2022 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW) | 978-1-6654-9628-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICSTW55395.2022.00016

Towards improving explainability, resilience and performance of cybersecurity analysis of 5G/IoT networks (*work-in-progress paper*)

Manh-Dung Nguyen, Vinh Hoa La, Ana R. Cavalli, Edgardo Montes de Oca
Montimage EURL, France
{manhdung.nguyen, vinh_hoa.la, ana.cavalli, edgardo.montesdeoca}@montimage.com

***Abstract*—Artificial Intelligence (AI) is envisioned to play a critical role in controlling and orchestrating 5G/IoT networks and their applications, thanks to its capabilities to recognize abnormal patterns in complex situations and produce accurate decisions. However, AI models are vulnerable to adversarial attacks, thus the societal view is far from trustworthy as to its usage in safety critical areas relying on 5G/IoT networks. In this paper, we present ongoing work being done in the H2020 SPATIAL project that targets developing and evaluating AI-based modules for anomaly detection and Root Cause Analysis in the 5G/IoT context regarding different criteria, such as explainability, resilience and performance on a real 5G/IoT testbed.**

***Index Terms*—5G, IoT, Machine Learning, Explainable AI, Cybersecurity, Resilience**

I. Introduction

A. Context

5G and IoT networks hold the promise of delivering ultra-low latency, ultra-high throughput, ultra-high reliability, ultra-low energy usage, and massive connectivity. Achieving these will pave the way to a new breed of applications, including autonomous driving, industry 4.0, augmented and virtual reality, collaborative gaming, near realtime remote surgery, and telepresence. However, the diversity of services and the huge number of connected IoT devices, envisaged in the future networks, will introduce new and increasingly broad cybersecurity and privacy risks [1]. Thus, it is crucial to develop effective and sustainable security solutions that can deal with the evolving threat landscape in order to satisfy security requirements of 5G/IoT systems and applications.

Technical security measures toward 5G/IoT networks are quickly embracing a variety of machine learning (ML) algorithms as an effective approach to empower intelligent, adaptive and autonomous security management, allowing to tackle the growing complexities of the network. Indeed, AI has the potential of recognizing abnormal patterns from a large set of time-varying multi-dimensional data, and delivering faster and accurate decisions [2]. However, the adoption of AI methods in IoT and future mobile networks is still in its infancy. Thus, research efforts need to take into consideration diverse aspects of AI solutions and practical implementation issues to support both users and developers in effectively auditing the code and data of safety-critical systems.

B. Problem and Challenges

AI-based systems have three major issues in 5G/IoT domain: (1) lack of real-world datasets; (2) lack of explainability; and (3) lack of resilience against adversarial attacks.

Lack of real-world datasets. AI models, such as supervised ML, require large amounts of data with correct labels, so the quality of data has a great impact on the advancement of modern AI research. However, diversified real-world datasets in 5G/IoT networks is not readily available due to privacy issues that need to be followed by all telecom operators.

Lack of explainability. Currently, AI schemes applied in security solutions focus mainly on accuracy and performance (e.g., precision, recall, resource utilization) and do not readily offer an explanation of why a particular output is obtained. Explanation of a decision taken often becomes a critical requirement for the 5G network, especially because many critical services depend on the 5G infrastructure.

Lack of resilience against adversarial attacks. ML models are vulnerable to adversarial attacks [10] where adversarial inputs are small carefully-crafted perturbations in the test data built for fooling the underlying ML models into taking the wrong decisions. Robustness against adversarial attacks is a challenging problem as there does not exist a solution that ensures complete protection against this kind of attacks.

C. Proposal

In the on-going work presented here, we aim to tackle the above challenges by: (1) producing real-world datasets for AI training, especially for 5G and encrypted network traffic thanks to our open source 5Greplay tool [3]; (2) enabling explainability features of existing AI algorithms in our different AI-based systems, such as MMT-Probe for anomaly detection and MMT-RCA for Root Cause Analysis; and, (3) considering the security threats that emerge from the rapid adoption of AI algorithms in 5G/IoT networks. Existing challenges of AI in 5G/IoT networks motivate us to make the ML models more *accurate, explainable* and *robust* before they are integrated into complex systems. To do so, we provide a real 5G/IoT testbed that allows us to evaluate various AI techniques related to the explainability, resiliency and distribution of AI models. This will be done by assessing the techniques used today by

978-1-6654-9628-5/22/$31.00 ©2022 IEEE
DOI 10.1109/ICSTW55395.2022.00016

Montimage[1] in its MMT security monitoring framework for performing cybersecurity analysis and protection of 5G/IoT networks, encrypted traffic analysis and RCA, in the H2020 SPATIAL[2] project.

