DEA: Anomaly Detection in Smart Environments using Artificial Intelligence

Diego Moreira, Humberto Marques,
Joaquim Celestino Jr. and Rafael L. Gomes
State University of Ceara (UECE)
Fortaleza – CE, Brazil
Email: {diego.moreira,humberto.marques,
celestino,rafaellgom}@larces.uece.br

Aldri Santos and Michele Nogueira
Federal University of Parana (UFPR)
Curitiba – PR, Brazil
Email: {aldri,michele}@inf.ufpr.br

***Abstract*—Nowadays, new paradigms have been deployed, such as smart environments. A smart environment is composed by several heterogeneous Internet of Things devices that communicate with each other and with the Internet. The IoT devices are subject to anomalous behavior, due to security vulnerabilities or malfunction. Monitoring the behavior of these devices is a crucial task to guarantee a efficient performance of the network infrastructure. Within this context, this paper presents the DEA project, which aims to develop a system based on AI to monitor network traffic and to detect anomalies in smart environments, generating a profile of the network and detecting out of order behaviors (different from the behavior pattern).**

***Index Terms*—Smart Environments, Anomaly Detection, Artificial Intelligence.**

I. Introduction

The current structure of the human society claims for environments more intelligent for tasks like infrastructure management, resources and services for the end users. This idea of smart environments can be applied in several contexts, such as Smart Cities, Smart Buildings, Smart Homes, Industry 4.0, E-Health, etc [1]. Each of these contexts have specific services that aim to improve the quality of the life of the users and/or to evolve the execution of daily tasks (automation, equipment checking, patient monitoring and others).

Smart environment is composed of several Internet of Things (IoT) devices (such as sensors and actuators) and mobile devices of end users (notebooks, smartphones, tablets, etc) [2]. The presence of these IoT devices of multiple origins (heterogeneous) in the environments is incrementally increasing during the last few years, bringing new challenges related to the management of the smart environments.

One of the existing challenges in smart environments is related to the monitoring and security of the network, due to existing security vulnerabilities of IoT devices. This security vulnerabilities, as well as possible malfunction, cause anomalies in the network [3]. In the DEA project, anomaly is defined as a non standard behavior of network traffic in the infrastructure [4].

Network anomalies directly affect the Quality of Service (QoS) of the applications running in the top of network infrastructure. As an example of this scenario, several cyberattacks performed in the Internet occurred through the infection of IoT devices [5]. Therefore, it is necessary a solution to identify anomalies in smart environments. The application of proactive policies for early anomalies detection a crucial aspect to guarantee reliability and security for the services, as well as to avoid wastage of resources and financial losses.

Within this context, this paper presents the DEA project, which aims to develop a system to monitor the network traffic and to detect anomalies in smart environments. The proposed system is based on Artificial Intelligence (AI), since it will generate a network profile and will detect possible anomalies through the identification of main features and distinct behaviors from the standard. The DEA project is part of a partnership between Microsoft and Brazilian National Research and Educational Network (RNP-Brazil), allowing the development of the proposed system using the Microsoft Azure Platform.

The remainder of this paper is organized as follows. Section II details the issues related to smart environments and it analyzes several existing related work. Section III describes DEA project and its goals, while Section IV presents the results of the preliminary experiments performed. Finally, Section V concludes the paper.

II. Background and Related Work

Nowadays new requirements for Internet services arose, together with the deployment of smart environments applying the principles of Internet of Things (IoT), such as Smart Homes (SH), Smart Cities (SC), Smart Buildings (SB), Industry 4.0, E-Health systems, among others[6]. These smart environments aim to evolve the daily services, through IoT devices that connect to the Internet and communicate with each other using wireless technologies [7].

Smart environments are characterized by several devices, where each kind of devices has their particularities: personal devices to communicate with the Internet, automation devices to perform actions inside the environment, monitoring devices to check the state of environment and others. All those devices are interconnected by a wireless network. However, smart environments are complex to manager due to the growing of the number of devices and the huge amount of protocols utilized by these devices [8].

978-3-903176-23-2 © 2019 IFIP

Within this context, next, this section will present several related works, focusing on anomalies detection by machine learning techniques in IoT and traditional networks. Machine learning is widely used to recognize patterns, forming different profiles for each characteristic. These proposals vary from traditional TCP/IP networks to IoT networks, classifying attacks or attackers.

