vehicles, and the system components utilization for attack propagation viz., communication routers and IoT-connected plug-in electric vehicles.

5. Application of Deep Reinforcement Learning in Cyber Defense

The limitation of above applications in reference to cyber security is the usage of less sophisticated network models or single layer intrusion detection. IoT networks are connected with numerous algorithms, regardless of malicious and secure nodes. Proposed reinforcement learning intrusion models address the issue by using deep learning models. In features, policy determination also improves the cybersecurity of IoT systems. Reinforcement learning is applied to face the intrusions generated by adversaries but fails to focus on reducing the

complexity of well-ness tracking across the big IoT networks including IoT connected autonomous vehicles network [refs: 895940c4-44b3-4a8d-92e8-24b0b52eaf55].

[16] [26] Development of several applications and cyber attacks in Internet of Things (IoT) are attainable due to a fast interaction of some nodes registering some critical vehicular details. Foundational literature publications considers the application of reinforcement learning in cyber-security of IoT networks. In [refs: 38d6adce-f5fd-46aa-b160-cad22961f2b5,3abf9c37-bf8b-4278-8dd6-bb58ba6e7164 search a multi-layer neural networking intrusion detection model in IoT systems. Supervised or semi-supervised frameworks for intrusion detection in IoT connected autonomous vehicles networks are provided by Deep Learning or Neural Networks models. In the security aspect, inappropriate data exchange between IoT subjects can amplify cybersecurity threats in connection with the automotive industry. Secure communication methods like Secure Multiparty Computation (SMC) and Blockchain help to provide security services in IoT networks.

6. Case Studies

[ref: 1970cd60–746f-41b8–8065-7b8e62ef4932]We show in the simulations and real-time experiments that even the simplest agent is able to instantly decentralize multiple system–communication take the right decisions against both single and compound attacks. Comparing the DRL-based system against both the central algorithm and classic scheduling for various attack scenarios, we observe that DRL-based systems result in the highest defence/communication throughputs, the least average vehicular allocation error, and the fastest learning rate.

[ref: 1970cd60–746f-41b8–8065-7b8e62ef4932]We implement a multi-task learning-based deep learning system to adaptively protect a set of 12 autonomous vehicles over a 433.6 m urban road stretch from cyber-physical attacks. We simulate network attacks selectively jamming the sensor-critical speed, LiDAR signature, and wireless reliability over varying intensities. We model a tight coupling between the potential of warping vehicle trajectory via sensor interface and delayed consequences via wireless communication tampering aiming at stealth and systematical attacks. The counteraction system combines the sensor and communication metrics via IoT-based task allocation as a medium access control, edge-implemented reinforcement learning. The learning benefits from map availability for the agent to leverage the context in the observations.

[ref: 895940c4–44b3-4a8d-92e8-24b0b52eaf55][ref: 1970cd60–746f-41b8-8065-7b8e62ef4932]In this case study, we propose a general methodology to use DRL for adaptive cyber defence in IoT-connected autonomous vehicle networks. We integrate the safety needs with autonomous vehicle sensors, communications, and high-definition maps into a MDP and apply DRL to select the optimal protection action. The experimental results demonstrate the effectiveness of the proposed network configuration and motivate the need for further research on the application of DRL solutions for IoT-connected autonomous vehicles.

6.1. Simulation Setup

Authors needed to simulate the environment the best they could. To start let's briefly describe the industrial simulator based on Veins, as the rest will be explained with the points where that is important. The simulator is a powerful tool that allows us to model a specific (multilane urban) environment with a number of vehicles and a number of access nodes (RSUs) in this environment [5]. It is based on OMNeT++, that is a modular programmable simulation environment, has real-time support, and is Java-like and C++-like extension language, used by more than 60.000 downloaders. The Veins framework is a part of the simulations. It a model library mixed with INET that represents a vehicular-networking popular cluster. The outcoming events are given by handleMessage. But these vehicle models have fundamental limitations, so adaptation for their own vehicle model should be carried out, like a Cooperative Adaptive Cruise Control modelled in CarNetOMNeT++. OMNeT++ only uses C++, so we could use the Carla CAR simulating tool for Pre-Simulation Design.

