ZEAL EDUCATION SOCIETY's

ZEAL COLLEGE OF ENGINEEIRNG AND RESEARCH,
NARHE, PUNE

DEPARTMENT OF COMPUTER ENGINEERING

SEMESTER-I

[A.Y. : 2022 - 2023]

CYBER SECURITY AND DIGITAL FORENSICS(410244(C))

LABORATORY MANUAL

Institute and Department Vision and Mission

INSTITUTE VISION	To impart value added technological education through pursuit of academic excellence, research and entrepreneurial attitude.
INSTITUTE MISSION	**M1:** To achieve academic excellence through innovative teaching and learning process. **M2:** To imbibe the research culture for addressing industry and societal needs**.** **M3:** To provide conducive environment for building the entrepreneurial skills. **M4:** To produce competent and socially responsible professionals with core human values.

DEPARTMENT VISION	To emerge as a department of repute in Computer Engineering which produces competent professionals and entrepreneurs to lead technical and betterment of mankind.
DEPARTMENT MISSION	**M1:** To strengthen the theoretical and practical aspects of the learning process by teaching applications and hands on practices using modern tools and FOSS technologies. **M2:** To endeavour innovative interdisciplinary research and entrepreneurship skills to serve the needs of Industry and Society. **M3:** To enhance industry academia dialog enabling students to inculcate professional skills. **M4:** To incorporate social and ethical awareness among the students to make them conscientious professionals.

Department
Program Educational Objectives(PEOs)

PEO1: To Impart fundamentals in science, mathematics and engineering to cater the needs of society and Industries.

PEO2: Encourage graduates to involve in research, higher studies, and/or to become entrepreneurs.

PEO3: To Work effectively as individuals and as team members in a multidisciplinary environment with high ethical values for the benefit of society.

Savitribai Phule Pune University

Fourth Year of Computer Engineering (2019 Course)

410244(C): Cyber Security & Digital Forensics Laboratory

Teaching Scheme:	Credit	Examination Scheme:
PR: 04 Hours/Week	02	TW: 25 Marks PR: 50 Marks

Course Objectives:

- To enhance awareness cyber forensics.
- To understand issues in Cyber Crime and different attacks
- To understand underlying principles and many of the techniques associated with the digital forensic practices
- To know the process and methods of evidence collection
- To analyze and validate forensic data collected.
- To apply digital forensic knowledge to use computer forensic tools and investigation report writing.

Course Outcomes:

On completion of the course, student will be able to-

CO1: Analyze threats in order to protect or defend it in cyberspace from cyber-attacks.

CO2: Build appropriate security solutions against cyber-attacks.

CO3: Underline the need of digital forensic and role of digital evidences.

CO4: Explain rules and types of evidence collection

CO5: Analyze, validate and process crime scenes CO6: Identify the methods to generate legal evidence and supporting investigation reports.

List of Assignments

Sr. No.	TITLE
	Group A
01	Write a program for tracking Emails and Investigating Email Crimes. i.e. Write a program to analyze e-mail header
02	Implement a program to generate and verify CAPTCHA image
03	Write a computer forensics application program for Recovering permanent Deleted Files and Deleted Partitions.
04	Write a program for Log Capturing and Event Correlation
05	Study of Honeypot.
Group B	
Mini-Projects/ Case Study (Any two)	
01	Mini Project- Design and develop a tool for digital forensics of images
02	Mini Project- Design and develop a tool for digital forensics of audio
03	Mini Project- Design and develop a tool for digital forensics of video
04	Mini Project- Design a system for the analysis of cyber crime using various cyber forensics techniques and compare each technique with respect to integrity, confidentiality, availability

GROUP A: ASSIGNMENT NO 1

Title:

Email Header Analysis

Problem Statement:

Write a program for Tracking Emails and investigating Email Crimes, i.e. to write a program to analyze e-mail header.

Aim:

To study for Tracking Emails, Investigating Email Crimes and program to analyzer e-mail header

Objective:

To develop a program for analyze e-mail header.