Hamamoto et al. [4] developed a combined scheme of genetic algorithms and Fuzzy logic for anomalies detection. In this combined scheme, the genetic algorithm is responsible for tracing a digital signature of the network, and the Fuzzy logic is used to determine when the network is on the attack from the signatures generated previously. The proposal in reference [4] demonstrated good performance when the Fuzzy logic is adjusted according to an anomaly detection database with features of traditional networks, which was populated previously. In this way, this proposal is exclusive for traditional networks, compromising its application over smart environments.

Pajouh et al. [9] presented a machine learning based anomaly classification model, focusing on R2L (Remote-to-Local) and U2R (User-to-Root) attacks detection, that steals a user's credentials, obtaining access to the victim's machine. In reference [9], Naive Bayes and KNN techniques were tested, where the training of these machine learning methods and the classification was performed in the backbone of a IoT network. However, this proposal needs to executed under the IoT device and the processing requirements of the IoT devices were not considered.

Doshi et al. [1] designed a pipelined based middlebox for IoT networks, which captures network packets to identify botnet devices and DDoS attacks. The middlebox uses different machine learning methods (KNN, Random Forests, Decision Trees, SVM, and Neural Networks), trained based on a database about the data shared by the devices, to classify the network packets. Nevertheless, the applicability of this proposal is restricted to well-know stationary networks, since it has no dynamic adjustment capacity.

Hodo et al. [10] proposed a multi-level perceptron neural network to perform the detection of DoS and DDoS attacks in IoT networks, following the KDD99 database. However, this proposal is limited to DoS and DDoS. Additionally, for the evaluation of the proposal, the authors used only five safe nodes content production and one attacker nodes, limiting the experiments and the analysis of the proposal in a realistic scenario.

Diro et al. [3] presented an attack detection system based on deep learning machines and a cloud structure, where the collected information is received and processed on a master node. This proposal checks possible anomalies in the network, reporting the observations to an edge node to perform a global update and re-propagation.

To evaluate the proposal, the NSL-KDD database was used to obtain excellent results with deep learning models. This work, despite using an edge computing architecture, performs all the training in a centralized way, the characteristics of different networks are not represented, and not using interactive learning.

Another architecture proposal for anomaly detection through Deep Leaning is presented in [5]. To analyze the network, the architecture is presented where both the messages exchanged between the IoT devices and the Gateway as well as the messages between Gateway and Cloud are monitored. All captured packets are sent to an analyzer that is responsible for training a Deep neural network that will classify the new packets later.

The system is able to identify the moment when the attack is started, in response to the tests performed, but does not perform a continuous study, presenting the results of this classification. All training and sorting are performed on a device connected to the network, which is not possible in all scenarios.

Deep Autoencoders was used in [11] to solve the problem of botnet detection in IoT. These types of autoencoders networks are able to create unsupervised patterns. Twenty-three characteristics were used for training the method. To perform network sniffing, a port mirroring was used and for the simulation of a Botnet infection, 9 IoT devices with different functions were connected.

The method is efficient in the execution time and classification of Botnets, but a feasibility study in IoT devices was not performed. Only one type of attack has been studied and classified.

III. DEA Project

In general, smart environments have a knowledge about the presence IoT devices, that usually are heterogeneous. Each device profile follows certain functionalities and, consequently, singular network behavior for a specific class/type of device. For example, temperature sensors perform periodical transmissions for a server, updating a set of data. Similarly, video cameras for safety monitoring constant transmit captured images. On the other hand, mobile devices of end users follow a less predictable behavior: in a given moment the user is just send text messages (small volume of traffic generation) and in another moment this same user is watching a on demand video in high definition (very high volume of traffic generation).

In the described approach, the information of the network traffic is collect, monitored and the behavior profile is dynamically generated based on features like amount of active flows, packet size, source/destination address (Layer 2 and 3), previous known characteristics device (when available), traffic volume (devices individually and network as a whole), among others. Additionally, the proposed system will consider the behavior of the devices that are part of the same class/type. From the profile definition, it is possible to detect unknown anomalies, as well as to adapt to new features aggregated to the smart environments. Figure 1 illustrates the application of the proposed system.

