Cybersecurity in TIC-Scientist Network Infrastructures by Honeypots: Catching Cyber Threat Passively

Juan Luis Martin Acal, Pedro A. Castillo Valdivieso, Gustavo Romero López, and Juan Julián Merelo Guervós

Springer-Verlag, Computer Science Editorial III, Postfach 10 52 80,
69042 Heidelberg, Germany
jlmacal@correo.ugr.es
{pcv, gustavo}@ugr.es
jjmerelo@gmail.com
http://www.springer.de/comp/lncs/index.html

Abstract. It should be a balance between security concerns and the right to privacy in the use of IT infrastructures. Scientist networks have a high risk of attack as a source of valuable information and resources because they were designed with an open and decentralized philosophy, in favour of the transmission of knowledge when security wasn't a critical section[3]. Private and scientific information have a enormous value for an attacker but the end user is worried about his privacy too. For this reason passive detection methods in cybersecurity like honeypots are the cornerstone in its defence plan. We expose the practical case of the University of Granada in the application of honeypots for the detection and study of intrusions which avoid intrusive techniques like the direct analysis of the traffic through networking devices. Also we expose the difficulties founded in the information's analysis gathered. Finally we propose the application of machine learning to overcome them in future works.

1 Introduction

From the earliest days, networks have been experiencing an increasing number of attacks. Nowadays, the number of attacks increases continuously[1][2]. Scientist networks are a special and interesting case, in one hand there is a strong demand of security in the network and the services which are listening. On the other hand, the end users demand privacy in his network traffic. In this scene the honeypots have an important role in the detection and protection against cyber attacks.

1.1 Cyber-Space and Cyber-Threats

The cyberspace is a virtual space that wraps all types of digital communication infrastructures and the entities that use them. The hostile actions from these entities against the security and safety of the information and others entities are ciberthreats. Internet is the most popular inhabitant of this space and for years we have seen how the number and complexity of attacks against information and resources has increased. This increase is motivated by for economic, politic or military interests or by the same entities interested in exercise a bigger control over communication freedom in the cyberspace[2][4].

1.2 Scientist Networks

In contrast to the corporate networks which usually have grown from the inside to outside and which have most hosts behind the demilitarized zone (DMZ), the scientist networks were born with a open philosofy without focusing on security but on technical requirements due to the limited number of public IPs, were expanding private services to the intranet.[3]

The information related to research, patents, computer and human resources is a juicy target for hostile agents. Also, the big size of the DMZ makes it prone to a massive attack and increases the possibility of finding a security breach or hidden advance vectors of attack.

1.3 Privacy and Passive Sensors

A honeypot is a trap that exposes itself, while is scanned, probed or compromised by a hostile entity, the trap collect information about the malicious activity. We differentiate between hierarchy and interaction in our taxonomy. The hierarchy is the complexity goes from a simple service like SSH, through a network, to a cloud. The interaction is the degree of fidelity in the response of the trap and goes from low to high.

TABLE HERE.

There isn't a ideal configuration of features because is the nature of the threats and the infrastructure which we want to protect, the key for a correct selection of them. In a software development environment, high interactions is used for test a new product with *fuzzers* or another type of pentesting[1] tool in order to discover potential vulnerabilities. On the other hand, low interaction honeypots are used like intrusion detection systems, warning about activity of scans or jumping attempts from compromised internal hosts. Both share a common point: they are not intrusive with the network traffic.

2 Deployment of a Security System Based in Honeypots

The architecture of the system is divided in two fronts: detection and management of the ciberthreats. The detection front usually are based on honeypots, one the most valuable tools at hand for this purpose.

Sensors were deployed in different production subnets and each content honeypot software. Specifically Dionaea[5] and Kippo[6] which are low and medium interaction honeypot respectively. Each sensor has local data bases for save the information attacks efficiently in space while is waiting for its saved in the collector in order to keep the information by duplicate and not to increase the network traffic in case of massive scans or attacks. Obviously, in the time space between information dumps each sensor sends by telegram incidents defined by the security operator like critical. The collector is a corporate database that feeds the incidents management system and is the core of all information analysis.

PONER GRAFICO

[1] Penetration Testing.

3 Data analyse

For three years each sensor collected information of more of half million of connections. The information's analysis shows the next points:

- External attacks are more frequent than internal attacks.
- In one hand the most frequent type of external attacks was weak credentials disclosure. On the other hand the most frequent type of internal attacks was malware propagation.
- Countries outside OTAN are the most active in the process of scanning and searching for vulnerabilities but curiously most of the intrusions come from OTAN member or member candidate countries. It is important to note that this data is dependent on the geolocation of where they are taken.

The increment of attacks in this decade is a proven fact and the data cached shows us how externals attacks are the most frequent kind of attacks. This matches with studies of big security IT enterprises[**?**]. When we studied in detail the information, we saw many attempts of connection, some from scans to the network infrastructure and others looking for exploit vulnerabilities or services without strong credentials. About the latter ones is necessary to emphasize those that showed a more advanced level in the process of intrusion because were linked to Advance persistent threats (APT). One of the greatest dangers for IT infrastructures of governments, public administrations and companies are advance persistence threats. A cyberthreat is persistent if it is continuous in time and establishes monitoring and control mechanisms with hostile agent, and it is advanced because uses mechanisms in order to hide its activity in the system. Usually APT are related with cyberspying and elite groups of cybercrime and they are attacks directed against a specific infrastructure. For this reason it is a priority to detect and study them.

