Analysis of computer attacks through Honeypots in the
District University Francisco José de Caldas

Miguel A. Leguizamón-Páez\textsuperscript{1,8}, María A. Bonilla-Díaz\textsuperscript{1}, Camilo A. León-Cuervo\textsuperscript{1}

\textsuperscript{1}Universidad Distrital Francisco José de Caldas, Facultad Tecnológica, Ingeniería en Telemática,
Bogotá, Colombia

\textsuperscript{8}maleguizamonp@correo.udistrital.edu.co, mabonillad@correo.udistrital.edu.co, caleonc@correo.udistrital.edu.co

Recibido: 03 de septiembre 2019 – Aceptado: 01 de enero de 2020

Abstract

This article is the result of the work developed for the design and implementation of Honeypots as a complementary alternative to the existing computer security scheme at Universidad Distrital Francisco José de Caldas as a project that also contributes to the analysis and detection of attacks to the network security and the elements of computer type in the institution.

For the development of this article, we worked using the PDCA cycle (Plan - Do - Check - Act). It is a model based on process management and continuous improvement of them, with a simple application and used properly, can help a lot in the realization of activities, both productive and administrative, in a more organized and effective way. Therefore, adopting the PDCA cycle provides a simple guide for the management of activities and processes, the basic structure of a system and it is applicable to any organization.

After the implementation of Cowrie and HoneyPy, it was possible to identify different patterns and ways to attack, guiding the configuration of a script in the IDS (Intrusion Detection System) server, allowing with the stored logs to create rules and implement it over Iptables. This fact allows become the IDS Server a node into a network of sensors feeding the database globally for an investigation of the attacks on all connected and configured computers, obtaining information to make a major analysis for the final user.

The design of infrastructure with honeypots, implemented at the District University Francisco José de Caldas allows finding security failures belonging to university’s servers due to computing attacks. A new network distribution was designed for registering information about the different attacks and enabling effective solutions to be apply to the university.

Keywords: Cyberattack, Cybersecurity, Honeypots, Vulnerabilities.
Resumen

El presente artículo es resultado del trabajo desarrollado para el diseño e implementación de Honeypots como una alternativa complementaria al esquema de seguridad informática existente en la Universidad Distrital Francisco José de Caldas, proyecto que a su vez contribuye en el análisis y detección de ataques a la seguridad de la red y demás elementos de tipo informático en la institución.

Para el desarrollo de este artículo, trabajamos utilizando el ciclo PDCA (Planificar - Hacer - Verificar - Actuar). Es un modelo basado en la gestión de procesos y su mejora continua, con una aplicación simple y utilizada adecuadamente, puede ayudar mucho en la realización de actividades, tanto productivas como administrativas, de una manera más organizada y efectiva. Por lo tanto, la adopción del ciclo PDCA proporciona una guía simple para la gestión de actividades y procesos, la estructura básica de un sistema y es aplicable a cualquier organización.

Después de la implementación de Cowrie y HoneyPy, fue posible identificar diferentes patrones y formas de ataque, guiando la configuración de un script en el servidor IDS (Intrusion Detection System), permitiendo con los registros almacenados crear reglas e implementarlo en Iptables. Este hecho permite que el servidor IDS se convierta en un nodo en una red de sensores que alimentan la base de datos globalmente para una investigación de los ataques en todas las computadoras conectadas y configuradas, obteniendo información para realizar un análisis complejo para el usuario final.

El diseño de infraestructura con honeypots, implementado en la Universidad del Distrito Francisco José de Caldas, permite encontrar fallas de seguridad pertenecientes a los servidores de la universidad debido a ataques informáticos. Se diseñó una nueva distribución de red para registrar información sobre los diferentes ataques y permitir establecer soluciones efectivas.

Palabras clave: Ataques Informáticos, Honeypots, Seguridad Informática, Vulnerabilidades.