II. Background

A. Potential of AI for Cybersecurity in 5G/IoT

This section introduces some potential applications of AI for cybersecurity in 5G/IoT networks.

a) Anomaly Detection in (Encrypted) Network Traffic: In the 5G/IoT context, an early detection and prediction of potential anomalous behaviors in the network enables fast reaction to them, preventing financial loss, malicious damage and service degradation. AI has been shown to help detect hidden or abnormal traffic patterns that can lead to security threats or service unavailability in 5G/IoT networks. For instance, [6] leverages the clustering algorithm DBSCAN to effectively detect anomalies caused by radio attenuation and SDN misconfiguration for self-healing of 5G Radio Access Networks (RAN). [8] applies different ML models to predict attacks and anomalies in IoT systems and finds that Random Forest techniques perform comparatively better than others.

The growing popularity of traffic encryption increases user security and privacy at the individual level, but also becomes a big challenge for performing traffic analysis. This raises the need for advanced analysis techniques based on other criteria, such as network packet and flow behavior analysis. With the introduction of network encryption protocols, such as Transport Layer Security (TLS), the accuracy and efficiency of conventional Network Intrusion Detection Systems (NIDS) using rule and signature-based monitoring detection methods is greatly reduced. Moreover, the variety and dynamicity of network malware poses a significant challenge on traffic monitoring tools in terms of flexibility and generalization of their algorithms. Advanced ML techniques, like Deep Learning (DL), is a highly desirable approach for mobile services traffic classification, and particularly encrypted traffic analysis [9].

b) Root Cause Analysis (RCA): RCA plays a vital role in the Risk Management process to accurately identify the cause of faults or security incidents in different domains, such as IT operations and telecommunications. The root cause diagnosis becomes highly intractable or even impossible because of the complexity and heterogeneity of emerging mobile networks (e.g., introducing virtualised functions, Software-Defined Networking), coupled with the increasing number of Key Performance Indicators (KPIs) and data related to end-users, devices, services and networks. Also, human-based mitigation actions become more challenging and time consuming in complex 5G/IoT systems. This leads to the need of an automated tool helping humans to troubleshoot a system and determine which events are causally connected and which are not. AI has been recognized as an appealing option for fostering automated Root Cause Analysis, thanks to its ability to process a large amount of data, uncover complex non-linear relationships within the data, and deliver faster and accurate conclusions. For example, Zhang et al. [5] proposed a DL-based Root Cause Analysis of faults in a cellular RAN by leveraging supervised classification and unsupervised clustering.

c) Moving Target Defense: Moving Target Defense (MTD) has emerged as an effective proactive security solution to address the problem of exploring and exploiting the unchanging vulnerability surface. The MTD can be established through various implementations including, IP address shuffling, Virtual Machine migration, network path diversification, and replication of software or network resources to increase the attacker's effort and cost. The flexibility and dynamicity opportunities provided by virtualization (i.e., Network Function Virtualization) and programmability (i.e., Software Defined Networking) will foster the implementation of MTD mechanisms in 5G/IoT networks, leading to more resilient networks. AI techniques, including game theory, genetic algorithms and ML, have been considered highly relevant to devise smart MTD mechanisms that can intelligently decide changes to make on the network and service configuration in order to meet the security/performance trade-off.

B. Overview of AI-based Attacks

An attack against AI systems can be classified as *white-box*, *grey-box* or *black-box*, depending on whether the adversary has, full, some or no knowledge about the training data, the learning algorithm and its hyper-parameters, respectively. An adversary can attack the ML models during its training phase and in its testing phase as well. In *poisoning attacks*, attackers seek to influence the outcome of the learning phase to their advantage by tampering the data or the learning algorithms used. In *evasion attacks*, the attacker aims at perturbing the learned model so that attacks are considered legitimate or remain undetected. This is commonly known as adversarial attacks. For instance, Goodfellow et al. [10] proposed one of the first and most popular adversarial attacks to fool a neural network using a gradient-based optimization method.