The first step in the setup is building the environment, which is defined in the class AITUCyberClass. This class contains information about the state, reward provided, information on the network, and if the vehicle has or has not a traffic jam. In conclusion, this class serves as the Environment (OpenAI Gym) for the autonomous vehicle. Then, the agent (the deep Q-Learning agent) was built according to the article by Ajodhia, et al. [6]. This agent learns the best policy in an environment (using the Q-table). This is the learning process divided in exploration and exploitation.

6.2. Results and Analysis

The rapidly increasing scale of autonomous vehicles, and increasingly complex forms of communication could also make the present systems and technologies unsustainable, and

vehicle communication networks may suffer increasingly complex cyber attacks in the future. A large number of researchers have established solutions to improve security and system robustness, and many intrusion detection and reinforcement learning-based systems have been used in different fields such as in-vehicle communication networks, news articles, smart grids and the industrial Internet [27]. However, there has been little development in models where autonomous vehicles are connected to the IoT. After comprehensive analysis of the potential attack scenarios faced by the network of autonomous vehicles combined with their communication networks, and after analyzing existing problems and breakthroughs, this paper designed a network security model for the communication of autonomous vehicles in an IoT, and proposed a new intelligent network security solution, reinforcement learning-based adaptive intrusion detection system RAIDS.

Powerful applicability: RAIDS can cooperate with different network security equipments and adapt to various states of the network automatically; Strong robustness: proposing a new network security test environment by using Mininet and SODA to simulate different types of network attacks and normal network situation. Under this environment, the new proposed solution is more comprehensively tested and proven to have very strong robustness. RAIDS is applied to the real network system of vehicle platooning, demonstrating that this system operates well in a real network communication environment [1].

Cyber attacks on the communication networks of autonomous vehicles can lead to traffic safety incidents, economic loss, and other severe consequences [6]. It can be difficult to defend against cyberattacks due to a variety of attack strategies, and even though intrusion detection systems (IDS) can identify some types of cyberattacks, the effectiveness of many cyberattack detection products is not high. This paper describes a reinforcement learning-based system for adaptive IDS (RAIDS), which has these advantages:

7. Challenges and Future Directions

In future research, we endeavor to set the foundation for measures, useful tools, and data-driven assessment processes to monitor for newly introduced strains of malwares in systems. This includes making adjustments to prevention security policies, user awareness trainings, and including traces of malicious cyber-physical systems within adversarial datasets for autonomous security checking. Furthermore, it is possible to contemplate the extension of the eCybersecurity strategy to other human or social factors in V2X environments. We plan to

consider combined user identity, mobility, and other complex topics, from persona or profiling orientations, to construct new models for anomaly detection and reinforcement learning applications [23]. Other expansion areas include contributions of user and automotive traces, electronic devices stored data, models, and skins, MCP integration, supervised and unsupervised clustering, association rules, and selected algorithms and skill collections into comprehensive data for evaluating and establishing security levels. We also intend to deepen testing compatibility, multipath validation, and other sophisticated studies to apply in complex real-world situations. We hope that this work encourages a broader, open and deeper interwoven interaction of researchers in various disciplines for constructing effective and coherent AI defense solutions universally across modern data-intrinsically deployed systems, particularly IoT-connected AVs, and nurtures the formation of international research initiatives and platforms for interdisciplinary cooperation.

In this work, we leverage knowledge from evolving research thrusts in machine learning, data mining, pervasive computing, smart cities, security, and software engineering to emerge pioneering contributions in the context of AIoVNs to (1) detect states and gauge risks associated with the IoT-connected AVs, (2) adaptively characterize normal/abnormal activities ("attacks") related to the driving patterns and safety constraints of the AVs, and (3) take appropriate real-time cybersecurity actions using DRL-based decision-making models [28]. Our research directions are exploratory and multidisciplinary, with a focus on the intersection of AI and autonomous vehicles future technology infrastructures, including an encompassing analysis of the risks associated with ill-intentioned interventions. Ultimately, our purpose is to develop an advanced support system to tackle the above security and cybersecurity challenges in future AIoVNs, focusing on the interplay of pervasive computing and data-driven decision-making security plan and evaluation in the context of autonomous vehicles and smart cities.