1. Introduction

The way that humans communicate has been changing through time and that is how people get knowledge, making way to impressive discoveries and inventions; in addition, it allowed many of the wars narrated in the history books. Precisely, in the search for knowledge, the values that allow the community to live without conflicts have been transgressed; even today, information is attacked with different purposes and using multiple methods.

Historically, information has represented an invaluable asset, who has the possibility of manipulating any data is possessor of great power. Information systems and information that is managed are the most important resource of any organization; it is for the information how competitive management and project management strategies are chosen.

Therefore, it is necessary to protect all types of information, in order to use it completely and at the same time, it can be protected and kept safe against any possible eventualities involving loss, distortion and/or use by unwanted entities.

This document contains theoretical information about computer security and Honeypots. The Honeypots seek to be compromised; the data collected are valuable because they are the product of intrusions made by entities that want to alter the network for criminal purposes. The Honeypots are designed to identify the methods, the motives and modus operandi of the attacker who arrives at this trap. They are flexible and adaptable tools; their implementation does not demand higher costs, instead, it represents a good of multiple benefits for any organization, identifying unknown vulnerabilities and discovering the risk of affection of their systems.

A Honeypot is a system designed to analyze how hackers use their methods to try to enter a system, analyze vulnerabilities and alter, copy or destroy their data or all of it. By learning their tools and methods, systems can then be better protected. They can consist of different applications, one of them is used to capture the
intruder or learn how they act without them knowing that they are being watched.

Honeypots in their most basic form are fake information servers, strategically positioned in a test network, which are fed with false information that is disguised as files of a confidential nature. In turn, these servers are initially configured in such a way that it is difficult, but not impossible, to be penetrated by a computer attacker, deliberately exposing them and making them highly attractive to a target hacker. Finally, the server is enabled with information monitoring and tracking tools, so that each step and trace of activity of a "hacker" can be recorded in a logbook that indicates these movements in detail.

The main functions of a Honeypot are:

- Divert the attacker's attention from the real network of the system, so that the main information resources are not compromised.
- Capture new viruses or worms for further study.
- Form profiles of attackers and their preferred attack methods, similar to that used by a police corporation to build the file of a criminal based on his modus operandi.
- Know new vulnerabilities and risks of the different operating systems, environments and programs which are not yet properly documented. (1)

Then, has been implemented Honeypots in the network of the Universidad Distrital Francisco José de Caldas, whose objective is support the learning process about computer attacks in a practical way being an input for the interpretation of the investigating entity, allowing identifying the modes of operation, the types of attacks, and others. In such a way the necessary methods can be determined to fight against any attack and create the processes to treat vulnerabilities that appear in the forms of communication that are currently used.

2. Methodology

For the development of this article, we worked using the PDCA cycle (Plan - Do - Check - Act); this is a dynamic cycle that can be used into the processes of any organization, allowing an integral improvement of competitiveness of products and services, continuously improving quality, reducing costs and prices, optimizing productivity, increasing market share and the profitability of the organization (2). It is a model based on process management and continuous improvement of them, with a simple application and used properly, can help a lot in the realization of activities, both productive and administrative, in a more organized and effective way. Therefore, adopting the PDCA cycle provides a simple guide for the management of activities and processes, the basic structure of a system and it is applicable to any organization.

The adoption of the PDCA cycle promotes the practice of management in favor of opportunities for the organization to improve the performance of its processes. Once an area of opportunity has been identified, the change can be planned and carried out. Then the results of the implementation of the change are verified and according to these results, we act to adjust the change or to start the cycle again planning new changes (3). It can be seen in Figure 1.

![Figure 1. PDCA Cycle(3)](image-url)
For each stage the following is done:

- **Planning**: Gathering information is carried out on the topics about computer security, attacks, vulnerabilities, Intrusion Detection System (IDS) and the main topic: Honeypots. Analyzing concept, structure, design, implementation, configuration and installation parameters, in order to establish the processes related to this job.