Anomalies are detected when it is identified significant changes (out of the pattern) in the network profile, considering individual devices, in a group of device with similar characteristics and the network as whole. As an example, in

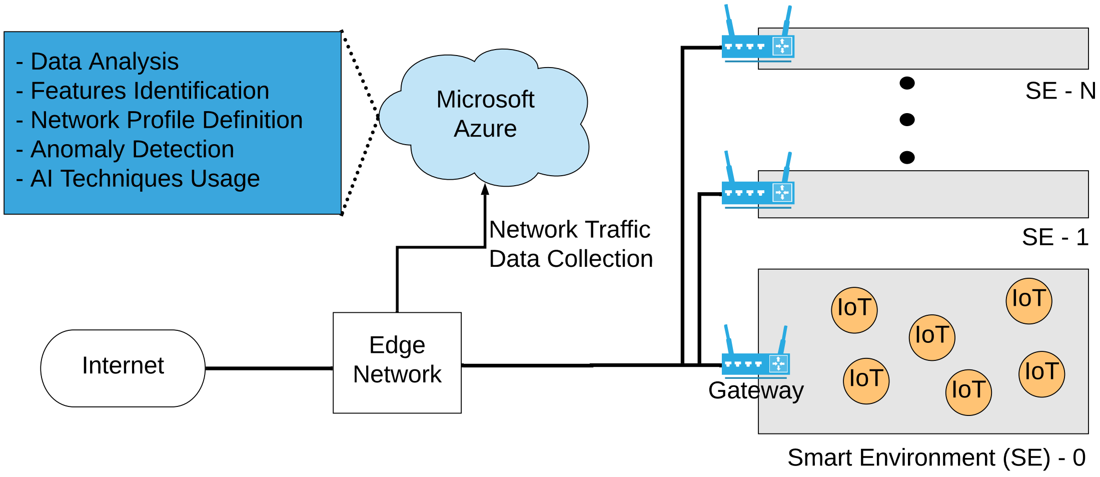

Fig. 1. DEA Scenario.

a situation where security video cameras change drastically they behavior, increasing the size of the transmitted packets and the transmission rate to an specific destination in the network or the Internet. Probably, this situation represents an anomaly related to a Distributed Denial-of-Service (DDoS) attack. Thus, the usage of the proposed system will enable the detection of these kind of situation, allowing the prevention and mitigation of cyberattacks.

In DEA Project AI techniques will be applied to dynamically define the profile of the network (i.e., the usual behavior), as well as the detection of anomalies, where the Microsoft Azure platform will be the basis. The Microsoft Azure platform allows the creation, testing and deployment of Machine Learning solutions from models generated and upgraded automatically. Therefore, it is flexible and customizable, since the supporting and integration of open-source technologies, as well as the adaptation of features of existing machine learning approaches

In this way, it is possible to insert features representing the characteristics of the devices in the smart environment in the Microsoft Azure platform to generate the profile of network behavior, using several AI techniques such as clusterization, neural networks, support vector machines, and others. Additionally, from the techniques cited, it is possible to perform combination: supervised and not supervised, cascade of techniques, etc. Regarding scalability issues, the Microsoft Azure platform is integrated with cloud computing, enabling the capacity to analyze the collected data and to detect anomalies in the network addressing requirements of computation time and resources limit.

IV. Preliminary Experiments and Perspectives

In order to evaluate the developed anomaly detection in the DEA project, several experiments will be performed in the existing network infrastructure of the Laboratory of Computer Networks and Security (LARCES) of State University of CearÃ¡ (UECE). The experiments will have datasets of network traffic as input. These datasets, with and without network attacks, collected in a test environment of the Federal University of Parana (UFPR).

Thus, several initial experiments were performed to evaluate the anomaly detection capacity of the existing AI techniques, as well as to improve the current studies related to the Microsoft Azure platform (mainly the Machine Learning Studio module). These initial experiments used the NSL-KDD database with information about real network traffic, including parameters about the packets, flows and devices.

First, considering the features of the network traffic in the database, a analysis of the parameters to identify the most suitable features (i.e., the information available in the context of the DEA project and information that has high potential to characterize the network traffic profile). As result of this analysis, the following features were used in the AI training process: network service used (HTTP, FTP, SMTP, Telnet, etc), connection flag, data volume in a connection, number of connections for the same destination (IP and/or port) and percentage of connections for the same network service (and its variation too).

From the features identification process, the following AI techniques were applied in the experiments: Recurrent Neural Network (RNN), Decision Tree, Random Forests, K-Nearest Neighbors (KNN). The preliminary results of accuracy of the evaluated techniques are presented in Figure 2.