8. Conclusion and Implications for Practice

From a broader perspective, our paper envisions a new cyber defense approach, where the current deadline-oriented security strategies are replaced by those that learn from new attack patterns [23]. The dynamic nature of the new threats to the underlying systems should motivate the design of adaptive security strategies endowed with learning abilities, similar to those recently applied to machine learning models, called adversarial machine learning. Here

we focused on a reinforcement learning (RL) approach since it has been successfully applied to dene the optimal defense strategies against intelligent attackers. Furthermore, RL allows to capture the sequential nature of the defense in interacting with the attacker at the proof of concept level. This approach has been successfully applied also to design Intrusion Detection System to detect and respond to cyberattacks targeting the integrity of the sensing and communications stacks of the IoT [2]. However, our model can be extended in several directions to increase the realism of the learning process underlying the AV's defense. For instance, in this work, we showed that the AV will act following the will of the attacker, thus being forced to slow down multiple times. A more effective AV's policy can be jointly designed by evaluating the security and the control performance for several choices of the dynamics parameters pairs using reinforcement learning techniques.

In this work, we proposed a novel adaptive cyber defense strategy for Internet of Things (IoT)-connected autonomous vehicle (AV) networks [29]. In particular, we leveraged an adversarial deep reinforcement learning framework to approximate both the defense and attack policies by two deep neural networks. We then analyzed the interactions between the AV and the attacker as a repeated Stackelberg game. In this game, the AV's policy is influenced by the state of the environment and the EWMA speed estimate, while the attacker is affected by the actions taken by the AV and its defense policy. Due to their different aims, the agents act so as to maximize their own returns, thus being involved in an adversarial process. Moreover, we allowed the defender not only to spend an action to recover the speed control system from attacks, but also to invest another action to adapt the dynamics controller of the AV, so as to eventually help the AV robustness against future attacks. In this sense, our defense policy is made adaptive to the attacks crafted by the attacker, and this allows the AV to constantly upgrade its defense strategy against unknown or innovative attacks. Realistic simulations confirmed that the adopted deep RL algorithm effectively enhances the resilience of the entire automotive control stack to different data injection attacks and enables an efficient merging of multiple sensors observations to track the speed of the AV.

9. References

1. [1] R. Singh Rathore, C. Hewage, O. Kaiwartya, and J. Lloret, "In-Vehicle Communication Cyber Security: Challenges and Solutions," 2022. ncbi.nlm.nih.gov
2. [2] S. K. B Sangeetha, P. Mani, V. Maheshwari, P. Jayagopal et al., "Design and Analysis of Multilayered Neural Network-Based Intrusion Detection System in the Internet of Things Network," 2022. ncbi.nlm.nih.gov
3. [3] S. Mahadik, P. M. Pawar, and R. Muthalagu, "Efficient Intelligent Intrusion Detection System for Heterogeneous Internet of Things (HetIoT)," 2023. ncbi.nlm.nih.gov
4. Mahammad Shaik, et al. "Envisioning Secure and Scalable Network Access Control: A Framework for Mitigating Device Heterogeneity and Network Complexity in Large-Scale Internet-of-Things (IoT) Deployments". Distributed Learning and Broad Applications in Scientific Research, vol. 3, June 2017, pp. 1-24, https://dlabi.org/index.php/journal/article/view/1.
5. Tatineni, Sumanth. "Beyond Accuracy: Understanding Model Performance on SQuAD 2.0 Challenges." *International Journal of Advanced Research in Engineering and Technology (IJARET)* 10.1 (2019): 566-581.
6. Vemoori, V. "Towards Secure and Trustworthy Autonomous Vehicles: Leveraging Distributed Ledger Technology for Secure Communication and Exploring Explainable Artificial Intelligence for Robust Decision-Making and Comprehensive Testing". *Journal of Science & Technology*, vol. 1, no. 1, Nov. 2020, pp. 130-7, https://thesciencebrigade.com/jst/article/view/224.
7. [7] T. H. H. Aldhyani and H. Alkahtani, "Attacks to Automatous Vehicles: A Deep Learning Algorithm for Cybersecurity," 2022. ncbi.nlm.nih.gov
8. [8] J. Egger, A. Pepe, C. Gsaxner, Y. Jin et al., "Deep learning—a first meta-survey of selected reviews across scientific disciplines, their commonalities, challenges and research impact," 2021. ncbi.nlm.nih.gov
9. [9] Y. Wang, A. Smahi, H. Zhang, and H. Li, "Towards Double Defense Network Security Based on Multi-Identifier Network Architecture," 2022. ncbi.nlm.nih.gov
10. [10] S. Oesch, P. Austria, A. Chaulagain, B. Weber et al., "The Path To Autonomous Cyber Defense," 2024. [PDF]