C. Explainable AI (XAI)

XAI [4] is a promising set of technologies that increases the AI black-box models' transparency to explain why certain decisions were made. While AI plays a critical role in 5G/IoT networks, XAI is crucial to enhance trust and transparency for enabling their use in critical networks. In the following, we introduce some of the most popular XAI techniques.

Local Interpretable Model-agnostic Explanations (LIME) is a widely popular technique used in interpreting outputs of black-box models in several fields and applications. LIME gives a local explanation, which means that it considers a subset of data when approximating explanations for model predictions. *Shapley Additive Explanations* (SHAP) is an XAI technique that identifies the importance of each feature value in a certain prediction using popular cooperative game theory technique. *Permutation Feature Importance* is a global XAI method that measures the increase in the prediction error of

[1]https://www.montimage.com

[2]https://spatial-h2020.eu/project/

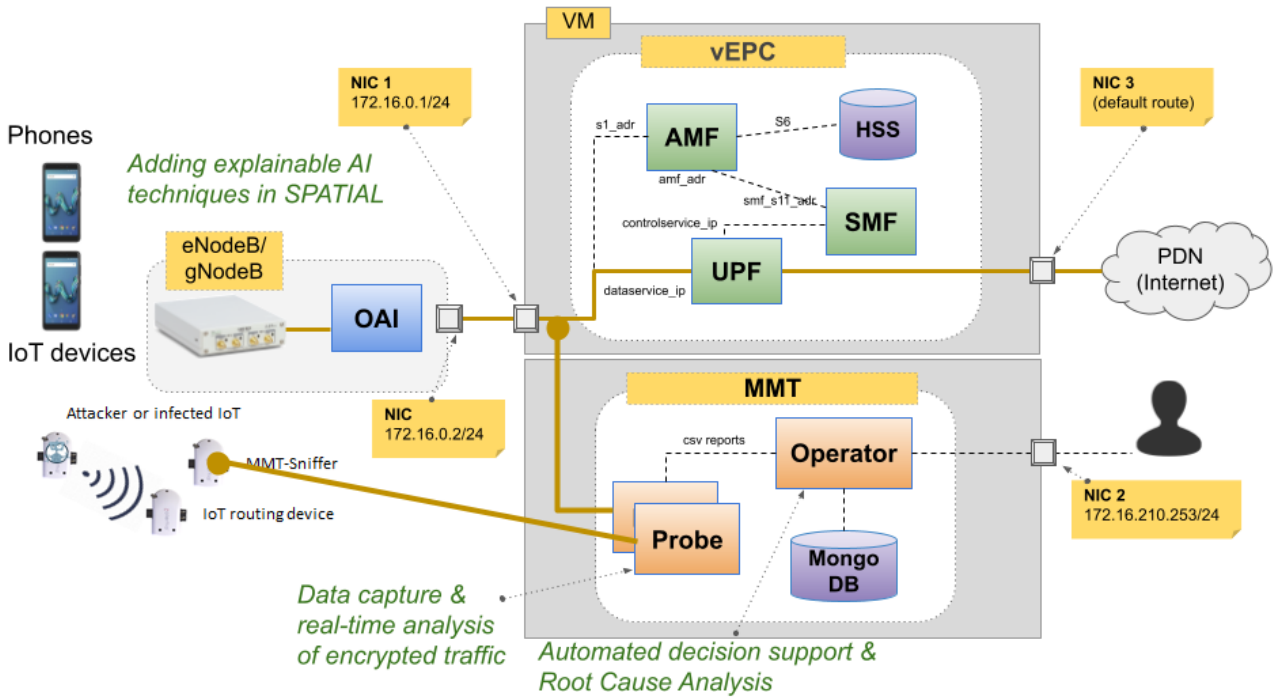

Fig. 1. Our real 5G/IoT testbed.

the model after one permutes the feature's tabular values. To assess how important a specific feature is, we compare the initial model with the new model on which the feature's values are randomly shuffled.