Theory :

1) **Mail: -** anything that's delivered to your mail box or post office box — letters, bills, packages, magazines, or anything else that's sent through the postal service. Email is the internet's version of mail.
2) **Mail header:** Email Headers are lines of metadata (data about data) attached to each email that contain lots of useful information for a forensic investigator.
3) **Providing information about the sender and recipient:** An email header tells who sent the email and where it arrived. Some markers indicate this information, like “From:” — sender’s name and email address, “To:” — the recipient’s name and email address, and “Date:” — the time and date of when the email was sent. All of these are mandatory indicators. Other parts of the email header are optional and differ among email provide service.
4) **From –** It displays who the message is from, however, this can be easily forged and can be the least reliable.
5) **Subject -** This is what the sender placed as a topic of the email content.
6) **Date -** This shows the date and time the email message was composed.
7) **To -** This shows to whom the message was addressed, but may not contain the recipient's address.
8) **Return-Path-** The email address for return mail. This is the same as "Reply-To:".
9) **Envelope-To-** This header shows that this email was delivered to the mailbox of a subscriber whose email address is user@example.com.
10) **Delivery Date -** This shows the date and time at which the email was received by your (mt) service or email client.
11) **Received -** They form a list of all the servers/computers through which the message traveled in order to reach you. The received lines are best read from bottom to top. That is, the first "Received:" line is your own system or mail server. The last "Received:" line is where the mail originated. Each mail system has their own style of "Received:" line. A "Received:" line typically identifies the machine that received the mail and the machine from which the mail was received.
12) **Message-id-** A unique string assigned by the mail system when the message is first created. These can easily be forged.

13) **Mime-Version -** Multipurpose Internet Mail Extensions (MIME) is an Internet standard that extendsthe format of email.
14) **Content-Type-** Generally, this will tell you the format of the message, such as html or plaintext.
15) **X-Spam-Status-** Displays a spam score created by your service or mail client.
16) **X-Spam-Level-** Displays a spam score usually created by your service or mail client.
17) **Message Body-** This is the actual content of the email itself, written by the sender.

Algorithm-:

```
Step 1 : Start
Step 2 : Import re library
Step 3 : For regstr in args
            match obj = re.search(regstr in args)
            if match obj == true then
                print match obj.group(0)
            else print "Not found"
Step 4 : print("email file name")
            Enter path as filename
Step 5 : Open filename as fo
Step 6 : data = fo.read()
Step 7 : match re(data, "mime-version", "Date: ", "Sub, ")
Step 8 : close file
Step 9 : End
```

Code-:

```
import re
def matchre(data, *args):
    for regstr in args:
        matchObj = re.search( regstr+'.*', data, re.M|re.I)
        if matchobj:
            print(matchobj.group(0).lstrip().rstrip())
        else:
            print("No",regstr,"found")
print("Email Header Program by XXXXX") #X=name of email-owner
filename= input("Enter path for email header file\n");
fo = open(filename, "r") #fo-filehandle
data=fo.read()
matchre(data, "MIME-version", "Date:","Subject:","delivered-to: ", "From:","to:")
fo.close()
```

Conclusion-: Successfully implemented program for E-mail header

Questions-:

1) What is email header?
2) What is meant by tracking emails?
3) What do you analyze an email header?
4) What information analyst can get from email header?
5) What is importance of email header analysis?

Date:	
Marks obtained:	
Sign of course coordinator:	
Name of course Coordinator :	

GROUP A : ASSIGNMENT NO 2

Title: -
Generate and Verify CAPTCHA

Problem Statement: -

Implement a program to generate and verify CAPTCHA image

Aim: -
To write a program code to generate and verify CAPTCHA image.

Objective: -
To develop a program which can generate and validate CAPTCHA image.

Theory Concepts: -

1) A CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a test to determine whether the user is human or not.
2) It is a security measure known for its challenge-response authentication. It is computer program or system used to distinguish humans from machine inputs (bots), typically usedfor spamming and web-crawling.
3) It is basically consisting of some distorted or overlapping letters and numbers that a user then has to submit via a form field. The distortion of the letters makes it difficult for bots to interpret the text and prevented access until the characters are verified.
4) This CAPTCHA type relies on a human's ability to generalize and recognize novel patterns based on variable past experience. In contrast, bots can often only follow set patterns or inputrandomized characters.
5) CAPTCHA offers protection from remote digital entry by making sure only a human being with the right password can access your account. CAPTCHA works because computers can create a distorted image and process a response, but they can't read or solve the problem the way a human must to pass the test.
6) Uses of CAPTCHA-:
   1. Maintaining poll accuracy
   2. Limiting registration for services
   3. Preventing false comments
   4. To ensure real human accessing the site and thus preventing Bots interactions.