11. [11] S. Ullah, M. A. Khan, J. Ahmad, S. Shaukat Jamal et al., "HDL-IDS: A Hybrid Deep Learning Architecture for Intrusion Detection in the Internet of Vehicles," 2022. ncbi.nlm.nih.gov
12. [12] A. Demontis, M. Pintor, L. Demetrio, K. Grosse et al., "A Survey on Reinforcement Learning Security with Application to Autonomous Driving," 2022. [PDF]
13. [13] N. Parvez Farazi, T. Ahamed, L. Barua, and B. Zou, "Deep Reinforcement Learning and Transportation Research: A Comprehensive Review," 2020. [PDF]
14. [14] E. Bates, V. Mavroudis, and C. Hicks, "Reward Shaping for Happier Autonomous Cyber Security Agents," 2023. [PDF]
15. [15] X. H. Nguyen, X. D. Nguyen, H. H. Huynh, and K. H. Le, "Realguard: A Lightweight Network Intrusion Detection System for IoT Gateways," 2022. ncbi.nlm.nih.gov
16. [16] T. Thi Nguyen and V. Janapa Reddi, "Deep Reinforcement Learning for Cyber Security," 2019. [PDF]
17. [17] P. Vaishno Mohan, S. Dixit, A. Gyaneshwar, U. Chadha et al., "Leveraging Computational Intelligence Techniques for Defensive Deception: A Review, Recent Advances, Open Problems and Future Directions," 2022. ncbi.nlm.nih.gov
18. [18] M. Karkheiran, "Yukawa Textures From Singular Spectral Data," 2021. [PDF]
19. [19] L. Yu, S. Huo, K. Li, and Y. Wei, "A Collision Relationship-Based Driving Behavior Decision-Making Method for an Intelligent Land Vehicle at a Disorderly Intersection via DRQN," 2022. ncbi.nlm.nih.gov
20. [20] J. Mern, K. Hatch, R. Silva, J. Brush et al., "Reinforcement Learning for Industrial Control Network Cyber Security Orchestration," 2021. [PDF]
21. [21] B. Ben Elallid, A. Abouaomar, N. Benamar, and A. Kobbane, "Vehicles Control: Collision Avoidance using Federated Deep Reinforcement Learning," 2023. [PDF]
22. [22] X. Xiong and L. Liu, "Combining Policy Gradient and Safety-Based Control for Autonomous Driving," 2023. [PDF]
23. [23] H. Cao, W. Zou, Y. Wang, T. Song et al., "Emerging Threats in Deep Learning-Based Autonomous Driving: A Comprehensive Survey," 2022. [PDF]
24. [24] A. Uprety and D. B. Rawat, "Reinforcement Learning for IoT Security: A Comprehensive Survey," 2021. [PDF]

25. [25] K. Istiaque Ahmed, M. Tahir, M. Hadi Habaebi, S. Lun Lau et al., "Machine Learning for Authentication and Authorization in IoT: Taxonomy, Challenges and Future Research Direction," 2021. ncbi.nlm.nih.gov
26. [26] Y. Deng, T. Zhang, G. Lou, X. Zheng et al., "Deep Learning-Based Autonomous Driving Systems: A Survey of Attacks and Defenses," 2021. [PDF]
27. [27] J. Wiebe, R. Al Mallah, and L. Li, "Learning Cyber Defence Tactics from Scratch with Multi-Agent Reinforcement Learning," 2023. [PDF]
28. [28] D. K. Galloway, J. J. M. in 't Zand, J. Chenevez, L. Keek et al., "The Influence of Stellar Spin on Ignition of Thermonuclear Runaways," 2018. [PDF]
29. [29] A. Olivares-Del Campo, S. Palomares-Ruiz, and S. Pascoli, "Implications of a Dark Matter-Neutrino Coupling at Hyper-Kamiokande," 2018. [PDF]