- **Do**: The prototype of the Honeypot comes into operation. The Honeypot is joined with the network infrastructure already existing in the university, afterwards the necessary elements have to be configured, logs are implemented and that will supply all the information for the analysis of the captured data.

- **Check**: At this point, the functional tests are done, and it is verified that all processes are fulfilling their purpose.

- **Act**: It is decided to give a solution to the nonconformities found in the “check” step, any corrective action is carried out and the improvement plans are established.

After this, the results of the implementation of the planned changes are verified and according to these, we act to adjust our course or to start the cycle again by planning new changes.

### 2.1. Analysis and Implementation

Currently, the Campus of the District University is made up of twenty-four branches across Bogotá city, some owned, others in loan and lease contract (4). We can find these branches throughout different locations. The support and laying of the network are implemented under a star topology. The department “Red de Datos UDNET” manages this network.

Below the diagrams for the data link and the network topology that connects the different university branches (5), can be seen in Figure 2.

The Figure 2 above shows that the university uses broadband links ADSL (Asymmetric Digital Subscriber Line) to the building that allocates PIGA (in Spanish: Plan Institucional...
de Gestión Ambiental) and a radio-link to the UGI (in Spanish: Unión General de Inversiones) building and besides, there is a radio link that provides connectivity to the building on the branch “Calle 64” as shown in the Figure 3.

The Engineering Department is the central node of the District University located in the Teusaquillo’s neighborhood; from here, the connectivity is given to the different branches shown in the both previous figures (Figure 2 and 3). This connection is established by MPLS (Multiprotocol Label Switching) and the Internet access is centralized in the Engineering department as well using a dedicated channel at 2 Gbps (Gigabit per second). The connection between the communication rooms is given by fiber optic, mostly to 10G according to the service provider (ETB, in Spanish: Empresa de Telecomunicaciones de Bogotá). There are known five Data Centers situated in different branches, in the Aduanilla de Paiba, in Macarena A, in Technology department, in Bosa el Porvenir and the main one in the Management Center, Olimpo. The Figure 4 allows identifying the data channels, their characteristics and the established topology.

Otherwise, one of the different services that the university uses is referring to an important resource that connect directly all the users linked to the institution, this is the e-mail; this resource represents one of the main methods for the execution of attacks talking about computer security, called Phishing.

Additional to the applications illustrated in a Figure 5, the which shows that the e-mails is the most used for the university community, is necessary to stand out that for this application the 99,95% don’t have two-step verification enabled, opening the possibility of intrusion about your information (6).

Together with the “Red de Datos UDNET” and its monthly reports that allow us to identify the attacks to the network and it is supported by Check Point Software Technologies and Kaspersky Lab. It was possible to get information from September 2017; this shows events under a critical risk classification, such SIPVicious Security Scanner, which is a product that searches vulnerabilities into the servers with

![Figure 3. Data Link connections in near university branches](image)

5 / 13
25,968 replicas; another resource, PHP Web Shell Generic Backdoor used by attackers for executing an arbitrary code, or use the server as a bot (apheresis of robot), which is an application that executes an automated task, for new attacks, this has 3443 replicas. It relevant to say there are 134 infected devices with 3876 different files that mostly contain malware of the Trojan type and worms (5).

In Latin America, statistics such as those published by the international cybersecurity company Kaspersky Lab in September 2016, indicated at least every second twelve malicious malware attacks are presented. The average keeps the region far behind Asia and Africa and even some European countries in terms of malware, but according to Kaspersky in Brazil almost half of the computers analyzed was threatened (49.9%), meanwhile in Peru, Bolivia, Chile, Mexico and Colombia, like show in Table 1, those affected were approximately 4 out of 10.