Algorithm: -

```
Step 1: Import essential libraries
Step 2: generate random_string as captcha
Step 3: Open an image for background in captcha and assign an appropriate fontto it. Add theRandom_string that
            you generated on the image.
Step 4: Display Captcha ImageStep 5:
Take input from user
Step 6: Validate
                if random_string == user_inputdisplay
                    "Captcha Verified"
                else
                    display "Wrong captcha"
```

Code: -

```
# Importing essential librariesimport
ipywidgets as widgets
from PIL import Image, ImageDraw, ImageFontimport random
import string

def random_string_generator():return
    ''.join(random.choices(
            string.ascii_letters + string.digits, k=6))

def create_captcha_image(random_string):img =
    Image.open('captcha.png','r')

 # Call draw Method to add 2D graphics in an imagefinal_image
 = ImageDraw.Draw(img)

 # Custom font style and font size
 myFont = ImageFont.truetype('GoodbyeCrewelWorldNF.ttf', 65)

 # Add Text to an image
 final_image.text((30, 10), random_string, font=myFont, fill =(255, 0, 0))
 img.save("captcha_pic.png")
 img.close()

 # Display edited image
 with open("captcha_pic.png", "rb") as file:image =
        file.read()
        return widgets.Image(
            value=image, format='png',
            width=300, height=400 )

random_string = random_string_generator()

create_captcha_image( random_string )

captcha_entered_by_user=input("Enter Captcha : ")

# Validating Captcha
if (random_string==captcha_entered_by_user):
    print("Captcha Verified")
else :
    print("Wrong Captcha")
```

Conclusion: To learn about CAPTCHA image and also implement a code for generating image CAPTCHA and its validation. Successfully implemented program for Image CAPTCHAand its validation.

Viva Questions-:

1) What are the different types of captcha?
2) Is captcha hackable? Where Captcha is used?
3) Why do you think every website should implement captcha?
4) Is there any alternative to captcha? What is it?
5) Why is there a need to use captcha in verification?

Date:	
Marks obtained:	
Sign of course coordinator:	
Name of course Coordinator :	

GROUP A: ASSIGNMENT NO 3

Title: -

Recovering permanent Deleted Files and Deleted Partitions.

Problem Statement: -

Write a computer forensic application program for Recoveringpermanent Deleted Files and Deleted Partitions.

Aim: -

To recover permanent Deleted Files (Digital Evidence) from system with the help of acomputer forensic application program.

Objective-:

To learn about methods and techniques to recover deleted digital evidence/files.

Theory Concepts: -

Deleted Files:

Deleting files is one of the easiest, convenient, and foremost ways to destroy the evidence. The principle of file recovery of deleted files is based on the fact that Windows does notwipe the contents of the file when it's being deleted. Instead, a file system record storing the exact location of the deleted file on the disk is being marked as "deleted" and the disk space previously occupied by the deleted file is then labeled as available – but not overwritten with zeroes or other data.

- The deleted file can be retrieved by analyzing the contents of the recycle bin as they are temporarily stored there before being erased.
- If the deleted files have no trace in the recycle bin like in case of the "Shift+Delete" command, then, in that case, you can use commercial recovery tools to recover the deleted evidence. One such example commercial tool is disk drill or Disk internals Partition Recovery.
- Looking for characteristic signatures of known file types by analyzing the file system and/or scanning the entire hard drive, one can successfully recover:
  - Files that were deleted by the user.
  - Temporary copies of Office documents (including old versions and revisions ofsuch documents).
  - Temporary files saved by many applications.
  - Renamed files.
- Information stored in deleted files can be supplemented with data collected from other sources. For example, the "chatsync" folder in Skype stores the internal data that may contain chunks and bits of user conversations. This means if the "chatsync" folder exists there is a possibility to recover user chat's even if the Skype database is deleted. Many tools exist for this purpose like Belk Soft Evidence 2020

Formatted Hard Drives:
Recovery of the data from the formatted hard drive depends upon a lot of parameters. Information from the formatted hard drive may be recoverable either using data carving technology or by using commercial data recovery tools. There are two possible ways to format a hard drive: Full Format and Quick Format.

SSD Drives:
SSD means Solid-State Drives represent a new storage technology.

- They operate much faster than traditional drives.
- They employ a completely different way of storing information internally, which makesit much easier to destroy information and much more difficult to recover it.

Conclusion-: By using disk drill a computer forensic application program we learn abouthow to recover our deleted data, permanent Deleted Files and Deleted Partitions, etc.

Questions-:
1) What is computer forensic
2) What is meant by digital evidence.
3) Examples of digital evidence.
4) What is type of computer forensics
5) What is importance of recovering data

Date:	
Marks obtained:	
Sign of course coordinator:	
Name of course Coordinator :	

GROUP A: ASSIGNMENT NO 4

Title: -
Log Capturing and Event correlation

Problem Statement: -
Write a program for Log Capturing and Event Correlation

Objectives:

- To understand the event log concepts
- To understand the places where event logs are maintained.
- To capture and analyze the log files.
- To correlate the events

Theory Concepts:-

Event – a change in the system state, e.g., a disk failure; when a system component (application, network device, etc.) encounters an event, it could emit an event message that describes the event.