III. Improving the Explainability, Resilience and Performance of Cybersecurity Analysis of 5G/IOT Networks

A. Overview of Our Testbed

Figure 1 shows the overall architecture of our real 5G/IoT testbed. This pilot corresponds to a 5G/IoT solution consisting of an gNodeB based on a Software-Defined Radio, a portable 5G Core solution, and the MMT security monitoring framework. The framework is based on distributed and extensible MMT-Probes that analyze (encrypted) network traffic from both the mobiles and IoT devices. The MMT-Operator is a Web application that allows further analysis and acts as a decision point for determining the causes of breaches and triggering corresponding countermeasures. Our AI-based tools, including MMT-Probe for cybersecurity analysis of (encrypted) network traffic and MMT-RCA for Root Cause Analysis, will later be evaluated with respect to the improved performance, transparency and precision they bring to the security analysis and algorithms.

IoT testbed. Our IoT testbed [7] includes a set of equipment forming an IoT 6LoWPAN network: several Zolertia RE-Motes, one Raspberry Pi and accessories. A border router mote acts as the gateway collecting sensed data from the other motes, and forwards the reports via the USB line to the server deployed in the Raspberry Pi. The IoT network consists of normal clients reporting sensed data every 10 seconds, and one (or several) attacker(s) behaving interchangeably in one of three modes: (1) *normal mode* reporting data every 10 seconds, (2) *denial of service (DoS) attack mode* reporting data 100 times faster (10 messages/second), and (3) *dead mode* not reporting data at all. An IoT device, MMT-Sniffer, performs sniffing tasks by capturing network traffic and piping it via the USB line to the Linux-based machine where MMT-Probe is deployed to analyze the traffic and extract the metrics for our MMT-RCA. A Raspberry Pi feeds the motes in terms of batteries, hosting the server dealing with the sensed data and receiving the sniffed traffic which is then analyzed by the MMT-Probe.

5G testbed. EPC-in-a-Box[3] platform represents a 4G/5G network core commercialized by Montimage and Cumucore[4]. It is a ready-to-use appliance allowing the creation of a full end-to-end 4G/5G network in 5 minutes. It can be used not only for testing but also to create a small-scale mobile network in order to provide mobile connection in industry, white or gray zones. It basically consists of 3 main building blocks: Radio Access Network (RAN), Evolved Packet Core (EPC) or 5G Core, and Montimage Monitoring Tool (MMT). Once deployed, the testbed platform creates a 4G/5G network allowing commercial of the shelf User Equipments (UEs) to connect. After being successfully attached, the UEs have access to the IP services in PDN or from the internet such

[3]https://www.montimage.com/products#EPC-in-a-box
[4]https://www.cumucore.com

as browsers, Web applications, VoIP video calls, etc. All the traffic between the RAN and the EPC is captured and analyzed in real-time by MMT-Probe to ensure that the defined security properties are respected. MMT-Operator supports automated decision and reaction in the case an anomaly is detected.

B. XAI-based Root Cause Analysis

In this section, we introduce our similarity-based machine learning approach implemented in MMT-RCA that takes into account highly granular monitoring indicators (e.g., statistics extracted from the logs, metrics, network traffic, and any data that could identify the system state), and performs deep analysis to assess the similarity of a newly observed event reflecting the current status of the system and each learned one saved in the historical database. RCA enables systematizing the experience in dealing with incidents to build a historical database and verify whether a newly detected incident is similar enough to an observed one with known causes.

To evaluate the RCA enabler, we performed experiments on our real IoT testbed. To improve the accuracy and explainability of our MMT-RCA, we employed feature selection, which is one of the core concepts in ML that tremendously impacts the model performance. Indeed, it is common that the data collected in a complex changing system is too complicated or redundant. In other words, there might be some irrelevant or less important attributes (i.e., noise) contributing less to the target variable. Removing the noise helps not only to improve the accuracy but also reduce the training time. It is the first and most essential step that should be performed automatically based on the feature selection techniques, or manually by system experts. Our tool, MMT-RCA, has been integrated with the following feature selection techniques: (1) *Univariate feature selection* that is based on univariate statistical tests; and, (2) *Recursive feature elimination* for selecting features by recursively considering smaller and smaller sets of features.