Based on the data presented in Figure 2, it is possible to note the viability of the proposed system in DEA project, since the AI techniques are capable to identify anomalies in the network infrastructure. In the same way, Each AI technique has a specific accuracy, enhancing the necessity to research possible approaches to be used to anomaly detection of smart environments. The final result of DEA project is an context-aware adaptive anomaly detection system based on AI. Additionally, the developed system can be deployed under existing network monitoring services, such as Ipe network of RNP in Brazil.

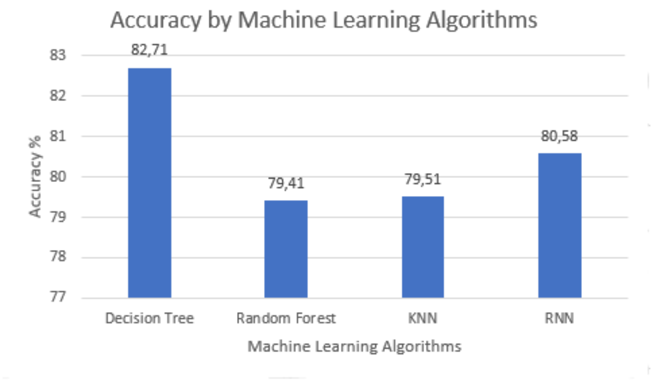

Fig. 2. Results.

V. Conclusion

Smart environments are composed of IoT (sensors, actuators and others) and end users devices (notebooks, smartphones, tablets, etc) that communicate with each other and with the Internet. The IoT devices are subject to anomalous behaviors (out of order execution), due to security vulnerabilities or malfunction. Monitoring the behavior of these devices is a crucial task to guarantee a efficient performance of the network infrastructure. Within this context, the DEA project aims to develop a system based on AI to monitor network traffic and to detect anomalies in smart environments, generating a profile of the network and detecting out of order behaviors (different from the behavior pattern).

Acknowledgment

The authors would like to thank RNP and Microsoft for their financial support.

References

[1] R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning ddos detection for consumer internet of things devices," in *2018 IEEE Security and Privacy Workshops (SPW)*. IEEE, 2018, pp. 29–35.

[2] H. Li, K. Ota, and M. Dong, "Learning iot in edge: deep learning for the internet of things with edge computing," *IEEE Network*, vol. 32, no. 1, pp. 96–101, 2018.

[3] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for internet of things," *Future Generation Computer Systems*, vol. 82, pp. 761–768, 2018.

[4] A. H. Hamamoto, L. F. Carvalho, L. D. H. Sampaio, T. Abrão, and M. L. Proença Jr, "Network anomaly detection system using genetic algorithm and fuzzy logic," *Expert Systems with Applications*, vol. 92, pp. 390–402, 2018.

[5] O. Brun, Y. Yin, J. Augusto-Gonzalez, M. Ramos, and E. Gelenbe, "Iot attack detection with deep learning," in *ISCIS Security Workshop*, 2018.

[6] B. L. R. Stojkoska and K. V. Trivodaliev, "A review of internet of things for smart home: Challenges and solutions," *Journal of Cleaner Production*, vol. 140, pp. 1454–1464, 2017.

[7] S. Feng, P. Setoodeh, and S. Haykin, "Smart home: Cognitive interactive people-centric internet of things," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 34–39, February 2017.

[8] M. A. Chaqfeh and N. Mohamed, "Challenges in middleware solutions for the internet of things," in *2012 International Conference on Collaboration Technologies and Systems (CTS)*, May 2012, pp. 21–26.

[9] H. H. Pajouh, R. Javidan, R. Khayami, D. Ali, and K.-K. R. Choo, "A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in iot backbone networks," *IEEE Transactions on Emerging Topics in Computing*, 2016.

[10] E. Hodo, X. Bellekens, A. Hamilton, P.-L. Dubouilh, E. Iorkyase, C. Tachtatzis, and R. Atkinson, "Threat analysis of iot networks using artificial neural network intrusion detection system," in *2016 International Symposium on Networks, Computers and Communications (ISNCC)*. IEEE, 2016, pp. 1–6.

[11] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, A. Shabtai, D. Breitenbacher, and Y. Elovici, "N-baiotânetwork-based detection of iot botnet attacks using deep autoencoders," *IEEE Pervasive Computing*, vol. 17, no. 3, pp. 12–22, 2018.