Table 1. Statistical analysis of computer attacks in Latin America (6)

<table>
<thead>
<tr>
<th>Country</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>49.9</td>
</tr>
<tr>
<td>Peru</td>
<td>41.9</td>
</tr>
<tr>
<td>Bolivia</td>
<td>41.8</td>
</tr>
<tr>
<td>Chile</td>
<td>40.0</td>
</tr>
<tr>
<td>Mexico</td>
<td>39.9</td>
</tr>
<tr>
<td>Colombia</td>
<td>39.3</td>
</tr>
<tr>
<td>Guatemala</td>
<td>37.5</td>
</tr>
<tr>
<td>Ecuador</td>
<td>36.1</td>
</tr>
</tbody>
</table>
In total, using as main source service cloud-based Kaspersky Security Network, recorded more than 398 million attacks of those type of programs between August 2015 and August 2016. (6)

With the above, the proposal of this article aims to achieve the objectives of the computing security, in order to protect the information administered in the different University’s offices against any attack. In the main objectives, stand out:

To protect the resources of computing systems, but making protect information a critical fact, including equipment, infrastructure, use of applications, and others.

- To guarantee the appropriate use of the resources and applications of the system.
- To restrict the losses and achieve the correct system recovery after a security incident.
- To respect the legal framework and the requirements expressed in the contracts. (7)

2.2. Infrastructure and Implementation design

The word telematics was use for the first time in France, in the 70's, and is the result of the fusion of the terms: telecommunications and computing. Telecommunications is the area that transports information between places in a transparent way for the final user, and computing is the area that manages the information keeping it available with the content it requires for this user. (8)

According to this definition, is important to understand, even in transport and management of information, attacks can come true affecting any of the pillars of computing security, this could be involve an economic outlay to cover expenses of repair and remediation of the information systems and affected infrastructures, which can cause the ruin of people and companies.

This project seeks an alternative that provides support information for administrative decision-making and educational ways, being efficient to identify types and methods of attackers, making easier to deal with the challenges and techniques of computer intruders. Reto Baumann proposes a Honeypot as "a resource that aims to be a real target. A Honeypot is expected to be attacked or compromised. Its main goal is to distract the attackers and obtain information about them and their attacks". (9)

An infrastructure with Honeypots was projected to be implemented in the Universidad Distrital Francisco José de Caldas, to find security flaws into the university's servers because of computing attacks. The design was based on a distribution with honeypots in the network making possible to record the different computing attacks, in order to collect information about them and after their analysis, establish patterns of behavior of the attackers and the methods they use; looking for the implementation of effective solutions proposed from the research and innovation in the academic practice at the University.

Concerning the disposition of the elements into the network infrastructure for the implementation of the Honeypots, it is necessary to highlight these tools need minimal resources, do not require complex architectures and the use of multiple equipment, neither modify the physical infrastructure to use them.

Below, in Figure 6, is shown the distribution of the network in the Management Center Olimpo, the Data Center that is the central node, before the implementation of the honeypots.

The Management Center Olimpo houses communication’s equipment owned by the university and leased from the Internet Service Provider (ISP), together with the server

<table>
<thead>
<tr>
<th>Country</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Venezuela</td>
<td>36.0</td>
</tr>
<tr>
<td>Uruguay</td>
<td>30.0</td>
</tr>
<tr>
<td>Argentina</td>
<td>29.5</td>
</tr>
</tbody>
</table>
equipment centralized. Providing different network services such as the Institutional Portal Web (IPW), e-mails, File Transfer Protocol (FTP), mailing lists, domain, DNS (Domain Name System), DHCP (Dynamic Host Configuration Protocol), Firewall, proxy, virtual classrooms, also, hosting equipment of research groups, work groups and areas like accounting and computing. (10)

In the implemented design, we used Honeypots classified by their implementation environment for research, because their purpose is to serve as an educational resource for the study of different threats and attack patterns. Concerning its definition by the level of interaction, the use of low-interaction Honeypots was determined in order that its use will be simple in the practical computing security exercises in which common services are emulated. Two honeypots were placed, before and after the proxy to define the relevant information about the access attempts to the network by SSH (Secure Shell) and Telnet connections, to document the attacks carried out by IPs (Internet Protocol directions) belonging to the internal network.

