engineering proceedings

MDPI

Proceeding Paper

Low Cost Automated Security Audit System †

Pedro Fernández-Arruti *, Julio J. Estévez-Pereira *, Francisco J. Nóvoa, Jose C. Dafonte and Diego Fernández

Research Center of Information and Communication Technologies (CITIC), Campus de Elviña, s/n, 15071 A Coruña, Spain; francisco.javier.novoa@udc.es (F.J.N.); carlos.dafonte@udc.es (J.C.D.); diego.fernandez@udc.es (D.F.)

* Correspondence: pedro.fernandez-arruti@udc.es (P.F.-A.); julio.jairo.estevez.pereira@udc.es (J.J.E.-P.)

† Presented at the 4th XoveTIC Conference, A Coruña, Spain, 7–8 October 2021.

Abstract: In recent years, a quick transition towards digitization has been observed in most organizations. Along with it, certain inherent problems have appeared, such as the increase in cyber threats. Large organizations are able to adapt easily, but this does not happen with small and medium-sized companies. Currently, there are very few solutions aimed at fulfilling the needs of these small enterprises, so we have worked on a tool for them. Our tool is capable of displaying key, easy-to-interpret information related to each organization's network assets. To achieve this, we used passive and active analysis techniques and successfully evaluated the viability of using machine learning techniques to get more meaningful information. All of the information obtained is displayed in a simple web application, which is designed to be used by managers in organizations without them needing to handle complex concepts and vocabulary.

Keywords: network audit; passive analysis; active analysis; machine learning

Citation: Fernández-Arruti, P.; Estévez-Pereira, J.J.; Nóvoa, F.J.; Dafonte J.C.; Fernández, D. Low Cost Automated Security Audit System. *Eng. Proc.* **2021**, *7*, 58. https://doi.org/10.3390/engproc2021007058

Academic Editors: Joaquim de Moura, Marco A. González, Javier Pereira and Manuel G. Penedo

Published: 28 October 2021

1. Introduction

Organizations of all sizes are now significantly reliant upon information technology and networks for the operation of their business activities. Therefore, they have the added requirement of ensuring that their systems and data are appropriately protected against security breaches. However, there is evidence to suggest that security practices are not strongly upheld within small and medium-sized enterprise (SME) environments [1].

There are different approaches in the literature that attempt to address this problem. However, many of them require those responsible for organizations to handle complex concepts and vocabulary and provide results that managers of this type of organization do not know how to interpret.

Our project involves building a modular tool that implements the creation of an inventory of the organization's assets (final and intermediate devices, active services, and identification of application-layer protocols) and an information visualization through a dashboard (providing key information to the organization's managers, indicating the technical risk of the organization). In addition, we evaluate the viability of machine learning techniques for offering advanced knowledge of the state of the network from the data collected by using unsupervised exploration techniques. There are non-functional characteristics that are key to the success of our tool: a low-cost, scalable, modular, and easy-to-use solution.

2. State of the Art

Nowadays, there are many solutions for carrying out network audits. Many of them produce satisfactory results, allowing their customers to improve the security of their networks. However, only a reduced number of them are both low-priced and easy to use. It is for this reason that small and medium-sized businesses cannot afford a secure network infrastructure.

An example of a currently available wired network audit tool is the Raspberry Pwn [2]. This tool is an open-source software created by the company Pwnie Express, which is aimed at detecting vulnerabilities in a network using a Raspberry. The disadvantage of this tool is its maintenance, since it has not been revised since 2014. On the other hand, there is a project called Wireless Attack Toolkit (WAT) [3] that allows one to convert a Raspberry Pi into a security auditing system for different types of networks. Its main disadvantage is the same as in the previous case, as the last revision of its code was in 2016.

3. Materials and Methods

3.1. Architectural Design

This project is about designing and building a scalable, low-cost, and easy-to-use system that performs audits on corporate networks with minimal intervention by the end user. For this, two types of analysis are performed: a passive analysis, which consists in a device that passively listens to the network traffic and makes an inventory of the active devices on the network, as well as the protocols that make some kind of broadcast communication on it; and an active analysis, which detects each asset's operating system, hostname, IP, and the status of its ports and services.

Moreover, we not only show the retrieved data, but also process them in order to generate a dataset on which unsupervised machine learning techniques can be applied. Concretely, we use data involving services and protocols to train a self-organizing map (SOM) [4] that clusterizes our samples, providing an easy-to-understand and very visual way to distinguish the devices in our network, which are atypical when it comes to the services running or protocols used.

To build a system that implements the desired functionalities and meets the non-functional requirements of low cost, scalability, and ease of use, we propose the design that can be seen in Figure 1.