- /var/log/messages
- /var/log/maillog : Mail
- /var/log/httpd : Apache webserver messages

Event correlation – a conceptual interpretation procedure where new meaning is assigned to a set of events that happen within a predefined time interval.
Event correlation is a technique for making sense of a large number of events and pinpointing the few events that are really important in that mass of information. Event correlation is performed by a special utility called as event correlator.
Examples:

- if 10 *login failure* events occur for a user within 5 minutes, generate a *security attack* event.
- if both *device internal temperature too high* and *device not responding* events have been observed within 5 seconds, replace them with the event *device down due to overheating*.

Event correlation can be decomposed into four steps: event filtering, event aggregation, event masking and root cause analysis.

1. Event filtering: - Event filtering consists in discarding events that are deemed to be irrelevant by the event correlator.
2. Event aggregation: - Event aggregation (also known as event de-duplication) consists in merging duplicates of the same event.
3. Event masking: - Event masking consists in ignoring events pertaining to systems that are downstream of a failed system.

4. Root cause analysis: - It consists in analyzing dependencies between events, based for instance on a model of the environment and dependency graphs, to detect whether some events can be explained by others.

Conclusion: -

Successfully implemented log capturing and event correlation program

Viva Questions :-

1) What are logs? What is log capturing?
2) Where do you find process in your PC?
3) What is event correlation? Give an example.
4) What are the steps in event correlation process?

Date:	
Marks obtained:	
Sign of course coordinator:	
Name of course Coordinator :	

GROUP A: ASSIGNMENT NO 5

Title-:
Study of Honeypot

Problem Statement-:
To study working of Honeypot.

Aim-:
To understand and study Honeypot.

Objective-:
To understand working of honeypot and to understand design and implementation ofa honeypot.

Theory Concepts:

A honeypot is a computer system that is set up to act as a decoy to lure cyber-attackers, and to detect, deflect or study attempts to gain unauthorized access to information systems.

Generally, a honeypot consists of data (for example, in a network site) that appears to be a legitimate part of the site, but is actually isolated and monitored, and that seems to contain information or a resource of value to attackers, who are then blocked. This is similar to the police baiting a criminal, conducting undercover surveillance, and finally punishing the criminal.

A honeypot is a security resource whose value lies in being probed, attacked or compromised.

Honeypots are weapons against spammers, honeypot detection systems are spammer-employed counter-weapons. As detection systems would likely use unique characteristics of specific honeypots to identify them.

Honey Pots can be setup inside, outside or in the DMZ of a firewall design or even in all of the locations although they are most often deployed inside of a firewall for control purposes. In a sense, they are variants of standard Intruder Detection Systems (IDS) but with more of a focus oninformation gathering and deception.

Honeypots can be classified based on their deployment (use/action) and based on their level of involvement. Based on deployment, honeypots may be classified as

1. **Production Honeypots**: - Production honeypots are easy to use, capture only limited information, and are used primarily by corporations. Production honeypots are placed inside the production network with other production servers by an organization to improve their overall state of security. Normally, production honeypots are low-interaction honeypots, which are easier to deploy. They give less information about the attacks or attackers than research honeypots.
2. **Research Honeypots**: **-** Research honeypots are run to gather information about the motives andtactics of the Black hat community targeting different networks. These honeypots do not add direct value to a specific organization; instead, they are used to research the threats that organizations face and to learn how to better protect against those threats. Research honeypots are complex to deploy and maintain, capture extensive information, and are used primarily by research, military,or government organizations.

Based on design criteria, honeypots can be classified as:

1. **Pure Honeypots**:-
   A pure honeypot refers to a full-scale system running on various servers. It completely mimics the production system. Within a pure honeypot is data made to look confidential, as well as "sensitive" user information, which have a number of sensors used to track and observe attacker activity.
2. **High-interaction honeypots**:-
   A computer system with a complete operating system (OS) installation and real network services running on it would be an example of high interaction honeypot. In this case, attackersconnecting to it would get their responses from the real services and they might be able to break into the system and take over it should these services present any vulnerability.
3. **Low-interaction honeypots**:-
   A piece of software capable of emulating network services would be an example of low interaction honeypot. Attackers connecting to it would get responses similar to the real servicesbut they would never be able to abuse these fake services using an exploit they may have codedfor their real counterparts.