Theoretically, the honeypots are located before and after the Firewall and in the DMZ (demilitarized zone), according to their location the previous configurations for their implementation are defined (1). Referring to the distribution of the network in the university, the firewall works together with the proxy, enabling the common functions referring to authorization, blocking and redirection of connections and/or ports. The attacks were registered with the use of Cowrie and HoneyPy.

Cowrie is a medium interaction honeypot designed to record brute force attacks and the shell interaction performed by the attacker (11). HoneyPy, is a low-to-medium interaction honeypot, written using Python, created to be adaptive, making possible to add new emulations of services (add-ons) to protocols based on TCP (Transmission Control Protocol) or UDP (User Datagram Protocol). Plugins can be created to emulate UDP or TCP based services to provide more interaction. All activity is logged to a file by default (12).

Cowrie is a useful tool to define the relevant information about the access attempts to the network by SSH and Telnet connections, there are a lot of information on the Web allowing it to use and configuration, in addition there are regular updates because it is in a continue development. HoneyPy using HoneyDB, available into the other equipment connected in the network being an input to identify global attacks in research honeypots, collecting data from HoneyPy sensors based on the number of events, connections, data received, data transferred, feeding the HoneyDB registers.

Finally, the complete scheme can be seeing in the Figure 7, showing the implementation of the Honeypots before and after the proxy, trying to give a complete analysis that help to determine the internal and external attacks in the university network.

The illustration below shows one Honeypot before the proxy, aiming to collect information about connection attempts by unused IPs. The first Honeypot use Cowrie to define the relevant information about the access attempts to the network by SSH and Telnet connections. The
Honeypot after the proxy allows collecting data about the attacks carried out by internal IPs in the network.