Figure 2 summarizes the results when different Feature Selection models are employed. There are six attributes, namely (1–3), (5–7), which are considered significant according to all the models. Four attributes (8), (10), (11), and (12) are concluded to be not relevant and can be left out. The attributes (4) and (9) are recommended by some models and not by others. To summarize, feature selection, similarity learning and Bayesian networks are currently employed in our MMT-RCA in the 5G/IoT context, but these techniques need to be improved to make them more resilient to attacks by considering adversarial attacks, transparency and explainability by applying more common XAI methods, such as LIME or SHAP, and privacy-awareness to protect the privacy and legitimate concerns of the users when applying AI techniques.

	Univariate feature selection			Recursive feature elimination	
Ref.	Chi-square test	f-test	Mutual information classification test	Logistic regression model	Random forest model
(1)	true	true	true	1	true
(2)	true	true	true	2	true
(3)	true	true	true	2	true
(4)	true	false	true	6	false
(5)	true	true	true	3	true
(6)	true	true	true	2	true
(7)	true	true	true	2	true
(8)	false	false	false	9	false
(9)	false	true	true	6	false
(10)	false	false	false	8	false
(11)	false	false	false	10	false
(12)	false	false	false	10	false

Fig. 2. Feature Selection results given the following attributes (1) Network throughput, (2) Throughput at devices, (3) Traffic transmitted on links, (4) Number of routing-related packets, (5) Transmission delay, (6) CPU usage, (7) Memory usage, (8) Battery level, (9) Power consumption, (10) Average packet size, (11) Probe ID, (12) Protocol ID.

IV. Conclusion and Future Work

Security and privacy are uncompromising necessities for modern and future global networks environments such as 5G/IoT. This ongoing work discusses at a high-level the potential applications and limitations of AI in 5G/IoT networks. To improve the explainability of AI models, we will apply well-known XAI techniques and new ones developed by the SPATIAL project partners into our AI-based modules. Furthermore, in view of increasing the resilience to AI threats, we will consider and implement different evasion and poisoning attacks against ML models used in cybersecurity detection and management.

Acknowledgment

This work was supported in part by the European Union's Horizon 2020 research and innovation programs through the SPATIAL project (Grant agreement ID: 101021808).

References

[1] ENISA, "*Threat Landscape for 5G Networks.*" Nov. 2019.

[2] Camps-Mur et al, "*AI and ML – Enablers for Beyond 5G Networks.*" 5G-PPP White Paper, 2021.

[3] Zujany Salazar, Huu Nghia Nguyen, Wissam Mallouli, Ana R. Cavalli, Edgardo Montes de Oca: 5Greplay: a 5G Network Traffic Fuzzer - Application to Attack Injection. ARES 2021: 106:1-106:8.

[4] Arrieta et al, "*Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI.*" Information Fusion 58 (2020): 82-115.

[5] W. Zhang et al, "*Self-Organizing Cellular Radio Access Network with Deep Learning.*" IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2019.

[6] J. Ali-Tolppa, S. Kocsis, B. Schultz, L. Bodrog and M. Kajo, "Self-healing and Resilience in Future 5G Cognitive Autonomous Networks." ITU Kaleidoscope: Machine Learning for a 5G Future (ITU K), 2018.

[7] Vinh Hoa La, Edgardo Montes de Oca, Wissam Mallouli, Ana R. Cavalli, "*A Framework for Security Monitoring of Real IoT Testbeds.*" ICSOFT 2021: 645-652.

[8] Mahmudul Hasan, Md. Milon Islam, Md Ishrak Islam Zarif, M.M.A. Hashem, "*Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches.*" Internet of Things, Volume 7, 2019

[9] P. Wang, X. Chen, F. Ye and Z. Sun, "*A Survey of Techniques for Mobile Service Encrypted Traffic Classification Using Deep Learning.*" IEEE Access, vol. 7, pp. 54024-54033, 2019.

[10] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. "*Explaining and harnessing adversarial examples.*" arXiv preprint arXiv:1412.6572, 2014.