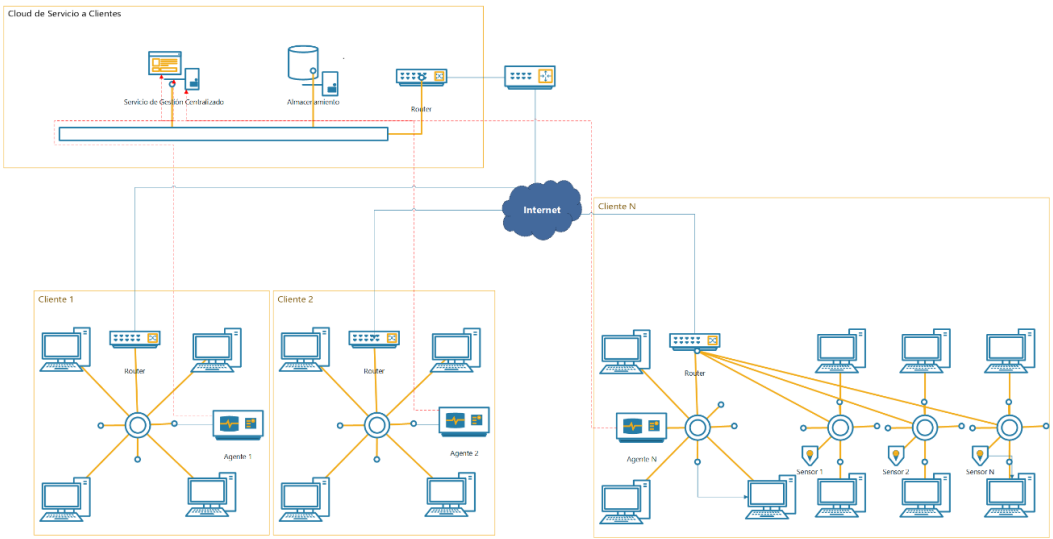

Figure 1. System architecture.

The system relies on three different types of elements. First of all, we have a hardware agent connected to the client network. If the network has multiple subnets, sensors will be placed to collect information from each segment. These sensors will send the added information to the agent of its organization. Finally, a server is used to store the data related to the operation of the network.

3.2. Software and Hardware

To achieve the objectives discussed above, we employ hardware devices based on a low-cost board where we run our application code. This code is written in Python 3 and

uses well-known network tool libraries, such as Scapy and Nmap. In addition, the Django framework is utilized to create the web interface. Finally, for performing the machine learning tasks, we relied mainly on Numpy, Pandas, and MiniSom.

4. Results

As a result of the implementation of the tool, a network scanning software was obtained and is executed periodically every hour. The information collected is reflected in a web interface.

The data obtained through passive analysis are the following: IPv4 addresses, IPv6 addresses, device network adapter manufacturer data (OUI), device name, the last time it was detected, domain names (through LLMNR and MDNS), operating system (through the TTL value of the packets), and the broadcast protocols used by each host.

On the other hand, the data obtained for each of the hosts detected by the active analysis are the following: IPv4 addresses, IPv6 addresses, computer name, possible operating systems that run on it, and open ports and services that run on said ports.

Finally, concerning the use of SOMs to clusterize the detected devices according to both their services and protocols, we found that the results are promising. To reach this conclusion, we relied on two metrics: average quantization error and topographic error measurement. In this project, we reached the values 0.30 and 0.15, respectively, for these metrics.

Although being anomalous is not the same as posing a threat, it is interesting for security purposes to discover and analyze devices that are different from others according to the topological distance between the clusters defined.

5. Discussion

After developing the first version of our tool, we came to the conclusion that it is possible to build a low-cost product that performs security audits in networks of small organizations. Our solution provides to SMEs a much-needed cybersecurity solution that is exclusively oriented to them and, therefore, affordable.

When contemplating future work, we plan to use agents as devices that not only perform network audits, but also carry out continuous monitoring. This is intended to perform network anomaly detection on a day-to-day basis by creating normal network profiles against which to compare network traffic at all times. We think that this is a very promising line of work, as good anomaly prevention could translate into effective attack prevention.

Author Contributions: Conceptualization J.C.D. and F.J.N.; methodology, J.C.D. and F.J.N.; software, P.F.-A. and J.J.E.-P.; validation, J.C.D., D.F. and F.J.N.; formal analysis, P.F.-A.; research, P.F.-A. and J.J.E.-P.; resources, P.F.-A. and J.J.E.-P.; data curation, P.F.-A. and J.J.E.-P.; writing—original draft preparation, P.F.-A. and J.J.E.-P.; writing—review and editing, P.F.-A., J.J.E.-P., J.C.D., D.F. and F.J.N.; visualization, P.F.-A. and J.J.E.-P.; supervision, J.C.D., D.F. and F.J.N. All authors have read and agreed to the published version of the manuscript.

Funding: CITIC, as a Research Center accredited by the Galician University System, is funded by "Consellería de Cultura, Educación e Universidade from Xunta de Galicia", supported in an 80% through ERDF, ERDF Operational Programme Galicia 2014–2020, and the remaining 20% by "Secretaría Xeral de Universidades (Grant ED431G 2019/01). This work was also funded by the research consolidation grant ED431B 2021/36, Art.83 collaboration F19/17, the Ministry of Economy and Competitiveness of Spain, and the FEDER funds of the European Union (Project PID2019-111388GB-I00).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kurpjuhn, T. The SME security challenge. *Comput. Fraud. Secur.* **2015**, *3*, 1–3. [CrossRef]
2. Raspberry Pwn. Available online: https://github.com/pwnieexpress/raspberry_pwn (accessed on 29 July 2021).

3. Wireless Attack Toolkit (WAT). Available online: https://sourceforge.net/projects/piwat/ (accessed on 29 July 2021).
4. Del Coso, C.; Fustes, D.; Dafonte, C.; Nóvoa, F.J.; Rodríguez-Pedreira, J.M.; Arcay, B. Mixing numerical and categorical data in a Self-Organizing Map by means of frequency neurons. *Appl. Soft Comput.* **2015**, *36*, 246–254. [CrossRef]