Malware propagation usually belong to advanced and persistent threats and come accompanied by a multivector attack. For example, the malware infect through the ms08_67 and immediately attempting connect with a command ad control server (C&C). In this point, start to protect itself through mutations of his own code, in order to hide from antivirus software, communicate through common protocols like HTTP to avoid rules for outgoing traffic firewall, it use encrypted communication by SSL to prevent the interception of information or generate thousands names of random domain to prevent the take down of the botnet.

But not always is easy follow the clue for rebuilding of a multivector attack. Usually the exploitation of SSH or MySQL weak credentials is the first step to gain the control or access to data in a server, but only a very reduced part shows a clever behavior behind the attack. Between hundred of thousand of connections only a few ones shows access to information to a service, then it use this informatión against other services and finally jumping to other hosts. A bit more frequent is the attempt to privilege elevation but the common behavior is use the base vulnerability in order to use his network and computational resources as soon as possible, in tasks like miner Litecoin[8], increase the number of nodes for other scans of networks, for a future deny of service attack or use the compromised host like a anonymous proxy.

When we rebuild the trace of the attack, the first advanced behavior that we find is the use of different hosts for scans and attacks and others for the intrusions and the explotations. The attack starts to scan subnets usually from countries without collaborations accord, in our case China. Then of the detection of the vulnerabilities, the exploitation is from Europe or United States. Obviously is not possible to be totally sure about the origin of the intrusion because we dodn't detect the use of proxies, but there isn't sense the use from countries like Chine for hidden its source in the intrusion stage but not in the scan stage. It has more relevance the inverse theory, the attack come from countries in a increase technology stage and many resources, where the culture of security isn't still high, finally the intrusion attempt is from a nearer geolocation of the attacker.

4 Strengths and Weaknesses of Honeypots

The strengths of honeypot are were:

- It was not intrusive with network traffic, remained the privacy of infrastructure users. This is a important point because any try of to catch indirect traffic of network would been seen as a threat by other users and a infringement of the use conditions of the network and legality.
- The computational and economic resources needed for passive detection are lower because we have only the traffic belong to a potential cyber threat. This an alternative approach to other solutions more expensive like intrusion detection systems based in hardware.
- Cyberthreats like advanced malware uses ciphered communications in order to dodge detection systems in the network layer, the only way to catch information its from inside of the compromised node. This is essential if we want analyse how persistence cibertheraths monitors the compromised host and what informations sent outside to its C&C network.

The weakness of the honeypot are:

- There are many cyberthreats focused in the network layer, usually related with deny of services and spoofing. This information is very valuable because this kind of attack are a very important element not only in simple vector attacks, in multivector advanced and persistent attacks too. Honeypots only fetch information from the application layer so they lose essential information for reconstruction of complex attacks.
- The use of passive sensors in a security system must be planned with some extra considerations respect the use of active methods of detection. Those considerations cover strategies of deception and hiding of the sensors and politics of migrations in the infrastructure for avoiding it's location.
  - Like others deception tools, honeypots must show itself like interesting target for a attacker and avoid them to be easily recognizable by fingerprint techniques. Default installations and configurations in low an medium interaction honeypot are easily detected by a human attacker or a intelligent threat like advanced malware.

- When attacker has knowledge of the infrastructure, honeypots are easily dodged so must be deployed together with politics of use like change its subnets or IP every so often.
- High interaction honeypots are dangerous in production environments because the monitored sensor is completely real and all his potential in order to attack its periphery. Usually they are deploying in isolated subnets with outgoing traffic strongly restricted in company of others honeypots, that configuration is called honeynet.

5 Improving Honeypots, Hybrid Machine Learning S.I.E.M

VECTORES HONYPOT VS CONJUTOS DE DATOS DEL LEARNING MACHINE GROUP FALTA CHICA, CAPA DE RED, TCPDUMP O IDS PILLANDO DATA

6 Conclusions and Future Works

HAY QUE METER MACHINE LEARNING Y MINERIA DE DATOS DE LA BUENA A ESTOS DISPOSITIVOS.

References

1. ESSET Latino América: Tendencias 2015: El mundo corporativo en la mira. January 2015. URL: http://http://www.welivesecurity.com/wp-content/uploads/2015/01/tendencias_2015_eset_mundo_corporativo.pdf Note: Chapter 2, page 6
2. CNI-Centro Critográfico Nacional: Informe de Amenazas CCN-CERT IA-03/14: Ciberamenazas 2013 y Tendencias 2014. October 2014. URL:https://www.ccn-cert.cni.es/publico/dmpublidocuments/CCN-CERT_IA-03-14-Ciberamenazas_2013_Tendencias_2014-publico.pdf Note: Prologue, page 7,8
3. Subdirección General de Organización y Automación, Secretaía General Técnica, Ministerio de Educación y Ciencia: Proyecto IRIS. November 1985. URL:https://www.rediris.es/rediris/historia/programa-iris.pdf
4. CNI-Centro Critográfico Nacional: CCN-CERT IA-09/15 Ciberamenazas_2014 Tendencias_2015 - Resumen Ejecutivo April 2015. URL:https://www.ccn-cert.cni.es/publico/dmpublidocuments/IE-Ciberamenazas2014-Tendencias-2015.pdf
5. Mark Schloesser: URL: http://dionaea.carnivore.it/
6. desaster: https://github.com/desaster
7. Verizon Enterprise: 2015 Data Breach Investigations Report 2015. http://www.verizonenterprise.com/resources/reports/rp_data-breach-investigation-report-2015_en_xg.pdf
8. https://litecoin.org/

https://www.schneier.com/blog/archives/2008/06/it_attacks_insi.html http://www.verizonenterprise.com/resources/repo breach-investigations-report-2013_en_xg.pdf