Other Types of Honeypots

Malware Honeypot: It is used for luring malwares.

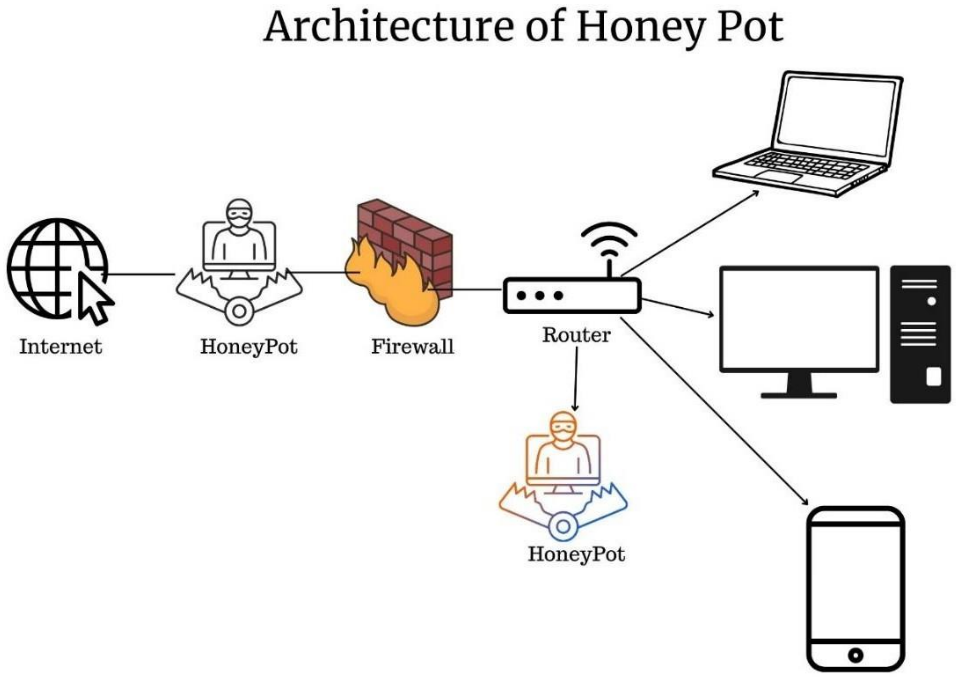
Spam Honeypot: Used to attract spammers by using open proxies and mail relays.

Database Honeypot: A database honeypot is used to make decoy databases to attract database-specific attacks like SQL injections, which illicitly manage data.

Client Honeypot: Client honeypots attempt to lure in malicious servers that attackers use while hacking clients. They pose as a client to observe how an attacker makes modifications to a server during the attack.

Honeynet: Honeynets consist of a network of honeypots. With different kinds of honeypots forming a honeynet, several types of attacks can be studied, such as distributed denial-of-service (DDoS) attacks, attacks to a content delivery network (CDN), or a ransomware attack.

Architecture of Honeypot-:



The architecture of an intrusion detection and prevention solution should be distributed, unless thenetwork is designed in a dedicated way. The architecture presented in this paper takes into accountthe possibility of using distributed approach. Such an approach is capable of detecting multiple types of attacks basing on the combination of two or more solutions.

It is possible to position honeypot systems on local area network (LAN), Internet, and demilitarized zone (DMZ). Each of these positioning scenarios has advantages and honeypots arestructured in ways and shapes that do not risk network security, depending on where they are located. Depending upon the use cases, honeypots are located at different parts of the network. Whenever the attacks try to attack these honeypots then alerts can be triggered. These honeypots also maintain logs and traces activities of the attackers so that we can figure out the intension of the attackers, we can predict future attacks and prepare for them accordingly. We come to know various details of the attackers like systems used by them, different configurations, IP addresses, etc.

Conclusion: -

Honeypots offer benefits that no other technology can provide. Although honeypotsdo not directly protect any systems, they improve the overall security posture of a network by allowing the network administrators to quickly identify systems from which attacks are being launched.

Viva Questions-:

1) What is honeypot?
2) Can you give me some examples of real-world scenarios where honeypots were used tosuccessfully identify malicious activity?
3) How do you use honeypots in your organization's IT security strategy?
4) What are some common attacks against web applications that honeypots can help prevent?
5) Do you think honeypots are completely secure from attack? Why or why not?
6) How do you use honeypots in your organization's IT security strategy?

Date:	
Marks obtained:	
Sign of course coordinator:	
Name of course Coordinator :	