**Figure 7. Final distribution with the Honeypots in Olimpo**

About the configuration, for example, in the first Honeypot that use Cowrie, has been use the shell commands in Figure 8:

```plaintext
# ./bin/cowrie-cowrie
# virtualenv cowrie
# pip install python-

# Installing scripts on python-shell...
# python shell
# print

# ./bin/cowrie-cowrie
# virtualenv cowrie
# pip install python-

# Installing scripts on python-shell...
# python shell
# print

# cowrie-logviewer.py
# cowrie-direct-tpip.request, cowrie-direct-tpip.data

# cowrie-logviewer.py
# cowrie-direct-tpip.request, cowrie-direct-tpip.data

# Cowrie logviewer, Figure 9, we can view the attacks classified by days bringing information about the events captured, showing the origin IP making connection with the port, and a summary of the activities carried out in the connection.

**Figure 8. Shell commands for the Cowrie configuration**

For the edit `cowrie-logviewer.py` for the script configuration. There are several variables at the top of the file that can change:

- "log_path" – the path to the cowrie's log directory
- "dl_path" – the path to the cowrie dl directory (downloads)
- "maxmind_path" - the path of the MaxMind GeoLite 2 database. The default value is "maxmind/GeoLite2.Country.mmdb"
- "bind_host" – the IP address to which the web server should bind, by default 0.0.0.0
- "bind_port" – the port that the web server should listen, by default 5000
- "min_upload_size" – The minimum file in bytes should appear on the “Files Uploaded” page. The default value is 1024
- "debug" – if you want to debug messages, set this to True. Default False
- "use_gzip" – if you don’t want gzip compression, set it to False. Default True.
- "filter_events" – List of log events to filter. Default "["cowrie.direct-tpip.request', 'cowrie.direct-tpip.data']"

**3. Results and discussion**

In the Cowrie logviewer, Figure 9, we can view the attacks classified by days bringing information about the events captured, showing the origin IP making connection with the port, and a summary of the activities carried out in the connection.

The logviewer, Figure 10, can show statistics allowing identify which country do the recurrent attacks, establishing a pattern limiting the access of IPs pool depending the country, knowing credentials used by the attacker to try enter into the server and with possible not safe enough passwords used by services administrators, to protect the server.

Establishing the attack patterns is not the only action to do, also is important to know the vulnerabilities in the university's servers, therefore, the honeypots implemented capture the malicious files that tried to take advantage of a security fail, allowing to see how the malicious scripts are developed and protect radically the servers.

With the connection HoneyPy to HoneyDB we can obtain the main IP addresses making connections with a specific port, obtaining two graphics classifying the activity by protocol and services. It can be seen in the Figure 11.
There are different websites where we can find data related to the selected IP address. These sites are:

- **Cymon** – It is the biggest open tracker of malware, phishing, botnets, spam and other threats.
- **DSheild** – Provides “who-is” information (Who is it?) and show if the IP appears in any of the Internet Storm Center threat lists.
- **OTX** – Threat data platform.
- **Twitter** – Tweets mentioning the IP address some time.
- **Google** – Search results containing the IP address.
- Virus Total – Provides a report if a malware or malicious domain are related with the IP address.
- Spamhaus – Provides a report about if the IP address is in the Spamhaus blocking list.
- SpamCop – Provides a report if the IP address be in the spam email block-list.
- Senderbase – Provides the reputation about email and web for the IP address.

There are other searching tools like:
- Session – This tool shows the activity and data of events captured by the honeypots, in the information compiled we will find the details of what the attacking host was trying to do.
- Shodan - Displays banner information (metadata sends back client by the server) collected by that page.
- Project Honeypot – If the IP address is in the project Honeypot database.
- Threat crowd – Threat search engine.

After the implementation of Cowrie and HoneyPy, it was possible to identify different patterns and ways to attack, guiding the configuration of a script in the IDS (Intrusion Detection System) server, allowing with the stored logs to create rules and implement it in Iptables, because the IPs was identified like the highest connotation data associated with the different attacks. This fact allows become the IDS Server a node into a network of sensors feeding the database globally for an investigation of the attacks on all connected and configured computers, obtaining information to make a complex analysis for the final user.

4. Conclusions

The amount of threats and vulnerabilities present in the environment that make the information highly susceptible, must be mitigated and limited in order to eradicate and eliminate any leakage or mishandling of information. Therefore, it is necessary to identify all the flaws in the information systems and make the best effort to protect the information against any unwanted entity with possible criminal purposes, such as those who have been detected in previous times.
by left evidence of bad use.

The design of infrastructure with honeypots, implemented at the Universidad Distrital Francisco José de Caldas allows finding security failures belonging to university's servers due to computing attacks. A new network distribution was designed for registering information about the different attacks and enabling effective solutions to be set, proposed in the research and innovation of the academic practice in the university.

Concerning the system adaptability, the highlight advantages of the Honeypots are their minimum requirements for the implementation, which ensures that more tools of this type can be added, either by physical equipment or use of virtual machines, being able to differentiate the configurations of each Honeypot.

However, it is necessary to take into account that honeypots are passive tools, that is, if they do not receive an attack, they will have no use for the purpose of this work, and these are not tools that are corrective in the face of an attack, their proactive functions, not reactive, its purpose is oriented solely to the analysis of computer attacks, as a basis for subsequent decision-making.

In an industrial and / or commercial environment that demands extreme security, it is necessary that the analysis obtained from an attack in real time can be countered in the shortest possible time, preventing the attacker from perceiving that he is in a trap and proceed to the execution of harmful alternatives to the primary network of the organization.

Finally, this project can be implemented in any productive area related to the university network management by Red de Datos UDNET bringing an important approach for research groups related. Contributing to recognized projects such as HoneyProject, providing results of the attack types that reach the university network and even in the development of more honeypots, allowing identify a extend diversity in the patterns and in the application of solutions for decision making these attacks.

5. References


